



UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE CIÊNCIA POLÍTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA POLÍTICA
DOUTORADO EM CIÊNCIA POLÍTICA

MAX STABILE

GOOGLE E AS ELEIÇÕES BRASILEIRAS

BRASÍLIA

2024

MAX STABILE

GOOGLE E AS ELEIÇÕES BRASILEIRAS

Tese apresentada ao Programa de Pós-Graduação em Ciência Política do Instituto de Ciência Política da Universidade de Brasília, como requisito parcial à obtenção do título de doutor em Ciência Política. Área de Concentração: Política e Instituições

Orientadora: Marisa von Bülow

BRASÍLIA

2024

Ficha catalográfica elaborada automaticamente,
com os dados fornecidos pelo(a) autor(a)

SS775g Stabile, Max
GOOGLE E AS ELEIÇÕES BRASILEIRAS / Max Stabile;
orientador Marisa von Bülow. -- Brasília, 2024.
218 p.

Tese(Doutorado em Ciência Política) -- Universidade de
Brasília, 2024.

1. Google. 2. Eleições. 3. SERP. 4. histórico de
navegação. 5. motores de busca. I. von Bülow, Marisa,
orient. II. Título.

MAX STABILE

GOOGLE E AS ELEIÇÕES BRASILEIRAS

Tese apresentada ao Programa de Pós-Graduação em Ciência Política do Instituto de Ciência Política da Universidade de Brasília, como requisito parcial à obtenção do título de doutor em Ciência Política. Área de Concentração: Política e Instituições

Aprovada em: 21 de junho de 2024

BANCA EXAMINADORA

Marisa von Bülow (Orientadora)
Universidade de Brasília

Luciana Fernandes Veiga
Universidade Federal do Estado do Rio de
Janeiro

Ernesto Calvo
University of Maryland

Mathieu Turgeon
University of Western Ontario

Às minhas fortalezas: Deborah, Ana Sofia e Elis.

AGRADECIMENTOS

Agradeço imensamente por toda a ajuda recebida ao longo da produção desta tese. Sou grato à professora Marisa von Bülow pela orientação impecável, precisa e exigente, mas sempre acompanhada de palavras motivadoras e acolhedoras ao longo destes anos. Também agradeço ao grupo RESOCIE, onde, em 2018, iniciamos as primeiras coletas sobre o Google, e aos colegas Alexandre Arns e Lorena Vilarins, que sempre incentivaram este trabalho.

Muitas pessoas me auxiliaram ao longo dos últimos cinco anos. Destaco Jeferson Ferreira e Maria Claudia da Silva, que me ajudaram na coleta de dados em 2020; Andressa Kniess e Luiza Aikawa, pela assistência na etapa teórica; Williams Cuccovia e Rebeca Garcia, por me ajudar com a visão do mercado sobre o Google; Amanda Ávila, por me ajudar no questionário da pesquisa; Giulliano Cesar, um super parceiro no trabalho de análise dos dados; Rafael Neves, meu querido amigo que revisou o capítulo de Diff-in-Diff; e Kaio Ribeiro pela compilação final desta tese e sua formatação em Markdown e \LaTeX .

Agradeço aos meus dois mestres: Prof. Carlos Batista, pelo incentivo constante, e Prof. Wladimir Gramacho, que, junto com o Centro de Pesquisa em Comunicação Política e Saúde Pública (CPS) da UnB, com quem realizamos tantas parcerias produtivas nos últimos anos, tornou possível, por meio de um financiamento conjunto, a aplicação do painel desta pesquisa.

Expresso minha gratidão a todo o time do Instituto Brasileiro de Pesquisa e Análise de Dados (IBPAD), especialmente à minha querida sócia e amiga Jaqueline Buckstegge, pelas dicas, conselhos e por permitir que o IBPAD financiasse parte da coleta dos dados desta pesquisa.

Por fim, agradeço a Deus e ao imenso apoio da minha família nestes últimos anos, especialmente Deborah Celentano. Seu amor e suporte foram fundamentais para esta jornada.

RESUMO

Esta tese investiga o papel do Google nas eleições brasileiras de 2020, focando na intermediação de informações políticas em uma eleição subnacional. O estudo foi conduzido com uma abordagem metodológica que combina: análise quantitativa de 518 mil resultados de buscas coletados (*SERP - Search Engine Results Page*) entre agosto e novembro de 2020 em uma amostra de municípios brasileiros; duas rodadas de *survey* em painel longitudinal com 2771 indivíduos, representativos da população brasileira (antes e após as eleições); e a coleta do histórico de buscas de parte dos respondentes. Este trabalho responde a quem, como e quando o Google é relevante nas eleições, demonstra quais foram os principais sites exibidos para diferentes termos de busca e, por meio de um “quase-experimento”, avalia a existência de interferência por parte do Google nos resultados de pesquisa ao início do período eleitoral. Os resultados indicam que eleitores com forte preferência política são os que mais realizam buscas político-eleitorais, desafiando as teorias que defendem o *Search Engine Manipulation Effect*; mostram que o Google é crucial na busca de informações sobre candidatos, especialmente nos últimos três dias antes da eleição; destacam que 10 sites concentraram 60% dos resultados das buscas político-eleitorais do período; e sugerem indícios de possível interferência do Google no início do período eleitoral.

Palavras-chave: eleições, Google, SERP, histórico de navegação, motores de busca, Brasil

ABSTRACT

This thesis investigates the role of Google in the Brazilian elections of 2020, focusing on the mediation of political information in a subnational election. The study was conducted with a methodological approach that combines: quantitative analysis of 518,000 search results (SERP - Search Engine Results Page), collected between August and November 2020 in a sample of Brazilian municipalities; two rounds of longitudinal panel surveys with 2,771 individuals, representative of the Brazilian population (before and after the elections); and the collection of the search history of some of the respondents. This work answers for whom, how, and when Google is relevant in elections, maps the main sites displayed for different search terms, and, through a “quasi-experiment”, assesses the existence of interference by Google in search results at the start of the electoral period. The results indicate that voters with strong political preferences are the ones who most frequently conduct political-electoral searches, challenging the Search Engine Manipulation Effect hypothesis; show that Google is crucial for accessing information about candidates, especially in the last three days before the election; highlight that 10 sites accounted for 60% of the political-electoral search results during the period; and suggest signs of possible Google interference at the beginning of the electoral period.

Keywords: elections, Google, SERP, browsing history, search engines, Brazil

LISTA DE FIGURAS

Figura 1 – Exemplo de uma posição 0 em um resultado de busca	50
Figura 2 – Resultado da busca: patroa	54
Figura 3 – Resultado da busca: lula coroação	55
Figura 4 – Trecho de resultado da busca retroativa para “candidata a prefeita”	63
Figura 5 – Processo de captura massiva dos SERPs	66
Figura 6 – Resultado da busca "where am i"sem nenhum parâmetro de localização	67
Figura 7 – Resultado da busca "where am i"utilizando-se o parâmetro UULE para Serra da Saudade/MG	67
Figura 8 – Resumo das etapas de captura	71
Figura 9 – Distribuição da variável - Uso Monitorado do Google	83
Figura 10 – Distribuição das Buscas Políticas (Uso Declarado)	85
Figura 11 – Média de Uso Monitorado por Uso Declarado	86
Figura 12 – Análise de Correspondência Múltipla	92
Figura 13 – Árvore de distribuição hierárquica	94
Figura 14 – Coeficientes dos Modelos de Regressão Logística	102
Figura 15 – Trigrama dos principais termos buscados pelos usuários classificados na categoria Informações sobre Candidatos/Partidos	108
Figura 16 – Trigrama dos principais termos buscados pelos usuários classificados na categoria Processos/Procedimentos Eleitorais	110
Figura 17 – Trigrama dos principais termos buscados pelos usuários classificados na categoria Buscas Gerais sobre Eleições	111
Figura 18 – Quantidade média de buscas de cada tipo de busca classificada por diferentes perfis	113
Figura 19 – Evolução das buscas no período eleitoral	115
Figura 20 – Série de tempo das diferentes buscas classificadas a partir do histórico de uso	117
Figura 21 – Visão gráfica do método de diferenças em diferenças	146
Figura 22 – Histograma da variável população	149
Figura 23 – Histograma da variável Veículos de mídia	150
Figura 24 – Residuals Monitorado	184
Figura 25 – Residuals Monitorado	184

LISTA DE TABELAS

Tabela 1 – Conjunto de expressões regulares para definição de uma busca política/eleitoral	64
Tabela 2 – Período de coleta dos resultados de busca do Google	66
Tabela 3 – Quantidade de buscas realizadas por tipo de abordagem	68
Tabela 4 – Distribuição dos municípios com buscas gerais sobre eleições	68
Tabela 5 – Termos de busca que os usuários fariam para buscar mais informações sobre os candidatos	69
Tabela 6 – Tabela dos termos de busca políticos para cada tipo de município	70
Tabela 7 – Resultados para a pergunta: Você já sabe ou já ouviu falar sobre quem são os(as) candidatos(as) ao cargo de prefeito(a) em sua cidade?	76
Tabela 8 – Resultados para a pergunta: E você já sabe em quem você vai votar? (Apenas para quem conhecia algum candidato)	76
Tabela 9 – Resultados para a pergunta: Você votou no primeiro turno em algum(a) candidato(a) ao cargo de prefeito(a) municipal?	77
Tabela 10 – Resultados para a pergunta: Qual é o seu partido político de preferência?	78
Tabela 11 – Resultados para a pergunta: Em qual partido político você nunca votaria?	79
Tabela 12 – Quantitativos gerais do histórico de busca dos usuários	82
Tabela 13 – Resultados para a pergunta: Ao longo do período eleitoral, com qual frequência você fez cada um desses tipos de busca sobre eleições no Google?	84
Tabela 14 – Variável dependente: Uso declarado	88
Tabela 15 – Quadro resumo das variáveis	89
Tabela 16 – Distribuição dos perfis	94
Tabela 17 – Resultados para o Modelo Monitorado	97
Tabela 18 – Resultados para o Modelo Declarado	99
Tabela 19 – Classificação dos termos utilizados nas buscas políticas a partir do histórico de uso	112
Tabela 20 – Exemplo de URLs e subdomínios vinculados ao Uol	122
Tabela 21 – Os 10 domínios mais frequentes nos resultados de buscas políticas	123
Tabela 22 – Os 10 domínios mais frequentes nos resultados de buscas não políticas	124
Tabela 23 – Termos de busca utilizados nas Buscas Gerais	125
Tabela 24 – Os 10 domínios mais frequentes nos resultados do termo Eleições 2020	126

Tabela 25 – As 03 principais URLs para os sites mais frequentes nos resultados da busca com o termo Eleições 2020	127
Tabela 26 – Os 10 domínios mais frequentes nos resultados das buscas com termos sobre candidatos	129
Tabela 27 – As 03 principais URLs para os sites mais frequentes nos resultados da busca com termos sobre candidatos	129
Tabela 28 – Proporção acumulada nos dez primeiros sites para cada faixa de município	131
Tabela 29 – Os 10 domínios mais frequentes nos resultados das buscas feitas nas capitais brasileiras	132
Tabela 30 – As 03 principais URLs para os sites mais frequentes nos resultados das buscas realizadas nas capitais brasileiras	134
Tabela 32 – Os 10 domínios mais frequentes nos resultados das buscas temáticas	134
Tabela 33 – As 03 principais URLs para os sites mais frequentes nos resultados das buscas temáticas	135
Tabela 31 – Termos utilizados para simular buscas políticas temáticas	137
Tabela 34 – Concentração dos resultados nos dez principais sites	137
Tabela 35 – Quantidade e proporção de buscas (em 1000)	142
Tabela 36 – Procedimento de Diferenças-em-Diferenças (DiD)	143
Tabela 37 – Médias para os 10 primeiros resultados - Top 10	144
Tabela 38 – Médias para os 03 primeiros resultados - Top 03	144
Tabela 39 – Tipos de busca por grupo (tratamento e controle)	147
Tabela 40 – Distribuição das buscas nas regiões brasileiras	147
Tabela 41 – Quantidade de buscas por dispositivo móvel	148
Tabela 42 – Comparação dos modelos Top 10 e Top 3 domínios	153
Tabela 43 – Comparação dos modelos Top 10 e Top 3 domínios na janela de 7 dias	156
Tabela 44 – Resultados para a pergunta: Critério Brasil de Classificação Econômica	176
Tabela 45 – Resultados para a pergunta: Até que série você estudou?	176
Tabela 46 – Resultados para a pergunta: Você é...	176
Tabela 47 – Resultados para a pergunta: E quantos anos você tem?	177
Tabela 48 – Resultados para a pergunta: A sua raça ou cor é:	177
Tabela 49 – Resultados para a pergunta: Qual a sua religião, se tiver?	177
Tabela 50 – Região do entrevistado	178

Tabela 51 – Quantidade de buscas realizadas por tipo de abordagem	178
Tabela 52 – Quantidade de buscas não eleitorais realizadas	178
Tabela 53 – Descritivo dos municípios inclusos na captura	179
Tabela 54 – Termos utilizados nas buscas políticas	179
Tabela 55 – Quantidade de buscas eleitorais realizadas	180
Tabela 56 – Quantidade de buscas não eleitorais realizadas	181
Tabela 57 – VIF - Modelo Declarado	182
Tabela 58 – VIF - Modelo Monitorado	182
Tabela 59 – Model Statistics for Model declarado	183
Tabela 60 – Model Statistics for Model monitorado	183

SUMÁRIO

1	INTRODUÇÃO	16
1.1	Por que essa tese é relevante?	17
1.2	Objeto de pesquisa	19
1.3	Perguntas de pesquisa	19
1.3.1	<i>Pergunta 1 - Para quem o Google é relevante?</i>	20
1.3.2	<i>Pergunta 2 - Como e quando os eleitores utilizam o Google?</i>	20
1.3.3	<i>Pergunta 3 - O que o Google entregou de resultados em 2020?</i>	21
1.3.4	<i>Pergunta 4 - Há interferência do Google nos resultados?</i>	21
1.4	Como este trabalho está organizado	22
2	MARCO TEÓRICO	23
2.1	A escolha do voto e o comportamento político	23
2.1.1	<i>Diferentes teorias para um mesmo indivíduo</i>	24
2.1.1.1	<i>A abordagem sociológica: interações sociais e pertencimento a grupos</i>	24
2.1.1.2	<i>A abordagem psicológica: valores individuais</i>	25
2.1.1.3	<i>A abordagem racional: cálculo da utilidade na participação</i>	26
2.1.2	<i>A mídia e o comportamento político</i>	27
2.2	Comunicação política e informação	30
2.3	Internet e Política	32
2.4	Os estudos sobre o Google	35
2.4.1	<i>O que já se estudou sobre viés no Google</i>	37
2.4.2	<i>O Google e a desinformação</i>	42
2.4.3	<i>O estudo do Google no Brasil</i>	43
3	COMO O GOOGLE FUNCIONA	45
3.1	Como o Google diz que funciona	45
3.2	Como os profissionais de marketing dizem que o Google funciona	48
3.3	O que o Google disse que funciona para se defender	50
3.3.1	<i>Por que o Google é tão melhor que o Bing?</i>	53
3.4	Quando o Google mudou os resultados da busca	53
3.4.1	<i>O Google pode interferir nos resultados?</i>	53
3.4.1.1	<i>Patroa e mulher solteira: mudanças no dicionário</i>	54

3.4.1.2	<i>Google sugere Lula corrupção</i>	55
3.4.1.3	<i>Influência política</i>	56
3.4.1.4	<i>Estratégia controversa sobre pornografia infantil</i>	56
3.5	O desafio de se estudar o Google	57
4	METODOLOGIA	59
4.1	O desenho de pesquisa	59
4.1.1	<i>Painel longitudinal - pesquisas antes e depois das eleições</i>	60
4.1.2	<i>Histórico de navegação (web tracking) - durante o período eleitoral</i>	61
4.1.2.1	<i>Identificação das buscas político-eleitorais</i>	62
4.1.3	<i>Captura dos resultados do Google</i>	65
4.1.3.1	<i>O período pré-eleitoral e eleitoral</i>	66
4.1.3.2	<i>Buscas geolocalizadas</i>	66
4.1.3.3	<i>Os termos buscados</i>	67
4.2	Resumo	71
5	PARA QUEM O GOOGLE É RELEVANTE	72
5.1	O que pode explicar o maior ou menor uso do Google	73
5.2	O que pode explicar - Variáveis Independentes	75
5.2.1	<i>Conhecimento sobre os candidatos e participação</i>	75
5.2.1.1	<i>Nível de conhecimento das candidaturas</i>	75
5.2.1.2	<i>Se sabia em quem votar</i>	76
5.2.1.3	<i>Se votou no 1º turno</i>	77
5.2.2	<i>Preferência partidária</i>	77
5.2.2.1	<i>Partido de preferência</i>	78
5.2.2.2	<i>Partido no qual nunca votaria</i>	78
5.2.3	<i>Variáveis de controle</i>	80
5.3	Como medir o uso do Google - Variáveis Dependentes	81
5.3.1	<i>Uso monitorado do histórico de navegação</i>	82
5.3.2	<i>Uso declarado pelas respostas ao questionário</i>	84
5.3.3	<i>Comparação entre as duas variáveis</i>	85
5.3.3.1	<i>Quando se transforma uma escala ordinal em binária</i>	86
5.3.3.2	<i>Quando os dados desmentem as autodeclarações</i>	87
5.3.4	<i>Definição da variável dependente</i>	88

5.4	Diferentes perfis para se entender o envolvimento do eleitor	89
5.4.1	<i>O uso de análise de correspondência múltipla (ACM)</i>	90
5.5	Teste da hipótese	95
5.5.1	<i>Testes e robustez do modelo</i>	96
5.5.2	<i>Resultados do modelo monitorado</i>	97
5.5.3	<i>Resultados do modelo declarado</i>	99
5.5.4	<i>Comparação entre os modelos</i>	100
5.6	Considerações finais	103
5.6.1	<i>Diferentes perfis sob a ótica do comportamento político</i>	104
6	COMO E QUANDO O GOOGLE É RELEVANTE	107
6.1	Termos de busca enquanto preferência	107
6.1.1	<i>Informações sobre Candidatos/Partidos</i>	108
6.1.2	<i>Processos/Procedimentos Eleitorais</i>	109
6.1.3	<i>Buscas Gerais sobre Eleições</i>	110
6.2	Quantidade média de buscas para cada perfil	113
6.3	Quando o Google importa	115
6.3.1	<i>Quando os usuários buscam</i>	115
6.3.2	<i>Diferentes tipos de buscas em diferentes momentos</i>	117
6.4	Considerações finais	118
7	O QUE O GOOGLE ENTREGOU?	120
7.1	A qual domínio a URL pertence	121
7.1.1	<i>Resultados das buscas com termos políticos</i>	122
7.1.2	<i>Resultados das buscas com termos não-políticos</i>	123
7.2	Resultados diferentes para termos diferentes	125
7.2.1	<i>Resultados para o termo “Eleições 2020”</i>	126
7.2.2	<i>Resultados para os termos “candidatos + cidade”</i>	128
7.2.2.1	<i>Análise da concentração de busca</i>	131
7.2.3	<i>Buscas nas capitais brasileiras</i>	132
7.2.4	<i>Buscas temáticas</i>	134
7.3	Considerações finais	137
8	O GOOGLE INTERFERIU NOS RESULTADOS DE BUSCA?	139
8.1	Abordagem empírica: diferença nas diferenças (DiD)	141

8.1.1	<i>Grupos de tratamento e controle e períodos pré e pós-tratamento</i>	141
8.1.2	<i>Como o método DiD funciona</i>	142
8.1.3	<i>Variáveis dependentes</i>	143
8.1.4	<i>Resultados iniciais do DiD (sem variáveis explicativas)</i>	143
8.1.5	<i>Modelo com variáveis explicativas</i>	146
8.1.6	<i>Aplicando o DiD com variáveis de controle (regressão)</i>	150
8.1.6.1	<i>Estimação do Modelo</i>	151
8.1.6.2	<i>Janela de 7 dias</i>	151
8.1.6.3	<i>Análise dos modelos</i>	153
9	CONCLUSÃO	158
9.1	Contribuições metodológicas e técnicas	158
9.2	Contribuições para o campo do comportamento político	160
9.3	Contribuições para a área de Comunicação, Mídia e Jornalismo	160
9.4	Contribuições para os estudos a respeito do Google	161
9.5	Contribuições para os estudos de internet e política	162
9.6	E se o Google mudar, como ficam essas conclusões?	163
	REFERÊNCIAS	165
10	APÊNDICE METODOLÓGICO	174
10.1	Declaração de Uso de IA generativa e tecnologias assistidas por IA	174
10.1.1	<i>Prompt de correção ortográfica e classificação dos termos</i>	174
10.2	Distribuição Amostral e Tratamento das buscas	176
10.3	Descritivo das buscas realizadas	178
10.3.1	<i>Busca sobre candidatos</i>	179
10.3.2	<i>Busca Gerais</i>	179
10.3.3	<i>Busca não eleitorais</i>	181
10.4	Modelos do Capítulo 5	182
10.5	Questionários	185
10.5.1	<i>Questionário Onda 1: Antes da campanha eleitoral</i>	186
10.5.2	<i>Questionário Onda 2: Depois do 2 turno</i>	203

1 INTRODUÇÃO

A maneira como vivemos está mudando conforme mudamos o nosso modo de acessarmos informação, e as transformações tecnológicas têm sido fundamentais nesse processo. Há duas décadas, uma viagem a turismo seria comprada em empresas especializadas, por meio de pacotes de viagens, confiando ao agente de turismo o sucesso da viagem; hoje, cada vez mais, buscamos por produtores de conteúdo, influenciadores ou sites que podem nos ajudar no tipo de roteiro preferido; em sequência, buscamos por fotos e comentários reais de hotéis ou hospedagens em plataformas e supomos termos a plena informação do lugar, suas redondezas e preços. Tudo isso contribui com uma melhor capacidade de tomarmos boas decisões sobre uma próxima viagem. Estamos substituindo cada vez mais figuras como a do agente de turismo em nome de nossa autonomia de buscar informações e recomendações em alguma plataforma.

Todas essas mudanças, que muitas vezes são observadas apenas pelo apelo tecnológico, modificaram um ponto relevante em nossa relação social: a figura do mediador. Algumas plataformas de tecnologia, como as mídias sociais e diversos aplicativos e sites de serviços, prometem uma comunicação direta, ou seja, sem os mediadores tradicionais como jornais ou canais de televisão. Uma agência de turismo sempre foi uma mediação importante na tomada de decisão entre você e suas férias. O táxi sempre foi uma mediação entre você e o seu destino. A pessoa treinada para atender bem em lojas de departamento, sempre pronta para te recomendar a melhor televisão, sofá ou conjunto de painéis, sempre foi uma mediação entre você e a escolha do melhor produto a ser comprado.

O paradoxo desse processo é que a aparente sensação de menor mediação, provocada por uma melhor gestão da experiência do usuário com o meio tecnológico, acaba por se mostrar uma grande captação dos fluxos de intermediação (DeNardis, 2009; Loveluck, 2015). A promessa da possibilidade de fluxos sem mediação e de colaboração livre foi substituída por uma competição acirrada pela atenção dos usuários e, principalmente, concentrada por um leque limitado de plataformas, sites e aplicativos (Evangelista, 2018). No fundo, o agente de turismo virou uma *big tech* global que deu ao cidadão um campo de busca e uma recomendação de ação. Tudo com a promessa de que o cidadão sempre terá a melhor informação possível sobre o que ele quer fazer.

Se informação é relevante para a vida pública, sua distribuição é peça fundamental, e aqui também há uma mudança importante na mediação. Se antes havia a crença de que os fluxos informacionais poderiam ser descentralizados, fica cada vez mais evidente que essa visão

otimista já não existe mais (Hindman, 2009; Tufekci, 2014). Além do debate sobre mediação e concentração econômica (e.g, Srinivasan e Fish, 2017), há também o debate sobre a opacidade dos algoritmos que regem nosso acesso à informação diariamente (e.g, Mehl e Silva, 2017; O’Neil, 2016; Silva *et al.*, 2020; Silva, 2022) e sobre o papel de uma tecnologia a respeito da qual se pode questionar o paradigma de que é neutra e que simplesmente está fazendo uma otimização do seu tempo (e.g, Feenberg, 2018).

Esta tese investiga, por meio de uma metodologia ainda pouco usual no contexto científico brasileiro e internacional, a relevância das plataformas tecnológicas de mediação de informação durante os períodos eleitorais.

1.1 Por que essa tese é relevante?

Novos hábitos em um velho sistema político

Se a tecnologia está nos acostumando com processos interativos cada vez mais rápidos em nosso modo de nos relacionarmos e se estamos cada vez mais autônomos em nossa maneira de consumir informação para tomar decisões, como essas transformações estão impactando o modo como participamos da vida política? Afinal, as democracias liberais representativas não são, no fundo, um grande mediador entre a população e a necessidade de organização social?

É de se esperar, portanto, que haja algum atrito com o sistema político. Segundo o último relatório do V-Dem (Papada *et al.*, 2023), os níveis globais de democracia voltaram a valores equivalentes aos de 1986. Ainda que não se discuta aqui a causalidade direta entre crises dos sistemas democráticos e tecnologias de informação, é intuitivo pensarmos nesse atrito. Sistemas representativos, que costumam trazer resultados a longo prazo, estão enfrentando um ambiente informacional em transformação, nos quais as expectativas por respostas rápidas e eficientes são cada vez maiores. Essa discrepância pode construir a percepção de lentidão ou insatisfação com o processo democrático, algo particularmente desafiador para sociedades que estão se acostumando com modos de interação cada vez mais imediatos.

A capacidade massiva de produzirmos conteúdo - *mass self-communication* (Castells, 2013) - também criou, por outro lado, a imensa quantidade de dados e informações disponíveis. Temos novas maneiras de enviar e receber mensagens entre nossos contatos, podemos contar a nossa visão sobre o mundo de inúmeras formas, temos novos modos de avaliar e sermos avaliados a respeito de nosso trabalho ou de serviços que contratamos em um aplicativo.

Se todo mundo pode produzir conteúdo, como encontrá-lo?

A abundância de informação só não é um problema ainda maior pois conseguimos encontrar o que aparentemente nos é relevante de maneira rápida e eficiente. Isso se dá por meio dos buscadores, sites especializados em indexar o conteúdo disponível na web e que nos dão a sensação de que rapidamente encontraremos o que estamos procurando.

Esse poder fez do Google um hábito, um verbo. Ele é um dos agentes da capacidade de acessarmos informação velozmente, tornando-nos cada vez mais autônomos em relação à busca informacional. Em 2021, estima-se que o Google teve 1.077 trilhões de visitas (Kemp, 2022) e estima-se que em 2024 possui cerca de 96% de participação no mercado de buscas realizadas em dispositivos móveis e 82% das buscas em “desktops”, no mundo (Chaffey, 2024). Em termos comparativos, estima-se um número de 265,7 bilhões de visitas para o Facebook no período de 2021 (Idem). Ou seja, 4 vezes menos que o Google.

O mediador da humanidade

Em que pese sua importância, a relação entre o Google e a política tem sido pouco estudada. As novas tecnologias digitais têm sido objeto de estudo de diversos pesquisadores nas áreas sociais e de comunicação, mas essa crescente literatura tem se concentrado no estudo de plataformas de redes sociais como Facebook ou Twitter.

Os poucos estudos acadêmicos brasileiros sobre os possíveis impactos do Google foram produzidos pelo grupo Grupo de Pesquisa sobre Relações entre Sociedade e Estado (RESOCIE) da Universidade de Brasília (Vilarins L.; Stabile *et al.*, 2021; STABILE, 2021), do qual esta tese faz parte. Esses estudos apontam para a urgência de se analisar de maneira mais aprofundada e rigorosa o papel do mais importante mediador digital de informação política, além disso, mostram indícios de que, pelo menos em algumas circunstâncias, o Google pode ter impactos importantes na vida política em geral e em processos eleitorais em particular.

O Brasil foi escolhido como estudo de caso. Atrás apenas de EUA, Índia e Filipinas, o Brasil é a quarta maior democracia do mundo. Em termos de proporção de acesso à internet em 2021, o Brasil possui 90% dos domicílios conectados (IBGE, 2021). A literatura que pesquisa Google e eleições é quase exclusivamente focada nas eleições nos Estados Unidos ou norte global. Trazer o debate para outras democracias e estudá-las de modo empírico é importante, pois são espaços regulatórios diferentes, que podem moldar diferentes papéis de atuação para companhias globais como o Google.

A escolha das Eleições de 2020 como objeto de análise também é relevante, na medida em que a pandemia mundial de Covid-19, no começo daquele ano, com suas medidas

restritivas, acelerou diversos processos digitais e levou diversas campanhas a reforçarem suas presenças digitais.

1.2 Objeto de pesquisa

Este estudo se debruça sobre o papel desempenhado pelo motor de busca Google em contextos eleitorais. O Google transcende a sua função de mero repositório de informações globais, posicionando-se como ator capaz de influenciar substancialmente a disseminação e o acesso a informação.

O escopo desta pesquisa incide sobre o processo de mediação informativa realizado pelo Google durante as eleições subnacionais no Brasil em 2020. A coleta de dados teve início 45 dias antes do começo formal da campanha eleitoral, abrangendo, portanto, o período de pré-campanha. Esse processo prosseguiu ao longo da campanha oficial, que começou em 27 de setembro, e se estendeu para além da apuração dos votos do primeiro turno, em 15 de novembro. Nessa data-chave, 147 milhões de brasileiros foram às urnas para escolher seus representantes municipais em 5569 municípios. A fase de coleta de dados prosseguiu até o segundo turno, ocorrido em 29 de novembro, que envolveu os municípios com eleitorado superior a 200 mil habitantes.

1.3 Perguntas de pesquisa

A partir do amplo debate teórico sobre informação e democracia, esta tese investiga a seguinte questão de pesquisa: **qual o papel dos mediadores digitais no processo de aquisição de informações durante o ciclo eleitoral?**

Com um enfoque empírico nos resultados de pesquisa exibidos pelo Google nas eleições subnacionais brasileiras de 2020, esta análise é enriquecida por um painel longitudinal de participantes que responderam questionários antes e depois do período eleitoral. Uma parte relevante desses participantes também autorizou o monitoramento de suas atividades on-line em computadores e celulares ao longo do período.

A partir desses dados, esta pesquisa responde a quatro questões empíricas relevantes que guiam a análise:

1.3.1 Pergunta 1 - Para quem o Google é relevante?

Estudos anteriores a este (Epstein e Robertson, 2015; Epstein *et al.*, 2017) mostraram que um possível viés nos resultados do Google pode influenciar na decisão de voto do eleitor. Tratando-se do maior mediador informacional da atualidade, esse diagnóstico é preocupante por diversas razões. Entretanto, nenhum desses estudos se propõe a medir **para quem**, de fato, o uso do Google durante o período eleitoral é relevante. Essa é, portanto, a primeira pergunta deste trabalho, a partir da qual propomos a seguinte hipótese:

Hipótese 1: Eleitores com preferência eleitoral consolidada tendem a utilizar **mais** frequentemente o Google para buscar informações sobre eleições.

Essa hipótese está ancorada na teoria de *motivated reasoning* (Lodge e Taber, 2013), a qual afirma que as pessoas que mais buscam informação política são aquelas já engajadas e com a sua decisão política já formada. O ato de interagir e buscar por mais informações políticas, portanto, serve para reafirmar suas convicções ou buscar fontes para ajudar no trabalho de convencimento de outras pessoas.

1.3.2 Pergunta 2 - Como e quando os eleitores utilizam o Google?

A segunda pergunta de pesquisa busca entender **como** os eleitores utilizam o Google, e o faz por meio dos seus termos de busca. O Google é apenas uma referência para informações procedimentais do ato de votar, uma fonte de informação sobre os candidatos ou apenas uma fonte para encontrar os resultados das apurações? Para além de entender como, responder **quando** os eleitores buscam informações desenha um quadro importante e relevante sobre o uso do mediador informacional no ciclo eleitoral brasileiro.

Para essas duas perguntas, não foi desenvolvida nenhuma hipótese específica. Apresentamos uma abordagem exploratória dos dados com o intuito de entender, de maneira qualitativa, os termos de busca utilizados pelos usuários.

Ao responder essas duas perguntas: como e quando, contribuimos com parte da literatura de comportamento político a respeito de como diferentes perfis de eleitores buscam por informação (e se buscam) por meio de um instrumento de observação monitorado, como o histórico de buscas dos usuários.

1.3.3 Pergunta 3 - O que o Google entregou de resultados em 2020?

A terceira pergunta busca entender **o que** o Google exibiu para quem fez buscas político-eleitorais ao longo da campanha eleitoral de 2020, ano em que o termo “eleições 2020” foi eleito pelo próprio Google como um dos termos do ano, perdendo apenas para coronavírus e auxílio emergencial.

A análise é feita com dados de **search engine results page** (SERP) de mais de 500 mil buscas ao longo do período. A análise mostra em que medida os resultados estão concentrados em poucos portais e quais são esses sites.

Esta pergunta também é respondida de maneira exploratória, com o objetivo de compreender o perfil desses sites e ajudar a entender o que os eleitores viram ao pesquisar por termos político-eleitorais durante a campanha.

1.3.4 Pergunta 4 - Há interferência do Google nos resultados?

A quarta e última pergunta é se há uma interferência do Google nos resultados de buscas políticas ao longo da campanha eleitoral. Essa interferência é medida por meio de um quase-experimento desenhado para aferirmos se houve uma mudança de padrão nas entregas antes e depois do início da campanha eleitoral.

Para essa pergunta específica, formulamos a seguinte hipótese:

Hipótese 2 Há uma mudança na quantidade de diferentes domínios entregues pelo Google nos resultados entregues das buscas político-eleitorais ao se iniciar a campanha eleitoral.

Diferentemente das pesquisas realizadas por (Robertson *et al.*, 2018), que tiveram como objetivo buscar por alterações de vieses políticos nos resultados, essa hipótese busca medir uma mudança de padrão na entrega. Essa mudança de padrão considera uma alteração na concentração ou dispersão de diferentes sites nos resultados orgânicos de busca.

Assume-se essa hipótese de interferência porque representantes do Google, quando confrontados, admitiram mudar o que era entregue em determinados termos de busca (Poder360, 2023; Bergamo, 2023; Terra, 2020). Além disso, há um esforço ativo do Google em se posicionar como fonte de informação política ao participar de debates eleitorais e desenvolver campanhas publicitárias para que o eleitor utilize o Google. O esforço de se posicionar como uma fonte crível o tornará um objeto de menor discussão e envolvimento em medidas regulatórias.

1.4 Como este trabalho está organizado

Esta tese está organizada em sete capítulos, além desta introdução e da conclusão.

De maneira sumária, os conteúdos dos capítulos são os seguintes:

- **Capítulo 2 - Marco teórico:** Situa a tese nos debates teóricos da comunicação política e do comportamento político. Também revista a literatura específica sobre os impactos do Google.;
- **Capítulo 3 - Como o Google funciona:** Apresenta uma análise sobre como o buscador funciona, e uma breve revisão das polêmicas recentes envolvendo o Google, a indústria do marketing e a justiça dos Estados Unidos.;
- **Capítulo 4 - Metodologia:** Este capítulo apresenta o desenho de pesquisa, as fontes de pesquisa, os instrumentos de coleta e as estratégias de análise dos dados. Os apêndices metodológicos da tese complementam este capítulo;
- **Capítulo 5 - Para quem o Google é relevante:** Investiga as variáveis que influenciam o uso do Google, utilizando a análise fatorial para a criação de perfis de comportamento dos usuários e testando a hipótese por meio de uma regressão logística;
- **Capítulo 6 - Como e quando o Google é relevante:** Explora quais termos de busca os diferentes perfis identificados no capítulo anterior utilizaram e em qual momento da corrida eleitoral cada tipo de busca foi feita;
- **Capítulo 7 - O que o Google entregou em 2020:** Este capítulo analisa os resultados de busca dos termos políticos, abordando de maneira exploratória aspectos de concentração de mídia e dos principais sites que foram exibidos nas buscas;
- **Capítulo 8 - O Google interferiu nos resultados de busca?:** Este último capítulo de apresentação dos resultados da pesquisa descreve o quase-experimento realizado e, a partir dos dados coletados, discute a questão da interferência do Google nos resultados de busca ao longo do período analisado.
- **Conclusão:** A conclusão recapitula as principais contribuições desta tese para diferentes áreas do conhecimento e proporemos a criação de um observatório do Google nas eleições, além de temas para próximas pesquisas.

2 MARCO TEÓRICO

Este capítulo é dedicado ao estabelecimento do marco teórico que fundamenta a análise proposta nesta tese, a qual visa elucidar o papel e a relevância do Google em contextos eleitorais.

A partir da revisão da literatura teórica sobre comportamento político e comunicação política, selecionamos conceitos e estudos que nos proporcionassem visões específicas sobre como a informação importa e como eleitores tomam suas decisões.

Por meio dessa abordagem, pretendemos não apenas entender o papel do Google no vasto campo da comunicação política e do comportamento eleitoral, mas também destacar como a evolução tecnológica e a disseminação de informação digital transformam as práticas eleitorais contemporâneas. Este marco teórico serve, portanto, como alicerce para uma análise mais aprofundada e específica sobre o impacto do Google em contextos eleitorais, contribuindo para uma compreensão mais ampla das dinâmicas atuais de informação e influência política.

2.1 A escolha do voto e o comportamento político

Um dos objetivos desta tese é compreender para quem e de que maneira o Google se mostra significativo durante o processo eleitoral. Para alcançar esse objetivo, torna-se imprescindível mergulharmos nos estudos de comportamento político que visam entender o eleitor e suas motivações. A literatura de comportamento político é vasta e aborda as motivações do voto sob diversas perspectivas a partir das quais podemos examinar a decisão eleitoral. Reconhecer e integrar essas diferentes perspectivas é crucial para contribuirmos com a discussão sobre o papel das ferramentas mediadoras de informação, como o Google, no contexto eleitoral. A complexidade do comportamento do eleitor e a multiplicidade de fatores que influenciam sua decisão de voto exigem uma abordagem que considere a interseção dessas variadas teorias.

No capítulo 5, a análise da interação entre três variáveis em um modelo multivariado revelou-se particularmente elucidativa, superando a utilidade de abordagens que consideram essas variáveis de maneira isolada. As interações identificadas resultaram na definição de cinco perfis eleitorais distintos, cujas características podem ser amplamente interpretadas à luz da literatura consolidada na área do comportamento político. Esse achado sublinha a importância de uma abordagem integrada, que reconhece a interdependência entre diferentes fatores na determinação do comportamento eleitoral.

2.1.1 Diferentes teorias para um mesmo indivíduo

O campo do comportamento político se consolidou na Ciência Política a partir de discussões sobre a escolha do voto enquanto elemento central na participação democrática. Três vertentes teóricas principais têm sido amplamente debatidas pela literatura: a corrente sociológica, a teoria psicológica e a abordagem da escolha racional. Figueiredo (2008) ressalta que, embora todas elas se concentrassem em analisar o mesmo indivíduo sob a perspectiva do individualismo metodológico, cada uma adotou variáveis diferentes como primariamente explicativas para o comportamento observado. A Teoria Sociológica prioriza aspectos socioeconômicos e de classe na escolha do voto, considerando também aspectos culturais, regionais e religiosos como determinantes para as predisposições políticas de grupos e indivíduos. A Teoria Psicológica, por sua vez, examina como as características pessoais e a motivação emocional influenciam a decisão de voto, destacando a importância da lealdade partidária e da identificação de grupo. Por fim, a Teoria da Escolha Racional tem origens econômicas e aborda a decisão do voto como um processo de análise de custo-benefício, em que os eleitores avaliam propostas para determinar qual opção atende melhor aos seus interesses. Essas teorias, apresentadas a seguir de maneira bastante resumida, constituem a base para entendermos as variadas perspectivas acerca daquilo que molda a escolha do voto, tema que será explorado mais detalhadamente neste capítulo.

2.1.1.1 A abordagem sociológica: interações sociais e pertencimento a grupos

A corrente sociológica, com raízes na Universidade de Columbia e ilustrada nos trabalhos de Lazarsfeld *et al.* (1948) e Berelson *et al.* (1954), procura estabelecer padrões de distribuição de votos e comportamentos eleitorais vinculados a fatores sociais.

No contexto estadunidense dos anos 50 e 60, esses estudos argumentaram que um nível socioeconômico mais alto estava associado a uma maior propensão ao voto republicano, enquanto a religião emergia como um fator significativo, com católicos tendendo majoritariamente ao voto democrata, independentemente do status socioeconômico. Além disso, identificaram-se maiores proporções de republicanos em áreas rurais e entre mulheres, algo que nos últimos tempos, com Donald Trump, não é mais o mesmo.

Educação, religião e redes de socialização

Essa abordagem argumenta que votar é primariamente uma experiência coletiva, na qual indivíduos que compartilham o trabalho, o convívio ou o lazer tendem a votar de modo

similar. Assim, grupos com predisposições políticas semelhantes tendem a se associar, levando-os a uma uniformidade de opiniões e comportamentos políticos. Nos estudos contemporâneos brasileiros de comportamento eleitoral, o foco tem sido a relação do voto com variáveis como educação, influência midiática, religião e redes de socialização. Essas pesquisas buscam decifrar como fatores sociais influenciam a participação e preferência de voto e o papel das conexões sociais como estruturas de constrangimento ou apoio eleitoral. Tais estudos se alinham à teoria sociológica ao examinar como as interações sociais e o pertencimento a determinados grupos impactam as decisões eleitorais, refletindo a continuidade da influência das redes sociais sobre a escolha do voto (e.g, Carreirão, 2004; Bezerra e Mundim, 2011; Rennó, 2004).

2.1.1.2 *A abordagem psicológica: valores individuais*

Essa visão sobre a escolha do voto foi problematizada pela Escola de Michigan, fundadora da corrente psicológica do comportamento político. Em “*The American Voter*”, Campbell *et al.* (1960) argumentavam que a posição social, por si só, não era suficiente para decifrar o comportamento do eleitor, enfatizando que fatores psicológicos e sociais interagem para moldar e alterar atitudes políticas. Para eles, as inclinações políticas emergem tanto do contexto social como do histórico psicológico dos indivíduos (Figueiredo, 1990). Em busca de maior precisão, advogavam o uso de pesquisas de opinião para investigar a relação entre atributos pessoais e preferências políticas. Diferentemente da teoria sociológica, que examina o coletivo eleitoral, a corrente psicossocial foca nas composições individuais. Os proponentes dessa corrente reconhecem a importância do contexto social, mas o consideram apenas parte da equação que define a decisão de voto, salientando a interação entre valores individuais e influências sociais.

Identificação partidária importa

Dentro das interpretações oferecidas pela corrente psicológica, percebe-se a tendência dos eleitores de pautarem seus votos em uma identificação partidária, que frequentemente é transmitida pelo ambiente familiar ou durante a socialização. Os valores internalizados ao longo do processo de socialização exercem influência significativa nas escolhas políticas, implicando no direcionamento do voto para o partido que mais se alinha a esses valores. Segundo essa perspectiva, a ideologia não é entendida como o fator dominante na tomada de decisão do eleitor. Em vez disso, enfatiza-se o conceito de “identificação partidária”, que pode não ter vínculos diretos com ideologias específicas (Campbell *et al.*, 1960).

A maioria tende a tomar decisões inconstantes

Converse (2006), um expoente da teoria psicossocial, sustentava que as pessoas interagem com seu ambiente político com base em sua capacidade de conceitualização sobre ele. A instabilidade das opiniões frente a temas políticos relevantes desafia a capacidade preditiva da teoria. O pesquisador observou que apenas uma minoria politizada mantém crenças bem estruturadas que influenciam suas decisões de voto a longo prazo; a maioria dos eleitores tende a tomar decisões inconstantes. Para ele, a educação é um preditor da sofisticação conceitual dos eleitores, mas não assegura a coesão de um sistema de crenças, que também depende das motivações individuais para com a política.

2.1.1.3 A abordagem racional: cálculo da utilidade na participação

Assim como as duas perspectivas teóricas já apresentadas, a teoria da escolha racional também nasceu a partir de uma inspiração em outra área do conhecimento e foi incorporada ao contexto da Ciência Política. Também chamada de abordagem econômica do voto, essa corrente foi adaptada da teoria econômica para estudos eleitorais, fruto da tradição econômica da Escola de Chicago. A escolha racional tem como texto basilar o livro *Uma Teoria Econômica de Democracia*, de Downs (1957), que dialogou com a democracia procedimentalista de Schumpeter para retratar os eleitores como racionais e capazes de calcular a utilidade na participação eleitoral e no direcionamento do voto. Nessa perspectiva, o cálculo utilitário do voto é uma maneira de reforçar a legitimidade democrática em si. Downs introduz o conceito de **ideologia** como um mecanismo de simplificação do mundo, possibilitando a otimização de ganhos e a minimização de custos por parte do eleitor-consumidor.

Capacidade de fazer a avaliação retrospectiva e prospectiva

A escolha racional tem foco especial na participação eleitoral. Em termos de orientação do voto, consolidou-se em estudos de *issue voting*, que enfatizam a habilidade do eleitor em medir as distâncias entre suas próprias posições e as dos candidatos em relação a temas-chave (e.g, Fournier *et al.*, 2003). Estudos mais recentes da teoria racional incluem o conceito do voto econômico, que se baseia na avaliação retrospectiva e prospectiva dos governantes, considerando-se a situação econômica individual percebida pelo eleitor (Lewis-Beck *et al.*, 2011). Essa perspectiva foi incorporada no contexto das eleições presidenciais brasileiras, com análises que abordam variáveis explicativas como a ideologia (Carreirão, 2002), a avaliação governamental (Peixoto e Rennó, 2011) e a influência de programas sociais na decisão de voto

(e.g, Licio *et al.*, 2015).

Considerando-se o foco da pesquisa na relevância do provimento de informação e seus impactos nas preferências eleitorais, a próxima seção discutirá como cada abordagem teórica trata o tema, tendo em conta, particularmente, o conteúdo midiático ao qual os eleitores são expostos.

2.1.2 A mídia e o comportamento político

Para a teoria sociológica, há um **nível reduzido de influência** que conteúdos midiáticos podem exercer sobre a decisão de voto do eleitorado. Isso porque os fatores considerados mais determinantes para a predisposição política de cada indivíduo são dados de maneira relativamente estática: sua classe, religião ou cultura. Nessa perspectiva, a análise concentra-se mais em identificar padrões de interação entre eleitores e meios de comunicação do que em medir o impacto direto da mídia. A exposição constante à mídia pode elevar o interesse político dos indivíduos sem necessariamente alterar suas posturas pré-existentes. Esse fenômeno é conhecido como efeito de ativação e reforço, indicando que as campanhas políticas podem despertar o interesse político e intensificar as convicções existentes dentro dos grupos, ao mesmo tempo em que esses grupos tendem a rejeitar informações que contrariem seus interesses.

Eleitores participantes e críticos

Questionando a noção de que a mídia e a propaganda poderiam facilmente moldar as opiniões, a sociologia política propõe uma visão na qual o eleitor é tido como um agente ativo. Nesse contexto, o poder da mídia para manipular crenças é limitado, e os eleitores são retratados como participantes engajados e críticos dos discursos veiculados pela mídia (e.g, TELLES *et al.*, 2009).

Pleitos normais ou desviantes

Para compreendermos mais profundamente a relação entre o eleitor e a mídia na teoria sociológica, é importante distinguirmos dois tipos de eleições existentes nessa teoria: normais e desviantes. Nas eleições normais, os acontecimentos conjunturais exercem influência limitada na escolha do voto. Por outro lado, nas eleições desviantes, os elementos conjunturais presentes no debate político podem ter impactos significativos nas decisões dos eleitores, inclusive na decisão de votar ou não. Fatores como alienação e imitação são considerados pela literatura como determinantes para explicar comportamentos eleitorais desviantes, superando a identidade partidária ou lealdade a partidos. Cervi (2012) explica que a alienação em comportamentos

eleitorais desviantes ocorre quando eleitores de um grupo social são influenciados por discursos políticos que não refletem as demandas sociais de seu grupo original. Em contrapartida, o princípio da imitação descreve como, mesmo sem manipulação explícita, eleitores podem votar seguindo o padrão de grupos sociais ideologicamente dominantes. Castro (1994) aponta que o desinteresse dos eleitores nas disputas eleitorais tende a aumentar a incidência de casos desviantes motivados por imitação ou alienação.

Alienação enquanto afastamento da política

A Escola de Michigan também explorou o conceito de alienação política, porém de maneira conceitualmente distinta da corrente sociológica. Aqui, a alienação política não significa que o eleitor está apartado de demandas do grupo ao qual pertence porque está envolvido com discursos políticos alheios. Pelo contrário, a alienação é entendida por essa corrente como um afastamento da política como um todo. Entre os fatores que podem gerar esse desinteresse pela política estão a sensação de impotência do eleitor comum em influenciar o governo; a percepção de imprevisibilidade nas decisões políticas; a anomalia gerada quando as regras aceitas são violadas pela elite política e a rejeição às normas políticas por indivíduos isolados da sociedade. A premissa da teoria psicossocial é que, ao compreendermos os níveis de envolvimento ou alienação e a capacidade cognitiva dos indivíduos, é possível anteciparmos o comportamento dos grupos sociais.

A maior parte dos eleitores não se envolve com questões políticas

Essa perspectiva reconhece que o acesso à informação pode modificar a interpretação da realidade política e, conseqüentemente, influenciar a escolha eleitoral. Ainda assim, a Escola de Michigan apresenta uma visão que relativiza a influência da campanha e da mídia nas decisões eleitorais. Primeiramente, parte do pressuposto de que a maior parte do eleitorado não se envolve profundamente com questões políticas; em segundo lugar, considera que os poucos cidadãos politicamente ativos, que se mantêm informados pela imprensa, já possuem um alinhamento partidário definido que orienta suas ações. Portanto, para esses eleitores engajados, o contato constante com os meios de comunicação serve menos para mudar a intenção de voto e mais para consolidar o comprometimento com o partido com o qual já se identificam, um fenômeno que ecoa a percepção da sociologia política sobre o efeito de reforço da identidade partidária (TELLES *et al.*, 2009).

Raciocínio motivado e o foco seletivo em informações

Para melhor compreendermos como ocorre esse reforço, é importante discutirmos

o conceito de *motivated reasoning*, ou raciocínio motivado. Conforme explorado na teoria psicológica do voto, ele é o processo pelo qual um indivíduo busca e processa informações com o objetivo de formar uma atitude, sendo guiado não apenas por uma busca de precisão da informação, mas também por motivações direcionadas. Essas motivações podem ser o desejo de persuasão, a concordância com um grupo, a manutenção de uma imagem ou preservação de si mesmo, o que pode levar a um foco seletivo em informações que confirmem crenças existentes e à rejeição de dados contrários, independentemente da sua veracidade.

Conforme apontam Taber e Lodge (2006), uma das principais motivações primárias no processo de formação de opinião são as metas direcionais, que surgem quando uma pessoa está motivada a chegar a uma conclusão que se alinhe às suas predisposições, como a identificação com um partido político. Quando impulsionados por essas metas, os indivíduos tendem a valorizar mais as informações que corroboram suas crenças ou identidades sociais existentes e a desconsiderar ou refutar as que são contraditórias. Esse tipo de raciocínio motivado pode resultar em uma série de vieses cognitivos, como o viés de confirmação, o viés de desconfirmação e o efeito de atitude prévia, fazendo com que as pessoas vejam as evidências que confirmam suas opiniões anteriores como mais convincentes.

A identidade partidária é um fator crítico na formação da opinião pública para a perspectiva psicológica, e o raciocínio direcional muitas vezes é impulsionado pelo desejo do indivíduo de permanecer leal e coerente com seu partido, acentuando as diferenças com o partido adversário. Isso se torna evidente quando políticas são julgadas não pelo seu mérito, mas pela afiliação partidária de quem as propõe. Endossos partidários, por exemplo, podem influenciar significativamente as atitudes em relação a questões como aborto e redução da maioria penal, estimulando o raciocínio motivado partidário, especialmente em contextos nos quais há polarização política. Esse raciocínio motivado pode levar à negligência de informações relevantes que poderiam ser benéficas, mas que são descartadas por não se alinharem com as convicções prévias do indivíduo.

Informação ajuda a diminuir as incertezas

Esse tipo de raciocínio não é abarcado no contexto da teoria da escolha racional, para a qual os eleitores fazem suas escolhas baseados na avaliação dos benefícios esperados de cada candidato, seja ele da situação ou da oposição. A informação desempenha um papel crucial nesse processo, pois ajuda a diminuir as incertezas e serve como um guia para decisões conscientes. Diferentemente das abordagens anteriores, a escolha racional reconhece o valor das

campanhas políticas e dos meios de comunicação como fontes vitais de informação e veículos para a propagação de ideias, que os eleitores utilizam para fundamentar suas decisões (e.g, TELLES *et al.*, 2009).

Ao longo da segunda metade do século XX, a dinâmica política sofreu transformações significativas, alterando o papel de diversos atores nas democracias ocidentais, incluindo o papel dos meios de comunicação e das campanhas eleitorais. Uma observação notável nessa nova dinâmica é a redução dos eleitores que declaram lealdade partidária ou se envolvem com partidos políticos diretamente (Mainwaring, 2001). Paralelamente, observa-se um aumento no número de eleitores que começam a refletir sobre política e a tomar suas decisões de voto durante o período das campanhas nos meios de comunicação. Diante dessas mudanças, fatores conjunturais, ou de curto prazo, tornam-se relevantes para explicarmos o comportamento do voto, especialmente em sociedades latino-americanas com democracias em consolidação e partidos políticos menos institucionalizados. Nessas sociedades, é frequente a desconfiança nas organizações formais de representação política, superando até mesmo a desconfiança nas instituições políticas tradicionais.

A combinação de baixa identificação partidária com alta volatilidade do eleitorado amplia a relevância das campanhas televisivas nas estratégias partidárias, refletindo uma realidade em que as campanhas e os meios de comunicação tradicionais e digitais assumem um papel cada vez mais decisivo na influência das escolhas eleitorais.

2.2 Comunicação política e informação

O argumento de que cidadãos bem-informados são essenciais ao bom funcionamento do regime democrático já é amplamente aceito na literatura (e.g, Blassnig *et al.*, 2023; PRIOR, 2007). Pessoas mais instruídas tendem a ser mais tolerantes e mais consistentes com suas opiniões e atitudes políticas (PRIOR, 2007). Além disso, há quem defenda que a desigualdade informacional reforça desigualdades socioeconômicas, uma vez que, além de políticos prestarem mais atenção às demandas dos politicamente envolvidos, a falta de conhecimento pode fazer com que os cidadãos apoiem políticas prejudiciais a seus interesses (Carpini e Keeter, 1996).

A partir desses argumentos, alguns pesquisadores têm buscado compreender quais são os efeitos do atual cenário midiático digital, que oferece ao público múltiplas escolhas de conteúdo. Antes da expansão da TV por assinatura, da internet e dos serviços de streaming, as pessoas escolhiam o que assistir entre as poucas opções oferecidas pelos canais abertos.

Consequentemente, era comum que todos tivessem algum tipo de contato com os noticiários políticos, mesmo quando não se interessavam muito pelo tema. Hoje, com a multiplicidade de conteúdos disponibilizados pelo ambiente midiático, o poder de escolha das pessoas é maior e assim como os politicamente interessados podem consumir notícias a qualquer hora do dia, os não interessados podem desprezar esse tipo de conteúdo (PRIOR, 2007). No contexto brasileiro, por exemplo, durante décadas a população se informou pela TV aberta – especialmente pelos jornais da Rede Globo, que sempre manteve a maior audiência e o maior faturamento em comparação com os outros canais (Becker *et al.*, 2015). No entanto, desde o ano 2000, a TV brasileira vem perdendo audiência, principalmente entre os mais jovens Castro (2023).

De acordo com PRIOR (2007), a transição de um ambiente midiático de baixa escolha para um ambiente de alta escolha aumenta a lacuna de conhecimento entre os cidadãos, que antes era reduzida pela televisão aberta. Isso porque as pessoas que gostam de notícias aprendem mais sobre política em um ambiente de mídia de alta escolha, enquanto o conhecimento político cai para aqueles que preferem produtos de entretenimento. O autor comprova sua perspectiva por meio de uma pesquisa realizada entre 2002 e 2003 com uma amostra aleatória de residentes dos EUA. O trabalho trata da expansão da internet, mas foca na popularização da TV por assinatura.

PRIOR (2007) mostra que à medida que o número de casas com TV a cabo aumenta, o público de notícias dos canais abertos diminui, e isso não quer dizer que as pessoas ficaram mais desinteressadas em relação à política. *“Interest in politics may never have been as high as audience shares for evening news led us to believe. There was simply nothing else to watch”* (p. 135). Isto é, a TV aberta tem a capacidade de levar notícias a segmentos menos instruídos e mais desinteressados. Ademais, PRIOR (2007) afirma que, em contextos em que o voto não é obrigatório, como nos EUA, o acesso à TV a cabo e à internet se tornaram importantes preditores do comparecimento eleitoral. Foi o que mostrou a sua pesquisa realizada no início do século: um ávido pesquisador de notícias tinha quase duas vezes mais chances de ir às urnas do que um fã de entretenimento, quando ambos tinham acesso à internet e à TV por assinatura.

