



Universidade de Brasília

Instituto de Ciências Exatas  
Departamento de Ciência da Computação

# **Decisões Eficazes em Gestão de Pessoas: Uma Abordagem por Níveis de Maturidade em People Analytics**

Stênio Soares e Silva

Dissertação apresentada como requisito parcial para conclusão do  
Mestrado Profissional em Computação Aplicada

Orientador

Prof. Dr. João José Costa Gondim

Brasília  
2025

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

|       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|-------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| S676d | <p>Soares e Silva, Stênio</p> <p>Decisões Eficazes em Gestão de Pessoas: Uma Abordagem por Níveis de Maturidade em People Analytics / Stênio Soares e Silva; orientador João José Costa Gondim. Brasília, 2025. 115 p.</p> <p>Dissertação(Mestrado Profissional em Computação Aplicada) Universidade de Brasília, 2025.</p> <p>1. People Analytics. 2. Business Intelligence. 3. Analytics Maturity Levels. 4. People Management. 5. Prescriptive Analytics. I. Costa Gondim, João José, orient. II. Título.</p> |
|-------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|



Universidade de Brasília

Instituto de Ciências Exatas  
Departamento de Ciência da Computação

# **Decisões Eficazes em Gestão de Pessoas: Uma Abordagem por Níveis de Maturidade em People Analytics**

Stênio Soares e Silva

Dissertação apresentada como requisito parcial para conclusão do  
Mestrado Profissional em Computação Aplicada

Prof. Dr. João José Costa Gondim (Orientador)  
CIC/UnB

Prof. Dr. Marcos Fagundes Caetano    Prof. Dr. Laerte Peotta de Melo  
CIC/UnB    Banco do Brasil S.A.

Programa de Pós-graduação em Computação Aplicada

Brasília, 05 de Janeiro de 2025

# Dedicatória

À memória de meu pai e de minha mãe, que, embora não estejam mais entre nós, continuam a guiar meus passos com seus ensinamentos e amor eterno. Este trabalho é um tributo à força, sabedoria e apoio incondicional que sempre me proporcionaram. Que a luz deles siga iluminando meu caminho.

# Agradecimentos

A realização deste trabalho não seria possível sem o apoio e a colaboração de muitas pessoas, às quais expresso minha mais sincera gratidão.

Ao meu orientador, agradeço profundamente pelas valiosas orientações, pela paciência e pelo incentivo constante ao longo desta jornada. Sua experiência e dedicação foram fundamentais para a condução e conclusão deste trabalho.

Aos membros da banca, registro minha imensa gratidão pela disponibilidade, pelas contribuições e pelas reflexões que enriqueceram este projeto, ajudando a elevar a qualidade deste estudo.

Aos meus colegas do Programa de Pós-Graduação em Computação Aplicada (PPCA), agradeço pela parceria, pelas trocas de conhecimento e pelos momentos compartilhados, que tornaram esta caminhada mais leve e inspiradora.

Por fim, aos colegas do Banco do Brasil, sou grato pelo incentivo diário e pelas interações que contribuíram para o amadurecimento tanto profissional quanto acadêmico, refletindo diretamente na qualidade deste trabalho.

A todos que, de alguma forma, contribuíram para a concretização deste projeto, o meu muito obrigado.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES), por meio do Acesso ao Portal de Periódicos.

# Resumo

Esta dissertação apresenta o desenvolvimento de um framework para a aplicação dos níveis de maturidade analítica (Descritivo, Diagnóstico, Preditivo e Prescritivo) na gestão de pessoas, destacando a crescente relevância de uma abordagem orientada a dados em Recursos Humanos. O trabalho parte da premissa de que a adoção de técnicas de análise de dados, combinadas com modelos preditivos e prescritivos, além de ferramentas de *Business Intelligence*, pode apoiar significativamente a tomada de decisões em processos de atração, sucessão e retenção de talentos em ambientes corporativos.

Além de discutir os potenciais benefícios, a dissertação fornece estudos de caso que ilustram a aplicabilidade prática do framework em diferentes contextos organizacionais. Esses estudos de caso demonstram como todos os níveis de maturidade analítica têm capacidade de gerar valor para a organização nos processos relacionados à força de trabalho, conferindo objetividade a decisões que, por vezes, são subjetivas quando o tema é gestão de pessoas. O framework proposto não apenas otimiza processos, mas também promove uma gestão mais eficaz dos recursos humanos, alinhando as práticas de gestão de pessoas com os objetivos estratégicos das organizações.

**Palavras-chave:** *People Analytics, Analytics Maturity Levels, People Management, Business Intelligence, Prescriptive Analytics*

# Abstract

This dissertation presents the development of a framework for applying the levels of analytical maturity (Descriptive, Diagnostic, Predictive, and Prescriptive) in people management, highlighting the growing relevance of a data-driven approach in Human Resources. The work is based on the premise that the adoption of data analysis techniques, combined with predictive and prescriptive models and Business Intelligence tools, can significantly support decision-making processes in the attraction, succession, and retention of talent within corporate environments.

In addition to discussing the potential benefits, the dissertation provides case studies that illustrate the practical applicability of the framework across different organizational contexts. These case studies demonstrate how all levels of analytical maturity have the capacity to generate value for organizations in workforce-related processes, bringing objectivity to decisions that are often subjective in the realm of people management. The proposed framework not only optimizes processes but also promotes more effective human resource management by aligning people management practices with the strategic objectives of organizations.

**Keywords:** People Analytics, Analytics Maturity Levels, People Management, Business Intelligence, Prescriptive Analytics

# Sumário

|          |                                                                                                         |           |
|----------|---------------------------------------------------------------------------------------------------------|-----------|
| <b>1</b> | <b>Introdução</b>                                                                                       | <b>2</b>  |
| 1.1      | Contextualização do Tema . . . . .                                                                      | 2         |
| 1.2      | Justificativa . . . . .                                                                                 | 4         |
| 1.3      | Problema de Pesquisa . . . . .                                                                          | 4         |
| 1.4      | Objetivo Geral . . . . .                                                                                | 5         |
| 1.5      | Objetivos Específicos . . . . .                                                                         | 5         |
| <b>2</b> | <b>Fundamentação Teórica</b>                                                                            | <b>7</b>  |
| 2.1      | Metodologia . . . . .                                                                                   | 7         |
| 2.2      | Modelo de Maturidade do <i>Analytics</i> . . . . .                                                      | 8         |
| 2.2.1    | Evolução dos Níveis de Maturidade . . . . .                                                             | 9         |
| 2.2.2    | Desafios na Evolução dos Níveis de Maturidade . . . . .                                                 | 10        |
| 2.3      | Teoria de Decisão em Gestão de Pessoas . . . . .                                                        | 11        |
| 2.3.1    | Aplicações da Teoria de Decisão em Gestão de Pessoas . . . . .                                          | 12        |
| 2.3.2    | Desafios e Considerações Éticas . . . . .                                                               | 14        |
| 2.4      | Trabalhos Relacionados . . . . .                                                                        | 15        |
| 2.4.1    | Análise Descritiva em Gestão de Pessoas . . . . .                                                       | 15        |
| 2.4.2    | Análise Diagnóstica em Gestão de Pessoas . . . . .                                                      | 16        |
| 2.4.3    | Análise Preditiva em Gestão de Pessoas . . . . .                                                        | 17        |
| 2.4.4    | Análise Prescritiva em Gestão de Pessoas . . . . .                                                      | 18        |
| 2.4.5    | Integração de Estatística, Ciência de Dados e Inteligência Artificial<br>na Gestão de Pessoas . . . . . | 19        |
| 2.4.6    | Aplicações Práticas . . . . .                                                                           | 19        |
| 2.4.7    | Desafios e Considerações Éticas . . . . .                                                               | 20        |
| 2.4.8    | Futuro do <i>People Analytics</i> . . . . .                                                             | 21        |
| <b>3</b> | <b>Revisão da Literatura</b>                                                                            | <b>23</b> |
| 3.0.1    | <i>Conceitos Gerais</i> . . . . .                                                                       | 23        |
| 3.0.2    | <i>Framework I</i> . . . . .                                                                            | 24        |



|          |                                                                        |            |
|----------|------------------------------------------------------------------------|------------|
| 3.0.3    | <i>Framework</i> II . . . . .                                          | 26         |
| 3.0.4    | Considerações quanto aos <i>Frameworks</i> . . . . .                   | 27         |
| <b>4</b> | <b>Metodologia</b>                                                     | <b>28</b>  |
| 4.0.1    | Descrição do Método de Pesquisa . . . . .                              | 28         |
| 4.0.2    | Ferramentas e Técnicas Utilizadas . . . . .                            | 29         |
| 4.0.3    | Procedimentos para Coleta e Análise de Dados . . . . .                 | 30         |
| 4.0.4    | Modelo de <i>Framework</i> . . . . .                                   | 31         |
| <b>5</b> | <b>Desenvolvimento e Análise dos Resultados</b>                        | <b>34</b>  |
| 5.0.1    | Reunião com os Stakeholders e Definição dos Objetivos do Projeto .     | 34         |
| 5.0.2    | Seleção de atributos relevantes e Coleta de Dados . . . . .            | 36         |
| 5.1      | Nível de maturidade: Análise Descritiva . . . . .                      | 37         |
| 5.1.1    | Análise Passiva . . . . .                                              | 37         |
| 5.2      | Nível de maturidade: Análise Diagnóstica . . . . .                     | 39         |
| 5.2.1    | Verificação de Correlações . . . . .                                   | 39         |
| 5.2.2    | Análise das Correlações . . . . .                                      | 43         |
| 5.3      | Nível de maturidade: Análise Preditiva . . . . .                       | 49         |
| 5.3.1    | Caso de uso - Nível Preditivo: Identificação de talentos para sucessão | 50         |
| 5.3.2    | Conclusão do caso de uso preditivo . . . . .                           | 58         |
| 5.4      | Nível de maturidade Prescritivo . . . . .                              | 59         |
| 5.4.1    | Caso de uso - Alocação ideal de colaboradores - Contexto Geral . .     | 60         |
| 5.4.2    | Compreensão do Negócio . . . . .                                       | 61         |
| 5.4.3    | Compreensão dos Dados . . . . .                                        | 62         |
| 5.4.4    | Preparação de Dados . . . . .                                          | 65         |
| 5.4.5    | Avaliação de performance . . . . .                                     | 74         |
| <b>6</b> | <b>Conclusão</b>                                                       | <b>79</b>  |
| <b>7</b> | <b>Anexos</b>                                                          | <b>81</b>  |
| A        | Correlações Lineares . . . . .                                         | 81         |
| B        | Lista de Atributos . . . . .                                           | 95         |
| C        | Lista de Atributos . . . . .                                           | 97         |
|          | <b>Referências</b>                                                     | <b>101</b> |

# Lista de Figuras

|      |                                                                                                    |    |
|------|----------------------------------------------------------------------------------------------------|----|
| 4.1  | <i>Framework</i> - Fonte: O autor . . . . .                                                        | 33 |
| 5.1  | Painel de Gênero - Fonte: O autor . . . . .                                                        | 38 |
| 5.2  | Painel de Desligamentos - Fonte: O autor . . . . .                                                 | 39 |
| 5.3  | Painel Geracional - Fonte: O autor . . . . .                                                       | 40 |
| 5.4  | Exemplos de gráficos de dispersão para diferentes tipos de correlação. Fonte:<br>O autor . . . . . | 42 |
| 5.5  | Tempo de empresa x Salário - Fonte: O autor . . . . .                                              | 43 |
| 5.6  | Idade x Total de Certificações - Fonte: O autor . . . . .                                          | 44 |
| 5.7  | Pesquisa Clima x Dias não úteis trabalhados - Fonte: O autor . . . . .                             | 45 |
| 5.8  | Média Clima Organizacional x Dimensões Clima Organizacional - Fonte:<br>O autor . . . . .          | 46 |
| 5.9  | Idade x Dimensões Desempenho: - Fonte: O autor . . . . .                                           | 47 |
| 5.10 | Média Desempenho x Horas Extras: - Fonte: O autor . . . . .                                        | 48 |
| 5.11 | Atributos . . . . .                                                                                | 52 |
| 5.12 | Valores faltantes - Fonte: O autor . . . . .                                                       | 53 |
| 5.13 | Level de escolaridade - Fonte: O autor . . . . .                                                   | 53 |
| 5.14 | Variáveis Dummy - Fonte: O autor . . . . .                                                         | 54 |
| 5.15 | Balanceamento de classes - Fonte: O autor . . . . .                                                | 54 |
| 5.16 | Base de treino e teste - Fonte: O autor . . . . .                                                  | 55 |
| 5.17 | X e Y - Fonte: O autor . . . . .                                                                   | 55 |
| 5.18 | Recall - Fonte: O autor . . . . .                                                                  | 56 |
| 5.19 | Especificidade - Fonte: O autor . . . . .                                                          | 56 |
| 5.20 | Matriz de confusão - Fonte: O autor . . . . .                                                      | 57 |
| 5.21 | Attribute weights - Fonte: O autor . . . . .                                                       | 58 |
| 5.22 | Bases Necessárias - Fonte: O autor . . . . .                                                       | 63 |
| 5.23 | Modelo Geral - Fonte: O autor . . . . .                                                            | 66 |
| 5.24 | Exemplo de grafos - Fonte: O autor . . . . .                                                       | 70 |
| 5.25 | Vértices e Arestas - Fonte: O autor . . . . .                                                      | 71 |
| 5.26 | Commits entre grupos - Fonte: O autor . . . . .                                                    | 77 |

# Lista de Tabelas

|     |                                                                                |    |
|-----|--------------------------------------------------------------------------------|----|
| 3.2 | Principais Categorias do <i>People Analytics Effectiveness Wheel</i> . . . . . | 26 |
| 4.2 | Etapas do Processo CRISP-DM . . . . .                                          | 29 |
| 5.1 | Tabela de Macrotópicos em Ordem Alfabética . . . . .                           | 63 |
| 5.2 | Tabela de Experiências . . . . .                                               | 64 |
| A.1 | Correlação entre Variáveis . . . . .                                           | 81 |

# Lista de Acrônimos

**AI** *Artificial Intelligence* (Inteligência Artificial)

**BI** *Business Intelligence*

**CRISP-DM** *Cross-Industry Standard Process for Data Mining*

**HR** *Human Resources* (Recursos Humanos)

**IT** *Information Technology* (Tecnologia da Informação)

**KSAOs** *Knowledge, Skills, Abilities, and Other Characteristics* (Conhecimentos, Habilidades, Capacidades e Outras Características)

**RH** Recursos Humanos

**TI** Tecnologia da Informação

# Capítulo 1

## Introdução

### 1.1 Contextualização do Tema

Nos últimos anos, as áreas de negócios têm adotado extensivamente técnicas de estatística, mineração de dados, *big data* e aprendizagem de máquina para o reconhecimento de padrões, modelagem preditiva e outros objetivos, na busca da melhoria dos processos. A vasta disponibilidade de dados e o desenvolvimento de algoritmos cada vez mais avançados estão permitindo análises mais amplas e assertivas, com foco no apoio da tomada de decisão orientada a dados [6]. E nesse contexto a ciência de dados tem um papel central, permitindo que as organizações obtenham *insights* valiosos. A ciência de dados se inter-relaciona intimamente com conceitos como *big data* e negócio, formando um tecido conectivo entre as tecnologias de processamento de dados e a capacidade de tomar decisões orientada a dados [45]. A aplicação dessas tecnologias tem proporcionado grandes vantagens competitivas, impulsionando a eficiência operacional e a inovação. Empresas de diversos setores, desde o varejo até telecomunicações e finanças, utilizam a ciência de dados para otimizar processos, prever comportamentos de clientes e gerenciar recursos de forma mais eficiente. A capacidade de extrair conhecimento e informações úteis dos dados, fortalece a tomada de decisão e garante que as análises estejam alinhadas aos objetivos estratégicos das organizações [45]. Dessa forma, a adoção de uma abordagem baseada em dados não apenas melhora a precisão das decisões, mas também oferece uma base sólida para a inovação contínua e a adaptação às mudanças do mercado.

Entretanto, quando o foco se volta para a Gestão de Pessoas, observa-se um cenário diferente. Apesar do potencial transformador das técnicas de análise de dados, muitas áreas de recursos humanos ainda estão em estágios iniciais de adoção dessas tecnologias [1]. Em vez de se basearem em dados concretos, muitas decisões relacionadas à gestão de pessoas são tomadas com base em percepções subjetivas e achismos [46]. Em 1620, Francis Bacon argumentava que o conhecimento científico não deveria ser baseado apenas em

hipóteses intuitivas, segundo ele essa abordagem é limitada por estabelecer uma premissa antes de um experimento. Bacon defendia uma metodologia empírica, onde a observação e a experimentação desempenham papéis centrais na formação de conhecimento confiável [48].

O *People Analytics* visa preencher essa lacuna, proporcionando uma abordagem baseada em dados para a gestão de pessoas. Esta abordagem ainda permanece em seus estágios iniciais nas organizações; por essa razão, este estudo elenca quatro componentes principais que levam ao baixo nível de implementação: maturidade, *mindset*, organização e competências [1]. O componente maturidade do *People Analytics* possui como limitação a qualidade dos softwares utilizados no setor de Recursos Humanos, levando a dados sem integridade ou disponibilidade [8]. O componente *mindset* reflete a necessidade de pensamento estratégico, onde muitas vezes falta dentre os profissionais de RH a *soft-skills* para transformar dados em conhecimento acionável e estratégico somado à falta de conhecimento do tema pelo *board* da corporação. O componente organização é outro fator crítico, pois as funções de *People Analytics* frequentemente são exclusivas do setor de RH, sendo que o ideal talvez seja que este conceito esteja horizontalmente na organização, promovendo debates sobre uma gestão de pessoas orientada a dados para fora do RH. Finalmente, as competências necessárias para uma análise eficaz de dados de RH são variadas e complexas, incluindo habilidades em estatísticas, gerenciamento de dados, storytelling, técnicas de visualização de dados e entendimento profundo do negócio. A ausência dessas competências variadas em uma equipe coesa dificulta o progresso do *People Analytics* nas corporações [1].

Ao aplicar técnicas avançadas de análise de dados, estatística e inteligência artificial, é possível promover decisões orientadas a dados, mesmo em uma área subjetiva como a gestão de pessoas. Do ponto de vista operacional, a busca deve ser por analisar a atividade das pessoas e a utilização dos recursos para o bem da organização. Em suma, se mede o valor das pessoas, se mede a eficiência, eficácia e resultados de seus processos. É necessário saber quanto custa um processo, quanto tempo ele demanda, qual resultado se obtém de um determinado insumo, qual é a qualidade do resultado e como as pessoas se sentem a respeito. A análise permite a mensuração dos processos e apoia a tomada de decisões [22].

Este trabalho, portanto, busca criar um *framework* aplicável em organizações que possuam dados de pessoas com disponibilidade e integridade, de forma a explorar possibilidades, exemplificar com aplicações práticas a viabilidade e os benefícios do uso de soluções estatísticas, de ciência de dados e inteligência artificial na gestão de pessoas, contribuindo para a evolução da área de Recursos Humanos rumo a uma prática mais científica e fundamentada em dados.

## 1.2 Justificativa

A escolha do tema justifica-se pela crescente necessidade de modernizar e tornar mais eficientes os processos de gestão de pessoas, um desafio que se intensifica em um cenário corporativo cada vez mais competitivo e orientado por dados. As organizações enfrentam demandas crescentes por decisões rápidas e precisas, que requerem a integração de técnicas avançadas de análise de dados na gestão de recursos humanos. Neste contexto, a capacidade de adotar abordagens baseadas em dados não apenas potencializa a eficiência operacional, mas também se torna um diferencial estratégico para a sobrevivência e crescimento das empresas [27].

A relevância deste estudo reside em sua contribuição direta para a área de *People Analytics*, um campo emergente que combina ciência de dados, estatística e inteligência artificial para transformar dados de pessoas em *insights* acionáveis. Além disso, o trabalho aborda a lacuna entre o uso intuitivo de dados e a implementação sistemática de níveis de maturidade analítica (descritivo, diagnóstico, preditivo e prescritivo) em organizações, oferecendo uma estrutura prática e acessível.

As contribuições deste trabalho incluem: (1) o desenvolvimento de um *framework* que orienta organizações na aplicação dos quatro níveis de maturidade analítica na gestão de pessoas, (2) a exemplificação prática do uso de técnicas de análise de dados em estudos de caso reais, demonstrando sua viabilidade e benefícios, e (3) o incentivo à transformação cultural e tecnológica em Recursos Humanos, promovendo uma tomada de decisão mais fundamentada e alinhada aos objetivos estratégicos das organizações.

## 1.3 Problema de Pesquisa

O problema de pesquisa que este trabalho se propõe a investigar é: Como estruturar um *framework* em *People Analytics* que seja aplicável a organizações de diferentes setores e contextos, considerando suas especificidades organizacionais e práticas de gestão de pessoas, e que atenda aos quatro níveis de maturidade em *Analytics*, promovendo uma tomada de decisão mais fundamentada e orientada por dados na gestão de pessoas?

Este trabalho defende que, apesar das variações entre tipos de organizações — como setor, porte, cultura organizacional e níveis de maturidade tecnológica —, é possível desenvolver um *framework* genérico e flexível que, com ajustes específicos, se adapte a essas realidades. Para tanto, o *framework* proposto baseia-se em princípios universais de análise de dados e gestão estratégica de pessoas, estruturados de forma modular, permitindo sua implementação gradual e customizável conforme as necessidades e características de cada organização.

Além disso, busca potencializar a tomada de decisão através do fortalecimento da capacidade organizacional de utilizar dados para suportar decisões mais informadas, precisas e alinhadas aos objetivos estratégicos. O *framework* não apenas melhora a qualidade das informações disponíveis, mas também promove a integração de diferentes níveis de maturidade analítica (descritivo, diagnóstico, preditivo e prescritivo), assegurando que as decisões em gestão de pessoas sejam tomadas com maior embasamento técnico e estratégico.

## 1.4 Objetivo Geral

Desenvolver um *framework* prático para a aplicação dos diferentes níveis de maturidade em *People Analytics* (Análise Descritiva, Análise Diagnóstica, Análise Preditiva e Análise Prescritiva) [35], com o objetivo de oferecer uma abordagem estruturada que auxilie organizações a aprimorarem a gestão de pessoas por meio de decisões mais informadas e orientadas por dados. O *framework* será exemplificado por meio de casos práticos baseados em dados de profissionais de tecnologia de um grande banco brasileiro, demonstrando sua aplicabilidade e os benefícios de sua utilização.

## 1.5 Objetivos Específicos

1. Demonstrar a viabilidade do uso da estatística, da ciência de dados e da inteligência artificial no apoio à tomada de decisões nos processos de gestão de pessoas.
2. Demonstrar a viabilidade para se transitar no nível de maturidade Análise Descritiva em processos de gestão de pessoas, através da construção de *dashboards* com métricas de recursos humanos, consultadas em tempo real pelos *stakeholders*.
3. Demonstrar a viabilidade para se transitar no nível de maturidade Análise Diagnóstica em processos de gestão de pessoas, através da exploração com técnicas de correlação e por comparativos de técnicas de amostragem.
4. Demonstrar a viabilidade para se transitar no nível de maturidade Análise Preditiva, através da utilização de técnicas de machine learning na criação de modelo preditivo para processos de gestão de pessoas.
5. Demonstrar a viabilidade para se transitar no nível de maturidade Análise Prescritiva em processos de gestão de pessoas, através da utilização de *solvers* para alocação ótima de novos funcionários, comparando diferentes métodos de otimização.



Por fim, este capítulo apresentou a contextualização do tema, justificativa, problema de pesquisa e os objetivos deste trabalho, destacando a relevância da aplicação de técnicas avançadas de análise de dados na gestão de pessoas. A estrutura do documento é organizada da seguinte forma: no Capítulo 2 é abordado o embasamento teórico diretamente ligado ao problema de pesquisa, no Capítulo 3 é realizada uma revisão da literatura, abordando os principais conceitos, teorias e trabalhos relacionados ao tema *People Analytics* e níveis de maturidade analítica. No Capítulo 4, é detalhada a metodologia adotada para o desenvolvimento do *framework* proposto. O Capítulo 5 apresenta o desenvolvimento do *framework*, incluindo os estudos de caso que ilustram sua aplicabilidade prática e seus resultados. Por fim, no Capítulo 6, são discutidos os resultados obtidos, as limitações do estudo e as sugestões para trabalhos futuros.

Com essa estrutura, busca-se construir uma base sólida para a compreensão do tema, ao mesmo tempo que se evidencia a aplicabilidade prática e as contribuições do *framework* proposto. O próximo capítulo explora a fundamentação teórica e a revisão de literatura, que sustentam os conceitos e estratégias apresentados neste trabalho.

# Capítulo 2

## Fundamentação Teórica

### 2.1 Metodologia

O desenvolvimento de projetos baseados em análise de dados requer uma abordagem estruturada que permita transformar dados brutos em informações úteis e acionáveis. Nesse contexto, o *Cross-Industry Standard Process for Data Mining* (CRISP-DM) é amplamente reconhecido como um dos frameworks mais robustos e versáteis para conduzir projetos de mineração de dados. Proposto por Chapman et al. (2000), o CRISP-DM foi projetado como um padrão metodológico flexível, aplicável a diferentes indústrias e domínios, proporcionando uma estrutura consistente para guiar o processo de extração de conhecimento a partir de dados [10].

O CRISP-DM organiza o processo analítico em seis etapas principais: (i) **compreensão do negócio**, onde os objetivos organizacionais são definidos e traduzidos em problemas analíticos; (ii) **compreensão dos dados**, que envolve a coleta e análise inicial para avaliar a qualidade e relevância dos dados disponíveis; (iii) **preparação dos dados**, que consiste na limpeza, transformação e integração dos dados para adequação às técnicas analíticas; (iv) **modelagem**, que inclui a aplicação de algoritmos de aprendizado de máquina e estatística para identificar padrões e gerar previsões; (v) **avaliação**, etapa em que os modelos desenvolvidos são validados quanto à sua eficácia e aplicabilidade ao problema de negócio; e (vi) **implantação**, em que as soluções analíticas são implementadas no ambiente organizacional e os resultados são monitorados.

O caráter iterativo do CRISP-DM é um de seus principais diferenciais, permitindo que os analistas retornem a etapas anteriores quando necessário, a fim de garantir que as soluções desenvolvidas estejam alinhadas aos objetivos de negócio e que os resultados sejam relevantes para a tomada de decisão. No campo de *People Analytics*, a aplicação do CRISP-DM tem se mostrado especialmente relevante, uma vez que os dados relacionados a recursos humanos apresentam desafios como heterogeneidade, confidencialidade

e variabilidade temporal. A estrutura proposta pelo CRISP-DM fornece um guia claro e replicável para superar esses desafios e extrair insights que auxiliem na gestão estratégica de pessoas.

Neste trabalho, o CRISP-DM será utilizado como referencial teórico para planejar e conduzir as análises propostas, servindo como um arcabouço para integrar conceitos de ciência de dados com práticas organizacionais. Além disso, sua flexibilidade permitirá a adaptação dos processos às particularidades dos dados organizacionais e aos objetivos específicos da pesquisa [10].

## 2.2 Modelo de Maturidade do *Analytics*

O modelo de maturidade do Analytics geralmente é dividido em quatro níveis principais: Descritivo, Diagnóstico, Preditivo e Prescritivo.

**Análise Descritiva:** Este nível envolve a coleta e descrição de dados históricos para entender o que aconteceu no passado. Ferramentas de Business Intelligence (BI) são amplamente utilizadas nesta fase para criar relatórios e dashboards que oferecem uma visão clara e compreensível das operações da organização [14].

**Análise Diagnóstica:** Neste nível, as organizações começam a investigar o motivo dos eventos ocorridos. Técnicas estatísticas são aplicadas para identificar causas e efeitos, permitindo uma compreensão mais profunda dos dados e dos fatores que influenciam os resultados [15].

**Análise Preditiva:** Utilizando técnicas de Machine Learning e modelos estatísticos avançados, este nível foca na previsão de eventos futuros. A análise preditiva permite às organizações antecipar tendências e comportamentos, proporcionando uma vantagem competitiva ao preparar-se para futuros desafios [24]. Aplicar modelos estatísticos preditivos aos dados de RH requer conhecimento de estatísticas e a capacidade (e experiência) para entender e interpretar os significados dos resultados das análises. Atualmente, poucas funções de RH utilizam efetivamente as técnicas analíticas estatísticas disponíveis. Embora as equipes de métricas e analytics de RH processem e relatem grandes quantidades de dados relacionados a pessoas, poucas aplicam técnicas estatísticas que permitam fazer inferências preditivas. [17]

**Análise Prescritiva:** O nível mais avançado, onde as organizações utilizam algoritmos de otimização e simulação para sugerir ações que maximizem os resultados desejados. A análise prescritiva não só prevê o que pode acontecer, mas também recomenda as melhores ações a serem tomadas [24].

### 2.2.1 Evolução dos Níveis de Maturidade

A evolução dos níveis de maturidade do Analytics pode ser traçada desde a simples coleta e relatórios de dados até a aplicação de técnicas sofisticadas de inteligência artificial. Inicialmente, as organizações concentram-se em coletar dados e apresentá-los de maneira descritiva, oferecendo uma visão clara do que está acontecendo no ambiente organizacional. Esse primeiro passo, embora básico, é essencial para construir uma base sólida sobre a qual as análises mais complexas serão desenvolvidas.