BAUM (2003) apresenta uma perspectiva um tanto divergente sobre os novos ambientes midiáticos e introduz à discussão uma nova variável: o aumento de programas classificados como infoentretenimento (ou *infoentertainment*). No livro *Soft news goes to war*, ele defende a tese de que a expansão do infoentretenimento fez com que o público norte-americano prestasse mais atenção a assuntos políticos. Em primeiro lugar, é importante esclarecer que o autor compreende que a realização de qualquer ação política – como ir votar ou assistir a um debate

eleitoral – pressupõe um cálculo do cidadão a respeito dos seus custos e dos seus benefícios. Em outras palavras, ele defende que o consumo de informações políticas também pode ser explicado com base na teoria da escolha racional (Downs, 1999), e mesmo que muitas pessoas possuam conhecimento político limitado, elas são capazes de tomar decisões racionais de maneira heurística, utilizando atalhos cognitivos, como experiências de vida e sugestões de líderes de opinião. O que o infoentretenimento causa, segundo BAUM (2003), é a redução dos custos da obtenção de informações. Assistindo a uma série ou a um *talk show*, por exemplo, as pessoas podem ter acesso a informações que, até a década de 1980, estavam restritas aos noticiários. Para comprovar a sua teoria, ele mostra que, de acordo com as pesquisas de opinião, os norte-americanos se preocuparam mais com a Guerra do Vietnã (1955-1975) do que com a Guerra do Golfo (1990-1991). No entanto, mais pessoas estavam dispostas a expressar uma opinião durante a crise do Golfo Pérsico – que, ao contrário do conflito no Vietnã, foi tema de programas de entretenimento. *“Political information might thus become a free bonus, or an incidental by-product, of paying attention to entertainment-oriented information”* (p. 30).

Esses dois trabalhos precursores – o livro de Markus Prior e o de Matthew Baum – são relevantes porque discutem pontos fundamentais ao tema desta tese: a maneira como as pessoas consomem informações vem se modificando e isso gera efeitos em suas atitudes políticas. Assim como os programas de infoentretenimento estudados por BAUM (2003) reduzem os custos da obtenção de informações, a internet também tem o potencial de fazê-lo. Afinal, sem sair de casa e em poucos cliques as pessoas podem encontrar as informações e os conteúdos que procuram (Roberts, 2006).

2.3 Internet e Política

Para aprofundarmos a discussão sobre os impactos das mudanças no acesso à informação e às fontes de informação no novo ambiente informacional, é crucial discutirmos como essas mudanças são mediadas por fatores além do controle do eleitor. A noção de intermediação, conforme abordada por Loveluck (2015), revela que, apesar da internet ter sido inicialmente celebrada por sua capacidade de diminuir os intermediários entre a informação e o usuário, acabou por facilitar a ascensão de novos intermediadores. Esses novos agentes, principalmente grandes corporações tecnológicas, organizam e controlam os fluxos de informação, exercendo uma influência significativa sobre o conteúdo visto e consumido. Loveluck descreve um modelo de captação que coloca empresas como o Google no centro desse ecossistema informacional,

com algoritmos gerenciando o vasto mar de dados on-line e decidindo o que ganha visibilidade.

A discussão sobre a mudança de paradigmas na disseminação da informação na era da internet deve considerar a perspectiva de Lessig (1999) sobre a infraestrutura da internet. Lessig enfatiza que a regulação da internet não se dá apenas por meio de leis e normas sociais, mas também com as escolhas de design e do código de software que fundamentam sua arquitetura. Essas escolhas determinam as possibilidades e as restrições no ambiente on-line. A arquitetura tecnológica da internet, incluindo protocolos de rede e o código HTML, oferece uma liberdade de conexão que, na prática, é limitada pelos motores de busca que exercem um papel significativo de *gatekeeping*, orientando e restringindo o comportamento de busca dos usuários. A liberdade teórica dos hiperlinks é contraposta por processos sociais que os distribuem em padrões de *winner take all*, concentrando atenção e tráfego em poucos sites.

Hindman Hindman (2009) complementa essa discussão ao afirmar que, contrariando as promessas iniciais de um espaço digital aberto e democrático, a internet reproduziu e até intensificou as práticas de mercado existentes, destacando a persistência de portões e porteiros digitais (*digital gatekeepers*) que filtram a informação. Essa filtragem não elimina a exclusividade na vida política, mas altera a barreira da exclusividade da produção para a filtragem da informação política.

A literatura crítica sobre a internet aprofunda essa questão ao analisar como o capitalismo de vigilância e os novos corretores de poder utilizam as informações para persuasão e engenharia social (Zuboff, 2015). Ela alerta para as práticas de coleta de dados em massa e a modelagem computacional sofisticada que permitem campanhas eficazes e obscuras de influência (Tufekci, 2014). A capacidade de captar, analisar e utilizar grandes volumes de dados pessoais para moldar o comportamento humano representa um desafio para as normas democráticas e levanta questões sobre a autonomia e a privacidade. O’Neil (2016) também chama a atenção para o poder dos algoritmos nas plataformas on-line, apontando-os como “armas de destruição matemática”. Ela alerta para o risco de características nocivas serem “inoculadas” nos algoritmos, que podem se espalhar como um vírus, ampliando preconceitos e danos sociais em larga escala. Evangelista (2018) vai além, argumentando que, apesar da internet ter incorporado princípios da cultura hacker, abandonou os ideais de liberdade em seu desenvolvimento. As plataformas atuais, segundo ele, promovem um capitalismo de vigilância, no qual a informação é controlada por aqueles que já detêm poder, seja a velha mídia ou os novos gigantes tecnológicos.

Portanto, se antes a escolha da informação que chegava até as pessoas era limitada

por meios de comunicação de massa, hoje a internet possibilita um acesso mais amplo, mas não necessariamente mais livre ou democrático. O papel dos algoritmos e das grandes corporações na internet não é apenas um fenômeno de mercado, mas também uma consequência direta da estrutura tecnológica subjacente. Essa estrutura, embora possa parecer neutra, na verdade tem o poder de influenciar profundamente quais vozes são ouvidas e quais informações são acessadas. Os algoritmos e as plataformas digitais têm um poder sem precedentes de filtrar e direcionar a informação, o que destaca a necessidade de uma compreensão mais crítica desses novos intermediários da era digital.

Se na literatura de comportamento político há clareza de que a informação importa e que pode influenciar como as pessoas votam, a discussão sobre estrutura e algoritmos nas plataformas se tornam ainda mais relevante. A democratização da informação na internet, especialmente no que diz respeito aos motores de busca como o Google, é uma promessa ainda em xeque. Essa promessa está sujeita às forças de mercado e aos poderes consolidados que moldam o comportamento individual e coletivo em um ambiente cada vez mais mediado por algoritmos e interesses corporativos.

Além disso, a posição do Google como um dos principais *gatekeepers* na era digital tem atraído a atenção de governos e organizações em todo o mundo, preocupados com as implicações dessa intermediação para a democracia e a liberdade de informação. O relatório apresentado pelo Comitê Gestor da Internet no Brasil (Bülow, 2023), que discute a remuneração do jornalismo por plataformas digitais, destaca a crescente influência de plataformas como o Google na distribuição de conteúdo jornalístico e na formação da opinião pública, sublinhando a necessidade de políticas que assegurem uma distribuição justa de receitas geradas pela informação on-line. A atenção dada a essas questões reflete uma crescente conscientização sobre como os algoritmos e as políticas corporativas não apenas filtram e direcionam o conteúdo, mas também têm o potencial de afetar a sustentabilidade do jornalismo e a diversidade de vozes na esfera pública. A discussão em torno da remuneração do conteúdo jornalístico por plataformas digitais é emblemática a respeito dos desafios enfrentados na busca por um equilíbrio entre os interesses comerciais das grandes corporações tecnológicas e a necessidade de se preservar a integridade e a pluralidade da informação no ambiente digital.

Se a preocupação com o poder de intermediação das plataformas digitais já é evidente no contexto da informação midiática, ela se torna ainda mais crítica durante os períodos eleitorais. A transparência nos gastos de partidos e candidaturas, especialmente com mídia on-line, é uma

questão fundamental que tem ganhado destaque nas campanhas eleitorais digitais. No entanto, conforme Dommett e Power (2024) argumentam, a transparência ainda é limitada e existem lacunas significativas. Isso é particularmente preocupante no Brasil, onde o financiamento público de campanhas significa que o próprio Estado contribui para os gastos eleitorais. Se já enfrentamos desafios de transparência e prestação de contas nesse cenário, a situação é ainda mais alarmante em relação aos serviços principais e rotineiros das campanhas, nos quais a opacidade e a falta de controle podem ser ainda mais acentuadas e preocupantes.

A escolha do Google como objeto de estudo desta tese se justifica pela sua posição dominante entre os motores de busca e sua capacidade de influenciar o acesso e a distribuição de informações. Diferentemente de plataformas como o Facebook, nas quais o acesso à informação é frequentemente determinado pelas interações sociais e pela publicação de amigos, o Google opera principalmente por meio de algoritmos que classificam e apresentam informações com base em critérios específicos.

Essa distinção é crucial para entendermos as diferentes dinâmicas de acesso à informação na internet. Enquanto as redes sociais podem criar câmaras de eco e bolhas de filtro baseadas nas conexões sociais, os motores de busca como o Google têm o potencial de moldar a visibilidade e a acessibilidade de informações de uma maneira mais abrangente e menos previsível. Portanto, ao focar no Google, este estudo busca compreender como essa poderosa ferramenta de busca influencia a democratização da informação e quais são as implicações dessa influência para a sociedade.

2.4 Os estudos sobre o Google

Nos últimos 20 anos, os estudos sobre como mecanismos de busca impactam o processo democrático começaram com as pesquisas sobre a importância do ranking e o comportamento dos usuários na plataforma. Essas pesquisas pioneiras descobriram a imensa importância que o ranking detinha e, principalmente, a confiança que os usuários davam aos primeiros resultados (Keane e Brien, 2008). Em sequência, pesquisadores se intrigaram com a falta de transparência dos algoritmos utilizados para se criar essas ordenações e começaram então as pesquisas sobre como os resultados eram entregues pelos buscadores e como funcionavam as personalizações para cada usuário (Roesner *et al.*, 2012). Há uma ideia importante de que parte da eficiência do algoritmo de busca é a capacidade de entregar algo personalizado para cada usuário. Essa característica chamou a atenção dos pesquisadores para o possível efeito bolha

como consequência (Pariser, 2011).

Os achados destes estudos, que se baseavam em coletas massivas de resultados do Google, não encontraram de forma conclusiva personalizações de resultados que reforçassem o efeito bolha para perfis ideológicos. Por outro lado, concluíram que as principais diferenças nos resultados são explicadas pela a localização física dos usuários e se os usuários estavam ou não logados nas contas do Google que estavam fazendo a busca (Hannak *et al.*, 2013; Kliman-Silver *et al.*, 2015).

Mais recentemente, ganharam notoriedade internacional as pesquisas lideradas por Epstein (Epstein e Robertson, 2015; Epstein *et al.*, 2017) sobre como rankings enviesados podem mudar a opinião dos usuários, trabalhos com os quais uma parte importante desta pesquisa pretende dialogar. Seguindo os estudos sobre personalização e preocupados com as conclusões de Epstein, algumas publicações focaram na tentativa de encontrar viés partidário nas buscas e as evidências encontradas mostram que pode haver algum viés, ainda que pequeno (Kulshrestha *et al.*, 2017; Robertson *et al.*, 2018). Em 2016, um projeto liderado pela Quello Center e financiado pela Google abordou as questões sobre a importância do buscador em conjunto com outros meios de comunicação. Suas principais conclusões mostraram que o efeito “echo chamber” é super estimado. (Dubois e Blank, 2018; Dutton *et al.*, 2017)

Sendo o Google uma das ferramentas de busca mais utilizadas por usuários do mundo todo (Unkel e Haim, 2021; Trielli e Diakopoulos, 2022). Essas plataformas se tornaram essenciais no atual cenário midiático de alta escolha. “*The rapid growth of the volume of available information dictates the need for large-scale filtering and ranking of information sources. Without automated mechanisms [...], users would be overwhelmed by the abundance of information*” (URMAN *et al.*, 2021, p. 2). No contexto brasileiro, o Google foi selecionado como a marca mais influente em 2021 (Forbes, 2022). Naquele ano, estima-se que cerca de 97% das pesquisas on-line no país tenham sido realizadas com a utilização do Google Chrome. Isso faz com que esses buscadores se tornem centrais na formação da opinião pública. Eles podem impactar as percepções dos usuários por meio da seleção de conteúdos e por meio dos rankings apresentados (Zade *et al.*, 2022).

Thus, search engines may be considered a decisive factor in political opinion formation and, subsequently, voting intentions. Given search engines’ algorithmic content curation, this leads to the question of what kind of information users will encounter when using search engines to find political information - and if this information is biased in certain ways (UNKEL; HAIM, 2021, p. 846).

Em resumo, essas plataformas atuam como grandes *gatekeepers*, exercendo um papel de mediação que se assemelha, mas não é idêntico, ao desempenhado pelo jornalismo profissional

no passado (Vos, 2019). Historicamente, a disseminação de informações sempre envolveu uma variedade de atores, incluindo panfletos apócrifos, comunicação boca a boca e propaganda de campanhas políticas e movimentos sociais. No entanto, o que distingue as *big techs* é a sua capacidade de selecionar e filtrar informações de uma miríade de fontes disponíveis na internet com base em algoritmos sofisticados. Diferentemente do jornalismo, que é guiado por princípios normativos (Steiner *et al.*, 2022), as plataformas digitais operam sob lógicas comerciais e algorítmicas, o que pode ter implicações distintas para a formação da opinião pública.

Steiner *et al.* (2022) enfatizam que tanto o *gatekeeping* humano, quanto o algorítmico podem ser tendenciosos – inclusive, a maneira como a imprensa seleciona suas histórias é um tema amplamente pesquisado (Clayman e Reisner, 1998; Chang e Lee, 1992; McCombs, 2005). Porém, o problema com os mecanismos de pesquisa é a falta de transparência sobre seus critérios de seleção. No caso do Google, estima-se que existam mais de 200 parâmetros internos (a exemplo do idioma e da localização do usuário) e externos (como a popularidade de um determinado site) (Unkel e Haim, 2021). Desse modo, a literatura internacional recente vem se dedicando a examinar empiricamente o viés de seleção das plataformas de buscas.

2.4.1 O que já se estudou sobre viés no Google

Simulando o comportamento humano por meio de agentes automatizados, Unkel e Haim (2021) investigaram os resultados do Google durante a eleição federal da Alemanha em 2017. O objetivo dos autores era verificar como os partidos foram representados e quais fontes foram oferecidas pelo Google. A pesquisa deu origem a mais de 3.400 URLs pertencentes a 538 domínios. Os pesquisadores codificaram os *hosts* em tipos de mídia e, para cada resultado da pesquisa, verificaram a menção a partidos políticos e candidatos. Os autores concluem: “*at least in the case of the German Federal Election 2017, however, our findings indicate that the fear of algorithmic constraints in the context of online search might be overrated*¹” (p. 854). Isto é, naquele contexto de pesquisa, os resultados do Google em relação às menções partidárias retrataram de perto a campanha eleitoral e o seu resultado. Sobre os tipos de mídia, quase metade dos *hosts* eram de agências de notícias.

Steiner *et al.* (2022) também pesquisaram a diversidade dos resultados das plataformas de buscas no contexto alemão, mas chegaram a resultados um tanto diferentes. Os autores

¹ “Pelo menos no caso das eleições federais alemãs de 2017, no entanto, nossas descobertas indicam que o medo de restrições algorítmicas no contexto da pesquisa on-line pode ser superestimado” (Tradução própria).

utilizaram uma metodologia distinta. Em primeiro lugar, selecionaram dez temas relevantes ao país entre novembro de 2015 e junho de 2016: o Brexit; as mudanças climáticas; um poema de Jan Böhmermann, que insultava o presidente turco, Recep Tayyip Erdogan; a intervenção militar na Síria; um julgamento de membros do NSU, grupo de extrema direita; a série jornalística sobre crimes financeiros internacionais, Panamá Papers; refugiados; as eleições regionais no estado de Rhineland-Palatinate; o TTIP, proposta de livre comércio entre União Europeia e Estados Unidos, e o caso de agressões sexuais na virada de 2015 para 2016 na cidade de Colônia. Ademais, cinco mecanismos de buscas foram analisados: Google, Bing, Ask, Ixquick e DuckDuckGo. A coleta de dados foi realizada manualmente por uma codificadora que considerou os primeiros dez artigos com hiperlink de cada *Search Engine Results Page* (SERP). A amostra compreende 500 artigos (5 motores de busca, 10 consultas, 10 artigos). O objetivo do texto é verificar em que medida os resultados compreendem a diversidade de informações (como um fato ou um ponto de vista relacionado a uma determinada questão) e de oradores (opiniões de pessoas, grupos e instituições).

Resultados menos diversos no topo da busca

Um dos resultados da análise é de que tanto a diversidade de informações quanto a diversidade de oradores aumentam quando são incluídos os primeiros cinco resultados orgânicos. Ou seja, existe um risco de os usuários não obterem informações diversas quando consideram somente o primeiro resultado – o que já foi demonstrado pela literatura que acontece com frequência (Pan *et al.*, 2007), embora esse tópico demande pesquisas mais recentes. O grau de diversidade também muda de acordo com os temas pesquisados: o tema do TTIP apresentou maior diversidade de informações e de oradores; já o poema de Böhmermann, Panama Papers, Brexit, Síria, NSU e agressões em Colônia formaram o grupo intermediário com diversidade moderada; a eleição regional de Rhineland-Palatinate e o tema dos refugiados apresentaram baixíssima diversidade e o tópico das mudanças climáticas apresentou diversidade moderada para informações e baixa para oradores. Por fim, os autores não encontraram diferenças estatisticamente significativas entre as cinco plataformas analisadas.

Há diferenças entre os diferentes mecanismos de busca

Assim como Unkel e Haim (2021), Urman *et al.* (2021) também criaram agentes virtuais que simularam o comportamento de navegação dos usuários para investigar os resultados de plataformas de buscas em um contexto eleitoral. Os autores analisaram Google, Baidu, Bing, DuckDuckGo, Yahoo e Yandex, durante as primárias de 2020 nos Estados Unidos, utilizando as palavras-chave “eleições americanas”, “donald trump”, “joe biden” e “bernie sanders”. Os

autores enfatizam que a filtragem e a classificação de informações realizadas pelas *big techs* são de grande relevância – especialmente em períodos de campanha política – porque podem influenciar o comportamento eleitoral dos cidadãos. Sendo assim, torna-se fundamental que essas empresas digitais forneçam aos usuários informações consistentes, confiáveis e livres de vieses. No entanto, Urman *et al.* (2021) encontram algumas discrepâncias nos resultados referentes às primárias de 2020. Além das diferenças entre os resultados produzidos por mecanismos de pesquisa distintos (o que já era esperado pois cada plataforma emprega seus próprios algoritmos de filtragem e classificação), os autores encontraram variações nos resultados de um mesmo mecanismo de pesquisa.

*Visto que tal variação ocorre sob condições não personalizadas, ela aparenta ser aleatória e, muito provavelmente, é atribuída ao fato de que os algoritmos dos motores de busca adaptam constantemente suas respostas para fornecer os resultados vistos como os mais relevantes para os usuários no momento dado. Essa adaptação constante dos resultados significa que usuários do mesmo motor de busca provavelmente receberão resultados diferentes mesmo quando realizam buscas no mesmo momento e nenhuma personalização está envolvida. (URMAN *et al.*, 2021, p. 11).*

Maior volatilidade nos resultados de Donald Trump

Além disso, os resultados referentes aos dois candidatos democratas, Joe Biden e Bernie Sanders, foram mais estáveis quando comparados aos resultados de Donald Trump. Contudo, considera-se importante comentar que essas diferenças entre os candidatos podem não ser um indício de viés, como chegam a argumentar Urman *et al.* (2021). Isto é, os resultados referentes a Donald Trump podem ter se mostrado mais voláteis porque, naquele momento, ele exercia o cargo de Presidente da República e, portanto, estava sujeito ao constante escrutínio do jornalismo profissional. Por fim, um último resultado importante do texto é que o Google priorizou conteúdos de grupos de mídia tradicionais, o que corrobora os achados de Unkel e Haim (2021).

Agregador de resultados de busca

Nemoto e Klyuev (2021) se dedicam a refletir sobre soluções a respeito da formação de bolhas, das câmaras de eco (situação em que ideias e crenças são recorrentemente repetidas e, conseqüentemente, amplificadas e reforçadas) e da própria disseminação de informações falsas (o que pode ocorrer quando as pessoas não têm acesso a informações contraditórias). Esses autores afirmam que as pessoas precisarão receber novos insumos à alfabetização midiática ano após ano. No entanto, o desenvolvimento de ferramentas que funcionem de modo mais transparente também é importante. Sendo assim, eles apresentam um aplicativo chamado MosaicSearch, sistema de buscas de lógica aberta que propõe resultados menos tendenciosos

para usuários. Especificamente, o aplicativo coleta os resultados de pesquisa de diferentes ferramentas (Google, Yahoo ou Yandex) e reconstrói os resultados a partir de parâmetros transparentes e mais compreensíveis aos usuários.

Grupos políticos diferentes usam termos de busca diferentes

Além dos vieses dos resultados (saída), a literatura especializada também tem investigado os vieses de entrada, ou seja, a maneira como os próprios usuários realizam suas pesquisas. Trielli e Diakopoulos (2022), por exemplo, analisaram quais foram os termos de pesquisa utilizados por liberais e conservadores quando queriam encontrar informações sobre candidatos e quais foram os resultados do Google, no contexto da eleição de 2018 nos EUA. O objetivo era responder a duas questões: usuários de diferentes espectros ideológicos buscam informações sobre política de maneiras distintas? Essa diferença é suficiente para revelar resultados de pesquisa substancialmente divergentes? A preocupação dos autores de deus em relação ao possível efeito de filtro de bolha, isto é, a possibilidade de os algoritmos reagirem às preferências e aos perfis dos usuários, oferecendo-lhes somente conteúdos que reafirmam suas posições políticas, limitando a exposição a informações contraditórias. Os resultados mostram que membros de diferentes grupos ideológicos utilizam, de fato, diferentes termos de pesquisa. No entanto, isso não é suficiente para que o Google produza resultados substancialmente divergentes. Isto é, o efeito de filtro de bolha é mínimo. Ao contrário, Trielli e Diakopoulos (2022) enfatizam que ocorre um efeito de *mainstreaming*, “[...] *in which users see a highly similar set of curated media, regardless of how their ideology may have skewed their choice in search queries*”(p. 13)².

Viés de seleção

Um outro exemplo de trabalho que analisa os vieses de entrada é o de Blassnig *et al.* (2023), que utilizaram *surveys* e históricos de pesquisa no Google coletados por meio de doações para compreender quais eram as consultas feitas pelos usuários a respeito de referendos ocorridos na Suíça em 2021. Em primeiro lugar, os autores descobriram que os participantes da pesquisa quase não utilizaram o Google para obter informações sobre os referendos – dentre cerca de 150 mil pesquisas, somente 90 estavam relacionadas ao tema. Como já esperado, os mais jovens e os politicamente mais interessados foram os mais propensos a realizar tais consultas. Sobre os termos utilizados, de modo geral, eles eram bastante neutros, não gerando vieses nos resultados do Google. No entanto, os autores concluem que na hora de escolher quais sites acessar, os defensores dos referendos visitaram sites explicitamente favoráveis, enquanto os opositores

² Disponível em: <https://www.camara.leg.br/noticias/673694-projeto-do-senado-de-combate-a-noticias-falsas-chega-a-camara/>. Acesso em: 14 jan. 2024

visitaram sites explicitamente contrários. Isto é, mesmo que o Google apresente resultados diversos, os próprios usuários tendem a acessar conteúdos que reafirmem seus posicionamentos. Em resumo, mesmo que as plataformas digitais intensifiquem, em alguns contextos, a formação de bolhas ideológicas, esse fenômeno comunicativo não é causado ou determinado somente pelas ferramentas de comunicação. Antes da internet, as pessoas já tendiam a debater assuntos políticos com aqueles que pensavam de maneira semelhante.

A localização importa

Outro trabalho relevante para esse campo de estudos é o de Ulloa *et al.* (2022), que analisa o desempenho das plataformas de buscas em relação à apresentação de notícias recentes. Os autores comparam assuntos que estavam sendo discutidos no momento (eleições de 2020 nos EUA, coronavírus, conflito entre Armênia e Azerbaijão sobre o controle da região de Nagorno-Karabakh e protestos na Polônia relacionados à proibição do aborto) com tópicos estáveis, como Primeira Guerra Mundial e Holocausto. As ferramentas estudadas foram Google, Bing, DuckDuckGo, Yahoo e Baidu e a coleta de dados ocorreu a cada 21 minutos, entre 07h31 de 3 de novembro e 11h10 de 4 de novembro de 2020, a partir de duas regiões (Oregon/EUA e Frankfurt/Alemanha). Todo o processo foi realizado de modo automatizado. Ulloa *et al.* (2022) concluíram que as eleições dos EUA e o coronavírus geraram mais conteúdos novos quando comparados à Primeira Guerra Mundial e ao Holocausto. No entanto, o mesmo não ocorreu com os protestos na Polônia e com o conflito entre Armênia e Azerbaijão. Ou seja, esses temas atuais, mas localizados, foram mais negligenciados. Além disso, mais novidades foram exibidas em Oregon do que em Frankfurt, dado o papel que a localização desempenha nos resultados de pesquisa, e as novidades para “joe biden” foram maiores do que para “donald trump” após a declaração de Biden como vencedor. Por fim, sobre as plataformas, o Google exibiu mais novidades do que o Yahoo em todos os casos e menos novidades que o Bing em 4 de 6 casos.

Até aqui, é possível percebermos que há uma amplitude de temas de pesquisa relevantes a esse campo de estudos – vieses de entrada e de saída, apresentação de novas informações, diferenças entre os assuntos das consultas, entre os termos utilizados, diferenças entre empresas, localização dos usuários, efeitos das pesquisas sobre os usuários etc. E os trabalhos ainda são poucos (devido à contemporaneidade do problema de pesquisa) e voltados a poucos países (especialmente os EUA), além de não apresentarem conclusões definitivas.

2.4.2 O Google e a desinformação

Um outro tema que vem sendo abordado em relação às plataformas de buscas é: em que medida essas ferramentas impulsionam o compartilhamento de informações falsas? Zade *et al.* (2022), por exemplo, investigam se o Google serviu como *gateway* para conteúdos falsos sobre fraudes eleitorais nos EUA. Ao longo de 2020, o país vivenciou uma onda de descrença no sistema eleitoral, motivada por teorias conspiratórias – assim como vem ocorrendo no Brasil nos últimos anos (Marques *et al.*, 2023).

A coleta de dados foi feita por meio da versão paga de uma API do Google (SerpApi), que extrai informações do SERP em tempo real, com a opção de se escolher um local de pesquisa específico (dez regiões dos EUA foram mobilizadas). O recorte temporal foi de 5 de outubro a 3 de dezembro de 2020 e os termos foram divididos em duas categorias: termos neutros, como “eleição presidencial”, e termos que remetem a teorias conspiratórias, como “fraudes eleitorais”. Analisando as manchetes das pesquisas coletadas, os autores descobriram que, quando os usuários utilizam termos neutros de pesquisa, o Google reúne informações confiáveis e que não têm potencial para minar a confiança da população no sistema eleitoral: “[...] *for searches based on general election terms, Google did a relatively good job of surfacing relevant content without leading users to misleading arguments that negatively impacted civic trust in the election processes*”³ (Zade *et al.*, 2022). No entanto, o mesmo não ocorre quando os usuários pesquisam ativamente por teorias conspiratórias, pois, nesse caso, o Google fornece aos usuários histórias falsas sobre as eleições. Os autores também concluíram que os conteúdos falsos tendem a aparecer mais em vídeos, que há algumas empresas de grupos da *legacy media* contribuindo com a desconfiança da população sobre o sistema eleitoral (The Washington Post e Fox News) e que a localização do usuário não moderou a qualidade das informações fornecidas pelo Google.

O trabalho de Harb *et al.* (2022) também trata do tema da desinformação em mecanismos de pesquisa. No entanto, esses autores não analisaram os resultados do Google, mas os termos utilizados com mais frequência pelos usuários da plataforma. Diferentemente de todos os outros textos aqui apresentados, Harb *et al.* (2022) tratam do caso brasileiro. Esses autores analisam como a população fez consultas sobre o coronavírus durante 2020 e 2021 por meio de dados do Google Trends, que mostra as palavras mais utilizadas pelos usuários. Além disso, eles relacionam os resultados com características socioeconômicas e políticas das regiões do país.

³ “[...] o Google fez um trabalho relativamente bom ao trazer à tona conteúdo relevante sem levar os usuários a argumentos enganosos que impactaram negativamente a confiança cívica nos processos eleitorais” (Tradução própria).

Os dados mostram que os brasileiros pesquisaram constantemente sobre a prevenção do coronavírus atrelada ao consumo de determinados alimentos (como o alho, por exemplo). Tratamentos ineficazes, como o uso da cloroquina, também foram amplamente pesquisados ao longo do período. No que se refere às vacinas, as buscas ganharam destaque a partir de abril de 2020. Termos que indicam desinformação aparecem com frequência (por exemplo: vacina mata, vacina causa covid). As variáveis socioeconômicas dos estados (renda mensal, nível de escolaridade) não tiveram impacto estatisticamente significativo sobre as pesquisas infodêmicas. Contudo, estados com maiores índices de votação no ex-presidente Jair Bolsonaro apresentaram mais pesquisas com termos relacionados à desinformação. Como a análise se restringiu ao comportamento dos usuários, não é possível saber se o Google fornecia conteúdos de desinformação sobre o coronavírus quando as pessoas buscavam ativamente por eles. No entanto, é importante destacarmos que, durante a pandemia, o Google implementou medidas para promover informações confiáveis, incluindo a exibição de um painel com links diretos para o site da Organização Mundial da Saúde (OMS) e informações destacadas sobre prevenção e tratamento da Covid-19 em suas pesquisas relacionadas ao coronavírus (TecnoBlog, 2020).

Outra vez, é importante enfatizar que os trabalhos que investigam as plataformas de buscas ainda são poucos e restritos a poucos países. Além da pesquisa de Harb *et al.* (2022), que integra a área da Ciência da Computação, pouco se sabe a respeito dos efeitos do Google sobre as opiniões e as atitudes dos brasileiros – comprovando o caráter precursor e relevante desta tese.

2.4.3 O estudo do Google no Brasil

Como já mencionado na introdução desta tese, em fevereiro de 2024 não encontramos nenhum resultado na plataforma Scielo sobre “Google e eleições”. Os 35 resultados referentes a “Google e política” mencionam o Google Acadêmico (Scholar), como fonte das revisões sistemáticas, e o Google Forms, como ferramenta de pesquisa.

O mais recente trabalho publicado sobre o Google faz parte do grupo de pesquisa no qual esta tese está inserida. O trabalho Vilarins L.;Stabile *et al.* (2021) foi a primeira publicação do grupo de pesquisa com dados referentes a 2018, em um primeiro desenho de pesquisa para se coletar resultados do Google. O estudo referente à coleta de 2018 simulou diferentes perfis em uma abordagem inovadora, à época, de se criar perfis de espectros ideológicos diferentes. Algumas conclusões desse primeiro estudo serão retomadas nesta pesquisa: a) a forte concentração dos resultados em mídias tradicionais e b) o destaque para uma diferença

entre a entrega de veículos reconhecidamente ideológicos (Brasil 247) para o candidato do PT “Fernando Haddad”.

Este capítulo apresentou uma revisão das principais teorias que buscam explicar o comportamento eleitoral, com foco nas discussões sobre acesso à informação e comunicação política. A consolidação das tecnologias digitais nos últimos anos apresenta novos desafios para essas perspectivas teóricas, à medida que diversas expectativas sobre seu papel descentralizador não foram realizadas e aspectos de transparência se perderam com o tempo.

A relação entre essas duas literaturas serve como pano de fundo para a análise dos capítulos empíricos desta tese. Esse enquadramento é essencial ao discutirmos para quem, quando e como o Google é relevante, além de analisar o padrão de entrega do Google para milhões de usuários.

3 COMO O GOOGLE FUNCIONA

Em pouco minutos, uma informação recém-publicada por um portal de notícias aparece como resultado de uma busca do Google. O processo para que o buscador encontre o endereço e decida colocar o novo conteúdo como elegível entre os seus resultados tem ao menos três versões: como o Google diz que funciona, como os especialistas no tema dizem que funciona e como os documentos internos da empresa revelam que funciona.

3.1 Como o Google diz que funciona

Em um documento disponibilizado em janeiro de 2024, o Google declara incluir e ordenar a maior parte das páginas listadas em seu buscador automaticamente por meio de rastreadores que exploram a rede e procuram novas inserções (Google Developers, 2024). Portanto, apesar de existir um guia de boas práticas a ser executado para se tornar relevante dentro das buscas, a empresa declara que as listagens são mais automatizadas do que manuais e que “o Google não garante que vai rastrear, indexar ou exibir a página, mesmo que ele siga os Fundamentos da Pesquisa Google” (Idem).

O mecanismo de busca executa, em resumo, três etapas: primeiro o rastreamento, em que a ferramenta encontra o endereço e faz o download de textos, imagens e vídeos; depois a listagem, em que os arquivos de análise sobre a primeira etapa são armazenados em um banco de dados do Google, e por fim a terceira etapa, que é a exibição dos resultados da pesquisa, em que um usuário faz uma busca e a ferramenta tenta retornar alguma informação relevante para a consulta. No entanto, a empresa alerta de maneira vaga que nem todas as páginas passam por todos os estágios. Abaixo serão detalhadas cada uma das etapas.

Etapa de Rastreamento

Na etapa de rastreamento, o Google procura por novas páginas ou atualizações. Para isso, segundo o *Guia detalhado sobre como a Pesquisa Google funciona*, existem alguns atalhos como percorrer páginas já conhecidas pela empresa para encontrar atualizações e descobrir links para outros conteúdos. Ainda, os donos dos sites podem enviar um arquivo conhecido como sitemap para o Google rastrear. Esses arquivos de mapa do site relacionam as páginas e arquivos mais importantes do domínio e adicionam informações úteis aos buscadores como idiomas disponíveis para as páginas e qual foi a última vez em que foram atualizadas.

O programa que faz essa operação de rastreamento e verificação é conhecido como

Googlebot, definido pela empresa como um “rastreador para computadores que simula um usuário no computador” (Developers, 2024b). Ele acessa a página tentando simular a navegação de um usuário a partir de um navegador Chrome.

O dono de site que deseja não ter o endereço encontrado nos resultados da ferramenta de busca pode bloquear o rastreador e impedir a listagem; o robô também não consegue avançar após páginas que exigem log in.

Indexação

Com o novo endereço rastreado, o Google pode ou não verificar o que está dentro da página. Não está claro quais são as condicionais que disparam a etapa de análise do conteúdo. Nessa segunda etapa, chamada pela empresa de indexação, uma das primeiras ações é determinar se o conteúdo é original ou uma cópia de outro disponível na internet. A intenção do buscador é exibir na ferramenta de busca a concepção canônica.

Além da originalidade do conteúdo, o Google coleta indicadores sobre a página que podem influenciar a exibição dos resultados de pesquisa, mas nem todos são de conhecimento público. “Alguns indicadores incluem o idioma da página, o país em que o conteúdo está localizado, a usabilidade da página, entre outros” (Google Developers, 2024).

Em vez da transparência quanto aos indicadores analisados para listar as páginas, o Google decidiu elencar os três problemas mais comuns para uma página não ser indexada em seu mecanismo de busca. Dois deles são pragmáticos: o endereço está sendo bloqueado porque o robô rastreador não consegue acessá-lo ou o design do site pode dificultar a indexação. No caso do design, podemos citar: páginas sem título ou descrição, códigos sem compatibilidade, páginas com erros de conteúdo não encontrado e lentidão do carregamento, entre outros. Ainda sobre o design, aspectos técnicos da construção de uma página e endereço podem fazer diferença na entrega do conteúdo no resultado de pesquisa, de acordo com o Google. “Quando todos os resultados são muito parecidos, um conteúdo mais acessível pode ter um melhor desempenho” (Google, 2024).

A acessibilidade mencionada pelo Google tem relação com a experiência na página do conteúdo, por exemplo: se o conteúdo pode ser exibido em dispositivos móveis, se é fácil visualizar o conteúdo principal da página, se existem muitos anúncios intrusivos na página e se o carregamento da página é rápido, entre outros (Google Developers, 2024).

O terceiro motivo para um conteúdo não ser indexado é mais subjetivo. O Google afirma que a sua página pode não ser indexada por “baixa qualidade do conteúdo” (Google

Developers, 2024), mas não existe uma explicação satisfatória a respeito do que é um conteúdo de baixa qualidade. A empresa tem as políticas de spam para a Pesquisa Google na Web, onde são detalhados os comportamentos que levam a uma classificação mais baixa ou à não indexação, mas não há garantias, por parte da empresa, de que seu conteúdo será considerado de qualidade caso não se enquadre nessas características negativas explicitadas nas políticas de spam⁴.

A exibição dos resultados

Por fim, a terceira etapa da funcionalidade do mecanismo de busca é a exibição do resultado da pesquisa do usuário. Não são divulgados pela empresa todos os fatores que influenciam o resultado da busca: “A relevância é determinada por centenas de fatores” (Google Developers, 2024). Mesmo que uma página seja rastreada e indexada, isso não significa que ela vai aparecer nos resultados da pesquisa. Ainda segundo a empresa, um dos principais motivos para isso pode ser a baixa qualidade do conteúdo, mas, para além disso, o Google adiciona um novo motivo: o de que o conteúdo da página pode ser irrelevante para as consultas dos usuários.

Para ajudar os criadores de conteúdo a fazerem um material relevante ao seu mecanismo de busca, o Google sugere que esse conteúdo seja útil, confiável e que priorize as pessoas. A empresa sugere uma autoavaliação do trabalho a partir de algumas perguntas, que dizem respeito a: originalidade do conteúdo, se o conteúdo é baseado em fontes, se o conteúdo tem erros de ortografia ou estilo, se a descrição do assunto é completa ou significativa o suficiente e se a informação é relevante o suficiente para ser impressa (Developers, 2024a).

Não está documentado em nenhum material público e de livre acesso do Google como os algoritmos da empresa avaliam a utilidade, confiança ou relevância do conteúdo exibido nos resultados do mecanismo de busca. Ou seja, não é possível, por exemplo, sabermos de que modo o algoritmo avalia a ortografia de um texto ou qual a taxa de erro ortográfico razoável para que o texto não seja penalizado. Ao mesmo tempo, a companhia menciona esse requisito como um item importante para se garantir a qualidade de conteúdo.

O Google, portanto, não explica com detalhes como é o seu algoritmo de exibição dos rankings. O que se sabe é que, ao longo do tempo, ele evoluiu para dar conta da alta competição de diferentes sites e continuar entregando um serviço de qualidade para o usuário. Essa “caixa-preta” fez com que surgisse uma comunidade com inúmeros profissionais de marketing que neste contexto acabam por especular possibilidades do funcionamento do algoritmo e desenvolvem teorias para dar conta desse ambiente competitivo.

⁴ Disponível em: <https://www.camara.leg.br/noticias/673694-projeto-do-senado-de-combate-a-noticias-falsas-chega-a-camara/>. Acesso em: 14 jan. 2024

3.2 Como os profissionais de marketing dizem que o Google funciona

Diante da incerteza a respeito das regras de indexação e resultados de busca do Google, alguns tópicos precisam ser deduzidos. Por exemplo, é sabido que o Googlebot, mencionado anteriormente, dá preferência à procura de endereços em sites construídos para dispositivo móvel (Developers, 2024b). Portanto, é impreterível que, na atualidade, para se ter uma boa colocação nos mecanismos de busca, o site seja amigável para celulares.

No entanto, isso não está descrito explicitamente nos principais guias do mecanismo de busca do Google. É nesse momento que se dá a função do profissional de SEO - sigla em inglês para otimização de mecanismo de busca (*search engine optimizer*) -, profissional cuja função é testar, documentar e observar mudanças nos algoritmos de busca - principalmente do buscador do Google -, com o objetivo de fazer com que sites e conteúdos tenham mais chances de aparecer em melhores posições nas buscas de interesse relevantes.

No guia de iniciante da Moz, umas das empresas que se dedica a estudar o mecanismo de busca, estão os sete principais passos para que seamos bem-sucedidos no ranqueamento dos resultados de busca, em ordem de importância: 1. Seja rastreável para os robôs indexadores; 2. Faça um conteúdo que responda à pergunta de quem busca; 3. Escolha palavras-chave que atraiam robôs e buscadores; 4. Garanta a experiência do usuário, com carregamento rápido e fácil navegação; 5. Compartilhe conteúdo útil com citações e boa abrangência do tema; 6. Faça um bom título, URL e descrição, para gerar uma taxa de clique alta e 7. Faça marcadores de palavras-chave para ganhar destaque na primeira página de resultados (Moz, 2024a).

Conteúdo de qualidade

Enquanto as sugestões do Google para a melhoria do conteúdo estão focadas na utilidade e relevância, os especialistas recomendam estratégias mais pontuais como a adaptação de palavras-chave, para que os robôs indexadores sejam atraídos, e a adequação de títulos, para que sejam mais chamativos, além da construção de conteúdo focada na resposta à pergunta. A qualidade do conteúdo, para o Google, aparece como um dos principais motivadores da melhoria na indexação, ao passo que na lista da Moz o mesmo tópico aparece em quinto lugar. O terceiro, sexto e sétimo passos do guia da Moz sequer são mencionados nas indicações do Google de como a busca funciona.

Palavras-Chave são importantes

Sobre o terceiro passo: ao configurar um novo conteúdo, quem o escreve pode decidir utilizar algumas palavras-chave que apontem do que se trata aquele texto. O guia da Moz indica

aqui que se faça uma estratégia que atraia não apenas usuários, mas também os robôs indexadores. Uma das maneiras de se fazer isso é ao usarmos algumas ferramentas pagas e gratuitas - como o Google Trends - que mostram o volume de busca que algumas palavras-chave estão recebendo. Assim, o usuário pode produzir conteúdos que utilizem tópicos em alta propositalmente.

URLs amigáveis

Entre as estratégias indicadas no sexto passo do guia da Moz está a construção de URLs - sigla para *uniform resource locator* - amigáveis. URL é o termo popular e ao mesmo tempo técnico para nos referirmos ao endereço de rede que localiza um recurso informático. Pode ser, por exemplo, o caminho para um site ou um arquivo no computador.

Os especialistas indicam que o endereço da URL importa e pode ter influência, ainda que secundária, na posição em que a página vai aparecer nos resultados de busca. “Uma URL bem feita leva humanos e ferramentas de busca a um fácil entendimento de qual destino eles vão ser levados” (Moz, 2024b). Esse ponto será revisitado no capítulo 6 para discutirmos os resultados de quais sites aparecem nas buscas políticas no período eleitoral de 2020.

Segundo a Moz, uma boa URL deve ser concisa, simples, relevante, precisa e, para além disso, deve conter a palavra-chave do tema central de que trata o conteúdo principal da página. Para exemplificar, suponha que o mesmo conteúdo foi publicado nos dois endereços abaixo:

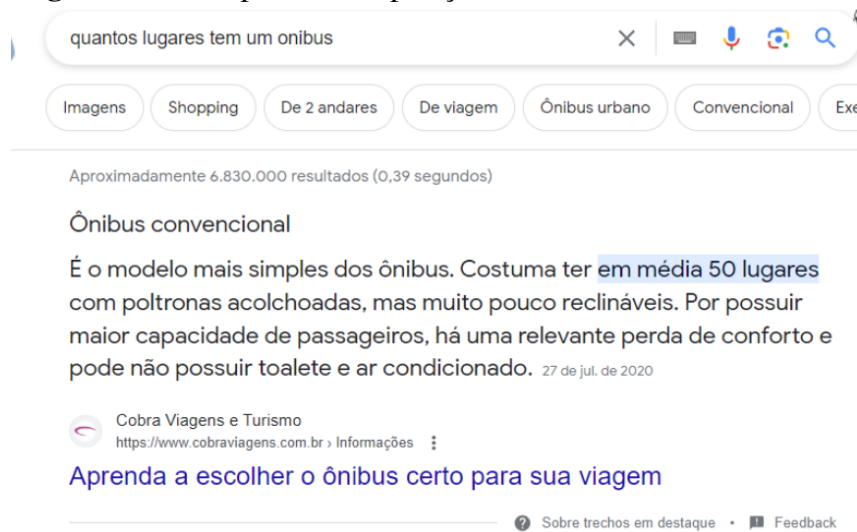
- (A) blog.com.br/debate-filosofia-platao-republica
- (B) blog.com.br/d3b4t3-12

Nesse caso, de acordo com o proposto pelos especialistas, a URL A é melhor do que a B. Mesmo sem o título do conteúdo, no caso do primeiro endereço, é possível termos uma ideia a respeito do que se trata a página que será acessada. Por conta disso, os profissionais da área de SEO acreditam que o Google privilegia esse tipo de construção nos resultados das buscas.

Responda a perguntas de maneira clara

Por fim, o sétimo passo se refere a uma configuração que é feita por quem publica o conteúdo. No momento da publicação, é possível ajustar-se algumas configurações e ter-se a chance de aparecer na posição 0 dos resultados de busca. Para isso é necessária atenção às perguntas usuais que os usuários fazem ao Google e preparar conteúdos que respondam precisa e rapidamente a essas perguntas. A imagem 1 exemplifica:

Figura 1 – Exemplo de uma posição 0 em um resultado de busca



Fonte: elaborado pelo autor, baseado em captura de tela da plataforma Google realizada em 08 de abril de 2024.

3.3 O que o Google disse que funciona para se defender

Uma ação fundamental: clique dos usuários

Para os proprietários de um site entenderem com maior precisão seu desempenho no Google, existe uma ferramenta da própria Alphabet chamada Google Search Console. Dentro dessa ferramenta, é possível acompanharmos cada URL do site e saber quantas vezes elas apareceram nos resultados de busca do google (impressões), com quais palavras-chave, qual a posição média da página e quantos cliques recebeu. A ferramenta também exibe uma métrica conhecida como *click through rate* (CTR), que é a divisão simples dos valores de quantas vezes o endereço foi exibido por quantas vezes isso gerou um clique.

A presença dessa métrica dentro de uma ferramenta oficial do Google sugere que a interação com a URL exibida no resultado da busca era um índice avaliado pelo algoritmo. Mas até novembro de 2023 não existia nenhum documento oficial que validasse essa hipótese ou dimensionasse o peso desse índice dentro do ranqueamento dos resultados na busca.

Julgamento do Google

Em janeiro de 2023, o Departamento de Justiça dos Estados Unidos decidiu processar o Google por monopólio de anúncios e buscas (Department of Justice, 2023). A alegação foi de que, nos últimos quinze anos, o Google agiu de maneira anticompetitiva e excludente. A tática seria neutralizar ou eliminar outros concorrentes de tecnologia por meio de aquisições e utilizar seu domínio de mercado na publicidade para fazer veículos e anunciantes usarem seus produtos,

frustrando, assim, a possibilidade de competição.

O julgamento começou em setembro do mesmo ano e em novembro o Departamento de Justiça dos Estados Unidos divulgou documentos internos da companhia que explicavam com mais detalhes como funciona realmente a busca da ferramenta de pesquisa do Google.

O material divulgado pelo jornalista Danny Goodwin, editor do site Search Engine Land (especializado na cobertura do Google e mais um entre as dezenas de sites especializados em SEO), carece de contexto completo, pois não foi divulgado na íntegra e alguns trechos foram editados. Também não se sabe quem ou qual área foi responsável pela criação da apresentação na empresa. No entanto, o editor do site organizou os slides que mais impactavam o tema do mecanismo de pesquisa e os achados são relevantes para o tema.

A principal descoberta é que o Google afirma, nos documentos internos da companhia, existirem três grandes pilares para a ordem em que as buscas vão aparecer no resultado da pesquisa: corpo do texto, ou seja, aquilo que o conteúdo do documento diz; âncora, que é aquilo que outras páginas dizem sobre aquele documento, e, por fim, interações do usuário (Goodwin, 2023).

Em uma nota de rodapé do documento, as interações dos usuários são explicadas como prioritariamente cliques, mas também podem incluir métricas de atenção como deslizes para o lado, caso o formato seja de carrossel, ou voltar a digitar uma nova busca.

O Google entende cada vez menos dos documentos indexados

Aqui, a novidade não está na confirmação de que o Google utiliza os cliques como critério para ranquear a relevância do conteúdo exibido no resultado da busca, mas sim em como os utiliza. Primeiramente, apesar da sua ampla documentação, a empresa afirma no julgamento que não entende completamente os documentos exibidos nas buscas (conteúdo das páginas, arquivos etc.), portanto observa como as pessoas reagem a eles e registram essas reações. Ou seja, se um documento recebe uma reação positiva, é considerado bom, se negativa, provavelmente é ruim. Isso mantém uma ilusão de que o Google entende os documentos, enquanto, na verdade, atualmente são as interações dos usuários que validam os conteúdos apresentados nos resultados da busca.

Ferramenta perpétua de experimentação

Nessa sequência, cada usuário se beneficia da reação anterior de outro usuário e contribui com sua reação para a experiência do próximo. Nesse sentido, o achado é que o Google montou uma ferramenta perpétua de treinamento, avaliação e experimento de dados, com a qual

a qualidade e a relevância do conteúdo, que o Google diz valorizar como primordial para a indexação e o ranqueamento, são mensuradas em grande parte pela reação do próprio usuário ao conteúdo.

O usuário é parte fundamental do aprendizado

Existe, então, um aprendizado bidirecional com os usuários. Enquanto o Google oferece resultados sobre o mundo, os usuários ensinam o que é o mundo para o Google. Por exemplo, um clique pode indicar que uma imagem é melhor do que um resultado da web (um direcionamento para um site). Esse aprendizado bidirecional é descrito como a fonte da “mágica” do Google.

Em resumo, o Google sempre adotou ao menos dois pilares essenciais para a classificação de um site dentro do seu buscador: critérios técnicos e qualidade do conteúdo. Sobre os critérios técnicos, nada mudou. A ênfase em usabilidade, carregamento da página, design responsivo e experiência do usuário continuam, no entanto, a incógnita sobre como é avaliada a qualidade do conteúdo, que estava apenas parcialmente explicada, agora ganha uma camada a mais.

O Google não sabe o que é um conteúdo de qualidade

Nos guias do Google sempre houve o apontamento de boas práticas para se garantir a qualidade do conteúdo e algumas diretrizes que orientavam o usuário a respeito do que de fato seria considerado SPAM, mas nunca houve nada afirmando veementemente o que seria um conteúdo de qualidade. Com os novos documentos, é possível entender o motivo da falta de precisão em entregar essa informação, pois o que aparenta é que o próprio Google não sabe o que é um conteúdo de qualidade, ainda que saiba se o documento possui uma boa entrega (se é bem-feito para dispositivos móveis, se possui títulos claros, se é rápido etc.). Na realidade, a empresa depende do comportamento dos usuários para entender melhor sobre a qualidade de um conteúdo.

Os usuários sabem o que é um bom conteúdo

A avaliação do usuário emergiu como um aspecto essencial na classificação dos conteúdos. O comportamento dos usuários - incluindo cliques, tempo de permanência em uma página e interações - passou a ser um indicador primordial da relevância e qualidade dos resultados. Essa mudança reflete uma transição do Google para uma abordagem mais orientada ao usuário, na qual suas preferências e comportamento tornam-se fatores decisivos na classificação.

Uma estratégia para não ser responsabilizado no tribunal?

É importante entendermos que os documentos apresentados pelo Google também podem sugerir uma estratégia deliberada para atribuir a responsabilidade do ranking ao comportamento dos usuários. Essa abordagem pode ser uma tentativa de distanciar a empresa da responsabilidade direta sobre os resultados de busca, enfatizando uma lógica de que os resultados são apenas um aglomerado do comportamento em massa da própria audiência do site.

3.3.1 Por que o Google é tão melhor que o Bing?

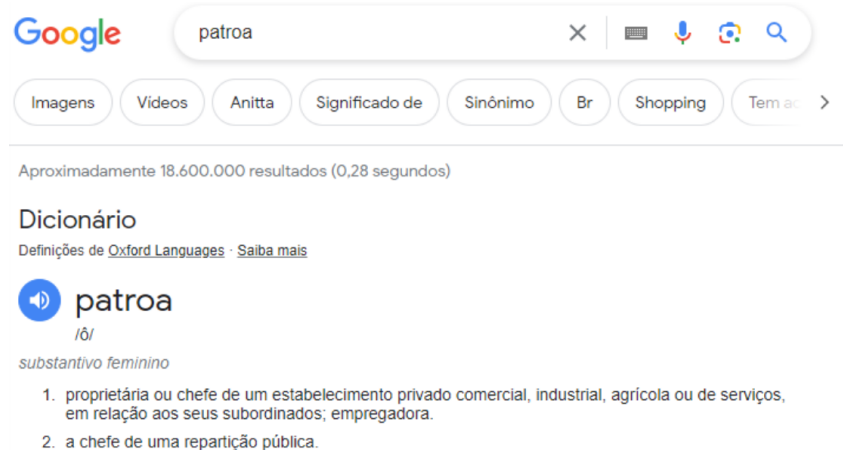
Ao mesmo tempo, essa estratégia de usar os usuários como parte do algoritmo distancia o Google de outros buscadores como o Bing ou o Duck Duck GO, que não têm o mesmo volume de buscas e, portanto, mesmo que tivessem um bom algoritmo de rastreamento e indexação, não teriam a mesma base de dados para refinar e aprimorar os resultados de busca.

O que o histórico da empresa e os documentos disponibilizados apontam é que a evolução dos algoritmos de busca do Google refletem uma jornada de adaptação constante e aprimoramento competitivo. A empresa partiu de uma abordagem com mais critérios técnicos para uma rede de estratégias orientadas aos usuários, na qual a avaliação e o comportamento da audiência desempenham um papel central. Essa transição também sugere uma estratégia sutil de gerenciamento de percepções e responsabilidades, posicionando o Google como um facilitador, mais do que um regulador do ecossistema de informação on-line.

3.4 Quando o Google mudou os resultados da busca

3.4.1 O Google pode interferir nos resultados?

Mas, se tudo é um grande algoritmo que responde ao longo dos anos a diferentes estímulos gerados por bilhões de acessos, o Google pode mudar o resultado de busca? Sim, ele pode. Essa pergunta é importante pois pressupomos essa possibilidade no capítulo 6 desta tese. A seguir, elencamos alguns poucos exemplos de vários já amplamente documentados pela mídia internacional e brasileira.

Figura 2 – Resultado da busca: patroa

Fonte: elaborado pelo autor, baseado em captura de tela da plataforma Google realizada em 21 de janeiro de 2024.

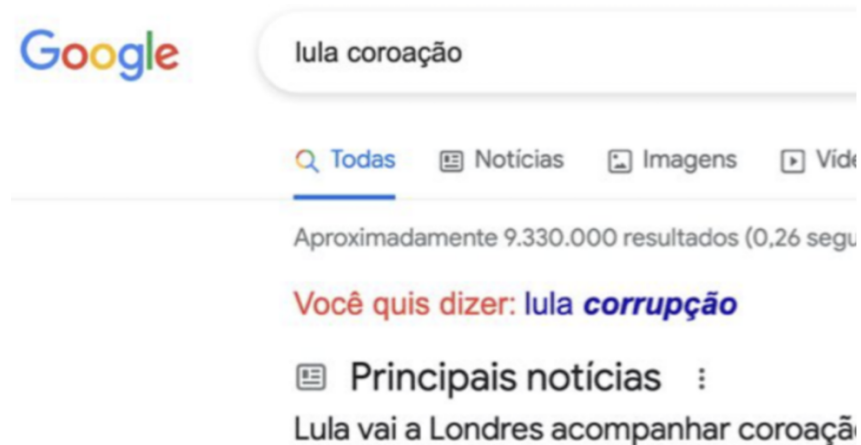
3.4.1.1 *Patroa e mulher solteira: mudanças no dicionário*

Em setembro de 2020, ao procurarmos o termo “mulher solteira” no buscador do Google o retorno à pesquisa apresentaria o significado de “prostituta” e “meretriz”. O caso ganhou repercussão na mídia após artistas como Luísa Sonza publicarem o resultado da busca em suas redes sociais (Terra, 2020).

O recurso de dicionário do Google usa as definições fornecidas pela Oxford Languages, dona do dicionário Oxford, que optou por retirar a definição de mulher solteira dos resultados. Hoje, ao pesquisarmos o termo, os primeiros resultados que aparecem são vídeos e não mais uma definição a partir do dicionário. “A pesquisa mostrou que a definição exibida ainda reflete o uso da expressão em algumas regiões do país, mas percebemos que essa forma de apresentar poderia levar a uma compreensão equivocada e, por isso, foi retirada”, disse o comunicado da Oxford Languages (Terra, 2020).

Um caso parecido ocorreu com o termo “patroa”, antes ao buscarmos o termo a definição que aparecia era “mulher do patrão” e “dona de casa”. Após o caso ganhar popularidade por meio de publicações da cantora Anitta, a Oxford Languages, junto ao Google, mudou os resultados da busca (Terra, 2020). Hoje, ao pesquisarmos o termo patroa, o resultado é algo similar ao significado de patrão (ver figura 2).

O Google afirma trabalhar com conteúdo fornecido por parceiros e, portanto, isenta-se de remover ou editar as definições fornecidas. Nos dois casos citados anteriormente, a companhia afirmou que as decisões de alteração foram tomadas pela Oxford Languages e refletidas nos resultados de dicionário exibidos na busca (Terra, 2020).

Figura 3 – Resultado da busca: lula coroação

Fonte: Poder360 (2023).

3.4.1.2 Google sugere Lula corrupção

Em maio de 2023, às vésperas da coroação do Rei Charles 3º, ao buscarmos por “Lula coroação” o Google sugeria uma correção para a busca “lula corrupção”. Ao mesmo tempo, ao buscarmos “Bolsonaro coroação” o Google sugeria uma correção para “bolsonaro coroação” (Bergamo, 2023).

A questão foi entendida como um problema e corrigida rapidamente após ganhar repercussão nos canais de pessoas como o influenciador Felipe Neto e o deputado federal André Janones (Avante-MG).

Ao contrário do primeiro caso, em que o conteúdo era fornecido por parceiro, o componente de correção de busca do Google não é feito declaradamente por nenhuma empresa terceira e, nesse caso, a mudança foi feita totalmente pela empresa.

A nota de comentário do Google foi sucinta e pouco explicativa: “Nossos sistemas automatizados não levam em conta ideologia política em seu funcionamento. Embora este seja um recurso gerado automaticamente, temos políticas que se aplicam e estamos tomando medidas para resolver o problema” (Poder360, 2023).

3.4.1.3 *Influência política*

O grupo de pesquisa NetLab, da Universidade Federal do Rio de Janeiro (UFRJ), levantou evidências de que o Google apresentou resultados de busca enviesados para usuários que pesquisassem por termos relacionados ao Projeto de Lei 2630/2020, conhecido como PL das Fake News. O texto institui a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet para criar medidas de combate à disseminação de conteúdo falso nas redes sociais e nos serviços de mensagens privadas⁵.

Ao fazer a busca no Google, o usuário era conduzido à pesquisa por PL da Censura, termo usado pela oposição à regulamentação das plataformas, evitando o nome oficial ou o nome usado pela imprensa. Ainda que não se trate dos resultados de uma busca, é importante citarmos que o Google também exibiu em sua página inicial um texto com o título: “O PL das Fake News pode aumentar a confusão sobre o que é verdade ou mentira no Brasil”, uma posição abertamente contra o projeto de lei.

Diante desses fatos, o grupo de pesquisa concluiu que as estratégias usadas pelo Google sugerem que a companhia se aproveita “de sua posição de liderança no mercado de buscas para propagar suas ideias e influenciar negativamente a percepção dos usuários sobre o projeto de lei em prol de seus interesses comerciais” (NetLab, 2023).

3.4.1.4 *Estratégia controversa sobre pornografia infantil*

Em novembro de 2013, em notícia do Mail Online, o Google, enfrentando crescente pressão pública e midiática, anunciou uma mudança significativa em seu motor de busca para combater a pornografia infantil na internet. A decisão, que afetou mais de 100.000 pesquisas, visava a bloquear resultados de busca relacionados ao abuso infantil em todo o mundo. Eric Schmidt, então presidente da empresa, declarou que o Google havia “escutado” as preocupações, promovendo ajustes para prevenir links que levassem a materiais de abuso sexual infantil em seus resultados de busca⁶.

A implementação dessas restrições começou nos países de língua inglesa, com a promessa de expansão para 158 outros idiomas em um prazo de seis meses. Além disso, o Google desenvolveu uma tecnologia para marcar vídeos ilegais, permitindo a remoção de cópias

⁵ Disponível em: <https://www.camara.leg.br/noticias/673694-projeto-do-senado-de-combate-a-noticias-falsas-chega-a-camara/>. Acesso em: 14 jan. 2024

⁶ Disponível em: <https://www.dailymail.co.uk/news/article-2509036/Google-blocks-child-porn-Internet-giant-axes-links-sex-abuse-websites.html>

duplicadas na internet. Simultaneamente, a Microsoft, operadora dos mecanismos de busca Bing e Yahoo, anunciou medidas semelhantes.

Contudo, essa abordagem do Google gerou controvérsias. Embora a iniciativa pareça uma resposta direta a um problema grave, ela levanta questionamentos sobre sua eficácia e as verdadeiras motivações da empresa. A decisão do Google de envolver mais ativamente seus usuários no processo de identificação e sinalização de conteúdo inapropriado, em vez de confiar apenas em métodos internos, sugere uma possível externalização de responsabilidades. Esse movimento pode ser interpretado como uma tentativa da empresa de se distanciar de falhas na moderação de conteúdo, transferindo parte da carga para os usuários.

Além disso, a abordagem do Google no combate à pornografia infantil levanta preocupações sobre a transparência e os critérios utilizados para se definir o que é bloqueado ou marcado. A falta de clareza sobre esses processos gera dúvidas quanto à possibilidade de censura excessiva ou de influência indevida nas informações acessíveis ao público.

3.5 O desafio de se estudar o Google

Neste capítulo exploramos a evolução dos algoritmos de busca do Google, destacando a complexidade crescente e a adaptação constante às demandas dos usuários e às mudanças tecnológicas. Inicialmente focado na análise de links entre sites, o Google evoluiu para incorporar critérios técnicos de usabilidade, qualidade do conteúdo e, mais recentemente, o comportamento do usuário como fatores determinantes na classificação dos resultados de busca. Essa trajetória evidencia uma transição para uma abordagem mais orientada ao usuário, na qual o feedback e as interações dos usuários tornam-se centrais na definição da relevância e da qualidade dos resultados.

A análise das divergências entre as declarações oficiais do Google e as apresentações em julgamentos sugere uma estratégia deliberada de atribuir a responsabilidade do ranqueamento ao comportamento dos usuários. Essa abordagem pode ser interpretada como uma tentativa de se minimizar a responsabilidade direta da empresa sobre os resultados de busca, enfatizando-se a ideia de que o Google é um facilitador, mais do que um regulador, do ecossistema de informação on-line.

Estudar o Google e seus algoritmos de busca é um desafio complexo, dada a natureza de “caixa-preta” desses sistemas. Ao longo dos anos, o funcionamento interno do Google claramente mudou, mesmo que a interface e o modo de uso pelos usuários tenham permanecido

relativamente constantes. Essa complexidade e opacidade tornam o estudo do Google um campo fértil para investigações acadêmicas, mas também exigem abordagens metodológicas rigorosas e adaptativas para desvendarmos os mecanismos subjacentes e entendermos seu impacto no acesso e na disseminação da informação.

À medida que avançamos para o próximo capítulo metodológico, é crucial reconhecermos a complexidade inerente ao estudo dos algoritmos de busca do Google. Dada a natureza opaca e em constante evolução desses sistemas, nossas conclusões serão, em certa medida, suposições baseadas nas informações disponíveis, refletindo a falta de transparência na operação dessa “caixa-preta”. Longe de encontrarmos uma lógica única ou definitiva, será necessário analisarmos os resultados à luz das diferentes explicações a respeito de como o Google funciona. Esse é um exercício científico desafiador, que exige uma abordagem aberta e adaptativa, capaz de lidar com a dinâmica dos algoritmos de busca e as múltiplas interpretações possíveis de seu funcionamento. No capítulo metodológico, enfatizaremos o caráter aberto e exploratório deste trabalho.

4 METODOLOGIA

Esta tese aborda o impacto das plataformas tecnológicas de mediação de informação durante períodos eleitorais, com ênfase no papel do Google nas eleições subnacionais brasileiras de 2020. A escolha das eleições de 2020 no Brasil, marcadas pela pandemia de covid-19, como estudo de caso é relevante para compreendermos a influência do Google em um contexto de intensificação da presença digital das campanhas. A metodologia que será apresentada é pioneira no cenário científico brasileiro e internacional e cobre o processo de mediação informativa desde o período pré-campanha até um mês após a apuração dos resultados.

A presente pesquisa visa a responder a quatro perguntas principais: a) para quem o Google é relevante durante o período eleitoral?, b) como e quando os eleitores utilizam o Google?, c) o que foi exibido pelo Google nas buscas político-eleitorais? e d) houve mudança de padrão na entrega de resultados com o início da campanha eleitoral? Neste capítulo abordaremos todas as técnicas e fontes de dados utilizadas.

4.1 O desenho de pesquisa

Dada a relativa escassez de pesquisas anteriores focadas no Google, a ausência de uma API oficial disponibilizada pela plataforma e a natureza complexa das questões de pesquisa levantadas, que foram abordadas no capítulo anterior, adotou-se uma abordagem metodológica diversificada, incorporando-se uma ampla gama de técnicas e ferramentas de investigação.

A pesquisa envolveu a realização de dois *surveys* aplicados a amostras representativas da população brasileira, nos quais um total de 2771 participantes foram consultados em dois momentos distintos. Desse total de participantes, 850 respondentes concederam ao estudo seus históricos de busca na internet, tanto de computadores quanto de dispositivos móveis. Além disso, foi implementada uma metodologia específica para se registrar os resultados de buscas apresentadas pelo Google, o que culminou na realização de 550 mil buscas e na coleta de mais de 46 gigabytes de dados. Esse vasto volume de informações demandou o emprego de técnicas computacionais avançadas para o processamento e análise dos dados coletados.

Este trabalho não se limita apenas à coleta e análise de dados, mas também enfatiza a integração e apresentação desses dados de modo transparente e acessível. Para isso, esta tese foi escrita em RMarkdown, permitindo uma interação direta entre as bases de dados e os códigos de análise e garantindo a disponibilidade completa deste estudo em seu respectivo repositório

on-line.

Será detalhado passo a passo o procedimento metodológico adotado, destacando-se os links para os códigos específicos de cada etapa do processo de coleta e, também, será indicado onde se pode encontrar, no apêndice metodológico ao final desta tese, as tabelas descritivas e as bases de dados utilizadas. É importante salientarmos que a versão on-line deste trabalho oferece ao leitor a oportunidade de interagir com os dados por meio de tabelas interativas, enriquecendo significativamente a experiência de exploração dos resultados, algo que não seria possível em formatos impressos ou no tradicional PDF.

Essa estrutura metodológica nos permite alcançar alguns objetivos importantes: a) que se possa auxiliar demais pesquisadores a estudarem o Google, b) que todas as análises realizadas neste estudo sejam replicáveis para qualquer pesquisador, c) que o texto possa ser lido de modo mais próximo e interativo e, por fim, d) que, dentre os possíveis efeitos deste trabalho, esteja a indicação de como um monitor do Google pode ser desenvolvido.

4.1.1 Painel longitudinal - pesquisas antes e depois das eleições

O primeiro método de coleta de dados, atrelado à primeira pergunta de pesquisa (para quem o Google é relevante?), foram dois *surveys* on-line, em um desenho de painel, longitudinais em duas aplicações. A vantagem de se fazer duas pesquisas em momentos diferentes está em se medir de maneira mais precisa informações pré-eleitorais, como o conhecimento do eleitor a respeito dos candidatos, e, após as eleições, avaliar esse mesmo respondente e o seu grau de participação e engajamento nas eleições.

O recrutamento e a administração dos convites foram feitos pela empresa Netquest, que possui painel com mais 1,5 milhão de pessoas cadastradas e que está frequentemente envolvida em pesquisas acadêmicas. As duas ondas de aplicação da pesquisa foram em parceria com o Centro de Pesquisa em Comunicação Política e Saúde Pública da Universidade de Brasília (Gramacho *et al.*, 2021; Gramacho e Turgeon, 2021; Mundim *et al.*, 2023).

A coleta dos dados desta pesquisa foi feita com o apoio do Instituto Brasileiro de Pesquisa e Análise de Dados (IBPAD), responsável pela construção dos questionários na plataforma *open source* LimeSurvey, em hospedagem própria e dedicada ao projeto. A empresa contratada responsável pelos convites aos painelistas foi a Netquest. Esses convites consideraram que qualquer indivíduo que estivesse em sua base de dados e com idade eleitoral - a partir de 16 anos - poderia responder à pesquisa. Sendo assim, a fim de se controlar a amostra, para

que fosse o mais próxima possível do universo de eleitores brasileiros com acesso à internet, foram criadas cotas que refletissem tal universo. As cotas eram referentes a idade, gênero e região do Brasil combinados, além de classe social. Para este trabalho, não consideramos a técnica de pesquisa on-line como uma pesquisa probabilística, portanto, não entraremos em cálculos de margem de erro. A diferença da amostra que respondeu à pesquisa para a população brasileira reside basicamente na diferença de escolaridade, com uma menor representação do Ensino Fundamental. O questionário completo da pesquisa pode ser conferido no apêndice 10.5 desta pesquisa, assim como as tabelas de distribuição da amostra no apêndice 10.2.

A primeira onda de pesquisa (w1) teve a participação de 2771 brasileiros entre os dias 21 e 30 de setembro de 2020. A segunda onda (w2) foi aplicada entre os dias 04 e 09 de dezembro de 2020, com o objetivo de obtermos pelo menos 2000 respostas daqueles que haviam participado da primeira onda. Por fim, todos os respondentes da primeira onda foram convidados, e houve 2017 respostas.

A primeira onda da pesquisa teve como objetivo aferir o conjunto de preferências eleitorais desses eleitores na véspera das campanhas eleitorais, o interesse no pleito eleitoral e o conhecimento sobre os possíveis candidatos. A segunda onda teve como objetivo aferir o nível de engajamento e interação na campanha eleitoral, se o usuário teve alguma mudança em suas preferências eleitorais e entender seus hábitos de consumo de mídia on-line e off-line em relação ao pleito eleitoral, apresentados no apêndice 10.2.

Com o questionário pós-eleições (w2), aferimos a variável dependente 1 (vd1) sobre o volume de vezes que o usuário utilizou o Google para buscar informações político-eleitorais. Evidentemente, o *survey* mediu a variável de maneira autodeclarada. A pergunta feita aos painelistas foi: “ao longo do período eleitoral, com qual frequência você fez cada um desses tipos de busca sobre eleições no Google?”.

4.1.2 Histórico de navegação (web tracking) - durante o período eleitoral

Perguntar a um eleitor **quantas vezes ele acessou o Google** e questionar **se ele fez buscas com temas políticos** adentra uma limitação da pesquisa quantitativa. Primeiramente, porque as buscas que os usuários fazem no Google, na maior parte das vezes, têm caráter rotineiro, dificilmente mensuradas de maneira correta se dependermos exclusivamente da memória dos eleitores. Em segundo lugar, porque o conceito de “buscas políticas” pode variar bastante entre os respondentes. Com o intuito de contornarmos essa limitação específica, o esforço de coleta

empírica dessa etapa da pesquisa utilizou os históricos de buscas realizadas no Google por uma amostra dentre os mesmos painelistas que responderam aos *surveys*.

A Netquest, mesma empresa fornecedora do painel de respondentes dos *surveys*, possui uma fonte de dados denominada Behavioural Data. A base da empresa conta com os dados de mais de 10 mil brasileiros, que consentem em fornecer essas informações em troca de incentivos. Os dados são detalhados e permitem, por exemplo, saber quais sites a pessoa acessou, quais buscas foram feitas e, dessas buscas, quais sites foram abertos como resultado. Os dados fornecidos podem advir de desktop, mobile e tablets e informam o tempo de uso e a abertura de sites e aplicativos. Todos os dados utilizados estão disponíveis no repositório desta tese para fins de reprodução do trabalho.

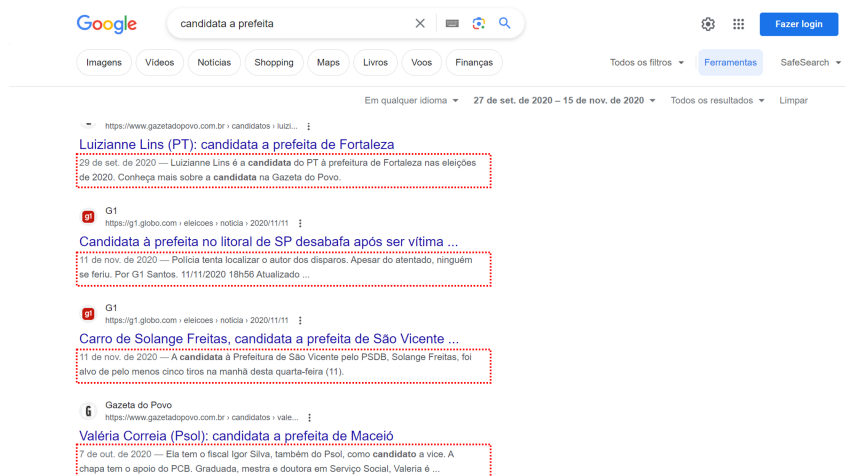
Foram obtidos os históricos de buscas de 826 respondentes ao longo de todo o período eleitoral até o 1º turno: entre os dias 27/09/20 e 15/11/20. Essa base de dados corresponde a 300 mil buscas, com 101 mil termos únicos de busca.

4.1.2.1 Identificação das buscas político-eleitorais

Dentre mais de 100 mil termos, como podemos identificar aqueles que são político-eleitorais? Em um primeiro momento, esse desafio de classificação poderia ter sido tratado de maneira clássica na literatura: amostragem e classificação manual. Amostrarmos, porém, cerca de 3 mil termos e classificá-los de modo manual tomaria quase o mesmo tempo de uma abordagem computacional, e não seríamos capazes de criar um indicador tão preciso para cada respondente a respeito de quantas vezes, de fato, ele utilizou o Google (nossa principal variável dependente, responsável por responder à primeira pergunta desta pesquisa).

Uma segunda abordagem possível seria um conjunto de pequenas regras para que se identificasse termos como “eleições” ou “candidatos” nos termos de busca. Essa segunda abordagem também não se mostraria eficiente, pois, mesmo que as pessoas escrevam de maneira errada, o Google, ainda assim, mostra os resultados esperados. Por exemplo: se uma pessoa buscar por “eleições 2020” ou “eleisoos 2020”, o Google provavelmente entregará retornos de pesquisa ainda associados ao termo correto. Por essa razão, seria muito difícil fazer um conjunto de regras com termos próximos ou acertar todas as variações incorretas de algum termo que ainda seja de uma busca política-eleitoral. Além disso, o eleitor pode buscar o nome de um candidato local ou algum fato local ligado explicitamente a alguma candidatura, tornando impossível prever-se possíveis termos.

Figura 4 – Trecho de resultado da busca retroativa para “candidata a prefeita”



Fonte: elaborado pelo autor, baseado em captura de tela da plataforma Google.

A solução, portanto, para se identificar quais buscas tinham caráter político-eleitoral, foi replicar essas buscas de modo retroativo. Coletou-se, portanto, o que havia sido indexado pelo Google no período da campanha eleitoral para cada termo, utilizando-se uma API de raspagem dos resultados do Google (maiores detalhes sobre a API serão explicados na próxima seção). Após a coleta de todos os resultados, a estratégia foi definir um conjunto de 21 expressões regulares de busca, levando-se em conta as variações de gênero e número. Foi criado, então, um *score* que contaria quantas vezes os 21 termos aparecem (dão *match*) no conteúdo desses resultados. As expressões regulares utilizadas nessa classificação são apresentadas na tabela 1. Esse processo computacional altamente complexo trouxe a capacidade de simular de maneira muito próxima o que o eleitor viu e, ainda que haja limitações e possíveis falhas, garantiu a possibilidade de classificação dos mais de 100 mil termos, evidenciando que, dentre as 352.726 buscas, 0,83% eram buscas com fins político-eleitorais.

Ao anexarmos a quantidade de buscas político-eleitorais realizadas por cada respondente à base de dados dos *surveys*, foi possível utilizá-la como uma variável dependente no modelo, o que será chamado de “uso monitorado” no capítulo correspondente. A análise dessa variável dependente ao longo dos dias também será útil para respondermos à questão: quando o Google é relevante?

IA Generativa para classificar as diferentes buscas político-eleitorais

Após a meticulosa identificação de um conjunto abrangente de 1700 termos relacionados ao contexto político-eleitoral, adotou-se uma estratégia avançada de análise, utilizando-se a mais recente tecnologia de inteligência artificial da OpenAI, especificamente o ChatGPT 4.0.

Tabela 1 – Conjunto de expressões regulares para definição de uma busca política/eleitoral

Regex	Referência
"lei[cç][ãa]o(es)?!",	Eleição, Eleições
"eleitor(alallais)?!",	Eleitor, Eleitoral, Eleitorais
"vota[cç][ãa](oloeslndo)?!",	Votação, Votar, Votando
"candidat(alosluraslarlou)?!",	Candidato, Candidatura, Candidatar
"partid(ariolarioslarismo)?!",	Partidário, Partidários, Partidarismo
"pol[ií]tic(aloloslas)?!",	Política, Político, Políticos
"campanha(s)?(eleitoral(is)?)?!",	Campanha, Campanhas, Campanha Eleitoral
"debate(s)?ldebatendoldebatido!",	Debate, Debates, Debatendo, Debatido
"urna(s)?(eletrônica(s)?)?!",	Urna, Urnas, Urna Eletrônica
"prefeit(olosluraluraslo eleito)?!",	Prefeito, Prefeitos, Prefeitura
"veread(oloeslançalanças)?!",	Vereador, Vereadores, Vereança
"tre(s)?ltribunal(is)? regionais eleitorais!",	TRE, Tribunais Regionais Eleitorais
"coliga[cç][ãa](oloes)?lcoligado(s)?!",	Coligação, Coligações, Coligado
"propaganda(s)? eleitoral(is)?!",	Propaganda Eleitoral, Propagandas Eleitorais
"zona(s)? eleitoral(is)?!",	Zona Eleitoral, Zonas Eleitorais
"se[cç][ãa](oloes)? eleitoral(is)?!",	Seção Eleitoral, Seções Eleitorais
"mesario(s)?lmesariado!",	Mesário, Mesários, Mesariado
"apura[cç][ãa](oloes)?lapurado(s)?!",	Apuração, Apurações, Apurado
"segundo turnolturno(s)? final(is)?!",	Segundo Turno, Turno Final
"eleitor(eslallais)?!",	Eleitor, Eleitoral, Eleitorais
"pesquisa(s)? (de opiniaoeleitoral(is)?)",	Pesquisa Eleitoral, Pesquisas de Opinião

Fonte: elaborado pelo autor.

Essa abordagem permitiu uma exploração profunda dos padrões de busca dos eleitores, revelando não apenas “o quê”, mas também “como” diferentes perfis de usuários recorrem ao Google para buscar informações eleitorais. Inicialmente, a rede neural foi empregada para uma revisão ortográfica criteriosa dos termos coletados. Esse passo inicial assegurou a correção de quaisquer imprecisões ortográficas, incluindo-se erros de digitação, omissões de acentuação, espaços inadequados e outros equívocos comuns, garantindo assim a integridade e a uniformidade da base de dados para análises subsequentes.

Com a base de dados devidamente refinada, procedeu-se à classificação dos termos em categorias pré-definidas por meio de *prompts* de classificação. em três diferentes categorias.

Os *prompts* de comando utilizados para a correção ortográfica e a classificação temática dos termos, assim como uma amostra da base de dados classificada, estarão disponíveis no apêndice 10.1.1.

4.1.3 Captura dos resultados do Google

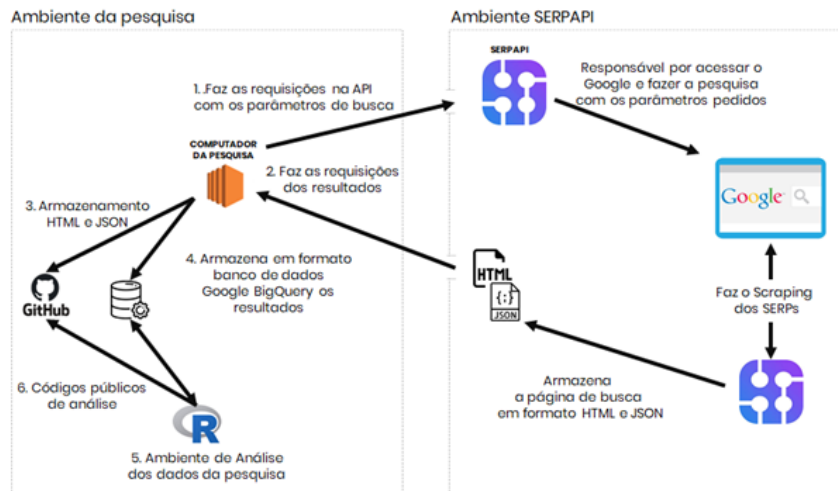
Para responder às perguntas 3 e 4 deste trabalho, foi necessária uma abordagem inovadora de captura dos resultados exibidos pelo Google ou, como a literatura chama, uma search engine results page (SERP).

O desenvolvimento dessa técnica de captura se deu a partir da experiência desenvolvida em 2018 para as eleições presidenciais brasileiras (e replicada para as eleições presidenciais argentinas em 2019) pelo grupo de pesquisa liderado pela Professora Marisa von Bülow em parceria com o grupo de pesquisa da Ciência da Computação da UnB, que desenvolveu um sistema de coleta de SERP para usuários logados. Para mais detalhes, ver (Stabile *et al.*, 2019; Valente e Franco, 2018). Essa primeira experiência demonstrou a imensa dificuldade de se coletar dados do Google. A necessidade de se ter estabilidade e vencer os diferentes protocolos feitos para impedir processos de coleta de dados fez com que, para as eleições de 2020, a técnica de coleta fosse diferente.

Em 2020 optou-se por se utilizar uma plataforma comercial chamada SERPAPI, que fornece API para a requisição desses dados. Ao se utilizar essa ferramenta, garantiu-se a estabilidade no processo de coleta, a facilidade no processo de armazenamento dos resultados e, principalmente, a capacidade de se emular buscas geolocalizadas para cada município brasileiro - variável que outros estudos (Hannak *et al.*, 2013; Kliman-Silver *et al.*, 2015) já mostraram ser a mais impactante na mudança das entregas do Google.

Resumidamente, o sistema funcionou da seguinte maneira:

1. Uma máquina hospedada em um ambiente virtual da amazon AWS executou os códigos responsáveis pela requisição à API do SERPAPI;
2. Os resultados em formato JSON e HTML do SERPAPI foram armazenados nesse servidor;
3. Os arquivos JSON contendo os resultados foram transformados e armazenados em um banco de dados Google BigQuery (plataforma especialmente desenhada para dados massivos);
4. Criou-se nesse momento uma base de dados de resultados cuja unidade de observação era cada link e sua posição nos resultados orgânicos e uma base de dados das pesquisas cuja unidade de observação era a pesquisa.

Figura 5 – Processo de captura massiva dos SERPs

Fonte: elaborado pelo autor.

Após a definição da estratégia de captura, o próximo passo foi a definição do que seria buscado e em qual período.

4.1.3.1 O período pré-eleitoral e eleitoral

Foram definidos três períodos de coleta de dados: o primeiro, pré-eleitoral, ocorreu entre os dias 28 de agosto e 26 de setembro; o segundo ocorreu entre os dias 27 de setembro (data de início da campanha eleitoral) e 15 de novembro (dia de realização do primeiro turno) e o terceiro período se deu entre os dias 16 de novembro e 29 de novembro (data de realização do segundo turno).

Tabela 2 – Período de coleta dos resultados de busca do Google

Período de coleta	Início	Fim	Dias de coleta
Pré-eleitoral	28 de agosto	26 de setembro	30
1º Turno	27 de setembro	15 de novembro	50
2º Turno	16 de novembro	29 de novembro	14

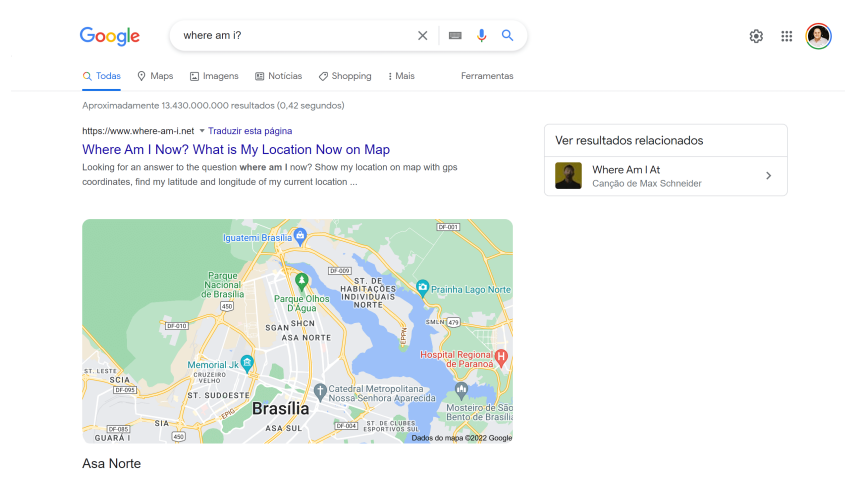
Fonte: elaborado pelo autor.

4.1.3.2 Buscas geolocalizadas

Todas as buscas realizadas foram feitas emulando-se o acesso a partir de algum município específico. Isso só é possível porque o Google utiliza um parâmetro em suas buscas que define cada região com um código diferente, chamado UULE. Por exemplo, o menor município brasileiro, segundo o IBGE, é Serra da Saudade (MG), com 781 habitantes. O seu código UULE é: (

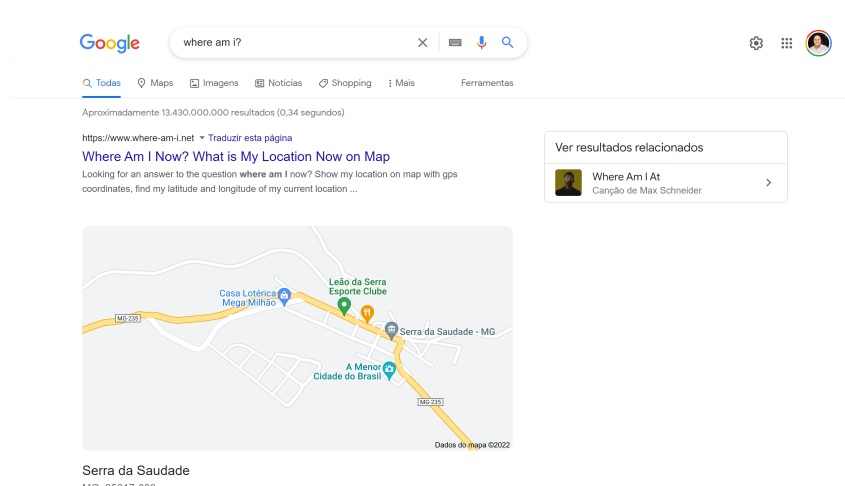
w+CAIQICItU2VycmEgZGEgU2F1ZGFkZSxTdGF0ZSBvZiBNaW5hcyBHZXJhaXMsQnJhemls). Sendo assim, foi possível utilizar o código UULE para se simular acessos advindos dos diferentes municípios brasileiros. As figuras a seguir mostram os resultados: 1. sem o parâmetro e 2. com o parâmetro de busca para o resultado da pergunta “Where am I?” no Google:

Figura 6 – Resultado da busca "where am i" sem nenhum parâmetro de localização



Fonte: elaborado pelo autor, a partir de captura de tela da plataforma Google realizada em 30 jun. 2020.

Figura 7 – Resultado da busca "where am i" utilizando-se o parâmetro UULE para Serra da Saudade/MG



Fonte: elaborado pelo autor, a partir de captura de tela da plataforma Google realizada em 30 jun. 2020.

4.1.3.3 Os termos buscados

De maneira geral, foram realizados quatro diferentes tipos de buscas, sendo dois com termos político-eleitorais, o que chamamos de “Buscas Eleitorais”, e dois de diversos temas,

que chamaremos de “Buscas não eleitorais”.

Tabela 3 – Quantidade de buscas realizadas por tipo de abordagem

Abordagem	Freq.	Prop.
1. Eleitorais: Sobre candidatos	26.874	5,2%
2. Eleitorais: Termos sobre eleições	348.884	67,1%
3. Não eleitorais: Termos genéricos	27.586	5,3%
4. Não eleitorais: Termos oriundos da população survey	116.771	22,5%

Fonte: elaborado pelo autor.

Buscas sobre candidatos: foram realizadas buscas específicas para os candidatos de todas as capitais brasileiras. Durante o período pré-eleitoral, ainda sem a definição formal de candidatos, os termos de busca foram construídos com base nos questionários de pesquisa de intenção de voto registrados no Tribunal Superior Eleitoral (TSE) brasileiro. Durante o período do 1º turno, os termos foram ajustados para que refletissem apenas os candidatos de cada capital. Nas capitais em que houve segundo turno, durante esse período, as buscas foram atualizadas para que fossem considerados os dois candidatos concorrentes. Para cada capital foram coletados, em média, 14 nomes de candidatos, e essas buscas foram realizadas diariamente, com os municípios em ordem aleatória, uma vez ao dia, às 20h. Ao todo foram 26.874 buscas realizadas nas 26 capitais. A tabela com os termos dessas buscas encontra-se no apêndice 10.3.1.

Buscas gerais sobre eleições: foram realizadas buscas gerais sobre as eleições municipais para todas as capitais, para todos os 70 municípios que poderiam ter segundo turno e para todos os 190 municípios acima de 110 mil habitantes. Essas três primeiras faixas de municípios representam onde moram 55% dos habitantes brasileiros. Dentre os 5283 municípios restantes, foram sorteados 286, sendo esses proporcionais a outras três faixas populacionais.

Tabela 4 – Distribuição dos municípios com buscas gerais sobre eleições

Tipos de municípios	Quantidade	População(em milhões)	População(em % - acum)	Amostra
Capitais	26	47,1	22,8%	26
Municípios com 2º turno	70	34,3	39,3%	70
Mais de 110 mil habitantes	190	32,4	55,0%	190
Mais de 45 mil habitantes	455	31,1	70,0%	95
Mais de 20 mil habitantes	1.032	30,3	84,6%	93
Até 20 mil habitantes	3.796	32,0	100,0%	98
Total	5.569	207,1		572

Fonte: elaborado pelo autor, com base em dados do IBGE (2020).

Para cada município, foram feitas ao menos três buscas com termos relacionados

a eleições. A definição dos termos de busca foi construída com a aplicação de um survey on-line realizado pelo Instituto Brasileiro de Pesquisa e Análise de Dados (IBPAD) por meio da empresa de recrutamento online NetQuest. A pesquisa contou com 2 mil pessoas em uma amostra representativa da população on-line brasileira entre os dias 20 e 27 de agosto de 2020 com a seguinte pergunta: “Se você tivesse que fazer agora uma busca no Google sobre os candidatos em sua cidade, como faria? (Por favor digite nos 2 campos abaixo como se você estivesse no Google. Em cada campo dê um exemplo diferente do tipo de busca que faria)”. Os resultados preliminares, após pequenos ajustes de português e devidas tabulações, foram:

Tabela 5 – Termos de busca que os usuários fariam para buscar mais informações sobre os candidatos

Termos de busca	Frequência	Prop.
Candidatos a prefeito em (CIDADE)	263	26%
Candidatos em (CIDADE)	145	15%
Eleições 2020	84	8%
Nome do Candidato	68	7%
Histórico/ Biografia/ Currículo (CANDIDATO)	62	6%
Propostas/ Plano de governo (CANDIDATOS)	45	5%
Ficha limpa (CANDIDATOS)	43	4%
Candidato a Prefeitura de (CIDADE)	34	3%
Candidato a Prefeito	32	3%
Nenhum/ Não sei/ Não tem interesse	38	4%
Outros	467	47%

Fonte: elaborado pelo autor.

Os três primeiros termos foram escolhidos, ainda que refletissem uma clara distinção de gênero no modo como as buscas são feitas. Além desses termos, outras buscas foram realizadas para diferentes faixas de tamanho de município. Todos os termos da tabela a seguir foram buscados diariamente após as 21h, primeiramente de modo aleatório para cada município.

Tabela 6 – Tabela dos termos de busca políticos para cada tipo de município

Termos de busca	Municípios
candidatos a prefeito em <city_name>	Todos os 572
candidatos em <city_name>	Todos os 572
Eleições 2020	Todos os 572
candidatas a prefeita em <city_name>	Capitais, 2º Turno e >110 mil hab.
candidatas em <city_name>	Capitais, 2º Turno e >110 mil hab.
Qual é o melhor candidato?	Apenas capitais
Qual é o melhor candidato para a saúde?	Apenas capitais
Qual é o melhor candidato para fazer ter mais segurança?	Apenas capitais
Qual é o melhor candidato para fazer ter mais empregos?	Apenas capitais
Qual é o candidato com melhores propostas para a educação?	Apenas capitais
candidato ficha-limpa OR candidato fichalimpa	Apenas capitais
candidato do bolsonaro	Apenas capitais
candidato do lula	Apenas capitais
candidato antipt	Apenas capitais
candidato do pt	Apenas capitais
candidato denunciado	Apenas capitais
fraude nas urnas	Apenas capitais

Fonte: elaborado pelo autor.

Faltando 15 dias para o 1º turno, esses termos foram buscados para todos os municípios brasileiros. A tabela com todos os termos utilizados nas buscas gerais e as quantidades de buscas eleitorais realizadas estão no apêndice 10.3.2.

Buscas não eleitorais

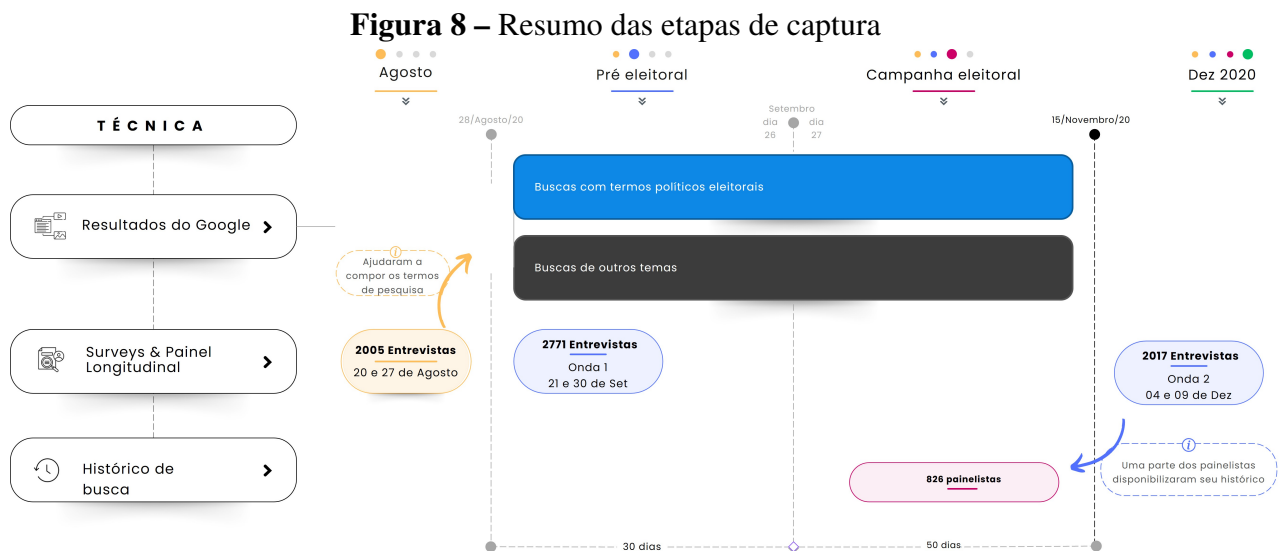
Também foi realizado um conjunto de buscas não eleitorais, com a finalidade de serem um grupo controle dos SERPs de natureza eleitoral. Diversos outros termos não políticos foram buscados. Esses termos genéricos foram definidos com base em relatórios do Google sobre os principais termos buscados nos últimos anos. As buscas também foram feitas de modo geolocalizado nas cidades amostradas neste projeto. O procedimento de escolha para cada cidade foi aleatório: para cada busca, uma cidade diferente. Os tipos de busca e suas quantidades estão no apêndice 10.3.1.

O *survey* que embasou a escolha dos termos eleitorais também incluiu uma pergunta com o intuito de se criar um conjunto de termos que pudessem ser utilizados como controle. A pergunta feita foi: “Pensando agora sobre qualquer assunto, quais foram as duas últimas buscas mais frequentes que você fez no Google? (Por favor digite nos 2 campos abaixo como se você estivesse no Google. Em cada campo dê um exemplo diferente do tipo de busca que você fez, independentemente do tema)”.

No total, foram 1509 termos utilizados, do mesmo modo como os usuários os digitaram no *survey* on-line, e que foram escolhidos aleatoriamente para os 572 municípios, diariamente, em três rodadas de captura. Assim, cada município, além de três buscas políticas, teve, pelo menos, três buscas não políticas. Os termos estão elencados no apêndice 10.3.3 deste trabalho.

4.2 Resumo

O quadro a seguir mostra um resumo didático de todas as etapas de captura em uma janela temporal no ano 2020.



Fonte: elaborado pelo autor.

5 PARA QUEM O GOOGLE É RELEVANTE

O termo Search Engine Manipulation Effect (SEME) se tornou conhecido com a série de publicações de Epstein que têm como tema “Google e política”, as mais relevantes nos últimos anos, e que concluem, por meio de experimentos, que um resultado enviesado de busca pode ter um efeito de pelo menos 20% na intenção de voto de eleitores indecisos. Esse valor pode ser ainda maior quando usuários não percebem o viés e desconhecem em maior medida o candidato (Epstein e Robertson, 2015; Epstein *et al.*, 2017). Ainda que Epstein tenha se tornado um forte crítico do Google e utilize suas pesquisas nesse debate, elas não abordam quem de fato utiliza o Google para se informar politicamente e, mais ainda, quem o toma como fonte para decidir em quem votar. O primeiro objetivo específico desta tese, portanto, é entender **para quem** o uso do Google durante o período eleitoral é relevante.

Como já abordado, os estudos prévios raramente fazem uma conexão com a área de comportamento político e é com esse debate que este trabalho procura contribuir. Ainda que as pesquisas anteriores tenham utilizado as eleições como estudo de caso, seu objetivo era entender como aquela aquisição informacional mediada pelo buscador impactaria no comportamento político, ou seja, não houve a discussão sobre que tipo de comportamento político preestabelecido pelo eleitor faria com que ele buscasse em maior ou menor medida aquela ferramenta para a sua aquisição informacional. O segundo objetivo específico deste trabalho é analisar **como** o Google é utilizado por diversos perfis.

Lodge e Taber 2006 definem como *motivated reasoning* a teoria segundo a qual o perfil das pessoas que mais buscam informação política são aquelas já engajadas e com a sua decisão política já formada. Essa definição, já preestabelecida, seria determinante no “filtro” para a melhor seleção da informação política. O ato de interagir e buscar por mais informações políticas, portanto, serviria para reafirmar suas convicções ou buscar fontes com o intuito de ajudar no trabalho de convencimento de outras pessoas.

Há, ainda, a corrente que aborda o fato de que, ainda que a informação seja relevante, são os atalhos cognitivos, principalmente vindos de pessoas próximas (Baker *et al.*, 2006; Huckfeldt *et al.*, 1995; Turgeon e Rennó, 2010), que de fato são essenciais para a formação política, diminuindo assim a importância das informações e enfraquecendo a máxima de que, nesse caso, “quanto mais informação, melhor”.

Desse modo, um terceiro e quarto objetivos específicos se delimitam e dizem respeito a tentar responder **como** o Google é utilizado e **quando**. A partir dessas duas premissas a respeito

dos aspectos da informação sobre o comportamento político, a principal hipótese deste capítulo é:

H1: Eleitores com preferência eleitoral consolidada tendem a utilizar **mais** frequentemente o Google para buscar informações sobre eleições.

Caso essa hipótese não seja refutada e se constate que apenas usuários que já possuem uma decisão de voto definida ou que estão ativamente envolvidos no processo de convencimento eleitoral são os principais usuários dos buscadores do Google, tal achado traz importantes implicações para o debate atual.

Em primeiro lugar, será necessário reavaliarmos a relevância das informações obtidas por meios digitais, especialmente em plataformas que exigem uma atuação ativa do usuário na busca de informações, como é o caso dos buscadores e sites. Esse ponto poderia desafiar diretamente as conclusões de Epstein, sugerindo que, apesar da possível influência de resultados de busca tendenciosos, não são todos os eleitores que buscam ativamente por informações eleitorais na internet. Aqueles que o fazem, possivelmente, já possuem uma opinião pré-formada e conhecimento sobre os candidatos, o que potencialmente reduziria o impacto de qualquer efeito que resultados de busca enviesados teriam.

Por outro lado, se a hipótese for refutada, tal resultado contribuirá significativamente para a compreensão do papel do Google e do consumo de informações na política. A descoberta de que o Google serve como uma fonte de informação substancial para eleitores indecisos e com preferências políticas menos definidas abriria novas perspectivas para pesquisas futuras. Essas pesquisas poderiam utilizar os resultados de busca do Google como um *proxy* para analisar preferências eleitorais, indo além do conteúdo das mídias tradicionais e da propaganda eleitoral. Tal achado enriqueceria o debate sobre a existência de câmaras de eco e o papel dos mediadores digitais na política (Dubois e Blank, 2018), além de iluminar o impacto real que os serviços de busca têm na formação das preferências políticas dos eleitores.

5.1 O que pode explicar o maior ou menor uso do Google

A relevância do Google, como já destacado na introdução deste trabalho, permanece significativa para milhões de brasileiros, sendo considerado uma fonte confiável tanto para informações sobre candidatos quanto para se verificar a veracidade de informações. Contudo, até que ponto o Google detém essa relevância na formação da opinião política ainda é uma questão em aberto. Embora muitos apontem o Google como uma referência essencial na busca

e qualificação de informações, é possível que, na hora da decisão do voto, influências mais próximas, como colegas de trabalho, familiares ou amigos, tenham um papel mais decisivo. Este estudo, com seu desenho de pesquisa específico, não visa a determinar qual é a fonte de informação mais relevante, mas sim identificar para quem o Google é especialmente significativo.

Para responder a essa pergunta, o estudo adota um desenho de pesquisa descrito detalhadamente na seção metodológica, combinando duas fontes de dados com abordagens analíticas distintas: 1) um *survey* em painel longitudinal e 2) o histórico de navegação dos usuários.

Essa combinação é interessante devido à sua complementaridade. Os dados oriundos do *survey*, embora úteis, encontram limites epistemológicos ao quantificar ações ou hábitos cotidianos (Alvarez e Li, 2022; Ansolabehere e Hersh, 2012; Georgiev *et al.*, 2022). No contexto desta pesquisa, por exemplo, perguntar a alguém quantas vezes utilizou o Google para pesquisar temas políticos/eleitorais pode esbarrar no viés da resposta socialmente desejável, além de depender da memória do respondente, que pode não refletir a real frequência de uso. Outras perguntas, como “Por quantos minutos você dirigiu hoje ouvindo rádio?” ou “Quantas vezes você se exercitou nos últimos seis meses?”, normalmente são abordadas em termos de preferência ou por questões generalistas sobre períodos ou escalas de intensidade.

Buscando uma abordagem mais precisa, esta pesquisa utilizou os dados de histórico de buscas de uma parte dos respondentes do *survey* em painel longitudinal. O dado oferece um nível de confiabilidade sobre o real uso e os hábitos de informação on-line dos indivíduos. Além disso, a análise do histórico de navegação permite um olhar mais concreto e imparcial sobre o comportamento informacional, superando as limitações das autoavaliações, que podem ser influenciadas por memórias seletivas ou pela tendência de responder de maneira socialmente aceitável. Essa abordagem, ao capturar dados do comportamento on-line, proporciona compreender como e quando o Google é utilizado no contexto político-eleitoral.

Adicionalmente, ao cruzarmos dados do histórico de navegação com as respostas do *survey* em painel, é possível identificarmos padrões e discrepâncias entre o que as pessoas relatam e o que efetivamente fazem em termos de consumo de informação política. Esse tipo de análise, baseada em métodos múltiplos, oferece uma rica oportunidade para se explorar não apenas as preferências declaradas dos usuários, mas também suas ações reais, fornecendo um panorama mais completo e revelador do papel do Google na formação da opinião política dos eleitores brasileiros.

A fim de analisarmos para quem o Google é importante, criamos um modelo multivariado para testar a hipótese deste capítulo de que o Google é importante para quem já possui um maior conhecimento político. Nas próximas seções, explicaremos as variáveis desse modelo e, na sequência, faremos a análise desse modelo.

5.2 O que pode explicar - Variáveis Independentes

As variáveis independentes são aquelas que pretendem explicar o modelo de análise. Se a hipótese é que “eleitores com preferência eleitoral consolidada tendem a utilizar **mais** frequentemente o Google para buscar informações sobre eleições”, o desafio é instrumentalizar o conceito “preferência eleitoral” e o que seria a sua consolidação em termos objetivos para se medir e averiguar sua relação com a variável dependente. Apresentaremos as variáveis que pretendem instrumentalizar “preferência eleitoral”.

5.2.1 *Conhecimento sobre os candidatos e participação*

Neste bloco, será operacionalizada uma parte essencial do nosso modelo ao medirmos o nível de conhecimento prévio dos eleitores e sua participação efetiva nas eleições. Essa análise nos ajudará a entender como a familiaridade com os candidatos pode influenciar o uso de plataformas de busca para a aquisição de informações durante o período eleitoral e, conseqüentemente, para validar empiricamente a relação proposta em nossa hipótese central.

5.2.1.1 *Nível de conhecimento das candidaturas*

A compreensão do nível de conhecimento dos eleitores sobre os candidatos é uma variável independente em nosso modelo. A tabela 7 revelou que, uma semana antes do início oficial da campanha, 35,7% dos entrevistados já estavam bem informados, reconhecendo três ou mais candidatos à prefeitura de sua cidade. Contrastando com esse dado, 26,6% dos participantes indicaram não ter conhecimento sobre nenhum dos candidatos.

Esse espectro de familiaridade com os candidatos é essencial para a análise, uma vez que reflete não apenas o interesse prévio e a exposição aos discursos políticos, mas também pode antecipar o comportamento eleitoral subsequente, incluindo a busca por informações adicionais e a decisão final nas urnas. O nível de conhecimento prévio pode influenciar a maneira como os eleitores interagem com as plataformas de busca como o Google: enquanto alguns podem usá-las

para confirmar suas opções já consideradas, outros podem buscar informações que ajudem a formar ou modificar suas preferências eleitorais. Portanto, essa variável desempenha um papel significativo em nosso estudo, pois ajuda a traçar um perfil mais detalhado do eleitorado e sua interação com a informação política no ambiente digital.

Tabela 7 – Resultados para a pergunta: Você já sabe ou já ouviu falar sobre quem são os(as) candidatos(as) ao cargo de prefeito(a) em sua cidade?

	Freq.	Prop.
Não sei de nenhum(a) candidato(a) a prefeito(a)	738	26,6%
Sei de um(a) candidato(a) a prefeito(a)	419	15,1%
Sei de dois(duas) candidatos(as) a prefeito(as)	557	20,1%
Sei de três ou mais candidatos(as) a prefeito(a)	989	35,7%
Não respondeu	68	2,5%
Total	2771	100%

Fonte: elaborado pelo autor.

5.2.1.2 Se sabia em quem votar

A pergunta sobre a intenção de voto (tabela 8) foi direcionada especificamente aos eleitores que indicaram conhecer ao menos uma candidatura e constitui uma segunda variável em nosso modelo. A inclusão dessa questão é fundamental para entendermos a dinâmica entre o conhecimento dos candidatos e a formação da intenção de voto, permitindo-nos investigar se o conhecimento prévio se traduz em decisão eleitoral concreta ou se ainda existe um espaço significativo para a influência de novas informações no período que antecede a votação. A pergunta foi realizada na 1ª onda do painel. Ela captura o grau de certeza do eleitorado quanto

Tabela 8 – Resultados para a pergunta: E você já sabe em quem você vai votar? (Apenas para quem conhecia algum candidato)

	Freq.	Prop.
Não pretendo votar	281	14,3%
Ainda não sei em quem vou votar	712	36,2%
Tenho alguma ideia, mas ainda estou em dúvida	416	21,2%
Sim, tenho certeza em quem vou votar	549	27,9%
Não respondeu	7	0,4%
Total	1965	100%

Fonte: elaborado pelo autor.

à sua escolha eleitoral, refletindo diretamente o nível de decisão política já estabelecido. Os resultados mostraram que 36,2% dos entrevistados ainda estavam indecisos sobre seu voto, enquanto 27,9% afirmaram ter certeza em quem votariam. Adicionalmente, 21,2% indicaram ter alguma ideia, mas permaneciam em dúvida.

5.2.1.3 *Se votou no 1º turno*

A última variável independente deste bloco é referente à participação no 1º turno, onde 70,4% da amostra afirmou ter votado em um(a) candidato(a) no primeiro turno das eleições municipais. Essa pergunta foi feita na segunda onda do painel, realizada, portanto, depois do 2º turno, entre os dias 04 e 09 de dezembro de 2020, e o resultado descritivo pode ser conferido na tabela 9.

Tabela 9 – Resultados para a pergunta: Você votou no primeiro turno em algum(a) candidato(a) ao cargo de prefeito(a) municipal?

	Freq.	Prop.
Sim, votei em um(a) candidato(a)	1420	70,4%
Votei em branco	63	3,1%
Anulei	148	7,3%
Não fui votar	333	16,5%
Não respondeu	53	2,6%
Total	2017	100%

Fonte: elaborado pelo autor.

5.2.2 *Preferência partidária*

Para além do engajamento e da participação direta nas eleições, foram desenvolvidas duas variáveis adicionais relacionadas às preferências políticas e partidárias dos respondentes. A primeira é baseada em uma questão clássica da ciência política que indaga sobre o partido político de preferência do eleitor. A segunda questiona qual partido o eleitor jamais escolheria para votar. Dada a peculiaridade do contexto político em 2020, em que o presidente Jair Bolsonaro, apesar de ter sido eleito pelo PSL em 2018, encontrava-se sem filiação partidária, decidiu-se incluir a categoria “Partido do presidente Bolsonaro” como uma opção de resposta distinta.

5.2.2.1 Partido de preferência

A tabela 10 indica a distribuição das preferências partidárias entre os entrevistados antes do início do período eleitoral em 2020. A maioria, com 75,8%, declara não ter um partido político de preferência, evidenciando um distanciamento ou uma falta de identificação dos eleitores com as estruturas partidárias existentes. O partido do presidente Jair Bolsonaro, no momento do levantamento, figura em segundo lugar, com 7,5% da preferência dos participantes, seguido pelo Partido dos Trabalhadores (PT), que conta com 7,1%. Observa-se uma queda significativa na proporção de preferência para os demais partidos, com o NA (Não Aplicável ou resposta não fornecida) recebendo 3,2% das indicações.

Partidos tradicionais como o PSOL e o PSDB aparecem subsequentemente, com 1,8% e 1,5%, respectivamente, sugerindo uma preferência menor em comparação aos líderes do levantamento. O PSL, partido pelo qual o presidente Jair Bolsonaro foi eleito em 2018, e o PDT apresentam 1,1% e 0,9% das preferências, enquanto o DEM e o MDB, ambos partidos com longa história no cenário político brasileiro, são citados por apenas 0,5% dos entrevistados cada um.

Tabela 10 – Resultados para a pergunta: Qual é o seu partido político de preferência?

	Freq.	Prop.
PT	196	7,1%
PSL	30	1,1%
MDB	14	0,5%
PSDB	41	1,5%
DEM	15	0,5%
PSOL	49	1,8%
PDT	26	0,9%
Partido do presidente Jair Bolsonaro	209	7,5%
Não tenho partido de preferência	2101	75,8%
Não respondeu	90	3,2%
Total	2771	100%

Fonte: elaborado pelo autor.

5.2.2.2 Partido no qual nunca votaria

Além da pergunta de partido de preferência, perguntamos também, na 1ª onda da pesquisa, em qual partido o eleitor **nunca** votaria. Os resultados podem ser conferidos

na tabela 11. A maior proporção, 38,2%, indica não ter um partido específico que rejeitaria categoricamente, o que pode sugerir um distanciamento em relação à política. O Partido dos Trabalhadores (PT) aparece como o partido com a maior rejeição específica, com 32,6% dos entrevistados afirmando que nunca votariam nele, uma indicação de que o partido possui uma imagem negativa considerável para uma parcela significativa do eleitorado.

O partido do presidente Jair Bolsonaro é mencionado por 14,8% dos participantes como uma opção na qual nunca votariam, refletindo uma rejeição específica que não é tão expressiva quanto a do PT, mas ainda assim notável. Seguindo a lista, o PSOL é citado por 3,6% dos entrevistados e o PSL por 3,4%, mostrando uma rejeição mais moderada a esses partidos. Os partidos PSDB, MDB e DEM apresentam níveis de rejeição ainda menores, com 2,1%, 0,9% e 0,9% respectivamente, enquanto o PDT é o partido com a menor rejeição específica, mencionado por apenas 0,4% dos respondentes. Esses dados indicam que a rejeição partidária varia consideravelmente, com alguns partidos enfrentando níveis de aversão mais acentuados por parte dos eleitores. Ao compararmos as tabelas referentes às preferências partidárias e à rejeição

Tabela 11 – Resultados para a pergunta: Em qual partido político você nunca votaria?