À medida que a organização avança, o próximo nível envolve a análise diagnóstica, que vai além de simplesmente descrever os dados para investigar as causas subjacentes aos padrões observados. Nesse estágio, a organização começa a entender o 'porquê' por trás dos dados, identificando os fatores que influenciam os resultados e possibilitando uma compreensão mais profunda do contexto organizacional. Esse entendimento é crucial para que as empresas possam abordar problemas e oportunidades de maneira mais eficaz.

O terceiro nível de maturidade é o preditivo, onde as organizações utilizam os dados e *insights* diagnósticos para prever futuras tendências e comportamentos. Neste estágio, técnicas como modelagem estatística e machine learning são empregadas para antecipar eventos, como a rotatividade de funcionários ou a necessidade de novas contratações. A análise preditiva permite que as organizações sejam mais proativas, preparando-se para desafios futuros antes que eles se manifestem plenamente.

Finalmente, as organizações mais maduras atingem o nível prescritivo, onde os *insights* preditivos são usados para recomendar ou até automatizar decisões. Nesse estágio, ferramentas avançadas de inteligência artificial e otimização são aplicadas para sugerir ações específicas que maximizem os resultados desejados, como a melhor alocação de recursos humanos ou a implementação de programas de treinamento personalizados.

No entanto, é fundamental ressaltar a importância de não pular etapas neste processo de evolução. Cada nível de maturidade serve como um alicerce para o próximo. Saltar diretamente de um estágio descritivo para o prescritivo, sem passar pelo diagnóstico e preditivo, pode levar a decisões mal fundamentadas e, potencialmente, prejudiciais. Por exemplo, sem uma análise diagnóstica adequada, as causas subjacentes dos padrões observados podem ser mal interpretadas, levando a previsões imprecisas e recomendações equivocadas. Da mesma forma, sem uma base preditiva sólida, as recomendações prescritivas podem carecer da precisão necessária para serem eficazes, resultando em ações que não otimizam os resultados ou, pior, causam danos à organização.

Portanto, a evolução dos níveis de maturidade analítica deve ser vista como um processo progressivo e interdependente, onde cada estágio constrói sobre o anterior, permitindo que as organizações desenvolvam uma capacidade analítica robusta e confiável. O desenvolvimento gradual e ordenado desses níveis de maturidade assegura que as orga-

nizações possam tomar decisões cada vez mais informadas, precisas e estratégicas, maximizando o valor dos dados e evitando os riscos associados a uma adoção precipitada de técnicas avançadas sem a devida fundamentação.

### 2.2.2 Desafios na Evolução dos Níveis de Maturidade

A transição entre os níveis de maturidade analítica é uma jornada complexa que apresenta desafios significativos, exigindo atenção cuidadosa a vários fatores críticos:

**Investimento em Tecnologia:** À medida que as organizações avançam para níveis mais altos de maturidade analítica, a necessidade de ferramentas e tecnologias avançadas torna-se cada vez mais premente. Essas tecnologias não apenas suportam análises sofisticadas, mas também integram dados de diferentes fontes, automatizam processos analíticos e permitem a aplicação de técnicas como machine learning e inteligência artificial. Investir em tecnologia não é apenas uma questão de adquirir novos sistemas, mas também de garantir que as infraestruturas de TI sejam escaláveis, seguras e capazes de lidar com a crescente complexidade dos dados e das análises.

**Desenvolvimento de Competências:** A evolução para níveis mais avançados de maturidade analítica depende da disponibilidade de profissionais altamente qualificados em ciência de dados, estatística e análise. Esses profissionais são essenciais para interpretar os dados de maneira significativa e transformar *insights* analíticos em decisões estratégicas. No entanto, a escassez de talentos nessas áreas pode representar um desafio significativo. As organizações precisam investir no desenvolvimento contínuo de competências, tanto por meio da contratação de especialistas quanto pela capacitação de sua força de trabalho existente. Programas de treinamento, parcerias com instituições educacionais e o incentivo à aprendizagem contínua são estratégias fundamentais para suprir essa demanda.

**Gestão de Mudanças:** A adoção de uma cultura orientada por dados exige mudanças organizacionais e comportamentais profundas. Isso inclui não apenas a implementação de novas tecnologias e processos, mas também a transformação da mentalidade dos colaboradores e líderes, promovendo uma abordagem mais analítica e baseada em evidências na tomada de decisões. A gestão eficaz dessas mudanças é crucial para o sucesso da evolução nos níveis de maturidade analítica. Resistências naturais à mudança, como a relutância em abandonar práticas tradicionais ou o ceticismo em relação à análise de dados, precisam ser abordadas de maneira proativa. Isso pode ser feito por meio de programas de comunicação, treinamento e liderança eficaz que inspirem confiança no valor da análise de dados e incentivem a adesão a novas práticas.

A transição bem-sucedida entre os níveis de maturidade analítica requer uma abordagem integrada que combine investimentos em tecnologia, desenvolvimento de competências e gestão de mudanças organizacionais. Ao enfrentar esses desafios com uma estratégia

bem planejada, as organizações podem garantir que suas capacidades analíticas evoluam de maneira robusta, permitindo decisões mais informadas e impulsionando o sucesso no longo prazo.

## 2.3 Teoria de Decisão em Gestão de Pessoas

A Teoria de Decisão é uma área multidisciplinar que combina elementos de estatística, psicologia, economia e gestão para estudar como decisões são tomadas e como podem ser melhoradas. Em gestão de pessoas, a Teoria de Decisão oferece ferramentas e *Frameworks* para melhorar a tomada de decisões estratégicas relacionadas à gestão de talentos, recrutamento, retenção e desenvolvimento de funcionários.

A Teoria de Decisão é uma área interdisciplinar que abrange diversos conceitos fundamentais, essenciais para entender como as decisões são tomadas, tanto em contextos individuais quanto organizacionais. Esses conceitos não só oferecem uma base teórica sólida, mas também têm aplicações práticas significativas em campos como economia, gestão, psicologia e ciência de dados.

**Decisões Racionais:** A premissa básica da decisão racional é que os tomadores de decisão são agentes racionais que fazem escolhas que maximizam a utilidade esperada. Este conceito é amplamente utilizado em modelos econômicos e de gestão para prever o comportamento humano em diferentes cenários. A ideia de racionalidade, no entanto, não é isenta de críticas; há debates contínuos sobre as limitações da racionalidade humana, especialmente em situações complexas onde as informações são incompletas ou as emoções desempenham um papel significativo [32]. Apesar dessas críticas, o modelo de decisão racional permanece um pilar central na análise econômica e em muitas estratégias de tomada de decisão corporativa.

**Teoria da Utilidade:** A Teoria da Utilidade propõe que as decisões podem ser avaliadas com base na utilidade esperada, onde a utilidade é uma medida subjetiva de satisfação ou benefício que um tomador de decisão obtém de um determinado resultado. Este conceito é fundamental na economia e na teoria da escolha racional, fornecendo uma estrutura para a tomada de decisões que leva em conta tanto a probabilidade quanto o valor dos resultados potenciais. A Teoria da Utilidade Esperada, desenvolvida por Von Neumann e Morgenstern, formalizou essa abordagem, oferecendo uma base matemática para decisões sob risco e incerteza [53]. Embora amplamente aceito, este modelo também enfrenta desafios, particularmente em situações onde as preferências dos indivíduos não seguem os axiomas da utilidade esperada, como evidenciado em experimentos de economia comportamental.

**Teoria da Decisão sob Incerteza:** Este conceito foca em como os tomadores de decisão lidam com a incerteza e o risco ao fazer escolhas. Em muitos casos, as decisões devem ser feitas com base em informações incompletas ou ambíguas, onde os resultados futuros não podem ser previstos com certeza. A Teoria da Decisão sob Incerteza explora diferentes estratégias que os indivíduos e organizações podem adotar nessas situações, como o uso de árvores de decisão, análise de cenários e a regra de minimax [49]. Essas ferramentas permitem que os tomadores de decisão modelem as possíveis consequências de suas escolhas e identifiquem a opção que minimiza o risco ou maximiza o ganho esperado. Entretanto, a eficácia dessas ferramentas depende da qualidade das informações disponíveis e da capacidade do tomador de decisão de avaliar corretamente as probabilidades associadas a diferentes cenários.

**Processos Heurísticos:** Ao contrário das teorias que assumem racionalidade completa, os processos heurísticos exploram como as pessoas realmente tomam decisões no mundo real. Heurísticas são atalhos mentais que simplificam o processo de tomada de decisão, permitindo que as pessoas façam escolhas rápidas e geralmente eficazes sem realizar uma análise completa de todas as opções possíveis. No entanto, essas heurísticas podem levar a vieses e erros sistemáticos, como o viés de confirmação, a ancoragem e a aversão à perda [50]. O estudo desses processos é crucial para entender as limitações da racionalidade humana e para desenvolver intervenções que ajudem a mitigar os efeitos desses vieses na tomada de decisões.

Esses conceitos, quando combinados, oferecem uma visão abrangente da Teoria de Decisão, destacando tanto os métodos ideais de tomada de decisão quanto as realidades práticas enfrentadas pelos tomadores de decisão no mundo real. Ao compreender essas nuances, os gestores e analistas podem aplicar essas teorias de forma mais eficaz, adaptando-as ao contexto específico de suas organizações e às características individuais de seus colaboradores.

### **2.3.1 Aplicações da Teoria de Decisão em Gestão de Pessoas**

A Teoria de Decisão oferece uma estrutura robusta para abordar diversos desafios na gestão de pessoas, permitindo que as organizações tomem decisões mais informadas e estratégicas. Suas aplicações práticas são amplas e cobrem desde a seleção de talentos até o desenvolvimento e retenção de funcionários, sempre com o objetivo de alinhar as necessidades organizacionais com os interesses e habilidades dos colaboradores.

**Recrutamento e Seleção:** No contexto de recrutamento e seleção, a Teoria de Decisão pode ser utilizada para avaliar candidatos de forma mais holística e objetiva, integrando múltiplos critérios em um modelo de decisão estruturado. Isso permite que os gestores ponderem fatores como habilidades técnicas, experiência profissional e adequação cultural

de maneira sistemática, reduzindo o impacto de vieses inconscientes no processo de seleção. Modelos de decisão multicritério, como o método de análise hierárquica (AHP), podem ser especialmente úteis ao permitir a comparação direta entre candidatos com diferentes perfis [30]. Além disso, técnicas de decisão preditiva podem ser aplicadas para identificar candidatos com maior probabilidade de sucesso a longo prazo na organização, utilizando dados históricos de desempenho e progressão de carreira.

**Desenvolvimento de Carreira:** As decisões sobre programas de desenvolvimento e treinamento também podem ser substancialmente aprimoradas por meio da aplicação da Teoria de Decisão. Ao utilizar modelos de decisão, os gestores podem avaliar diferentes opções de desenvolvimento considerando tanto os benefícios esperados para os funcionários quanto o retorno sobre o investimento para a organização. Por exemplo, ao decidir entre várias opções de treinamento, a análise pode considerar fatores como o aumento esperado de competências, a relevância para os objetivos estratégicos da empresa e o custo de oportunidade [25]. Essa abordagem não só otimiza o processo de desenvolvimento de carreira, mas também garante que os recursos sejam alocados de forma eficiente, promovendo um crescimento sustentável tanto para os colaboradores quanto para a organização.

**Planejamento de Sucessão:** O planejamento de sucessão é uma área crítica onde a Teoria de Decisão desempenha um papel fundamental. Identificar futuros líderes e preparar planos de sucessão eficazes requer uma avaliação cuidadosa de diversos fatores, como o potencial de liderança, a prontidão dos candidatos e as necessidades futuras da organização. A Teoria de Decisão permite que os gestores comparem diferentes estratégias de sucessão, considerando o impacto de cada uma no longo prazo [47]. Modelos como a análise de decisão multicritério e a teoria da utilidade esperada podem ser aplicados para avaliar as opções disponíveis e selecionar a estratégia que maximiza o alinhamento entre as necessidades organizacionais e o desenvolvimento individual dos talentos internos. Além disso, a aplicação dessas técnicas ajuda a mitigar riscos associados à perda de líderes chave, garantindo uma transição suave e sustentada.

**Retenção de Talentos:** Retenção de talentos é outro domínio onde a Teoria de Decisão pode ser aplicada com grande eficácia. Modelos de decisão podem ser utilizados para identificar os principais fatores que influenciam a retenção de funcionários, como satisfação no trabalho, oportunidades de crescimento e equilíbrio entre vida pessoal e profissional. Ao analisar esses fatores, as organizações podem desenvolver políticas e programas direcionados para aumentar a satisfação dos colaboradores e reduzir a rotatividade [29]. Além disso, a Teoria de Decisão pode auxiliar na segmentação de grupos de funcionários com diferentes níveis de risco de saída, permitindo a implementação de estratégias de retenção personalizadas. Isso não apenas ajuda a manter os talentos críticos dentro da organização, mas também contribui para a construção de um ambiente de trabalho mais positivo.



e engajador.

O futuro da Teoria de Decisão em gestão de pessoas está intimamente ligado aos avanços em tecnologia e análise de dados. Ferramentas de inteligência artificial e machine learning oferecem novas oportunidades para melhorar a precisão e eficácia das decisões. Além disso, a crescente ênfase na ética e responsabilidade social corporativa destaca a importância de decisões transparentes e justas [28].

Em resumo, a Teoria de Decisão fornece uma base sólida para enfrentar os desafios da gestão de pessoas, oferecendo ferramentas analíticas que melhoram a qualidade das decisões em várias áreas críticas. Ao aplicar esses conceitos, as organizações podem não apenas otimizar seus processos de gestão de talentos, mas também garantir que suas decisões estejam alinhadas com seus objetivos estratégicos, promovendo um ambiente de trabalho mais eficiente e satisfatório para todos os envolvidos.

### **2.3.2 Desafios e Considerações Éticas**

Embora a aplicação da Teoria de Decisão em gestão de pessoas traga inúmeros benefícios, ela também apresenta desafios significativos e questões éticas que precisam ser cuidadosamente considerados para garantir que as decisões tomadas sejam não apenas eficazes, mas também justas e responsáveis.

**Complexidade das Decisões:** As decisões em gestão de pessoas são inerentemente complexas, frequentemente envolvendo múltiplos stakeholders com interesses e prioridades diversas. Essa complexidade torna a modelagem e a análise das decisões um desafio considerável. A gestão de pessoas lida com variáveis quantitativas e qualitativas, como habilidades técnicas, potencial de liderança, satisfação no trabalho e alinhamento cultural. Modelar essas variáveis de forma precisa e integrá-las em um processo de decisão coerente requer uma abordagem multifacetada que reconheça a complexidade das interações humanas dentro das organizações [5]. Além disso, a natureza dinâmica dos ambientes organizacionais, onde as condições e os contextos podem mudar rapidamente, exige que os modelos de decisão sejam flexíveis e adaptáveis, o que pode aumentar ainda mais a complexidade do processo.

**Vieses e Heurísticas:** Outro desafio significativo na aplicação da Teoria de Decisão em gestão de pessoas é a propensão dos tomadores de decisão humanos a vieses cognitivos e heurísticas. Esses vieses podem distorcer a análise racional e levar a decisões subótimas ou até prejudiciais. Por exemplo, o viés de confirmação pode levar os gestores a buscar e interpretar informações de maneira que confirme suas crenças pré-existentes, enquanto o viés de ancoragem pode fazer com que se fixem em uma informação inicial ao avaliar opções subsequentes [3]. Mitigar esses vieses exige conscientização e a implementação de processos estruturados de tomada de decisão que incorporem salvaguardas, como revisões

por pares, análise de dados objetiva e o uso de algoritmos de apoio à decisão que podem neutralizar alguns desses vieses. No entanto, é importante lembrar que, embora os modelos e algoritmos possam reduzir a influência dos vieses, eles próprios podem ser influenciados por vieses nos dados de treinamento ou nas suposições subjacentes, exigindo uma atenção contínua à justiça e à precisão.

**Ética e Justiça:** As considerações éticas são centrais na aplicação da Teoria de Decisão em gestão de pessoas, especialmente porque essas decisões afetam diretamente a vida e o bem-estar dos funcionários. As decisões devem ser tomadas de forma a respeitar os princípios de transparência, equidade e respeito pelos direitos individuais dos trabalhadores [13]. Isso significa que as organizações devem garantir que os processos de decisão sejam justos, tanto em termos de processo quanto de resultado, e que todas as partes interessadas sejam tratadas com dignidade e respeito. A transparência é crucial para que os funcionários compreendam os critérios e processos que influenciam as decisões que os afetam, o que pode ajudar a construir confiança e legitimidade. Além disso, as considerações éticas também envolvem a responsabilidade de minimizar qualquer potencial dano que possa surgir das decisões, como discriminação ou exclusão, e de promover um ambiente de trabalho inclusivo e justo.

Por fim, à medida que as organizações continuam a adotar e integrar a Teoria de Decisão em suas práticas de gestão de pessoas, é essencial que elas desenvolvam uma compreensão profunda dos desafios e considerações éticas envolvidas. Isso não apenas garante que as decisões tomadas sejam eficazes e informadas, mas também que reflitam os valores e princípios fundamentais da organização, promovendo um ambiente de trabalho mais justo, ético e produtivo para todos.

## **2.4 Trabalhos Relacionados**

A literatura existente oferece uma base sólida para a compreensão e implementação dessas técnicas em diferentes contextos organizacionais.

### **2.4.1 Análise Descritiva em Gestão de Pessoas**

A Análise Descritiva representa o primeiro nível de maturidade no campo do analytics, concentrando-se na coleta, organização e apresentação de dados históricos para fornecer uma compreensão clara do que aconteceu no passado. Este nível de análise é fundamental para estabelecer uma base sólida sobre a qual decisões mais complexas e informadas podem ser tomadas.

Diversos estudos têm demonstrado a eficácia da Análise Descritiva na gestão de recursos humanos e veremos na prática essa aplicação descritiva na sessão de Desenvolvimento

desta dissertação. Conforme destaca [21], a importância de métricas e relatórios detalhados como ferramentas essenciais para monitorar o desempenho dos funcionários, identificar áreas de melhoria e apoiar a gestão no reconhecimento de padrões e tendências que podem impactar a organização. Através da análise descritiva, os gestores podem obter *insights* valiosos sobre aspectos como a rotatividade de funcionários, a assiduidade e o desempenho em diferentes departamentos, permitindo uma visão mais estruturada e objetiva da força de trabalho.

No contexto de TI, a utilização de dashboards interativos se torna um recurso poderoso ao permitir que *stakeholders* acessem informações em tempo real, facilitando a visualização de dados e a rápida identificação de anomalias ou oportunidades de melhoria [27]. Esses dashboards não apenas agregam valor ao processo de tomada de decisão ao fornecer uma representação visual clara e imediata dos dados, mas também promovem uma cultura de transparência e responsabilidade ao tornar as informações acessíveis a todos os níveis da organização. A capacidade de interagir com os dados de forma dinâmica permite que os gestores ajustem rapidamente suas estratégias, respondendo de maneira ágil às mudanças nas condições organizacionais.

A Análise Descritiva, portanto, não só oferece uma compreensão detalhada do que ocorreu no passado, mas também estabelece as bases para a evolução para níveis mais avançados de maturidade analítica, onde a previsibilidade e a prescrição se tornam possíveis. Ao dominar a Análise Descritiva, as organizações criam uma plataforma sólida para a exploração de análises mais complexas, garantindo que as decisões sejam baseadas em uma compreensão clara e fundamentada dos dados históricos.

## **2.4.2 Análise Diagnóstica em Gestão de Pessoas**

A Análise Diagnóstica aprofunda-se além da mera descrição dos dados, buscando entender as razões subjacentes aos eventos e padrões observados. Enquanto a Análise Descritiva foca no *o quê*, a Análise Diagnóstica concentra-se no *porquê*, oferecendo uma compreensão mais profunda dos fatores que influenciam os resultados organizacionais.

Técnicas como análise de correlação, regressão e métodos comparativos de amostragem são amplamente empregadas para identificar causas subjacentes e padrões complexos nos dados. Essas técnicas permitem que os gestores desvendem relações causais entre variáveis, facilitando a identificação dos principais fatores que impactam aspectos críticos da gestão de pessoas, como a rotatividade e o engajamento dos funcionários. Por exemplo, um estudo conduzido por [5] destaca a importância da Análise Diagnóstica para compreender os determinantes da rotatividade e do engajamento, fornecendo *insights* que permitem às organizações desenvolver estratégias mais eficazes para reter talentos e aumentar a satisfação dos colaboradores. Este trabalho conecta-se ao estudo de Boudreau ao explorar

como técnicas analíticas, como correlações, podem ser empregadas para compreender e abordar desafios semelhantes. A Análise Diagnóstica proposta neste estudo avança nesse contexto ao integrar ferramentas estatísticas e técnicas de amostragem, demonstrando sua aplicabilidade prática em dados reais de Recursos Humanos, especialmente no setor de tecnologia. Essa conexão ressalta a viabilidade de traduzir métodos acadêmicos em soluções práticas para problemas organizacionais.

Em um ambiente de TI, onde a competição por talentos técnicos é particularmente acirrada, essas técnicas diagnósticas se mostram ainda mais valiosas. A Análise Diagnóstica pode ajudar a identificar as causas subjacentes da alta rotatividade entre desenvolvedores e outros profissionais técnicos, como fatores relacionados à satisfação no trabalho, oportunidades de crescimento, ou desequilíbrio entre vida pessoal e profissional [2]. Ao compreender essas dinâmicas, as organizações podem implementar intervenções mais direcionadas e eficazes, como programas de desenvolvimento de carreira ou ajustes na carga de trabalho, para reduzir a saída de talentos e fortalecer o engajamento.

A Análise Diagnóstica, portanto, é um componente essencial para a tomada de decisões mais informadas e estratégicas em gestão de pessoas. Ela oferece uma base robusta para o desenvolvimento de políticas e práticas que não apenas abordam os sintomas dos problemas organizacionais, mas também atacam suas causas raízes, permitindo uma gestão de pessoas mais proativa e eficaz.

### **2.4.3 Análise Preditiva em Gestão de Pessoas**

Envolve o uso de modelos estatísticos avançados e algoritmos de machine learning para prever eventos futuros com base em padrões identificados em dados históricos. Essa abordagem permite às organizações antecipar e se preparar para desafios e oportunidades, tornando-se uma ferramenta poderosa no campo da gestão de pessoas.

A Análise Preditiva é amplamente utilizada para prever a rotatividade de funcionários, identificar talentos emergentes e planejar sucessões de maneira mais eficaz. Por exemplo, [37] demonstrou como modelos preditivos podem ser aplicados para calcular a probabilidade de demissões voluntárias em grandes organizações, permitindo que os gestores intervenham de forma proativa para reter funcionários. Ao prever quais funcionários têm maior risco de deixar a empresa, as organizações podem tomar medidas direcionadas, como oferecer pacotes de retenção, ajustes na remuneração ou oportunidades de desenvolvimento de carreira, para mitigar a rotatividade e manter seus melhores talentos. Esse tipo de abordagem se dá principalmente por modelos de classificação supervisionados, como regressão logística.

Na dinamicidade do mercado de trabalho atual, as habilidades e competências exigidas mudam rapidamente, os modelos preditivos desempenham um papel crucial ao prever a

demanda futura por habilidades específicas. Isso permite que as organizações planejem sua força de trabalho de maneira mais estratégica, garantindo que tenham as competências necessárias para atender às demandas futuras do mercado [38]. Ao antecipar as necessidades de habilidades e ajustar as estratégias de recrutamento e treinamento em conformidade, as empresas podem manter uma vantagem competitiva e evitar lacunas críticas de competências.

A Análise Preditiva, portanto, não apenas fornece *insights* sobre o que pode acontecer no futuro, mas também capacita os gestores a tomar decisões informadas que alinhem os recursos humanos com as metas estratégicas da organização. Ao integrar essas previsões em suas práticas de gestão de pessoas, as organizações podem melhorar significativamente sua capacidade de reter talentos, desenvolver líderes futuros e garantir que estejam preparadas para as mudanças no ambiente de negócios.

#### 2.4.4 Análise Prescritiva em Gestão de Pessoas

Representa o nível mais avançado de maturidade no analytics, focando na utilização de algoritmos de otimização para recomendar ações que maximizem os resultados desejados. Diferente das análises descritivas e preditivas, que se concentram em entender o que aconteceu e prever o que pode acontecer, A Análise Prescritiva avança ainda mais ao recomendar as ações mais adequadas a serem implementadas, considerando diversos cenários e restrições.

Estudos como o de [14] destacam a aplicação da Análise Prescritiva na alocação eficiente de recursos e no planejamento estratégico, demonstrando como essa abordagem pode transformar a tomada de decisão em diversas áreas organizacionais. Ao considerar múltiplas variáveis e restrições, os modelos prescritivos oferecem recomendações que ajudam os gestores a tomar decisões mais informadas e alinhadas com os objetivos estratégicos da organização. Por exemplo, em situações de planejamento de pessoal, a Análise Prescritiva pode sugerir a combinação ideal de habilidades e recursos para atender a uma demanda específica, garantindo que os recursos sejam utilizados da forma mais eficaz possível.

Ferramentas de otimização podem ser aplicadas para alocar novos funcionários de maneira eficiente, balanceando as cargas de trabalho entre as equipes e maximizando a produtividade geral [18]. Além disso, essas ferramentas podem ajudar a identificar as alocações que minimizam custos ou reduzem o tempo de entrega de projetos críticos, contribuindo para uma gestão mais eficaz e ágil dos recursos de TI.

A Análise Prescritiva, portanto, não apenas responde ao o que e por que dos dados, mas também ao como — ou seja, como uma organização pode alcançar seus objetivos da forma mais eficiente possível. Ao implementar soluções prescritivas, as organizações podem

não apenas otimizar suas operações, mas também se adaptar rapidamente a mudanças e desafios, garantindo uma vantagem competitiva sustentável no mercado.

### 2.4.5 Integração de Estatística, Ciência de Dados e Inteligência Artificial na Gestão de Pessoas

A integração de estatística, ciência de dados e inteligência artificial está revolucionando a gestão de pessoas, proporcionando uma abordagem mais precisa, informada e baseada em dados para a tomada de decisões. Essa convergência de tecnologias permite que as organizações obtenham *insights* mais profundos e acionáveis sobre seu capital humano, otimizando processos essenciais como recrutamento, retenção e desenvolvimento de funcionários.

No contexto do recrutamento, a ciência de dados e a inteligência artificial permitem a análise de grandes volumes de candidatos, identificando aqueles com maior probabilidade de sucesso com base em dados históricos e preditivos, conforme será demonstrado na sessão Desenvolvimento, subseção Nível de Maturidade Prescritivo. Na área de desenvolvimento de funcionários, a integração dessas disciplinas facilita a criação de programas de treinamento personalizados que respondem às necessidades específicas de desenvolvimento de competências, aumentando o engajamento e o desempenho.

Essa integração não apenas melhora a eficácia das decisões, mas também transforma a cultura organizacional, promovendo uma mentalidade orientada por dados em toda a empresa [54]. Ao alavancar a estatística, a ciência de dados e a inteligência artificial, as organizações podem construir um sistema de gestão de pessoas mais robusto e adaptável, capaz de responder rapidamente às mudanças do mercado e às demandas dos funcionários.

### 2.4.6 Aplicações Práticas

*People Analytics* oferece uma vasta gama de aplicações práticas que podem transformar a gestão de pessoas em diversos níveis organizacionais. Estas aplicações vão além das simples análises de dados, integrando técnicas avançadas de ciência de dados e inteligência artificial para oferecer *insights* estratégicos e operacionais.

**Recrutamento e Seleção:** A utilização de algoritmos sofisticados permite a análise detalhada de currículos e perfis de candidatos, prevendo a adequação dos candidatos às posições oferecidas com base em dados históricos de contratações bem-sucedidas. Além disso, técnicas como análise preditiva e machine learning podem ser aplicadas para identificar candidatos com maior potencial de sucesso a longo prazo, considerando não apenas suas habilidades técnicas, mas também fatores como compatibilidade cultural e trajetória profissional.

Estudos de alocação aplicaram métodos de otimização, como a programação linear inteira mista, para distribuir recursos e pessoal no setor de saúde, ressaltando a importância de modelos matemáticos na gestão eficiente de recursos humanos [57]

**Retenção de Talentos:** A análise de dados desempenha um papel crucial na retenção de talentos, identificando os principais fatores que levam à rotatividade, como insatisfação com a liderança, falta de oportunidades de crescimento ou desajuste entre as expectativas dos funcionários e as ofertas da empresa. Com esses *insights*, as organizações podem desenvolver estratégias de retenção mais eficazes, tais como programas de reconhecimento, planos de carreira personalizados e melhorias no ambiente de trabalho, visando reduzir a saída de colaboradores-chave.

**Desenvolvimento e Treinamento:** A avaliação contínua das necessidades de treinamento é outra área em que *People Analytics* se destaca. Ao analisar dados de desempenho e feedback dos funcionários, as empresas podem identificar lacunas de habilidades e planejar programas de desenvolvimento personalizados que atendam às necessidades individuais e organizacionais. Isso não apenas otimiza o investimento em treinamento, mas também garante que os funcionários adquiram as competências necessárias para contribuir de forma mais eficaz para os objetivos da empresa.