	Freq.	Prop.
PT	904	32,6%
PSL	93	3,4%
MDB	24	0,9%
PSDB	59	2,1%
DEM	24	0,9%
PSOL	101	3,6%
PDT	10	0,4%
Partido do presidente Jair Bolsonaro	409	14,8%
Não tenho partido que nunca votaria	1059	38,2%
Não respondeu	88	3,2%
Total	2771	100%

Fonte: elaborado pelo autor.

a partidos políticos dos entrevistados, observamos um contraste notável. Enquanto uma vasta maioria de 75,8% dos eleitores não expressa preferência por nenhum partido específico, quando questionados sobre rejeição, a indecisão diminui consideravelmente, com 38,2% afirmando não ter um partido em que nunca votariam. Esse contraste sugere que, embora os eleitores possam não se identificar com uma legenda específica, eles têm mais clareza sobre as opções que definitivamente não considerariam.

O Partido dos Trabalhadores (PT), que contava em 2020 com 7,1% da preferência dos entrevistados, apresenta um índice de rejeição substancialmente mais alto, com 32,6% dos participantes afirmando que nunca votariam nesse partido. Por outro lado, o partido do presidente Jair Bolsonaro, que detém a preferência de 7,5% dos respondentes, também aparece com uma taxa significativa de rejeição de 14,8%.

No modelo multivariado que será desenvolvido, as variáveis relativas à preferência partidária e à rejeição a partidos serão tratadas por meio de variáveis binárias, ou *dummy*. Essa metodologia nos permite discernir a influência da preferência partidária específica, ou de uma rejeição declarada a determinados partidos, sobre a frequência de uso do Google para a busca de informações políticas.

Será levado em consideração, portanto, se os entrevistados identificam algum partido como de sua preferência ou se existe algum partido que categoricamente não receberia seu voto.

5.2.3 *Variáveis de controle*

Em um modelo multivariado, a inclusão de variáveis de controle é importante para garantirmos a robustez das análises e a precisão dos resultados. Variáveis de controle, como renda, escolaridade, raça, gênero e religião, são fundamentais pois permitem que o modelo estatístico ajuste-se a partir de fatores que podem influenciar a variável dependente, mas que não são o foco principal da pesquisa.

Ao controlarmos essas características, é possível isolarmos o efeito específico das variáveis de interesse - no caso, preferência política consolidada - sobre o comportamento de busca no Google. A ausência dessas variáveis de controle poderia resultar em estimativas enviesadas, uma vez que outros fatores não considerados poderiam explicar parte da variação na variável dependente. Assim, ao incluirmos essas variáveis no modelo, reduzimos o risco de se atribuir incorretamente a influência de variáveis omitidas àquelas que estão sendo efetivamente examinadas, aumentando assim a confiabilidade das conclusões do estudo.

A inclusão de variáveis de controle específicas como renda, escolaridade, raça, gênero e religião em um modelo multivariado é justificada pelo reconhecimento de que essas características socioeconômicas e demográficas podem exercer influência significativa no comportamento eleitoral e no acesso e uso de informações políticas.

Por exemplo, a renda pode afetar a disponibilidade de recursos tecnológicos e, conseqüentemente, a facilidade de acesso ao Google para pesquisas políticas. A escolaridade,

por sua vez, está frequentemente associada ao nível de habilidade crítica e de informação política, o que pode influenciar como as pessoas utilizam o Google durante as eleições. As variáveis de raça e gênero nos permitem analisar de que modo diferentes grupos sociais interagem com as informações políticas, considerando-se as diversas experiências e contextos que podem moldar tais interações. Região é fundamental em um país tão diverso, principalmente em pleitos sub-nacionais, objeto deste estudo. Por fim, a religião pode ser um fator determinante nas preferências políticas e na mobilização em torno de questões eleitorais, afetando potencialmente a maneira como os indivíduos procuram informações.

Ao controlar esses fatores, o modelo pode proporcionar uma visão mais clara da relação entre as variáveis de interesse principal e o uso do Google para informações políticas, descontaminando os efeitos de outros determinantes sociais e individuais.

Vale ressaltarmos que, ainda que não se considere probabilística e representativa da população brasileira a amostra desta pesquisa, o controle de cota na aplicação foi rigoroso, o que trouxe uma amostra semelhante à da população brasileira em termos de classe, gênero, faixa etária e região. A exceção se deu para a variável educação, que está sub-representada no nível de analfabetismo e Ensino Fundamental. Esse é um aspecto importante, mas compreensível pela técnica de coleta (painel on-line) e pelo modo de aplicação do instrumento de pesquisa (questionário autoaplicado), que foram escolhas essenciais para se responder à pergunta de pesquisa.

Além dessas variáveis, também foi incluída a variável “uso do google”, que é a contagem de quantas buscas o painalista que forneceu os dados históricos fez.

O descritivo das variáveis de controle estão no apêndice 10.2 deste trabalho.

5.3 Como medir o uso do Google - Variáveis Dependentes

Variáveis dependentes são aquelas que se busca explicar por meio da representação de efeitos ou resultados que são hipoteticamente influenciados por outras variáveis (independentes). No contexto de uma análise empírica, espera-se que essas medidas de resultado variem em resposta a diferentes condições experimentais ou correlações observadas.

No presente estudo, abordamos duas variáveis dependentes que capturam a frequência das buscas políticas do entrevistado, cada uma delas derivada de dois processos de coleta distintos, mas que se referiam às mesmas pessoas.

Quantitativos gerais: a 1ª aplicação da pesquisa na véspera do início do período

eleitoral teve o total de 2.771 respondentes. A segunda, em que todos os participantes iniciais foram convidados a responder, teve 2.017 respostas. Desses últimos que responderam à segunda rodada, 982 eram painelistas que possuíam históricos de buscas disponíveis.

5.3.1 *Uso monitorado do histórico de navegação*

A primeira variável foi construída a partir de uma metodologia que envolve o exame de dados do histórico de navegação dos usuários. Esse processo envolveu a classificação de buscas como “eleitorais” ou não, baseando-se na análise de conteúdo dos resumos exibidos nos resultados de pesquisa do Google.

Esse método, melhor detalhado na seção metodológica, fornece uma medida da frequência de buscas políticas efetivamente realizadas, mitigando as limitações associadas aos dados autodeclarados.

Portanto, essa variável reflete o número quantitativo de buscas eleitorais identificadas para cada respondente do painel, agregando rigor empírico à investigação do comportamento informacional dos eleitores ao longo do período eleitoral.

Esses 982 fizeram um total de **352.726** buscas, com **101.367** diferentes termos/*queries* de busca.

Após a identificação de quais dessas buscas se referiam àquelas político-eleitorais (conforme já explicado na seção metodológica), encontramos um total de 2.952 buscas, com 1.098 diferentes termos/*queries* de busca. Todas essas buscas político-eleitorais foram realizadas por 312 respondentes.

A tabela 12 mostra os quantitativos para buscas, termos únicos e painelistas que disponibilizaram seus dados. **Os números encontrados são altos ou baixos?**

Tabela 12 – Quantitativos gerais do histórico de busca dos usuários

	Buscas	Termos	Painelistas
Geral	352.726	101.367	982
Política-Eleitoral	2.952	1.098	312
(%)	0.84 %	1.08 %	31.77 %

Fonte: elaborado pelo autor.

Essa simples pergunta pode ser respondida a partir de uma comparação com o único estudo que possui um desenho próximo ao desta pesquisa. O trabalho de Blassnig *et al.* (2023), que utilizou o histórico de buscas de eleitores suíços para entender o uso do Google em referendos

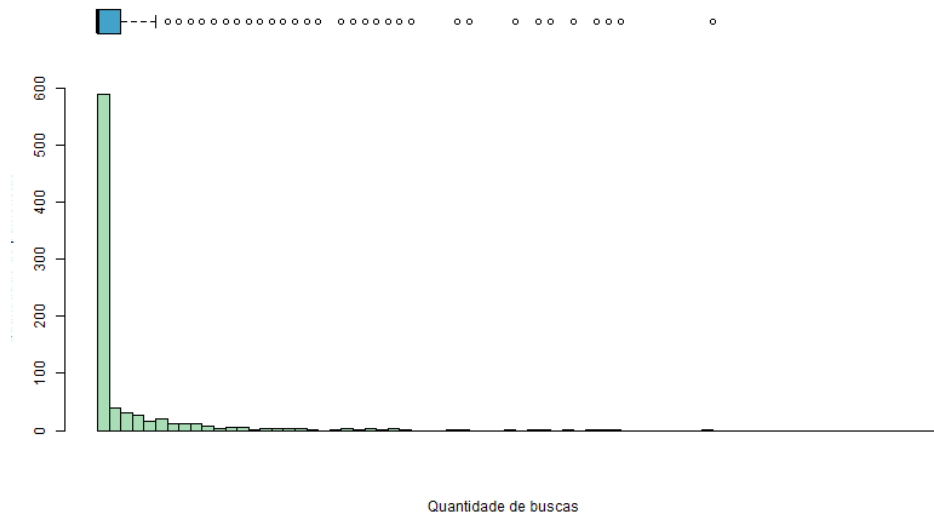
no país europeu encontrou valores menores. Dos 128 usuários monitorados, apenas 27 fizeram alguma busca política-eleitoral (21%) e dos 117.739 mil termos de busca, apenas 90 faziam referência a um uso político-eleitoral (0.07%).

Criação da variável

Ao final de todo o processo, criou-se a variável dependente chamada “Uso monitorado”, que indica se a pessoa fez a quantidade de buscas eleitorais que foram identificadas no histórico de busca de cada um dos painelistas. O gráfico 12 mostra a distribuição dessa variável.

Transformação em uma variável binária

Figura 9 – Distribuição da variável - Uso Monitorado do Google



Fonte: elaborado pelo autor.

Como é possível notar, o gráfico 9 mostra de que modo a distribuição da quantidade de buscas com base no histórico dos painelistas não segue uma distribuição normal. Para utilizar essa variável em nosso modelo, seria importante ajustá-lo, por exemplo utilizando-se a função LOG. Para evitar essa transformação e facilitar a leitura comparativa entre esta variável com a próxima, optamos por transformá-la em uma variável binária onde 1 é o painalista que fez alguma busca política (37,7%) e 0 aquele que não fez alguma busca.

5.3.2 *Uso declarado pelas respostas ao questionário*

A segunda variável é baseada em dados autodeclaratórios obtidos por meio de um questionário em que os respondentes avaliam a frequência com que realizaram buscas relacionadas a eleições no Google, numa escala ordinal que varia de “Nunca” a “Sempre”. A natureza autodeclarada dessa variável requer consideração cuidadosa de possíveis vieses, como a desajustabilidade social ou a imprecisão na autopercepção. Uma maneira inicial de avaliarmos

Tabela 13 – Resultados para a pergunta: Ao longo do período eleitoral, com qual frequência você fez cada um desses tipos de busca sobre eleições no Google?

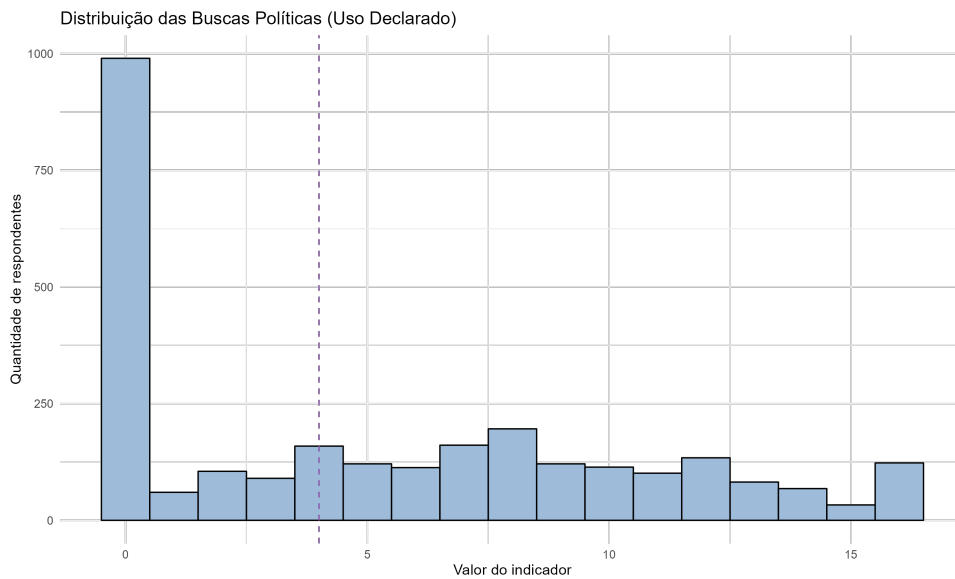
	Nunca	Raramente	Algumas vezes	Frequentemente	Sempre	Sem resposta
Checar se uma informação que recebeu era verdadeira	20,4%	12,2%	22,1%	17,1%	26,5%	1,7%
Para se informar sobre os candidatos	21,3%	15,2%	29,2%	14,7%	17,9%	1,7%
Para se informar sobre procedimentos da votação (horários, locais, como justificar etc.)	23,5%	20,4%	26,7%	13,9%	13,8%	1,7%
Para se informar sobre quem estava à frente nas pesquisas	24,1%	17,8%	25,9%	16,2%	14,2%	1,7%

Fonte: elaborado pelo autor.

a tabela 13 e de determinarmos a abordagem para a construção da variável dependente é a elaboração de um índice. Uma possibilidade é construí-lo a partir da soma de respostas a quatro questões, em que cada uma é avaliada em uma escala ordinal de frequência de uso. Essa escala é quantificada da seguinte maneira: “Nunca” equivale a 0, “Raramente” a 1, “Algumas vezes” a 2, “Frequentemente” a 3 e “Sempre” a 4. Portanto, o indicador é uma medida discreta, com valores possíveis variando de 0 (nenhum uso) a 16 (uso máximo em todas as categorias).

Essa abordagem pressupõe que as distâncias entre as categorias são uniformes, o que é uma suposição comum, porém limitada, em pesquisas que convertem dados ordinais em quantitativos. Assim, a variável “Uso declarado” serve como um indicador agregado da intensidade do uso do Google para fins eleitorais, refletindo a autoavaliação dos respondentes quanto à frequência de suas buscas relacionadas às eleições.

O gráfico 10 mostra a distribuição desse indicador.

Figura 10 – Distribuição das Buscas Políticas (Uso Declarado)

Fonte: elaborado pelo autor, com base no histórico dos painelistas.

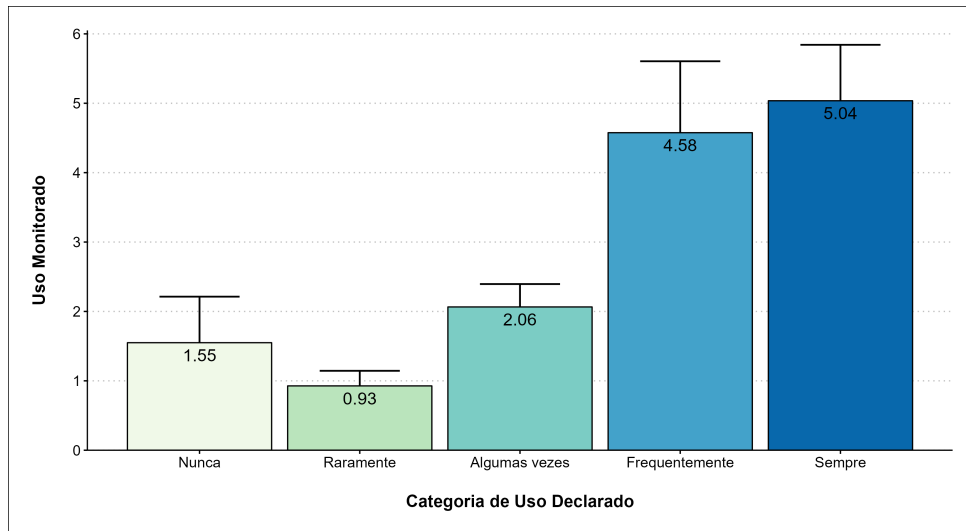
5.3.3 Comparação entre as duas variáveis

No intuito de realizarmos uma análise comparativa mais refinada entre as percepções autodeclaradas e os comportamentos reais de busca no Google, criamos uma nova variável: “Uso Declarado Categorizado”. Essa variável é derivada do indicador apresentado na seção anterior, que representa a soma dos valores ordinais atribuídos às respostas dos participantes sobre a frequência de uso do Google para diferentes tipos de buscas relacionadas às eleições.

A categorização foi realizada utilizando-se intervalos específicos que correspondem às categorias originalmente utilizadas no questionário: “Nunca” (valores menores que 1), “Raramente” (1 a 4), “Algumas vezes” (5 a 8), “Frequentemente” (9 a 12) e “Sempre” (mais de 12).

Essa estratégia permite agrupar os respondentes de maneira que isso reflita mais precisamente a frequência autorreportada de suas buscas eleitorais no Google, facilitando a comparação com a variável discreta que mede a quantidade real de buscas políticas realizadas. A inclusão da categoria “Nunca” é particularmente relevante, pois captura especificamente os indivíduos que não utilizaram o Google para esse fim durante o período eleitoral.

O teste de ANOVA foi realizado para verificarmos se existem diferenças estatísticas significativas entre as médias de “Uso Monitorado” nas diferentes categorias de “Uso Declarado”. O resultado foi significativo, indicando a existência de diferenças notáveis entre os grupos.

Figura 11 – Média de Uso Monitorado por Uso Declarado

Fonte: elaborado pelo autor.

Esses resultados sugerem uma correlação positiva entre a frequência autodeclarada de uso do Google para buscas eleitorais e o número de buscas políticas reais. A categoria “Sempre” destaca-se, indicando que os usuários que declaram utilizar frequentemente o Google para buscas eleitorais demonstram um número maior de buscas políticas reais em comparação com as outras categorias.

5.3.3.1 Quando se transforma uma escala ordinal em binária

Os resultados deste estudo revelam nuances metodológicas significativas para o tratamento de dados coletados por escalas ordinais. Frequentemente, em pesquisas comportamentais - seja na investigação de padrões de consumo de mídia, hábitos políticos ou outros comportamentos recorrentes - utilizam-se escalas de frequência para se captar a regularidade de determinadas ações. As respostas a essas escalas, embora ricas em informação, muitas vezes são convertidas em variáveis binárias (dicotomizadas) para simplificar a análise estatística, especialmente quando se pretende aplicar modelos multivariados como regressões logísticas ou lineares.

A dicotomização é adotada por pesquisadores, principalmente, para facilitar a interpretação dos modelos estatísticos. Ao transformar uma variável ordinal em binária, simplificamos o espectro de respostas possíveis para um formato de “presença” ou “ausência” do fenômeno em questão. No entanto, esse processo de simplificação deve ser conduzido com cautela. A escolha de qual categoria representará o valor “1” (indicativo de presença ou resultado positivo) não deve ser arbitrária, mas sim embasada em evidências empíricas e teóricas. No caso do presente estudo, os respondentes que declararam usar o Google “Sempre” para buscas eleitorais

não apenas demonstraram um comportamento consistente com suas declarações, mas também tiveram essa autodeclaração confirmada por meio de dados objetivos de uso. Esse achado apoia a decisão de se considerar a resposta “Sempre” como a categoria positiva na conversão em variável binária.

Além disso, é fundamental que pesquisadores estejam atentos aos possíveis vieses de autopercepção. Respostas como “Nunca”, “Raramente” e “Algumas vezes” podem ser influenciadas tanto pela falibilidade da memória humana quanto pela tendência de fornecer respostas socialmente desejáveis. Isso é agravado quando os comportamentos investigados são frequentes e rotineiros, como é o caso do uso do Google, tornando-se quase automáticos e menos suscetíveis à autoanálise consciente.

Em contextos em que a precisão e a fidelidade dos dados são cruciais, é imprescindível que métodos de coleta mais confiáveis sejam empregados. Isso pode incluir, por exemplo, o uso de registros de atividades ou dados de rastreamento, que podem fornecer um contraponto valioso às autodeclarações, especialmente para se aferir a veracidade das respostas dos extremos das escalas de frequência.

5.3.3.2 *Quando os dados desmentem as autodeclarações*

Um segundo aspecto intrigante emerge ao considerarmos que indivíduos que negaram veementemente o uso do Google para pesquisas políticas, conforme suas autodeclarações, tiveram, não obstante, um registro histórico de buscas que contradiz essa negativa. A média dessas buscas, ainda que discrepante do declarado, supera o valor nulo. Tal contradição pode ser interpretada sob duas luzes distintas, ambas com implicações metodológicas profundas.

Por um lado, essa divergência nos dados sugere um desafio inerente à mensuração de atividades habituais. A frequência com que se recorre a serviços como o Google pode se perder na rotina diária, tornando-se uma ação quase inconsciente e, por isso, difícil de ser relatada com precisão em retrospectiva. A natureza automática e recorrente do uso de ferramentas digitais pode, assim, obscurecer a autoavaliação e a autopercepção da frequência real de uso.

Por outro lado, a onnipresença do Google como plataforma de pesquisa é inegável; sua integração no tecido do dia a dia é tão completa que muitas interações podem passar desapercibidas ou ser esquecidas. Isso pode resultar em uma subestimação sistemática do uso efetivo nas autodeclarações dos usuários.

A discrepância entre o número de buscas autorreportadas e as registradas nos dados

de navegação ressalta a necessidade vital de se empregar abordagens metodológicas que possam capturar de maneira objetiva o comportamento dos usuários. Métodos que envolvem o monitoramento direto das atividades on-line, por exemplo, oferecem um contraponto factual e robusto às limitações das métricas auto declaradas e podem fornecer um panorama mais fidedigno das práticas dos indivíduos na esfera digital.

5.3.4 Definição da variável dependente

Diante dos testes realizados, optamos por refinar a variável “Uso Autodeclarado” para que representasse um alto grau de engajamento na busca de informações eleitorais no Google. A variável passou a ser, então, binária, definida de modo que o valor 1 indique que o respondente se engajou frequentemente ou sempre nas buscas relacionadas às eleições no Google, conforme representado por um valor acima de 9 no indicador criado. Isso corresponde às categorias “Frequentemente” (valores de 9 a 12) e “Sempre” (valores acima de 12). Por outro lado, o valor 0 passou a ser atribuído aos respondentes que se enquadraram nas categorias “Nunca”, “Raramente” ou “Algumas vezes”, representando um uso menor ou inexistente do Google.

Essa codificação binária ainda que diminua a possibilidade explicativa nos ajudará a comparar os dois modelos que serão construídos com as duas diferentes variáveis e ganharemos em termos de compreensão.

Tabela 14 – Variável dependente: Uso declarado

	Freq.	Prop.
Não fizeram buscas	1995	72,0%
Fizeram buscas	776	28,0%
Frequentemente e Sempre		
Total	2771	100%

Fonte: elaborado pelo autor.

Tabela 15 – Quadro resumo das variáveis

Variáveis Dependentes		
Variável	Descrição	Tipo e Características
d_sum_gglesearch_w2	Uso Declarado	Binária - Assume valor de Referência para valores acima de 9 no indicador ordinal que agrega respostas de frequência de uso
d_pol_searches	Uso Monitorado	Binária - Assume valor 1 para qualquer respondente que teve alguma busca identificada

Fonte: elaborado pelo autor.

5.4 Diferentes perfis para se entender o envolvimento do eleitor

A análise de dados em Ciências Sociais e Comportamentais frequentemente recorre a modelos multivariados para investigar a relação entre diversas variáveis e seus impactos sobre fenômenos de interesse. Apesar de sua ampla aplicabilidade e robustez, modelos multivariados tradicionais enfrentam críticas quanto à sua capacidade de fornecer uma visualização mais intuitiva do indivíduo dentro do espectro analítico.

Uma limitação significativa desses modelos reside na dificuldade de se representar a multidimensionalidade das interações humanas de uma forma que se destaque a singularidade e a complexidade dos perfis individuais. A ênfase em estimativas pontuais e associações entre variáveis, embora informativa, muitas vezes subestima a riqueza dos padrões comportamentais e as nuances que caracterizam a experiência humana.

Além disso, a abordagem predominante nos modelos multivariados tende a abstrair os indivíduos a conjuntos de variáveis, perdendo, em certa medida, a conexão direta com as manifestações concretas de suas escolhas e preferências. Essa abstração pode limitar nossa capacidade de interpretar os resultados de uma maneira que ressoe as realidades vividas pelos indivíduos estudados.

Em contraste, técnicas que priorizam a identificação de padrões e grupos dentro dos dados oferecem uma oportunidade para reenquadrarmos nossa análise em uma perspectiva que valoriza a heterogeneidade individual e facilita a compreensão das estruturas subjacentes que orientam o comportamento humano.

Nesse contexto, a possibilidade de agruparmos os dados em grupos emerge como

uma ferramenta didática e intuitiva, permitindo a identificação de grupos de indivíduos que compartilhem características semelhantes, adotando-se como base uma análise de suas interações variáveis.

Essa abordagem pode complementar as limitações dos modelos multivariados no que tange a visualização e interpretação dos dados, mas também proporciona uma perspectiva rica e detalhada sobre a configuração dos grupos, refletindo de maneira mais fiel a complexidade dos fenômenos em estudo.

Ao priorizarmos a formação de grupos baseada na semelhança de perfis em um espaço multidimensional, a clusterização nos permite visualizar o “indivíduo” de forma mais concreta e didática, destacando-se as relações intrínsecas que muitas vezes são obscurecidas em análises mais tradicionais.

Portanto, optamos por investigar a interação entre as três variáveis do primeiro bloco de variáveis para testarmos a Hipótese 1: Conhecimento sobre os candidatos e participação, com o objetivo de criarmos perfis analíticos mais interessantes para analisarmos “para quem o Google é relevante”.

5.4.1 O uso de análise de correspondência múltipla (ACM)

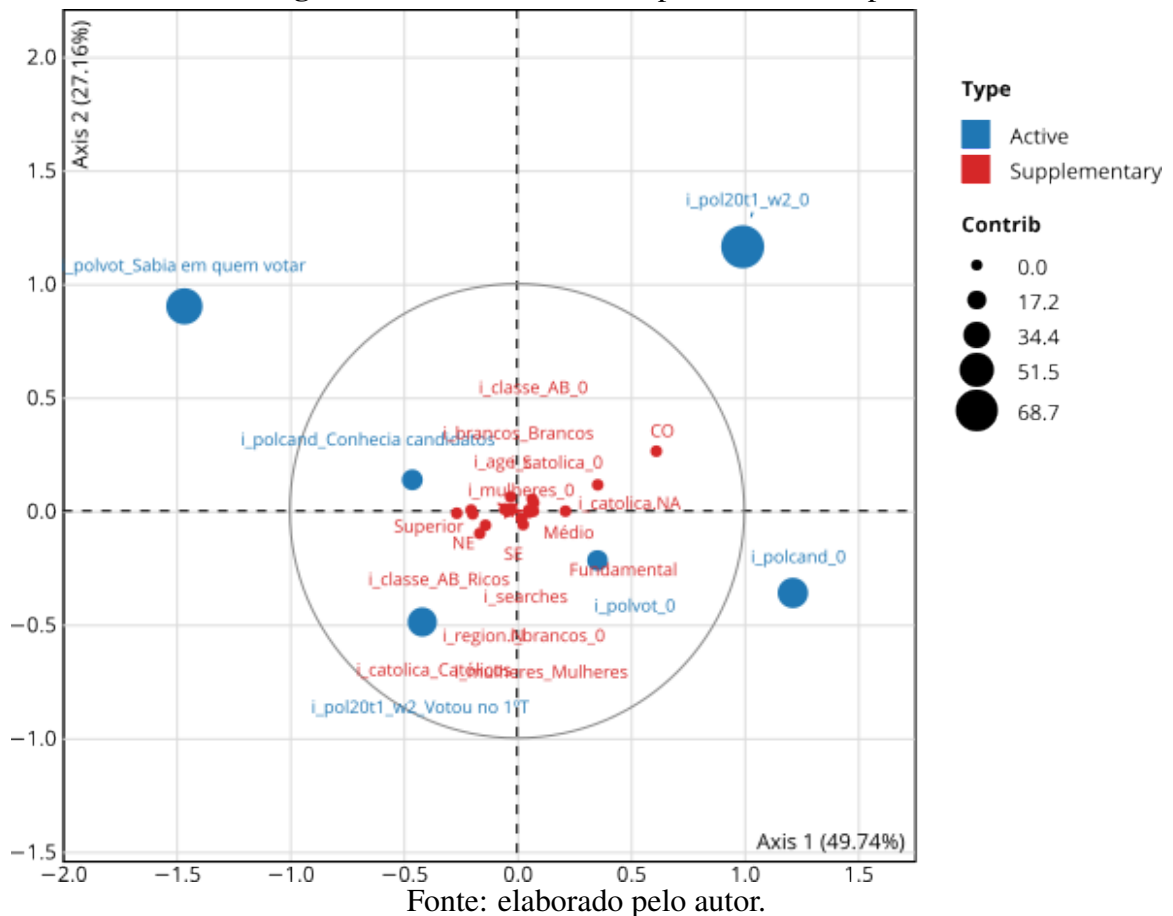
A escolha metodológica de empregar a análise de correspondência múltipla (ACM) (Lê *et al.*, 2008) seguida de uma clusterização hierárquica, ao invés de se aplicar diretamente as três variáveis específicas no modelo, foi motivada por duas razões: a interação entre as variáveis independentes se mostraram mais úteis do que a análise sem o uso da sua interação; a maior capacidade analítica dos agrupamentos realizados.

Primeiramente, a ACM é uma técnica exploratória de redução de dimensionalidade adequada para variáveis categóricas, permitindo a identificação e interpretação das relações entre múltiplas categorias de variáveis. Essa técnica é particularmente valiosa em contextos em que as interações entre variáveis são complexas e não facilmente discerníveis. A ACM facilita a visualização dessas relações em um espaço de dimensões reduzidas, revelando padrões e estruturas ocultas que poderiam permanecer obscuros em análises diretas das variáveis.

Ao aplicar a ACM, conseguimos transformar um conjunto de variáveis categóricas em um conjunto de dimensões latentes ou fatores, que capturam a variação máxima entre os perfis dos respondentes. Esses fatores, por sua vez, oferecem uma representação simplificada, mas abrangente, das interações complexas entre as variáveis.

A análise das dimensões resultantes permite uma compreensão mais matizada das dinâmicas subjacentes ao conjunto de dados, indo além da análise superficial que poderia ser obtida pelo exame direto das variáveis.

Figura 12 – Análise de Correspondência Múltipla



O gráfico 12 é um mapa de fatores da análise de correspondência múltipla (ACM), uma técnica utilizada para se explorar e visualizar as relações entre diversas variáveis categóricas. As duas dimensões principais representadas (dimensão 1 no eixo horizontal e dimensão 2 no eixo vertical) capturam a maior parte da variação nos dados, com a dimensão 1 explicando 49,74% e a dimensão 2 explicando 27,26% da inércia total. O gráfico gerado por essa análise, conhecido como biplot, mostra as variáveis ativas, que são as principais responsáveis pela variação nos dados e definem os eixos do gráfico, bem como as variáveis suplementares, que são adicionadas posteriormente e não influenciam na formação dos eixos.

As variáveis ativas são aquelas utilizadas na análise para a construção do espaço multidimensional no qual os dados são projetados.

Nesta pesquisa, as variáveis ativas foram “Conhecer algum candidato”, “Saber em quem iria votar” e “Se votou no 1º Turno das eleições”.

Por outro lado, as variáveis suplementares são adicionadas à análise após a definição dos eixos principais, servindo para enriquecer a interpretação dos resultados sem alterar a sua configuração. No contexto deste estudo, foram as variáveis de controle, já explicadas na seção

anterior. Assim, elas nos ajudam a contextualizar os padrões observados nas variáveis ativas sem influenciar diretamente a estrutura do espaço de dados da MCA.

A análise das variáveis ativas pode revelar quais fatores são mais decisivos nas decisões políticas dos eleitores. Pode também ajudar a identificar grupos de eleitores com perfis semelhantes ou divergentes com base em suas práticas de busca de informação e conhecimento político.

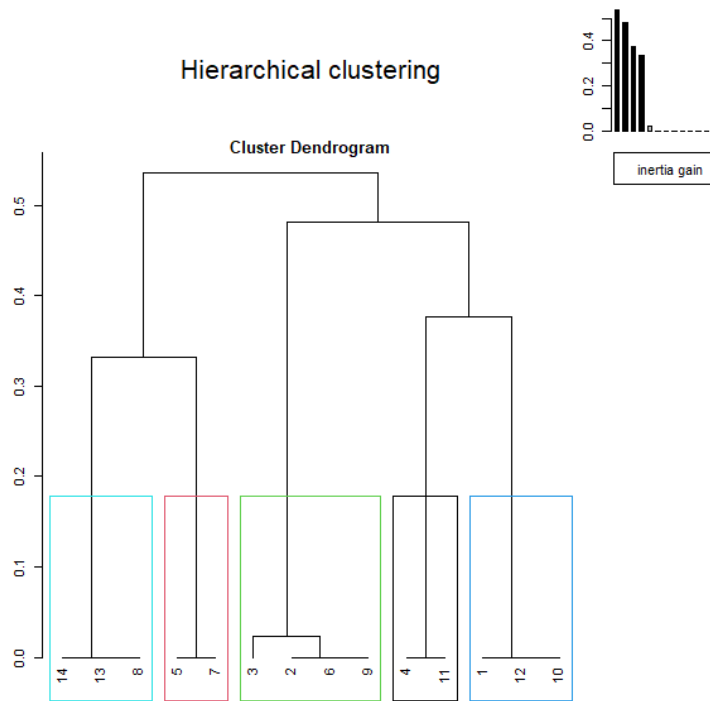
Analisando-se o gráfico, podemos identificar agrupamentos e oposições entre as variáveis. Por exemplo, “i_polvot_Sabia em quem votar” está oposto a “i_polvot_0”, indicando uma distinção clara entre os respondentes que já sabiam em quem votar e aqueles que não sabiam. Da mesma maneira, “i_polcand_Conhecia candidatos” e “i_polcand_0” estão em lados opostos da dimensão 1, sugerindo uma distinção entre os respondentes que conheciam os candidatos e os que não conheciam. A variável “i_pol20t1_w2_Votou no 1T” (votou no primeiro turno) aparece próxima ao eixo central, sugerindo que essa ação pode não ser tão distintiva para separar os respondentes em termos de suas outras características políticas. Em suma, o mapa de fatores da ACM oferece uma representação visual das relações complexas entre as atitudes e comportamentos eleitorais dos respondentes.

Quanto às variáveis suplementares, mesmo que não variem significativamente em relação aos componentes principais, o fato de serem centradas no gráfico sugere que, isoladamente, não discriminam entre os perfis de eleitores identificados pelas variáveis ativas. A ausência de oscilação significativa nas variáveis suplementares sugere que a combinação das variáveis ativas é suficientemente robusta para explicar a variabilidade nos dados.

Subsequentemente, a utilização de uma clusterização hierárquica sobre as dimensões derivadas da MCA permite a identificação de grupos de indivíduos com perfis semelhantes de maneira mais natural e fundamentada. Essa abordagem proporciona uma visualização intuitiva do modo como os indivíduos se agrupam com base nas dimensões latentes, facilitando a identificação de padrões de agrupamento e a determinação do número ótimo de clusters, neste caso, cinco grupos distintos de indivíduos.

Portanto, a combinação dessas técnicas - MCA seguida de clusterização hierárquica - oferece uma abordagem robusta e sofisticada para desvendarmos a complexidade e a multidimensionalidade das preferências e comportamentos dos indivíduos.

Essa metodologia não apenas maximiza a variação explicada em um espaço de dimensões reduzidas, mas também permite uma segmentação dos respondentes em grupos

Figura 13 – Árvore de distribuição hierárquica

Fonte: elaborado pelo autor.

homogêneos, com base na interação entre as variáveis, em vez de suas características isoladas.

O dendrograma ilustrado na figura 13 mostra a existência de diferentes clusters, identificados por linhas coloridas que cortam o dendrograma em diferentes alturas. Onde a linha corta as ramificações verticais do dendrograma está indicada a formação de clusters. Por exemplo, a linha preta corta o dendrograma na parte inferior, indicando um ponto de corte que define um maior número de clusters menores.

A correlação das variáveis dependentes para cada cluster identificou 5 diferentes perfis que são o resultado de três diferentes variáveis e que podem ser visualizados na tabela 16.

Tabela 16 – Distribuição dos perfis

	Freq.	Prop.
1. Não votou no 1ºT e não conhecia algum candidato	260	12,9%
2. Não votou no 1ºT e conhecia algum candidato	293	14,5%
3. Votou no 1ºT e não conhecia algum candidato	295	14,6%
4. Votou no 1ºT e Conhecia algum candidato	776	38,5%
5. Sabia em quem iria votar	393	19,5%
Total	2017	100%

Fonte: elaborado pelo autor.

De maneira intuitiva, a classificação dessas variáveis mostra uma “escada” na participação eleitoral e nos ajuda a pensar os eleitores brasileiros em diferentes categorias⁷.

5.5 Teste da hipótese

O objetivo deste capítulo é entender para quem o Google é relevante. Essa análise possui a seguinte hipótese:

H1: Eleitores com preferência eleitoral consolidada tendem a utilizar **mais** frequentemente o Google para buscar informações sobre eleições.

Ao longo deste capítulo, demonstramos as variáveis dependentes que medirão a preferência eleitoral consolidada e que buscam explicar o maior ou menor uso do Google para buscas políticas. A interação de três dessas variáveis (se votou no 1º turno, se conhecia algum candidato e se sabia em quem votar) foi resumida em uma única variável chamada “Perfil do Eleitor”, com cinco modalidades distintas. Duas outras, ter algum partido pelo qual tem preferência, também fazem parte de como estamos medindo a preferência eleitoral consolidada.

A técnica que será adotada no modelo é o uso de uma regressão logística com a seguinte especificação:

$$\begin{aligned} \log\left(\frac{p}{1-p}\right) = & \beta_0 + \beta_1 \times \text{Perfil do Eleitor} \\ & + \beta_2 \times \text{Ter Partido de Preferência} + \beta_3 \times \text{Ter um partido que nunca votaria} \\ & + \beta_4 \times \text{ClasseAB} + \beta_5 \times \text{Branco} + \beta_6 \times \text{Católicos} \\ & + \beta_7 \times \text{Região} + \beta_8 \times \text{Nível de Educação} + \beta_9 \times \text{Mulheres} + \beta_{10} \times \text{Idade} \\ & + \beta_{11} \times \text{Uso do Google} \end{aligned} \tag{5.1}$$

onde: - p é a probabilidade de uso do Google para informação política-eleitoral, - β_0 é o intercepto do modelo, - $\beta_1, \beta_2, \dots, \beta_{10}$ são os coeficientes que representam a influência de cada variável independente.

⁷ Dedico este pequeno trecho do trabalho ao querido Prof. Glaucio Ary Dillon Soares (f. 14/06/21), que tanto nos incentivou enquanto alunos a pensar em novos conceitos e a explorar novos jeitos de pensar.

5.5.1 Testes e robustez do modelo

Multicolinearidade

A análise dos fatores de inflação da variância (VIF) para os modelos logísticos monitorado e declarado revela que as variáveis independentes não sofrem de multicolinearidade severa, uma vez que todos os valores de VIF estão abaixo do limiar comumente aceito de 10 (disponível no apêndice 10.4).

Os valores de VIF mais altos, observados em ambas as análises, particularmente para Classe AB e Idade, no modelo monitorado, e Classe AB, no modelo declarado, sugerem uma moderada correlação com outras variáveis independentes, mas não a ponto de comprometer a integridade dos modelos. Isso indica que as variáveis podem ser mantidas nos modelos sem que causem distorções significativas nos coeficientes estimados.

Ajuste do modelo

O teste de Hosmer-Lemeshow, utilizado para se avaliar a adequação do ajuste dos modelos, mostra resultados distintos para os modelos monitorado e declarado. Para o modelo monitorado, o p-valor de 0.062 sugere uma adequação de ajuste marginalmente aceitável ao nível de significância de 0.05, indicando que o modelo tem um bom ajuste aos dados, embora esteja próximo do limite para rejeição da hipótese de bom ajuste.

Para o modelo declarado, com um p-valor de 0.54, indica claramente um bom ajuste do modelo aos dados observados, sugerindo que as discrepâncias entre as frequências observadas e esperadas são pequenas.

Resíduos padronizados

Com base nos gráficos de resíduos padronizados dos modelos logísticos monitorado e declarado, pode-se afirmar que os modelos parecem ter um bom ajuste aos dados. A ausência de padrões nos gráficos de resíduos e a sua distribuição em torno de zero sugerem que as suposições subjacentes do modelo logístico, como homocedasticidade e independência, não foram violadas. A falta de resíduos extremamente altos ou baixos indica que não há observações individuais que estão tendo um impacto desproporcional sobre o modelo.

Após verificarmos a consistência do modelo, a próxima etapa é a leitura e interpretação de seus coeficientes.

5.5.2 Resultados do modelo monitorado

Tabela 17 – Resultados para o Modelo Monitorado

Variável	Odds Ratio	Intervalo de Confiança	Valor-p
Uso do Google	1.00	1.00, 1.00	< 0.001 ***
Classe AB	1.07	0.71, 1.61	0.74
Branços	1.11	0.78, 1.57	0.57
Católicos	1.15	0.82, 1.62	0.41
Região			0.42
N	-	—	—
NE	0.82	0.36, 1.96	
SE	0.78	0.34, 1.82	
S	1.04	0.42, 2.64	
CO	0.47	0.15, 1.41	
Educação			0.22
Fundamental	-	—	—
Médio	1.30	0.68, 2.60	
Superior	1.71	0.85, 3.62	
Mulheres	0.75	0.54, 1.04	0.082 *
Idade	1.01	1.00, 1.02	0.19
Perfil			< 0.001 ***
1. Não votou e não conhecia	-	—	—
2. Não votou e conhecia	3.02	1.51, 6.28	
3. Votou e não conhecia	2.23	1.05, 4.82	
4. Votou e conhecia	3.69	1.97, 7.25	
5. Sabia em quem iria votar	4.28	2.14, 8.92	
Ter Partido de Pref.	1.21	0.79, 1.86	0.38
Ter Partido Nunca Votaria	0.83	0.58, 1.18	0.30
Observations	1,962		
Log Likelihood	-1,225.918		
Akaike Inf. Crit.	2,487.836		

Note: *p<0.1; **p<0.05; ***p<0.01

Fonte: elaborado pelo autor.

- *Uso do Google*: o coeficiente é estatisticamente significativo com um valor-p inferior a 0,001. No entanto, o *odds ratio* (OR) é de 1,00, o que sugere ter sido a variável provavelmente usada como referência ou controle. Em modelos logísticos, um OR de 1,00 indica que, para essa variável específica, não há aumento nem diminuição na probabilidade da variável dependente ocorrer em comparação com a categoria de referência.
- *Perfil do Eleitor*: essa é uma variável importante do modelo. Conforme esperado, há cinco perfis que demonstram uma maior proximidade do eleitor com as eleições;
 - *Não votou (no 1º turno) e não conhecia (os candidatos)*: serve como categoria de referência para as comparações seguintes;

- *Não votou e conhecia*: com um OR de 3,02 significa que, mantendo-se todas as outras variáveis constantes, os indivíduos que não votaram mas conheciam algum candidato têm aproximadamente três vezes mais chances de usar o Google para informações eleitorais em comparação com aqueles que não votaram e não conheciam candidatos;
- *Votou e não conhecia*: com um OR de 2,23, os indivíduos que votaram mas não conheciam os candidatos têm mais do que o dobro da probabilidade de usar o Google para informações eleitorais em comparação com a categoria de referência;
- *Votou e conhecia*: essa categoria tem um OR de 3,69, indicando que esses indivíduos têm quase quatro vezes mais chances de usar o Google para informações eleitorais do que os da categoria de referência;
- *Sabia em quem iria votar*: é a categoria com o OR mais alto, 4,28. Isso indica que os indivíduos que já haviam decidido seu voto tinham a probabilidade quatro vezes maior de usar o Google para buscar informações eleitorais em comparação com a categoria de referência;
- *Mulheres*: o OR é de 0,75, com um valor-p de 0,082, o que é marginalmente significativo (geralmente, um valor-p abaixo de 0,1 é considerado como tal). Isso sugere que as mulheres têm 25% menos chances do que os homens (a categoria de referência implícita) de usar o Google para informações eleitorais. No entanto, devido ao valor-p estar na margem da significância, essa conclusão deve ser tratada com cautela.

O modelo monitorado indica que a variável “Perfil do Eleitor” é um preditor significativo do uso do Google para informações eleitorais. Especificamente, aqueles que conhecem os candidatos ou têm um candidato preferido estão mais propensos a utilizar o Google para informações eleitorais. Isso sustenta a hipótese de que eleitores com preferência eleitoral consolidada tendem a buscar mais informações no Google sobre eleições.

É importante ressaltarmos que o modelo utiliza *odds ratios* para quantificar as associações. Um OR acima de 1 indica maior probabilidade, enquanto um OR abaixo de 1 indica menor probabilidade.

5.5.3 Resultados do modelo declarado

Tabela 18 – Resultados para o Modelo Declarado

Variável	Odds Ratio	Intervalo de Confiança	Valor-p
Classe AB	1.28	1.00, 1.63	0.048*
Branços	1.07	0.86, 1.31	0.55
Católicos	0.90	0.74, 1.10	0.29
Região			0.22
N	—	—	
NE	1.32	0.88, 1.99	
SE	1.23	0.83, 1.83	
S	1.05	0.67, 1.66	
CO	0.85	0.48, 1.47	
Educação			0.19
Fundamental	—	—	
Médio	1.16	0.87, 1.54	
Superior	1.36	0.97, 1.90	
Mulheres	1.07	0.88, 1.29	0.51
Idade	0.99	0.99, 1.00	0.056
Perfil			< 0.001***
1. Não votou e não conhecia	—	—	
2. Não votou e conhecia	1.73	1.12, 2.69	
3. Votou e não conhecia	2.20	1.44, 3.40	
4. Votou e conhecia	3.46	2.39, 5.10	
5. Sabia em quem iria votar	4.60	3.08, 6.97	
Ter Partido de Pref.	1.47	1.15, 1.88	0.002**
Ter Partido Nunca Votaria	1.31	1.06, 1.62	0.012*
Observations	1,962		
Log Likelihood	-1,225.918		
Akaike Inf. Crit.	2,487.836		

Note: *p<0.1; **p<0.05; ***p<0.01

Fonte: elaborado pelo autor.

O modelo declarado confirma e expande as tendências observadas, com algumas diferenças e mais variáveis significativas:

- *Classe AB*: a análise revela um OR de 1,28 (valor-p = 0,048), sugerindo que pessoas pertencentes à classe AB têm uma probabilidade 28% maior de utilizar o Google para informações político-eleitorais, comparativamente à classe de referência, que pode ser a classe C, D ou E;
- *Perfil do Eleitor*: aqui, as subcategorias mostram um aumento escalonado na probabilidade de uso do Google, conforme o nível de engajamento eleitoral aumenta:
 - *Não votou e conhecia*: com um OR de 1,73 (valor-p < 0.001), indica um aumento de

73% na probabilidade de usar o Google para buscas eleitorais em comparação com aqueles que não votaram e não conheciam nenhum candidato;

- *Votou e não conhecia*: o OR de 2,20 implica um aumento de 120% (um pouco mais que o dobro) na probabilidade;
- *Votou e conhecia*: com um OR de 3,46, a probabilidade é 246% maior;
- *Sabia em quem iria votar*: esse grupo, com um OR de 4,60, tem uma probabilidade 360% maior de usar o Google para tais buscas;
- *Ter Partido de Preferência*: o OR de 1,47 (valor-p = 0,002) sugere que ter um partido político de preferência aumenta a probabilidade de usar o Google para buscas político-eleitorais em 47%;
- *Ter Partido que Nunca Votaria*: com um OR de 1,31 (valor-p = 0,012), indica que pessoas que possuem um partido no qual nunca votariam têm uma probabilidade 31% maior de realizar buscas político-eleitorais no Google;
- *Idade*: apresenta um OR de 0.99 (valor-p = 0.056), o que sugere uma tendência de que, com o aumento da idade, há uma ligeira diminuição na probabilidade de uso do Google para informações político-eleitorais. Esse resultado está na margem da significância estatística.

Os resultados do “modelo declarado” indicam que fatores socioeconômicos (Classe AB) e ter preferência política eleitoral consolidada, aqui em todas as variáveis (Perfil do Eleitor, Ter Partido de Preferência e Ter Partido que Nunca Votaria) estão associados ao uso do Google para buscas relacionadas às eleições. Especificamente, o Perfil do Eleitor mostra uma relação positiva e gradativa com o uso do Google, alinhando-se com a ideia de que o interesse e envolvimento eleitoral podem motivar a busca por informações e, portanto, confirmando a hipótese.

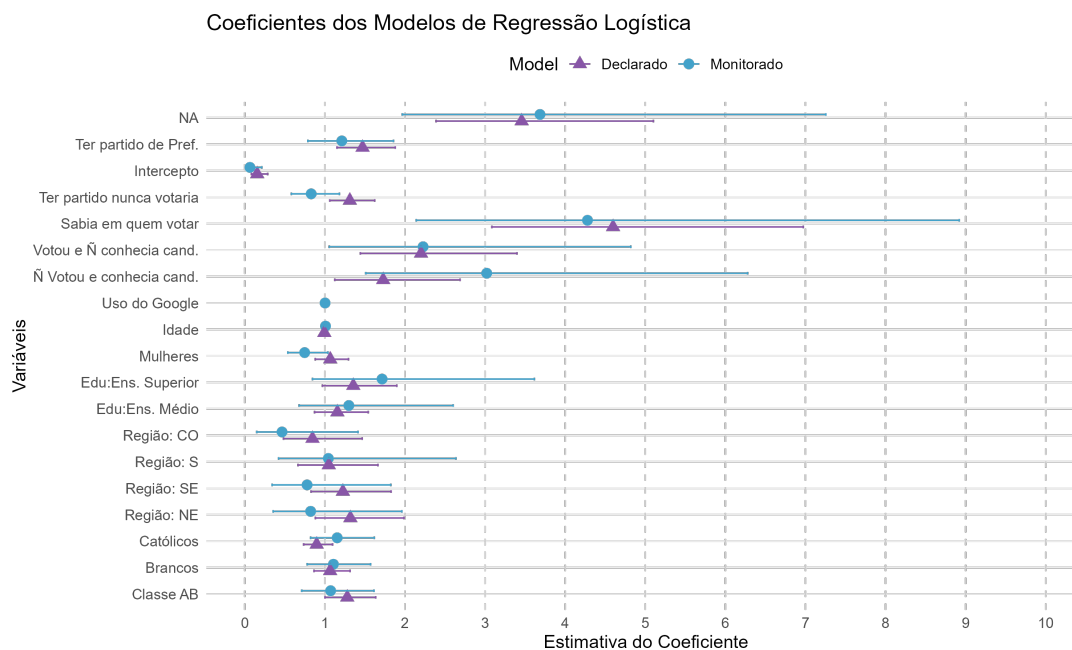
5.5.4 Comparação entre os modelos

A única variável que aparece como significativa em ambos os modelos é “Perfil do Eleitor”. Isso indica que, independentemente do método de coleta de dados — seja pela observação direta do comportamento on-line ou por meio de autodeclarações em questionários —, o nível de envolvimento e conhecimento político do eleitor está consistentemente associado ao uso do Google para buscas político-eleitorais. Essa consistência reforça a robustez dessa variável como um indicador-chave no estudo do comportamento eleitoral e na busca de informações políticas.

Algumas outras variáveis não foram constantes em todos os modelos:

- Classe AB: é significativa apenas no modelo declarado. Uma hipótese que pode explicar é a tradicional questão da desejabilidade social. Em outras palavras, os respondentes da Classe AB podem relatar comportamentos que acreditam ser consistentes com seu status social ou que sentem que são esperados deles.
- Ter Partido de Preferência e Ter um Partido que Nunca Votaria: são duas variáveis significativas apenas no modelo declarado. Há duas hipóteses possíveis para que essas duas variáveis apareçam apenas nesse modelo: a primeira delas é a própria desejabilidade social de um perfil que queira se manifestar “mais atuante”. O mesmo indivíduo que diz fazer buscas no Google diz ter preferência partidária. Uma segunda hipótese, não muito longe da primeira, é o contexto do questionário, que pode ser observado no apêndice deste trabalho. Foi um questionário com diversas perguntas que podem ter feito o eleitor “se engajar” politicamente e criar uma avaliação sobre a política brasileira. Ambas as hipóteses, ainda que muito próximas, podem mostrar como o instrumento de pesquisa pode, de fato, influenciar de que maneira a pessoa responde.

A variável “Mulheres” é marginalmente significativa no modelo monitorado, mas não no modelo declarado, ao contrário de “Idade”, que aparece como marginalmente significativa apenas no modelo declarado. Devido à significância marginal e seu baixo coeficiente, uma hipótese possível é a de que isso talvez seja apenas uma questão da amostra pesquisada, mas que serve de atenção para próximos estudos.

Figura 14 – Coeficientes dos Modelos de Regressão Logística

Fonte: elaborado pelo autor.

Um modo de se analisar a figura 14 é focar na diferença entre os modelos apenas para a principal variável “Perfil do Eleitor”.

- *Não votou e não conhecia*: como é a categoria de referência, não há o dado;
- *Não votou e conhecia*: é onde há a maior diferença. No modelo monitorado, o OR é de 3,02, indicando que esses eleitores são aproximadamente três vezes mais propensos a usar o Google para informações político-eleitorais. No modelo declarado, o OR é de 1,73, sugerindo um aumento menor, de cerca de 73%. É a maior diferença entre os modelos, de 1,29;
- *Votou e não conhecia*: há uma pequena diferença entre os modelos, com um OR de 2,23 no modelo monitorado e de 2,20 no modelo declarado, uma diferença de -0,03. Isso sugere uma concordância entre o comportamento real de busca e o comportamento reportado;
- *Votou e Conhecia*: no modelo monitorado, o OR é de 3,69, enquanto no modelo declarado é de 3,46, uma diferença de -0,23. Os eleitores que votaram e conheciam candidatos têm uma alta probabilidade de usar o Google em ambos os contextos, mas essa probabilidade é ligeiramente menor conforme o reportado nos questionários.
- *Sabia em quem iria votar*: para eleitores que tinham conhecimento claro de seu voto, o OR no modelo monitorado é de 4,28, comparado a um OR mais alto de 4,60 no modelo declarado, com uma diferença de +0,32. Isso pode sugerir que a intenção declarada de voto

se traduz em uma probabilidade ligeiramente maior de busca de informações no Google, conforme reportado em questionários.

Olhando os quatro perfis, a maior diferença está naquele que não votou. Uma possível hipótese para ser investigada em trabalhos futuros é analisar-se o fato de que a pergunta sobre ter ido votar foi imediatamente anterior às perguntas sobre o uso do Google. O respondente, ao dizer que não votou, pode ter superestimado de maneira retroativa sua não-atividade política, mais do que de fato possa ter acontecido.

5.6 Considerações finais

Search Engine Manipulation Effect (SEME) é superestimado

Os resultados deste capítulo mostram, de maneira clara, que o Search Engine Manipulation Effect (SEME) não existe, ou no mínimo é superestimado. Ou que pode até existir, mas não deve acontecer na realidade, pois um dos pressupostos do método experimental é que as pessoas não sejam pré-tratadas em relação àquilo a que elas serão expostas. Ou seja, não é adequado fazer um experimento sobre algo a respeito do que as pessoas já possuem opinião, pois o efeito que buscamos medir com o tratamento exposto dificilmente será mensurável (Druckman e Leeper, 2012; Coppock, 2018; Slothuus, 2015). Por isso, em seu desenho inicial, Epstein 2015 faz com eleitores americanos a avaliação de uma eleição de primeiro ministro da Austrália e, quando repete o experimento na Índia sobre uma eleição local, faz a ressalva de que, para quem tinha alguma familiaridade, o efeito é menor.

O que os dados evidenciaram aqui foi: quem busca por informação política é quem já conhece em maior medida algum dos candidatos ou já sabe em quem votar. Portanto, uma manipulação do Google dificilmente mudaria a opinião das pessoas no mundo real, fora dos laboratórios experimentais.

Google é relevante para indecisos

Um segundo dado relevante, para além da confirmação do raciocínio motivado, é que o Google também é relevante para eleitores que ainda não decidiram seu voto. A análise dos *odds ratio*, feita na seção anterior, mostra que o perfil que votou e não conhecia o candidato tem, aproximadamente, nos dois modelos, duas vezes mais chances de fazer alguma busca política-eleitoral. Esse número é a metade do perfil que já sabia em quem votar. Esse comparativo mostra que há uma importante análise a ser feita em estudos posteriores especificamente sobre o uso do Google em perfis com alguma proximidade eleitoral, mas que não tenham uma preferência

política tão consolidada assim.

A diferença entre autopercepção e um uso monitorado

A diferença entre os resultados dos modelos uso monitorado e uso declarado resalta o impacto que o método de coleta de dados tem nos resultados da pesquisa. Enquanto o modelo monitorado pode refletir o comportamento real mais precisamente, o modelo declarado pode capturar a autopercepção e as atitudes dos indivíduos, que são igualmente importantes no entendimento do comportamento eleitoral. As respostas dos questionários podem ser influenciadas pelo contexto das perguntas, pela percepção do que é socialmente desejável e pela tendência dos respondentes de apresentarem-se de uma maneira que esteja em conformidade com suas autoidentidades políticas. As discrepâncias observadas reforçam a importância de utilizarmos múltiplas abordagens metodológicas para obtermos uma compreensão abrangente do comportamento de busca de informação política.

Google é invisível ou o usuário não consegue responder um hábito cotidiano?

Neste capítulo, abordamos também a complexidade de se medir hábitos por meio de questionários. A discrepância observada entre as respostas autorrelatadas e os dados de uso monitorado sugere que há indivíduos que realizam buscas político-eleitorais no Google sem reconhecer essas ações como tal. Isso levanta a hipótese da ubiquidade do Google: a ferramenta está tão integrada ao cotidiano que sua utilização ocorre de maneira quase automática, fazendo com que os usuários não tenham plena consciência de seu uso específico (Coppock, 2018).

As decisões a respeito de quais tecnologias e aplicativos utilizar transcendem a funcionalidade prática, refletindo o status social e o estilo de vida, o que pode mascarar a percepção consciente do uso frequente dessas ferramentas (Graham, 2012). Isso sugere que, em vez de uma falha na memória ou atenção dos usuários, a onipresença do Google nos hábitos diários pode ofuscar sua visibilidade como uma entidade separada e identificável na realização de buscas específicas. Portanto, ao medir o comportamento de busca, deve-se considerar o papel das tecnologias que, embora onipresentes, são percebidas não como plataformas individuais, mas como extensões da atividade diária normal.

5.6.1 Diferentes perfis sob a ótica do comportamento político

Uma contribuição deste estudo foi a capacidade de se utilizar dados de painel e criar cinco diferentes perfis. Um exercício importante é a reflexão sobre como esses perfis dialogam com a literatura de comportamento político.

Não votou e não conhecia os candidatos - Desengajados

Esse perfil, usado como referência nos modelos, é caracterizado por um baixo engajamento e interesse político. Segundo a Teoria da Escolha Racional (Downs, 1957), a decisão de não votar reflete um cálculo racional no qual o custo percebido do voto supera os benefícios esperados, especialmente na ausência de conhecimento suficiente sobre os candidatos. Psicologicamente (Castro, 1994), a falta de conhecimento pode intensificar a sensação de alienação e desilusão com o sistema político. Sociologicamente (Bezerra e Mundim, 2011), a falta de informação reflete uma desconexão social mais ampla, na qual o voto parece menos relevante devido ao isolamento social ou limitações de mídia.

Não votou e conhecia os candidatos - Céticos

Esse perfil indica uma maior probabilidade de busca por informações políticas no Google em relação ao anterior, mesmo que isso não se traduza em votação. A Teoria da Escolha Racional explicaria isso como uma decisão ainda baseada em custo-benefício, mas com um eleitor mais informado ponderando que nenhum candidato atende adequadamente às suas necessidades (Lewis-Beck *et al.*, 2011). Psicologicamente, pode refletir uma alienação em relação à política, indicando que o conhecimento dos candidatos não mitigou o desencanto ou a desconfiança (Campbell *et al.*, 1960). Sociologicamente, a não-participação, apesar do conhecimento, pode ser entendida como uma resposta à percepção de falta de representatividade (Castro, 1994; Cervi, 2012).

Votou e não conhecia os candidatos - Conformistas

Esse grupo tem uma chance de fazer uma busca no Google de 2,23 e 2,20 nos modelos monitorado e declarado, respectivamente, mostrando que, mesmo sem um conhecimento prévio, a decisão de votar estimula a busca por informações. A Teoria da Escolha Racional vê isso como uma tentativa de se maximizar os benefícios da votação por meio de uma escolha mais informada (Fournier *et al.*, 2003; Peixoto e Rennó, 2011). Psicologicamente, a busca por informações pode ser uma maneira de se reduzir a incerteza e formar uma conexão mais forte com o processo eleitoral.

Votou e conhecia os candidatos - Indecisos

Com chances de 3,69 e 3,46, esse perfil mostra a maior probabilidade de buscar informações, refletindo maior interesse político. A Teoria da Escolha Racional sugere que esses eleitores fazem uma avaliação cuidadosa para otimizar os benefícios de sua escolha eleitoral (Fournier *et al.*, 2003). Psicologicamente e sociologicamente, o conhecimento prévio dos

candidatos e suas políticas fortalece a identificação e a lealdade partidária, influenciando a busca por informações que confirmem suas preferências e reforcem sua decisão.

Sabia em quem iria votar - Decididos

Esse perfil tem as maiores chances - 4.28 e 4.60 - nos modelos realizados, indicando uma forte tendência a buscar informações para confirmar ou justificar a decisão de voto já tomada. A alta busca de informações sugere um processo de *motivated reasoning* em que os eleitores, firmes em suas decisões, utilizam as informações para reforçar suas crenças Taber e Lodge (2006). Sociologicamente, as buscas podem refletir o interesse por informações que validem a adequação do candidato às demandas sociais do eleitor.

6 COMO E QUANDO O GOOGLE É RELEVANTE

O capítulo anterior concentrou-se na análise quantitativa dos eleitores que utilizam o Google para informações políticas. Agora, avançamos para a próxima fase desta investigação: uma avaliação qualitativa do conteúdo das buscas. Este capítulo avança na compreensão das interações dos eleitores com o processo eleitoral por meio da análise do que eles buscaram no Google durante as eleições brasileiras de 2020.

Aqui, a metodologia também combina os dados declarados das pesquisas realizadas com os históricos de busca. Os termos de busca foram classificados e nos fornecem uma base para compreendermos as motivações dos eleitores e suas respostas aos eventos políticos.

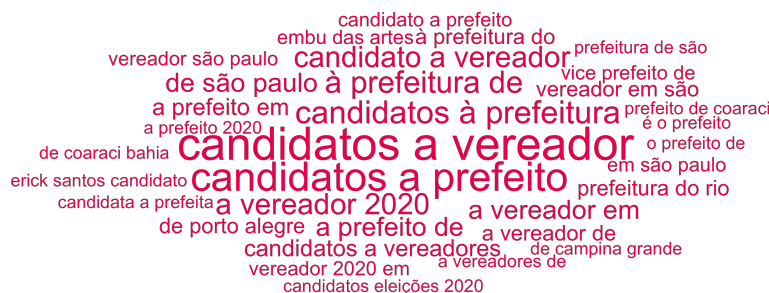
6.1 Termos de busca enquanto preferência

Após o processo de identificação a respeito de se um termo de busca fazia referência ou não a uma busca política, e também após a criação da variável dependente “uso monitorado”, utilizada no capítulo anterior, o próximo passo foi avançarmos em uma compreensão mais qualitativa a respeito do que se tratavam esses termos de busca.

Conforme já mencionado na seção metodológica, utilizamos aqui uma inteligência artificial (IA) generativa para fazer a correção ortográfica dos termos e procedermos com uma classificação em categorias distintas. A categorização desses termos, muitas vezes bem específicos, como o nome de candidatos, deu-se ao utilizarmos também os resultados principais que o Google entregou a partir dos *snippets* desses resultados. A ideia aqui, portanto, é classificar os termos em três grandes categorias, a partir das quais se possa inferir a respeito do que se referiam, com base nos resultados.

Obviamente, há um limite nessa análise, pois estamos inferindo que os resultados, em conjunto com o termo da busca, podem nos dizer qual era o intuito do eleitor ao realizá-la. Cabe ressaltarmos que essa aproximação nem sempre está correta. Por exemplo, a partir dessa metodologia, se alguém buscar por “eleições 2020” com o intuito de saber o seu local de votação, mas encontrar apenas os resultados das votações, essa busca será classificada como o mesmo tipo daquela realizada por alguém que buscou por “eleições 2020” e tinha o intuito de encontrar tão somente a data da realização do primeiro turno. Ainda assim, analisar os termos de busca utilizados pode nos ajudar a entender o que os diferentes perfis buscaram ao longo do período eleitoral.

Figura 15 – Trigramas dos principais termos buscados pelos usuários classificados na categoria Informações sobre Candidatos/Partidos



Fonte: elaborado pelo autor.

As três categorias são:

- Informações sobre Candidatos/Partidos;
- Processos/Procedimentos Eleitorais;
- Buscas Gerais sobre Eleições.

6.1.1 *Informações sobre Candidatos/Partidos*

Esta categoria compreende termos de buscas com o possível intuito de obter informações sobre candidatos ou partidos políticos. Os termos poderiam ser relativos às propostas políticas de cada candidato ou partido, seu histórico de votação, biografias e outras informações pessoais ou profissionais. Exemplos típicos incluem a digitação direta do nome de candidatos, buscas por termos específicos que remetam à trajetória política dos indivíduos ou consultas sobre propostas de governo. Essa classificação evidencia um interesse direcionado do eleitor a compreender e avaliar as opções eleitorais com base em critérios específicos.

A nuvem de palavras abaixo ilustra os principais “trigramas” (conjunto de 3 palavras mais frequentes) nas buscas classificadas nessa categoria.

A figura 15 é uma representação conhecida como “nuvem de palavras”, que foi gerada a partir de trigramas para a categoria Informações sobre Candidatos/Partidos. Ela oferece uma visualização dos termos frequentemente pesquisados em relação às eleições municipais. Trigramas como “candidato a vereador”, “candidatos a prefeito” e “prefeitura de [nome da cidade]” indicam um interesse dos eleitores em informações específicas sobre candidaturas e contextos locais. Esses termos refletem a busca dos usuários por detalhes das campanhas e perfis dos candidatos.

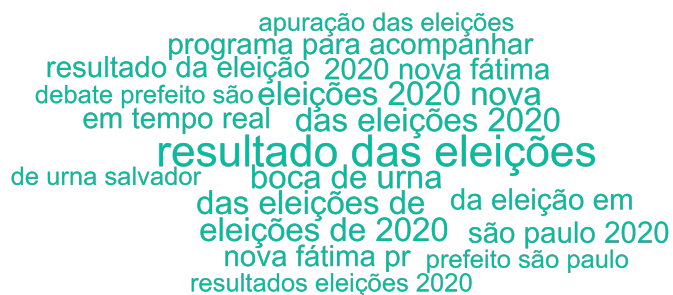
A presença de certos trigramas sugere foco em temas eleitorais e pode espelhar o interesse público durante períodos críticos do calendário eleitoral. A recorrência de nomes de lugares, como “São Paulo” ou “Porto Alegre”, aponta para um engajamento dos eleitores nessas áreas.

A nuvem de palavras está sendo utilizada aqui como uma ferramenta visualmente orientada e tem o propósito de servir como ilustração qualitativa. Seu uso é destinado a fornecer uma visão geral dos dados e não se propõe a uma análise léxica detalhada.

6.1.2 Processos/Procedimentos Eleitorais

Engloba termos de busca que se concentram nos aspectos procedimentais do processo eleitoral. Eleitores que realizam essas buscas procuram informações práticas sobre como exercer seu direito de voto, incluindo instruções sobre locais e datas de votação, requisitos de elegibilidade e informações sobre a obtenção de certidões ou agendamentos em tribunais eleitorais. Essa categoria reflete a busca por orientações claras e precisas que possibilitem ao eleitor cumprir suas obrigações cívicas de maneira informada e eficaz.

Figura 17 – Trigrama dos principais termos buscados pelos usuários classificados na categoria Buscas Gerais sobre Eleições



Fonte: elaborado pelo autor.

como “eleições 2020” ou “debates na televisão”. Essa classificação revela um interesse por parte do eleitorado em manter-se informado sobre o cenário eleitoral de modo mais holístico, incluindo o acompanhamento de eventos políticos de destaque e a compreensão das tendências e desdobramentos gerais das eleições.

Na nuvem de palavras da figura 17, referente às “Buscas Gerais sobre Eleições”, destacam-se termos que refletem o interesse dos usuários em resultados e acontecimentos em tempo real nas eleições de 2020. O foco em “resultado das eleições” e variações como “eleições 2020 nova” e “apuracão das eleições” evidencia uma busca ativa por informações atualizadas e por detalhes específicos dos eventos eleitorais. A menção a locais, como “nova fátima pr” e “prefeito são paulo”, aponta para o interesse regional nas eleições e sugere uma procura por informações sobre os resultados em cidades específicas.

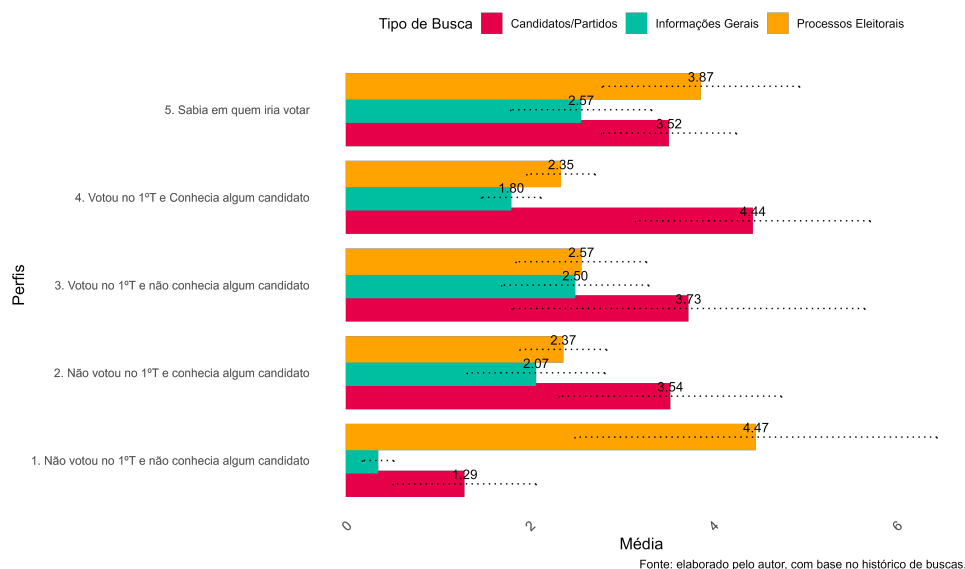
A recorrência de expressões relacionadas à pesquisa “boca de urna” e ao acompanhamento dos resultados (“programa para acompanhar”) indica uma busca por previsões e tendências eleitorais. Essas observações, extraídas dos trigramas, demonstram uma população engajada na monitorização do processo eleitoral e interessada nas implicações dos resultados para suas comunidades e para o cenário político mais amplo.

Tabela 19 – Classificação dos termos utilizados nas buscas políticas a partir do histórico de uso

	Freq.	Prop.
1-Informações sobre Candidatos/Partidos	1343	45,5%
2- Processos/Procedimentos Eleitorais	953	32,3%
3-Buscas gerais sobre eleições	656	22,2%
Total	2952	100%

Fonte: elaborado pelo autor.

A tabela 19 sintetiza o comportamento de busca dos usuários por temas eleitorais nas 2.952 consultas. A maioria das buscas, 45,5%, procurava “Informações sobre Candidatos/Partidos”, evidenciando um interesse específico na informação para decisão do voto. A categoria “Processos/Procedimentos Eleitorais” constituiu 32,3% das buscas. A categoria “Buscas Gerais sobre Eleições”, que abarcava questões mais amplas e resultados eleitorais, representou 22,2% do total. Esses dados quantitativos, além de ilustrarem a distribuição do interesse público, também ajudam a compreender a natureza qualitativa das informações que os eleitores buscavam no período eleitoral.

Figura 18 – Quantidade média de buscas de cada tipo de busca classificada por diferentes perfis

Fonte: elaborado pelo autor.

6.2 Quantidade média de buscas para cada perfil

Se no capítulo anterior constatamos que a probabilidade de se fazer uma busca política-eleitoral é maior para os perfis com preferência política mais consolidada, a diferença entre os tipos de buscas pode nos trazer uma visão mais qualitativa dos diferentes objetivos entre aqueles com visão política mais consolidada.

Uma maneira de fazermos essa análise é observarmos, dentre aqueles que fizeram alguma busca política, qual foi a quantidade média de buscas para cada tipo de termo analisado, conforme o gráfico que consta na figura 18.

A análise das médias e das diferenças significativas entre os grupos de votantes revela alguns padrões. O grupo que “Não votou no primeiro turno e não conhecia algum candidato” mostrou as menores médias em todas as categorias, exceto na de “Processos Eleitorais”, indicando um engajamento mais baixo tanto em relação aos candidatos e às informações gerais, mas preocupado, muito provavelmente, nos processos eleitorais de justificativa de seu voto.

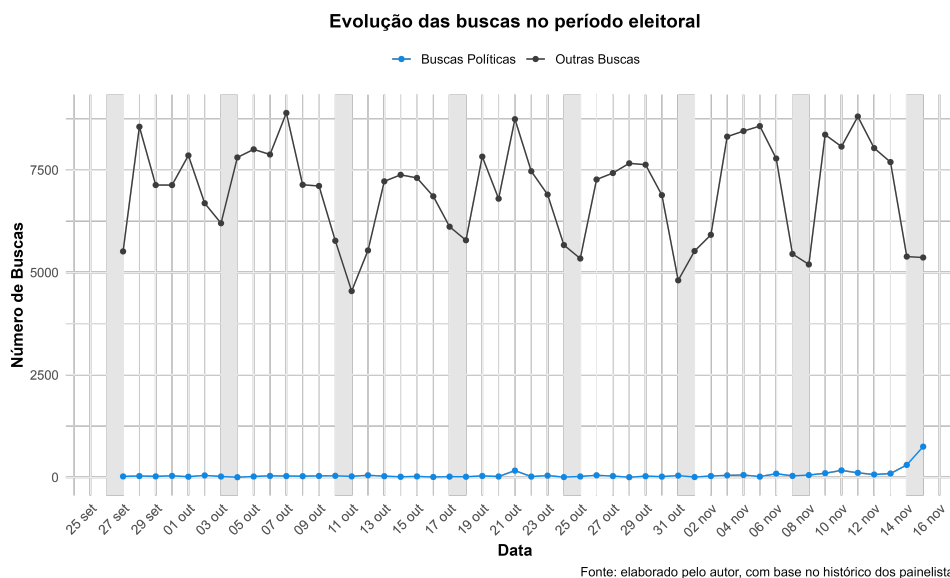
O grupo que “Votou no primeiro turno e conhecia algum candidato” não apenas apresentou as maiores médias na pesquisa sobre candidatos e processos eleitorais, como, se analisado em conjunto com o perfil “Votou e não conhecia algum candidato”, dá indícios a respeito de que maneira o Google pode ser utilizado para a tomada de decisão sobre em quem votar.

Como já esperado, o grupo “Sabia em quem iria votar” se sobressaiu nas buscas por informações gerais e sobre processos eleitorais, com a maior média nesse último quesito, sugerindo que, mesmo com uma decisão de voto formada, o interesse por um entendimento mais amplo do contexto eleitoral e dos seus mecanismos é marcante. As diferenças significativas identificadas, especialmente entre aqueles com conhecimento prévio sobre candidatos ou decisões de voto formadas e os menos engajados, sublinham o impacto do conhecimento e da decisão eleitoral no comportamento de busca por informações.

Em resumo, esses dados nos permitem concluir, novamente, que o interesse por informações sobre o processo eleitoral é maior entre os eleitores que já têm um conhecimento prévio ou que estão firmes em suas decisões de voto. Entretanto, evidencia que a busca por candidatos/partidos foi maior entre os usuários que já conheciam algum dos candidatos mas não sabiam em quem votar.

Esse dado qualifica de maneira importante as conclusões do capítulo anterior. Primeiro por mostrar que os eleitores que já sabiam em quem votar, aparentemente, fizeram mais buscas, mas buscaram sobre as eleições de modo amplo e generalista. Ou seja, é muito provável que estivessem interessados nas eleições enquanto acontecimento político.

Dos eleitores que votaram, aqueles que já conheciam algum candidato foram os perfis que mais buscaram sobre os candidatos, uma média quase 100% maior do que aqueles que não conheciam nenhum candidato, mas ainda assim votaram. Esse dado é um indício de como o Google pode ser especialmente útil para o público que possui algum interesse em política mas ainda não decidiu o seu voto.

Figura 19 – Evolução das buscas no período eleitoral

Fonte: elaborado pelo autor.

6.3 Quando o Google importa

Após investigarmos *quem* recorre ao Google para informações políticas e *como* essas buscas são realizadas, é crucial compreendermos *quando* esse fenômeno adquire maior relevância. A temporalidade das buscas políticas no Google não apenas reflete, mas também influencia a dinâmica eleitoral, marcando momentos em que a informação se torna uma ferramenta poderosa na formação da opinião pública e na decisão eleitoral.

Este segmento da pesquisa dedica-se a analisar o fluxo de buscas políticas ao longo do período eleitoral, identificando picos de interesse e correlacionando-os aos eventos-chave da agenda política. Por meio de uma metodologia quantitativa, analisaremos os dados de busca diária para desvendarmos quando os eleitores buscam ativamente por informações políticas no Google e quais fatores impulsionam essa procura.

Com a análise do fluxo de buscas por dia, pretendemos contribuir para uma compreensão mais aprofundada de como o Google, enquanto mediador de informação, desempenha um papel crucial em períodos eleitorais, funcionando como um termômetro da atenção política dos eleitores ao longo do tempo.

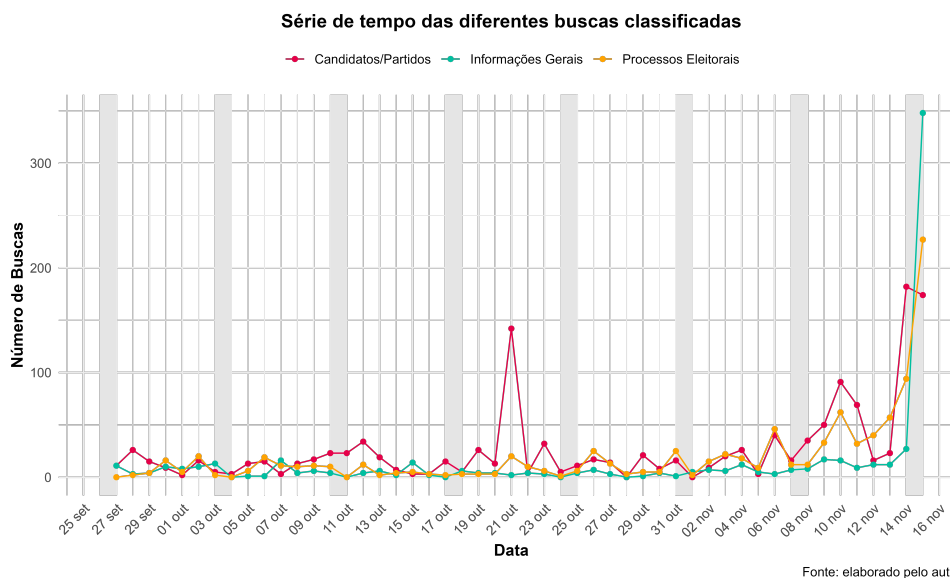
6.3.1 Quando os usuários buscam

A análise do gráfico que traça a evolução das buscas ao longo do período eleitoral (27 de setembro a 15 de novembro de 2020) revela padrões comportamentais interessantes.

Observa-se uma redução perceptível nas buscas durante os finais de semana, conforme ilustrado pelas áreas cinzentas. Esse fenômeno sugere que, nos dias tradicionalmente reservados ao descanso e ao lazer, há uma menor atividade no buscador.

Destaca-se a tendência de um aumento nas buscas não políticas às quartas-feiras, indicando um pico semanal de engajamento dos usuários com o buscador. Um ponto de inflexão notável nas buscas políticas em azul é o pico em 21 de outubro, e o seu maior volume de buscas políticas ocorreu durante o fim de semana das eleições. Apenas o último final de semana correspondeu a 13,38% de todas as buscas realizadas no período.

Esses dados não apenas delineiam o comportamento informacional dos usuários na plataforma de busca, mas também reforçam a premissa de que o acesso à informação é dinâmico e se molda em resposta ao contexto político vigente, realçando o papel do Google como um agente fundamental no processo eleitoral.

Figura 20 – Série de tempo das diferentes buscas classificadas a partir do histórico de uso

Fonte: elaborado pelo autor.

6.3.2 Diferentes tipos de buscas em diferentes momentos

O gráfico analisado apresenta a evolução das buscas relacionadas ao período eleitoral, distinguindo as três diferentes categorias classificadas e apresentadas anteriormente: informações gerais, candidatos/partidos e processos eleitorais. A linha referente às informações gerais, que é identificada pela cor verde-clara, demonstra um padrão de menor volume de buscas ao longo do período observado, com uma exceção notável no último dia do intervalo. Esse pico representa uma busca intensa por resultados eleitorais, sublinhando a natureza desses termos como veículos para a obtenção de notícias relacionadas às eleições.

Por outro lado, a linha vermelha, que mapeia as buscas por candidatos e partidos, mostra variações mais significativas, com vários picos ao longo do tempo. O ápice dessa categoria ocorre na véspera da eleição, em 14 de outubro, indicando um interesse agudo por informações sobre os candidatos, seguido por um segundo pico no dia da eleição. Isso sugere um comportamento de busca por informações de última hora por parte dos eleitores, possivelmente para tomada de decisão final ou esclarecimento de dúvidas. O pico do dia 21 de outubro se dá pela busca de um mesmo usuário com a temática de “candidatos ridículos 2020”.

A linha amarela, que ilustra as buscas sobre procedimentos eleitorais, atinge seu pico no dia 15 de novembro, imediatamente após a eleição. Esse aumento sugere que eleitores estavam em busca de informações sobre como justificar o voto e localizar referências do Tribunal Superior Eleitoral (TSE), provavelmente em um esforço para encontrar os resultados eleitorais ou

para compreender melhor o processo de justificativa. No dia anterior, 14 de novembro, véspera da eleição, houve também uma busca significativa por informações sobre o horário de votação e justificativa de voto, o que indica uma procura por orientações práticas para a participação no pleito.

6.4 Considerações finais

Os termos de busca podem representar preferências de informação dos usuários. Ao selecionar palavras-chave específicas, o eleitor não apenas busca dados alinhados aos seus interesses, mas também introduz essas preferências nos sistemas de busca, influenciando os resultados retornados (Baeza-Yates, 2018).

Esse fenômeno destaca a presença de viés na própria ação de buscar, conforme discutido por White e Horvitz White (2013), dentre outros, que apontam como as crenças prévias e diferenças demográficas moldam os comportamentos de busca on-line, tendendo à seleção de conteúdos que confirmam as atitudes políticas dos usuários (Knobloch-Westernwick *et al.*, 2015).

O papel do Google e das redes sociais na busca de informações políticas on-line (Kaye e Johnson, 2002) é particularmente relevante, pois as consultas podem refletir as preferências políticas dos usuários, funcionando como um indicativo viável do interesse público, podendo ser utilizadas para analisarmos comportamentos e tendências (Whyte, 2016; Borra e Weber, 2012). Portanto, esses aspectos fundamentam a necessidade de investigarmos como os algoritmos podem influenciar a visibilidade de informações e candidatos políticos, conforme já visto no Capítulo 2, em que discutimos os estudos sobre a representação de candidatos políticos nos resultados de busca (Trielli e Diakopoulos, 2019; Muddiman, 2013).

Este capítulo, entretanto, não se aprofundou da mesma maneira que boa parte dos estudos mencionados aqui, que fizeram uma análise dos termos de buscas em relação a um evento único: eleições presidenciais americanas. Devido à multiplicidade de eventos, ou de diversas eleições em nível sub-nacional, esta análise sobre o viés do que foi buscado seria algo complexo e, principalmente, fora do objetivo desta tese.

Google como fonte de informação sobre candidaturas

Entretanto, há contribuições importantes para o campo. A primeira delas é a compreensão de maneira ampla sobre aquilo que o eleitorado busca no Google. As três categorias utilizadas para a classificação dos termos explicitam que o Google importa, sim, para a busca de informação sobre candidaturas e partidos, o que na amostra coletada analisada é a principal

razão de uso. Em um segundo momento, informações sobre procedimentos eleitorais, como justificativa de voto e informação a respeito de onde votar, são relevantes. Além disso, ao fim das eleições, foi a fonte utilizada para se ter informação sobre as apurações.

A informação na última hora

Uma contribuição também importante, e que não é enfatizada nos estudos sobre o Google, é: quando ele é utilizado. Os dados desta pesquisa mostram que a última semana concentra quase um terço de todas as buscas realizadas, mas não da mesma maneira para todos os eleitores.

7 O QUE O GOOGLE ENTREGOU?

No tradicional relatório do Google sobre quais foram as buscas mais frequentes no ano, em 2020, o termo “Eleições 2020” esteve em 3º lugar (Google, 2020). O termo só não esteve à frente de dois temas amplamente correlacionados àquele ano de início da pandemia mundial de covid-19: coronavírus e auxílio emergencial.

É interessante notar que o Google sempre divulga a lista de termos buscados e que isso é bem repercutido, com diversas notícias, “veja o que as pessoas mais buscaram este ano” ou “saiba as dúvidas mais frequentes dos brasileiros na internet” são títulos comuns na mídia. Entretanto, são escassas as críticas sobre qual, de fato, é a base de dados desses relatórios e quais são seus números reais. Perguntas óbvias não são feitas, como: o termo coronavírus é o agrupamento de termos próximos como “corona”? Qual é a quantidade exata de vezes que o termo foi utilizado e por quantos usuários diferentes? O que as pessoas consumiram de informação depois de terem feito a busca? Como analisado no capítulo anterior, o volume de buscas com termos que se relacionam a candidatos, possivelmente, pode ter sido maior que simplesmente “eleições 2020”. Ainda que haja a divulgação de relatórios públicos, seguem obscurecidos os detalhes a respeito de como o Google, maior mediador informacional da história da humanidade, contabiliza esses dados.

Saber a quantidade exata de buscas feitas e de pessoas que as fizeram é um árduo exercício. Como abordado no Capítulo 3, sobre como o Google funciona, o setor de marketing digital recorre frequentemente a uma ampla variedade de sites “não endossados pelo Google”, os quais oferecem informações para profissionais de todo o mundo priorizarem qual conteúdo irão publicar. A prática é comum no universo de *search engine optimize* (SEO) e, principalmente, para profissionais que irão desenvolver campanhas pagas utilizando os resultados do próprio buscador⁸.

Apesar de a prática científica desencorajar o uso de dados provenientes de “caixas-pretas”, o mercado de marketing digital tem se apoiado nesses dados há anos para aprimorar a eficácia de suas estratégias no Google. De acordo com uma dessas fontes de mercado, o termo “eleições 2020” registrou 673 mil buscas em setembro, 1,2 milhão em outubro e alcançou um

⁸ Há indícios de que antivírus gratuitos podem fornecer dados de navegação dos usuários. Essas amostras são, supostamente, utilizadas para aprimorar os serviços oferecidos e, em alguns casos, compartilhadas com terceiros para fins diversos. A fonte específica dessas aplicações on-line utilizadas no mercado de SEO, no entanto, muitas vezes, permanece não especificada, levantando questões sobre privacidade e segurança digital. Referência: (<https://g1.globo.com/tecnologia/blog/seguranca-digital/post/antivirus-gratuitos-compartilham-dados-de-navegacao-do-usuario.html>).

pico de 9,3 milhões em novembro. Já o termo “coronavírus” atingiu seu ápice com 124 milhões de buscas em abril de 2020 (Mangools, 2024).

Esse quantitativo reforça a relevância do objetivo deste capítulo: saber o que milhões de brasileiros visualizaram quando buscaram informações sobre as eleições de 2020. Conforme já explicado na seção metodológica, esse esforço foi realizado por meio da coleta de **518.661** buscas. Para todas essas buscas, foram coletados os dez primeiros resultados orgânicos, o que alcançou um total de **5.186.610** resultados, que foram compostos de **354.525** links únicos.

Ao longo da seção metodológica, buscamos explicar como foi feito o desenho de coleta, com a preocupação de que fosse robusto o suficiente para que esse conjunto de buscas pudesse representar, da maneira mais próxima o possível, a realidade que milhões de eleitores brasileiros tiveram ao utilizar o Google como fonte de busca de informação. Esta seção aborda, portanto, quais foram os resultados das buscas realizadas ou, conforme a literatura especializada denomina, em inglês: Search Engine Results Page (SERP).

Como o foco do nosso estudo não é examinar o conteúdo específico de cada uma dessas páginas, procuramos entender a origem e a natureza dos links. Para isso, o primeiro passo foi identificar o domínio de cada URL (endereço da web).

7.1 A qual domínio a URL pertence

Um domínio na internet é como um endereço que aponta para um local específico na web, semelhante a um endereço que nos leva a uma casa ou estabelecimento específicos. Os domínios são fundamentais porque nos informam sobre a origem de uma página da internet, indicando se ela pertence a um portal de notícias, um site corporativo ou uma plataforma de mídia social, entre outros.

Para ilustrar, considere o exemplo do Universo Online (Uol), um dos maiores portais brasileiros de conteúdo. Esse portal possui um domínio principal, “uol.com.br”, e vários subdomínios associados a ele. Cada subdomínio representa uma seção ou categoria específica dentro do portal. Por exemplo, “economia.uol.com.br” é um subdomínio dedicado a notícias econômicas, enquanto “noticias.uol.com.br/politica/” foca em notícias políticas. Assim, ao identificarmos o domínio de um link, podemos compreender a que portal está associado.

A seguir, a tabela 20 apresenta alguns exemplos de links e seus respectivos domínios, demonstrando como diferentes URLs pertencem ao mesmo domínio principal:

Analisar o domínio é uma abordagem que nos permite entender melhor a distribuição

Tabela 20 – Exemplo de URLs e subdomínios vinculados ao Uol

URL (Página web)	Subdomínio	Domínio
https://noticias.uol.com.br/politica/	noticias.uol.com.br	uol.com.br
https://economia.uol.com.br/cotacoes/cambio/	economia.uol.com.br	uol.com.br
https://paranaportal.uol.com.br/eleicoes-2020/paulo-opuszka-candidato-curitiba-2020	paranaportal.uol.com.br	uol.com.br
https://www.uol.com.br/vivabem/noticias/bbc/2020/08/29/jogo-sujo-vacina-coronavirus.htm	uol.com.br	uol.com.br
http://m.jornaldotempo.uol.com.br/	jornaldotempo.uol.com.br	uol.com.br

Fonte: elaborado pelo autor.

e a representatividade de diferentes fontes de informação na internet, oferecendo uma visão mais agregada e organizada dos resultados das buscas. Nas bases de dados coletadas em todas as pesquisas, foi realizado um tratamento nos endereços de URL para que estivessem atrelados a um domínio. A quantidade de diferentes domínios encontrada, portanto, foi de **15.763**. Sendo que apenas para as buscas políticas (realizadas com termos políticos e eleitorais) foi um total de **5.969** domínios diferentes.

Neste trabalho, utilizaremos os termos “link”, “URL”, “página na web” ou apenas “página” para nos referirmos ao conteúdo mais específico coletado: o resultado para onde o Google apontou. Também utilizaremos o termo “site” e “domínio”.

A primeira pergunta relevante para entendermos os dados é: quais foram os sites que mais apareceram nos resultados de busca durante o período eleitoral de 2020?. Nesta seção mostraremos quais foram esses sites, fazendo a separação entre as buscas em que os termos eram político-eleitorais (buscas políticas) e, a título comparativo, as buscas não políticas, com uma composição de diversos tipos de termos de busca.

7.1.1 Resultados das buscas com termos políticos

O primeiro achado relevante é que a estatística da presença desses sites mostra uma concentração extremamente alta. A tabela 21 é um descritivo simples da quantidade de vezes que um domínio apareceu entre os 10 primeiros resultados exibidos pelo Google em todas as buscas orgânicas **com temas políticos** deste trabalho. A coluna Freq. é a contagem de vezes em que aquele domínio esteve entre os 10 resultados exibidos. A posição média é a média das diferentes posições em que os links de cada domínio tiveram em cada busca. Um domínio com uma média próxima de 1 mostra que ele esteve em média como o primeiro posicionado. A coluna Prop. é a proporção relativa da frequência e Prop. Acum. é o somatório dessas proporções.

A tabela 21 nos ajuda a entender o cenário dos resultados orgânicos para as buscas políticas. Sites como “diariocidade.com”, “omunicipio.com.br” e “descubraonline.com”, apesar

Tabela 21 – Os 10 domínios mais frequentes nos resultados de buscas políticas

Ranking	Domínio	Freq.	Posição Média	Prop.	Prop Acum.
1	gazetadopovo.com.br	402.460	6,3	10,8%	10,8%
2	jus.br	344.017	4,9	9,2%	19,9%
3	globo.com	295.759	4,6	7,9%	27,8%
4	diariocidade.com	245.442	2,8	6,6%	34,4%
5	estadao.com.br	224.606	4,7	6,0%	40,4%
6	otempo.com.br	201.923	3,8	5,4%	45,8%
7	uol.com.br	184.089	5,9	4,9%	50,7%
8	descubraonline.com	181.395	5,5	4,8%	55,6%
9	omunicipio.com.br	98.590	7,2	2,6%	58,2%
10	ebc.com.br	68.035	7,8	1,8%	60,0%
	Todos os demais domínios	1.564.919	7,6	41,8%	100,0%

Fonte: elaborado pelo autor.

de menos conhecidos, alcançam posições de destaque, ficando à frente de gigantes da mídia brasileira. Também é digno de nota o fato do site “gazetadopovo.com.br” assumir a liderança e o Portal Uol, um dos maiores e mais tradicionais portais on-line do Brasil, ocupar a sétima posição.

O site “jus.br”, que representa os sites da justiça brasileira, ocupa o segundo lugar, mas com uma posição média mais privilegiada do que sites como “diariocidade.com”, indicando que, quando aparecem, é mais provável que estejam no topo dos resultados. O portal “globo.com”, pertencente ao maior conglomerado de mídia do país, mantém-se firme em terceiro lugar.

A presença da “ebc.com.br” no décimo lugar destaca a empresa pública de comunicação do Brasil nesse ranking. O fato de que os dez primeiros colocados concentram 60,0% dos resultados é um indicativo claro de uma forte concentração. Esse resultado corrobora o trabalho de Vilarins L.;Stabile *et al.* (2021), feito com os resultados de pesquisa nas eleições presidenciais de 2018, e é um número pouco menor que o encontrado por (Nechushtai e Lewis, 2019) nas eleições americanas de 2016, apenas com o Google News, em que 5 portais correspondiam a 69% das recomendações.

7.1.2 Resultados das buscas com termos não-políticos

A análise das buscas não políticas revela nuances interessantes quando comparada à tabela anterior. Embora os grandes portais de notícias ainda dominem o cenário, há uma presença mais marcante das mídias sociais. A proporção acumulada dos dez primeiros sites nessas buscas

é de 24,9%, que é inferior à observada nas buscas políticas, sugerindo uma menor concentração. Esse fenômeno era esperado, dada a variedade de temas abordados nesta parte da pesquisa, que vai além das questões políticas, permitindo uma maior diversidade de fontes nos resultados de busca.

Tabela 22 – Os 10 domínios mais frequentes nos resultados de buscas não políticas

Ranking	Domínio	Freq.	Posição Média	Prop.	Prop Acum.
1	uol.com.br	88.652	5,60	6,10%	6,1%
2	globo.com	72.956	4,60	5,10%	11,2%
3	wikipedia.org	53.956	5,00	3,70%	14,9%
4	pinterest.com	27.123	5,40	1,90%	16,8%
5	abril.com.br	25.825	5,30	1,80%	18,6%
6	facebook.com	24.868	5,20	1,70%	20,3%
7	youtube.com	18.531	4,90	1,30%	21,6%
8	terra.com.br	16.499	5,60	1,10%	22,8%
9	google.com	16.214	5,80	1,10%	23,9%
10	mercadolivre.com.br	15.073	4,80	1,00%	24,9%
	Todos os demais domínios	1.098.786	7,10	76,10%	100,0%

Fonte: elaborado pelo autor.

Primeiramente, ao analisarmos os resultados de buscas políticas (tabela 21) e não políticas (tabela 22), observamos diferenças na diversificação de domínios. As buscas **não políticas** apresentam uma maior variedade de domínios, como evidenciado pela proporção de 76,10% para “todos os demais domínios”, em contraste com 41,8% das buscas **políticas**. Essa diferença sugere uma concentração mais significativa de fontes de informação nas buscas políticas.

Em segundo lugar, a constante presença de grandes portais de notícias, como “globo.com” e “uol.com.br”, em ambas as categorias de busca, demarca sua grande influência, independentemente do tipo de conteúdo pesquisado. No entanto, a liderança desses portais varia entre as categorias. Enquanto “uol.com.br” aparece no topo das buscas não políticas, domínios como “gazetadopovo.com.br” e “jus.br” dominam nas buscas políticas. Os sites de mídias sociais também aparecem em maior medida nas buscas não políticas.

Para compreendermos mais profundamente as razões por trás da relevância desses domínios nas buscas políticas, realizaremos uma análise mais detalhada nas seções posteriores. Essa análise focará em identificar para quais termos buscados esses domínios se destacaram. Dessa maneira, será possível discernirmos se a prevalência de certos domínios está ligada a

temas políticos específicos ou se é resultado de estratégias de SEO mais eficientes, direcionadas a tópicos políticos. Essa investigação mais detalhada é fundamental para entendermos não apenas o quê, mas também o porquê da visibilidade desses domínios em buscas políticas.

Este panorama inicial nos ajuda a entender a distribuição dos domínios e o seu nível de concentração. Por si só, esta análise mostra como é relevante a concentração dos resultados do Google em um período tão sensível quanto o eleitoral. Entretanto, não nos ajuda a entender de modo qualitativo o que de fato o Google entregou nos resultados de pesquisa. Para avançarmos nessa pergunta, foi necessário desenvolvermos um processo de categorização temática dos domínios.

7.2 Resultados diferentes para termos diferentes

Nesta seção mostraremos como diferentes consultas podem levar a resultados de pesquisa distintos, enfatizando a necessidade de entendermos a dinâmica dos mecanismos de busca. Ao focar nessa diversidade, pretendemos mostrar que a seleção de termos pode alterar significativamente as informações que são apresentadas aos usuários, ressaltando a relevância de uma abordagem cuidadosa na escolha de palavras-chave para pesquisa.

Conforme já observado na seção de metodologia, a escolha das palavras de busca do que chamamos de “buscas gerais”, entre as buscas com termos políticos, refere-se à busca sobre “Candidatos + Nome da Cidade”. A razão dessa escolha se deve aos resultados de um *survey* realizado no início de 2020, em que perguntou-se como o eleitor buscaria informações sobre os candidatos de sua cidade.

Tabela 23 – Termos de busca utilizados nas Buscas Gerais

Termo	Qtd. de Municípios	Qtd. de Buscas
candidatos a prefeito em (CIDADE)	5.295	93.111
candidatos em (CIDADE)	5.295	90.527
Eleições 2020	5.295	90.425
candidatas a prefeita em (CIDADE)	286	20.494
candidatas em (CIDADE)	286	20.481

Fonte: elaborado pelo autor.

Além disso, escolhemos também um termo genérico referente a eleições: “Eleições

2020”, que, segundo o próprio Google, foi um dos três termos do ano e, conforme observado no capítulo anterior, foi um dos termos mais buscados pelo painel de respondentes que tiveram seus históricos de busca analisados.

Essas buscas gerais contribuíram de modo relevante para o total das buscas políticas, pois, faltando 15 dias para o pleito do primeiro turno, as buscas foram realizadas para a totalidade de municípios brasileiros, conforme já mencionado na seção metodológica.

7.2.1 Resultados para o termo “Eleições 2020”

Tabela 24 – Os 10 domínios mais frequentes nos resultados do termo Eleições 2020

Ranking	Domínio	Freq.	Posição Média	Prop.	Prop Acum.
1	jus.br	328.735	4,90	36,4%	36,4%
2	globo.com	105.970	3,20	11,7%	48,1%
3	uol.com.br	88.133	5,70	9,7%	57,8%
4	estadao.com.br	72.050	5,00	8,0%	65,8%
5	ebc.com.br	64.689	7,80	7,2%	72,9%
6	gazetadopovo.com.br	62.520	8,70	6,9%	79,9%
7	r7.com	49.866	7,80	5,5%	85,4%
8	elpais.com	27.681	3,00	3,1%	88,4%
9	nsctotal.com.br	22.963	6,50	2,5%	91,0%
10	otempo.com.br	15.868	6,50	1,8%	92,7%
	Todos os demais domínios	81.643	7,70	9,0%	100,0%

Fonte: elaborado pelo autor.

A tabela 24 mostra quais domínios mais apareceram nos resultados de busca para o termo “Eleições 2020”. Sites oficiais da justiça brasileira estiveram presentes em mais de um terço de todas as urls entregues. Na sequência, aparecem sites do Grupo Globo, com 11,7%, e Grupo Uol, com 9,7%. Esse cenário já é, qualitativamente, bem diferente do que foi observado na tabela 21, pois com esse termo de busca há uma prevalência maior de sites mais conhecidos, sendo Globo e Uol os principais, o que é coerente com o tamanho e importância deles na audiência on-line brasileira (Journalism, 2022).

Entender as principais URLs que contribuíram para que esses sites estivessem tão bem posicionados pode elucidar algumas questões a respeito desse bom posicionamento. A tabela 25 mostra os três links mais entregues para esses 10 sites. Afinal, do que se tratavam esses resultados?

Tabela 25 – As 03 principais URLs para os sites mais frequentes nos resultados da busca com o termo Eleições 2020

Domínio	Freq.	Link
ebc.com.br	29.749	https://agenciabrasil.ebc.com.br/eleicoes-2020
ebc.com.br	10.282	https://agenciabrasil.ebc.com.br/tags/eleicoes-2020
ebc.com.br	3.293	https://agenciabrasil.ebc.com.br/eleicoes-2020/noticia/2020-11/eleicoes-2020-ultimo-dia-para-pedir-segunda-do-titulo-de-eleitor?amp
elpais.com	5.542	https://brasil.elpais.com/internacional/2020-11-07/resultado-das-eleicoes-eua-2020-joe-biden-eleito-presidente-kamala-harris-eleita-vice.html
elpais.com	2.789	https://brasil.elpais.com/internacional/2020-11-08/eleicoes-eua-2020-ao-vivo-repercussos-da-vitoria-de-biden.html?outputType=amp
elpais.com	2.780	https://brasil.elpais.com/internacional/2020-11-03/eleicoes-eua-2020-trump-x-biden-ao-vivo.html?outputType=amp
estadao.com.br	64.599	https://politica.estadao.com.br/eleicoes/2020
estadao.com.br	2.742	https://politica.estadao.com.br/noticias/eleicoes,como-votar-saiba-tudo-sobre-o-dia-da-votacao-nas-eleicoes-2020,70003510167.amp
estadao.com.br	1.603	https://politica.estadao.com.br/noticias/eleicoes,como-votar-saiba-tudo-sobre-o-dia-da-votacao-nas-eleicoes-2020,70003510167
gazetadopovo.com.br	45.966	https://www.gazetadopovo.com.br/eleicoes/2020/
gazetadopovo.com.br	8.797	https://www.gazetadopovo.com.br/tudo-sobre/eleicoes-2020/
gazetadopovo.com.br	1.449	https://www.gazetadopovo.com.br/eleicoes/2020/pesquisa-tendencias-sp-rio-recife-bh-fortaleza-prefeito-eleicoes-2020/amp/
globo.com	74.565	https://g1.globo.com/politica/eleicoes/2020/
globo.com	5.347	https://g1.globo.com/google/amp/politica/eleicoes/2020/noticia/2020/09/12/eleicoes-2020-conheca-regras-e-saiba-o-que-candidato-e-eleitor-podem-e-nao-podem-fazer.ghtml
globo.com	3.138	https://g1.globo.com/politica/eleicoes/2020/noticia/2020/09/12/eleicoes-2020-conheca-regras-e-saiba-o-que-candidato-e-eleitor-podem-e-nao-podem-fazer.ghtml
jus.br	72.745	https://www.tse.jus.br/eleicoes/eleicoes-2020/eleicoes-2020
jus.br	62.356	https://www.tre-mg.jus.br/eleicoes/eleicoes-2020/eleicoes-2020
jus.br	25.483	http://www.tse.jus.br/eleicoes/eleicoes-2020/eleicoes-2020
nsctotal.com.br	9.884	https://www.nsctotal.com.br/eleicoes/2020
nsctotal.com.br	2.741	https://www.nsctotal.com.br/noticias/eleicoes-2020-o-que-saber-tres-dias-antes-da-votacao?amp=1
nsctotal.com.br	1.688	https://www.nsctotal.com.br/noticias/eleicoes-2020-data-regras-prazos-calendario-eleitoral?amp=1
otempo.com.br	9.225	https://www.otempo.com.br/hotsites/eleicoes-2020
otempo.com.br	2.849	https://www.otempo.com.br/mobile/hotsites/eleicoes-2020/eleicoes-2020-nao-sabe-em-quem-votar-confira-a-ficha-de-todos-os-candidatos-1.2412574?amp
otempo.com.br	1.600	https://www.otempo.com.br/hotsites/eleicoes-2020/eleicoes-2020-nao-sabe-em-quem-votar-confira-a-ficha-de-todos-os-candidatos-1.2412574
r7.com	49.860	https://noticias.r7.com/eleicoes-2020
r7.com	5	https://www.google.com/amp/s/noticias.r7.com/brasil/servidores-precisam-se-afastar-ate-hoje-para-concorrer-as-eleicoes-2020-15082020%3Famp
r7.com	1	https://noticias.r7.com/brasil/servidores-precisam-se-afastar-ate-hoje-para-concorrer-as-eleicoes-2020-15082020?amp
uol.com.br	48.964	https://www1.folha.uol.com.br/poder/eleicoes/2020
uol.com.br	11.672	https://www1.folha.uol.com.br/poder/eleicoes/2020/
uol.com.br	3.002	https://noticias.uol.com.br/internacional/ultimas-noticias/2020/11/06/recontagem-de-votos-na-georgia-comecara-em-1-de-dezembro.amp.htm

Fonte: elaborado pelo autor.

As URLs entregues pelo Google para o termo “Eleições 2020” durante o período eleitoral de 2020 podem ser analisadas em duas vertentes:

A. Páginas de conteúdo generalista:

– **EBC (ebc.com.br)**: destaca-se com várias páginas altamente indexadas, oferecendo uma

cobertura ampla das eleições 2020, desde informações gerais até notícias específicas, como o último dia para pedir a segunda via do título de eleitor;

- **Globo.com, Uol.com.br e Estadao.com.br:** apresentam links com altas frequências, direcionando para seções dedicadas às eleições de 2020, sugerindo uma cobertura extensiva e generalista do evento;
- **Jus.br:** inclui links diretos para o Tribunal Superior Eleitoral (TSE) e o Tribunal Regional Eleitoral de Minas Gerais (TRE-MG), ambos fornecendo informações oficiais e abrangentes sobre as eleições, indicando um conteúdo mais institucional e formal.

B. Domínios com conteúdo mais específico:

- **ElPais.com:** foca nas eleições dos EUA de 2020, com artigos sobre os resultados e repercussões da vitória de Biden, destacando-se pela cobertura de um evento eleitoral específico fora do Brasil. Isso mostra como o interesse pelo termo “eleições 2020” não se limitou ao contexto nacional;
- **GazetadoPovo.com.br e NSCTotal.com.br:** também oferecem conteúdos específicos, como pesquisas de tendências em cidades brasileiras e informações detalhadas sobre o calendário eleitoral e regras, respectivamente, indicando um esforço para fornecer análises mais profundas e orientações práticas aos eleitores;
- **OTempo.com.br e R7.com:** apresentam conteúdos que vão desde fichas completas dos candidatos até notícias específicas sobre prazos legais para candidaturas, refletindo uma abordagem que busca atender às necessidades informativas específicas dos eleitores.

As temáticas abordam desde aspectos gerais do processo eleitoral até detalhes específicos sobre candidaturas e legislação eleitoral. Domínios como EBC, Globo.com e Estadao.com.br serviram como fontes generalistas, oferecendo uma ampla gama de informações sobre as eleições. Em contraste, sites como ElPais.com e GazetadoPovo.com.br se concentraram em aspectos mais específicos ou em eventos eleitorais em outros países, demonstrando a variedade de interesses e necessidades informativas do público durante o período eleitoral.

7.2.2 Resultados para os termos “candidatos + cidade”

A busca com termos envolvendo candidatos tem uma particularidade importante: a inclusão do nome da cidade em todas as consultas. Concomitantemente a esse fato, conforme já explicado na seção metodológica, todas as buscas foram geolocalizadas para simularem os resultados específicos para cada município, o que garantiu que, para cada busca, além do nome

da cidade, o buscador entendesse que a busca estava sendo feita naquela cidade.

Tabela 26 – Os 10 domínios mais frequentes nos resultados das buscas com termos sobre candidatos

Ranking	Domínio	Freq.	Posição Média	Prop.	Prop Acum.
1	gazetadopovo.com.br	315.095	5,90	14,00%	14,0%
2	diariocidade.com	245.296	2,80	10,90%	24,9%
3	descubraonline.com	181.323	5,50	8,10%	33,0%
4	otempo.com.br	179.738	3,50	8,00%	41,0%
5	globo.com	131.050	5,40	5,80%	46,9%
6	estadao.com.br	124.817	4,40	5,60%	52,4%
7	omunicipio.com.br	97.765	7,20	4,40%	56,8%
8	todapolitica.com	51.246	7,60	2,30%	59,0%
9	diariorio.com	51.077	6,70	2,30%	61,3%
10	uol.com.br	48.699	6,90	2,20%	63,5%
	Todos os demais domínios	868.723	7,70	38,70%	100,0%

Fonte: elaborado pelo autor.

É nesse contexto de buscas que alguns sites chamam atenção na tabela 26: “diariocidade.com”, “otempo.com.br”, “descubraonline.com” e “omunicipio.com.br”, que são sites não tão conhecidos do público brasileiro, mas que estiveram muito presentes nos resultados de busca.

Para tentarmos entender as razões da presença desses sites nessa lista, uma análise de quais páginas de fato contribuíram para essa presença pode ser esclarecedora. A tabela 27 mostra o descritivo dos principais links desses domínios que apareceram para essas buscas. Ao analisarmos essas páginas, é possível levantarmos algumas possíveis hipóteses para a sua relevância.

Tabela 27 – As 03 principais URLs para os sites mais frequentes nos resultados da busca com termos sobre candidatos

Domínio	Freq.	Link
descubraonline.com	173	https://www.descubraonline.com/eleicoes/2020/sp/itatiba/candidatos/amp/
descubraonline.com	143	https://www.descubraonline.com/eleicoes/2020/sp/cabreuva/candidatos/amp/
descubraonline.com	138	https://www.descubraonline.com/eleicoes/2020/rs/sapucaia-do-sul/candidatos/amp/
diariocidade.com	227	https://www.diariocidade.com/ba/serrinha/eleicoes/2020/candidatos/amp/
diariocidade.com	200	https://www.diariocidade.com/sp/franco-da-rocha/eleicoes/2020/candidatos/amp/
diariocidade.com	187	https://www.diariocidade.com/pr/colombo/eleicoes/2020/candidatos/amp/
diariorio.com	172	https://diariorio.com/os-pre-candidatos-a-prefeito-de-duque-de-caxias-para-as-eleicoes-2020/amp/

diariodorio.com	152	https://diariodorio.com/os-pre-candidatos-a-prefeito-de-nova-iguacu/amp/
diariodorio.com	132	https://diariodorio.com/os-pre-candidatos-a-prefeito-de-sao-joao-de-meriti-para-as-eleicoes-2020/amp/
estadao.com.br	138	https://politica.estadao.com.br/eleicoes/2020/candidatos/mg/santa-luzia/prefeito
estadao.com.br	128	https://politica.estadao.com.br/eleicoes/2020/candidatos/rj/bom-jesus-do-itabapoana
estadao.com.br	127	https://politica.estadao.com.br/eleicoes/2020/candidatos/sp/santana-de-parnaiba
gazetadopovo.com.br	315	https://especiais.gazetadopovo.com.br/eleicoes/2016/itapecerica-da-serra-sp/prefeito/
gazetadopovo.com.br	263	https://especiais.gazetadopovo.com.br/eleicoes/2016/passos-mg/prefeito/
gazetadopovo.com.br	262	https://especiais.gazetadopovo.com.br/eleicoes/2016/alvorada-rs/prefeito/
globo.com	285	https://g1.globo.com/google/amp/sp/vale-do-paraiba-regiao/eleicoes/2020/noticia/2020/08/24/conheca-os-pre-candidatos-a-prefeitura-de-sao-jose-dos-campos-ja-declarados.ghtml
globo.com	217	https://g1.globo.com/google/amp/mg/minas-gerais/eleicoes/2020/noticia/2020/09/29/candidatos-a-prefeito-de-betim-contagem-ibirite-e-nova-lima-nas-eleicoes-2020-veja-quem-sao.ghtml
globo.com	215	https://g1.globo.com/google/amp/ce/ceara/eleicoes/2020/noticia/2020/09/19/candidatos-a-prefeito-em-itapipoca-maranguape-iguatu-e-quixada-veja-quem-sao.ghtml
omunicipio.com.br	91	https://omunicipio.com.br/eleicoes-2020-saiba-quem-sao-os-seis-candidatos-que-disputarao-prefeitura-de-brusque/amp/
omunicipio.com.br	75	https://omunicipio.com.br/saiba-quais-sao-os-possiveis-candidatos-prefeitura-de-brusque-em-2020/amp/
omunicipio.com.br	64	https://omunicipio.com.br/eleicoes-2020/maranhao/sao-domingos-do-maranhao/
otempo.com.br	135	https://www.otempo.com.br/hotsites/eleicoes-2020/candidatos/itapecerica-da-serra-sp/prefeito
otempo.com.br	131	https://www.otempo.com.br/mobile/politica/lideranca-de-murta-se-amplia-na-disputa-em-vespasiano-1.301622?amp
otempo.com.br	118	https://www.otempo.com.br/mobile/hotsites/eleicoes-2020/eleicoes-em-betim-cidade-tera-sete-candidatos-a-prefeito-neste-ano-1.2386253?amp
todapolitica.com	201	https://www.todapolitica.com/eleicoes-2016/candidatos-sao-domingos-do-azeitao-ma/
todapolitica.com	200	https://www.todapolitica.com/eleicoes-2016/candidatos-tomar-do-geru-se/
todapolitica.com	173	https://www.todapolitica.com/eleicoes-2016/candidatos-sao-joao-de-pirabas-pa/
uol.com.br	229	https://atarde.uol.com.br/politica/eleicoes/noticias/2125705-teixeira-de-freitas-nao-tem-uma-obra-de-relevancia-dele-diz-precandidato-sobre-prefeito
uol.com.br	199	http://eleicoes.uol.com.br/2012/candidatos/2012/prefeito/ba/10011948-fernao.htm
uol.com.br	182	https://atarde.uol.com.br/politica/eleicoes/noticias/2124198-olhar-futuro-tres-nomes-da-oposicao-ja-estao-confirmados-na-corrida-eleitoral-em-lauro-de-freitas

Fonte: elaborado pelo autor.