**Desempenho e Produtividade:** O monitoramento contínuo do desempenho dos funcionários, utilizando métricas quantitativas e qualitativas, permite identificar tanto áreas de melhoria quanto talentos emergentes. Com *People Analytics*, as empresas podem realizar análises mais detalhadas do impacto de diferentes fatores, como liderança, recursos e ambiente de trabalho, na produtividade dos colaboradores. Além disso, essas análises possibilitam o reconhecimento e a valorização dos principais talentos, promovendo um ambiente de trabalho mais motivador e alinhado com os objetivos estratégicos da organização.

Essas aplicações demonstram como *People Analytics* pode ser um diferencial competitivo, proporcionando às organizações uma compreensão mais profunda de sua força de trabalho e permitindo uma tomada de decisão mais informada e precisa em todas as áreas da gestão de pessoas.

## 2.4.7 Desafios e Considerações Éticas

Embora *People Analytics* ofereça inúmeros benefícios, também apresenta desafios significativos, especialmente em termos de privacidade, ética e a correta interpretação dos dados. A coleta e análise de dados sensíveis dos funcionários devem ser conduzidas com extrema responsabilidade, respeitando os direitos individuais e garantindo que todos os processos sejam transparentes. Isso inclui a adoção de práticas robustas de segurança de dados para proteger a confidencialidade das informações e assegurar que o consentimento

dos funcionários seja obtido de maneira clara e informada, evitando qualquer forma de coerção ou falta de clareza sobre como seus dados serão utilizados.

Além disso, a interpretação dos dados apresenta um desafio particular. A análise deve ser conduzida com um enfoque cuidadoso para evitar a introdução de preconceitos e discriminação, intencionais ou não. É crucial que os modelos analíticos sejam desenvolvidos e revisados com atenção para vieses implícitos que podem influenciar os resultados e, por extensão, as decisões de gestão de pessoas [51]. Isso requer um compromisso contínuo com a revisão crítica dos algoritmos e das metodologias utilizadas, garantindo que as decisões baseadas em dados não perpetuem ou reforcem desigualdades preexistentes.

Outro aspecto ético relevante é a necessidade de transparência na comunicação dos resultados e das decisões derivadas de *People Analytics*. As organizações devem estar dispostas a explicar claramente aos seus funcionários como os *insights* obtidos serão utilizados e quais impactos isso pode ter em suas carreiras e no ambiente de trabalho. Essa transparência não só reforça a confiança dos colaboradores no uso de suas informações, mas também promove um ambiente de trabalho mais justo e equitativo, onde as decisões são percebidas como legítimas e fundamentadas em dados objetivos.

Finalmente, as organizações devem estar atentas ao equilíbrio entre a utilização de *People Analytics* para melhorar a eficiência e o bem-estar organizacional e a necessidade de preservar a dignidade e os direitos dos funcionários. Em um cenário onde os dados são cada vez mais valorizados, é essencial que as práticas de *People Analytics* sejam guiadas por princípios éticos sólidos que priorizem o respeito à privacidade, a equidade e a integridade dos indivíduos. Isso assegura que os benefícios da análise de dados sejam realmente traduzidos em vantagens para todos os envolvidos, sem comprometer valores fundamentais que sustentam a confiança e a coesão dentro da organização.

## 2.4.8 Futuro do *People Analytics*

O futuro de *People Analytics* é extremamente promissor, impulsionado pelos avanços contínuos em tecnologias de Inteligência Artificial (IA), Machine Learning e análise de big data. À medida que essas tecnologias se tornam mais acessíveis e poderosas, espera-se que as organizações se tornem cada vez mais dependentes dessas ferramentas para tomar decisões estratégicas baseadas em dados, transformando a forma como gerenciam seus recursos humanos.

A possibilidade de uma integração entre *People Analytics* e outras áreas funcionais da empresa, como marketing, finanças e operações, permitirá uma visão mais holística do impacto das pessoas no desempenho organizacional, possibilitando uma tomada de decisão mais alinhada com os objetivos estratégicos da organização. A análise de dados poderá



revelar correlações anteriormente invisíveis entre diferentes áreas, permitindo intervenções mais eficazes e *insights* que atravessam fronteiras funcionais.

Além disso, o desenvolvimento de técnicas de análise preditiva e prescritiva abrirá novas fronteiras para uma gestão de pessoas ainda mais proativa e personalizada. Essas técnicas permitirão prever com maior precisão tendências de comportamento dos colaboradores, como probabilidade de turnover, necessidade de desenvolvimento de competências específicas, e até mesmo o impacto de mudanças organizacionais no moral e na produtividade. A capacidade de antecipar desafios e oportunidades permitirá às organizações não apenas reagir, mas também se preparar com antecedência, implementando estratégias que maximizem o potencial de sua força de trabalho.

Com a evolução das técnicas prescritivas, as organizações poderão não apenas prever o futuro, mas também recomendar as melhores ações a serem tomadas para alcançar os resultados desejados. Isso inclui sugestões automatizadas para alocação de recursos, planejamento de sucessão e até mesmo intervenções personalizadas para melhorar o bem-estar dos funcionários. Essas recomendações baseadas em dados terão o potencial de reduzir os custos operacionais, melhorar a satisfação dos colaboradores e, em última análise, impulsionar o desempenho geral da empresa.

Entretanto, junto com esses avanços, surgem também novos desafios. A crescente dependência de algoritmos e IA para a tomada de decisões levanta questões éticas e de governança que precisarão ser abordadas com seriedade. A transparência nos processos, a explicabilidade dos modelos utilizados e a prevenção de vieses serão fundamentais para garantir que *People Analytics* continue a ser uma ferramenta confiável e justa.

Por fim, é provável que o futuro de *People Analytics* envolva uma maior personalização das experiências dos funcionários. Utilizando dados de múltiplas fontes, as organizações poderão criar perfis detalhados de seus colaboradores, permitindo uma gestão de carreira que leva em consideração não apenas as habilidades e o desempenho, mas também as preferências individuais, as motivações e até mesmo o bem-estar emocional. Isso poderá resultar em ambientes de trabalho mais satisfatórios e produtivos, onde cada colaborador é tratado como um indivíduo único com necessidades e potenciais específicos.

Em resumo, o futuro de *People Analytics* promete transformar profundamente a forma como as organizações gerenciam seu capital humano, proporcionando uma abordagem mais inteligente, eficiente e humana na gestão de pessoas [18].

# Capítulo 3

## Revisão da Literatura

### 3.0.1 *Conceitos Gerais*

*People Analytics*, também conhecido como *HR Analytics* ou *Workforce Analytics* ou *Talent Analytics*, refere-se ao uso de dados, ferramentas estatísticas e técnicas analíticas para melhorar a gestão de pessoas nas organizações [54]. Este campo emergente combina princípios de gestão de recursos humanos com avanços em ciência de dados, visando transformar dados sobre funcionários em *insights* acionáveis para informar decisões estratégicas. Basicamente nasce da interseção de 4 áreas do conhecimento: Estatística, Método Científico, Estratégia de Pessoas e Computação Aplicada [54].

O HR Analytics é definido como a aplicação sistemática de modelagem utilizando estatísticas inferenciais aos dados existentes relacionados a pessoas em Recursos Humanos, com o objetivo de informar julgamentos sobre possíveis fatores causais que influenciam indicadores-chave de desempenho relacionados a RH [17]. De maneira simplificada, são utilizadas técnicas avançadas de estatística e análise quantitativa nas informações relacionadas às pessoas dentro das organizações. Isso permite testar modelos estatísticos e prever fatores que podem impulsionar alto desempenho ou causar por exemplo a saída de um funcionário da organização [31]. Além disso, quando apropriado, pode-se aplicar modelos preditivos para fazer previsões tangíveis sobre resultados específicos, como comportamentos de funcionários ou da organização, sob determinadas condições.

O termo *People Analytics* refere-se amplamente ao uso de uma abordagem orientada por dados para lidar com decisões e práticas relacionadas aos funcionários. O conceito *People Analytics* pode ser usado tanto como a função organizacional na qual a coleta de dados, análises e traduções ocorrem, quanto como um conjunto de práticas que utilizam dados dos funcionários para informar e auxiliar os processos de tomada de decisão e atividades dos funcionários em toda a organização [44].

*People Analytics* é definido como o uso de análise de dados para gerenciar e melhorar

as práticas de recursos humanos. Esta abordagem vai além da simples coleta e armazenamento de dados, enfatizando a análise e a interpretação das informações para apoiar decisões estratégicas em diversas áreas, como recrutamento, retenção, desenvolvimento de talentos e planejamento de sucessão. Ao concentrar-se na compreensão profunda dos dados, é possível obter *insights* valiosos que fortalecem as práticas de gestão e potencializam o desempenho organizacional [2].

Esta abordagem envolve o uso de métodos estatísticos e tecnologias de big data para analisar dados relacionados aos funcionários, incluindo informações demográficas, desempenho, engajamento e outros atributos relevantes [46].

O HR Analytics é uma das tendências mais recentes no campo de Recursos Humanos, visando realizar análises fundamentadas em dados. Essa abordagem orienta a tomada de decisões com base em evidências, buscando uma gestão prática que incorpora rigor científico ao RH. Além disso, o HR Analytics complementa as decisões intuitivas com uma perspectiva objetiva [46].

A importância de *People Analytics* reside na sua capacidade de fornecer uma visão baseada em dados sobre o comportamento e desempenho dos funcionários, permitindo às organizações tomar decisões mais informadas e estratégicas. Essa abordagem pode levar a uma melhor alocação de recursos, aumento da eficiência operacional e melhoria na satisfação e engajamento dos funcionários (Levenson, 2018).

### 3.0.2 *Framework I*

O artigo *People Analytics Effectiveness: Developing a Framework* [41] propõe um *Framework* para avaliar e melhorar a eficácia das equipes de *People Analytics*. O *Framework*, denominado *People Analytics Effectiveness Wheel*, identifica quatro categorias principais de ingredientes que uma equipe de *People Analytics* precisa para ser eficaz: Recursos Habilitadores, Produtos, Gestão de *stakeholders* e Estrutura de Governança, conforme tabela Tabela 3.2:

| Categoria                     | Descrição                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|-------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Recursos Habilitados          | <p>Incluem três subtemas principais:</p> <ul style="list-style-type: none"> <li>• <b>Suporte da alta administração:</b> Fornece os recursos financeiros e o apoio político necessários.</li> <li>• <b>Dados e infraestrutura:</b> Essenciais para análises precisas e úteis.</li> <li>• <b>KSAOs (Conhecimentos, Habilidades e Capacidades):</b> Envolvem habilidades em estatísticas, gestão de dados, visualização, storytelling, RH e psicologia.</li> </ul> |
| Produtos                      | <p>Os produtos entregues podem ser divididos em:</p> <ul style="list-style-type: none"> <li>• <b>Ferramentas de monitoramento:</b> Dashboards e scorecards para identificar problemas e tomar decisões informadas.</li> <li>• <b>Pesquisa organizacional:</b> Aborda questões específicas e fornece <i>insights</i> contextuais.</li> <li>• <b>Cultura baseada em evidências:</b> Garante decisões fundamentadas em dados e análises.</li> </ul>                |
| Gestão de <i>Stakeholders</i> | <p>Abrange a interação com:</p> <ul style="list-style-type: none"> <li>• <b>Profissionais de RH:</b> Combinam intuição com dados objetivos.</li> <li>• <b>Alta administração:</b> Necessita de <i>insights</i> para apoiar a estratégia de negócios.</li> <li>• <b>Gerentes de linha:</b> Implementam políticas baseadas em <i>insights</i>.</li> <li>• <b>Funcionários:</b> Exigem privacidade e confiança no uso dos dados.</li> </ul>                        |

Continua na próxima página...

| Categoria               | Descrição                                                                                                                                                                                                                                                                                                                            |
|-------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Estrutura de Governança | <p>Abrange:</p> <ul style="list-style-type: none"> <li>• <b>Gestão de dados:</b> Procedimentos para garantir qualidade e segurança.</li> <li>• <b>Questões éticas:</b> Asseguram uso justo e transparente dos dados.</li> <li>• <b>Colaboração interna:</b> Facilita o trabalho entre subequipes de análise e relatórios.</li> </ul> |

Tabela 3.2: Principais Categorias do *People Analytics Effectiveness Wheel*

### 3.0.3 *Framework II*

O artigo denominado Uma Proposta de *Framework* para a Execução e Estruturação do *People Analytics* [42] apresenta um *Framework* para a implementação e gestão de *People Analytics* no contexto de Gestão de Recursos Humanos (GRH). Este *Framework* visa integrar práticas de gestão de pessoas com as necessidades estratégicas das organizações, destacando-se por seu enfoque multidisciplinar e por abordar tanto a gestão do capital humano quanto a gestão da tecnologia da informação (TI).

O *Framework* proposto por [42] é estruturado para abordar três constructos principais:

- **Arquiteturas de Capital Humano (ACHs):** Indicadas pelas horizontais de execução. ACHs que demandam maior foco na gestão das relações de trabalho podem se valer da criação de instrumentação analítica menos sofisticada (indicadores e dashboards) como primeiro passo de implantação.
- **Papel da TI:** Dado pelas horizontais de estruturação, consolidando a organização das etapas de gestão do conhecimento e as questões essenciais de estruturação de *People Analytics*.
- **Descentralização de Processos e Poder de Decisão:** A difusão de *People Analytics* torna mais crítico o entendimento sobre o funcionamento dos algoritmos de aprendizado de máquina, mantendo a flexibilização para a adoção de modelos descentralizados e centralizados.

Este *Framework* é apresentado como uma proposta flexível, que pode ser ajustada conforme a maturidade da organização e suas necessidades específicas. Ele visa ordenar os

principais pontos dos *Frameworks* existentes, servindo como guia para pesquisadores e praticantes na implementação e gestão de *People Analytics*.

### **3.0.4 Considerações quanto aos *Frameworks***

O observado na revisão da literatura é que nenhum dos frameworks identificados alinha-se aos objetivos específicos deste estudo. Enquanto os *Frameworks* mencionados anteriormente enfatizam principalmente a governança, este trabalho foca no desenvolvimento de um *Framework* voltado não para práticas de governança e gerenciamento de TI, mas para a aplicação prática da Ciência de Dados.

# Capítulo 4

## Metodologia

### 4.0.1 Descrição do Método de Pesquisa

O método de pesquisa adotado neste trabalho segue o modelo CRISP-DM (Cross-Industry Standard Process for Data Mining), devido à sua flexibilidade, abordagem iterativa e ampla aplicabilidade em projetos que envolvem análise de dados. Este modelo foi escolhido por oferecer uma estrutura clara que conecta as etapas técnicas de manipulação e modelagem de dados aos objetivos de negócio, facilitando a aplicação prática dos resultados em contextos organizacionais diversos. Além disso, o CRISP-DM se destaca por ser um padrão consolidado, utilizado em projetos de ciência de dados por sua capacidade de lidar com problemas complexos e garantir soluções replicáveis e alinhadas com os desafios do negócio. No contexto deste trabalho, ele permite integrar de forma eficiente os quatro níveis de maturidade analítica (Descritivo, Diagnóstico, Preditivo e Prescritivo) às necessidades específicas de gestão de pessoas.

Este processo é composto por seis fases sequenciais, conforme Tabela 4.2:

| Fase                    | Descrição                                                                                                                                                      |
|-------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Entendimento do Negócio | Compreensão dos objetivos do negócio e dos requisitos do projeto. Define os objetivos específicos e os critérios de sucesso que orientarão as próximas etapas. |
| Entendimento dos Dados  | Coleta inicial e exploração dos dados disponíveis. Avaliação da qualidade dos dados e identificação de lacunas ou inconsistências.                             |

Continua na próxima página...

| Fase                 | Descrição                                                                                                                                                                                                                                       |
|----------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Preparação dos Dados | Limpeza, transformação e integração dos dados necessários para a modelagem. Inclui correção de valores ausentes, normalização de variáveis e criação de novos atributos relevantes.                                                             |
| Modelagem            | Aplicação de técnicas de mineração de dados, como algoritmos de machine learning (árvores de decisão, regressão logística, redes neurais, etc.). Teste de várias técnicas para selecionar a melhor em termos de precisão e robustez.            |
| Avaliação            | Verificação se o modelo atende aos critérios definidos na fase de entendimento do negócio. Avaliação dos insights gerados para garantir utilidade e acionabilidade. Ajustes podem ser necessários, refletindo a natureza iterativa do processo. |
| Implantação          | Implementação dos modelos em ambiente de produção. Inclui automação, criação de dashboards, integração em sistemas existentes, documentação e monitoramento contínuo para garantir desempenho ao longo do tempo.                                |

Tabela 4.2: Etapas do Processo CRISP-DM

#### 4.0.2 Ferramentas e Técnicas Utilizadas

Para a execução desta pesquisa, serão utilizadas diversas ferramentas e técnicas de análise de dados e machine learning, incluindo:

- **Power BI:** Utilizado para criação de painéis e visualização de dados [56].
- **Python:** Linguagem de programação utilizada para análise de dados e construção de modelos preditivos, com bibliotecas como Pandas, Scikit-learn e TensorFlow [52].
- **SQL:** Para manipulação e consulta de dados em bancos de dados relacionais [34].
- **Métodos Estatísticos:** Incluindo análise de regressão, testes de hipóteses e análise de correlação para a fase de análise diagnóstica [9].
- **Algoritmos de Machine Learning:** Como árvores de decisão, regressão logística e redes neurais para a criação de modelos preditivos [4].
- **Métodos de otimização:** Para otimização e alocação de recursos na análise prescritiva [36].



### 4.0.3 Procedimentos para Coleta e Análise de Dados

Os procedimentos para coleta e análise de dados seguem as fases do CRISP-DM:

- **Entendimento do Negócio:**
  - Definição dos objetivos e requisitos do projeto junto aos stakeholders.
  - Identificação dos processos de gestão de pessoas que serão analisados, incluindo atração, sucessão, formação e retenção de colaboradores.
- **Entendimento dos Dados:**
  - Coleta de dados relevantes dos sistemas de RH do empresa alvo do case aplicado, incluindo dados de colaboradores, métricas de desempenho, histórico de treinamento e registros de desligamentos [11].
  - Análise exploratória inicial para identificar padrões, tendências e possíveis problemas de qualidade nos dados.
- **Preparação dos Dados:**
  - Limpeza e pré-processamento dos dados, incluindo tratamento de valores ausentes, normalização e transformação de variáveis [39].
  - Criação de variáveis derivadas e seleção de atributos relevantes para a modelagem.
- **Modelagem:**
  - Utilização de ferramentas DataViz citadas na sessão 3.0.2, para criação de dashboards que apoiem o processo da tomada de decisões, como employee churn, distribuição de gênero, Head Count, entre outros.
  - Utilização de técnicas de correlação linear para análise exploratória de dados quantitativos. A técnica utilizada no case foi o Coeficiente de correlação de Pearson, gerando insights valiosos para a organização.
  - Comparação da eficiência de tipos de amostragem para pesquisas de pessoas. No case exemplo, foi aplicado na Pesquisa de Clima Organizacional o comparativo entre Amostragem Aleatória Simples e Amostragem por conglomerados.
  - Aplicação de técnicas de machine learning para construção de modelos preditivos de processos de pessoas. No case exemplo, foi criado um modelo de ascensão profissional e risco de desligamento [20].

- Implementação de métodos de otimização para atingimento de nível prescritivo [36]. No case exemplo, o princípio de maximização de fluxo por grafos foi utilizado para a melhor alocação de novos colaboradores, em áreas de maior conhecimento e interesse.

- **Avaliação:**

- Avaliação dos modelos descritivos: Ampla utilização por parte dos stakeholders.
- Avaliação da etapa de diagnóstico: Conversão de insights em práticas pra melhoria da gestão de pessoas.
- Avaliação dos modelos preditivos utilizando métricas como acurácia, precisão, recall e F1-score [4].
- Avaliação dos modelos prescritivos: No case exemplo, verificar diferença no total de pontuação entre alocações aleatórias e a alocações com a solução de otimização.
- Validação dos resultados junto aos stakeholders e ajustes nos modelos conforme necessário.

#### 4.0.4 Modelo de *Framework*

Na Fig.4.1, são apresentadas as etapas e as relações do *Framework* proposto, com foco no atingimento dos quatro níveis de maturidade analítica (Descritivo, Diagnóstico, Preditivo e Prescritivo). Este modelo foi desenvolvido para ser flexível e aplicável a diferentes tipos de organizações, permitindo ajustes às necessidades específicas de cada negócio.

O *Framework* é composto por quatro etapas principais: Entendimento do Negócio, Entendimento dos Dados, Aplicação de Técnicas Analíticas e Avaliação. A etapa Aplicação de Técnicas Analíticas é customizada de acordo com o nível de maturidade analítica (Descritivo, Diagnóstico, Preditivo e Prescritivo), sendo organizada para garantir que o progresso entre os níveis seja gradual e estruturado.

Diferentemente do modelo CRISP-DM, a etapa de Implantação foi excluída deste estudo. Essa decisão foi tomada porque, embora as avaliações realizadas demonstrem a viabilidade do uso das técnicas propostas, a implantação depende de decisões estratégicas específicas de cada organização, o que extrapola o escopo desta pesquisa.

A seguir, cada uma dessas etapas é detalhada:

1. **Entendimento do Negócio:** A primeira etapa envolve a identificação dos problemas a serem resolvidos e a definição clara dos objetivos organizacionais. Nesta fase, é crucial alinhar as expectativas dos *stakeholders* e garantir que o esforço analítico esteja orientado para atender às metas estratégicas do negócio.

2. **Entendimento dos dados:** Após a definição do problema, os dados necessários são coletados, organizados e preparados para análise. Isso inclui limpeza de dados, integração de diferentes fontes e a criação de variáveis que suportem os níveis de maturidade analítica. A qualidade e a representatividade dos dados nesta etapa têm impacto direto na eficácia do modelo.
3. **Aplicação de Técnicas Analíticas:** Nesta fase, são aplicadas as técnicas específicas de cada nível de maturidade:
  - **Descritivo:** Dashboards e relatórios para visualização de dados históricos e compreensão de tendências.
  - **Diagnóstico:** Análises estatísticas, como correlação e regressão, para identificar causas de padrões observados.
  - **Preditivo:** Modelagem com algoritmos de *machine learning* para prever eventos futuros, como risco de turnover.
  - **Prescritivo:** Uso de *solvers* e otimização para recomendar ações, como alocação eficiente de recursos.
4. **Avaliação e Validação:** Os resultados gerados pelas análises são avaliados para garantir que atendam aos critérios definidos na primeira etapa. Isso inclui validações quantitativas (como métricas de precisão e acurácia) e qualitativas (relevância e aplicabilidade dos *insights*).
5. **Implantação e Monitoramento:** Após a validação, os modelos e as soluções desenvolvidas são implementados no ambiente organizacional. É fundamental estabelecer métricas de monitoramento contínuo para avaliar o impacto das ações e garantir que os resultados se mantenham consistentes ao longo do tempo.

O *Framework* proposto busca integrar de maneira prática as capacidades analíticas da organização com suas necessidades estratégicas, promovendo o uso de dados como base para decisões mais informadas e assertivas. Essa abordagem iterativa e ajustável permite que diferentes tipos de organizações, independentemente de seu setor ou nível de maturidade atual, adaptem o modelo para maximizar o impacto em seus processos de gestão de pessoas.

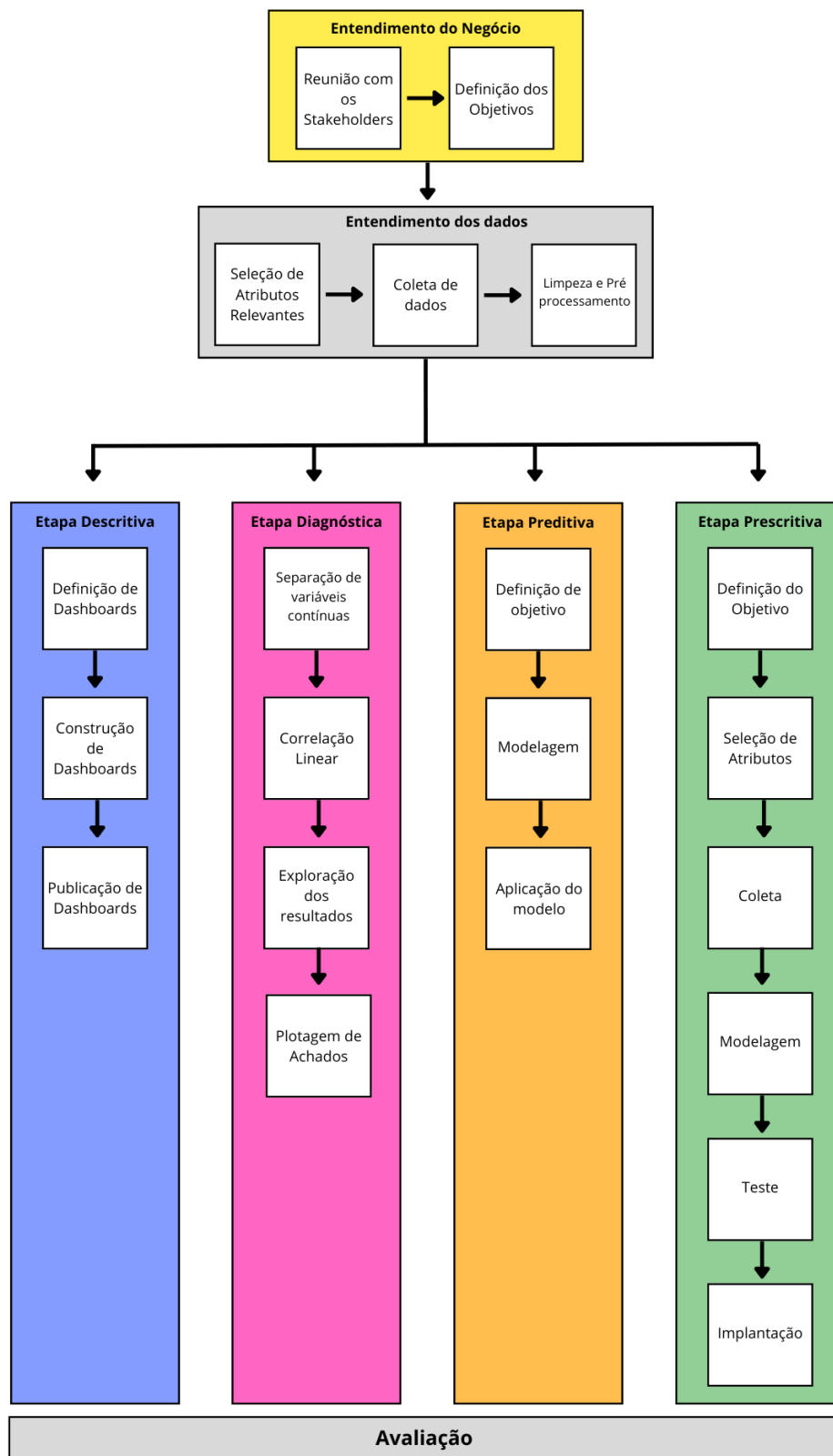


Figura 4.1: *Framework* - Fonte: O autor

# Capítulo 5

## Desenvolvimento e Análise dos Resultados

Neste trecho, são descritas as etapas do framework proposto, conforme listadas no cronograma, e como foram aplicadas no case objeto da pesquisa. O estudo refere-se à área de tecnologia de uma grande empresa pública brasileira:

### 5.0.1 Reunião com os Stakeholders e Definição dos Objetivos do Projeto

No contexto do CRISP-DM (Cross-Industry Standard Process for Data Mining), a fase inicial de entendimento do negócio é fundamental para o sucesso de qualquer projeto de análise de dados. Uma etapa crucial dentro desta fase é a realização de reuniões com os stakeholders, que desempenham um papel vital na compreensão profunda das necessidades de negócio e na definição dos objetivos do projeto.

Para padronizar essas reuniões, sugere-se o uso de um formulário estruturado, que pode incluir as seguintes seções:

#### 1. Identificação do Projeto:

- Nome do projeto
- Objetivo principal
- Equipe responsável

#### 2. Stakeholders Presentes:

- Nome
- Cargo/Função

- Papel no projeto

### **3. Necessidades e Requisitos:**

- Quais são os principais problemas a serem resolvidos?
- Quais dados e informações são essenciais?
- Quais são os indicadores-chave de desempenho (KPIs) relacionados?

### **4. Restrições e Limitações:**

- Quais são as principais limitações operacionais ou de recursos?
- Existem restrições de tempo ou orçamento?

### **5. Oportunidades Identificadas:**

- Quais são os ganhos esperados com o projeto?
- Como os resultados podem impactar positivamente a organização?

### **6. Próximos Passos e Ações:**

- Quais atividades serão realizadas após a reunião?
- Quem será responsável por cada atividade?

### **7. Métricas de Sucesso:**

- Quais métricas serão usadas para avaliar o sucesso do projeto?
- Como essas métricas estão alinhadas aos objetivos organizacionais?
- Há benchmarks ou metas específicas a serem atingidas?

Essa estrutura permite que as reuniões sejam conduzidas de forma organizada e eficiente, garantindo que todas as informações relevantes sejam capturadas e documentadas.