A análise da tabela 27 revela uma estratégia comum entre os sites listados: a otimização de seu conteúdo para pesquisas locais. A inclusão do nome da cidade e do termo “candidatos” nas URLs é uma prática recorrente, alinhando-se com a otimização para mecanismos de busca discutida no terceiro capítulo desta tese, no qual afirmamos que URLs claras e com conteúdo geograficamente localizado são favorecidas pelo Google.

O grupo Globo se destaca por sua ampla capilaridade e, de toda a lista, talvez seja a única organização com real capacidade de fazer uma cobertura jornalística que cubra boa parte dos municípios brasileiros. Os portais otempo.com.br e diariodorio.com aparecem com maior ênfase em uma cobertura para os municípios mineiros e fluminenses, respectivamente. Em contraste, sites como todapolitica.com.br, omunicipio.com.br, descubraonline.com e diariocidade.com não possuem essa cobertura jornalística. Em vez disso, parecem ter adotado uma abordagem de “simulação” de notícias locais, criando conteúdo automatizado específico para cada localidade. O diariocidade.com, por exemplo, adaptou dados do TSE em uma URL amigável, uma prática que, embora alinhe-se com critérios algorítmicos, desvia-se das boas práticas informacionais esperadas para a cobertura de um evento de tamanha importância.

Levanta-se, então, a questão sobre a adequação dessa estratégia, especialmente

quando consideramos a possibilidade de dar visibilidade a pequenos sites regionais, que oferecem diálogo direto com os candidatos locais. Em municípios com menor cobertura midiática, é provável que os sites regionais tenham sido preteridos em favor de conteúdo automatizado, considerado relevante pelo algoritmo.

Essa observação sublinha a importância da seleção criteriosa de informações políticas. Conforme discutido anteriormente, eleitores indecisos recorrem ao Google para obter informações. No entanto, a análise sugere que a qualidade da informação, embora concentrada em poucos domínios, pode ser facilmente manipulada para apresentar páginas locais aparentemente relevantes, mas que, na realidade, podem não oferecer o valor informativo esperado por quem busca informações autênticas sobre o cenário político local.

7.2.2.1 *Análise da concentração de busca*

A concentração dos domínios para os dez primeiros sites nessa categoria de buscas é de 63,5%. Um valor menor que para o termo anterior. Para explorarmos um pouco mais a variação desse termo, a tabela 28 traz uma análise comparada com as diferentes faixas de municípios, apenas para essas buscas.

Tabela 28 – Proporção acumulada nos dez primeiros sites para cada faixa de município

Faixa Populacional dos Municípios	Prop. Acumulada
01. Capitais brasileiras	46,5%
02. Municípios com segundo turno	42,6%
03. Municípios com mais de 110 mil hab.	46,8%
04. Mais de 45	63,6%
05. Municípios com mais de 20 mil hab.	73,5%
06. Municípios com até 20 mil hab.	82,3%

Fonte: elaborado pelo autor.

Fica claro que, para municípios menores, a concentração é maior. Para estudos futuros, esse é um tema que pode ser explorado ao se adicionarem outros fatores que podem ser melhor explorados em modelos multivariados.

7.2.3 Buscas nas capitais brasileiras

As buscas realizadas nas capitais brasileiras focaram-se, na pré-campanha, no nome dos possíveis candidatos que disputariam o pleito, já ao longo da campanha, em todos os candidatos registrados no TSE. A análise dos principais domínios obtidos mostra uma diferença relevante em relação às buscas gerais, observadas na seção anterior.

Tabela 29 – Os 10 domínios mais frequentes nos resultados das buscas feitas nas capitais brasileiras

Ranking	Domínio	Freq.	Posição Média	Prop.	Prop Acum.
1	globo.com	25.589	4,60	9,5%	9,5%
2	facebook.com	25.580	4,60	9,5%	19,0%
3	wikipedia.org	16.298	2,80	6,1%	25,1%
4	gazetadopovo.com.br	13.566	5,70	5,0%	30,2%
5	estadao.com.br	10.576	5,50	3,9%	34,1%
6	instagram.com	9.698	5,90	3,6%	37,7%
7	camara.leg.br	9.472	2,40	3,5%	41,2%
8	jusbrasil.com.br	7.526	6,60	2,8%	44,0%
9	uol.com.br	7.040	6,50	2,6%	46,6%
10	twitter.com	4.810	5,80	1,8%	48,4%
	Todos os demais domínios	143.395	7,50	53,4%	100,0%

Fonte: elaborado pelo autor.

A análise das buscas nas capitais brasileiras durante o período eleitoral destaca um foco nos nomes dos candidatos, com uma diversidade de domínios indexando conteúdo diretamente relacionado a esses indivíduos. Isso reflete uma estratégia de conteúdo voltada para os concorrentes a cargos públicos, uma abordagem esperada, dada a natureza das buscas.

Links do “camara.leg.br” direcionaram para perfis de deputados, enquanto o Estadão (“estadao.com.br”) incluiu links para candidatos de eleições anteriores (2018), sugerindo a relevância contínua de informações históricas sobre os candidatos. Plataformas sociais como Facebook, Instagram, Twitter e Wikipedia forneceram acessos diretos a perfis oficiais e informações detalhadas sobre os candidatos, evidenciando a ampla gama de fontes de informação utilizadas pelos eleitores.

Gazeta do Povo, Globo.com (G1) e JusBrasil destacam-se pela criação automatizada de conteúdo sobre os candidatos, indicando o uso de páginas dedicadas a compilar informações eleitorais. Essa estratégia sugere que candidatos com longa presença institucional tendem a ter uma visibilidade mais favorável no Google, beneficiando-se de publicações oficiais em plataformas como a Câmara dos Deputados.

A posição média da Wikipedia na lista de domínios ressalta sua importância como principal referência para o Google na busca por nomes de candidatos, sublinhando o papel central da plataforma na disseminação de informações políticas.

Tabela 30 – As 03 principais URLs para os sites mais frequentes nos resultados das buscas realizadas nas capitais brasileiras

Domínio	Freq.	Link
camara.leg.br	116	https://www.camara.leg.br/deputados/160635
camara.leg.br	113	https://www.camara.leg.br/deputados/204479
estadao.com.br	112	https://politica.estadao.com.br/eleicoes/2018/candidatos/ro/deputado-estadual/dr-breno-mendes,55123
estadao.com.br	87	https://politica.estadao.com.br/eleicoes/2018/candidatos/ap/governador/cirilo-fernandes,17
estadao.com.br	87	https://politica.estadao.com.br/eleicoes/2018/candidatos/pb/deputado-estadual/anisio-maia,13000
facebook.com	96	https://m.facebook.com/brenomendesadvogado/?locale2=pt_BR
facebook.com	95	https://m.facebook.com/jarbassoster.acre/?locale2=pt_BR
gazetadopovo.com.br	90	https://especiais.gazetadopovo.com.br/eleicoes/2016/boa-vista-rr/vereador/professor-linoberg-18123/
gazetadopovo.com.br	84	https://especiais.gazetadopovo.com.br/eleicoes/2016/natal-rn/prefeito/rosalia-fernandes-16/
gazetadopovo.com.br	84	https://www.gazetadopovo.com.br/tudo-sobre/fernando-francischini/
globo.com	111	https://g1.globo.com/politica/politico/marcelo-lelis/
globo.com	107	https://g1.globo.com/politica/politico/joaquim-rocha/
instagram.com	84	https://www.instagram.com/danielzendoacre/?hl=pt-br
instagram.com	84	https://www.instagram.com/fernandapsol/?hl=pt-br
instagram.com	84	https://www.instagram.com/fernandofreitascs/?hl=pt-br
instagram.com	84	https://www.instagram.com/manueladavila/?hl=pt-br
jusbrasil.com.br	103	https://www.jusbrasil.com.br/topicos/28823669/roberto-duarte
jusbrasil.com.br	99	https://www.jusbrasil.com.br/topicos/52462968/rosalia-fernandes
twitter.com	62	https://mobile.twitter.com/majordenice
twitter.com	61	https://mobile.twitter.com/hermanomorais
uol.com.br	58	https://terceirotempo.uol.com.br/que-fim-levou/mazinho-3336
uol.com.br	53	https://terceirotempo.uol.com.br/que-fim-levou/ze-ricardo
wikipedia.org	98	https://pt.m.wikipedia.org/wiki/Samuel_Costa
wikipedia.org	96	https://pt.m.wikipedia.org/wiki/Alfredo_Nascimento

Fonte: elaborado pelo autor.

7.2.4 Buscas temáticas

As buscas temáticas tinham como objetivo simular possíveis perguntas, dúvidas ou termos mais específicos, que não fossem genéricos como as buscas gerais ou tão específicos como nomes de candidatos e as buscas nas capitais brasileiras. Essas buscas temáticas estavam subdivididas em alguns temas específicos:

Tabela 32 – Os 10 domínios mais frequentes nos resultados das buscas temáticas

Ranking	Domínio	Freq.	Posição Média	Prop.	Prop Acum.
---------	---------	-------	---------------	-------	------------

1	uol.com.br	40.217	5,3	12,40%	12,4%
2	globo.com	33.150	6,2	10,20%	22,6%
3	estadao.com.br	17.163	4,7	5,30%	27,9%
4	gazetadopovo.com.br	11.279	3,2	3,50%	31,4%
5	jus.br	9.485	4,3	2,90%	34,3%
6	solides.com.br	7.934	4,4	2,40%	36,8%
7	em.com.br	7.302	5,8	2,30%	39,0%
8	politize.com.br	7.274	3,8	2,20%	41,3%
9	kenoby.com	7.005	6,1	2,20%	43,4%
10	abril.com.br	6.523	6,4	2,00%	45,5%
	Todos os demais domínios	183.271	7,4	56,60%	100,0%

Fonte: elaborado pelo autor.

Os domínios mais referenciados abordam diferentes aspectos das eleições. Por exemplo, “abril.com.br” e “politize.com.br” fornecem orientações sobre como escolher o melhor candidato, enfatizando a importância de se avaliar propostas para educação. “Estadao.com.br” destaca as propostas dos candidatos para a educação e também traz o perfil do então Presidente da República, Jair Bolsonaro.

Plataformas como “globo.com” discutem a veracidade de informações sobre fraude nas urnas e apresentam dados sobre candidatos “ficha-limpa”, enquanto “jus.br” oferece artigos acadêmicos e informações legais relacionadas às eleições. “Kenoby.com” e “solides.com.br” se destacam por abordarem temas de recrutamento e seleção, sugerindo uma analogia entre escolher candidatos a empregos e candidatos políticos.

Tabela 33 – As 03 principais URLs para os sites mais frequentes nos resultados das buscas temáticas

Domínio	Freq.	Link
abril.com.br	1353	https://super.abril.com.br/cultura/como-escolher-o-melhor-candidato/amp/
abril.com.br	976	https://super.abril.com.br/cultura/como-escolher-o-melhor-candidato/
abril.com.br	608	https://veja.abril.com.br/politica/o-dilema-dos-candidatos-do-pt-a-prefeito-usar-ou-nao-a-imagem-de-lula/amp/
em.com.br	1167	https://www.em.com.br/app/noticia/politica/2020/08/20/interna_politica,1177924/amp.html
em.com.br	712	https://www.em.com.br/app/noticia/politica/2020/08/20/interna_politica,1177924/lula-diz-que-e-plenamente-possivel-pt-nao-ter-candidato-em-2022.shtml
em.com.br	390	https://www.em.com.br/app/noticia/politica/2020/07/04/interna_politica,1162740/amp.html

estadao.com.br	2617	https://infograficos.estadao.com.br/focas/por-minha-conta/materia/as-propostas-dos-candidatos-para-a-educacao
estadao.com.br	2529	https://politica.estadao.com.br/eleicoes/2018/candidatos/br/presidente/jair-bolsonaro,17
estadao.com.br	1575	https://politica.estadao.com.br/noticias/eleicoes,bolsonaro-tira-do-psdb-dominio-do-voto-anti-pt,70002504226.amp
gazetadopovo.com.br	1421	https://especiais.gazetadopovo.com.br/eleicoes/2018/candidatos/presidente/jair-bolsonaro?amp=true
gazetadopovo.com.br	1065	https://www.gazetadopovo.com.br/republica/fraude-nas-urnas-2018-justica-cobra-provas-bolsonaro/amp/
gazetadopovo.com.br	684	https://www.gazetadopovo.com.br/republica/lula-insinua-candidatura-em-2022-de-onde-vem-a-confianca-do-petista/amp/
globo.com	2647	https://oglobo.globo.com/fato-ou-fake/mensagens-com-conteudo-fake-sobre-fraude-em-urnas-eletronicas-se-espalham-nas-redes-23134205
globo.com	1450	https://oglobo.globo.com/brasil/eleicoes-2010/portal-ficha-limpa-com-prestacoes-de-contas-de-candidatos-ja-esta-no-ar-4999632
globo.com	1180	https://valor.globo.com/google/amp/politica/noticia/2018/10/07/jornal-valor-empresarios-aderem-a-onda-anti-pt-e-veem-economia-melhor.ghtml
jus.br	1738	http://www.tse.jus.br/o-tse/escola-judiciaria-eleitoral/publicacoes/revistas-da-eje/artigos/revista-eletronica-eje-n.-4-ano-5/digressoes-sobre-as-doacoes-de-campanha-oriundas-de-pessoas-juridicas
jus.br	851	https://www.tse.jus.br/o-tse/escola-judiciaria-eleitoral/publicacoes/revistas-da-eje/artigos/revista-eletronica-eje-n.-4-ano-5/digressoes-sobre-as-doacoes-de-campanha-oriundas-de-pessoas-juridicas
jus.br	850	http://www.tse.jus.br/eleicoes/eleicoes-2018/denuncias-eleitorais
kenoby.com	1060	https://kenoby.com/blog/roteiro-de-entrevista/amp/
kenoby.com	654	https://kenoby.com/blog/roteiro-de-entrevista/
kenoby.com	546	https://kenoby.com/blog/como-avaliar-um-candidato/amp/
politize.com.br	1080	https://www.politize.com.br/propostas-dos-candidatos-a-presidencia-para-a-educacao/
politize.com.br	982	https://www.politize.com.br/propostas-dos-candidatos-a-presidencia-para-a-educacao/amp/
politize.com.br	850	https://www.politize.com.br/bom-candidato-o-que-e/
solides.com.br	4076	https://blog.solides.com.br/testes-para-selecionar-um-candidato/
solides.com.br	2323	https://blog.solides.com.br/melhores-ferramentas-de-recrutamento-e-selecao/
solides.com.br	1499	https://blog.solides.com.br/processo-seletivo/
uol.com.br	2333	https://congressoemfoco.uol.com.br/especial/noticias/exclusivo-conheca-os-candidatos-que-cairam-na-lei-da-ficha-limpa/
uol.com.br	1710	https://congressoemfoco.uol.com.br/especial/noticias/exclusivo-os-candidatos-barrados-pela-ficha-limpa-em-2014/
uol.com.br	1415	https://www1.folha.uol.com.br/amp/poder/2018/10/empresarios-bancam-campanha-contra-o-pt-pelo-whatsapp.shtml

Fonte: elaborado pelo autor.

Tabela 31 – Termos utilizados para simular buscas políticas temáticas

Termo	Qtd. de Municípios	Nr. de Buscas
Qual é o candidato com melhores propostas para a educação?	284	2.617
candidato do pt	286	2.616
candidato do bolsonaro	284	2.614
Qual é o melhor candidato para fazer ter mais empregos?	284	2.611
candidato antipt	286	2.611
Qual é o melhor candidato?	286	2.610
candidato ficha-limpa OR candidato fichalimpa	286	2.610
candidato denunciado	284	2.609
Qual é o melhor candidato para fazer ter mais segurança?	285	2.606
Qual é o melhor candidato para a saúde?	283	2.597
candidato do lula	284	2.596
fraude nas urnas	96	2.318
Como votar?	26	233
Quem está na frente nas eleições	26	233
candidato do psl	26	233
candidato do psdb	26	232
candidato do psol	26	232
Por que devemos votar	26	230

Fonte: elaborado pelo autor.

Tabela 34 – Concentração dos resultados nos dez principais sites

Tipo de Busca	Concentração nos 10 Domínios
Termo: Eleições 2020	92,7%
Termos “candidatos + cidade”	63,5%
Busca nas capitais	48,4%
Buscas temáticas	45,5%

Fonte: elaborado pelo autor.

7.3 Considerações finais

Quando se pesquisa com termos político-eleitorais, o Google concentra seus resultados em poucos sites. A estatística encontrada neste trabalho é de 60% para o conjunto de todas as buscas políticas, mas varia entre 45,5% e 92,7%, o que chama a atenção por ser um número destacadamente elevado. Esses dados podem ser observados na tabela 34, a seguir: **A concentração na mídia tradicional**

Esse resultado dialoga com trabalhos anteriores, como (Trielli e Diakopoulos, 2019; Nechushtai e Lewis, 2019; Muddiman, 2013; Vilarins L.;Stabile *et al.*, 2021), mas avança substancialmente no intuito de entendermos como é o comportamento do Google para diferentes termos. Para um termo generalista como “eleições 2020” (termo do ano de 2020), o Google concentrou quase 92,7% de sua entrega em dez sites e uma entrega predominante de mídias tradicionais.

Esse indicador de concentração é um dos achados constantes em todos os trabalhos relacionados ao Google e que, no próximo capítulo, será trabalhado como um indicador importante para medirmos a interferência do buscador no processo eleitoral.

A possibilidade de enganar o algoritmo

Para a importante busca sobre candidatos em uma determinada cidade, o Google entregou sites que automatizaram a produção de conteúdo a partir de bases de dados públicas, com baixa qualidade informacional e que simulavam conteúdo jornalístico local. Os resultados aqui mostram como é possível “enganar” o algoritmo do Google ao simular-se uma geolocalização por meio de palavras-chave bem localizadas nos textos e na URL amigável (um ponto já discutido no capítulo 3 sobre como o Google funciona).

Concentração maior em municípios menores

Um outro achado relevante, e até então não explorado na literatura, é que o Google concentra mais resultados para veículos menores. Esse dado, concomitante ao fato de que é possível simular uma entrega regional e ser privilegiado pelo Google, leva-nos a concluir que em municípios menores há uma entrega de conteúdo com menor qualidade informacional por parte do algoritmo. Um ponto para se explorar em estudos futuros é se há a existência de pequenos produtores de conteúdo que poderiam ser melhor exibidos nessas buscas e que não estão sendo.

Candidaturas com presença institucional se saem bem

Quando as buscas trazem o nome do candidato, aqueles que tiverem presenças institucionais (prefeituras, câmara dos deputados ou outros sites de mesma natureza) possuem maiores chances de terem esses resultados bem posicionados, assim como as páginas de suas mídias sociais também são importantes.

Candidatas se referem a vagas de emprego?

Uma análise qualitativa que poderá ser feita em estudos futuros é a hipótese de que o Google entrega para o termo feminino de candidato, “candidata”, páginas relacionadas a candidatura de empregos ou RH. O que pode mostrar, como já discutido no Capítulo 3, a possibilidade do Google realizar entregas equivocadas.

8 O GOOGLE INTERFERIU NOS RESULTADOS DE BUSCA?

Este capítulo se propõe a investigar se existe alguma mudança no padrão dos resultados entregues pelo Google nas buscas realizadas ao longo do processo eleitoral brasileiro de 2020. A pesquisa realizada por Epstein e Robertson (2015), Epstein *et al.* (2017) parte do pressuposto de que se houver alguma manipulação nos resultados de pesquisa com o intuito de favorecer alguma possível candidatura, o eleitor que tiver um baixo nível de conhecimento poderá sofrer com o *search engine manipulation effect* (SEME). Na pesquisa realizada por Robertson *et al.* (2018), que se pode usar como parâmetro para esse tipo de situação, a conclusão é de que houve sim um viés na entrega dos resultados do Google nas eleições presidenciais americanas de 2018, ainda que em níveis muito baixos.

Se houve algum tipo de viés na entrega dos resultados em um evento da magnitude das eleições dos Estados Unidos, com duas campanhas altamente digitais e profissionalizadas, como fica a atuação do Google durante o provimento de resultados das buscas político-eleitorais nas eleições subnacionais brasileiras? Como já analisado no capítulo 6, o Google é sim uma fonte de busca sobre candidatura e, como observado no capítulo 5, é relevante para eleitores indecisos.

Ao mesmo tempo, diante de recentes discussões e posicionamentos de plataformas como X e Meta, é de se esperar que alguma ação o Google possa tomar para se isentar de possíveis discussões a respeito de algum favorecimento para alguma campanha. No trabalho realizado pelo grupo de pesquisa Resocie (Stabile *et al.*, 2019; Valente e Franco, 2018), nas eleições nacionais de 2018, e como observado no capítulo anterior, constatou-se uma forte preponderância de um conjunto limitado de mídias tradicionais. Esses achados geram uma pergunta relevante, que é a origem principal deste capítulo: haveria alguma interferência do Google privilegiando os grandes portais e diminuindo a entrega para diferentes sites na web? Afinal de contas, poderia ser uma estratégia privilegiar a entrega de grandes portais para evitar discussões sobre possíveis resultados que favoreceriam algum candidato, mitigando os riscos de *fake news*.

Não é sobre o conteúdo

Neste trabalho não será analisado o aspecto do conteúdo daquilo que é entregue nos resultados do Google, como o viés ideológico, que é algo relativamente comum na literatura de mídia e política, o que, inclusive, foi realizado por Robertson *et al.* (2018). Isso seria especialmente desafiador em um contexto subnacional e pluripartidário como o brasileiro. O

foco, portanto, é medir de maneira mais quantitativa a possível interferência do Google. Para tanto, interessa conhecermos o conceito de interferência algorítmica, que é definido como qualquer alteração intencional nos algoritmos de entrega dos resultados de busca que não pode ser explicada por alguma mudança no contexto, seja ela na produção de conteúdo (estão falando mais sobre) ou pelo comportamento dos usuários (estão buscando mais, por exemplo).

É sobre a quantidade de diferentes domínios

Um dos aspectos mais estudados sobre o Google é a sua imensa concentração dos resultados nas “mãos de poucos” sites (Hindman, 2009; Unkel e Haim, 2021; Urman *et al.*, 2021; Trielli e Diakopoulos, 2022). Uma maneira, portanto, de medirmos algum padrão de resultados que não seja o seu aspecto qualitativo de conteúdo é focarmos na **alteração da quantidade de diferentes domínios** nos resultados de busca com termos político-eleitorais. Por exemplo: se antes de se iniciar o período de campanha eleitoral havia uma distribuição média de 9 diferentes domínios para o termo “eleições 2020” e imediatamente na sequência do início do período eleitoral a média diminuiu, é possível aferirmos que houve alguma mudança nesse padrão.

Interferência implica causalidade

Quando escrevemos que há “interferência”, escolhemos essa palavra pois estamos buscando a ideia de que há alguma causalidade entre a mudança do padrão e o início imediato do período de campanha eleitoral e, novamente, que ela não é causada pela mudança de contexto ou dos usuários. Ainda que seja uma delimitação bem pontual, o desafio para estudá-la é gigante, principalmente em uma plataforma “caixa-preta”, como já mencionamos, e sem abertura para a comunidade científica.

Hipótese

Para verificar, portanto, se houve alguma interferência do Google nos resultados durante as campanhas eleitorais, adotamos a seguinte hipótese:

H1: há uma mudança na quantidade de diferentes domínios entregues pelo Google nos resultados entregues para as buscas político-eleitorais ao se iniciar a campanha eleitoral.

Assume-se essa hipótese pois é razoável que o Google queira se proteger de alguma maneira durante um evento com tamanha expressão e que, como depois fora divulgado, teve um dos seus termos elencados como a palavra do ano de 2020. Para além desses pontos, há também a iminência de leis e regulação das plataformas digitais, que é a maneira com que o Tribunal Superior Eleitoral tem buscado responsabilizar ou se aproximar das plataformas.

8.1 Abordagem empírica: diferença nas diferenças (DiD)

Para avaliarmos a hipótese proposta, investigaremos a diversidade no número de domínios exibidos nos resultados de buscas eleitorais em comparação com as buscas não eleitorais. Uma abordagem inicial poderia ser compararmos a média da quantidade de domínios distintos em ambos os tipos de busca durante o período eleitoral. No entanto, essa análise seria simplista, pois precisamos considerar dois aspectos importantes: a diferença já existente no número de domínios entre buscas eleitorais e não eleitorais e a variação temporal natural nos padrões de busca. Portanto, uma comparação direta ou uma análise antes e depois seriam insuficientes, pois seriam influenciadas por mais fatores do que o efeito específico que buscamos captar.

Para superarmos esse desafio, empregaremos o método das diferenças em diferenças (DiD), que é uma abordagem quase-experimental que tem como intuito verificar a causalidade ao comparar as diferenças entre os dois tipos de buscas após ajustar pela diferença pré-existente antes do período eleitoral. O método DiD envolve a definição de dois grupos, “tratamento” e “controle”, e um ponto de corte temporal que estabelece os períodos “antes” e “depois” do tratamento. Além disso, o método permite a inclusão de variáveis explicativas adicionais para se aumentar a precisão da análise.

Neste estudo, aplicamos o DiD para investigar como as buscas eleitorais, grupo “tratamento”, influenciam a quantidade de domínios distintos nos resultados de busca do Google em comparação com as buscas não eleitorais, grupo “controle”.

8.1.1 Grupos de tratamento e controle e períodos pré e pós-tratamento

As definições precisas dos grupos de tratamento e controle são:

- **1. Tratamento:** buscas no Google que envolvem termos eleitorais;
- **2. Controle:** buscas não eleitorais, o que nos fornece uma comparação na mudança de período.

Por sua vez, o que delimita o início efetivo da intervenção (ou tratamento) é a data de início da campanha eleitoral de 2020, especificamente:

- **1. Pré-eleitoral:** entre os dias 28 de agosto e 26 de setembro de 2020, configurando 30 dias de coleta;
- **2. Campanha eleitoral:** entre os dias 27 de setembro, data de início da campanha eleitoral,

até 15 de novembro de 2020, 50 dias de coleta, portanto.

Além disso, em prol da comparabilidade, tivemos que restringir a nossa amostra àquelas cidades que foram pesquisadas no período anterior e posterior ao início da campanha, bem como restringir a coleta apenas até o 1º turno, no dia 15 de novembro.

No apêndice 10.3.2 deste trabalho há o detalhamento de todas as buscas e suas tipologias, enquanto que na seção metodologia há um descritivo de como foram as abordagens técnicas de captura desses dados.

A seguir, temos as quantidades de buscas para cada uma das abordagens:

Tabela 35 – Quantidade e proporção de buscas (em 1000)

Tipo	1º Turno	Pré-Campanha
Buscas Não Políticas (Controle)	128,6 (56%)	91,7 (69%)
Buscas Políticas (Tratamento)	102,2 (44%)	42,1 (31%)
Total	230,7 (100%)	133,8 (100%)

Fonte: elaborado do autor

8.1.2 Como o método DiD funciona

A chave para o método DiD reside na comparação de duas diferenças:

- **1. Diferença no período pré-eleitoral:** primeiramente, comparamos a variável dependente (quantidade de diferentes domínios) durante o período pré-eleitoral entre as buscas eleitorais (tratamento) e as buscas não eleitorais (controle). Essa diferença nos fornece a diferença pré-existente entre os referidos grupos;
- **2. Diferença no período de campanha:** em seguida, comparamos a variável dependente durante o período de campanha do primeiro turno entre as buscas eleitorais (tratamento) e as buscas não eleitorais (controle). Essa diferença nos fornece a diferença pós-intervenção, que contém tanto o efeito que buscamos quanto a diferença pré-existente entre os grupos.

O DiD consiste em subtrair a diferença observada no período pré-eleitoral da diferença observada no período de campanha do primeiro turno, ou seja, o método consiste em limpar a diferença pós-tratamento da diferença preexistente. Se essa diferença resultante for estatisticamente significativa, podemos inferir que a intervenção (buscas eleitorais) teve um efeito causal sobre a variável de interesse. Subjacente ao método está a hipótese de que, uma vez controlado pelos valores pré-tratamento e outras variáveis explicativas, qualquer diferença residual (em

Tabela 36 – Procedimento de Diferenças-em-Diferenças (DiD)

	1º Turno	Período Pré-Eleitoral	Diferença
Tratamento (Buscas Eleitorais)	B	A	B - A
Controle (Buscas Não eleitorais)	D	C	D - C
Diferença	B - D	A - C	Resultado = (B - A) - (D - C)

Fonte: elaborado pelo autor.

características observáveis ou não) entre os grupos se deve ao tratamento. Essa hipótese é comumente chamada de “tendência comum” (Wooldridge, 2010; Cameron e Trivedi, 2005). A tabela a seguir ilustra as etapas do procedimento.

8.1.3 Variáveis dependentes

Desenvolvemos duas variáveis dependentes para testarmos a nossa hipótese, ambas estão relacionadas à quantidade de domínios distintos apresentados nos resultados das busca orgânicas do Google. Ao todo, foram 14.208 domínios quando observamos o período específico abordado na análise DiD. As variáveis dependentes são:

- **Domínios nos três primeiros resultados (Top 03):** número de domínios distintos que apareceram nos três primeiros resultados como resultado da busca. Ao todo foram 4.843 diferentes domínios. Como já mencionado, os 3 primeiros links correspondem a cerca de 95% dos cliques feitos nos resultados de busca orgânica;
- **Domínios nos dez primeiros resultados (Top 10):** número de domínios distintos que apareceram nos dez primeiros lugares como resultado da busca. A ideia dessa variável é trazer mais robustez à análise ao permitir a comparação do efeito das possíveis mudanças nas entregas do Google na primeira página de resultados.

As duas variáveis nos dão visões complementares. Para testarmos a sobreposição entre elas, aplicamos testes de correlação nos quais observamos uma correlação positiva moderada, com um coeficiente de 0.6. Esse resultado sugere uma relação significativa, porém não extremamente forte, entre as variáveis. Tal achado reforça a importância de se considerar ambas as medidas.

8.1.4 Resultados iniciais do DiD (sem variáveis explicativas)

Como abordagem inicial, conduzimos o método DiD sem variáveis explicativas; desse modo, seremos capazes de comparar a magnitude do efeito estimado com as variáveis explicativas.

Tabela 37 – Médias para os 10 primeiros resultados - Top 10

	1º Turno	Período Pré-Eleitoral	Diferença
Tratamento (Buscas Eleitorais)	7,28	7,14	0,14
Controle (Buscas não eleitorais)	8,44	8,39	0,06
Diferença	0,14	0,06	0,09

Fonte: elaborado pelo autor.

Tabela 38 – Médias para os 03 primeiros resultados - Top 03

	1º Turno	Período Pré-Eleitoral	Diferença
Tratamento (Buscas Eleitorais)	2,33	2,19	0,14
Controle (Buscas não eleitorais)	2,54	2,52	0,02
Diferença	0,14	0,02	0,13

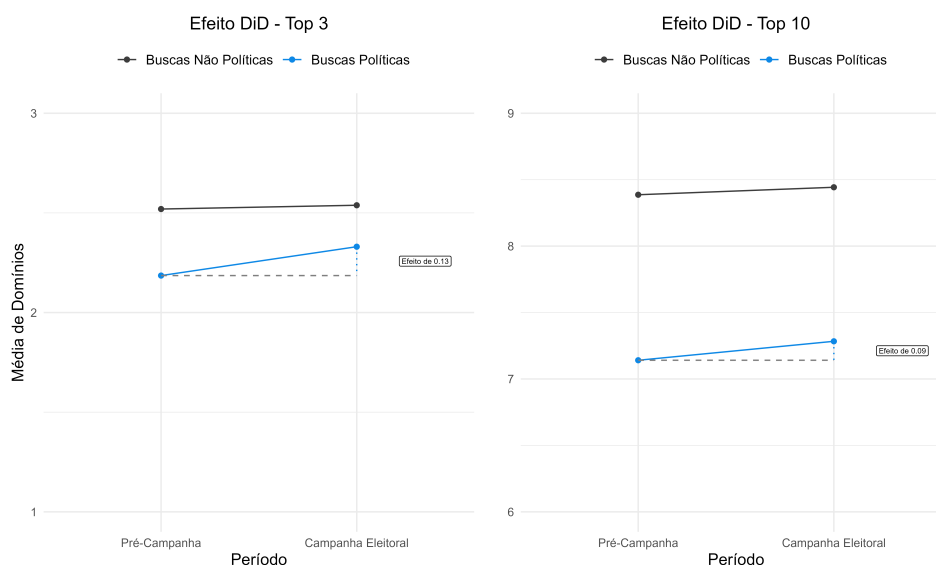
Fonte: elaborado pelo autor.

As tabelas 37 e 38 apresentam os resultados para as variáveis Top3 e Top10, substituindo as letras da tabela anterior pelos valores reais da variável dependente, que se referem à quantidade de diferentes domínios nos dez primeiros resultados.

Quando analisamos a variável dependente construída com base nos diferentes domínios exibidos nos resultados de cada busca no Google na primeira página (top 10 resultados), encontramos um efeito de 0.09, o que indica que houve um aumento na quantidade média de domínios após o início do período eleitoral.

Ao analisar a variável dependente construída com base nos diferentes domínios exibidos nos resultados de cada busca do Google apenas pelos 3 primeiros resultados, encontramos um efeito de 0.13, mas que também indica de maneira consistente um aumento na quantidade média de domínios após o início do período eleitoral. No entanto, nota-se que, como o valor estimado está na mesma unidade que as variáveis dependentes, a magnitude do efeito é pequena.

Uma outra abordagem didática para entendermos o método DiD é visual. A figura 21 ilustra o funcionamento do método:

Figura 21 – Visão gráfica do método de diferenças em diferenças

Fonte: elaborado pelo autor.

A linha escura indica a trajetória do grupo controle (buscas não políticas), que se mantém com um leve crescimento, enquanto a linha azul-claro mostra o grupo tratamento. A linha pontilhada indica como deveria ser o crescimento da linha clara se ela seguisse o mesmo padrão do grupo controle, a diferença entre a linha azul-claro e a linha pontilhada é a ilustração da diferença em diferenças.

8.1.5 Modelo com variáveis explicativas

Para aprimorarmos a precisão da nossa análise empírica, incluímos variáveis de controle que podem afetar os resultados das buscas, mas que não foram diretamente influenciadas pelo início do período eleitoral. A seleção dessas variáveis foi baseada na busca de fatores relacionados aos resultados das buscas, mas que permanecem independentes do contexto eleitoral.

As variáveis independentes incluídas são:

Tipo de busca

Conforme discutido no capítulo anterior, o planejamento desta pesquisa foi elaborado para abranger uma diversidade de abordagens nos termos pesquisados. A necessidade de controlarmos os tipos de busca se justifica pela sua ampla variedade e abrangência, fatores que podem ser cruciais na precisão das estimativas. A distribuição dos diferentes tipos de buscas foi organizada entre categorias político-eleitorais:

Tabela 39 – Tipos de busca por grupo (tratamento e controle)

	Freq.	Prop.
Controle - Buscas aleatórias da população	116.759	32,0%
Controle - Principais buscas do Google	27.582	7,6%
Tratamento - Nome dos candidatos nas capitais	23.406	6,4%
Tratamento - Buscas políticas gerais	164.452	45,1%
Tratamento - Perguntas político-eleitorais	32.408	8,9%
Total	364.607	100%

Fonte: elaborado pelo autor.

Região geográfica

Conforme evidenciado pela literatura (Hannak *et al.*, 2013; Kliman-Silver *et al.*, 2015), a geolocalização é o principal fator que influencia os resultados fornecidos pelo mecanismo de busca. Como detalhado na seção de metodologia, todas as buscas realizadas neste estudo foram geolocalizadas em uma amostra representativa de diversos municípios brasileiros. A distribuição dessas buscas por região é apresentada na tabela a seguir:

Tabela 40 – Distribuição das buscas nas regiões brasileiras

Regiões	Freq.	Prop.
Centro-Oeste	29.356	8,1%
Nordeste	97.584	26,8%
Norte	44.694	12,3%
Sudeste	132.881	36,4%
Sul	60.092	16,5%
Total	364.607	100%

Fonte: elaborado pelo autor.

Dispositivo móvel

Além da disposição geográfica, o Google nos últimos anos penalizou sites que não eram *mobile first*. Isso significa que sites otimizados para dispositivos móveis passaram a ter prioridade nos resultados de busca, refletindo a crescente tendência de uso da internet em *smartphones* e *tablets*. A importância dessa otimização está relacionada à experiência do usuário, velocidade de carregamento e acessibilidade em telas menores.

Considerando-se essa tendência, a variável de dispositivos utilizados foi programada

a priori no processo de coleta de dados. Optamos por uma proporção aleatória de 2 buscas em dispositivos móveis para cada busca em desktop, com o intuito de simular, assim, o comportamento médio dos usuários e capturar possíveis variações nos resultados de busca devido à otimização *mobile first*.

Tabela 41 – Quantidade de buscas por dispositivo móvel

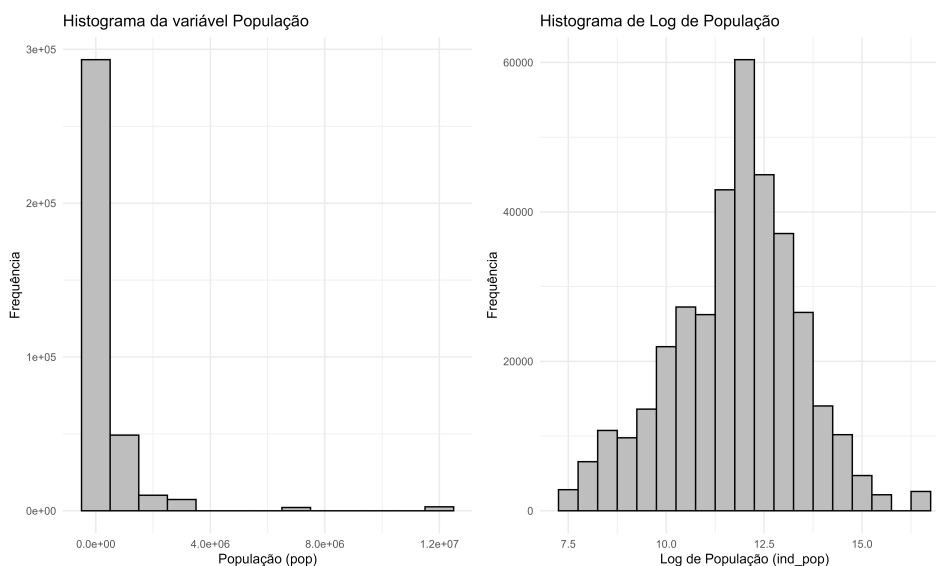
Dispositivo Móvel	Freq.	Prop.
desktop	106.118	29,1%
mobile	258.489	70,9%
Total	364.607	100%

Fonte: elaborado pelo autor.

População

O tamanho dos municípios onde as buscas foram simuladas pode influenciar de maneiras distintas a possível interferência do Google. Cidades maiores podem receber maior atenção e, conseqüentemente, experimentar algum tipo de ajuste algorítmico mais significativo. Por outro lado, cidades menores podem ter uma demanda de pesquisa menor, resultando em um incentivo reduzido para que grandes portais publiquem ou discutam determinados tópicos.

Para minimizar o impacto da variação do tamanho dos municípios, aplicamos a transformação logarítmica à variável correspondente. Essa abordagem permite uma comparação mais precisa e equilibrada entre os grupos de controle e tratamento ao longo do tempo. Sem essa normalização, as variações extremas nos tamanhos dos municípios poderiam distorcer os resultados, afetando a interpretação dos efeitos estimados pelo modelo.

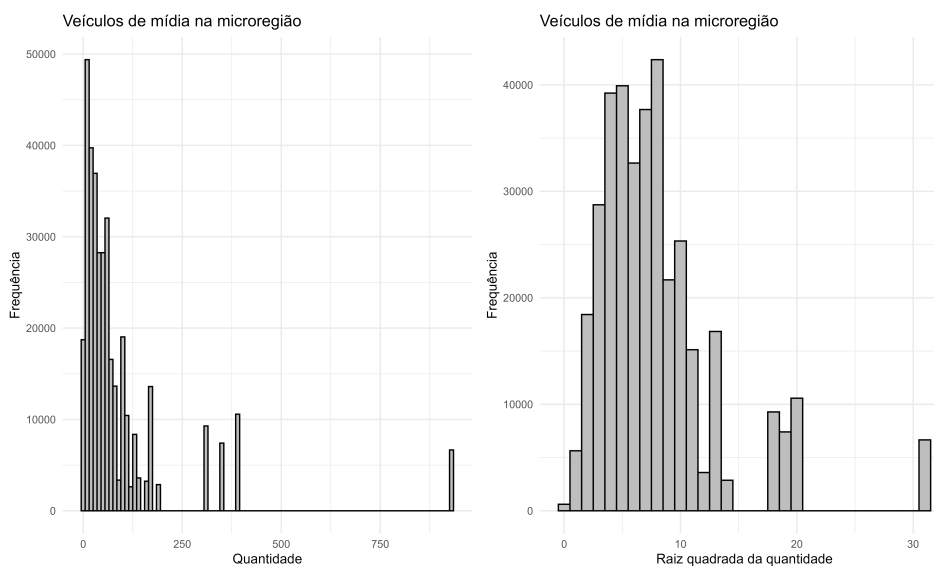
Figura 22 – Histograma da variável população

Fonte: elaborado pelo autor.

Quantidade de veículos de mídia

Para controlarmos os resultados, utilizamos dados do Atlas da Notícia (PROJOR - Instituto para o Desenvolvimento do Jornalismo 2022) a fim de calcularmos a soma de veículos de comunicação na microrregião do município onde a busca foi simulada. O objetivo dessa abordagem é utilizarmos a quantidade de produtores de conteúdo oficial como um *proxy* para a oferta de informação na região. Presume-se que, em locais com uma maior oferta de produção de conteúdo, haja uma maior diversidade de domínios nos resultados das buscas.

Para lidarmos com a variação na quantidade de veículos e considerarmos a existência de locais com estimativa zero, aplicamos a transformação da raiz quadrada à variável. Essa transformação ajuda a reduzir a amplitude dos valores, facilitando uma análise mais equilibrada e precisa do impacto dessa variável nos resultados do estudo.

Figura 23 – Histograma da variável Veículos de mídia

Fonte: elaborado pelo autor.

8.1.6 Aplicando o DiD com variáveis de controle (regressão)

Embora a análise manual do método das diferenças em diferenças (DiD) seja uma abordagem interessante e didática para se estimar o efeito, ela possui limitações, como a incapacidade de se controlar o efeito de covariáveis e a ausência de uma medida de incerteza. Por isso, recorreremos à **regressão linear múltipla** para incorporarmos as variáveis independentes e obtermos medidas de erro estatístico. Essa abordagem é considerada padrão na aplicação do método DiD, uma vez que ambos estimam efeitos no nível médio e têm uma relação estreita.

Além disso, a regressão permite isolar o efeito do tratamento de outras influências potenciais, tornando os resultados mais confiáveis e interpretáveis. Ela também oferece a flexibilidade de modelar relações mais complexas entre as variáveis, considerando-se interações e não linearidades, o que enriquece a análise.

8.1.6.1 Estimação do Modelo

A equação a seguir representa o modelo estatístico que utilizamos para nossa análise

DiD:

$$\begin{aligned}
 \text{Quantidade de Diferentes Domínios} = & \beta_0 + \beta_1 \times \text{Indicador de busca política} \\
 & + \beta_2 \times \text{Indicador de período eleitoral} \\
 & + \beta_3 \times (\text{Indicador de busca política} \times \text{Indicador de período eleitoral}) \quad (\text{Termo de interação DiD}) \\
 & + \beta_4 \times \text{Indicador de tipo de busca} \\
 & + \beta_5 \times \text{Indicador de região geográfica} \\
 & + \beta_6 \times \text{Indicador de dispositivo móvel} \\
 & + \beta_7 \times \text{Log da população do município} \\
 & + \beta_8 \times \text{Raiz quadrada da quantidade de veículos na microrregião} \\
 & + \varepsilon
 \end{aligned}
 \tag{8.1}$$

Neste modelo, o termo de interação entre o indicador de busca política e o indicador de período eleitoral é fundamental, pois ele representa o resultado-chave do método DiD. Esse termo captura o efeito diferencial do período eleitoral em buscas políticas na concentração de domínios, ou seja, ele mede a mudança na concentração de domínios em buscas políticas, que pode ser atribuída especificamente ao início do período eleitoral.

No contexto do DiD, esse termo de interação é essencial para isolarmos o efeito do tratamento (nesse caso, o início do período eleitoral) das tendências gerais que afetam tanto o grupo de tratamento quanto o grupo de controle. Assim, a estimativa desse termo nos fornece uma medida precisa do impacto do período eleitoral nas buscas políticas, controlada por outros fatores que possam influenciar a concentração de domínios.

8.1.6.2 Janela de 7 dias

Para testarmos a robustez do nosso modelo, incorporamos uma análise adicional focada em uma janela de 7 dias ao redor do início do período eleitoral. Essa abordagem é inspirada na técnica de regressão descontínua, que é frequentemente utilizada para se avaliar os efeitos de intervenções em pontos específicos no tempo. Ao adotarmos essa estratégia, estamos sendo especialmente cautelosos, dada a importância do tema e a magnitude dos efeitos que estamos investigando, portanto decidimos submeter nosso modelo a um teste mais rigoroso.

A importância de se analisar essa janela de 7 dias reside na suposição de que as buscas realizadas imediatamente antes e depois do início do período eleitoral são mais comparáveis

em termos de interesse e comportamento do usuário. Isso nos permite isolar o efeito do início do período eleitoral, minimizando distorções que poderiam ser causadas por variações sazonais ou eventos específicos que ocorrem fora desse intervalo. Essa análise mais focada ajuda a garantir que as diferenças observadas sejam de fato atribuíveis ao início do período eleitoral, proporcionando uma avaliação mais precisa e confiável do impacto nas buscas políticas.

Tabela 42 – Comparação dos modelos Top 10 e Top 3 domínios

	<i>Dependent variable:</i>	
	depvar_qtd_domain_top10	depvar_qtd_domain_top03
	Modelo Top 10	Modelo Top 3
	(1)	(2)
Busca política (Tratamento)	-0.380*** (0.013)	0.032*** (0.006)
Período eleitoral	0.059*** (0.009)	0.020*** (0.004)
Classificação de busca não política por temas	-0.128*** (0.010)	-0.098*** (0.005)
Classificação de busca política em capitais	-0.045*** (0.013)	-0.195*** (0.006)
Classificação geral de busca política	-1.366*** (0.010)	-0.551*** (0.005)
Classificação geral de busca política temática		
Indicador de região: Nordeste	-0.077*** (0.010)	-0.042*** (0.005)
Indicador de região: Norte	-0.091*** (0.011)	-0.063*** (0.005)
Indicador de região: Sudeste	-0.108*** (0.010)	-0.046*** (0.005)
Indicador de região: Sul	-0.028*** (0.011)	-0.032*** (0.005)
Dispositivo móvel	-0.675*** (0.006)	-0.190*** (0.003)
População do município	0.068*** (0.002)	0.021*** (0.001)
Qtd de veículos na microrregião	0.007*** (0.001)	0.003*** (0.0003)
Efeito DiD: busca política * período eleitoral	0.200*** (0.011)	0.170*** (0.005)
Constante	8.145*** (0.024)	2.460*** (0.012)
Observations	364,607	364,607
R ²	0.241	0.107
Adjusted R ²	0.241	0.107

Note:

*p<0.1; **p<0.05; ***p<0.01

Fonte: elaborado pelo autor.

8.1.6.3 Análise dos modelos

Para todo o período

- **Busca política (tratamento):** no modelo de análise dos dez primeiros links orgânicos, ou primeira página de resultados do Google (Top 10), observa-se um coeficiente negativo significativo, indicando que buscas políticas tendem a apresentar menos diversidade nos domínios em comparação com buscas não políticas. No modelo dos três primeiros links (Top 3), o coeficiente é próximo de zero e não significativo, sugerindo que, para os três primeiros resultados, a natureza política da busca não afeta significativamente a diversidade dos domínios;
- **Período eleitoral:** ambos os modelos mostram um coeficiente positivo e significativo, indicando que, durante o período eleitoral, há uma maior diversidade de domínios nos

resultados de busca;

- **Tipo de busca:** é interessante notar que a classificação geral de busca política tem um coeficiente negativo significativo, e grande em ambos os modelos, o que sugere uma menor diversidade de domínios em buscas políticas gerais em comparação com outras buscas, o que é coerente com as tabelas descritivas observadas no capítulo anterior, em que a concentração para os principais domínios era significativamente maior;
- **Região geográfica:** as variações regionais são consistentes em ambos os modelos, com algumas regiões mostrando diferenças significativas na diversidade de domínios;
- **Dispositivo móvel:** esse coeficiente negativo e significativo em ambos os modelos sugere que buscas realizadas em dispositivos móveis tendem a apresentar menos diversidade de domínios;
- **População do município:** mostra um coeficiente positivo e significativo, indicando que municípios mais populosos apresentam maior diversidade de domínios nos resultados de busca;
- **Quantidade de veículos na microrregião:** esse coeficiente positivo e significativo sugere que uma maior quantidade de veículos de mídia na microrregião está associada a uma maior diversidade de domínios;
- **Mediana da audiência:** varia entre os modelos, com o modelo Top 10 mostrando um efeito significativo e o modelo Top 3 apresentando um efeito menor.

Resultado do coeficiente de interação (efeito DiD)

O modelo DiD examina o impacto do período eleitoral nas buscas políticas, medindo a diversidade de domínios exibidos nos resultados de busca do Google. A variável dependente no modelo Top 10 é a quantidade de diferentes domínios, variando de 1 a 10, enquanto no modelo Top 3 é a quantidade de diferentes domínios exibidos nas três primeiras posições de uma página do Google, variando de 1 a 3.

O coeficiente de interação entre “busca política” e “período eleitoral” é o foco principal desse modelo, pois ele capta o efeito combinado da natureza da busca (política ou não) e do período eleitoral. Em ambos os modelos, esse coeficiente é positivo e significativo, indicando que durante o período eleitoral as buscas políticas apresentaram um aumento na diversidade de domínios em comparação com outros períodos. Especificamente, o coeficiente de interação é **0,206** no modelo Top 10 e **0,173** no modelo Top 3, sugerindo que o efeito é mais pronunciado na diversidade dos 10 primeiros domínios do que nos 3 primeiros.

Para ilustrar o significado do coeficiente de interação, considere o seguinte exemplo: suponha que, fora do período eleitoral, a média de domínios distintos em buscas políticas seja de 5 no Top 10 e 2 no Top 3. Durante o período eleitoral, a presença do efeito combinado da busca política e do período eleitoral aumentaria essas médias para aproximadamente 5,206 no Top 10 e 2,173 no Top 3. Embora o aumento na diversidade de domínios seja numericamente pequeno, ele é estatisticamente significativo, indicando uma mudança consistente no comportamento das buscas durante o período eleitoral.

Convergência e divergência entre os modelos

Ao compararmos os dois modelos, observamos convergência em vários aspectos, como a significância e direção dos coeficientes para variáveis como dispositivo móvel, população do município e tipo de busca. No entanto, há divergência na magnitude dos coeficientes e na significância em algumas variáveis, como a mediana da audiência, que é significativa no modelo Top 10, mas não no modelo Top 3.

Isso sugere que a estrutura geral dos resultados de busca é consistente, ainda que para o Top 10 eles não sejam robustos. Como observamos, existem nuances específicas entre os primeiros três resultados e os primeiros dez.

Tabela 43 – Comparação dos modelos Top 10 e Top 3 domínios na janela de 7 dias

	<i>Dependent variable:</i>	
	depvar_qtd_domain_top03	depvar_qtd_domain_top10
	Modelo Top 10	Modelo Top 3
	(1)	(2)
Busca política	0.129*** (0.014)	-0.400*** (0.032)
Período eleitoral	-0.015* (0.009)	0.009 (0.019)
Buscas não políticas por temas	-0.099*** (0.011)	-0.177*** (0.025)
Buscas política em capitais	-0.272*** (0.015)	0.117*** (0.034)
Buscas políticas generalistas	-0.706*** (0.013)	-1.461*** (0.028)
Buscas políticas temáticas		
Região: Nordeste	-0.050*** (0.011)	-0.062** (0.024)
Região: Norte	-0.065*** (0.012)	-0.074*** (0.028)
Região: Sudeste	-0.045*** (0.011)	-0.103*** (0.024)
Região: Sul	-0.037*** (0.012)	-0.063** (0.026)
Dispositivo móvel	-0.199*** (0.006)	-0.686*** (0.013)
População do município	0.050*** (0.002)	0.099*** (0.005)
Veículos na microrregião	0.003*** (0.001)	0.010*** (0.001)
Efeito DiD: busca política * período eleitoral	0.030*** (0.011)	-0.169*** (0.025)
Constante	2.139*** (0.026)	7.798*** (0.057)
Observations	68,998	68,998
R ²	0.166	0.284
Adjusted R ²	0.165	0.284
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01	

Fonte: elaborado pelo autor.

Análise da Janela de 7 dias Na análise do modelo abrangendo o período eleitoral completo, observamos um aumento significativo na diversidade de domínios nas buscas políticas, conforme evidenciado pelos coeficientes positivos de interação entre a busca política e o período eleitoral. Especificamente, os coeficientes são 0.206 para o modelo Top 10 e 0.173 para o modelo Top 3, indicando uma maior diversidade de domínios durante o período eleitoral.

Entretanto, ao restringirmos a análise a uma janela de 7 dias em torno do início do período eleitoral, notamos uma mudança na tendência. No modelo Top 10, o coeficiente de interação torna-se negativo (-0.169) e significativo, sugerindo uma redução na diversidade de domínios. Por outro lado, no modelo Top 3, o coeficiente é positivo (0.030) e significativo, porém indica um aumento bem mais modesto na diversidade em comparação ao observado no modelo do período completo.

Essa variação nos coeficientes de interação revela que, ao focarmos exclusivamente

nos 7 dias em torno do início do período eleitoral, o padrão de diversidade de domínios em buscas políticas se altera. No modelo Top 10 há uma diminuição na diversidade, enquanto no modelo Top 3 ocorre um leve aumento, embora muito inferior ao verificado no período completo.

Os resultados dos modelos indicam a presença de alguma interferência, embora seu impacto numérico seja relativamente pequeno. Antes do período eleitoral, o Google apresentava, em média, 7,14 sites diferentes nas buscas políticas. Após o início da campanha, esse número aumentou para 7,34, resultando em um efeito de +2,8%. No entanto, entre os três primeiros links, a média inicial de 2,19 aumentou para 2,36, refletindo um efeito de 7,76% e uma menor concentração.

A interpretação dessa variação, seja em termos de concentração ou de aumento na diversidade, ainda requer investigações adicionais. Isso se deve, em parte, ao fato de que a análise da janela de 7 dias para os dez primeiros links revela uma maior concentração, um padrão oposto ao encontrado no período eleitoral completo.

9 CONCLUSÃO

Na difícil escolha entre aprofundar-se em questões específicas ou explorar amplamente o objeto de estudo, este trabalho optou pela segunda abordagem. A decisão consciente de analisarmos o Google sob uma perspectiva abrangente nos colocou diante das inerentes limitações de qualquer pesquisa que precisa chegar a uma conclusão dentro de prazos estabelecidos.