O uso deste formulário estruturado promove o alinhamento de expectativas entre os stakeholders ao detalhar claramente os objetivos, papéis e responsabilidades de cada participante. Além disso, ele possibilita a identificação precoce de restrições e oportunidades, permitindo que a equipe projete soluções mais adaptadas à realidade do negócio. Por fim, ao estabelecer um padrão de comunicação claro e documentado, o formulário contribui para a construção de um relacionamento baseado na confiança e colaboração entre os envolvidos, assegurando maior comprometimento com os resultados do projeto.

Reunir-se com os stakeholders no início do processo CRISP-DM não é apenas uma formalidade, mas uma etapa estratégica que estabelece as bases para o sucesso do projeto. Essa

interação permite um entendimento profundo das necessidades de negócio, alinha expectativas, e garante que os esforços de análise de dados estejam firmemente conectados aos objetivos organizacionais.

O caso em análise identificou algumas questões significativas que poderiam ser abordadas nas diversas frentes, a saber:

- Nível Descritivo: Falta de informações de pessoas em tempo real para a tomada de decisões, como *turnover*, *headcount* e distribuição de gênero e de geracional.
- Nível Exploratório: Este nível busca identificar e responder perguntas que, muitas vezes, nem mesmo a organização tem ciência de que existem. Ele envolve uma análise exploratória dos dados disponíveis, com o objetivo de revelar padrões, correlações e *insights* ocultos que possam trazer valor para a organização.
- Nível Preditivo: Prontidão de funcionários para sucessão de Analista Senior para Gerente de Times.
- Nível Prescritivo: Melhor alocação de novos funcionários oriundos de seleção externa (concurso público).

### 5.0.2 Seleção de atributos relevantes e Coleta de Dados

A criação de um framework para aplicação dos quatro níveis de maturidade do Analytics em gestão de pessoas, é fundamental garantir que o primeiro passo na busca por um modelo data-driven na gestão de talentos seja a criação ou posse de um banco de dados com integridade, autenticidade, disponibilidade e confidencialidade.

Os dados que subsidiarão este caso exemplo estão armazenados em ambiente IBM Db2 for z/OS. O IBM Db2 é um sistema de gerenciamento de banco de dados relacional (RDBMS) desenvolvido pela IBM, amplamente utilizado em ambientes corporativos para armazenamento, gerenciamento e acesso seguro a grandes volumes de dados. Este sistema suporta múltiplas linguagens de consulta, como SQL, e oferece recursos avançados para processamento de transações, análise de dados e integração com outras ferramentas e plataformas. Com sua alta escalabilidade e desempenho robusto, o Db2 for z/OS é particularmente adequado para ambientes de missão crítica que exigem alta disponibilidade e segurança. Com base nos objetivos do negócio do caso exemplo, foram utilizadas 58 atributos referentes a funcionários da ativa, listados no Anexo B e 85 atributos referentes a funcionários em processo de recrutamento, listados no Anexo C.

## 5.1 Nível de maturidade: Análise Descritiva

### 5.1.1 Análise Passiva

Na análise passiva, os dados são solicitados pelas partes interessadas, como gestores e analistas de RH, que identificam uma necessidade específica de informação. Essas partes interessadas entram em contato com a área de Recursos Humanos, que por sua vez acessa o banco de dados para recuperar os dados relevantes e fornecer informações. Este método, embora eficaz, pode ser limitado pela necessidade de solicitações individuais e pela dependência de intermediários para a recuperação de dados, o que pode gerar atrasos na obtenção da informação.

### Análise Ativa

A análise ativa representa uma abordagem mais avançada e proativa. Nesta forma de análise, utilizam-se ferramentas de visualização de dados (dataviz) para criar dashboards interativos e informativos que são disponibilizados continuamente para as partes interessadas. Esses dashboards são projetados para fornecer acesso em tempo real a informações críticas, garantindo total disponibilidade e integridade dos dados. Com a análise ativa, gestores e analistas de RH podem monitorar indicadores chave de desempenho, identificar tendências e tomar decisões informadas de maneira mais ágil e eficiente. A implementação de dashboards também permite a automação de relatórios, reduzindo a necessidade de intervenções manuais e melhorando a precisão e a confiabilidade das informações.

Em resumo, enquanto a análise passiva é reativa e depende de solicitações específicas para a obtenção de dados, a análise ativa é proativa e utiliza a tecnologia para fornecer informações em tempo real, promovendo uma cultura de tomada de decisões baseada em dados dentro da organização.

Dentro do escopo de casos exemplares, utilizando os dados descritos na seção 4.0.1 Coleta de Dados, apresentamos exemplos de alguns dashboards construídos na ferramenta Power BI da Microsoft. Esses dashboards têm a capacidade de fornecer informações valiosas para a tomada de decisão, suprimindo as necessidades apontadas na reunião com os Stakeholders:

- **Distribuição de Gênero**

- Tem por objetivo demonstrar a distribuição de gênero entre os funcionários (5.1), permitindo monitorar e promover a diversidade e a igualdade de gênero na organização, subsidiando políticas de atração mais equitativas. Este painel também supre a necessidade de headcounts.

- **Desligamentos:**



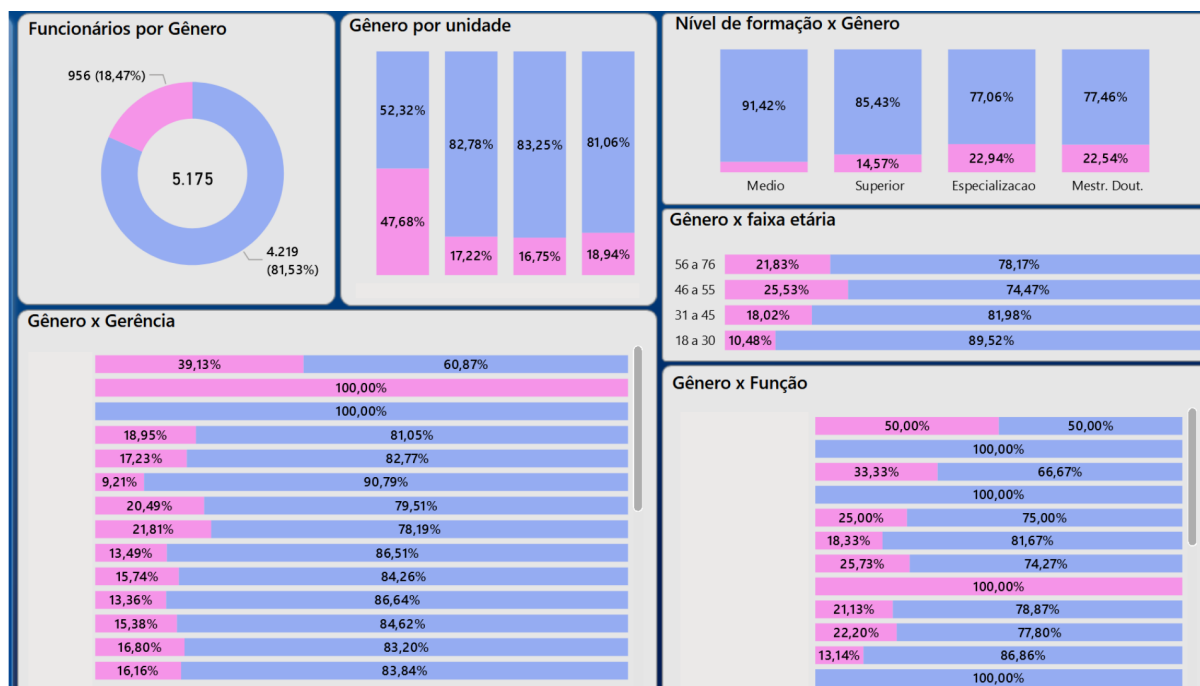


Figura 5.1: Painel de Gênero - Fonte: O autor

- Na fig. 5.2 é possível analisar os padrões de desligamento, identificando possíveis causas e áreas que necessitam de intervenção para reduzir a rotatividade. Alguns dados foram suprimidos da figura listada.

#### • Distribuição de Gerações:

- A fig.5.3 mostra a distribuição geracional e a partir deste dado é possível ajustar políticas de gestão com práticas mais adequadas para formação e motivação do clusters funcionais.

Neste nível de maturidade, o valor da entrega reside na capacidade de transitar do dado para a informação, da informação para a inteligência competitiva, e da inteligência competitiva para a estratégia. Por exemplo, no painel de distribuição de gênero (Fig.5.1), podemos observar uma maior concentração de mulheres nas faixas etárias mais altas. Essa tendência sugere que, a médio e longo prazo, haverá uma redução percentual de mulheres na organização em relação ao total de funcionários. Caso a organização identifique essa situação como uma questão estratégica, ela pode implementar um processo de atração focado no gênero feminino para manter ou aumentar a diversidade de gênero.



Figura 5.2: Painel de Desligamentos - Fonte: O autor

## 5.2 Nível de maturidade: Análise Diagnóstica

Este passo pode ser dividido em dois momentos: verificação de correlações entre as variáveis e inferências.

### 5.2.1 Verificação de Correlações

A análise dos dados em busca do impacto causal entre variáveis de amostras independentes (não pareadas) utiliza uma série de técnicas estatísticas, de acordo com o tipo de dado. Para a análise entre variáveis quantitativas, utiliza-se o *Coefficiente de Correlação Linear de Pearson* ([40]). Para a análise entre variáveis qualitativas, utiliza-se o *V de Cramer* ([12]). Para o cruzamento entre dados quantitativos e qualitativos, utiliza-se o  $R^2$  ([23]).

É importante frisar que o fato de duas variáveis serem fortemente correlacionadas não implica uma relação de causa e efeito entre elas ([19]). Havendo uma correlação significativa entre duas variáveis, deve-se considerar as seguintes possibilidades:

- Existe uma relação direta de causa e efeito entre as variáveis?
- Existe uma relação reversa de causa e efeito entre as variáveis?
- Existe uma terceira variável que possa ser a causa da correlação entre as variáveis?
- É possível que essa relação seja apenas coincidência?

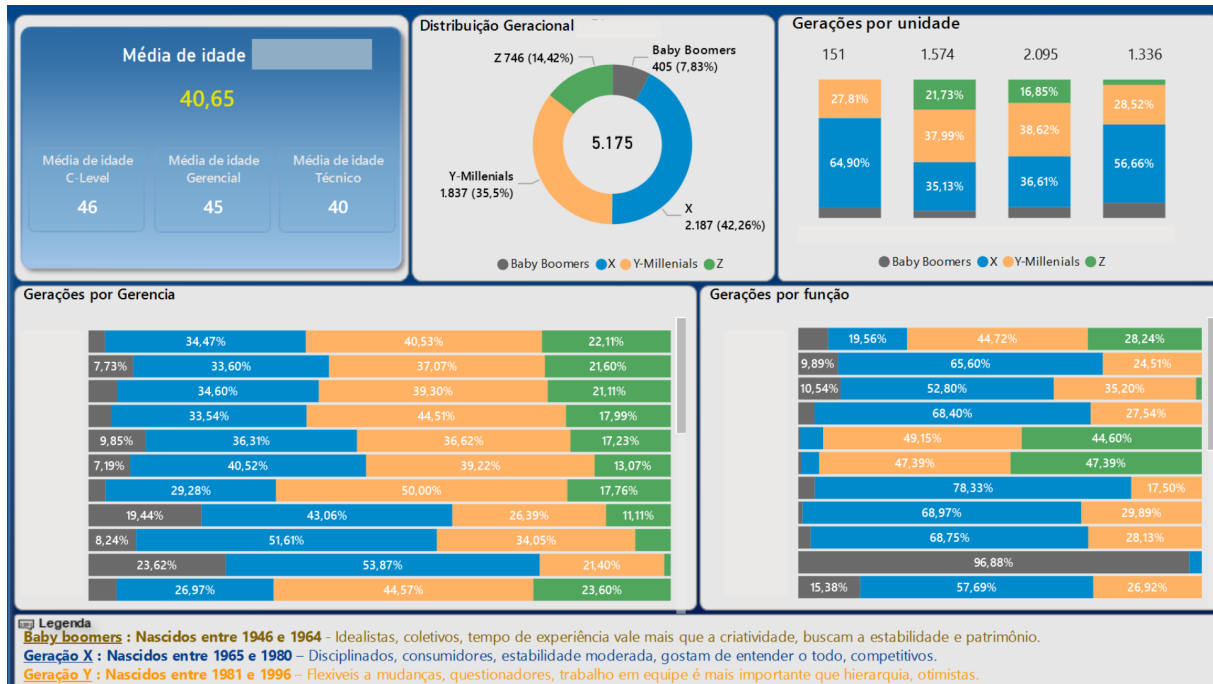


Figura 5.3: Painel Geracional - Fonte: O autor

Para verificação das correlações entre variáveis quantitativas, o Coeficiente de Correlação Linear de Pearson é a técnica sugerida deste framework. O Coeficiente de Correlação de Pearson é uma medida da força e direção de uma relação linear entre duas variáveis. [19].

A Fórmula utilizada para determinação deste coeficiente é:

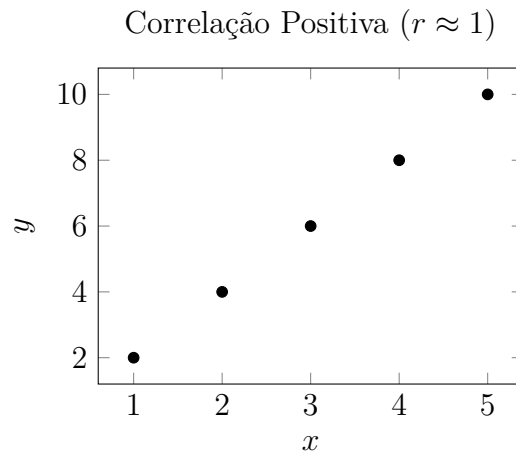
$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$

onde:

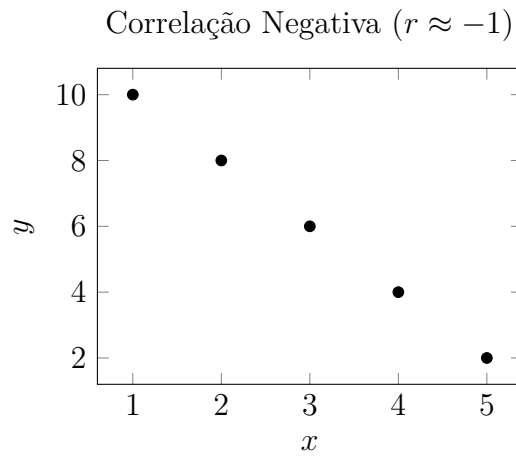
- $r$  é o coeficiente de correlação de Pearson,
- $n$  é o número de pares de valores,
- $\sum xy$  é a soma dos produtos de  $x$  e  $y$ , onde o  $x$  e  $y$  são as variáveis independentes que se busca verificar a existência de correlação, por exemplo, salário atual e nota da avaliação de desempenho,
- $\sum x$  é a soma dos valores  $x$ ,
- $\sum y$  é a soma dos valores  $y$ ,
- $\sum x^2$  é a soma dos quadrados dos valores  $x$ ,
- $\sum y^2$  é a soma dos quadrados dos valores  $y$ .

O Coeficiente de Correlação de Pearson ( $r$ ) terá valores que variam entre -1 e 1. Esses valores indicam a força e a direção da correlação entre duas variáveis:

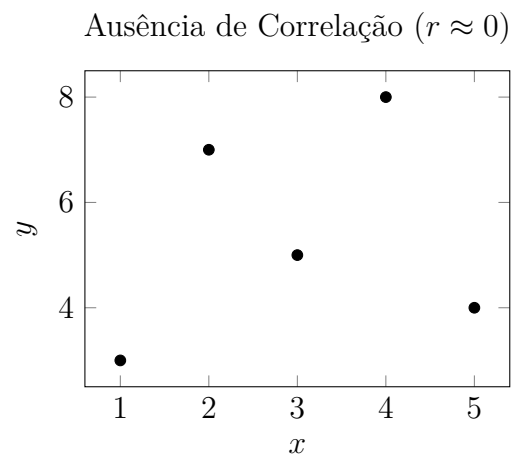
- Quanto mais próximo de -1, mais forte é a correlação negativa.
- Quanto mais próximo de zero, menor é a existência de correlação.
- Quanto mais próximo de 1, mais forte é a correlação positiva.



(a) Correlação Positiva. Fonte: O autor



(b) Correlação Negativa. Fonte: O autor



(c) Ausência de Correlação. Fonte: O autor

Figura 5.4: Exemplos de gráficos de dispersão para diferentes tipos de correlação. Fonte: O autor

Aplicando o Coeficiente de Correlação de Pearson ( $r$ ) nos dados qualitativos levantados, temos as correlações descritas no Anexo A.

## 5.2.2 Análise das Correlações

### Correlações Positivas

As correlações positivas indicam associações diretas entre as variáveis, sugerindo que o aumento de uma está associado ao aumento de outra. As principais correlações positivas identificadas são:

Interpretando os valores das correlações lineares apresentadas, é possível identificar algumas correlações interessantes, por exemplo:

- Tempo de empresa x salário: 0.43 (Figura 5.5)

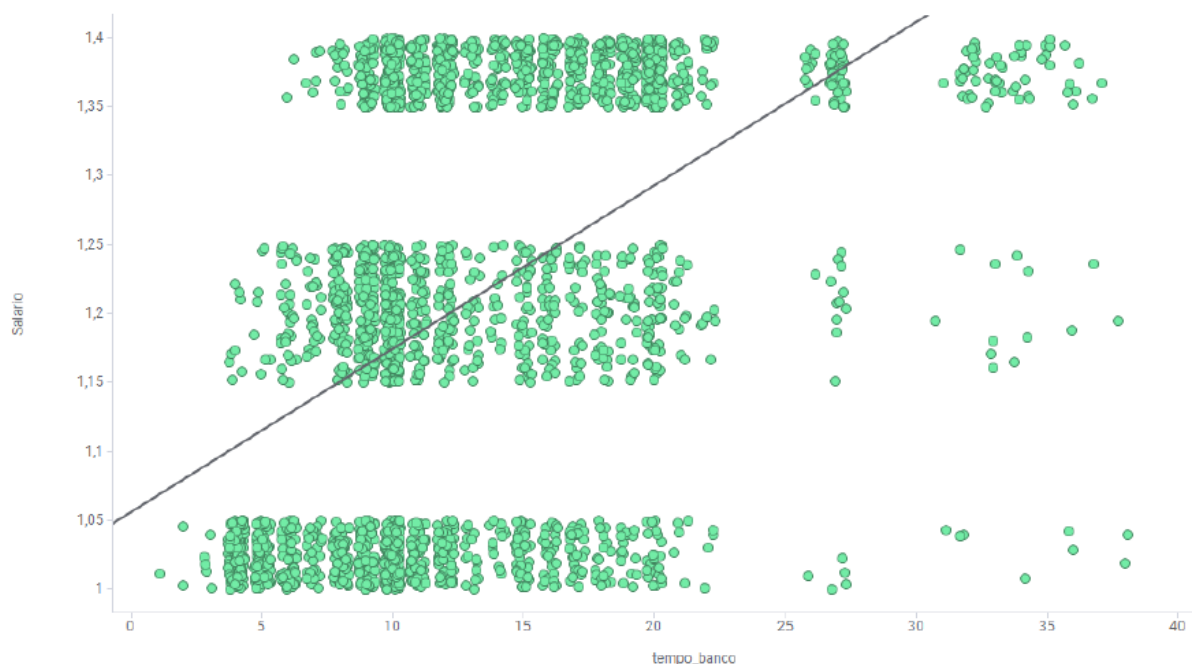


Figura 5.5: Tempo de empresa x Salário - Fonte: O autor

Observa-se uma correlação positiva moderada de 0,43 entre o tempo de empresa e o salário. Esse dado indica que os cargos mais altos tendem a ser ocupados por pessoas com mais tempo de organização. A base de dados analisada abrange analistas de TI em três níveis (júnior, pleno e sênior), sendo que um dos fatores que contribuem para a senioridade é a experiência adquirida ao longo do tempo na função. No gráfico de dispersão (figura 5.5), verifica-se que não há analistas sêniores com menos de 5 anos na organização. Mesmo quando ingressam na empresa com ampla experiência de mercado, esses profissionais passam novamente pelos níveis iniciais da função.

Essa observação levanta a questão de se a organização deveria adotar uma política de sucessão baseada mais em habilidades técnicas (hard skills) e interpessoais (soft skills), em vez de depender predominantemente do tempo de serviço. Essa abordagem poderia melhorar os processos de sucessão e alocação de talentos na empresa.

- Idade x Total de Certificações: -0.27 (Figura 5.6)

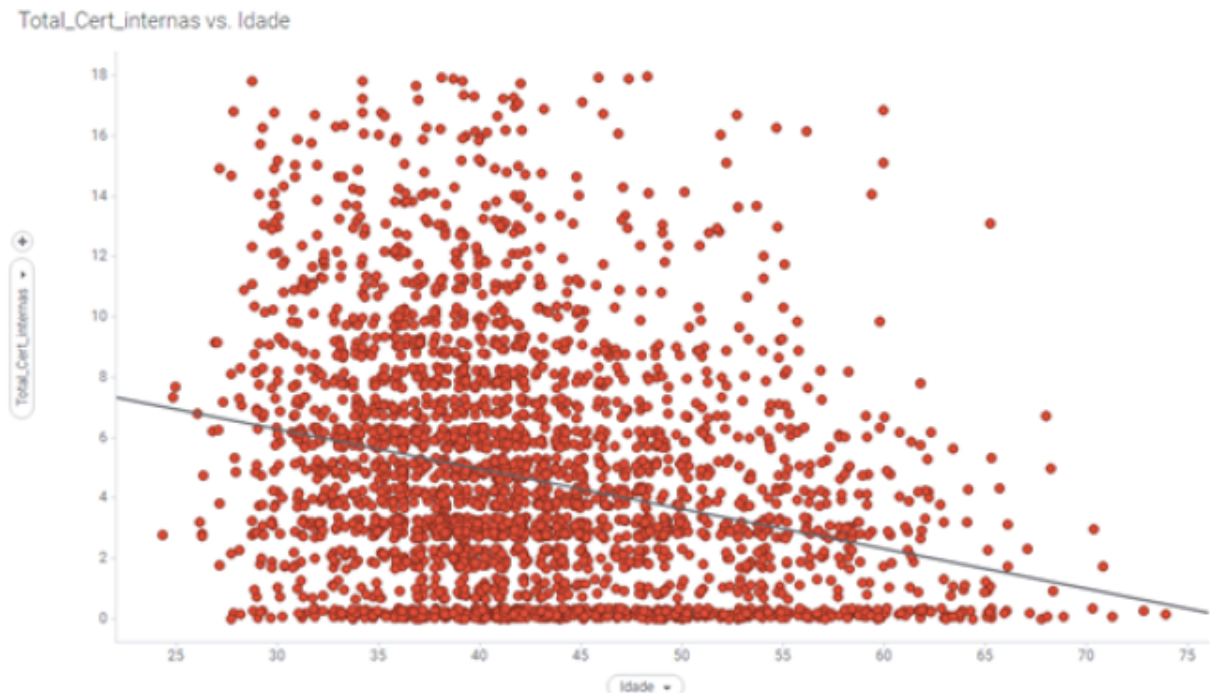


Figura 5.6: Idade x Total de Certificações - Fonte: O autor

Identifica-se uma correlação negativa moderada de -0,27 entre a idade dos colaboradores e a quantidade de certificações profissionais obtidas. Esse dado permite levantar algumas hipóteses, considerando que há uma correlação moderada entre idade e tempo de empresa (0,33). Colaboradores mais velhos tendem a ter mais tempo de empresa, o que pode resultar em possíveis acomodações na carreira, inclusive no interesse pelo desenvolvimento contínuo. Outra hipótese é que, dado o vínculo entre tempo de empresa e salário, colaboradores com mais tempo de organização podem já ter alcançado posições mais estratégicas e táticas, com menor foco em certificações. Apesar dessas hipóteses, o objetivo deste trabalho não é comprová-las, mas exemplificar o potencial do framework abordado.

- Pesquisa Clima x Dias não úteis trabalhados: -0.41 (Figura 5.7)

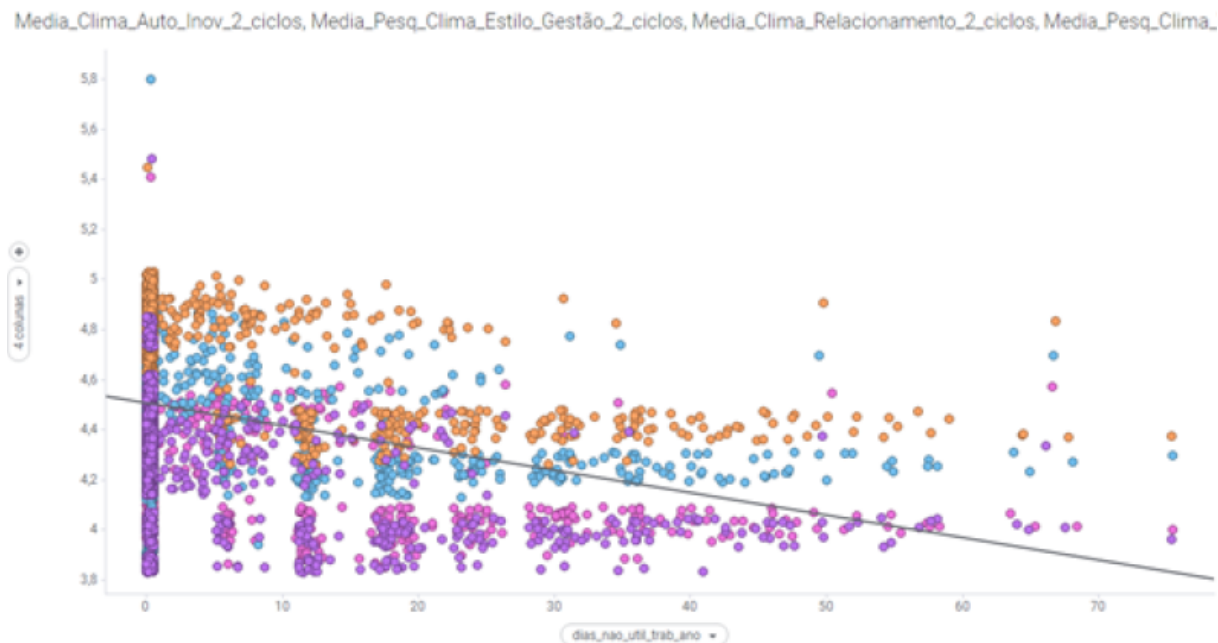


Figura 5.7: Pesquisa Clima x Dias não úteis trabalhados - Fonte: O autor

Nesta correlação negativa forte, identifica-se a possibilidade de que colaboradores que trabalham em dias não úteis apresentem menor satisfação na pesquisa de clima organizacional. Caso essa hipótese se comprove, caberia à organização investigar as causas dessa insatisfação e implementar políticas para revertê-la. Novamente, enfatiza-se que o objetivo não é comprovar hipóteses, mas demonstrar o potencial da etapa de diagnóstico e a aplicabilidade do framework.

- Média Clima Organizacional x Dimensões Clima Organizacional: [Dimensão financeira: 0,65], [Dimensão Clientes: 0,68], [Dimensão Aprendizado: 0,58], [Dimensão Processos Internos: 0,69], [Dimensão Socioambiental: 0,45] (Figura 5.8)



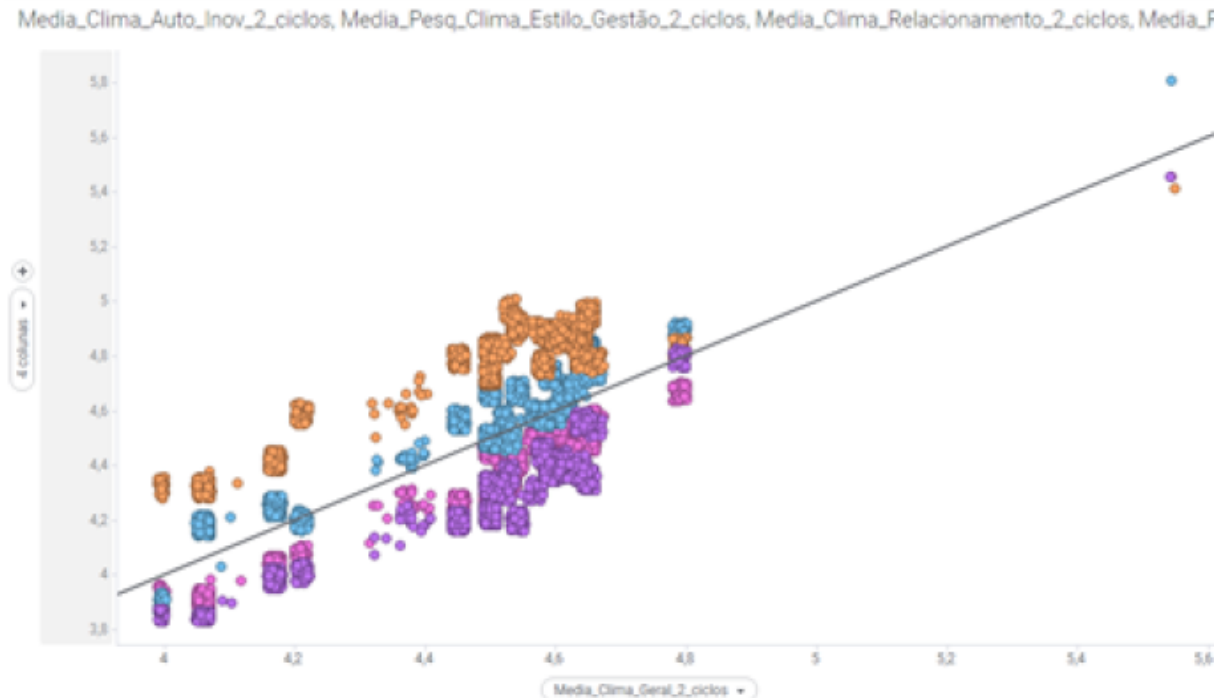


Figura 5.8: Média Clima Organizacional x Dimensões Clima Organizacional - Fonte: O autor

As correlações positivas fortes entre a média da pesquisa de clima e as dimensões da pesquisa indicam que a separação das dimensões pode ser pouco clara para os colaboradores. Eles tendem a avaliar todas as dimensões de forma uniforme, de acordo com seu nível de satisfação geral. Caso essa hipótese se confirme, a organização poderia buscar maior clareza sobre as dimensões com os colaboradores, visando uma pesquisa de clima mais detalhada e realista. Reitera-se que o objetivo desta abordagem é demonstrar a viabilidade do framework, e não validar hipóteses.