Por outro lado, a análise do papel do Google a partir de diferentes perguntas de pesquisa – quem utiliza, quando e como é relevante, o que o buscador entregou e se houve interferência no período eleitoral – permitiu que contribuíssemos com a agenda emergente de pesquisa sobre mecanismos de busca, os quais tornaram-se mediadores informacionais cruciais. No momento de conclusão desta tese, há uma escassez de estudos detalhados sobre o Google, que é, talvez, o maior mediador informacional da nossa história recente.

A complexidade de se pesquisar o Google é acentuada pela falta de acesso a APIs de dados e pelo mínimo ou nenhum incentivo da empresa à comunidade científica, como por exemplo por meio da abertura de dados para uso científico. Nas próximas seções, detalhamos as contribuições deste estudo para diversas áreas do conhecimento, bem como apresentaremos hipóteses e sugestões para estudos futuros que possam avançar nessa importante linha de pesquisa.

9.1 Contribuições metodológicas e técnicas

Técnica de captura e dados disponíveis publicamente

Este estudo contribui para a área de métodos e técnicas de pesquisa ao enfrentar o desafio de se estudar o Google. A pesquisa foi projetada para ser completamente replicável. Os códigos em Python utilizados para capturarmos os dados dos resultados do Google estão disponíveis no repositório on-line da tese.

Todos os dados da pesquisa, aproximadamente 46 gigabytes, estão acessíveis publicamente a outros pesquisadores interessados em explorar essa vasta base de informações.

Exemplo de processo analítico em R

Além da captura de dados, toda a análise e redação deste trabalho foram realizadas utilizando-se RMarkdown, adaptando-se um template de LaTeX específico para teses de doutorado. Os códigos para modelos e análises estão detalhadamente explicados e comentados no documento principal. Os códigos compilados se transformaram neste texto. Assim, garantimos

que outros pesquisadores possam seguir ou adaptar nossos métodos com facilidade.

Uso de quase-experimento

A introdução na literatura sobre estudos do Google de um modelo de quase-experimento utilizando análise de diferenças em diferenças, complementado por um teste de robustez adaptado de modelos descontínuos, representa uma inovação na área. Essa técnica foi crucial para atribuímos causalidade aos resultados, permitindo discussões fundamentadas (ainda que preliminares) sobre a possível interferência do Google nos resultados de busca.

Uso de históricos de navegação

De forma pioneira na ciência política brasileira e sul-americana, utilizamos históricos de buscas de indivíduos. Uma diferença marcante dessa amostra em relação a estudos internacionais é sua distinção em termos de representatividade, considerando-se a distribuição demográfica. É importante destacar que esses indivíduos não foram recrutados para um estudo específico de política, o que diferencia este trabalho da, até então, única abordagem anterior de mesma natureza, a de Blassnig *et al.* (2023).

Além disso, desenvolvemos um método para identificar e classificar buscas políticas em um conjunto de dados extenso, um exercício que pode ser extremamente útil para pesquisadores que enfrentarem desafios semelhantes de classificação automatizada de grandes volumes de dados.

Limitações de questionários e *surveys*

Este estudo também aborda as limitações dos dados provenientes de questionários, comparando-os com dados de comportamento real. Propusemos um método de agrupamento de escalas de respostas que revelou discrepâncias importantes, como a tendência de respostas socialmente desejáveis em termos de interesse político. Encontramos evidências de que até participantes que afirmaram não realizar buscas políticas, de fato, realizaram-nas. Esse fato levanta questões críticas sobre a precisão das autoavaliações em *surveys* e sua viabilidade para a captura de hábitos cotidianos.

Além de explorar mais a fundo as discrepâncias entre respostas em *surveys* e comportamentos reais, sugerimos que pesquisas futuras investiguem a viabilidade de questionamentos a respeito de hábitos tão integrados ao cotidiano. A precisão das respostas em questionários, especialmente sobre atividades frequentes e inconscientes como buscas na internet, é uma área rica para investigação adicional.

9.2 Contribuições para o campo do comportamento político

Raciocínio motivado e estudos do Google

Nos capítulos 5 e 6, demonstramos a relevância do Google para indivíduos com preferências políticas consolidadas, o que se alinha com a teoria de raciocínio motivado (Taber e Lodge, 2006). Apesar de sua ampla difusão na área de comportamento político, essa teoria raramente foi aplicada em estudos sobre o Google, que são geralmente conduzidos por pesquisadores em comunicação ou ciências computacionais. Este trabalho buscou integrar essas áreas, oferecendo uma nova perspectiva sobre o impacto do comportamento político no uso de mecanismos de buscas.

Diferentes Perfis

Identificamos cinco perfis distintos do eleitorado no capítulo 5, que foram essenciais para o modelo explicativo do uso do Google em pesquisas com termos político-eleitorais. Esses perfis, construídos a partir de um painel antes e após as eleições, podem ser explorados em futuras pesquisas para se medir níveis de participação e engajamento político, contribuindo com novos conceitos que podem ser utilizados em futuros estudos de comportamento político.

Como o eleitor busca por informação

A análise qualitativa de como os eleitores buscam por informações é um aspecto que merece atenção especial. No capítulo 6, a classificação dos termos de busca visou a compreender esse comportamento de maneira ampla, mas há uma grande possibilidade de se analisar esses termos de maneira qualitativa. Entre os eleitores que pesquisavam por candidatos, quais informações buscavam? Histórico, informações pessoais, propostas ou afiliação partidária são variáveis que podem ser correlacionadas a períodos e perfis específicos de eleitores. Esses dados oferecem uma base rica para entendermos as motivações e as necessidades informacionais do eleitorado.

9.3 Contribuições para a área de Comunicação, Mídia e Jornalismo

O que o Google entrega

No capítulo 7, analisamos a concentração das entregas do Google, observando uma maior concentração em municípios menores. Os resultados desse capítulo podem ser relevantes para pesquisadores interessados na intersecção entre mídia e política, especialmente no que diz respeito à concentração de veículos de comunicação, a existência de desertos de notícias e a

dominância de alguns grupos midiáticos.

Como posicionar uma campanha

Os achados desta pesquisa também oferecem diretrizes para pesquisadores de comunicação política a respeito de quais veículos de mídia e formatos de campanha tendem a ser mais eficazes em sua presença no Google. Pesquisas futuras também podem explorar como diferentes campanhas e partidos se beneficiam dos padrões de entrega do Google. Há melhores entregas para candidatos de partidos mais conservadores ou progressistas?

Como ser bem-sucedido em SEO

Uma outra discussão importante e que pode ser útil especialmente para jornalistas e veículos de mídia é: como obter posições melhores no ranking do Google no período eleitoral. Os dados mostraram que é possível, a partir de uma construção automatizada de páginas com URLs, “enganar” o algoritmo do Google e fazer com que um veículo pouco conhecido seja bem-sucedido nas buscas em todo o Brasil.

Essa hipótese de trabalho, inclusive, pode ser objeto de estudos futuros para se colocar à prova a possibilidade de se “enganar” o algoritmo e, principalmente, discutir-se questões sobre qualidade de entrega em todo o país.

9.4 Contribuições para os estudos a respeito do Google

O search engine manipulation effect é superestimado

Os resultados encontrados no capítulo 5 questionam a prevalência do *search engine manipulation effect* (SEME), sugerindo que sua ocorrência pode ser superestimada. Os dados coletados e analisados neste trabalho indicam que a influência do Google nas eleições pode ser inexistente, ou bem menor do que a sugerida em trabalhos anteriores.

O principal achado que sustenta essa conclusão é a comprovação da hipótese de que usuários com preferência política mais consolidada são os que mais utilizam o Google. Ou seja, os usuários que eventualmente estariam expostos a uma manipulação dos resultados já teriam um certo nível de preferência política, o que afetaria pouco ou nada a sua opinião política.

Essa hipótese é uma consequência lógica da incorporação da literatura de comportamento político aos estudos sobre o Google, algo que normalmente não acontece, pois quase sempre esses estudos são conduzidos por pesquisadores do campo da ciência de computação.

Google interfere nos resultados de busca

Com base na análise dos modelos de diferenças em diferenças (DiD) para as buscas

político-eleitorais durante a campanha eleitoral de 2020, observamos evidências que oferecem suporte à hipótese de uma mudança de padrão nos resultados entregues pelo Google. No entanto, é importante destacarmos que os coeficientes encontrados são numericamente baixos, sugerindo que, embora haja uma alteração no padrão, o efeito é relativamente pequeno.

A análise do modelo restrito à janela de 7 dias ao redor do início da campanha eleitoral revela uma mudança interessante: o coeficiente de interação, que capta o efeito combinado da busca política e do período eleitoral, muda de sinal no modelo que analisa os dez primeiros resultados orgânicos e diminui ainda mais no modelo restrito aos 3 principais resultados. Isso indica que a dinâmica da diversidade de domínios em buscas políticas pode variar significativamente em momentos específicos do período eleitoral.

Essas observações sugerem que, de fato, pode haver alguma mudança no padrão de entrega dos resultados pelo Google durante as eleições. Embora o efeito seja pequeno, dada a enorme quantidade de buscas realizadas, ele pode ter implicações relevantes. No entanto, é crucial reconhecermos que, dada a opacidade dos algoritmos do Google e as complexidades envolvidas na oferta e demanda de buscas políticas, é desafiador atribuímos essa diferença exclusivamente à interferência algorítmica. Como discutido nos capítulos teóricos, outros fatores, como mudanças na oferta de conteúdo ou um aumento na demanda por buscas políticas, podem também desempenhar algum papel.

Embora os resultados indiquem a existência de uma mudança no padrão dos resultados de buscas político-eleitorais durante o período eleitoral, a magnitude do efeito é limitada. Esses indícios justificam a necessidade de um monitoramento contínuo das práticas do Google para compreendermos melhor o tipo de interferência que pode ocorrer e as implicações dessas mudanças para a democracia e a liberdade de informação.

9.5 Contribuições para os estudos de internet e política

Interferência algorítmica

Este estudo levanta a questão de se o Google pode ter interferido nos resultados durante o processo eleitoral. Um ponto importante que pode ser objeto de pesquisas futuras é entender se a interferência é um efeito automatizado do algoritmo ou se resulta de uma decisão consciente.

Ademais, os dados coletados oferecem uma base para se explorar como o uso de termos específicos, como “candidatas”, , frequentemente relacionados a vagas de empregos em

vez de informações políticas, pode influenciar os tipos de resultados entregues pelo Google.

Disponibilização de códigos e estrutura para um observatório público

Uma das contribuições deste trabalho é a disponibilização dos códigos e do desenho metodológico necessário para se construir um observatório público dos resultados do Google. Isso se demonstra urgente, na medida em que as evidências demonstram a sua alta concentração de entrega de resultados em poucos portais, a possível interferência durante o período eleitoral e a conclusão de que usuários indecisos utilizam sim o Google como fonte de informação para a tomada de decisão sobre em quem votar.

É essencial desenvolver-se uma aplicação de abrangência nacional, considerando-se a influência geográfica nos resultados de busca. Este estudo forneceu os códigos e a metodologia (incluindo os códigos UULE discutidos na seção metodológica) para a coleta de dados em nível municipal.

Como modelo para a coleta de dados, utilizou-se a API da empresa SERPAPI. Embora o custo em dólar represente uma limitação, durante este estudo, a empresa ofereceu descontos significativos para projetos de pesquisa, o que pode facilitar o acesso a essa ferramenta para iniciativas científicas.

Esta tese também serve como referência para o tratamento dos dados coletados. A conversão de URLs em domínios permitiu uma análise eficaz da concentração mediática. No repositório do trabalho, estão disponíveis bases de dados que classificam os domínios por tipologias (veículos de comunicação, sites de compras etc.) e associações a grandes grupos de mídia, recursos valiosos tanto para o observatório quanto para futuras pesquisas.

9.6 E se o Google mudar, como ficam essas conclusões?

Nos próximos 10 ou 20 anos é bem provável que o formato que o mundo conhece hoje como buscador mude. Já durante a escrita deste trabalho existem sugestões como a Perplexity.ai, que utiliza inteligência artificial generativa para se construir parágrafos “mais didáticos” sobre uma busca. O eleitor terá um resumo dos resultados construído por uma IA e não precisará mais entrar nos resultados de busca.

Nesse contexto de resumo dos resultados, cresce ainda mais a necessidade de acompanharmos com quais critérios esses resumos estão sendo realizados e, principalmente, como o cidadão se relaciona com essa mediação. Independentemente de como os buscadores venham a funcionar, a necessidade de pesquisá-los continuará, seja para entendermos como eles estão

funcionando ou como eles moldam a percepção do eleitor sobre a realidade política.

REFERÊNCIAS

- ALVAREZ, R. M.; LI, Y. Survey attention and self-reported political behavior. **Public Opinion Quarterly**, v. 86, p. 793–811, 2022.
- ANSOLABEHERE, S.; HERSH, E. D. Validation: what big data reveal about survey misreporting and the real electorate. **Political Analysis**, v. 20, p. 437–459, 2012.
- BAEZA-YATES, R. Bias on the web. **Commun. ACM**, Association for Computing Machinery, New York, NY, USA, v. 61, n. 6, p. 54–61, may 2018. ISSN 0001-0782. Disponível em: <https://doi.org/10.1145/3209581>.
- BAKER, A.; AMES, B.; RENNO, L. R. Social context and campaign volatility in new democracies: Networks and neighborhoods in brazil's 2002 elections. **American Journal of Political Science**, v. 50, p. 382–399, 2006. ISSN 00925853.
- BAUM, M. A. **Soft News Goes to War: Public Opinion and American Foreign Policy in the New Media Age**. Princeton University Press, 2003. ISBN 9780691123776. Disponível em: <http://www.jstor.org/stable/j.ctt7sfmh>.
- BECKER, V.; GAMBARO, D.; FILHO, G. L. d. S. O impacto das mídias digitais na televisão brasileira: queda da audiência e aumento do faturamento. **Palabra Clave - Revista de Comunicación**, Universidad de la Sabana, v. 18, n. 2, p. 341–373, maio 2015. ISSN 2027-534X. Disponível em: <http://dx.doi.org/10.5294/pacla.2015.18.2.3>.
- BERELSON, B.; LAZARFELD, P.; MCPHEE, W. **Voting. A Study of Opinion Formation in a Presidential Campaign**. Chicago London Chicago University Press, 1954. Disponível em: <https://books.google.com.br/books?id=YOYfyQEACAAJ>.
- BERGAMO, M. Google associa lula à corrupção e bolsonaro a coraçã em busca sobre rei charles. **Folha de S.Paulo**, Maio 2023. Disponível em: <https://www1.folha.uol.com.br/colunas/monicabergamo/2023/05/google-associa-lula-a-corrupcao-e-bolsonaro-a-coracao-em-busca-sobre-rei-charles.shtml>.
- BEZERRA, H. D.; MUNDIM, P. S. Qual foi o papel das variáveis midiáticas na eleição presidencial de 2010? **Opinião Pública**, FapUNIFESP (SciELO), v. 17, n. 2, p. 452–476, nov. 2011. ISSN 0104-6276. Disponível em: <http://dx.doi.org/10.1590/S0104-62762011000200007>.
- BLASSNIG, S.; MITOVA, E.; PFIFFNER, N.; REISS, M. V. Googling referendum campaigns: Analyzing online search patterns regarding swiss direct-democratic votes. **Media and Communication**, Cogitatio, v. 11, n. 1, p. 19–30, jan. 2023. ISSN 2183-2439. Disponível em: <http://dx.doi.org/10.17645/mac.v11i1.6030>.
- BORRA, E.; WEBER, I. Political insights: Exploring partisanship in web search queries. **First Monday**, v. 17, 07 2012.
- BüLOW, M. V. **Relatório sobre Remuneração do Jornalismo pelas Plataformas Digitais**. 2023. Disponível em: https://cgi.br/media/docs/publicacoes/1/20230517100031/Estudo_Remuneracao_Jornalismo_pelas_Plataformas_Digitais.pdf.
- CAMERON, A. C.; TRIVEDI, P. K. **Microeconometrics: Methods and Applications**. [S. l.]: Cambridge University Press, 2005.

CAMPBELL, A.; CONVERSE, P. E.; MILLER, W. E.; STOKES, D. E. *et al.* The american voter. **New York: John Wiley**, 1960.

CARPINI, D. M. X.; KEETER, S. **What Americans Know about Politics and why it Matters**. [S. l.: s. n.], 1996.

CARREIRÃO, Y. A eleição presidencial de 2002: uma análise preliminar do processo e dos resultados eleitorais. **Revista de Sociologia e Política**, FapUNIFESP (SciELO), n. 22, p. 179–194, jun. 2004. ISSN 0104-4478. Disponível em: <http://dx.doi.org/10.1590/S0104-44782004000100013>.

CARREIRÃO, Y. d. S. Identificação ideológica e voto para presidente. **Opinião Pública**, FapUNIFESP (SciELO), v. 8, n. 1, p. 54–79, maio 2002. ISSN 0104-6276. Disponível em: <http://dx.doi.org/10.1590/S0104-62762002000100004>.

CASTELLS, M. **Communication Power**. OUP Oxford, 2013. ISBN 9780199681938. Disponível em: <https://books.google.com.br/books?id=3Q8oAAAAQBAJ>.

CASTRO. Determinantes do comportamento eleitoral: a centralidade da sofisticação política. Tese de Doutorado. IUPERJ, 1994.

CASTRO, D. **Audiência da Globo envelhece e fica mais pobre; jovens fogem para o streaming**. 2023. Disponível em: <https://noticiasdatv.uol.com.br/noticia/daniel-castro/audiencia-da-globo-envelhece-e-fica-mais-pobre-jovens-fogem-para-o-streaming-101631>.

CERVI, E. U. **Opinião pública e comportamento político**. [S. l.: s. n.], 2012.

CHAFFEY, D. **Search Engine Marketing Statistics 2024**. 2024. Disponível em: <https://www.smartinsights.com/search-engine-marketing/search-engine-statistics/>.

CHANG, T.-K.; LEE, J.-W. Factors affecting gatekeepers' selection of foreign news: A national survey of newspaper editors. **Journalism Quarterly**, SAGE Publications, v. 69, n. 3, p. 554–561, set. 1992. ISSN 0022-5533. Disponível em: <http://dx.doi.org/10.1177/107769909206900303>.

CLAYMAN, S. E.; REISNER, A. Gatekeeping in action: Editorial conferences and assessments of newsworthiness. **American Sociological Review**, SAGE Publications, v. 63, n. 2, p. 178, abr. 1998. ISSN 0003-1224. Disponível em: <http://dx.doi.org/10.2307/2657322>.

CONVERSE, P. E. The nature of belief systems in mass publics (1964). **Critical Review**, Informa UK Limited, v. 18, n. 1–3, p. 1–74, jan. 2006. ISSN 1933-8007. Disponível em: <http://dx.doi.org/10.1080/08913810608443650>.

COPPOCK, A. Avoiding post-treatment bias in audit experiments. **Journal of Experimental Political Science**, v. 6, p. 1–4, 2018.

DENARDIS, L. **Protocol Politics: The Globalization of Internet Governance**. MIT Press, 2009. ISBN 9780262258739. Disponível em: <https://books.google.com.br/books?id=BgVPAQAACAAJ>.

Department of Justice. **Justice Department Sues Google for Monopolizing Digital Advertising Technologies**. 2023. Accessed: 2024-05-24. Disponível em: <https://www.justice.gov/opa/pr/justice-department-sues-google-monopolizing-digital-advertising-technologies>.

DEVELOPERS, G. **Como criar conteúdo útil, confiável e que prioriza as pessoas**. 2024. Acessado em 14 de janeiro de 2024. Disponível em: <https://developers.google.com/search/docs/fundamentals/creating-helpful-content?hl=pt-br>.

DEVELOPERS, G. **O que é o Googlebot**. 2024. Disponível em: <https://web.archive.org/web/20240118025133/https://developers.google.com/search/docs/crawling-indexing/googlebot?hl=pt-br>.

DOMMETT, K.; POWER, S. Monitoring digital election campaigns: Assessing the transparency ecosystem in the united kingdom. **Politics**, v. 44, n. 1, p. 119–139, 2024. Disponível em: <https://doi.org/10.1177/02633957231156084>.

DOWNS, A. **An economic theory of democracy**. [S. l.]: Pearson, 1957.

DOWNS, A. **Teoria Econômica da Democracia, Uma**. [S. l.]: Edusp, 1999.

DRUCKMAN, J. N.; LEEPER, T. J. Learning more from political communication experiments: Pretreatment and its effects. **American Journal of Political Science**, v. 56, n. 4, p. 875–896, 2012. Disponível em: <https://EconPapers.repec.org/RePEc:wly:amposc:v:56:y:2012:i:4:p:875-896>.

DUBOIS, E.; BLANK, G. The echo chamber is overstated: the moderating effect of political interest and diverse media. **Information, Communication & Society**, Routledge, v. 21, n. 5, p. 729–745, 2018. Disponível em: <https://doi.org/10.1080/1369118X.2018.1428656>.

DUTTON, W. H.; REISDORF, B. C.; DUBOIS, E.; BLANK, G.; AHMAD, S.; ROBERTSON, C. **Search and Politics: and the United States**. 2017. Disponível em: <http://quello.msu.edu/research/the-part-played-by->.

EPSTEIN, R.; ROBERTSON, R. E. The search engine manipulation effect (seme) and its possible impact on the outcomes of elections. **Proceedings of the National Academy of Sciences of the United States of America**, National Academy of Sciences, v. 112, p. E4512–E4521, 8 2015. ISSN 10916490.

EPSTEIN, R.; ROBERTSON, R. E.; LAZER, D.; WILSON, C. Suppressing the search engine manipulation effect (seme). **Proceedings of the ACM on Human-Computer Interaction**, v. 1, 2017. ISSN 25730142.

EVANGELISTA, R. **Beyond machines of loving grace: Hacker culture, cybernetics and democracy**. Edições Sesc SP, 2018. ISBN 9788594931023. Disponível em: <https://books.google.com.br/books?id=kA9mDwAAQBAJ>.

FEENBERG, A. **Tecnologia, modernidade e democracia**. Inovatec Press, 2018. Disponível em: https://books.google.com.br/books?id=o_m3DwAAQBAJ.

FIGUEIREDO, M. O voto obrigatório (comportamento do eleitor brasileiro). **Cem anos de eleições presidenciais, Textos IDESP**, n. 36, 1990.

FIGUEIREDO, M. F. **A decisão do voto - democracia e racionalidade**. [S. l.]: Editora Ufmg, 2008.

FORBES, R. **Google é a marca mais influente do Brasil, diz estudo; Samsung e YouTube completam top 3**. 2022. Disponível em: <https://forbes.com.br/forbes-tech/2022/04/google-e-a-marca-mais-influente-do-brasil-diz-estudo-samsung-e-youtube/>.

FOURNIER, P.; BLAIS, A.; NADEAU, R.; GIDENGIL, E.; NEVITTE, N. Issue importance and performance voting. **Political Behavior**, Springer, v. 25, n. 1, p. 51–67, 2003. ISSN 01909320, 15736687. Disponível em: <http://www.jstor.org/stable/3657314>.

GEORGIEV, D.; CHRISTIE, R.; TORKAMANI, M.; SONG, R.; LIMOUSIN, P.; JAHANSHAH, M. Development and validation of a daily habit scale. **Frontiers in Neuroscience**, v. 16, 2022.

GOODWIN, D. 7 must-see google search ranking documents in antitrust trial exhibits. **Search Engine Land**, 2023. Acessado em 14 de janeiro de 2024. Disponível em: <https://searchengineland.com/google-search-ranking-documents-434141>.

GOOGLE. **Pesquisas do ano no Google – Google Trends**. 2020. <https://trends.google.com.br/trends/yis/2020/BR/>. Acessado em 17 de fevereiro de 2024.

GOOGLE. **How Search Works**. 2024. https://www.google.com/intl/en_uk/search/howsearchworks/how-search-works/. Acessado em 14 de janeiro de 2024.

Google Developers. **Guia detalhado sobre como a Pesquisa Google funciona**. 2024. Acessado em 14 de janeiro de 2024. Disponível em: <https://developers.google.com/search/docs/fundamentals/how-search-works?hl=pt-br>.

GRAHAM, R. **The Digital Practices of Mobile Phones**. 2012.

GRAMACHO, W.; TURGEON, M.; KENNEDY, J.; STABILE, M.; MUNDIM, P. S. Political preferences, knowledge, and misinformation about covid-19: The case of brazil. **Frontiers in Political Science**, Frontiers, v. 3, p. 36, 2021.

GRAMACHO, W. G.; TURGEON, M. When politics collides with public health: Covid-19 vaccine country of origin and vaccination acceptance in brazil. **Vaccine**, v. 39, n. 19, p. 2608–2612, 2021. ISSN 0264-410X. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0264410X21003960>.

HANNAK, A.; SAPIEZYŃSKI, P.; KAKHKI, A. M.; KRISHNAMURTHY, B.; LAZER, D.; MISLOVE, A.; WILSON, C. Measuring personalization of web search. **WWW 2013 - Proceedings of the 22nd International Conference on World Wide Web**, p. 527–537, 2013.

HARB, M. d. P. d. A. A.; SILVA, L. V. e.; VIJAYKUMAR, N. L.; SILVA, M. S. d.; FRANCÊS, C. R. L. An analysis of the deleterious impact of the infodemic during the covid-19 pandemic in brazil: A case study considering possible correlations with socioeconomic aspects of brazilian demography. **International Journal of Environmental Research and Public Health**, MDPI AG, v. 19, n. 6, p. 3208, mar. 2022. ISSN 1660-4601. Disponível em: <http://dx.doi.org/10.3390/ijerph19063208>.

HINDMAN, M. **The Myth of Digital Democracy**. Princeton University Press, 2009. ISBN 9781400837496. Disponível em: <https://books.google.com.br/books?id=iaq1Wd7PWR4C>.

HUCKFELDT, R. R.; SPRAGUE, J.; CHONG, D.; KUKLINSKI, J. H.; FELDMAN, S.; WYER, R. S. **Citizens, Politics and Social Communication: Information and Influence in an Election Campaign**. Cambridge University Press, 1995. ISBN 9780521452984. Disponível em: https://books.google.com.br/books?id=xdAB_0-c0ZcC.

IBGE. **Internet já é acessível em 90,0% dos domicílios do país em 2021**. 2021. Disponível em: <https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/34954-internet-ja-e-acessivel-em-90-0-dos-domicilios-do-pais-em-2021>.

JOURNALISM, R. I. for the Study of. **Digital News Report 2022: Brazil**. Oxford: University of Oxford, 2022. Online. Disponível em: <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022/brazil>.

KAYE, B. K.; JOHNSON, T. J. Online and in the know: Uses and gratifications of the web for political information. **Journal of Broadcasting & Electronic Media**, Routledge, v. 46, n. 1, p. 54–71, 2002. Disponível em: https://doi.org/10.1207/s15506878jobem4601_4.

KEANE, B. M. T.; BRIEN, M. O. Are people biased. **Communications of the ACM**, v. 51, p. 49–52, 2008.

KEMP, S. **Digital 2022: The World's Top Websites**. 2022. Disponível em: <https://datareportal.com/reports/digital-2022-worlds-top-websites>.

KLIMAN-SILVER, C.; HANNAK, A.; LAZER, D.; WILSON, C.; MISLOVE, A. Location, location, location: The impact of geolocation on web search personalization. **Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC**, v. 2015-Octob, p. 121–127, 2015.

KNOBLOCH-WESTERWICK, S.; JOHNSON, B. K.; SILVER, N. A.; WESTERWICK, A. Science exemplars in the eye of the beholder: How exposure to online science information affects attitudes. **Science Communication**, v. 37, n. 5, p. 575–601, 2015. Disponível em: <https://doi.org/10.1177/1075547015596367>.

KULSHRESTHA, J.; ESLAMI, M.; MESSIAS, J.; ZAFAR, M. B.; GHOSH, S.; GUMMADI, K. P.; KARAHALIOS, K. Quantifying search bias: Investigating sources of bias for political searches in social media. **Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW**, p. 417–432, 2017.

LAZARSELD, P.; BERELSON, B.; GAUDET, H. **The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign**. Columbia University Press, 1948. Disponível em: <https://books.google.com.br/books?id=S-lnIFR02FIC>.

LESSIG, L. **Code: And Other Laws Of Cyberspace**. London, England: Basic Books, 1999.

LEWIS-BECK, M. S.; NADEAU, R.; FOUCAULT, M. The compleat economic voter: New evidence from the uk. **SSRN Electronic Journal**, Elsevier BV, 2011. ISSN 1556-5068. Disponível em: <http://dx.doi.org/10.2139/ssrn.1904682>.

LICIO, E. C.; RENNÓ, L. R.; CASTRO, H. C. d. O. Bolsa família e voto na eleição presidencial de 2006: em busca do elo perdido. **Opinião Pública**, v. 15, n. 1, p. 31–54, out. 2015. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/op/article/view/8641319>.

LODGE, M.; TABER, C. S. **The Rationalizing Voter**. Cambridge University Press, 2013. ISBN 9780521763509. Disponível em: <https://books.google.com.br/books?id=tqI5GxWmIU4C>.

LOVELUCK, B. **Réseaux, libertés et contrôle: Une généalogie politique d'internet**. Armand Colin, 2015. ISBN 9782200612955. Disponível em: <https://books.google.com.br/books?id=XUy5CgAAQBAJ>.

Lê, S.; JOSSE, J.; HUSSON, F. Factominer: An r package for multivariate analysis. **Journal of Statistical Software**, v. 25, n. 1, p. 1–18, 2008. Disponível em: <https://www.jstatsoft.org/index.php/jss/article/view/v025i01>.

MAINWARING, S. **Sistemas partidários em novas democracias: o caso do Brasil**. Mercado Aberto, 2001. ISBN 9788528005349. Disponível em: <https://books.google.com.br/books?id=v2csAAAACAAJ>.

MANGOOLS. **KWfinder: Free Keyword Research & Analysis Tool**. 2024. <https://mangools.com/kwfinder/>. Acessado em 17 de fevereiro de 2024.

MARQUES, F. P. J.; FERRACIOLI, P.; COMEL, N.; KNISS, A. B. Who is who in fact-checked conspiracy theories? disseminators, sources, and the struggle for authority in polarized environments. **Journalism**, SAGE Publications, p. 146488492311655, mar. 2023. ISSN 1741-3001. Disponível em: <http://dx.doi.org/10.1177/14648849231165579>.

MCCOMBS, M. E. (Ed.). **The agenda-setting function of the press**. In: **OVERHOLSER, G., JAMIESON, K. H. (Eds.). The Press**. New York, NY: Oxford University Press, 2005.

MEHL, J. P.; SILVA, S. P. da Pereira da. **Cultura Digital, internet e apropriações políticas: Experiências, desafios e horizontes**. [S. l.: s. n.], 2017. ISBN 9788554730024.

MOZ. **The Beginner's Guide to SEO**. 2024. Acessado em 14 de janeiro de 2024. Disponível em: <https://moz.com/beginners-guide-to-seo>.

MOZ. **URLs**. 2024. Acessado em 14 de janeiro de 2024. Disponível em: <https://moz.com/learn/seo/url>.

MUDDIMAN, A. Searching for the next u.s. president: Differences in search engine results for the 2008 u.s. presidential candidates. **Journal of Information Technology & Politics**, Routledge, v. 10, n. 2, p. 138–157, 2013. Disponível em: <https://doi.org/10.1080/19331681.2012.707440>.

MUNDIM, P. S.; GRAMACHO, W.; TURGEON, M.; STABILE, M. Viés noticioso e exposição seletiva nos telejornais brasileiros durante a pandemia de covid-19. **Opinião Pública**, SciELO Brasil, v. 28, p. 615–634, 2023.

NECHUSHTAI, E.; LEWIS, S. C. What kind of news gatekeepers do we want machines to be? filter bubbles, fragmentation, and the normative dimensions of algorithmic recommendations. **Computers in Human Behavior**, v. 90, p. 298–307, 2019. ISSN 0747-5632. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0747563218303650>.

NEMOTO, Y.; KLYUEV, V. Tool to retrieve less-filtered information from the internet. **Information**, MDPI AG, v. 12, n. 2, p. 65, fev. 2021. ISSN 2078-2489. Disponível em: <http://dx.doi.org/10.3390/info12020065>.

NETLAB. **A guerra das plataformas contra o PL 2630**. 2023. Disponível em: <https://netlab.eco.ufrj.br/post/a-guerra-das-plataformas-contr-o-pl-2630>.

O'NEIL, C. **Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy**. Crown, 2016. ISBN 9780553418828. Disponível em: <https://books.google.com.br/books?id=NgEwCwAAQBAJ>.

PAN, B.; HEMBROOKE, H.; JOACHIMS, T.; LORIGO, L.; GAY, G.; GRANKA, L. In google we trust: Users' decisions on rank, position, and relevance. 2007. Disponível em: <https://academic.oup.com/jcmc/article-abstract/12/3/801/4582975>.

PAPADA, E.; ANGIOLILLO, F.; GASTALDI, L.; KÖHLER, T.; LUNDSTEDT, M.; NATSIKA, N.; NORD, M.; SATO, Y.; WIEBRECHT, F.; LINDBERG, S. I.; BURMEISTER, S.; GRAHN, S.; LIETHMANN, M.; PERNES, J.; RYDÉN, O.; VERKHOVTSEVA, M.; VON, J.; EDITORES, R.; ALTMAN, D.; HUERTAS-HERNÁNDEZ, S. **REPORTE DE LA DEMOCRACIA 2023**. 2023. Disponível em: www.v-dem.net.

PARISER, E. **The Filter Bubble: What The Internet Is Hiding From You**. Penguin Books Limited, 2011. ISBN 9780141969923. Disponível em: <https://books.google.com.br/books?id=-FWO0puw3nYC>.

PEIXOTO, V.; RENNÓ, L. Mobilidade social ascendente e voto: as eleições presidenciais de 2010 no brasil. **Opinião Pública**, FapUNIFESP (SciELO), v. 17, n. 2, p. 304–332, nov. 2011. ISSN 0104-6276. Disponível em: <http://dx.doi.org/10.1590/S0104-62762011000200002>.

PODER360. Google sugere "lula corrupção" ao pesquisar "lula coroação". **Poder360**, Maio 2023. Disponível em: <https://www.poder360.com.br/brasil/google-sugere-lula-corrupcao-ao-pesquisar-lula-coroacao/>.

PRIOR. **Cambridge studies in public opinion and political psychology: Post-broadcast democracy: How media choice increases inequality in political involvement and polarizes elections**. Cambridge, England: Cambridge University Press, 2007.

RENNÓ, L. R. Information and voting: Microfoundations of accountability in complex electoral environments. 2004. Disponível em: <http://d-scholarship.pitt.edu/8927/>.

ROBERTS, A. **Blacked out: Government Secrecy in the Information Age**. Cambridge, England: Cambridge University Press, 2006.

ROBERTSON, R. E.; JIANG, S.; JOSEPH, K.; FRIEDLAND, L.; LAZER, D.; WILSON, C. Auditing partisan audience bias within google search. **Proceedings of the ACM on Human-Computer Interaction**, v. 2, 2018. ISSN 25730142. Disponível em: <https://personalization.ccs.neu.edu/static/pdf/robertson-2018-pacmhci.pdf>.

ROESNER, F.; KOHNO, T.; WETHERALL, D. Detecting and defending against third-party tracking on the web. In: . {USENIX} Association, 2012. p. 155–168. ISBN 978-931971-92-8. Disponível em: <https://www.usenix.org/conference/nsdi12/technical-sessions/presentation/roesner>.

SILVA, T. **Racismo algorítmico: inteligência artificial e discriminação nas redes digitais**. Edições Sesc SP, 2022. ISBN 9786586111705. Disponível em: <https://books.google.com.br/books?id=BCZcEAAAQBAJ>.

SILVA, T.; EMICIDA; BENJAMIN, R.; TRINDADE, L.; ROSHANI, N.; BROCK, A.; GOMES, L.; ARAUJO, R.; JÚNIOR, J.; CARRERA, F.; BIRHANE, A.; KATEMBERA, S.; NEVES, T.; ALAMU, F.; AWORINDE, H.; ISHARUFE, W.; AVLE, S.; DARITY, W. **Comunidades, Algoritmos e Ativismos Digitais: olhares afrodiáspóricos**. [S. l.: s. n.], 2020. ISBN 978-65-86113-01-3.

SLOTHUUS, R. Assessing the influence of political parties on public opinion: the challenge from pretreatment effects. **Political Communication**, v. 33, p. 302–327, 2015.

SRINIVASAN, R.; FISH, A. **After the Internet**. Wiley, 2017. ISBN 9781509506187. Disponível em: <https://books.google.com.br/books?id=PBjZjwEACAAJ>.

STABILE, M.; ARNS, A.; MOURA, T.; GOMES, A. “e o google?” técnicas de coleta dos resultados de pesquisa do google. In: . [S. l.: s. n.], 2019. p. 1–15.

STABILE, M. V. B. M. **O velho não morreu, o novo já está aqui: informação e participação digital na era do bolsonarismo**. [S. l.]: Autêntica Editora, 2021. 481-501 p.

STEINER, M.; MAGIN, M.; STARK, B.; GEIß, S. Seek and you shall find? a content analysis on the diversity of five search engines’ results on political queries. **Information, Communication & Society**, Routledge, v. 25, n. 2, p. 217–241, 2022. Disponível em: <https://doi.org/10.1080/1369118X.2020.1776367>.

TABER, C. S.; LODGE, M. Motivated skepticism in the evaluation of political beliefs. **American Journal of Political Science**, [Midwest Political Science Association, Wiley], v. 50, n. 3, p. 755–769, 2006. ISSN 00925853, 15405907. Disponível em: <http://www.jstor.org/stable/3694247>.

TECNOBLOG. **Google Altera Busca, Lança Portal Coronavirus**. 2020. <https://tecnoblog.net/noticias/2020/03/25/google-altera-busca-lanca-portal-coronavirus/>. Acessado em: data de acesso aqui.

TELLES, H. D. S.; LOURENÇO, L. C.; STORNI, T. P. L. Partidos, campanhas e voto: como o eleitor decide nas municipais. **Sociedade e Cultura**, v. 12, n. 1, p. 91–116, ago. 2009. Disponível em: <https://revistas.ufg.br/fcs/article/view/6903>.

TERRA. Após polêmica, google altera significados de "mulher solteira" e "patroa". **Terra**, 2020. Disponível em: <https://www.terra.com.br/diversao/gente/apos-polemica-google-altera-significados-de-mulher-solteira-e-patroa,d6fcf3975fd0785dbb24911bd0a59d61srxoahen.html>.

TRIELLI, D.; DIAKOPOULOS, N. Search as news curator: The role of google in shaping attention to news information. In: **Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems**. New York, NY, USA: Association for Computing Machinery, 2019. (CHI '19), p. 1–15. ISBN 9781450359702. Disponível em: <https://doi.org/10.1145/3290605.3300683>.

TRIELLI, D.; DIAKOPOULOS, N. Partisan search behavior and google results in the 2018 u.s. midterm elections. **Information, Communication and Society**, Routledge, v. 25, n. 1, p. 145–161, 2022. Disponível em: <https://doi.org/10.1080/1369118X.2020.1764605>.

TUFEKCI, Z. Engineering the public: Big data, surveillance and computational politics. **First Monday**, v. 9, 2014. Disponível em: <https://journals.uic.edu/ojs/index.php/fm/article/view/4901/4097>.

TURGEON, M.; RENNÓ, L. Informação política e atitudes sobre gastos governamentais e impostos no brasil: Evidências a partir de um experimento de opinião pública. **Opinio Publica**, v. 16, p. 143–159, 2010. ISSN 01046276.

- ULLOA, R.; MAKHORTYKH, M.; URMAN, A.; KULSHRESTHA, J. **Novelty in news search: a longitudinal study of the 2020 US elections**. arXiv, 2022. Disponível em: <https://arxiv.org/abs/2211.04746>.
- UNKEL, J.; HAIM, M. Googling politics: Parties, sources, and issue ownerships on google in the 2017 german federal election campaign. **Social Science Computer Review**, SAGE Publications Sage CA: Los Angeles, CA, v. 39, n. 5, p. 844–861, 2021.
- URMAN, A.; MAKHORTYKH, M.; ULLOA, R. The matter of chance: Auditing web search results related to the 2020 u.s. presidential primary elections across six search engines. **Social Science Computer Review**, SAGE Publications, v. 40, n. 5, p. 1323–1339, abr. 2021. ISSN 1552-8286. Disponível em: <http://dx.doi.org/10.1177/08944393211006863>.
- VALENTE, A. F.; FRANCO, B. “dÁ um ‘google’ aÍ”: buscando a relevância do google nas eleições brasileiras de 2018. In: . [S. l.: s. n.], 2018. p. 1–18.
- VILARINS L.; STABILE, M. v. B. M.; MOURA T.; ARNS, A. G. A.; FONTENELLE, A. F. B. **GOOGLE E AS ELEIÇÕES BRASILEIRAS DE 2018**. 2021. 341 p.
- VOS, T. P. Journalists as gatekeepers. In: _____. **The Handbook of Journalism Studies**. Routledge, 2019. p. 90–104. ISBN 9781315167497. Disponível em: <http://dx.doi.org/10.4324/9781315167497-6>.
- WHITE, R. Beliefs and biases in web search. In: **Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval**. New York, NY, USA: Association for Computing Machinery, 2013. (SIGIR '13), p. 3–12. ISBN 9781450320344. Disponível em: <https://doi.org/10.1145/2484028.2484053>.
- WHYTE, C. E. Thinking inside the (black) box: Agenda setting, information seeking, and the marketplace of ideas in the 2012 presidential election. **New Media & Society**, v. 18, n. 8, p. 1680–1697, 2016. Disponível em: <https://doi.org/10.1177/1461444814567985>.
- WOOLDRIDGE, J. M. **Econometric Analysis of Cross Section and Panel Data**. [S. l.]: The MIT Press, 2010.
- ZADE, H.; WACK, M.; ZHANG, Y.; STARBIRD, K.; CALO, R.; YOUNG, J.; WEST, J. D. Auditing google’s search headlines as a potential gateway to misleading content: Evidence from the 2020 us election. **Journal of Online Trust and Safety**, Stanford Internet Observatory, v. 1, n. 4, set. 2022. ISSN 2770-3142. Disponível em: <http://dx.doi.org/10.54501/jots.v1i4.72>.
- ZUBOFF, S. Big other: Surveillance capitalism and the prospects of an information civilization. **Journal of Information Technology**, v. 30, n. 1, p. 75–89, 2015. Disponível em: <https://doi.org/10.1057/jit.2015.5>.

10 APÊNDICE METODOLÓGICO

10.1 Declaração de Uso de IA generativa e tecnologias assistidas por IA

Durante a preparação desta tese, foi utilizada a tecnologia **ChatGPT 4.0 da OpenAI**, a partir de dezembro de 2023, após o processo de coleta de dados e análises preliminares. O uso dessa ferramenta teve múltiplas finalidades:

1. **Programação dos gráficos em R:** auxílio na programação de gráficos em R, visando a uma representação visual mais eficaz e precisa dos dados analisados. A ferramenta proporcionou sugestões de código e ajudou na otimização de gráficos existentes, melhorando a interpretação e a apresentação dos resultados.
2. **Ajuda no processo de análise:** auxílio nas análises dos dados, oferecendo um protocolo a ser seguido sobre métodos estatísticos e técnicas de análise de dados.
3. **Apoio analítico dos modelos multivariados:** a ferramenta foi empregada na leitura inicial das tabelas analíticas dos modelos multivariados. O recurso ChatGPT auxiliou na interpretação dos coeficientes, apontando de modo sistemático as variáveis relevantes.
4. **Leitura de algumas tabelas:** auxiliou a descrição inicial de algumas tabelas, de modo que o texto se tornasse mais didático.
5. **Processo de classificação das buscas políticas:** auxiliou a classificação das buscas políticas, conforme detalhado no apêndice.
6. **Desenvolvimento de assistente específico:** com o auxílio da ferramenta ChatGPT, foi criado um assistente denominado “Assistente do Max”, que foi crucial para a categorização e análise das buscas políticas de maneira eficiente. O assistente pode ser acessado [aqui] (<https://chat.openai.com/g/g-0hjBcbQNn-assistente-do-max>).

Após a utilização dessa ferramenta/serviço, o autor revisou e editou o conteúdo conforme necessário e assume total responsabilidade pela publicação. O uso de uma IA generativa visou a enriquecer a qualidade analítica e à clareza expositiva do trabalho, contribuindo para uma apresentação mais robusta e fundamentada das análises realizadas.

10.1.1 *Prompt de correção ortográfica e classificação dos termos*

Prompt 1: Chat, você atuará como um corretor ortográfico da língua portuguesa. Preciso que você faça ajustes ortográficos nestes termos quando houver necessidade. Corrija escrita incorreta, falta de acentuações espaços ou itens incorretos para compreensão humana:

(TERMO)

Sua resposta será apenas o termo corrigido. Não precisa colocar pontuação ou aspas.

Prompt 2: Você é um cientista político que fará classificações dos termos de buscas realizados por diversos eleitores em 2020. Você irá classificar conforme essas 4 categorias: “1-Informações sobre Candidatos/Partidos”: Buscas focadas em informações específicas sobre candidatos ou partidos políticos, como políticas propostas, histórico de votação e biografias. Geralmente nome dos candidatos ou buscas específicas sobre candidatos ou termos gerais sobre candidatos na cidade. “2-Processos/Procedimentos Eleitorais”: Inclui buscas relacionadas a como e onde votar, datas de eleições e requisitos de elegibilidade, certidões, buscas por tribunais, agendamentos. “3-Buscas gerais sobre eleições”: abrange buscas sobre questões políticas amplas, apurações, resultados ou buscas generalistas sobre as eleições como eleições 2020 ou debates na televisão. “4-Não é política”: são os termos que não parecem ser buscas sobre eleições políticas. Você irá responder apenas uma categoria para cada texto e apenas a categoria. A busca é: (TERMO)

10.2 Distribuição Amostral e Tratamento das buscas

Tabela 44 – Resultados para a pergunta: Critério Brasil de Classificação Econômica

Classe	Freq.	Prop.
D-E	784	28,3%
C2	741	26,7%
C1	609	22,0%
B2	460	16,6%
B1	106	3,8%
A	69	2,5%
Não respondeu	2	0,1%
Total	2771	100%

Fonte: elaborado pelo autor, com base em Onda 1 do painel, set. 2020.

Tabela 45 – Resultados para a pergunta: Até que série você estudou?

Escolaridade	Freq.	Prop.
01. Fundamental	529	19,1%
02. Médio	1636	59,0%
03. Superior	606	21,9%
Total	2771	100%

Fonte: elaborado pelo autor, com base em Onda 1 do painel, set. 2020.

Tabela 46 – Resultados para a pergunta: Você é...

Gênero	Freq.	Prop.
Homem	1293	46,7%
Mulher	1478	53,3%
Total	2771	100%

Fonte: elaborado pelo autor, com base em Onda 1 do painel, set. 2020.

Tabela 47 – Resultados para a pergunta: E quantos anos você tem?

Faixa Etária	Freq.	Prop.
16 a 24 anos	422	15,2%
25 a 34 anos	586	21,1%
35 a 44 anos	577	20,8%
45 a 54 anos	472	17,0%
55 anos ou mais	714	25,8%
Total	2771	100%

Fonte: elaborado pelo autor, com base em Onda 1 do painel, set. 2020.

Tabela 48 – Resultados para a pergunta: A sua raça ou cor é:

Cor/Raça	Freq.	Prop.
Branca	1276	46,0%
Parda	1108	40,0%
Preta	302	10,9%
Amarela	70	2,5%
Indígena	15	0,5%
Total	2771	100%

Fonte: elaborado pelo autor, com base em Onda 1 do painel, set. 2020.

Tabela 49 – Resultados para a pergunta: Qual a sua religião, se tiver?

Religião	Freq.	Prop.
Católica	1037	37,4%
Evangélico	420	15,2%
Pentecostal		
Evangélico	156	5,6%
Neo-Pentecostal		
Não tem religião/Agnóstico	571	20,6%
Sou ateu/Não acredito em Deus	73	2,6%
Outros	421	15,2%
Não respondeu	93	3,4%
Total	2771	100%

Fonte: elaborado pelo autor, com base em Onda 1 do painel, set. 2020.

Tabela 50 – Região do entrevistado

Região	Freq.	Prop.
N	215	7,8%
NE	742	26,8%
SE	1209	43,6%
S	404	14,6%
CO	201	7,3%
Total	2771	100%

Fonte: elaborado pelo autor, com base em Onda 1 do painel, set. 2020.

10.3 Descritivo das buscas realizadas

Tabela 51 – Quantidade de buscas realizadas por tipo de abordagem

	Freq.	Prop.
1. Eleitorais: Sobre candidatos	26.874	5.2%
2. Eleitorais: Termos sobre eleições	347.446	67%
3. Não eleitorais: Termos genéricos	27.582	5.3%
4. Não eleitorais: Termos oriundos da população survey	116.759	22.5%

Fonte: elaborado pelo autor.

Tabela 52 – Quantidade de buscas não eleitorais realizadas

Abordagem	Tipo de Busca	Qtd de Buscas	Qtd de Municípios	Qtd de Termos
Eleitorais - Capitais	Sobre Candidatos	26.874	26	589
Eleitorais - Geral	Sobre corrupção	6.814	96	3
Eleitorais - Geral	Perguntas gerais	15.959	286	18
Eleitorais - Geral	Sobre polarização	9.635	96	7
Eleitorais - Geral	Termos eleitorais	315.038	5295	5
Não eleitorais	Survey	116.759	570	1509
Não eleitorais	Termos gerais	27.582	570	69

Fonte: elaborado pelo autor.

10.3.1 Busca sobre candidatos

10.3.2 Busca Gerais

Tabela 53 – Descritivo dos municípios incluídos na captura

Tipos de Municípios	Quantidade	População (em milhões)	População (em % - acum)	Amostra
Faixa 1 - Capitais	26	47.1	22.8%	26
Faixa 2 - Municípios com 2º turno	70	34.3	39.3%	70
Faixa 3 - Mais de 110 mil habitantes	190	32.4	55.0%	190
Faixa 4 - Mais de 45 mil habitantes	455	31.1	70.0%	95
Faixa 5 - Mais de 20 mil habitantes	1032	30.3	84.6%	93
Faixa 6 - Até 20 mil habitantes	3796	32.0	100.0%	98

Fonte: elaborado pelo autor.

Tabela 54 – Termos utilizados nas buscas políticas

Termos de busca	Frequência	Prop.
Candidatos a prefeito em (CIDADE)	263	26%
Candidatos em (CIDADE)	145	15%
Eleições 2020	84	8%
Nome do Candidato	68	7%
Histórico/ Biografia/ Currículo (CANDIDATO)	62	6%
Propostas/ Plano de governo (CANDIDATOS)	45	5%
Ficha limpa (CANDIDATOS)	43	4%
Candidato a Prefeitura de (CIDADE)	34	3%
Candidato a Prefeito	32	3%
Nenhum/ Não sei/ Não tem interesse	38	4%
Outros	467	47%

Fonte: elaborado pelo autor.

Tabela 55 – Quantidade de buscas eleitorais realizadas

Tipo de busca	Termo Buscado	Qtd de Buscas	Qtd de Municípios
Sobre corrupção	candidato ficha-limpa OR candidato fichalimpa	2.273	96
Sobre corrupção	candidato denunciado	2.275	96
Sobre corrupção	fraude nas urnas	2.266	96
Perguntas gerais	Como votar?	233	26
Perguntas gerais	Por que devemos votar	230	26
Perguntas gerais	Qual é o candidato com melhores propostas para a educação?	2.617	284
Perguntas gerais	Qual é o melhor candidato para a saúde?	2.597	283
Perguntas gerais	Qual é o melhor candidato para fazer ter mais empregos?	2.611	284
Perguntas gerais	Qual é o melhor candidato para fazer ter mais segurança?	2.606	285
Perguntas gerais	Qual é o melhor candidato?	2.610	286
Perguntas gerais	Quem está na frente nas eleições	233	26
Perguntas gerais	candidato ficha-limpa OR candidato fichalimpa	337	286
Perguntas gerais	candidato antipt	337	285
Perguntas gerais	candidato denunciado	334	282
Perguntas gerais	candidato do bolsonaro	336	284
Perguntas gerais	candidato do lula	333	281
Perguntas gerais	candidato do psdb	52	26
Perguntas gerais	candidato do psol	51	26
Perguntas gerais	candidato do psol	52	26
Perguntas gerais	candidato do pt	338	286
Perguntas gerais	fraude nas urnas	52	26
Sobre polarização	candidato antipt	2.274	96
Sobre polarização	candidato do bolsonaro	2.278	96
Sobre polarização	candidato do lula	2.263	96
Sobre polarização	candidato do psdb	180	26
Sobre polarização	candidato do psol	182	26
Sobre polarização	candidato do psol	180	26
Sobre polarização	candidato do pt	2.278	96
Termos eleitorais	candidatas a prefeita	8.026	286
Termos eleitorais	candidatas em	8.021	286
Termos eleitorais	candidatos a prefeito em (CIDADE)	77.048	5295
Termos eleitorais	candidatos em (CIDADE)	74.480	5295
Termos eleitorais	Eleições 2020	74.426	5295
Termos eleitorais	candidatos a prefeito em (CIDADE)	16.063	570
Termos eleitorais	candidatos em (CIDADE)	16.047	570
Termos eleitorais	Eleições 2020	15.999	570
Termos eleitorais	candidatas a prefeita	12.468	286
Termos eleitorais	candidatas em	12.460	286

Fonte: elaborado pelo autor.

10.3.3 Busca não eleitorais

Tabela 56 – Quantidade de buscas não eleitorais realizadas

Tipo de Busca	Sub-tipo de busca	Qtd de Buscas	Qtd de Muni- cipios	Qtd de Termos
Não eleitorais - Survey	survey	116.759	570	1509
Não eleitorais - Termos gerais	seguranca	909	452	2
Não eleitorais - Termos gerais	vestibular	909	452	2
Não eleitorais - Termos gerais	viagem	907	456	2
Não eleitorais - Termos gerais	emprego	904	456	2
Não eleitorais - Termos gerais	transito	903	444	2
Não eleitorais - Termos gerais	programacao_globo	902	464	2
Não eleitorais - Termos gerais	novela	901	452	2
Não eleitorais - Termos gerais	comprar	894	466	2
Não eleitorais - Termos gerais	covid	5.036	570	16
Não eleitorais - Termos gerais	jornais	451	315	1
Não eleitorais - Termos gerais	filme	449	309	1
Não eleitorais - Termos gerais	redes_sociais	449	302	1
Não eleitorais - Termos gerais	tematica	449	305	1
Não eleitorais - Termos gerais	saude	2.698	566	9
Não eleitorais - Termos gerais	esporte	2.257	560	5
Não eleitorais - Termos gerais	cantores	2.250	562	5
Não eleitorais - Termos gerais	educacao	1.806	550	4
Não eleitorais - Termos gerais	entretenimento	1.801	548	4
Não eleitorais - Termos gerais	celular	1.357	520	3
Não eleitorais - Termos gerais	emagrecimento	1.350	513	3

Fonte: elaborado pelo autor.

10.4 Modelos do Capítulo 5

Tabela 57 – VIF - Modelo Declarado

	GVIF	Df	GVIF ^{1/(2*Df)}
i_classe_AB	1.352785	1	1.163093
i_branços	1.225113	1	1.106848
i_catolica	1.058041	1	1.028611
i_region	1.269136	4	1.030240
i_edu	1.275153	2	1.062651
i_mulheres	1.034409	1	1.017059
i_age	1.213455	1	1.101569
i_persona_fact	1.144037	4	1.016963
i_partypref	1.115989	1	1.056404
i_partynev	1.145745	1	1.070395

Fonte: elaboração do autor

Tabela 58 – VIF - Modelo Monitorado

	GVIF	Df	GVIF ^{1/(2*Df)}
i_searches	1.125367	1	1.060833
i_classe_AB	1.552110	1	1.245837
i_branços	1.219369	1	1.104250
i_catolica	1.073187	1	1.035948
i_region	1.296416	4	1.032983
i_edu	1.283374	2	1.064359
i_mulheres	1.079764	1	1.039117
i_age	1.462006	1	1.209135
i_persona_fact	1.287656	4	1.032108
i_partypref	1.135198	1	1.065457
i_partynev	1.170795	1	1.082033

Fonte: elaboração do autor

Hosmer and Lemeshow goodness of fit (GOF) test

data: modelo_declarado_logit\$y, fitted(modelo_declarado_logit) X-squared = 6.9648,
df = 8, p-value = 0.5404

Hosmer and Lemeshow goodness of fit (GOF) test

data: modelo_monitorado_logit\$y, fitted(modelo_monitorado_logit) X-squared =
13.85, df = 8, p-value = 0.08576

Tabela 59 – Model Statistics for Model declarado

Statistic	Value
Log-Lik Intercept Only	-1343.992
Log-Lik Full Model	-1225.918
Deviance (1944)	2451.836
LR (17)	236.148
Prob > LR	0.000
McFadden's R^2	0.088
McFadden's Adj R^2	0.074
ML (Cox-Snell) R^2	0.110
Cragg-Uhler (Nagelkerke) R^2	0.150
McKelvey & Zavoina's R^2	0.113
Efron's R^2	0.080
Count R^2	0.638
Adj Count R^2	0.062
BIC	2588.307
AIC	2487.836

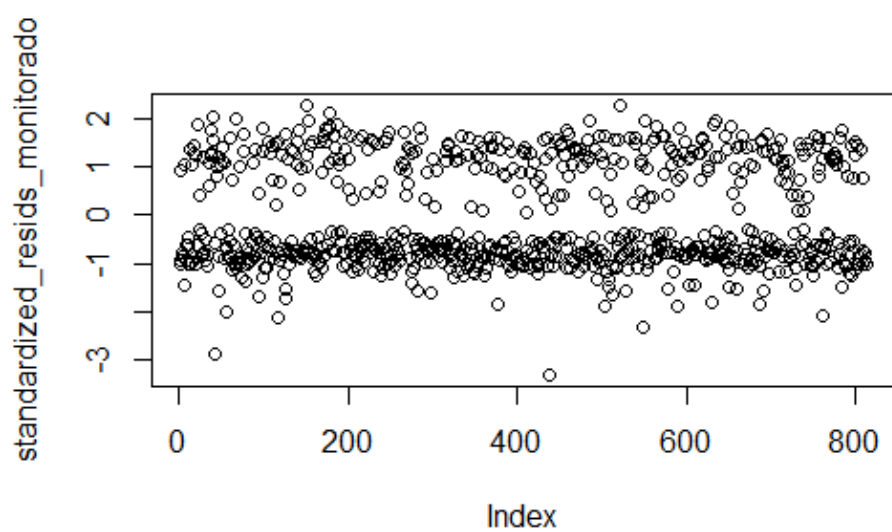
Fonte: elaboração do autor.