- Idade x Dimensões Desempenho: [Financeira: -0,10], [Cliente: -0,22], [Aprendizado: -0,23], [Processos Internos: -0,19] (Figura 5.9)

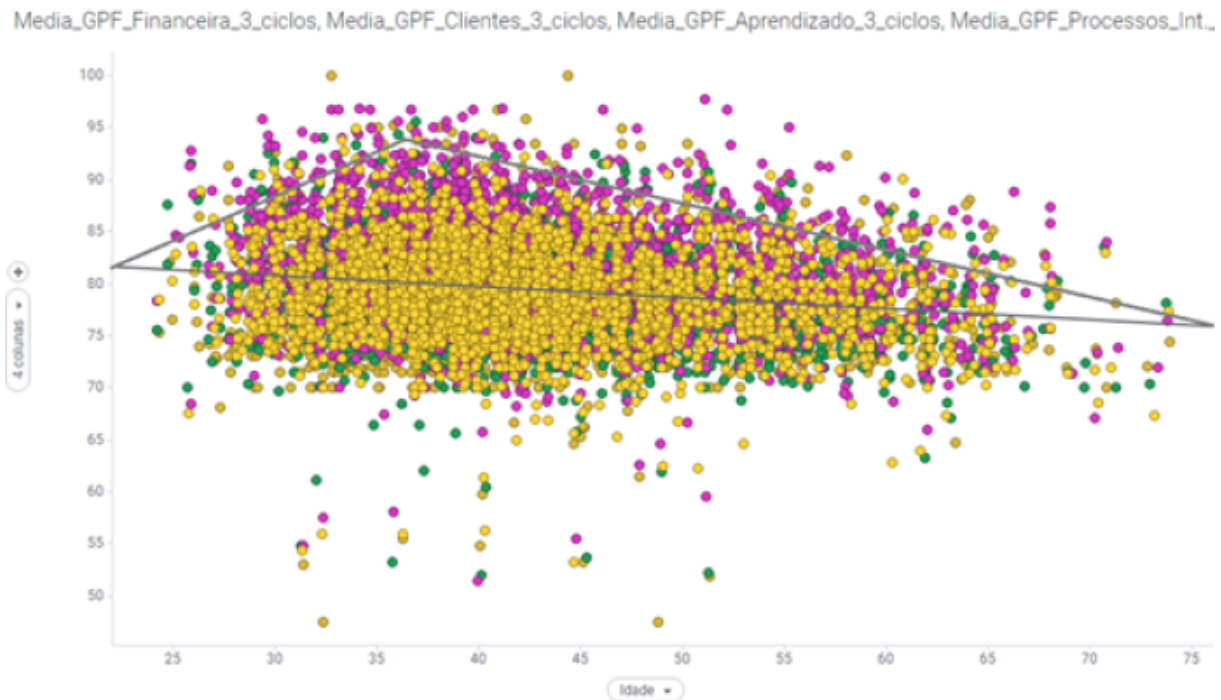


Figura 5.9: Idade x Dimensões Desempenho: - Fonte: O autor

Observa-se uma correlação negativa moderada entre a idade e as dimensões das avaliações de desempenho, sugerindo que, quanto maior a idade, menor a avaliação de desempenho. Uma hipótese possível é que, dado o vínculo entre tempo de empresa e salário, e entre idade e tempo de empresa, colaboradores mais velhos tendem a ocupar cargos mais altos. Esses cargos, que requerem habilidades mais avançadas, podem resultar em avaliações mais rigorosas. Mais uma vez, destaca-se que o objetivo não é testar essas hipóteses, mas demonstrar a capacidade do framework em gerar insights úteis para a organização.

- Média Desempenho x Horas Extras: 0,22 (Figura 5.10)

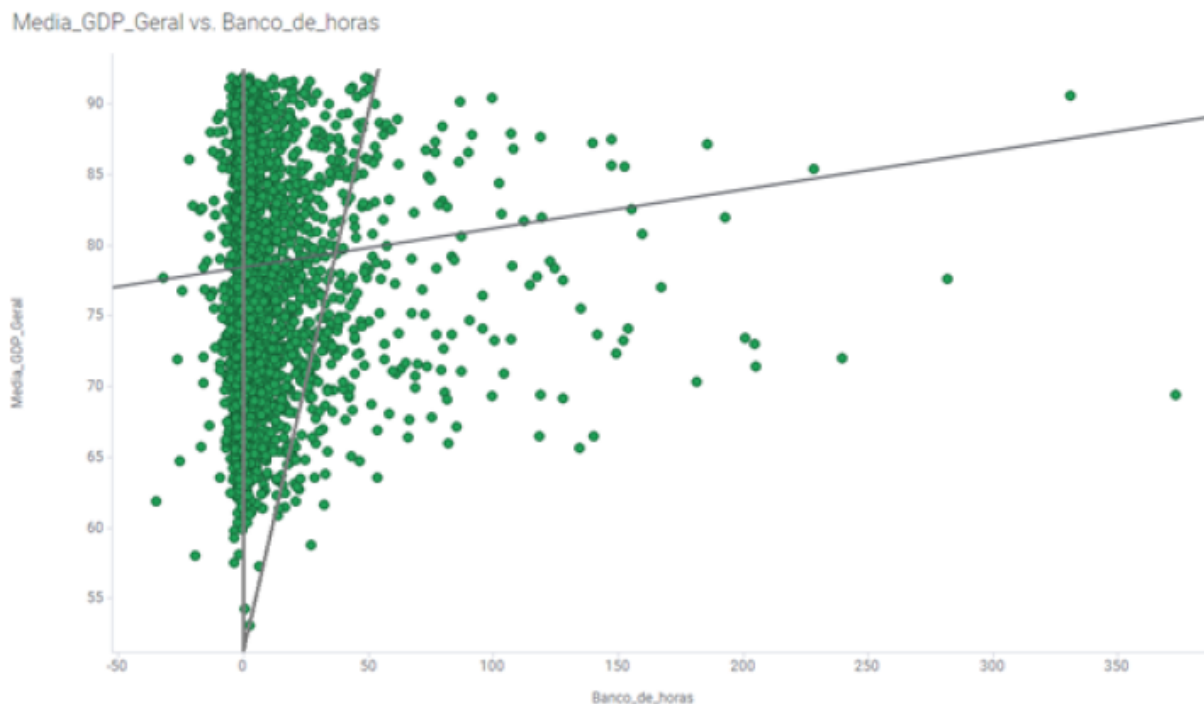


Figura 5.10: Média Desempenho x Horas Extras: - Fonte: O autor

Nesta correlação positiva moderada, verifica-se que a média do desempenho dos colaboradores é maior entre aqueles que realizam mais horas extras. Isso pode indicar que a organização direciona investimentos em horas extras para colaboradores com maior capacidade de entrega. Se essa hipótese for verdadeira, trata-se de uma demonstração de uma boa estratégia de alocação de recursos. Reitera-se que o objetivo é explorar o potencial do framework proposto em trazer insights relevantes para a organização.

### Visão Geral do Nível de Maturidade de Diagnóstico

Com base nas análises e metodologias descritas neste capítulo, foram identificadas diversas correlações que fornecem insights relevantes para a organização. Utilizando técnicas estatísticas, como o Coeficiente de Correlação Linear de Pearson, foram detectados padrões significativos que podem embasar futuras decisões estratégicas. As correlações apresentadas indicam tendências capazes de gerar hipóteses que, ao serem validadas ou não, têm o potencial de oferecer à organização respostas para questões que, muitas vezes, não eram previamente conhecidas.

Embora o objetivo deste trabalho não seja a validação de hipóteses, as inferências realizadas demonstram o potencial do framework proposto em oferecer uma visão clara e útil para a tomada de decisões. Além disso, foi apresentada uma abordagem complementar com foco na melhoria de processos por meio de técnicas estatísticas simples, permitindo

à organização avaliar se as estratégias utilizadas são as mais adequadas para o contexto analisado.

Conclui-se, assim, que a aplicação deste framework pode contribuir significativamente para um diagnóstico mais preciso e eficaz, auxiliando a organização na implementação de políticas e estratégias mais direcionadas e fundamentadas.

### 5.3 Nível de maturidade: Análise Preditiva

O próximo passo na jornada de maturidade do *People Analytics* é a aplicação de análises preditivas. Enquanto as análises descritivas e diagnósticas fornecem uma visão clara do que ocorreu no passado e identificam relações entre variáveis, as análises preditivas vão além, oferecendo a capacidade de prever eventos futuros e tendências com base em dados históricos e padrões observados.

As análises preditivas utilizam uma variedade de técnicas estatísticas e algoritmos de machine learning para construir modelos que podem antecipar resultados futuros. Esses modelos são treinados com dados históricos e, uma vez validados, são aplicados para prever cenários futuros. No contexto de *People Analytics*, essas previsões podem incluir tendências de rotatividade de colaboradores, probabilidade de promoção, níveis de engajamento futuro, entre outros.

A aplicação de análises preditivas traz diversos benefícios para a gestão de pessoas, permitindo que as organizações se antecipem a problemas potenciais e capitalizem oportunidades. Por exemplo, ao prever quais colaboradores estão mais propensos a deixar a empresa, as equipes de recursos humanos podem implementar estratégias de retenção direcionadas, minimizando o impacto da rotatividade. Da mesma forma, prever o desempenho futuro de colaboradores pode ajudar a identificar potenciais líderes e planejar sucessões de forma mais eficaz.

Para implementar análises preditivas de forma eficaz, é crucial que a organização possua uma infraestrutura de dados robusta, com dados de alta qualidade e bem integrados. Além disso, é necessário investir em ferramentas de análise avançadas e capacitar a equipe para utilizar essas ferramentas. A construção de um modelo preditivo também requer uma compreensão profunda dos fatores que influenciam os resultados que se deseja prever, além de uma abordagem iterativa para testar e refinar os modelos.

Nos próximos parágrafos, apresentaremos uma aplicação prática do framework de análises preditivas em *People Analytics*, demonstrando como essas técnicas podem ser implementadas para gerar insights valiosos e orientar decisões estratégicas na gestão de pessoas.

### **5.3.1 Caso de uso - Nível Preditivo: Identificação de talentos para sucessão**

No ambiente de negócios altamente competitivo de hoje, a aquisição de talentos e a cuidadosa seleção de profissionais qualificados são processos fundamentais para o sucesso e crescimento sustentável das organizações. Particularmente quando se trata de posições de liderança, fazer a escolha certa dos indivíduos desempenha um papel crucial na direção estratégica, eficiência operacional e no desenvolvimento de uma cultura organizacional positiva. Nesse contexto, técnicas avançadas de análise de dados e aprendizado de máquina têm desempenhado um papel cada vez mais importante no recrutamento e seleção de talentos, fornecendo insights valiosos e preditivos para identificar os candidatos mais adequados para posições-chave. Este artigo aborda especificamente o estudo de caso de uma empresa financeira pública onde técnicas de aprendizado de máquina foram empregadas para aprimorar o processo de recrutamento e sucessão de liderança em seu setor de TI. Abordagens como essas têm se mostrado eficazes no suporte à tomada de decisões em processos de seleção de talentos, auxiliando na identificação e avaliação de candidatos.

Um dos processos mais desafiadores e estratégicos dentro de uma organização é o recrutamento eficiente da força de trabalho adequada [43]. Esse desafio se torna ainda mais complexo quando o foco está na sucessão de liderança [26]. Posições de maior responsabilidade exigem processos de recrutamento mais cuidadosos, pois esses profissionais estarão envolvidos na tomada de decisões estratégicas e táticas, afetando diretamente o desempenho das organizações [33].

Em grandes empresas, os processos de sucessão podem apresentar ineficiências, tornando a escolha dos candidatos mais adequados para as posições disponíveis ainda mais complexa. Grandes organizações frequentemente possuem avaliações de desempenho não padronizadas, pois diferentes gestores têm percepções individuais com diferentes níveis de severidade e leniência. Além disso, há um ponto crítico quando se trata de profissionais técnicos, como aqueles em TI, uma vez que uma boa avaliação de desempenho nem sempre garante o potencial para ser um bom gestor, criando um problema duplo: perde-se um bom técnico e ganha-se um gestor com desempenho abaixo do esperado [16].

O recrutamento eficiente com foco na sucessão, especialmente para posições de liderança, pode trazer vários benefícios para o negócio, como: melhor gestão de processos, maior alinhamento com os objetivos organizacionais, maior motivação da equipe, redução da rotatividade, aumento da produtividade, entre outros. Por outro lado, escolhas inadequadas podem ter um alto custo para a organização, como: baixo desempenho, insatisfação dos colaboradores, sensação de desvalorização, aumento no número de demissões voluntárias, ambiente de trabalho comprometido, não atingimento dos objetivos organizacionais, entre outros. Para [16], ao recrutar para novas posições, a prioridade não deve ser ter um

grande número de candidatos, mas sim candidatos qualificados.

Partindo do estudo de caso de uma empresa financeira pública, este artigo visa verificar a viabilidade do uso de técnicas de mineração de dados e aprendizado de máquina na análise preditiva da prontidão de talentos internos para a posição de gerente de equipe, a partir do grupo de Analistas Sêniores de TI. Além disso, os parâmetros definidos pelo modelo serão avaliados por métricas estatísticas e objetivos de negócios, a fim de verificar se esses resultados atendem ou não aos parâmetros mínimos sugeridos para a obtenção da posição-alvo.

O uso de técnicas estatísticas e aprendizado de máquina nos desafios de atração, retenção e sucessão de recursos humanos é chamado de *People Analytics*, uma abordagem que utiliza dados, análise e métodos estatísticos para entender e otimizar a força de trabalho de uma organização. Em um nível amplo, *People Analytics* consiste em aplicar evidências nas decisões de gestão sobre pessoas [54].

Focando no conceito de *People Analytics*, o objetivo deste estudo é verificar a viabilidade do uso de aprendizado de máquina para identificar profissionais de TI prontos para a sucessão ao cargo de gerente de equipe. No estudo a seguir, identificou-se que os modelos apresentados trazem benefícios ao processo de tomada de decisão e podem servir como uma pré-seleção.

## **Compreensão do Negócio**

Esta etapa envolve entender os desafios e necessidades do negócio, identificando as necessidades sem vieses inconscientes. O objetivo do projeto foi definido como a criação de entregas valiosas para apoiar a tomada de decisões nos processos de sucessão de Analista Sênior de TI para Gerente de Equipe, por meio do uso de mineração e aprendizado de máquina. Dentro deste escopo, o modelo busca uma mediana da Taxa de Verdadeiros Negativos, ou seja, prevê parte dos Analistas de TI incorretamente, assumindo que estes seriam os mais preparados para a sucessão. O foco será então na busca de falsos negativos da classe 0 (Analistas).

## **Compreensão do Negócio**

Esta etapa consistiu em compreender os desafios e necessidades do negócio, identificando os requisitos sem vieses inconscientes. O objetivo do projeto foi definido como a criação de entregas valiosas para apoiar a tomada de decisões nos processos de sucessão de Analista Sênior de TI para Gerente de Equipe, por meio do uso de mineração de dados e aprendizado de máquina. Dentro deste escopo, o modelo busca uma mediana da Taxa de Verdadeiros Negativos (*True Negative Rate* – TNR), ou seja, prever parte dos Analistas

de TI incorretamente, assumindo que estes seriam os mais preparados para a sucessão. O foco está, portanto, na busca de falsos negativos da classe 0 (Analistas).

## Compreensão dos Dados

Esta etapa teve como objetivo a análise dos bancos de dados disponíveis que impactam o negócio, bem como de outros dados necessários e sua disponibilidade para o estudo. O conjunto de dados utilizado é do ano de 2020 e envolve 1.521 registros reais do departamento de TI de uma instituição financeira pública (sendo 1.161 Analistas Seniores e 360 Gerentes de Equipe, representando 100% do público avaliado) e 21 atributos, conforme descrito na Figura 5.11.

```
geral.columns  
  
Index(['gender', 'age', 'marital_status', 'company_tenure',  
      'level_of_education', 'job_description', 'department', 'dependents',  
      'vacation', 'days_absent', 'days_of_absence', 'punctuality_score',  
      'bonuses', 'days_off', 'total_absences', 'training_hours_per_week',  
      'state', 'salary', 'performance_review_3_cycles',  
      'climate_survey_review_3_cycles', 'climate_survey_evaluation'],  
      dtype='object')
```

Figura 5.11: Atributos

Para facilitar a compreensão dos dados, foi realizada uma análise exploratória utilizando a linguagem *Python*, a biblioteca *matplotlib.pyplot* e técnicas de distribuição visual, como histogramas, *boxplots* e o coeficiente de correlação de Pearson.

## Preparação de Dados

A etapa de preparação de dados é amplamente reconhecida como a mais trabalhosa em projetos de aprendizado de máquina. Esta fase, embora complexa, é crucial para garantir que os dados estejam em um formato que possa ser efetivamente interpretado e utilizado pelos algoritmos. Um cuidado meticuloso na preparação dos dados é essencial para a construção de modelos robustos e precisos, uma vez que mesmo o algoritmo mais avançado não compensa dados mal preparados. Nesta etapa, foram realizados os seguintes procedimentos:

### Verificação de Valores Ausentes

A identificação e o tratamento de valores ausentes são etapas cruciais no processo de preparação de dados dentro de projetos de ciência de dados, conforme delineado no modelo CRISP-DM. A ausência de dados pode levar a distorções, erros e conclusões imprecisas.

No entanto, o banco de dados analisado apresentou integridade completa, sem valores ausentes, como mostrado na Figura 5.12.

```
print("The number of missing values:",  
      geral.isnull().sum().sum())
```

```
The number of missing values: 0
```

Figura 5.12: Valores faltantes - Fonte: O autor

### Remoção de Atributos Correlacionados

A remoção de atributos correlacionados foi realizada para evitar multicolinearidade e redundância de informações, fatores que podem levar a estimativas imprecisas ou *overfitting*. Neste trabalho, foram removidos os atributos `salário`, devido à sua correlação direta com o atributo alvo `desc_cargo`, e `dias_ausentes` e `total_ausências`, que apresentavam correlação direta entre si.

### Remoção de Atributos Desnecessários

Com base na análise de negócios, os seguintes atributos foram removidos por não agregarem valor ao modelo: `férias`, `dias_de_folga`, `UF`, `pontualidade_eps`, `auxílios`, `dependentes`, `estado_civil` e `avaliação_da_pesquisa_de_clima`.

### Atribuição de Valores Sequenciais a Atributos Categóricos Ordenados

A identificação de atributos categóricos com ordem natural levou à conversão numérica sequencial, promovendo melhor interpretabilidade pelo modelo. O atributo `grau_de_instrução` foi tratado conforme o código apresentado na Figura 5.13.

```
geral['grau_instr'] =  
geral['grau_instr'].replace  
(["Medio", "Superior em Andamento", "Superior",  
 "Graduacao", "Especializacao", "Mestrado",  
 "Doutorado"], [1, 2, 2, 2, 3, 3, 4, 5])
```

Figura 5.13: Level de escolaridade - Fonte: O autor

### Transformação de Variáveis Categóricas Nominiais em Variáveis *Dummy*

Para transformar variáveis categóricas nominiais em um formato numérico, de forma a serem interpretadas pelo modelo, foi utilizada a técnica de variáveis *dummy*, conforme demonstrado na Figura 5.14.



```
#transforming remaining category variables into dummy
geral = pd.get_dummies(geral, columns = ['job_description'])
geral = pd.get_dummies(geral, columns = ['gender'])
geral = pd.get_dummies(geral, columns = ['departament'])
```

Figura 5.14: Variáveis Dummy - Fonte: O autor

## Recorte da Base de Gerentes

Como o objetivo do modelo preditivo era identificar analistas seniores aptos a se tornarem gerentes de alto desempenho, a base foi filtrada para incluir apenas gerentes com avaliação de desempenho acima de 4,0 (em uma escala de 1 a 5). A base final consolidou-se com 1.265 registros, sendo 104 gerentes e 1.161 analistas.

## Balanceamento de Classes

Para evitar viés na classificação devido à disparidade de distribuição entre classes, foi utilizada a técnica *RandomOverSampler*. Isso aumentou proporcionalmente o número de indivíduos na classe minoritária, conforme mostrado na Figura 5.15.

```
from imblearn.over_sampling import RandomOverSampler
sobreamostrador1 = RandomOverSampler(random_state=42)
x_treino4_oversampled1, y_treino4_oversampled1 =
sobreamostrador1.fit_resample(x_treino4, y_treino4)
```

Figura 5.15: Balanceamento de classes - Fonte: O autor

## Tratamento de *Outliers*

Nenhum *outlier* foi tratado, pois esses dados representam uma variabilidade natural e sua remoção poderia levar à perda de informações importantes.

## Modelagem

A variável alvo definida foi `desc_cargo`. Com base na análise de negócio, buscou-se identificar Analistas Seniores de TI mais aptos a se tornarem gerentes de alto desempenho. Para isso, apenas Gerentes de Equipe que apresentaram avaliação de desempenho superior à média foram mantidos na base.

Conforme mencionado anteriormente, os dados analisados são de 2020. Assim, foi possível promover uma validação no mundo real ao comparar o resultado do modelo com os Analistas que se tornaram gerentes entre a data da base e o momento atual. Dado o tamanho do conjunto de dados, optou-se por dividi-lo aleatoriamente em 50% para a base

de treinamento e 50% para a base de teste. A validação foi realizada utilizando o próprio cenário real, conforme ilustrado nas Figuras ?? e 5.17.

```
from sklearn.model_selection import train_test_split
teste, treino = train_test_split
(geral, test_size=0.5, random_state=42)
```

Figura 5.16: Base de treino e teste - Fonte: O autor

```
y_train = train['job_description_team_manager']
x_train = train.drop(['job_description_team_manager'], axis = 1)
y_test = test['job_description_team_manager']
x_test = test.drop(['job_description_team_manager'], axis = 1)
```

Figura 5.17: X e Y - Fonte: O autor

As seguintes técnicas preditivas foram utilizadas: *Random Forest Classifier*, *XGBoost* e Regressão Logística.

### Avaliação dos Resultados

Esta etapa consistiu na análise dos resultados obtidos, tanto das fases exploratórias quanto das preditivas. No contexto deste modelo, o foco principal foi a identificação de falsos positivos no grupo de Gerentes de Equipe. Esses falsos positivos correspondem a Analistas de TI classificados como gerentes, embora não ocupem essa função, mas que têm potencial para sucessão.

Portanto, métricas importantes para este estudo foram a Especificidade ou Taxa de Verdadeiros Negativos (*True Negative Rate* – TNR), que avalia falsos positivos na classe 0, e o *Recall*, que mede a precisão na identificação de verdadeiros positivos (gerentes). A métrica mais relevante, no entanto, foi a capacidade do modelo de prever um cenário ainda inexistente à época da análise, comparando as previsões com os analistas que efetivamente foram promovidos a gerentes após a extração da base. Entre os 582 analistas presentes na base de teste, 52 foram promovidos à posição de gerente de equipe ou similar até novembro de 2023.

Entre os modelos avaliados, dois apresentaram alta precisão e um, precisão média: *Random Forest* - 91,7%; *XGBoost* - 88,1%; Regressão Logística - 63%. Contudo, apenas a Regressão Logística apresentou um *recall* elevado (acertos na classe 1) combinado a uma especificidade média (acertos na classe 0). Nas Figuras 5.18 e 5.19, são apresentados os índices de *recall* e especificidade do modelo de Regressão Logística:

```
recall = recall_score(y_teste4,predictions4)
print("Recall:", recall)
```

Recall: 0.74

Figura 5.18: Recall - Fonte: O autor

```
#Especificity or True Negative Rate - TNR.
from sklearn.metrics import confusion_matrix
tn, fp, fn, tp = confusion_matrix(y_teste4,predictions4).ravel()
specificity = tn / (tn+fp)
print("True Negative Rate:", specificity)
```

True Negative Rate: 0.627147766323024

Figura 5.19: Especificidade - Fonte: O autor

Pode-se observar na Matriz de Confusão apresentada na Figura 5.20 que o modelo demonstrou maior eficiência na identificação correta de gerentes em comparação à identificação de analistas, como era esperado. O modelo previu corretamente 37 de 50 gerentes (74% de precisão) e apenas 365 de 582 analistas (62% de precisão).

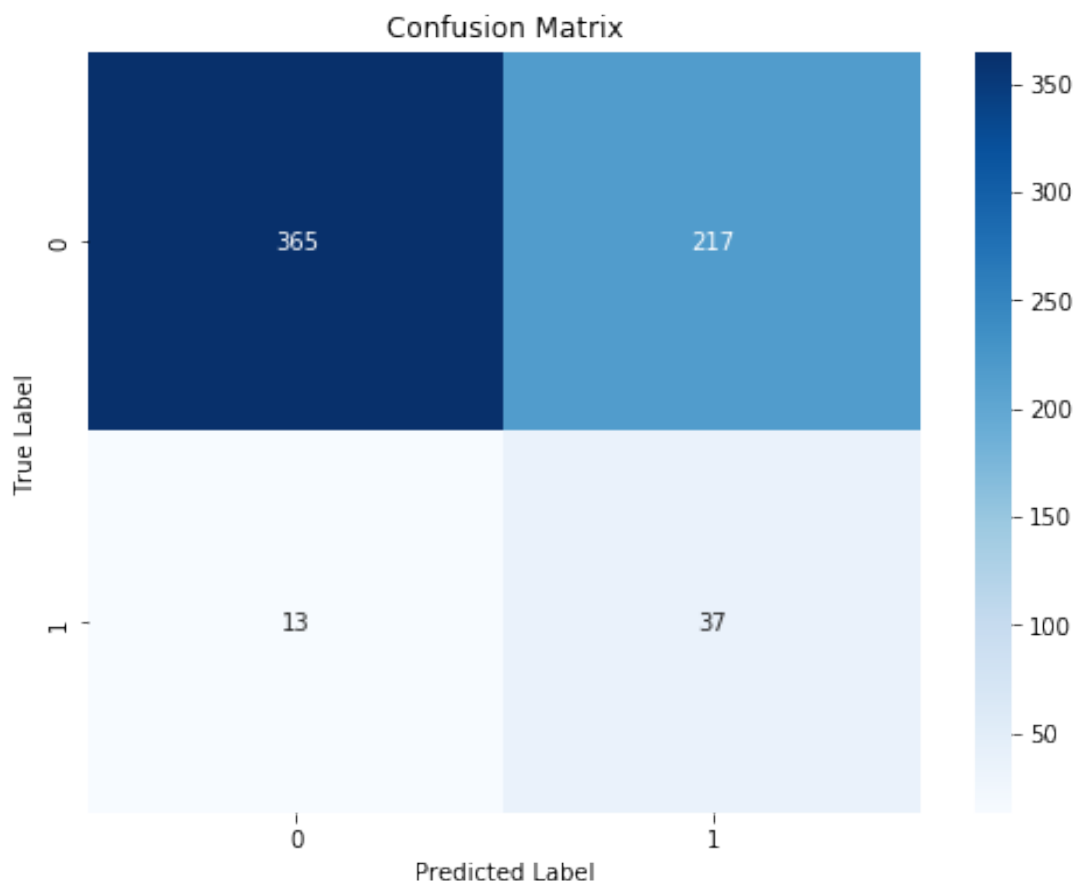


Figura 5.20: Matriz de confusão - Fonte: O autor

Considerando os dados de maio de 2023, entre os 582 analistas presentes no conjunto de teste, 52 foram promovidos à posição de gerente de equipe ou uma posição similar. A etapa seguinte da análise envolveu validar os 217 falsos positivos no mundo real, juntamente com os 52 indivíduos promovidos. Observou-se que 33 dos 52 gerentes estavam incluídos nos 217 falsos positivos, o que indica que, entre os falsos positivos identificados, aproximadamente 1 em cada 6 foi efetivamente promovido em um intervalo de três anos. Quando considerado o conjunto completo de dados de teste, esse número muda para 1 em 11 (52 promovidos de 582).

Os dados analisados indicam que o modelo se mostrou útil como ferramenta de apoio à decisão, exibindo uma eficiência quase duas vezes superior à proporção de promoções observadas no conjunto total de dados. Estudos futuros podem aprofundar essa análise. Além disso, a análise da importância dos pesos atribuídos a cada variável no modelo, como mostrado na Figura ??, revelou alinhamento com as políticas da empresa. Isso se evidenciou ao observar que o desempenho dos funcionários foi a variável com maior impacto, seguido por pesos atribuídos aos funcionários já em posições de gestão. Neste caso, o peso inverso para funções não técnicas reflete uma menor probabilidade de essas funções

serem escolhidas como gerentes de TI. Outro ponto relevante foi o peso ligeiramente maior atribuído ao gênero feminino, o que está em consonância com as políticas organizacionais que visam promover a liderança feminina.

|                        | importance |
|------------------------|------------|
| Performance_evaluation | 2.744795   |
| Area_Management        | 2.211889   |
| Gender_F               | 1.720580   |
| Gender_M               | 1.150773   |
| climate_survey         | 1.064236   |
| Area_Infrastructure    | 0.617848   |
| Education_level        | 0.069639   |
| Age                    | 0.058841   |
| Area_Construction      | 0.041616   |
| Company_ternure        | 0.032946   |
| Training_hours_sem     | 0.008826   |

Figura 5.21: Attribute weights - Fonte: O autor

### 5.3.2 Conclusão do caso de uso preditivo

Este estudo examinou a aplicação de técnicas de análise de dados e aprendizado de máquina para apoiar a tomada de decisões nos processos de sucessão para cargos de liderança no departamento de TI de uma empresa financeira pública. O principal objetivo foi identificar, entre os analistas seniores, aqueles com potencial para assumir posições de gerentes de equipe.

Os resultados obtidos confirmaram a eficácia do modelo proposto. Técnicas preditivas, incluindo *Random Forest Classifier*, *XGBoost* e Regressão Logística, foram aplicadas. Notavelmente, a Regressão Logística, embora não tenha apresentado a maior precisão, foi a que melhor respondeu ao desafio específico deste estudo. Falsos positivos na classe de gerentes, que normalmente são considerados erros em muitos estudos, neste caso, representaram analistas com potencial de sucessão. Essa abordagem revela uma inovação na metodologia, ao atribuir valor significativo aos falsos positivos.

Com base em um conjunto de dados de 2020, o modelo de Regressão Logística identificou 217 analistas seniores como potenciais gerentes de equipe. Ao validar essas previsões com dados do mundo real em maio de 2023, constatou-se que 33 dos 52 indivíduos promovidos a gerentes de equipe entre 2020 e 2023 estavam incluídos nos 217 inicialmente indicados pelo modelo.

Isso implica que, entre os 217 apontados como potenciais gerentes, a proporção de promoção foi de 1 para cada 6 indivíduos (33 de 217) ao longo de três anos. Em comparação,

considerando a base total de analistas, essa proporção foi de 1 para 11 (52 de 582). Assim, o modelo demonstrou quase o dobro da eficiência na identificação de potenciais gerentes em relação a uma seleção aleatória.

É relevante destacar que a base de validação utilizada para gerar as métricas de resultado não foi subtraída da base original de treinamento e teste, mas sim representou o cenário do mundo real. A base de treinamento e teste referia-se a 2020, enquanto a base de validação refletia os dados atuais. Esse aspecto reforça a aplicabilidade prática e a relevância dos resultados obtidos.

Adicionalmente, os atributos selecionados pelo modelo para identificar potenciais gerentes estavam alinhados com as políticas corporativas. A avaliação de desempenho, como esperado, foi o atributo com maior peso, evidenciando a importância do desempenho individual no processo de sucessão. Um ponto interessante foi o peso atribuído ao atributo "gênero feminino", que apresentou uma leve vantagem, refletindo a política da empresa de incentivar a liderança feminina. Isso demonstra que o modelo conseguiu capturar e refletir as políticas e prioridades da organização.

Em síntese, este estudo evidenciou o potencial das técnicas de análise de dados e aprendizado de máquina para apoiar os processos de sucessão de liderança. A abordagem inovadora de considerar falsos positivos como indicadores de liderança em potencial, a validação utilizando dados reais posteriores ao período de treinamento e teste, e a capacidade do modelo de refletir as políticas organizacionais representam avanços significativos. Contudo, é fundamental enfatizar que essas técnicas devem ser vistas como complementos ao julgamento humano, e não como substitutos. Recomenda-se a continuidade das pesquisas para explorar e refinar essas técnicas, bem como investigar outras possíveis aplicações. Além disso, é essencial que as organizações mantenham investimentos em pesquisa e desenvolvimento, a fim de aprimorar esses modelos e maximizar seu potencial.

## 5.4 Nível de maturidade Prescritivo

O objetivo desta seção é demonstrar um framework do mais alto nível de maturidade do *People Analytics*. Enquanto as etapas anteriores se concentravam em demonstrar o que acontece? (Análise Descritiva), as buscas pela causa ou o porque acontece? (Análise Diagnóstica), o que vai acontecer? (Análise Preditiva), o nível de maturidade descritivo busca trazer uma aplicação prática de forma a fazer acontecer. E justamente por isso essa etapa não possui uma fórmula pronta, necessitando de conhecimento técnico e conhecimento do negócio a um nível profundo o suficiente para ter a capacidade na busca de soluções customizadas frente aos desafios da organização. O caso prático apresentado como exemplo para o nível prescritivo aborda o problema de negócio da alocação de novos

funcionários nas equipes mais adequadas ao perfil dentro da organização, o case em questão foi aplicado em no processo de atração e onboarding de novos funcionários oriundos de seleção externa de um grande banco brasileiro.

#### **5.4.1 Caso de uso - Alocação ideal de colaboradores - Contexto Geral**

No contexto dinâmico dos ambientes empresariais contemporâneos, gerenciar efetivamente o talento por meio do alinhamento estratégico com os objetivos organizacionais é cada vez mais crucial. Este estudo apresenta uma estrutura abrangente para a otimização da alocação de talentos que integra três técnicas avançadas de otimização: Maximum Cardinality Matching, Otimização Híbrida e o Método Simplex. Projetada para alinhar as habilidades e interesses dos funcionários com as necessidades organizacionais, esta estrutura promete um ajuste estratégico que melhora tanto a satisfação individual quanto a produtividade organizacional. Uma aplicação prática é demonstrada por meio de um estudo de caso no Banco do Brasil S.A., onde 409 novos funcionários foram alocados usando a estrutura proposta. A eficácia do modelo é destacada por seu desempenho superior em alavancar as competências dos funcionários em comparação com métodos tradicionais baseados na intuição, obtendo resultados significativamente melhores do que alocações aleatórias. Este artigo contribui para o campo da Gestão de Recursos Humanos ao oferecer um modelo baseado em dados, repetível, que pode ser aplicado em vários contextos organizacionais para melhorar tanto o engajamento dos funcionários quanto os resultados empresariais.

O recrutamento eficaz da força de trabalho certa representa um dos processos mais estratégicos e desafiadores dentro de uma organização [43]. Empresas que se engajam em extensos processos de atração enfrentam o desafio crucial de garantir a alocação correta de novos talentos, maximizando o uso de suas habilidades. Diversas métricas são utilizadas para identificar talentos, geralmente categorizadas em três principais dimensões: desempenho, potencial e dinâmica de desempenho [7]. Este desafio é ainda mais acentuado com a chegada de novos funcionários, pois a organização ainda não possui métricas estabelecidas para aqueles que estão no início de sua jornada e na fase de integração.

Adicionado a esse desafio, há uma característica típica da gestão de pessoas: o uso da intuição como um elemento de validação nos processos de atração, sucessão e alocação de talentos [48].

Em 1620, Francis Bacon argumentou que o conhecimento científico não deveria se basear apenas em hipóteses intuitivas, pois estas podem ser restritivas ao estabelecer uma premissa antes mesmo de realizar um experimento, limitando assim o processo de raciocínio

para se adequar a essa premissa. Em resposta a essa limitação e em busca de uma gestão de pessoas baseada em dados, o conceito de *People Analytics* surgiu em 2008, em colaboração com Mike West [54]. Utilizando técnicas estatísticas, aprendizado de máquina e inteligência artificial, este conceito visa gerar dados e informações que apoiem a tomada de decisões na gestão de recursos humanos dentro das organizações.

Este estudo tem como objetivo desenvolver uma estrutura que possa ser aplicada em organizações que simultaneamente possuem um grande número de vagas, uma diversidade de posições e complexidade nos temas. Duas possíveis aplicações podem ser propostas: melhor alocação de mão-de-obra recém-contratada (principalmente aplicável em organizações onde o processo de atração é realizado por meio de concursos públicos, já que este modelo não inclui etapas de seleção como entrevistas técnicas) e a realocação de funcionários já presentes na organização (aplicável a todas as organizações). O estudo aborda a busca por alocação ótima de uma perspectiva global, alinhando as habilidades técnicas e os interesses dos funcionários com as necessidades da organização.

Para ilustrar a estrutura apresentada, a aplicação prática da técnica será discutida usando um caso real, alocando 409 novos funcionários em 707 vagas disponíveis na área de tecnologia do Banco do Brasil S.A., a maior instituição financeira da América Latina.

A estrutura consiste em quatro etapas: Mapeamento de habilidades técnicas, Mapeamento dos interesses dos funcionários, Mapeamento das necessidades organizacionais e alocação ótima. Comparativamente, serão utilizadas três técnicas de otimização distintas: Maximum Cardinality Matching, Otimização Híbrida (Heurística + Branch and cut + simplex primal + Crossover) e o Método Simplex.

## 5.4.2 Compreensão do Negócio

Tradicionalmente, em processos extensivos de recrutamento de talentos, a atribuição de candidatos a vagas de emprego é determinada por consenso entre os gestores. Um problema significativo com esse método manual é que líderes mais influentes muitas vezes garantem os melhores talentos, resultando em uma distribuição desigual de recursos humanos. Além disso, buscar uma alocação ideal que atenda tanto às necessidades dos novos funcionários quanto da organização é improvável com métodos manuais. Isso leva a uma redução na satisfação dos recém-contratados e na eficiência de utilização de suas habilidades e conhecimentos existentes.

O desafio deste estudo é baseado na capacidade de criar um método aplicável a qualquer organização com um grande número e complexidade de posições, reconhecendo a necessidade de uma melhor alocação técnica de seus funcionários, alinhando desempenho e satisfação. Como evidência da eficácia do método, será apresentado um estudo de caso real envolvendo a alocação de 409 novos funcionários de tecnologia no Banco do Brasil



S.A., a maior instituição financeira da América Latina. A estrutura proposta visa alcançar a alocação ótima dos funcionários, alinhando suas habilidades técnicas e interesses de trabalho com as necessidades da organização. O modelo é estruturado em quatro etapas distintas:

1º) Mapeamento das Necessidades Organizacionais: Esta etapa envolve trabalhar com a organização para identificar os requisitos específicos de conhecimento e assuntos para cada posição disponível.

2º) Mapeamento de Habilidades Técnicas: Esta etapa envolve identificar e documentar as habilidades técnicas dos funcionários que são relevantes e valiosas para a organização.

3º) Mapeamento dos Interesses dos Funcionários: Isso envolve mapear os interesses relacionados ao trabalho dos funcionários para entender melhor suas preferências e inclinações.

4º) Mapeamento de Mídias Sociais: Este conceito envolve a análise, por meio de autorização e fornecimento de links pelos funcionários, utilizando Web Scraping e IA Generativa para gerar informações que demonstrem níveis de conhecimento nas necessidades da organização.

5º) Alocação Ótima: Uma técnica de otimização é aplicada utilizando os dados coletados nas etapas anteriores para determinar a melhor alocação possível dos funcionários. Múltiplas alocações ótimas podem ser identificadas se vários arranjos produzirem a mesma pontuação final.

Esta abordagem estruturada garante uma avaliação abrangente tanto das capacidades dos funcionários quanto dos requisitos organizacionais, facilitando um desdobramento da força de trabalho estrategicamente alinhado.

### **5.4.3 Compreensão dos Dados**

Essencialmente, há 5 bases necessárias (fig. 5.22):

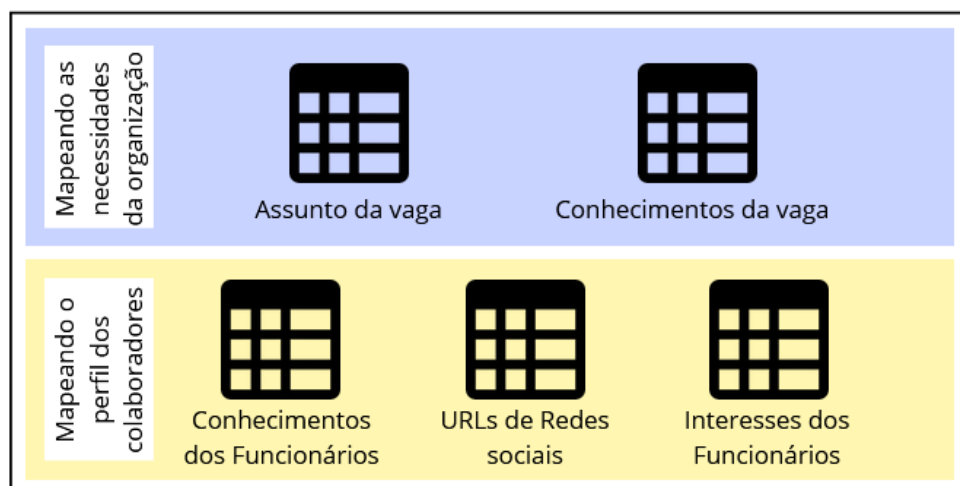


Figura 5.22: Bases Necessárias - Fonte: O autor

Essas tabelas podem ser agrupadas em 2 categorias:

1º) Tabelas para Mapeamento das Necessidades Organizacionais: Neste grupo, existem 2 tabelas, ambas preenchidas pela organização:

A) Macro Tópicos das Posições: Esta tabela reflete o mapeamento de macro tópicos para todas as posições disponíveis e pode ser aplicada a qualquer organização, de qualquer setor. Cada posição deve ter apenas um Macro Tópico. No caso apresentado, que demonstra a viabilidade da estrutura, os macro tópicos são os mostrados na Tabela 5.1.

| Macrotópicos                   |
|--------------------------------|
| Cientista de Dados             |
| Desenvolvedor Java/Cloud       |
| Desenvolvedor Mobile           |
| Desenvolvedor Web              |
| Especialista em Agilidade      |
| Especialista em Cibersegurança |
| Especialista em UX             |
| Governança de TI               |
| Infraestrutura de TI           |
| Mainframe                      |

Tabela 5.1: Tabela de Macrotópicos em Ordem Alfabética

B) Tópicos das Posições: Um nível abaixo dos macro tópicos, os tópicos das posições representam as ferramentas ou conhecimentos técnicos específicos necessários para as posições. Reitera-se que estes podem ser aplicados a qualquer organização de qualquer setor. Dependendo das necessidades e da complexidade das posições, há um número

mínimo e máximo de tópicos que cada posição pode ter. No exemplo, foi utilizado um mínimo de um assunto técnico e um máximo de três assuntos técnicos. Os tópicos utilizados foram os mostrados na Tabela 5.2.

|                         |                      |                    |                   |
|-------------------------|----------------------|--------------------|-------------------|
| Agile Coach             | Data Scientist       | JCL                | Public Clouds     |
| AI Ops                  | Design               | Kanban             | Puppet            |
| Analytics Tools         | Distributed Database | Kubernetes         | Python            |
| Android                 | Docker               | Mainframe CICS     | Quarkus           |
| Ansible                 | FinOps               | Mainframe DB2      | React Native      |
| Aplication Server       | Framework Angular    | Mainframe Hard     | Research          |
| Artificial Intelligence | Framework React      | Mainframe Software | Scala             |
| Assembly                | GO                   | Middleware         | Scrum             |
| AWS                     | Google               | Monitoring         | SER               |
| Azure                   | IA Tools             | Natural            | Signal Processing |
| C/C++                   | IOS                  | Net                | Spring Boot       |
| CI/CD                   | ISSO 20000           | NLP/LLMs           | Storage           |
| Cobit                   | IT Infrastructure    | Node JS            | Terraform         |
| Cobol                   | ITIL/ITSM            | Observability      | UX                |
| Computing Vision        | ITSM                 | Openstack          | Virtual Server    |
| CX                      | Java                 | Operating System   | Web Developer     |
| Data Engineer           | Java Script          | Others             | Writing           |

Tabela 5.2: Tabela de Experiências

## 2) Tabelas para Mapeamento de Perfis dos Funcionários:

Neste grupo, três tabelas foram utilizadas, todas preenchidas pelos funcionários:

A) Conhecimento dos Funcionários: Cada funcionário declara o número de anos de experiência profissional em cada um dos tópicos (estes devem ser os mesmos listados na tabela de Tópicos das Posições). Dependendo das necessidades do negócio, é definida uma escala para pontuar a experiência de cada respondente em cada tópico. No caso em consideração, foi utilizada uma escala de 0 a 6, onde “0” representa nenhuma experiência, “1” um ano de experiência, “2” dois anos de experiência, “3” três anos de experiência, “4” quatro anos de experiência, “5” cinco anos de experiência e “6” mais de cinco anos de experiência.

B) Interesses dos Funcionários: Cada funcionário deve declarar macro tópicos de interesse hierarquicamente. Cada organização deve limitar o número de acordo com as necessidades do negócio. Os itens desta tabela devem ser os mesmos que os da tabela de Macro Tópicos das Posições. No caso utilizado no artigo, cada funcionário declarou 3 tópicos de interesse, hierarquicamente (maior preferência, preferência média e menor preferência).

C) URLs das Redes Sociais: Os funcionários declaram URLs de seus perfis nas redes sociais que são capazes de demonstrar conhecimento nos tópicos de interesse da organização. No caso em consideração, foram solicitados 3 URLs: LinkedIn, GitHub e Currículo Lattes.

Esses dados serão processados para gerar uma pontuação para cada um dos itens na tabela de Tópicos das Posições.

#### **5.4.4 Preparação de Dados**

Do ponto de vista de ETL, o processo está focado em um único propósito: gerar uma tabela que descreva a pontuação de cada candidato para cada posição disponível, o que pode resultar em uma matriz quadrada ou não quadrada, dependendo do número de candidatos e vagas. Para chegar a esta tabela final (Tabela de Pontuação Final: Funcionários vs Vagas), a Figura 5.23 apresenta o framework para se chegar na alocação ótima em termos de aproveitamento de pontos dos novos funcionários.

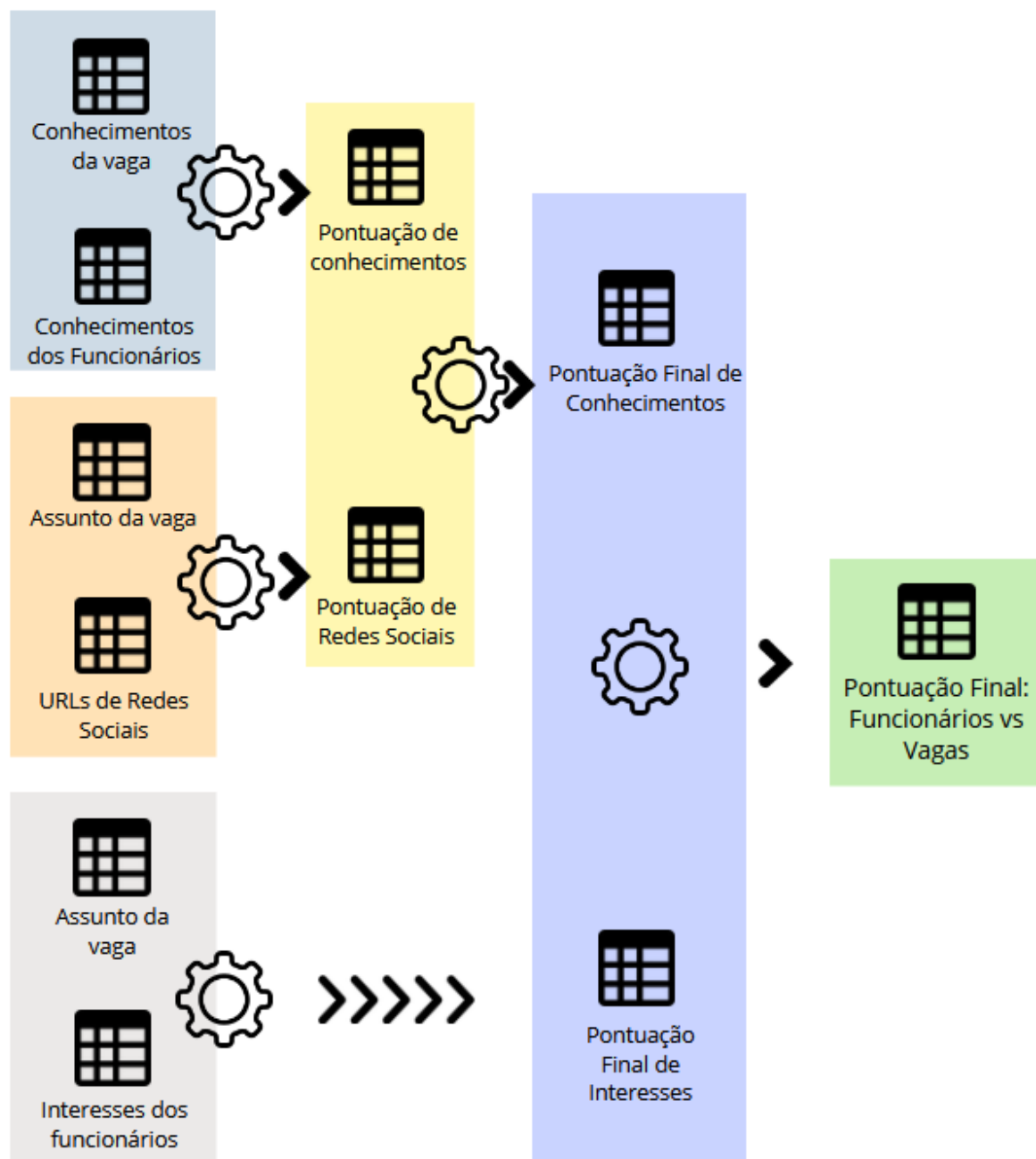


Figura 5.23: Modelo Geral - Fonte: O autor

De uma perspectiva global, para chegar à Tabela de Pontuação Final: Funcionários vs Vagas, são necessários os seguintes passos:

1º passo) Calcular uma pontuação parcial para cada candidato para cada vaga, cruzando as tabelas Tópicos das Posições e Conhecimento dos Funcionários. Essa pontuação representa o alinhamento do conhecimento declarado pelo funcionário com cada conjunto de tópicos definidos para cada vaga. No estudo em questão, a escala utilizada para cada correspondência de conhecimento varia de 0 a 6 pontos, seguindo a mesma escala do número

de anos declarados pelo candidato. Como cada vaga no caso pode registrar até 3 pontos de conhecimento, este item pode pontuar até 18 pontos por candidato/vaga.

2º passo) Calcular a pontuação parcial para cada candidato para cada vaga acessando os links da tabela URLs das Redes Sociais. O acesso é feito via web scraping capaz de acessar os links das redes sociais, e com esses dados, uma solução LLM é usada para analisar o percentual de alinhamento com a tabela Tópicos das Posições. No estudo mencionado, devido a potenciais alucinações, apenas aqueles que tinham mais de 30% de alinhamento em conhecimento específico foram considerados, usando o script empregado. A escala utilizada é de até 6 pontos se atingir 100% de alinhamento. Se não for de interesse da organização, é possível não utilizar esta base de dados e executar o processo apenas com as demais bases de dados.

A análise executada com o LLM requer a elaboração de um prompt adequado para este fim. O prompt deste estudo usou como base alguns padrões específicos para o ChatGPT ([55]). O primeiro orienta o LLM a assumir uma persona capaz de realizar a análise com um conhecimento específico. Para isso, foi elaborado o segmento *“O assistente é um especialista em contratação de profissionais de TI para um grande banco, avaliando o perfil dos profissionais, com o objetivo de indicar o candidato para um determinado cargo. Sua tarefa é emitir uma pontuação de 0 a 100 com base no nível de compatibilidade do profissional com um determinado perfil. Os perfis a serem considerados são ...”*. O texto foi complementado com os cargos necessários. Não foi necessário especificar as atribuições do cargo. Foi fornecido apenas o título, conforme mostrado na Tabela 5.1.

Para viabilizar a avaliação da pontuação atribuída pelo LLM, foi instruído que apresentasse a explicação para a nota, passando diretrizes sobre seu formato. Para facilitar a leitura das respostas do LLM e seu processamento via software, foi implementado um modelo de resposta a ser seguido pelo LLM *“Caso o profissional NÃO TENHA experiência na área, a pontuação DEVE SER 0. Apresente sua avaliação com a resposta exatamente nesta estrutura json: {role:nome do cargo score:NN%, reasoning:sua explicação sobre a pontuação}”*. As palavras em maiúsculas foram necessárias para reduzir a taxa de desvio de comando pelo LLM (ou seja, alucinação).

Os dados das redes sociais foram processados e adicionados ao final do prompt. Precedidos por este segmento *Esta é a experiência profissional do candidato: ...*, e este *Estes são os repositórios de código do candidato: ....* Para cada segmento, foram anexados, respectivamente, a experiência profissional do candidato conforme declarado por ele mesmo na rede social e os repositórios de projetos presentes no perfil.

3 padrões foram aplicados (*Persona, Template and Reflection*) para elaborar o prompt que orienta o LLM sobre como responder à tarefa atribuída a ele.

3º passo) Calcular a soma da Pontuação de Conhecimento e Pontuação das Redes Sociais,

que é uma soma simples dos tópicos anteriores, gerando a tabela Pontuação Final de Conhecimento, que descreve a pontuação final de cada candidato em cada vaga. No caso em questão, a pontuação máxima possível para cada candidato/vaga é de 24 pontos, somando os limites das tabelas anteriores.

4º passo) Calcular a pontuação da relação entre as tabelas Macro Tópicos das Posições e Interesses dos Funcionários. Essa pontuação ocorre à medida que há coincidência dos interesses declarados pelos funcionários e os macro tópicos da vaga. No caso em questão, foi decidido que cada funcionário poderia declarar até 3 interesses por ordem de preferência, onde a coincidência do macro tópico da vaga com a primeira preferência do candidato pontua 8 pontos, com a segunda preferência 5 pontos, e com a primeira coincidência 3 pontos. Essa pontuação é menor no teto e inferior à distribuição média da tabela de conhecimento porque, neste caso, optou-se pela priorização das necessidades da organização.

5º passo) Calcular a matriz final (Tabela Pontuação Final: Funcionários vs Vagas) que gerará a base a ser usada pelos otimizadores para localizar a distribuição ideal. Esse cálculo é feito somando as pontuações das tabelas Pontuação Final de Conhecimento e Pontuação Final de Interesses. No estudo de caso em questão, chega-se a uma matriz não quadrada 409:707 cruzando os 409 novos funcionários a serem alocados com as 707 vagas mapeadas, onde cada candidato tem uma pontuação em cada vaga, totalizando 289.163 registros.

O foco desta seção é buscar a melhor distribuição de funcionários nas vagas, com base na distribuição ideal da soma do conhecimento utilizado. Este estudo de caso tem por objeto um cenário de alocação ótima (maior aproveitamento possível dos conhecimentos e interesses, frente aos parâmetros das vagas disponíveis). Um mesmo grupo de candidatos frente a uma mesma demanda de vagas pode apresentar diversos cenários ótimos, dependendo da técnica ou otimizador utilizado, alcançando cenários distintos, mas com valores finais idênticos (máximo possível).

Em todas as abordagens, buscamos a seguinte função objetivo:

$$\text{Maximizar} \quad \sum_{(u,v) \in E} f(u, v)$$

Onde:

- $u$ : Representa um candidato do conjunto de candidatos  $U$ .
- $v$ : Representa uma vaga do conjunto de vagas  $V$ .
- $E$ : Conjunto de arestas (ou conexões) que ligam candidatos  $u$  às vagas  $v$ , representando a possibilidade de um candidato ser alocado a uma vaga.

- $f(u, v)$ : Função que indica o fluxo ou grau de conhecimento aproveitado quando o candidato  $u$  é alocado à vaga  $v$ . Este é o valor que queremos maximizar.

### Restrições:

#### 1. Restrição de Capacidade:

$$0 \leq f(u, v) \leq c(u, v) \quad \forall (u, v) \in E$$

Onde:

- $c(u, v)$ : Capacidade máxima de conhecimento que pode ser alocado de  $u$  para  $v$ . Essa variável pode ser definida com base no grau de conhecimento do candidato e nas exigências da vaga.

#### 2. Conservação do Fluxo:

$$\sum_{(u,i) \in E} f(u, i) = \sum_{(i,v) \in E} f(i, v) \quad \forall i \neq s, i \neq t$$

Onde:

- $i$ : Nó intermediário no fluxo. Este índice assegura que a soma do fluxo que entra em um nó intermediário seja igual à soma do fluxo que sai dele, garantindo equilíbrio.
- $s$ : Nó de origem no fluxo (nó inicial).
- $t$ : Nó de destino no fluxo (nó final).