Tabela 60 – Model Statistics for Model monitorado

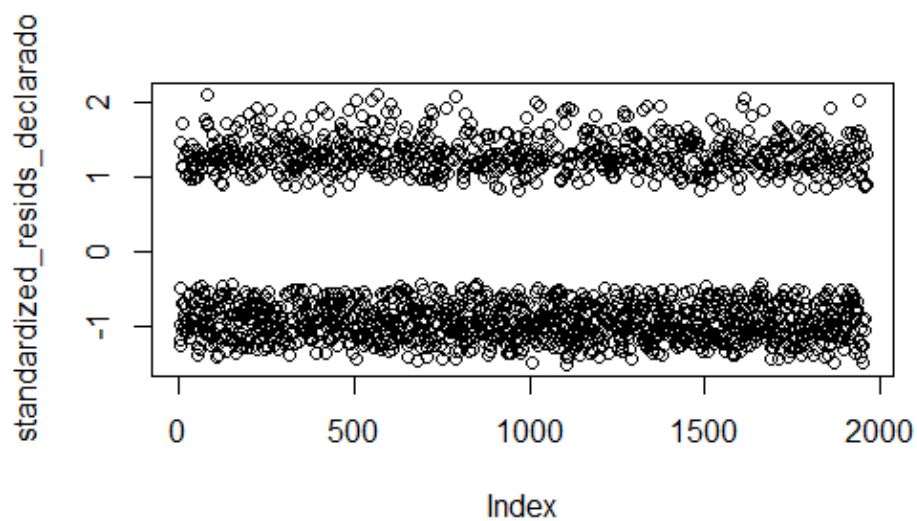
Statistic	Value
Log-Lik Intercept Only	-547.113
Log-Lik Full Model	-455.114
Deviance (793)	910.228
LR (18)	183.998
Prob > LR	0.000
McFadden's R^2	0.168
McFadden's Adj R^2	0.133
ML (Cox-Snell) R^2	0.087
Cragg-Uhler (Nagelkerke) R^2	0.208
McKelvey & Zavoina's R^2	0.317
Efron's R^2	0.203
Count R^2	0.725
Adj Count R^2	0.278
BIC	1037.518
AIC	948.228

Fonte: elaboração do autor.

Model Fit Statistics

Figura 24 – Residuals Monitorado

Fonte: elaboração do autor.

Figura 25 – Residuals Monitorado

Fonte: elaboração do autor.

10.5 Questionários

Questionário Onda 1: Antes da campanha eleitoral

1

Código da Variável	Texto	Descrição das Modalidades
id		Contagem automática de IDs únicas da plataforma de pesquisa
submitdate		Data de submissão do questionário na plataforma. Horário de Brasília (GMT -3)

BLOCO DE INTRODUÇÃO

INTRO	Olá! Nesta pesquisa queremos saber como as pessoas no Brasil se informam e quais são suas opiniões sobre alguns temas muito atuais. A pesquisa deve durar 15 minutos. Sua participação é totalmente voluntária e você pode escolher não participar ou encerrar sua participação a qualquer momento. Para obter mais informações sobre este estudo, incluindo detalhes sobre como seus dados serão coletados e armazenados, por favor, mande uma mensagem para este e-mail . Suas respostas são muito relevantes para que possamos entender melhor algumas questões sobre a opinião das pessoas no país. Portanto, é muito importante que você mesmo responda as perguntas, sem consultar outras pessoas ou a internet.	1. sim, eu concordo em participar da pesquisa. 0. não, eu não concordo em participar da pesquisa.
-------	--	--

linkrequest		ID único da empresa responsável pelo painel. Responsável pela conexão com a base de dados da Netquest
uf	Unidade da Federação Brasileira [data from Netquest]	Códigos IBGE para cada UF
region	Região [recoded from BR_UF]	Códigos IBGE para cada região: 1 "N" 2 "NE" 3 "SE" 4 "S" 5 "CO"
city	[data from Netquest]	Códigos IBGE para cada cidade
class	Critério Brasil de Classificação Econômica [data from Netquest]	6 "A" 5 "B1" 4 "B2" 3 "C1" 2 "C2" 1 "D-E".
edu	Educação	VALUE LABELS edu

Questionário Onda 1: Antes da campanha eleitoral

2

	[data from Netquest]	1 "Analfabeto / Primário incompleto (Analfabeto/ Até 3ª série Fundamental/ Até 3ª série 1º. Grau)" 2 "Primário completo / Ginásial incompleto (Até 4ª série Fundamental / Até 4ª série 1º. Grau)" 3 "Ginásial completo / Colegial incompleto (Fundamental completo/ 1º. Grau completo)" 4 "Colegial completo / Superior incompleto (Médio completo/ 2º. Grau completo)" 5 "Superior completo" 6 "Pós-graduação" 7 "Mestrado" 8 "Doutorado"
edubinned	[recoded from edu]	1 "01. Fundamental " 2 " 02. Até Médio " 3 "03. Até Superior"
useragent	[data from JavaScript]	[character] OS: Browser: Mobile: true/false Flash: yes/no check Cookies: true/false Viewport width: Viewport heigth: Screen Size: Full User Agent:

BLOCO - DADOS SOCIODEMOGRÁFICOS 1 - COTAS

Código da Variável	Texto	Descrição das Modalidades
gender	Você é...	0. Homem 1. Mulher 77. Outros
gender_other		Especificação de outros
age	E quantos anos você tem?	(Preenchimento até 2 dígitos numéricos máximo até 99 anos)
edu	Até que série você estudou?	00. Nunca frequentei a escola 11. Primário incompleto (até 3ª série do ensino fundamental) 12. Primário completo (4ª série do ensino fundamental) 13. Ginásio incompleto (até 7ª série do ensino fundamental) 14. Ginásio completo (8ª série do ensino fundamental) 21. Colegial incompleto (até 2ª série do ensino médio) 22. Colegial completo (3ª série do ensino médio) 31. Ensino universitário incompleto ou especialização (técnico após ensino médio) 32. Ensino universitário completo 33. Pós-graduação ou mais

Questionário Onda 1: Antes da campanha eleitoral

3

info	Como você avalia seu nível de informação sobre a Covid-19, de modo geral. Você se considera:	<ol style="list-style-type: none"> 1. Muito bem informado(a) 2. Razoavelmente informado(a) 3. Pouco informado(a) 4. Nada informado(a)
------	--	---

BLOCO - EXPERIMENTO 1

exp1rand	Aleatória de 1 a 7 para a escolha do tratamento	<ol style="list-style-type: none"> 1 “control” 2 “exp1indloss” 3 “exp1famloss” 4 “exp1socloss” 5 “exp1indgain” 6 “exp1famgain” 7 “exp1socgain”
----------	---	---

exp1	Você diria que tem: [recoded from next 7 responses]	<ol style="list-style-type: none"> 1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar
------	--	---

[if exp1rand = 1] exp1control controle	Considere que, em conjunto, empresas e pesquisadores vão conseguir produzir em breve uma vacina contra a Covid-19. Qual é a chance de você ir se vacinar contra o coronavírus quando a vacina estiver disponível para você? Você diria que tem:	<ol style="list-style-type: none"> 1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar
--	--	---

[if exp_arg_il = 2] exp1indloss2 (INDIVIDUAL, PERDA)	Considere que, em conjunto, empresas e pesquisadores vão conseguir produzir em breve uma vacina contra a Covid-19. Nesse caso, há algumas razões pelas quais cada um de nós deveria tomar essa nova vacina. Uma das mais importantes é que ela pode evitar que você contraia uma doença que mata. Pensando nisso, qual é a chance de você ir se vacinar contra o coronavírus quando a vacina estiver disponível para você? Você diria que tem:	<ol style="list-style-type: none"> 1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar
---	--	---

Questionário Onda 1: Antes da campanha eleitoral

4

<p>[if exp_arg_il = 3] exp1famloss</p> <p>(FAMÍLIA, PERDA)</p>	<p>Considere que, em conjunto, empresas e pesquisadores vão conseguir produzir em breve uma vacina contra a Covid-19. Nesse caso, há algumas razões pelas quais cada um de nós deveria tomar essa nova vacina. Uma das mais importantes é que ela pode evitar que você contraia uma doença que mata e transmite essa doença a pessoas da sua família. Pensando nisso, qual é a chance de você ir se vacinar contra o coronavírus quando a vacina estiver disponível para você? Você diria que tem:</p>	<p>1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar</p>
<p>[if exp_arg_il = 4] exp1socloss</p> <p>(SOCIEDADE, PERDA)</p>	<p>Considere que, em conjunto, empresas e pesquisadores vão conseguir produzir em breve uma vacina contra a Covid-19. Nesse caso, há algumas razões pelas quais cada um de nós deveria tomar essa nova vacina. Uma das mais importantes é que ela pode evitar que você contraia uma doença que mata e transmite essa doença a outras pessoas. Pensando nisso, qual é a chance de você ir se vacinar contra o coronavírus quando a vacina estiver disponível para você? Você diria que tem:</p>	<p>1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar</p>
<p>[if exp_arg_il = 5] exp1indgain</p> <p>(INDIVIDUAL, GANHO)</p>	<p>Considere que, em conjunto, empresas e pesquisadores vão conseguir produzir em breve uma vacina contra a Covid-19. Nesse caso, há algumas razões pelas quais cada um de nós deveria tomar essa nova vacina. Uma das mais importantes é que ela pode proteger sua vida contra essa doença. Pensando nisso, qual é a chance de você ir se vacinar contra o coronavírus quando a vacina estiver disponível para você? Você diria que tem:</p>	<p>1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar</p>
<p>[if exp_arg_il = 6] exp1famgain</p> <p>(FAMÍLIA, GANHO)</p>	<p>Considere que, em conjunto, empresas e pesquisadores vão conseguir produzir em breve uma vacina contra a Covid-19. Nesse caso, há algumas razões pelas quais cada um de nós deveria tomar essa nova vacina. Uma das mais importantes é que ela pode proteger sua vida contra essa doença e, assim, proteger a vida de pessoas da sua família que tenham contato com você. Pensando nisso, qual é a chance de você ir se vacinar contra o coronavírus quando a vacina estiver disponível para você? Você diria que tem:</p>	<p>1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar</p>

Questionário Onda 1: Antes da campanha eleitoral

5

<p>[if exp_arg_il = 7] exp1socgain</p> <p>(SOCIEDADE, GANHO)</p>	<p>Considere que, em conjunto, empresas e pesquisadores vão conseguir produzir em breve uma vacina contra a Covid-19. Nesse caso, há algumas razões pelas quais cada um de nós deveria tomar essa nova vacina. Uma das mais importantes é que ela pode proteger sua vida contra essa doença e, assim, proteger a vida de outras pessoas que tenham contato com você. Pensando nisso, qual é a chance de você ir se vacinar contra o coronavírus quando a vacina estiver disponível para você? Você diria que tem:</p>	<p>1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar</p>
--	---	---

PÁGINA - CONHECIMENTO SOBRE A COVID-19

<p><u>matriz</u> <u>conhecimento</u></p>	<p>Agora gostaríamos de saber sua opinião sobre algumas informações que circulam a respeito da Covid-19. Para cada frase a seguir, indique se você acha que ela é VERDADEIRA ou FALSA. Se você não souber dizer, não tem problema. Nesse caso, marque a opção “Não sei”. É importante que você não busque informações na internet para responder a estas questões. Responda simplesmente com o que você sabe ou ouviu falar. (RANDOMIZAR A ORDEM DAS FRASES)</p>	<p>2. Falso 1. Verdadeiro 96. Não sei</p> <p>(Alternativas abaixo. Terceira coluna é o gabarito que não constará no questionário aplicado)</p>
--	--	--

10	Ficar isolado em casa é uma forma eficaz de reduzir a propagação do vírus.	V
11	É possível contrair a Covid-19 se você comer alimentos contaminados.	F
12	O coronavírus pode ficar nas suas mãos se você não esfregá-las com sabão por pelo menos 20 segundos.	V
13	É possível contrair a Covid-19 se você coçar os olhos com as mãos contaminadas.	V
14	Se uma pessoa com Covid-19 tossir ou respirar perto de você, há risco de pegar a doença, mesmo que você e a outra pessoa estejam utilizando máscaras.	V

Questionário Onda 1: Antes da campanha eleitoral

6

15	A Covid-19 é uma doença sexualmente transmissível.	F
16	Pessoas com Covid-19 mas sem febre não transmitem o coronavírus a outras pessoas.	F
17	O coronavírus se espalha por meio de gotículas respiratórias de indivíduos infectados.	V
18	O uso de máscara ajuda a proteger contra o coronavírus.	V
19	Crianças não precisam se prevenir contra a Covid-19.	F
21	Diarreia e perda do paladar também podem ser sintomas de Covid-19.	V
22	Os sintomas mais comuns da Covid-19 são febre, cansaço e tosse seca.	V
23	Ao contrário do resfriado comum, nariz entupido, coriza e espirros são menos comuns em pessoas infectadas com Covid-19.	V
31	Os sintomas da Covid-19 podem demorar até 14 dias para aparecerem.	V
32	Tomar cloroquina previne a Covid-19.	F
33	Tomar cloroquina cura a Covid-19.	F
34	Existem medicamentos específicos para a prevenção da Covid-19.	F

INFO_Covid19	[recodificado da matriz_conhecimento]	01. Sim ou 02. Não 96. Prefiro não responder
--------------	---------------------------------------	--

PÁGINA - PROXIMIDADE COM A DOENÇA

Questionário Onda 1: Antes da campanha eleitoral

7

disease_01	[Câncer] Você tem ou já teve alguma destas doenças?	01. Sim ou 02. Não 96. Prefiro não responder
disease_02	[Doença renal crônica] Você tem ou já teve alguma destas doenças?	01. Sim ou 02. Não 96. Prefiro não responder
disease_03	[Doença do coração crônica] Você tem ou já teve alguma destas doenças?	01. Sim ou 02. Não 96. Prefiro não responder
disease_04	[Já sofreu transplante] Você tem ou já teve alguma destas doenças?	01. Sim ou 02. Não 96. Prefiro não responder
disease_05	[Já teve diabetes tipo 2] Você tem ou já teve alguma destas doenças?	01. Sim ou 02. Não 96. Prefiro não responder
disease_06	[Asma ou doença pulmonar grave (doença do sistema respiratório)] Você tem ou já teve alguma destas doenças?	01. Sim ou 02. Não 96. Prefiro não responder
disease_07	[Complicações respiratórias ou cardíacas por causa de obesidade] Você tem ou já teve alguma destas doenças?	01. Sim ou 02. Não 96. Prefiro não responder

disease_sum	[sum from all disease]	range 1 to 7
-------------	------------------------	--------------

<u>covhad</u>	Em relação à doença provocada pelo coronavírus, a Covid-19, marque por favor a opção a seguir que melhor se aplica a você.	0. Não tive a doença, mas nunca fiz o teste 1. Não tive a doença e já fiz o teste 2. Tive a doença, mas não tive nenhum dos sintomas 3. Tive a doença e senti sintomas leves 4. Tive a doença e senti sintomas graves
---------------	--	---

Questionário Onda 1: Antes da campanha eleitoral

8

covsome	Conhece alguém que contraiu a doença, ou seja, que fez o teste e deu positivo para Covid-19? [Múltipla escolha]	0. Não conheço ninguém 1. Sim, uma pessoa que mora comigo teve a doença 2. Sim, alguém da minha família ou amigo próximo teve a doença 3. Sim, conheço uma pessoa não tão próxima que teve a doença
covdead	E você conhece alguém que faleceu por causa da Covid-19? [Múltipla escolha]	0. Não conheço ninguém 1. Sim, uma pessoa que morava comigo 2. Sim, alguém da minha família ou amigo próximo 3. Sim, conheço uma pessoa não tão próxima que faleceu da doença
worried	E qual é o seu nível de preocupação com a doença?	1. Não estou nada preocupado(a) 2. Estou só um pouco preocupado(a) 3. Estou bastante preocupado(a) 4. Estou muito preocupado(a)

attention	Se você realmente estiver lendo o enunciado desta pergunta, selecione a opção Discordo um pouco.	1. Concordo muito 2. Concordo 3. Concordo um pouco 4. Não concordo nem discordo 5. Discordo um pouco - CONTINUAR 6. Discordo 7. Discordo muito
-----------	--	--

PÁGINA - PREFERÊNCIAS SOBRE PROVIDÊNCIAS

heatheco	<p>Todos os países do mundo estão enfrentando atualmente o coronavírus, que causa a doença Covid-19. Esta pandemia impõe importantes escolhas sociais, incluindo o equilíbrio entre a propagação do vírus e a atividade econômica.</p> <p>Algumas pessoas acreditam que as vidas devem ser protegidas a todo custo e que os governos devem impor o distanciamento social, enquanto não tenhamos uma vacina ou uma cura para o coronavírus. Suponha que coloquemos essas pessoas em 1 em uma escala de 7 pontos. Outros acreditam que a vida deve seguir e esse distanciamento social acarreta muitos problemas como a perda de empregos. Suponha que coloquemos as pessoas que pensam assim na outra extremidade da escala, em 7.</p> <p>Onde você se colocaria nessa escala, ou não pensou muito sobre isso?</p>	<p>1 ----- 2 ----- 3 ----- 4 ----- 5 ----- 6 ----- 7</p> <p>[configurar para exibir sempre em 7 colunas e fazer o teste em mobile]</p>
----------	---	--

Questionário Onda 1: Antes da campanha eleitoral

9

promatriz	[Matriz - Aleatória todas as frases] Agora queremos saber seu nível de concordância com algumas frases relacionadas à Covid-19.	1. Concordo muito 2. Concordo 3. Nem concordo nem discordo 4. Discordo 5. Discordo Muito [Sequência de frases abaixo. Frases devem aparecer em ordem aleatória.]
-----------	--	---

pro1mask	Eu evito o uso de máscara quando estou fora da minha casa
pro2crowd	Me incomoda muito ir a lugares onde há muita gente
pro3fam	Acho que as chances de pessoas da minha família e de amigos próximos pegarem a Covid-19 são maiores que as chances dos brasileiros em geral
pro4serious	A Covid-19 não é tão grave quanto se diz
pro5transport	Evitar o transporte público é importante para diminuir o número de pessoas que se infectam com o vírus
pro6school	O fechamento das escolas foi uma medida acertada para diminuir o número de pessoas infectadas
pro7close	O fechamento do comércio foi uma medida acertada para diminuir o número de pessoas infectadas
pro8quar	Sem a quarentena a pandemia da Covid-19 teria sido muito pior no país
pro9hands	Tenho higienizado minhas mãos com muito mais frequência que antes

PÁGINA - HÁBITOS DE MÍDIA

mediahab1_XX	Gostaríamos de saber um pouco mais sobre como você se informa. Desde o início da pandemia, com que frequência você se informou sobre a pandemia de Covid-19...	4. Muito frequentemente 3. Frequentemente 2. De vez em quando 1. Raramente 0. Nunca (Perguntas abaixo - Programar de forma aleatória)
--------------	--	--

Questionário Onda 1: Antes da campanha eleitoral

10

mediahab1_01	Com familiares ou amigos?
mediahab1_11	No Jornal da Record?
mediahab1_12	No Jornal Nacional, da Rede Globo?
mediahab1_21	Em jornais de circulação nacional como Folha de S. Paulo, O Estado de S. Paulo ou O Globo, mesmo que na internet?
mediahab1_25	Em algum portal da Internet, como o G1, Globo.com, R7, Terra ou UOL?
mediahab1_22	Em redes sociais como o Facebook, Instagram, YouTube ou Twitter?
mediahab1_23	Em mecanismos de busca na internet como o Google?
mediahab1_24	Em aplicativos como o WhatsApp?
mediahab1_13	Em meios externos (placas publicitárias, outdoor, ônibus, elevador, metrô)

PÁGINA - EXPERIMENTO 2

exp2rand	<p>Aleatório de 1 a 5 para a escolha do tratamento</p> <p>1 "Control"</p> <p>2 "exp2china"</p> <p>3 "exp2russia"</p> <p>4 "exp2usa"</p> <p>5 "exp2oxfeng"</p>
----------	---

exp2	Você diria que tem:	<p>1.Nenhuma chance de se vacinar</p> <p>2.Pouca chance de se vacinar</p> <p>3.Alguma chance de se vacinar</p> <p>4.Muita chance de se vacinar</p>
------	---------------------	--

Questionário Onda 1: Antes da campanha eleitoral

11

[if exp_country = 1] exp2control	Algumas vacinas contra a Covid-19 já estão em fase avançada de testes. Indique a seguir qual é a chance de você ir tomar a nova vacina contra o coronavírus quando ela estiver disponível para você. Você diria que tem:	1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar
[if exp_country = 2] exp2china	Algumas vacinas contra a Covid-19 já estão em fase avançada de testes. Uma delas está sendo produzida na China. Indique a seguir qual é a chance de você tomar a nova vacina contra o coronavírus que está sendo produzida na China. Você diria que tem:	1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar
[if exp_country = 3] exp2russia	Algumas vacinas contra a Covid-19 já estão em fase avançada de testes. Uma delas está sendo produzida na Rússia. Indique a seguir qual é a chance de você tomar a nova vacina contra o coronavírus que está sendo produzida na Rússia. Você diria que tem:	1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar
[if exp_country = 4] exp2usa	Algumas vacinas contra a Covid-19 já estão em fase avançada de testes. Uma delas está sendo produzida nos Estados Unidos. Indique a seguir qual é a chance de você tomar a nova vacina contra o coronavírus que está sendo produzida nos Estados Unidos. Você diria que tem:	1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar

Questionário Onda 1: Antes da campanha eleitoral

12

[if exp_country = 5] exp2oxfeng	Algumas vacinas contra a Covid-19 já estão em fase avançada de testes. Uma delas está sendo produzida na Universidade de Oxford, na Inglaterra. Indique a seguir qual é a chance de você tomar a nova vacina contra o coronavírus que está sendo produzida na Universidade de Oxford, na Inglaterra. Você diria que tem:	1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar
------------------------------------	--	---

PÁGINA - HÁBITOS DE MÍDIA 2

mediahab2_	Voltando a falar sobre como você tem se informado sobre a pandemia, gostaria de saber com que frequência você:	4. Muito frequentemente 3. Frequentemente 2. De vez em quando 1. Raramente 0. Nunca 96. Não sabe (Perguntas abaixo - Programar de forma aleatória)
------------	--	---

mediahab2_disagree	Leu algo do qual você DISCORDOU?
mediahab2_pandisa	Checkou uma fonte de notícias ou informação DIFERENTE das que você normalmente assiste ou acessa?
mediahab2_pandif	Procurou por pessoas em quem você CONFIA para ver o que elas diziam?
mediahab2_panditrust	Descobriu algo que MUDOU a sua opinião sobre o coronavírus ou a pandemia?

fn	Pensando agora sobre as diversas notícias e informações sobre a pandemia, você se lembra de ter visto algo que era possivelmente inventado, exagerado ou incorreto? [0 e 1 em ordem aleatória]	0 Não 1 Sim
----	---	----------------

Questionário Onda 1: Antes da campanha eleitoral

13

[if fn = 1] fnshare	<p> você chegou a compartilhar essa notícia ou informação? [1 e 2 em rodízio] </p>	<p> 1. Sim, mas não sabia que era inventada, exagerada ou incorreta 2. Sim, mesmo achando que era inventada, exagerada ou incorreta 0. Não 96. Não me lembro </p>
------------------------	---	--

[if fn = 1] fnwhere	<p> Onde foi o lugar que você mais recebeu ou viu essas notícias e informações que possivelmente eram inventadas, exageradas ou incorretas? [alternativas em rodízio] </p>	<p> 01. Com familiares ou amigos? 11. No Jornal da Record? 12. No Jornal Nacional, da Rede Globo? 21. Em jornais de circulação nacional como Folha de S. Paulo, Estado de S. Paulo ou O Globo, mesmo que na internet? 25. Em algum portal da Internet, como o G1, Globo.com, R7, Terra ou UOL? 22. Em redes sociais como o Facebook, Instagram, YouTube ou Twitter? 23. Em mecanismos de busca na internet como o Google? 24. Em aplicativos como o WhatsApp? 77. Outro lugar? </p>
------------------------	---	---

fnpte	<p> Você diria que as pessoas são capazes de compartilhar essas notícias ou informações sobre a pandemia mesmo sabendo que elas são possivelmente inventadas, exageradas ou falsas? [0 e 1 em ordem aleatória] </p>	<p> 0 Não 1 Sim 96 Não sei </p>
-------	--	---

BLOCO - PREFERÊNCIA PARTIDÁRIA E ELEIÇÕES

TEXT	<p>Agora mudando um pouco de assunto e falando sobre política e as próximas eleições</p>
------	--

evalcovid	<p> Na sua avaliação, a atuação do governo do presidente Jair Bolsonaro no combate à pandemia de Covid-19 está sendo: Ótima, Boa, Regular, Ruim ou Péssima? </p>	<p> 1. Ótima 2. Boa 3. Regular 4. Ruim 5. Péssima </p>
-----------	--	--

Questionário Onda 1: Antes da campanha eleitoral

14

evalgov	E no geral, você acha que o governo do presidente Jair Bolsonaro até aqui está sendo: ótimo, bom, regular, ruim ou péssimo?	<ol style="list-style-type: none"> 1. Ótimo 2. Bom 3. Regular 4. Ruim 5. Péssimo
partypref	Qual é o seu partido político de preferência? [Single Option and Randomize options 1-8]	<ol style="list-style-type: none"> 1. PT 2. PSL 3. MDB 4. PSDB 5. DEM 6. PSOL 7. PDT 8. Partido do presidente Jair Bolsonaro 77. Outro partido não listado, qual? [open – accept both letters and numbers] 9. Não tenho partido de preferência
partynev	Em qual partido político você nunca votaria? [Single Option and Randomize options 1-8]	<ol style="list-style-type: none"> 1. PT 2. PSL 3. MDB 4. PSDB 5. DEM 6. PSOL 7. PDT 8. Partido do presidente Jair Bolsonaro 77. Outro partido não listado, qual? [open – accept both letters and numbers] 9. Não tenho partido que nunca votaria
polcand	Você já sabe ou já ouviu falar sobre quem são os(as) candidatos(as) ao cargo de prefeito(a) em sua cidade?	<ol style="list-style-type: none"> 0. Não sei de nenhum(a) candidato(a) a prefeito(a) 1. Sei de um(a) candidato(a) a prefeito(a) 2. Sei de dois(duas) candidatos(as) a prefeito(as) 3. Sei de três ou mais candidatos(as) a prefeito(a)

Questionário Onda 1: Antes da campanha eleitoral

15

polvot	(Apenas para ≠ de 0 na anterior) E você já sabe em quem você vai votar?	0. Não pretendo votar 1. Ainda não sei em quem vou votar 11. Tenho alguma ideia, mas ainda estou em dúvida 12. Sim, tenho certeza em quem vou votar
--------	--	--

PÁGINA - EXPERIMENTO 3

exp3corr	Todos os países estão enfrentando o coronavírus, também conhecido como Covid-19. Para você, o coronavírus é: [Aleatório opções 1 e 2]	1. Um vírus que se originou em um mercado de animais vivos em Wuhan, na China, e que acabou se espalhando de maneira involuntária pelo mundo. 2. Um vírus que foi criado de propósito num laboratório chinês para aumentar o poder econômico da China. 3. Não sei qual é a verdade.
----------	--	---

[EXPERIMENTO: CORRECTION FOR THOSE WHO ANSWERED (2) IN SE1] [TWO VERSIONS – Randomize SE1C1 and SE1C2]

[if exp_corr = 2] exp3rand	Aleatório de 1 a 2, somente se 2 na exp3corr para a escolha dos tratamentos
-------------------------------	---

[if exp_corr = 2 & exp_corr_rand = 1] exp3corr1	Essa explicação que você escolheu é considerada uma teoria conspiratória, ou seja, é uma explicação que não está baseada em fatos comprovados. A única explicação baseada nos fatos é a de que o coronavírus se originou em um mercado de animais vivos em Wuhan, na China, e acabou se espalhando de maneira involuntária pelo mundo. Pensando nessas informações que acabou de ler, [DO NOT randomize option order]	1. Você muda de opinião e concorda que o coronavírus se originou em um mercado de animais vivos em Wuhan, na China, e acabou se espalhando de maneira involuntária pelo mundo. 2. Você prefere manter a sua opinião. 3. Você não sabe mais qual é a verdade.
[if exp_corr = 2 & exp_corr_rand = 2] exp3corr2	Essa explicação que você escolheu é considerada uma teoria conspiratória, ou seja, é uma explicação que não está baseada em fatos comprovados. Não há evidência que sugira que o coronavírus tenha sido	1. Você muda de opinião e concorda que o coronavírus se originou em um mercado de animais vivos em Wuhan,

Questionário Onda 1: Antes da campanha eleitoral

16

	criado de propósito num laboratório chinês para aumentar o poder econômico da China. Pensando nessas informações que acabou de ler, [DO NOT randomize option order]	na China, e acabou se espalhando de maneira involuntária pelo mundo. 2. Você prefere manter a sua opinião. 3. Você não sabe mais qual é a verdade.
--	---	--

5.

BLOCO - SOCIO DEMOGRÁFICO

troll1	Você dirigiu um carro sozinho antes dos 6 anos de idade?	0. Não 1. Sim
--------	--	------------------

marstat	Qual é seu estado civil?	0. Solteiro(a) 1. Casado(a) ou em União Estável 2. Divorciado(a) 3. Viúvo(a) 4. Separado(a)
anychild	Você tem filhos em idade escolar?	0. Não 1. Sim
adults	Quantas pessoas adultas moram com você?	0. Moro sozinho 1. 1 2. 2 3. 3 4. 4 ou mais pessoas
adultover60	[Exibir apenas se 1 na adults] Se 1: Essa pessoa tem 60 anos de idade ou mais?	0. Não 1. Sim
adultover60	[Exibir apenas se > 1 na adults] Alguma dessas pessoas tem 60 anos de idade ou mais?	0. Não 1. Sim
occup	Atualmente qual é a sua ocupação?	11. Trabalhador(a) com carteira assinada 12. Trabalhador(a) sem carteira assinada 13. Funcionário(a) público(a) 14. Empresário(a) 01. Desempregado(a) procurando emprego 02. Desempregado(a) mas não procuro emprego 03. Aposentado

Questionário Onda 1: Antes da campanha eleitoral

17

		04. Atividades domésticas 05. Estudante.
religion	Qual a sua religião, se tiver? [Randomize option 1 a 13]	1. Católica 2. Evangélico Pentecostal 3. Evangélico Neo-Pentecostal 4. Outra religião 14. Não tem religião/Agnóstico 15. Sou ateu/Não acredito em Deus
race	A sua raça ou cor é:	1. Branca 2. Parda 3. Preta 4. Amarela 5. Indígena
benef	Você ou alguém que mora com você recebe um benefício do Bolsa Família?	0. Não 1. Sim
benef2	Você ou alguém que mora com você recebe o auxílio emergencial da Covid-19?	0. Não 1. Sim
pea	[recoded from occup]	1. Trabalhador(a) com carteira assinada 1. Trabalhador(a) sem carteira assinada 1. Funcionário(a) público(a) 1. Empresário(a) 0. Desempregado(a) procurando emprego 0. Desempregado(a) mas não procuro emprego 0. Aposentado 0. Atividades domésticas 0. Estudante.

Questionário Onda 2: Depois do 2 turno

1

PÁGINA DE INTRODUÇÃO (sem menção à primeira rodada) e TCLE

Olá! Esta pesquisa pretende entender melhor como os brasileiros têm lidado com alguns temas da atualidade e quais são suas expectativas para o próximo ano. Sua participação é totalmente voluntária e você pode escolher não participar ou encerrar sua participação a qualquer momento. Para obter mais informações sobre este estudo, por favor mande uma mensagem para o e-mail contato@ibpad.com.br . Suas respostas são muito relevantes para que possamos entender melhor algumas questões sobre a opinião das pessoas no país. Portanto, é muito importante que você mesmo responda as perguntas, sem consultar outras pessoas ou a internet.

- sim, eu concordo em participar da pesquisa.
 não, eu não concordo em participar da pesquisa.

BLOCO - DADOS SOCIODEMOGRÁFICOS 1 - COTAS

Código da Variável	Texto	Descrição das Modalidades
gender_w2	Você é...	0. Homem 1. Mulher
age_w2	E quantos anos você tem?	(Preenchimento até 2 dígitos numéricos máximo até 99 anos)

uf_w2	Em qual Unidade da Federação brasileira você mora?	Listagem das UFs com código do IBGE
-------	--	-------------------------------------

mun_uf	Em que município da/de/do [UF]?	Listagem dos municípios para cada UF conforme código IBGE
--------	---------------------------------	---

BLOCO - INFO BRASIL

newsbr	Pensando no noticiário das últimas semanas sobre o Brasil , qual é o assunto ou a notícia que você mais se lembra? Anote a seguir sua resposta, de forma curta. E em segundo lugar?
--------	--

BLOCO - EXPERIMENTO 1

exp1rand	Random 1 to 6
----------	---------------

Questionário Onda 2: Depois do 2 turno

2

[if exp1rand = 1] exp1controllow	Em 2021, uma vacina segura contra a Covid-19 provavelmente deve estar aprovada e disponível para a população. Até aqui, uma das vacinas em fase mais adiantada de testes e aprovação tem eficácia de 60%. Ou seja, 6 em cada 10 pessoas que tomam a vacina ficam imunes ao coronavírus. Pensando nessa vacina, especificamente, você diria que tem:	1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar
[if exp1rand = 2] exp1controlhigh	Em 2021, uma vacina segura contra a Covid-19 provavelmente deve estar aprovada e disponível para a população. Até aqui, uma das vacinas em fase mais adiantada de testes e aprovação tem eficácia de 90%. Ou seja, 9 em cada 10 pessoas que tomam a vacina ficam imunes ao coronavírus. Pensando nessa vacina, especificamente, você diria que tem:	1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar
[if exp1rand = 3] exp1chinalow	Em 2021, uma vacina segura contra a Covid-19 provavelmente deve estar aprovada e disponível para a população. Até aqui, uma das vacinas em fase mais adiantada de testes e aprovação está sendo desenvolvida na China e tem eficácia de 60%. Ou seja, 6 em cada 10 pessoas que tomam a vacina ficam imunes ao coronavírus. Pensando nessa vacina, especificamente, você diria que tem:	1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar
[if exp1rand = 4] exp1chinahigh	Em 2021, uma vacina segura contra a Covid-19 provavelmente deve estar aprovada e disponível para a população. Até aqui, uma das vacinas em fase mais adiantada de testes e aprovação está sendo desenvolvida na China e tem eficácia de 90%. Ou seja, 9 em cada 10 pessoas que tomam a vacina ficam imunes ao coronavírus. Pensando nessa vacina, especificamente, você diria que tem:	1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar
[if exp1rand = 5] exp1uslow	Em 2021, uma vacina segura contra a Covid-19 provavelmente deve estar aprovada e disponível para a população. Até aqui, uma das vacinas em fase mais adiantada de testes e aprovação está sendo desenvolvida nos Estados Unidos e tem eficácia de 60%. Ou seja, 6 em cada 10 pessoas que tomam a vacina ficam imunes ao coronavírus. Pensando nessa vacina, especificamente, você diria que tem:	1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar

Questionário Onda 2: Depois do 2 turno

3

[if exp1rand = 6] exp1ushigh	Em 2021, uma vacina segura contra a Covid-19 provavelmente deve estar aprovada e disponível para a população. Até aqui, uma das vacinas em fase mais adiantada de testes e aprovação está sendo desenvolvida nos Estados Unidos e tem eficácia de 90%. Ou seja, 9 em cada 10 pessoas que tomam a vacina ficam imunes ao coronavírus. Pensando nessa vacina, especificamente, você diria que tem:	1. Nenhuma chance de se vacinar 2. Pouca chance de se vacinar 3. Alguma chance de se vacinar 4. Muita chance de se vacinar
---------------------------------	--	---

PÁGINA - PROXIMIDADE E PREOCUPAÇÃO COM A COVID

covhad2w	Em relação à doença provocada pelo coronavírus, marque por favor a opção a seguir que melhor se aplica a você.	0. Não tive a doença, mas nunca fiz o teste 1. Não tive a doença e já fiz o teste 2. Tive a doença, mas não tive nenhum dos sintomas 3. Tive a doença e senti sintomas leves 4. Tive a doença e senti sintomas graves
----------	--	---

worried2w	E qual é o seu nível de preocupação com a doença?	1. Não estou nada preocupado(a) 2. Estou só um pouco preocupado(a) 3. Estou bastante preocupado(a) 4. Estou muito preocupado(a)
-----------	---	--

Questionário Onda 2: Depois do 2 turno

4

PÁGINA - COOPERAÇÃO

promatriz	[Matriz - Aleatória todas as frases] Agora queremos saber seu nível de concordância com algumas frases relacionadas à Covid-19.	1. Concordo muito 2. Concordo 3. Nem concordo nem discordo 4. Discordo 5. Discordo Muito [Sequência de frases abaixo. Frases devem aparecer em ordem aleatória.]
-----------	--	---

pro1mask	Eu uso máscara quando estou fora da minha casa.
pro10distance	Eu tenho mantido uma distância mínima de dois metros em relação a outras pessoas, mesmo usando máscara.
pro9hands	Tenho higienizado minhas mãos com muito mais frequência que antes.

attention2w	Se você realmente estiver lendo o enunciado desta pergunta, selecione a opção Concordo.	1. Concordo muito 2. Concordo 3. Concordo um pouco 4. Não concordo nem discordo 5. Discordo um pouco 6. Discordo 7. Discordo muito
-------------	---	--

!!! Atenção - Configurar cota no valor de 0 para as opções que não são a 2 e "expulsar" o respondente com um redirecionamento para a Netquest !!!

BLOCO - EXPERIMENTO 2

exp2rand	Random 1 to 4
----------	---------------

[if exp2rand = 1] exp2brasil	Para terminarmos esta parte da pesquisa sobre a Covid-19, gostaríamos de fazer mais uma pergunta sobre a vacina contra essa doença. Alguns estudos estão sendo feitos sobre a intenção de vacinação da população brasileira. Um levantamento preliminar mostra que a maioria das pessoas no Brasil pretende se vacinar contra a Covid-19. E você? Você diria que tem:	1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar
---------------------------------	---	---

Questionário Onda 2: Depois do 2 turno

5

[if exp2rand = 2] exp2mun	Para terminarmos esta parte da pesquisa sobre a Covid-19, gostaríamos de fazer mais uma pergunta sobre a vacina contra essa doença. Alguns estudos estão sendo feitos sobre a intenção de vacinação da população brasileira. Um levantamento preliminar mostra que a maioria das pessoas em {cityname} pretende se vacinar contra a Covid-19. E você? Você diria que tem:	1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar
[if exp2rand = 3] exp2mungain	Para terminarmos esta parte da pesquisa sobre a Covid-19, gostaríamos de fazer mais uma pergunta sobre a vacina contra essa doença. Alguns estudos estão sendo feitos sobre a intenção de vacinação da população brasileira. Um levantamento preliminar mostra que a maioria das pessoas em {cityname} pretende se vacinar contra a Covid-19. Os médicos dizem que, se todos tomarmos essa vacina, ela pode salvar muitas vidas. E você? Você diria que tem:	1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar
[if exp2rand = 4] exp2munloss	Para terminarmos esta parte da pesquisa sobre a Covid-19, gostaríamos de fazer mais uma pergunta sobre a vacina contra essa doença. Alguns estudos estão sendo feitos sobre a intenção de vacinação da população brasileira. Um levantamento preliminar mostra que a maioria das pessoas em {cityname} pretende se vacinar contra a Covid-19. Os médicos dizem que, se todos tomarmos essa vacina, ela pode evitar muitas mortes. E você? Você diria que tem:	1.Nenhuma chance de se vacinar 2.Pouca chance de se vacinar 3.Alguma chance de se vacinar 4.Muita chance de se vacinar

BLOCO - AVALIAÇÃO BOLSONARO

Exposição de texto: "Agora mudando de assunto, vamos falar um pouco sobre política e as eleições municipais"

evalgov	Na sua avaliação, o governo do presidente Jair Bolsonaro está sendo até aqui: ótimo, bom, regular, ruim ou péssimo?	1. Ótimo 2. Bom 3. Regular 4. Ruim 5. Péssimo
---------	---	---

promatriz2	[Matriz - Aleatória todas as frases] Gostaríamos também que você avaliasse o desempenho do governo do presidente Jair Bolsonaro em diferentes áreas. Marque, para cada área: Ótimo, Bom, Regular, Ruim ou Péssimo. (ALEATORIZAR ordem dos itens.)	1. Ótimo 2. Bom 3. Regular 4. Ruim 5. Péssimo
------------	--	---

Questionário Onda 2: Depois do 2 turno

6

	[Sequência de frases abaixo. Frases devem aparecer em ordem aleatória.]
--	---

eval_unemp	no combate ao desemprego?
eval_violen	no combate à violência?
eval_corrup	no combate à corrupção?
eval_heath	no combate aos problemas da saúde de modo geral?
eval_covid	no combate à pandemia de Covid-19?
eval_edu	no combate aos problemas da educação de modo geral?
eval_eco	no combate à crise econômica de modo geral?

BLOCO - 1 Turno1

pol20t1 se não for do df	1. Você votou no primeiro turno em algum(a) candidato(a) ao cargo de prefeito(a) municipal?	<ol style="list-style-type: none"> 1. Sim, votei em um(a) candidato(a) 2. Votei em branco 3. Anulei 4. Não fui votar
-----------------------------	---	--

absten20t1a if pol20 == 2 & 3	Por que não votou? [Randomizar a ordem das opções 1 a 7]	<ol style="list-style-type: none"> 3. Achei que meu voto não faria diferença. 5. Não gostei de nenhum(a) candidato(a). 6. Meu(Minha) candidato(a) não tinha nenhuma chance de ganhar. 8. Outro motivo. Qual? (Pergunta aberta)
----------------------------------	---	--

absten20t1b if pol20 == 4	Por que votou em branco ou anulou no primeiro turno? [Randomizar a ordem das opções 1 a 7]	<ol style="list-style-type: none"> 1. Tive medo de contrair Covid-19, por causa da falta de cuidado das outras pessoas. 2. Tive medo de contrair Covid-19, por causa da falta de organização da Justiça Eleitoral. 3. Achei que meu voto não faria diferença. 4. Normalmente não costumo votar. 5. Não gostei de nenhum(a) candidato(a). 6. Meu(Minha) candidato(a) não tinha nenhuma chance de ganhar. 7. Estava fora do meu domicílio eleitoral. 8. Outro motivo. Qual? (Pergunta aberta)
------------------------------	---	---

BLOCO - 2 Turno

pol20t2 [se mun com 2t]	Você votou no segundo turno em algum(a) candidato(a) ao cargo de prefeito(a) municipal?	<ol style="list-style-type: none"> 1. Sim, votei em um(a) candidato(a) 2. Votei em branco 3. Anulei 4. Não fui votar
----------------------------	---	--

Questionário Onda 2: Depois do 2 turno

7

absten20t1a if pol20 == 2 & 3	Por que não votou? [Randomizar a ordem das opções 1 a 7]	3. Achei que meu voto não faria diferença. 5. Não gostei de nenhum(a) candidato(a). 6. Meu(Minha) candidato(a) não tinha nenhuma chance de ganhar. 8. Outro motivo. Qual? (Pergunta aberta)
----------------------------------	---	--

absten20t1b if pol20 == 4	Por que votou em branco ou anulou no primeiro turno? [Randomizar a ordem das opções 1 a 7]	1. Tive medo de contrair Covid-19, por causa da falta de cuidado das outras pessoas. 2. Tive medo de contrair Covid-19, por causa da falta de organização da Justiça Eleitoral. 3. Achei que meu voto não faria diferença. 4. Normalmente não costumo votar. 5. Não gostei de nenhum(a) candidato(a). 6. Meu(Minha) candidato(a) não tinha nenhuma chance de ganhar. 7. Estava fora do meu domicílio eleitoral. 8. Outro motivo. Qual? (Pergunta aberta)
------------------------------	---	---

BLOCO - POLÍTICA E ELEIÇÕES - PARTE 2

polact	Sobre a última campanha eleitoral, com qual frequência você realizou alguma das seguintes atividades?	4. Sempre 3. Frequentemente 2. Algumas vezes 1. Raramente 0. Nunca 96. Não sabe
--------	---	--

pol_social	Postou ou compartilhou algo nas redes sociais e WhatsApp apoiando ou criticando algum(a) candidato(a)
pol_talk	Conversou sobre a campanha política com seus amigos e familiares
pol_change	Tentou mudar o voto de alguém
pol_groups	Discutiu de forma ativa em grupos de política no WhatsApp sobre a campanha
pol_engaj	Acompanhou as redes sociais de algum(a) candidato(a), interagindo, comentando ou participando de lives

Questionário Onda 2: Depois do 2 turno

8

gglesearch se não for do df	Ao longo do período eleitoral, com qual frequência você fez cada um desses tipos de busca sobre eleições no Google?	4. Sempre 3. Frequentemente 2. Algumas vezes 1. Raramente 0. Nunca
--------------------------------	---	--

cand	Para se informar sobre os candidatos
surv	Para se informar sobre quem estava à frente nas pesquisas
check	Checar se uma informação que recebeu era verdadeira
vote	Para se informar sobre procedimentos da votação (horários, locais, como justificar, etc)

PÁGINA - HÁBITOS DE MÍDIA

<u>mediahabpoll</u> se não for do df	Gostaríamos de saber um pouco mais sobre como você se informa. Ao longo do período eleitoral, com que frequência você se informou sobre as candidaturas à prefeitura do seu município?	4. Sempre 3. Frequentemente 2. Algumas vezes 1. Raramente 0. Nunca (Perguntas abaixo - Programar de forma aleatória)
---	--	--

01	Conversando com amigos e familiares
11	Assistindo a programas eleitorais gratuitos na TV e no Rádio
14	Com veículos da sua cidade (rádio, jornais impressos e tv local)
25	Em algum portal da Internet, como o G1, Globo.com, R7, Terra ou UOL?
26	Em algum site de notícias da sua cidade
22	Por informações que viu em redes sociais como (Facebook, Instagram, YouTube ou Twitter) ou Whatsapp

<u>media_geral1</u> (para todos)	E de forma geral, que meio de comunicação você mais usa para se informar sobre o que acontece no Brasil? (ordem aleatória de alternativas de 1 a 5)	1. TV 2. Rádio 3. Jornais impressos e revistas 4. Sites de notícia na internet 5. Redes sociais e WhatsApp 6. Outro meio Anote aqui: _____
-------------------------------------	---	---

Questionário Onda 2: Depois do 2 turno

9

<u>media_geral2</u> (para todos)	E em segundo lugar, que meio de comunicação você mais usa para se informar sobre o que acontece no Brasil? (ordem aleatória de alternativas de 1 a 5)	<ol style="list-style-type: none"> 1. TV 2. Rádio 3. Jornais impressos e revistas 4. Sites de notícia na internet 5. Redes sociais e WhatsApp 6. Outro meio Anote aqui: _____
-------------------------------------	---	---

BLOGO INFO PAÍSES

	<p>Nos últimos meses, a imprensa trouxe muitas informações sobre as diferentes vacinas contra a Covid-19 que vêm sendo desenvolvidas por várias empresas e instituições.</p> <p>Queremos saber o quanto as pessoas no Brasil sabem sobre o país de origem de algumas dessas vacinas. Se você não souber, não tem problema. Nesse caso, marque a primeira opção: "Não sei".</p> <p>Pelo que você já sabe ou ouviu falar, em que país está sendo desenvolvida a vacina da...</p>	<ol style="list-style-type: none"> 1. Não sei 2. Alemanha 3. Brasil 4. China 5. Estados Unidos 6. Inglaterra 7. Rússia <p>(randômico de 2 a 7)</p>
--	--	---

ormoderna	Moderna?
orpfizer	Pfizer?
orbiontech	BioNTech?
orastrazeneca	AstraZeneca?
oroxford	Universidade de Oxford?
orsinovac	Sinovac?

Questionário Onda 2: Depois do 2 turno

10

Experimento 3 - Google - Página 1

exp3rand	Random 1 to 6
----------	---------------

[if exp3rand= 1] exp3uolgg1	<p>Durante a apuração dos resultados das eleições do primeiro turno, houve um atraso atípico por parte do Tribunal Superior Eleitoral (TSE) na totalização dos votos de alguns municípios.</p> <p>Ao fazer uma pesquisa no Google sobre “problemas na apuração do 1º turno”, um dos principais resultados é uma reportagem do Portal UOL que está na imagem abaixo:</p>	
--------------------------------	--	--

[if exp3rand= 2] exp3uolrs	<p>Durante a apuração dos resultados das eleições do primeiro turno, houve um atraso atípico por parte do Tribunal Superior Eleitoral (TSE) na totalização dos votos de alguns municípios.</p> <p>Uma publicação que foi muito compartilhada nas redes sociais sobre “problemas na apuração no 1º turno” é uma reportagem do Portal UOL que está na imagem abaixo:</p>	
-------------------------------	--	--

[if exp3rand= 3] exp3uolcontrol	<p>Durante a apuração dos resultados das eleições do primeiro turno, houve um atraso atípico por parte do Tribunal Superior Eleitoral (TSE) na totalização dos votos de alguns municípios.</p> <p>Uma das reportagens mais lidas no Portal UOL sobre “problemas na apuração do 1º turno” está na imagem abaixo:</p>	
------------------------------------	--	--

Questionário Onda 2: Depois do 2 turno

11

Exemplo das imagens do grupo 1 a 3 para Mobile

UOL NOTÍCIAS

PRODUTOS NOTÍCIAS ELEIÇÕES CARROS ECONOMIA FOLHA ESPORTE SPLASH TV E FAMOSOS UNIVERSA VIVABEM

POLÍTICA

Análise da PF contradiz versão de Barroso sobre ataque hacker nas eleições 2020

Do UOL, em São Paulo

Análise da Polícia Federal (PF) em conjunto com o Tribunal Superior Eleitoral (TSE) apontou que o ataque hacker que acessou dados pessoais de funcionários do tribunal foi mais amplo que o divulgado anteriormente por Luís Roberto Barroso e teria ocorrido neste ano.

Segundo as investigações, o criminoso acessou dados de 2020 e a suspeita é de que o ataque tenha ocasionado o atraso da apuração na noite do 1º turno das eleições municipais. A informação inicial dada por Barroso, porém, era de que o atraso aconteceu devido a um dos núcleos de processadores do supercomputador que processa a totalização ter falhado e ter sido preciso repará-lo.

[if exp3rand= 4] exp3gbbg!	<p>Durante a apuração dos resultados das eleições do primeiro turno, houve um atraso atípico por parte do Tribunal Superior Eleitoral (TSE) na totalização dos votos de alguns municípios.</p> <p>Ao fazer uma pesquisa no Google sobre “problemas na apuração do 1º turno”, um dos principais resultados é uma reportagem da Gazeta Brasil que está na imagem abaixo:</p>	
----------------------------	--	--

Questionário Onda 2: Depois do 2 turno

12

[if exp3rand= 5] exp3gbrs	<p>Durante a apuração dos resultados das eleições do primeiro turno, houve um atraso atípico por parte do Tribunal Superior Eleitoral (TSE) na totalização dos votos de alguns municípios.</p> <p>Uma publicação que foi muito compartilhada nas redes sociais sobre “problemas na apuração no 1º turno” é uma reportagem da Gazeta Brasil que está na imagem abaixo:</p>	
[if exp3rand= 6] exp3gbcontrol	<p>Durante a apuração dos resultados das eleições do primeiro turno, houve um atraso atípico por parte do Tribunal Superior Eleitoral (TSE) na totalização dos votos de alguns municípios.</p> <p>Uma das reportagens mais lidas na Gazeta Brasil sobre “problemas na apuração do 1º turno” está na imagem abaixo:</p>	

Exemplo das imagens do grupo 4 a 6 para e Mobile


f @ t v Q

[DESTAQUE](#) [AO VIVO](#) [POLÍTICA](#) [ECONOMIA](#) [ESPECIAIS](#) [BRASIL](#) [MUNDO](#) [OPINIÃO DO LEITOR](#) [MAIS](#)

DESTAQUES
POLÍTICA

Análise da PF contradiz versão de Barroso sobre ataque hacker nas eleições 2020

De Gazeta Brasil

Análise da Polícia Federal (PF) em conjunto com o Tribunal Superior Eleitoral (TSE) apontou que o ataque hacker que acessou dados pessoais de funcionários do tribunal foi mais amplo que o divulgado anteriormente por Luís Roberto Barroso e teria ocorrido neste ano.

Segundo as investigações, o criminoso acessou dados de 2020 e a suspeita é de que o ataque tenha ocasionado o atraso da apuração na noite do 1º turno das eleições municipais. A informação inicial dada por Barroso, porém, era de que o atraso aconteceu devido a um dos núcleos de processadores do supercomputador que processa a totalização ter falhado e ter sido preciso repará-lo.

A informação inicial dada por Barroso, porém, era de que os dados obtidos pelo hacker se referiam somente ao período de 2001 a 2010.

A suspeita é de que o ataque tenha acontecido antes de 1º de setembro porque o material não mostra informações registradas nos arquivos do TSE após o dia 2 daquele mês.

Questionário Onda 2: Depois do 2 turno

13

trust01	Quanto você confia na informação desta reportagem?	<ol style="list-style-type: none"> 1. Confio totalmente 2. Confio um pouco 3. Desconfio um pouco 4. Desconfio totalmente
---------	--	--

Experimento 3 - Google - Página 2

trust02gg1 [se exp3rand = 1 3 4 6]	Em geral, quanto você confia nos principais resultados da busca do Google?	<ol style="list-style-type: none"> 1. Confio totalmente 2. Confio um pouco 3. Desconfio um pouco 4. Desconfio totalmente
---------------------------------------	--	--

trust02rsl [se exp3rand = 2 3 5 6]	Em geral, quanto você confia em notícias muito compartilhadas em redes sociais?	<ol style="list-style-type: none"> 1. Confio totalmente 2. Confio um pouco 3. Desconfio um pouco 4. Desconfio totalmente
---------------------------------------	---	--

checktse	Na sua opinião, qual é a razão do atraso da apuração no primeiro turno? [Resposta única aleatório de 1 a 4]	<ol style="list-style-type: none"> 1. Devido ao ataque de hackers ao sistema do TSE 2. Devido a uma falha no supercomputador que totaliza os votos 3. Suspeitas de classificação errada nos dados de candidatos 4. Falta de energia elétrica na sede do TSE 67. Não sei
----------	---	--

Questionário Onda 2: Depois do 2 turno

14

BLOCO - SOCIO DEMOGRÁFICO

occup	Atualmente qual é a sua ocupação?	11. Trabalhador(a) com carteira assinada 12. Trabalhador(a) sem carteira assinada 13. Servidor ou funcionário(a) público(a) 14. Empresário(a) 01. Desempregado(a) procurando emprego 02. Desempregado(a) mas não procuro emprego 03. Aposentado 04. Atividades domésticas 05. Estudante.
-------	-----------------------------------	--

Debriefing	Atenção, a notícia exibida sobre os problemas do 1º turno das eleições municipais brasileiras não é verdadeira . A verdadeira razão dos atrasos foi um problema em um dos computadores do TSE responsável pela totalização dos votos. Por favor, indique a seguir que você leu esta informação.
	Estou ciente

10

PROJOR - Instituto para o Desenvolvimento do Jornalismo. 2022. “Atlas Da Notícia.” <https://www.atlas.jor.br/>.