#### 3. Não Negatividade:

$$f(u, v) \geq 0 \quad \forall (u, v) \in E$$

Onde:

- Esta restrição assegura que o fluxo seja sempre um valor não negativo, representando a quantidade real de conhecimento alocado.

Este modelo busca, assim, otimizar a alocação de talentos, respeitando as restrições do sistema e maximizando o uso eficiente dos conhecimentos disponíveis frente às necessidades organizacionais.

Neste contexto, discutiremos três técnicas de otimização distintas e faremos comparações entre elas: Maximum Cardinality Matching, Otimização Híbrida (Heurística + Branch and cut + simplex primal + Crossover) e o Método Simplex.



## Maximum Cardinality Matching

No contexto da alocação de candidatos a vagas de emprego com base no conhecimento declarado e nas necessidades específicas das vagas, a técnica de *Maximum Cardinality Matching* oferece um método estruturado para otimizar esse processo, utilizando a teoria dos grafos.

O *Maximum Cardinality Matching* em um grafo bipartido é utilizado para identificar a melhor correspondência possível entre dois conjuntos distintos — candidatos e vagas, neste caso —, de forma que o número de correspondências seja maximizado. Cada candidato e cada vaga são representados como vértices em dois conjuntos separados. As arestas entre esses vértices representam correspondências potenciais e podem ser ponderadas com base nas pontuações de compatibilidade, derivadas das qualificações dos candidatos e dos requisitos das vagas.

De maneira geral, as arestas são desenhadas entre todos os vértices de origem e destino (candidatos e vagas), resultando na melhor formação possível, onde cada vértice de origem pode ser correspondido a apenas um vértice de destino, e vice-versa. A Figura 5.24 ilustra o modelo de forma simplificada:

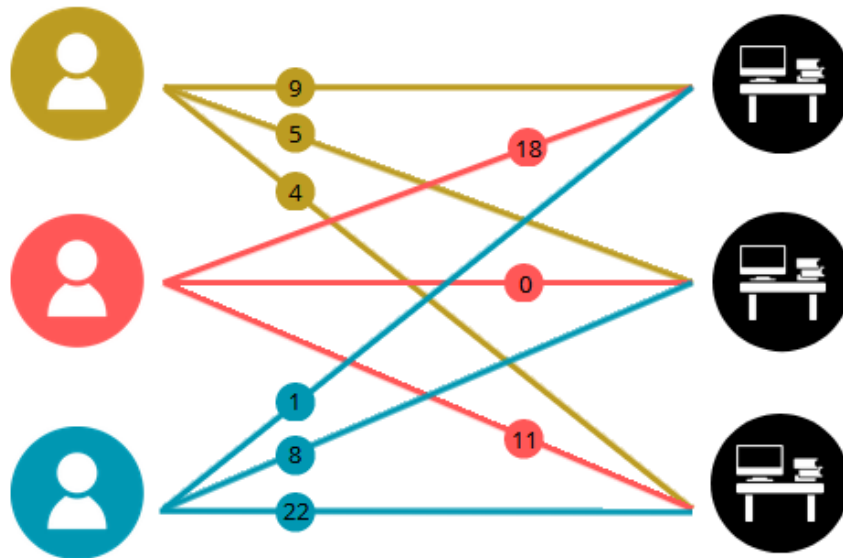


Figura 5.24: Exemplo de grafos - Fonte: O autor

Vantagens:

- Correspondência Ótima: Garante que o número máximo de candidatos seja correspondido a vagas, utilizando eficientemente os recursos disponíveis. Além disso, é capaz de lidar com matrizes retangulares, onde o número de candidatos e vagas pode ser diferente.

- Escalabilidade: Pode processar grandes conjuntos de dados de maneira eficaz, sendo adequado para organizações de grande porte, com um elevado número de candidatos e vagas.
- Flexibilidade: Permite a incorporação de diversos fatores no processo de correspondência, ajustando os pesos das arestas de acordo com critérios específicos, como nível de experiência, relevância das habilidades e outros requisitos do cargo.

Desvantagens:

- Implementação Complexa: Implementar esta técnica requer um conhecimento aprofundado de teoria dos grafos e algoritmos, o que pode ser intensivo tanto em termos de desenvolvimento quanto em computação, especialmente para conjuntos de dados muito grandes.

A Figura 5.25 ilustra um grafo que representa as alocações ideais de vértices e arestas para o estudo de caso. Embora a visualização seja útil, ela se torna confusa devido ao número elevado de vagas e funcionários envolvidos no modelo.

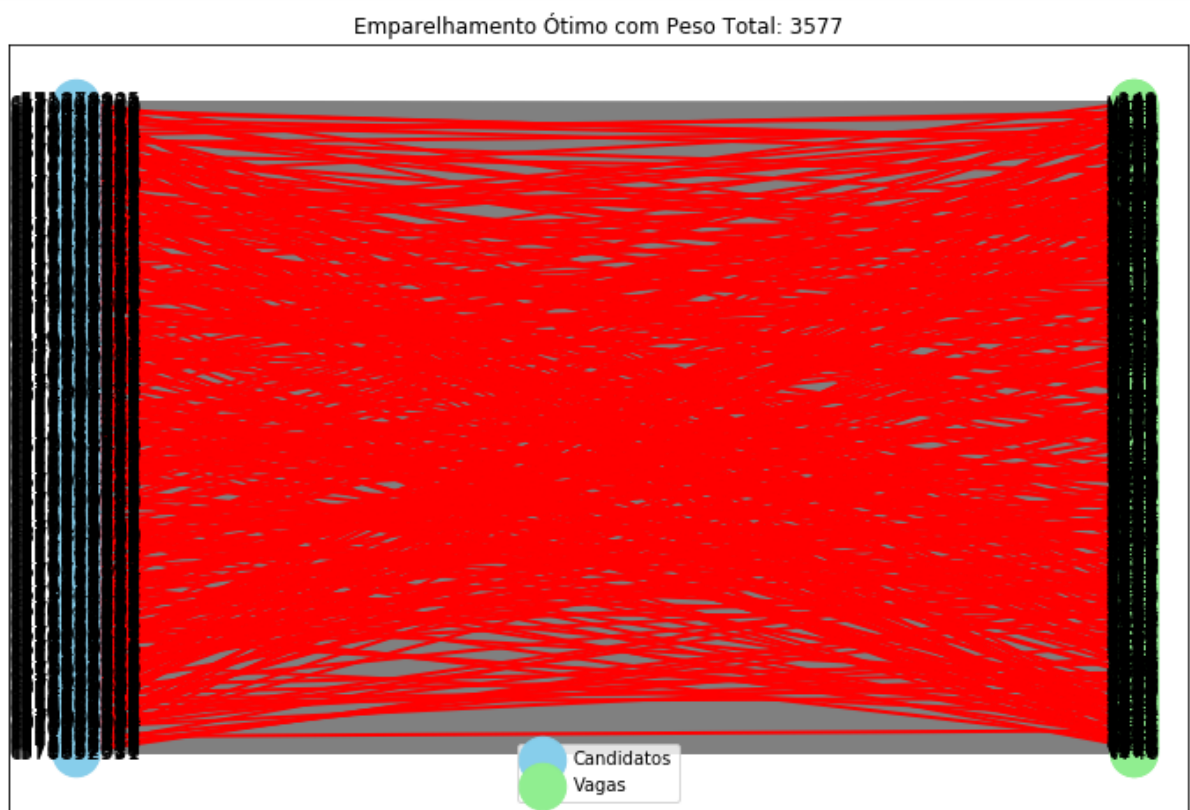


Figura 5.25: Vértices e Arestas - Fonte: O autor

A pontuação alcançada com a técnica de *Maximum Cardinality Matching* foi de 3577 pontos.

## Otimização Híbrida

A abordagem de otimização híbrida combina programação linear e heurísticas, adaptando-se para alcançar a alocação ideal de recursos. O problema é modelado como um problema de programação inteira binária, no qual as variáveis de decisão  $x[i, j]$  indicam se o candidato  $i$  é atribuído à vaga  $j$  (1 se atribuído, 0 caso contrário). As etapas desta técnica, com o uso do *Gurobi* como otimizador, são descritas a seguir:

- Heurísticas: Embora não explicitamente detalhadas no *script*, o *Gurobi* utiliza métodos heurísticos automaticamente durante o processo de otimização para encontrar soluções iniciais e reduzir o espaço de busca.
- Branch and Cut: Técnica baseada no algoritmo *branch-and-bound*, aprimorada com a adição de planos de corte. Essa abordagem reduz a região viável e, conseqüentemente, o espaço de busca, aumentando a eficiência na resolução de problemas de programação inteira.
- Simplex Primal: O algoritmo *simplex primal* é aplicado para resolver problemas de programação linear. Ele busca melhorar iterativamente a solução movendo-se ao longo das arestas da região viável definida pelas restrições, até encontrar o vértice ótimo.
- Crossover: Este método converte uma solução viável não básica (frequentemente obtida por métodos de barreira ou pontos interiores) em uma solução viável básica, essencial para certas análises pós-otimização.

### Vantagens:

- Escalabilidade: Este método é capaz de lidar com grandes conjuntos de dados de maneira eficiente, devido às robustas capacidades de otimização do *Gurobi*.
- Flexibilidade: A técnica permite a incorporação de restrições adicionais ou modificações no modelo com facilidade, como a alteração dos pesos de diferentes critérios ou a inclusão de novas restrições específicas.

### Desvantagens:

- Complexidade: A implementação requer um conhecimento sólido de técnicas de otimização e o uso de softwares especializados, como o *Gurobi*, que podem não estar prontamente disponíveis para todos os usuários.
- Intensivo em Recursos: A otimização, especialmente com grandes conjuntos de dados, pode demandar recursos computacionais significativos, resultando em tempos de execução mais longos.

A pontuação alcançada na abordagem de Otimização Híbrida foi de 3577 pontos.

---

**C) Método Simplex:** O Método Simplex é amplamente utilizado para maximizar ou minimizar uma função objetivo linear sujeita a restrições lineares. Neste estudo, foi empregada a função `linprog` do módulo `scipy.optimize`. A função objetivo, que busca

maximizar a soma das pontuações das alocações, foi minimizada no contexto do Simplex devido à convenção da função `linprog`. Assim, os valores foram passados com sinal negativo para adequar-se à metodologia.

Os passos básicos do funcionamento do Método Simplex são descritos a seguir:

- **Forma Padrão:** O problema é convertido para a forma padrão, o que geralmente significa que todas as restrições são transformadas em equações (utilizando variáveis de folga para substituir desigualdades por igualdades) e a função objetivo é expressa como uma função de maximização.
- **Solução Básica Inicial:** O processo inicia com uma solução básica inicial, frequentemente a solução trivial onde todas as variáveis de decisão são iguais a zero, enquanto as variáveis de folga (se presentes) assumem os valores constantes das restrições. Essa solução corresponde a um vértice do poliedro definido pelas restrições.
- **Movimento Entre Vértices:** O algoritmo progride de um vértice (ou ponto de canto) do espaço de solução para outro, seguindo as arestas do poliedro formado pelas restrições, com o objetivo de aumentar o valor da função objetivo (no caso de maximização).
- **Critério de Entrada e Saída:** Em cada iteração, uma variável de decisão não básica que pode aumentar a função objetivo é incluída na solução (critério de entrada), enquanto uma variável básica é removida (critério de saída). O critério para essa escolha baseia-se no teste da razão mínima (minimum ratio test), que garante a satisfação contínua das restrições.
- **Iteração:** O processo é repetido, substituindo uma variável básica por uma não básica a cada passo, e os valores das variáveis são recalculados para otimizar ainda mais a função objetivo. Cada etapa desse processo é chamada de "pivotagem".
- **Terminação:** O algoritmo é concluído quando não é possível aumentar a função objetivo (todos os coeficientes da função objetivo para as variáveis não básicas são não positivos no caso de maximização), indicando que a solução ótima foi encontrada, ou quando uma solução ilimitada é detectada (nenhuma variável pode ser aumentada sem violar as restrições).

#### Vantagens:

- **Eficiência e Velocidade:** Para problemas de programação linear de tamanho moderado, o Método Simplex é reconhecido por sua rapidez e eficiência.

- **Flexibilidade:** O método pode ser ajustado para lidar com diferentes tipos de restrições lineares, sendo aplicável a uma ampla variedade de problemas de otimização linear.

#### Desvantagens:

- **Limitação a Problemas Lineares:** O Método Simplex é aplicável apenas a problemas que podem ser formulados como funções objetivo lineares com restrições lineares. Não é adequado para problemas de otimização não linear sem transformações significativas.
- **Problemas de Grande Escala:** Embora eficiente para muitos casos, o Método Simplex pode se tornar computacionalmente caro e menos eficaz à medida que o tamanho do problema aumenta, especialmente em cenários de grande escala com um número elevado de variáveis e restrições.
- **Degeneração e Ciclagem:** Em situações específicas, o Método Simplex pode enfrentar desafios como degeneração (progresso muito lento) ou ciclagem (revisitação repetida de soluções), embora variantes modernas e técnicas, como o método de ponto interior, tenham contribuído para mitigar esses problemas.

Pontuação alcançada na abordagem do Método Simplex: 3577 pontos.

#### 5.4.5 Avaliação de performance

As técnicas aplicadas, incluindo Maximum Cardinality Matching, Otimização Híbrida e o Método Simplex, apresentaram resultados consistentes, alcançando o mesmo cenário ótimo, com uma pontuação total de 3577 pontos. Como esperado, essas abordagens, orientadas para a otimização máxima, demonstraram eficácia na maximização da utilização do conhecimento e interesses dos candidatos em alinhamento com as necessidades organizacionais.

Embora todas as técnicas tenham gerado pontuações idênticas, observaram-se variações significativas nas alocações específicas de candidatos às posições. Por exemplo, o Método Simplex e o Maximum Cardinality Matching não compartilharam nenhuma alocação idêntica. Entre o Método Simplex e a Otimização Híbrida, apenas um dos 409 candidatos foi alocado para a mesma posição. Em contrapartida, a Otimização Híbrida e o Maximum Cardinality Matching exibiram maior sobreposição, com 77 dos 409 candidatos alocados para as mesmas posições.

Ao analisar as coincidências de pontuação entre os candidatos, independentemente das posições específicas, verificou-se uma maior uniformidade:

- Método Simplex vs. Maximum Cardinality Matching: 351 coincidências de pontuação.
- Método Simplex vs. Otimização Híbrida: 340 coincidências de pontuação.
- Otimização Híbrida vs. Maximum Cardinality Matching: 340 coincidências de pontuação.

Em relação aos tempos de execução, avaliados em uma máquina padrão (processador i5 de 8ª geração, 8GB de RAM, 2GB GPU), os resultados obtidos foram os seguintes:

- Método Simplex: 20,1 segundos.
- Otimização Híbrida: 79,3 segundos.
- Maximum Cardinality Matching: 244,3 segundos.

Esses tempos são considerados aceitáveis, com o maior tempo de execução sendo de aproximadamente 4 minutos, o que é insignificante, dado que a execução ocorre uma única vez para promover a alocação ótima.

### **Avaliação dos Resultados - Caso de uso prescritivo**

Para metrificar o desempenho dos 409 candidatos alocados nas vagas indicadas, foi utilizada a quantidade de *commits* realizados nos primeiros 100 dias de trabalho como indicador de produtividade, considerando tratar-se de uma área de TI. Contudo, devido à diversidade de funções nas áreas de alocação, que abrangem desde desenvolvimento de softwares até infraestrutura, apenas 192 candidatos realizaram *commits* nesse período, constituindo o universo amostral observado para a análise. Não foi realizado tratamento de outliers na análise, pois entende-se que a variação na quantidade de *commits* reflete características individuais e específicas do comportamento humano no ambiente de trabalho.

A primeira abordagem utilizada para validação da hipótese alternativa foi o coeficiente de Pearson. O resultado foi um coeficiente de correlação de 0,201, indicando uma correlação positiva entre a pontuação de afinidade e o desempenho medido em *commits*.

A ANOVA foi aplicada para investigar diferenças significativas no número de *commits* entre diferentes níveis de afinidade. Para esta análise, as pontuações foram agrupadas em cinco categorias:

- Nenhuma Afinidade (0 pontos): 14 pessoas
- Baixa Afinidade (1 a 7 pontos): 33 pessoas

- Média Afinidade (8 a 12 pontos): 77 pessoas
- Alta Afinidade (13 a 20 pontos): 51 pessoas
- Altíssima Afinidade (acima de 20 pontos): 17 pessoas

Os resultados da ANOVA foram:

- F-Statistic: 3,511
- p-value: 0,0086

A estatística F de 3,511 representa a razão entre a variabilidade explicada pelo modelo (variabilidade entre os grupos) e a variabilidade não explicada (variabilidade dentro dos grupos). Um valor mais elevado da estatística F indica que a variabilidade entre as médias dos grupos é significativamente maior do que a variabilidade dentro dos grupos, sugerindo diferenças reais entre os grupos analisados.

Com um valor-p de 0,0086, que é significativamente menor do que o nível de significância comum de 0,05, a ANOVA indica que há uma diferença estatisticamente significativa entre os grupos de afinidade em relação à média de *commits*.

Conforme observado na Figura 5.26, os grupos de Alta Afinidade e Altíssima Afinidade apresentam médias superiores aos grupos de afinidade inferior, sendo que a média do grupo de Alta Afinidade (280,80) destaca-se de forma expressiva em relação às demais. Esse comportamento sugere que o framework proposto de alocação de candidatos, baseado em afinidade, foi eficaz ao identificar uma relação positiva entre o alinhamento das habilidades/interesses dos candidatos e o desempenho nas atividades.

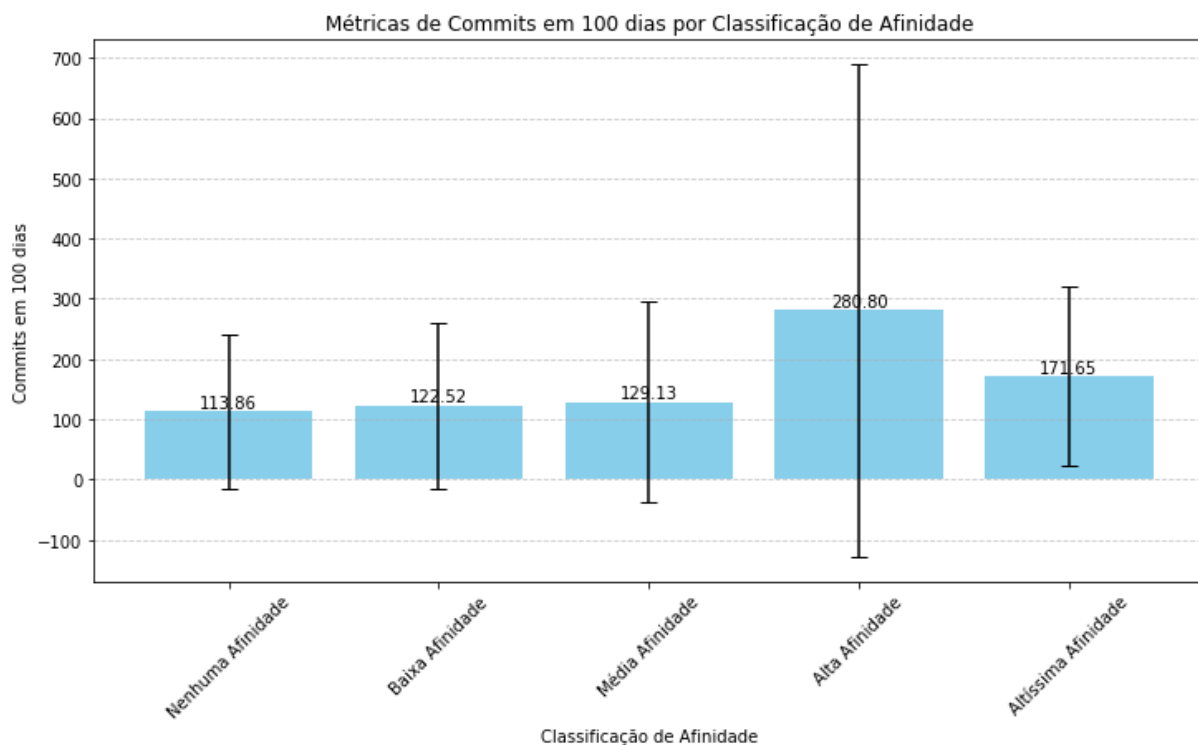


Figura 5.26: Commits entre grupos - Fonte: O autor

No entanto, a média da Altíssima Afinidade é inferior à da Alta Afinidade, o que pode indicar uma variabilidade no desempenho ou outros fatores contextuais. Ainda assim, a tendência geral respalda a lógica do framework, pois candidatos alocados em vagas com maior correspondência de afinidade tendem a apresentar um número superior de commits, evidenciando que a consideração de habilidades e interesses específicos para a alocação tem impacto positivo no desempenho inicial dos funcionários.

### Conclusão do Caso de Uso Prescritivo

Este estudo demonstrou a eficácia de uma abordagem orientada por dados na gestão de talentos, criando um *framework* capaz de ser implementado em qualquer organização. Usando um estudo de caso de alocação de novos funcionários no setor de tecnologia de um grande banco brasileiro como pano de fundo, a viabilidade do *framework* proposto na alocação de novos funcionários dentro de uma organização ficou evidente. A aplicação de métodos de otimização, como Correspondência Máxima de Cardinalidade, Otimização Híbrida e o Método Simplex, revelou uma capacidade notável de maximizar a eficiência da alocação de talentos, superando significativamente as alocações aleatórias em termos de aproveitamento do conhecimento e interesses dos candidatos, com um alcance de 84,2% mais pontos do que a melhor pontuação entre um milhão de simulações aleatórias.



Os resultados destacam que a incorporação de técnicas avançadas de análise e otimização pode transformar a gestão de talentos, permitindo um melhor alinhamento entre as habilidades dos funcionários e as necessidades das organizações. Este alinhamento não só melhora o desempenho e a satisfação dos funcionários, como também promove um ambiente organizacional mais eficiente e produtivo.

Além disso, a diversidade nos cenários de alocação observados entre as diferentes técnicas de otimização utilizadas sublinha a importância de escolher a ferramenta certa de acordo com as características específicas da organização e das posições disponíveis. Isso ressalta a relevância de uma abordagem estratégica e personalizada na alocação de talentos.

Para pesquisas futuras, seria interessante explorar a aplicabilidade deste modelo em diferentes contextos organizacionais e setores, bem como integrar outras variáveis relevantes, como habilidades interpessoais e adaptabilidade cultural, para uma compreensão ainda mais holística da gestão de talentos. A técnica proposta também tem potencial para a alocação de funcionários atualmente empregados em corporações, promovendo uma rotação eficiente entre as equipes.

A continuação deste trabalho pode abrir novos caminhos para uma gestão de recursos humanos mais estratégica e baseada em evidências, redefinindo como as organizações pensam e implementam suas políticas de alocação de talentos.

# Capítulo 6

## Conclusão

Este estudo propôs e validou um *framework* abrangente para aplicação dos níveis de maturidade analítica — descritivo, diagnóstico, preditivo e prescritivo — na gestão de pessoas, com foco em exemplificar sua viabilidade em uma organização do setor financeiro. Ao longo desta dissertação, foram exploradas técnicas estatísticas, de ciência de dados e inteligência artificial, demonstrando como essas abordagens podem transformar a tomada de decisões em recursos humanos, alinhando-se aos objetivos estratégicos organizacionais. Os resultados obtidos evidenciaram que o *framework* proposto não apenas é viável, mas também eficaz. No nível descritivo, a criação de *dashboards* permitiu uma visualização clara e objetiva de métricas de desempenho organizacional, promovendo maior transparência e agilidade na interpretação dos dados. A análise diagnóstica revelou relações causais relevantes entre atributos organizacionais, destacando fatores que impactam diretamente o desempenho e a satisfação dos colaboradores. No nível preditivo, os modelos de *machine learning* se mostraram eficazes na identificação de talentos emergentes. Por fim, a análise prescritiva, utilizando métodos de otimização, destacou-se na alocação eficiente de talentos, com um desempenho 84,2% superior em relação a alocações aleatórias.

A validação empírica do *framework* foi fundamental para evidenciar sua aplicabilidade prática. Por meio de estudos de caso, comprovou-se que as ferramentas analíticas não apenas oferecem suporte à decisão, mas também agregam valor à organização ao otimizar processos e alinhar as competências dos funcionários às necessidades organizacionais. A correlação positiva entre a afinidade das alocações e a produtividade, medida em *commits*, é um exemplo claro de como a abordagem proposta pode gerar *insights* valiosos e impactar diretamente os resultados empresariais.

Apesar dos avanços alcançados, este trabalho também identificou limitações que merecem atenção em estudos futuros. A diversidade das funções analisadas e a limitação na disponibilidade de dados completos sobre produtividade foram desafios enfrentados durante a análise. Além disso, o impacto de variáveis qualitativas, como habilidades interpessoais

e fatores culturais, ainda precisa ser explorado para uma compreensão mais holística da gestão de talentos.

Como sugestões para pesquisas futuras, recomenda-se expandir o escopo do *framework* para outros setores e tipos de organizações, explorando sua aplicação em contextos culturais e empresariais distintos. A integração de variáveis qualitativas e métricas de bem-estar também é essencial para refinar as técnicas analíticas e garantir que as decisões estejam alinhadas aos valores organizacionais e à satisfação dos colaboradores.

Em síntese, esta dissertação reforça o potencial transformador do *People Analytics* na gestão de pessoas, oferecendo um caminho estruturado e científico para o aprimoramento das práticas de recursos humanos. O *framework* proposto representa um passo significativo rumo à gestão baseada em dados, contribuindo para um ambiente organizacional mais eficiente, transparente e alinhado aos desafios do futuro.

# Capítulo 7

## Anexos

### A Correlações Lineares

Tabela A.1: Correlação entre Variáveis

| Par de Variáveis                                                                | Coeficiente de Correlação |      |
|---------------------------------------------------------------------------------|---------------------------|------|
| Total_abonos x ab_acordos_ant                                                   |                           | 0.98 |
| Media_Clima_Auto_Inov_2_ciclos x Media_Clima_Geral_2_ciclos                     |                           | 0.96 |
| Otimiza_Processo_Trabalho x Resultado_Competencia                               |                           | 0.95 |
| Media_Clima_Geral_2_ciclos x Media_Pesq_Clima_Estilo_Gestão_2_ciclos            |                           | 0.95 |
| Media_Pesq_Clima_Valorizacao_2_ciclos x Media_Clima_Geral_2_ciclos              |                           | 0.94 |
| Media_Clima_Geral_2_ciclos x Media_Clima_Relacionamento_2_ciclos                |                           | 0.93 |
| Autonomia x Resultado_Geral_9box                                                |                           | 0.92 |
| Media_Pesq_Clima_Valorizacao_2_ciclos x Media_Clima_Auto_Inov_2_ciclos          |                           | 0.91 |
| Qtde.Cert..Intermediárias x Total_Cert_internas                                 |                           | 0.90 |
| Media_Clima_Relacionamento_2_ciclos x Media_Clima_Auto_Inov_2_ciclos            |                           | 0.89 |
| Media_Pesq_Clima_Valorizacao_2_ciclos x Media_Pesq_Clima_Estilo_Gestão_2_ciclos |                           | 0.88 |

| <b>Par de Variáveis</b>                                                            | <b>Coeficiente de Correlação</b> |
|------------------------------------------------------------------------------------|----------------------------------|
| Media_Clima_Relacionamento_2_ciclos x Me-<br>dia_Pesq_Clima_Estilo_Gestão_2_ciclos | 0.87                             |
| saldo_folgas x folgas                                                              | 0.87                             |
| Resultado_Geral_9box x Adaptabilidade                                              | 0.86                             |
| Resultado_Competencia x Resultado_Geral_9box                                       | 0.86                             |
| Media_Pesq_Clima_Estilo_Gestão_2_ciclos x Me-<br>dia_Clima_Auto_Inov_2_ciclos      | 0.85                             |
| Resultado_Competencia x Eficiência                                                 | 0.84                             |
| Inovação x Resultado_Geral_9box                                                    | 0.83                             |
| Resultado_Geral_9box x Eficiência                                                  | 0.82                             |
| Resultado_Geral_9box x Colaboração                                                 | 0.81                             |
| Autonomia x Eficiência                                                             | 0.81                             |
| Otimiza_Processo_Trabalho x Resultado_Geral_9box                                   | 0.80                             |
| Media_Pesq_Clima_Valorizacao_2_ciclos x Me-<br>dia_Clima_Relacionamento_2_ciclos   | 0.79                             |
| Otimiza_Processo_Trabalho x Eficiência                                             | 0.79                             |
| Autonomia x Resultado_Competencia                                                  | 0.79                             |
| Inovação x Otimiza_Processo_Trabalho                                               | 0.79                             |
| Resultado_Competencia x Inovação                                                   | 0.78                             |
| Autonomia x Adaptabilidade                                                         | 0.77                             |
| Autonomia x Otimiza_Processo_Trabalho                                              | 0.77                             |
| Adaptabilidade x Eficiência                                                        | 0.75                             |
| Media_GPF_Processos_Int._3_ciclos x Me-<br>dia_GPF_Clientes_3_ciclos               | 0.75                             |
| Autonomia x Inovação                                                               | 0.74                             |
| Adaptabilidade x Inovação                                                          | 0.73                             |
| Media_GPF_Aprendizado_3_ciclos x Me-<br>dia_GPF_Processos_Int._3_ciclos            | 0.73                             |
| Adaptabilidade x Resultado_Competencia                                             | 0.72                             |
| Eficiência x Inovação                                                              | 0.72                             |
| Otimiza_Processo_Trabalho x Resul-<br>tado_Eixo_Desempenho                         | 0.70                             |
| Adaptabilidade x Colaboração                                                       | 0.70                             |
| Otimiza_Processo_Trabalho x Adaptabilidade                                         | 0.70                             |

| Par de Variáveis                               |   |        | Coeficiente de Correlação |
|------------------------------------------------|---|--------|---------------------------|
| Media_GPF_Financeira_3_ciclos                  | x | Me-    | 0.69                      |
| dia_GPF_Processos_Int._3_ciclos                |   |        |                           |
| Colaboração x Eficiência                       |   |        | 0.68                      |
| Colaboração x Resultado_Competencia            |   |        | 0.66                      |
| Autonomia x Colaboração                        |   |        | 0.66                      |
| Resultado_Eixo_Desempenho                      | x | Resul- | 0.65                      |
| tado_Competencia                               |   |        |                           |
| Media_GPF_Financeira_3_ciclos                  | x | Me-    | 0.65                      |
| dia_GPF_Clientes_3_ciclos                      |   |        |                           |
| tempo_banco x anuenio                          |   |        | 0.65                      |
| Dias.Absenteísmo x Dias.Ausências              |   |        | 0.65                      |
| Eficiência x Resultado_Eixo_Desempenho         |   |        | 0.64                      |
| Colaboração x Otimiza_Processo_Trabalho        |   |        | 0.64                      |
| Qtde.Cert..Básicas x Total_Cert_internas       |   |        | 0.63                      |
| Inovação x Colaboração                         |   |        | 0.59                      |
| Resultado_Eixo_Desempenho x Autonomia          |   |        | 0.58                      |
| Banco_de_horas x saldo_folgas                  |   |        | 0.58                      |
| Media_GPF_Financeira_3_ciclos                  | x | Me-    | 0.58                      |
| dia_GPF_Aprendizado_3_ciclos                   |   |        |                           |
| Resultado_Geral_9box                           | x | Resul- | 0.56                      |
| tado_Eixo_Desempenho                           |   |        |                           |
| Inovação x Resultado_Eixo_Desempenho           |   |        | 0.56                      |
| Media_GPF_Socioambiental                       | x | Me-    | 0.53                      |
| dia_GPF_Processos_Int._3_ciclos                |   |        |                           |
| Adaptabilidade x Resultado_Eixo_Desempenho     |   |        | 0.53                      |
| Dias.Absenteísmo x faltas_categoria            |   |        | 0.51                      |
| folgas x Banco_de_horas                        |   |        | 0.49                      |
| Resultado_Eixo_Desempenho x Colaboração        |   |        | 0.47                      |
| Media_GPF_Processos_Int._3_ciclos              | x | Resul- | 0.47                      |
| tado_Geral_9box                                |   |        |                           |
| Media_GPF_Socioambiental                       | x | Me-    | 0.47                      |
| dia_GPF_Clientes_3_ciclos                      |   |        |                           |
| Eficiência x Media_GPF_Processos_Int._3_ciclos |   |        | 0.46                      |
| Media_GPF_Processos_Int._3_ciclos              | x | Resul- | 0.46                      |
| tado_Competencia                               |   |        |                           |

| <b>Par de Variáveis</b>                                       | <b>Coefficiente de Correlação</b> |
|---------------------------------------------------------------|-----------------------------------|
| Media_GPF_Processos_Int._3_ciclos x Adaptabilidade            | 0.45                              |
| Media_GPF_Clientes_3_ciclos x Inovação                        | 0.45                              |
| Media_GPF_Financeira_3_ciclos x Media_GPF_Socioambiental      | 0.45                              |
| Autonomia x Media_GPF_Processos_Int._3_ciclos                 | 0.44                              |
| Qtde.Cert..Avançadas x Total_Cert_internas                    | 0.44                              |
| Media_GPF_Clientes_3_ciclos x Resultado_Geral_9box            | 0.44                              |
| Media_GPF_Clientes_3_ciclos x Resultado_Competencia           | 0.43                              |
| tempo_banco x Salario                                         | 0.43                              |
| Media_GPF_Processos_Int._3_ciclos x Otimiza_Processo_Trabalho | 0.43                              |
| Media_GPF_Socioambiental x Media_GPF_Aprendizado_3_ciclos     | 0.43                              |
| faltas_banco x faltas_categoria                               | 0.43                              |
| Media_GPF_Aprendizado_3_ciclos x Resultado_Geral_9box         | 0.41                              |
| Media_GPF_Clientes_3_ciclos x Otimiza_Processo_Trabalho       | 0.41                              |
| Media_GPF_Clientes_3_ciclos x Autonomia                       | 0.40                              |
| Resultado_Competencia x Media_GPF_Aprendizado_3_ciclos        | 0.39                              |
| Media_GPF_Aprendizado_3_ciclos x Adaptabilidade               | 0.39                              |
| Eficiência x Media_GPF_Clientes_3_ciclos                      | 0.39                              |
| Inovação x Media_GPF_Processos_Int._3_ciclos                  | 0.39                              |
| Media_GPF_Aprendizado_3_ciclos x Inovação                     | 0.38                              |
| faltas_categoria x Dias.Ausências                             | 0.38                              |
| Media_GPF_Clientes_3_ciclos x Adaptabilidade                  | 0.38                              |
| Resultado_Geral_9box x Media_GPF_Financeira_3_ciclos          | 0.37                              |
| Media_GPF_Financeira_3_ciclos x Resultado_Competencia         | 0.37                              |

| Par de Variáveis                                                  | Coefficiente de Correlação |
|-------------------------------------------------------------------|----------------------------|
| Media_GPF_Aprendizado_3_ciclos x Otimiza_Processo_Trabalho        | 0.37                       |
| Autonomia x Media_GPF_Aprendizado_3_ciclos                        | 0.36                       |
| Media_GPF_Processos_Int._3_ciclos x Colaboração                   | 0.36                       |
| Media_GPF_Financeira_3_ciclos x Otimiza_Processo_Trabalho         | 0.35                       |
| Autonomia x Media_GPF_Financeira_3_ciclos                         | 0.35                       |
| Adaptabilidade x Media_GPF_Financeira_3_ciclos                    | 0.34                       |
| Media_GPF_Financeira_3_ciclos x Eficiência                        | 0.34                       |
| tempo_empresa x Idade                                             | 0.33                       |
| Colaboração x Media_GPF_Aprendizado_3_ciclos                      | 0.33                       |
| Media_GPF_Financeira_3_ciclos x Inovação                          | 0.32                       |
| ab_acordo_atual x Total_abonos                                    | 0.32                       |
| Qtde.Cert..Intermediárias x Qtde.Cert..Básicas                    | 0.31                       |
| Media_GPF_Processos_Int._3_ciclos x Resultado_Eixo_Desempenho     | 0.31                       |
| nr_depend_ir x nr_depend_creche_novo                              | 0.30                       |
| dias_nao_util_trab_ano x folgas                                   | 0.30                       |
| Dias.Absenteísmo x faltas_banco                                   | 0.29                       |
| Qtde.Cert..Avançadas x Qtde.Cert..Intermediárias                  | 0.28                       |
| Media_GPF_Clientes_3_ciclos x Colaboração                         | 0.28                       |
| Horas.FIP x qtde_horas_treina_sem                                 | 0.28                       |
| Mais_2HE_dia_ano x Prorroga-cao_nao_negociada_ano                 | 0.27                       |
| Media_GPF_Financeira_3_ciclos x Colaboração                       | 0.27                       |
| faltas_banco x tempo_banco                                        | 0.25                       |
| anuenio x Idade                                                   | 0.24                       |
| Dias.Ausências x faltas_banco                                     | 0.24                       |
| Resultado_Competencia x Salario                                   | 0.23                       |
| Media_GPF_Clientes_3_ciclos x Resultado_Eixo_Desempenho           | 0.23                       |
| Media_Clima_Auto_Inov_2_ciclos x Entradas_nao_negociadas_ano      | 0.23                       |
| Media_Clima_Relacionamento_2_ciclos x Entradas_nao_negociadas_ano | 0.23                       |



| Par de Variáveis                      |   |                               | Coeficiente de Correlação |
|---------------------------------------|---|-------------------------------|---------------------------|
| Media_Clima_Geral_2_ciclos            | x | Entra-                        | 0.22                      |
| das_nao_negociadas_ano                |   |                               |                           |
| Otimiza_Processo_Trabalho             | x | Salario                       | 0.22                      |
| Media_GPF_Aprendizado_3_ciclos        | x | Resul-                        | 0.22                      |
| tado_Eixo_Desempenho                  |   |                               |                           |
| Salario                               | x | Resultado_Geral_9box          | 0.22                      |
| Autonomia                             | x | Salario                       | 0.22                      |
| Media_GPF_Financeira_3_ciclos         | x | Resul-                        | 0.22                      |
| tado_Eixo_Desempenho                  |   |                               |                           |
| Banco_de_horas                        | x | Me-                           | 0.22                      |
| dia_GPF_Processos_Int._3_ciclos       |   |                               |                           |
| Media_Pesq_Clima_Valorizacao_2_ciclos | x | Entra-                        | 0.21                      |
| das_nao_negociadas_ano                |   |                               |                           |
| Salario                               | x | Eficiência                    | 0.21                      |
| Banco_de_horas                        | x | Autonomia                     | 0.20                      |
| Media_GPF_Financeira_3_ciclos         | x | Me-                           | 0.20                      |
| dia_Geral_Placar_GDP_3_ciclos         |   |                               |                           |
| Resultado_Geral_9box                  | x | Banco_de_horas                | 0.20                      |
| Inovação                              | x | Salario                       | 0.20                      |
| Banco_de_horas                        | x | Eficiência                    | 0.20                      |
| Total_Cert_internas                   | x | Me-                           | 0.20                      |
| dia_GPF_Aprendizado_3_ciclos          |   |                               |                           |
| ab_acordo_atual                       | x | Adaptabilidade                | 0.19                      |
| saldo_folgas                          | x | Resultado_Geral_9box          | 0.19                      |
| Salario                               | x | Media_GPF_Financeira_3_ciclos | 0.19                      |
| Qtde.Cert..Intermediárias             | x | Me-                           | 0.19                      |
| dia_GPF_Aprendizado_3_ciclos          |   |                               |                           |
| Resultado_Geral_9box                  | x | ab_acordo_atual               | 0.19                      |
| Autonomia                             | x | ab_acordo_atual               | 0.19                      |
| Media_GPF_Clientes_3_ciclos           | x | Me-                           | 0.18                      |
| dia_Geral_Placar_GDP_3_ciclos         |   |                               |                           |
| folgas                                | x | Autonomia                     | 0.18                      |
| Entradas_nao_negociadas_ano           | x | Me-                           | 0.18                      |
| dia_Pesq_Clima_Estilo_Gestão_2_ciclos |   |                               |                           |
| Autonomia                             | x | saldo_folgas                  | 0.18                      |

| <b>Par de Variáveis</b>                                          | <b>Coefficiente de Correlação</b> |
|------------------------------------------------------------------|-----------------------------------|
| saldo_folgas x Adaptabilidade                                    | 0.18                              |
| Resultado_Geral_9box x folgas                                    | 0.18                              |
| Adaptabilidade x Banco_de_horas                                  | 0.18                              |
| Mais_2HE_dia_ano x Banco_de_horas                                | 0.18                              |
| Eficiência x ab_acordo_atual                                     | 0.18                              |
| Media_Geral_Placar_GDP_3_ciclos x Salario                        | 0.18                              |
| Media_GPF_Socioambiental x Media_Geral_Placar_GDP_3_ciclos       | 0.18                              |
| Resultado_Eixo_Desempenho x Salario                              | 0.17                              |
| <b>Parâmetros</b>                                                | <b>Correlação</b>                 |
| Media_GPF_Aprendizado_3_ciclos x Media_Geral_Placar_GDP_3_ciclos | 0.17                              |
| Banco_de_horas x Resultado_Competencia                           | 0.17                              |
| Horas.FIP x Colaboração                                          | 0.17                              |
| Resultado_Competencia x ab_acordo_atual                          | 0.17                              |
| Media_GPF_Aprendizado_3_ciclos x Qtde.Cert..Avançadas            | 0.17                              |
| Inovação x Qtde.Cert..Avançadas                                  | 0.17                              |
| Media_GPF_Processos_Int._3_ciclos x Salario                      | 0.17                              |
| folgas x Adaptabilidade                                          | 0.17                              |
| Media_GPF_Aprendizado_3_ciclos x ab_acordo_atual                 | 0.17                              |
| folgas x ab_acordo_atual                                         | 0.17                              |
| saldo_folgas x Media_GPF_Processos_Int._3_ciclos                 | 0.17                              |
| ab_acordo_atual x Banco_de_horas                                 | 0.17                              |
| Salario x Adaptabilidade                                         | 0.17                              |
| Media_GPF_Clientes_3_ciclos x Banco_de_horas                     | 0.16                              |
| Media_Clima_Relacionamento_2_ciclos x Banco_de_horas             | 0.16                              |
| saldo_folgas x Eficiência                                        | 0.16                              |
| Eficiência x saldo_folgas                                        | 0.16                              |
| Media_GPF_Processos_Int._3_ciclos x ab_acordo_atual              | 0.16                              |
| ab_acordo_atual x Media_GPF_Processos_Int._ciclos                | 0.16                              |

| <b>Par de Variáveis</b>                                                  | <b>Coeficiente de Correlação</b> |
|--------------------------------------------------------------------------|----------------------------------|
| Entradas_nao_negociadas_ano x Prorroga-<br>cao_nao_negociada_ano         | 0.16                             |
| Eficiência x folgas                                                      | 0.16                             |
| Media_GPF_Clientes_3_ciclos x ab_acordo_atual                            | 0.16                             |
| nr_depend_ir x tempo_banco                                               | 0.16                             |
| Media_GPF_Processos_Int._3_ciclos x Me-<br>dia_Geral_Placar_GDP_3_ciclos | 0.16                             |
| Qtde.Cert..Avançadas x Adaptabilidade                                    | 0.16                             |
| ab_acordo_atual x Otimiza_Processo_Trabalho                              | 0.16                             |
| Resultado_Competencia x Qtde.Cert..Avançadas                             | 0.15                             |
| Media_GPF_Processos_Int._3_ciclos x folgas                               | 0.15                             |
| ab_acordos_ant x anuenio                                                 | 0.15                             |
| Otimiza_Processo_Trabalho x Qtde.Cert..Avançadas                         | 0.15                             |
| anuenio x Salario                                                        | 0.15                             |
| Colaboração x Banco_de_horas                                             | 0.15                             |
| qtde_horas_treina_sem x Total_Cert_internas                              | 0.15                             |
| Qtde.Cert..Avançadas x Resultado_Geral_9box                              | 0.15                             |
| Media_GPF_Aprendizado_3_ciclos x Me-<br>dia_Geral_Placar_GDP_3_ciclos    | 0.17                             |
| Banco_de_horas x Resultado_Competencia                                   | 0.17                             |
| Horas.FIP x Colaboração                                                  | 0.17                             |
| Resultado_Competencia x ab_acordo_atual                                  | 0.17                             |
| Media_GPF_Aprendizado_3_ciclos x<br>Qtde.Cert..Avançadas                 | 0.17                             |
| Inovação x Qtde.Cert..Avançadas                                          | 0.17                             |
| Media_GPF_Processos_Int._3_ciclos x Salario                              | 0.17                             |
| folgas x Adaptabilidade                                                  | 0.17                             |
| Media_GPF_Aprendizado_3_ciclos x<br>ab_acordo_atual                      | 0.17                             |
| folgas x ab_acordo_atual                                                 | 0.17                             |
| saldo_folgas x Media_GPF_Processos_Int._3_ciclos                         | 0.17                             |
| ab_acordo_atual x Banco_de_horas                                         | 0.17                             |
| Salario x Adaptabilidade                                                 | 0.17                             |
| Media_GPF_Clientes_3_ciclos x Banco_de_horas                             | 0.16                             |

| Par de Variáveis                                  |   | Coeficiente de Correlação |
|---------------------------------------------------|---|---------------------------|
| Media_Clima_Relacionamento_2_ciclos               | x | 0.16                      |
| Banco_de_horas                                    |   |                           |
| saldo_folgas x Eficiência                         |   | 0.16                      |
| Eficiência x saldo_folgas                         |   | 0.16                      |
| Media_GPF_Processos_Int._3_ciclos                 | x | 0.16                      |
| ab_acordo_atual                                   |   |                           |
| ab_acordo_atual                                   | x | Me-0.16                   |
| dia_GPF_Processos_Int._3_ciclos                   |   |                           |
| Entradas_nao_negociadas_ano                       | x | Prorroga-0.16             |
| cao_nao_negociada_ano                             |   |                           |
| Eficiência x folgas                               |   | 0.16                      |
| Media_GPF_Clientes_3_ciclos x ab_acordo_atual     |   | 0.16                      |
| nr_depend_ir x tempo_banco                        |   | 0.16                      |
| Media_GPF_Processos_Int._3_ciclos                 | x | Me-0.16                   |
| dia_Geral_Placar_GDP_3_ciclos                     |   |                           |
| Qtde.Cert..Avançadas x Adaptabilidade             |   | 0.16                      |
| ab_acordo_atual x Otimiza_Processo_Trabalho       |   | 0.16                      |
| Resultado_Competencia x Qtde.Cert..Avançadas      |   | 0.15                      |
| Media_GPF_Processos_Int._3_ciclos x folgas        |   | 0.15                      |
| ab_acordos_ant x anuenio                          |   | 0.15                      |
| Otimiza_Processo_Trabalho x Qtde.Cert..Avançadas  |   | 0.15                      |
| anuenio x Salario                                 |   | 0.15                      |
| Colaboração x Banco_de_horas                      |   | 0.15                      |
| qtde_horas_treina_sem x Total_Cert_internas       |   | 0.15                      |
| Qtde.Cert..Avançadas x Resultado_Geral_9box       |   | 0.15                      |
| ab_acordo_atual x Inovação                        |   | 0.15                      |
| Otimiza_Processo_Trabalho x Banco_de_horas        |   | 0.15                      |
| Resultado_Competencia x saldo_folgas              |   | 0.15                      |
| Banco_de_horas x Media_GPF_Financeira_3_ciclos    |   | 0.14                      |
| Eficiência x Qtde.Cert..Avançadas                 |   | 0.14                      |
| Media_GPF_Financeira_3_ciclos                     | x | Me-0.14                   |
| dia_Clima_Relacionamento_2_ciclos                 |   |                           |
| Qtde.Cert..Intermediárias x qtde_horas_treina_sem |   | 0.13                      |
| saldo_folgas x Inovação                           |   | 0.13                      |

| Par de Variáveis                                                             | Coefficiente de Correlação |
|------------------------------------------------------------------------------|----------------------------|
| Banco_de_horas x Me-<br>dia_GPF_Aprendizado_3_ciclos                         | 0.13                       |
| ab_acordo_atual x dias_nao_util_trab_ano                                     | 0.13                       |
| Colaboração x saldo_folgas                                                   | 0.13                       |
| qtde_horas_treina_sem x dias_nao_util_trab_ano                               | 0.13                       |
| Banco_de_horas x Me-<br>dia_Pesq_Clima_Estilo_Gestão_2_ciclos                | 0.13                       |
| Media_Geral_Placar_GDP_3_ciclos x tempo_banco                                | 0.13                       |
| Media_GPF_Aprendizado_3_ciclos x<br>qtde_horas_treina_sem                    | 0.13                       |
| Media_Clima_Geral_2_ciclos x Banco_de_horas                                  | 0.13                       |
| ab_acordos_ant x ab_acordo_atual                                             | 0.13                       |
| Salario x Colaboração                                                        | 0.13                       |
| Media_GPF_Clientes_3_ciclos x Salario                                        | 0.13                       |
| faltas_categoria x anuenio                                                   | 0.13                       |
| folgas x Resultado_Competencia                                               | 0.12                       |
| saldo_folgas x Otimiza_Processo_Trabalho                                     | 0.12                       |
| ab_acordo_atual x Colaboração                                                | 0.12                       |
| Media_GPF_Financeira_3_ciclos x Me-<br>dia_Clima_Geral_2_ciclos              | 0.12                       |
| ab_acordo_atual x saldo_folgas                                               | 0.12                       |
| Resultado_Eixo_Competencias x Autonomia                                      | 0.12                       |
| Media_Clima_Relacionamento_2_ciclos x<br>Mais_2HE_dia_ano                    | 0.12                       |
| Qtde.Cert..Avançadas x Autonomia                                             | 0.12                       |
| saldo_folgas x Media_GPF_Clientes_3_ciclos                                   | 0.12                       |
| Total_Cert_internas x ab_acordo_atual                                        | 0.12                       |
| saldo_folgas x Media_Clima_Relacionamento_2_ciclos                           | 0.12                       |
| Resultado_Geral_9box x Horas.FIP                                             | 0.12                       |
| nr_depend_ir x Salario                                                       | 0.12                       |
| faltas_banco x anuenio                                                       | 0.12                       |
| Media_Clima_Relacionamento_2_ciclos x Me-<br>dia_GPF_Processos_Int._3_ciclos | 0.12                       |
| Inovação x Banco_de_horas                                                    | 0.12                       |
| Banco_de_horas x Resultado_Eixo_Desempenho                                   | 0.12                       |

| <b>Par de Variáveis</b>                                                 | <b>Coefficiente de Correlação</b> |
|-------------------------------------------------------------------------|-----------------------------------|
| ab_acordo_atual x Media_GPF_Financeira_3_ciclos                         | 0.12                              |
| Media_GPF_Clientes_3_ciclos x Qtde.Cert..Avançadas                      | 0.12                              |
| Total_Cert_internas x Total_abonos                                      | 0.12                              |
| folgas x Inovação                                                       | 0.12                              |
| Mais_2HE_dia_ano x saldo_folgas                                         | 0.12                              |
| Media_GPF_Financeira_3_ciclos x Media_Clima_Auto_Inov_2_ciclos          | 0.12                              |
| dias_nao_util_trab_ano x Total_abonos                                   | 0.12                              |
| folgas x Media_GPF_Clientes_3_ciclos                                    | 0.11                              |
| Total_abonos x folgas                                                   | 0.11                              |
| Qtde.Cert..Intermediárias x ab_acordo_atual                             | 0.11                              |
| Colaboração x folgas                                                    | 0.11                              |
| faltas_banco x Idade                                                    | 0.11                              |
| Total_abonos x ferias_aq                                                | 0.11                              |
| Resultado_Competencia x Horas.FIP                                       | 0.11                              |
| tempo_banco x faltas_categoria                                          | 0.11                              |
| Banco_de_horas x Prorrogaçao_nao_negociada_ano                          | 0.11                              |
| Media_GPF_Financeira_3_ciclos x Media_Pesq_Clima_Estilo_Gestão_2_ciclos | 0.11                              |
| Resultado_Geral_9box x Resultado_Eixo_Competencias                      | 0.11                              |
| Prorrogaçao_nao_negociada_ano x Resultado_Geral_9box                    | 0.11                              |
| Banco_de_horas x Media_Clima_Auto_Inov_2_ciclos                         | 0.11                              |
| dias_nao_util_trab_ano x Media_Geral_Placar_GDP_3_ciclos                | 0.11                              |
| Total_abonos x Qtde.Cert..Intermediárias                                | 0.11                              |
| Banco_de_horas x Entradas_nao_negociadas_ano                            | 0.11                              |
| Media_Clima_Geral_2_ciclos x Media_GPF_Processos_Int._3_ciclos          | 0.11                              |
| Otimiza_Processo_Trabalho x Media_GPF_Socioambiental                    | 0.11                              |

| Par de Variáveis                        |     | Coeficiente de Correlação |
|-----------------------------------------|-----|---------------------------|
| Media_GPF_Processos_Int._3_ciclos       | x   | 0.11                      |
| Qtde.Cert..Avançadas                    |     |                           |
| Media_GPF_Processos_Int._3_ciclos       | x   | 0.11                      |
| Media_Clima_Auto_Inov_2_ciclos          | Me- |                           |
| Media_Pesq_Clima_Valorizacao_2_ciclos   | x   | 0.11                      |
| Media_GPF_Financeira_3_ciclos           | Me- |                           |
| Mais_2HE_dia_ano                        | x   | 0.10                      |
| Autonomia                               |     |                           |
| Adaptabilidade                          | x   | 0.10                      |
| Horas.FIP                               |     |                           |
| ferias_aq                               | x   | 0.10                      |
| ab_acordos_ant                          |     |                           |
| folgas                                  | x   | 0.10                      |
| Otimiza_Processo_Trabalho               |     |                           |
| Resultado_Competencia                   | x   | 0.10                      |
| Media_GPF_Socioambiental                | Me- |                           |
| Resultado_Geral_9box                    | x   | 0.10                      |
| Mais_2HE_dia_ano                        |     |                           |
| Qtde.Cert..Intermediárias               | x   | 0.10                      |
| Media_GPF_Processos_Int._3_ciclos       | Me- |                           |
| Salario                                 | x   | 0.10                      |
| Banco_de_horas                          |     |                           |
| Media_GPF_Clientes_3_ciclos             | x   | 0.10                      |
| Prorroga-cao_nao_negociada_ano          |     |                           |
| Dias.Absenteísmo                        | x   | -0.10                     |
| Resultado_Competencia                   |     |                           |
| faltas_categoria                        | x   | -0.10                     |
| ab_acordo_atual                         |     |                           |
| Entradas_nao_negociadas_ano             | x   | -0.10                     |
| Media_Geral_Placar_GDP_3_ciclos         | Me- |                           |
| Media_Pesq_Clima_Estilo_Gestão_2_ciclos | x   | -0.11                     |
| Total_Cert_internas                     |     |                           |
| saldo_folgas                            | x   | -0.11                     |
| Idade                                   |     |                           |
| faltas_categoria                        | x   | -0.11                     |
| Resultado_Geral_9box                    |     |                           |
| faltas_banco                            | x   | -0.11                     |
| Resultado_Eixo_Desempenho               |     |                           |
| Media_GPF_Clientes_3_ciclos             | x   | -0.11                     |
| Dias.Absenteísmo                        |     |                           |
| Idade                                   | x   | -0.11                     |
| folgas                                  |     |                           |
| Media_GPF_Aprendizado_3_ciclos          | x   | -0.11                     |
| faltas_categoria                        |     |                           |
| Media_GPF_Clientes_3_ciclos             | x   | -0.11                     |
| tempo_banco                             |     |                           |
| faltas_categoria                        | x   | -0.11                     |
| Media_GPF_Clientes_3_ciclos             |     |                           |
| Banco_de_horas                          | x   | -0.11                     |
| Idade                                   |     |                           |
| Eficiência                              | x   | -0.11                     |
| faltas_categoria                        |     |                           |

| Par de Variáveis                                               |     | Coeficiente de Correlação |
|----------------------------------------------------------------|-----|---------------------------|
| Media_Pesq_Clima_Valorizacao_2_ciclos<br>qtde_horas_treina_sem | x   | -0.11                     |
| Autonomia x Dias.Ausências                                     |     | -0.11                     |
| ab_acordos_ant x Indice._de_Pontualidade                       |     | -0.11                     |
| Media_GPF_Processos_Int._3_ciclos<br>Dias.Absenteísmo          | x   | -0.12                     |
| faltas_banco x Media_GPF_Aprendizado_3_ciclos                  |     | -0.12                     |
| tempo_banco x Qtde.Cert..Básicas                               |     | -0.12                     |
| Media_GPF_Processos_Int._3_ciclos x faltas_banco               |     | -0.12                     |
| Dias.Ausências x Resultado_Competencia                         |     | -0.12                     |
| Otimiza_Processo_Trabalho x tempo_banco                        |     | -0.12                     |
| faltas_categoria x<br>Media_GPF_Processos_Int._3_ciclos        | Me- | -0.12                     |
| faltas_banco x Total_abonos                                    |     | -0.12                     |
| Inovação x faltas_categoria                                    |     | -0.12                     |
| Indice._de_Pontualidade x Total_abonos                         |     | -0.13                     |
| Colaboração x faltas_banco                                     |     | -0.13                     |
| faltas_categoria x Resultado_Competencia                       |     | -0.13                     |
| Inovação x tempo_banco                                         |     | -0.13                     |
| Otimiza_Processo_Trabalho x faltas_categoria                   |     | -0.13                     |
| Media_Geral_Placar_GDP_3_ciclos x Colaboração                  |     | -0.13                     |
| Media_GPF_Aprendizado_3_ciclos<br>Dias.Absenteísmo             | x   | -0.13                     |
| Autonomia x tempo_banco                                        |     | -0.13                     |
| tempo_banco x Eficiência                                       |     | -0.13                     |
| tempo_banco x Media_GPF_Aprendizado_3_ciclos                   |     | -0.13                     |
| Media_Clima_Geral_2_ciclos<br>qtde_horas_treina_sem            | x   | -0.13                     |
| tempo_banco x Total_Cert_internas                              |     | -0.13                     |
| folgas x Media_Pesq_Clima_Estilo_Gestão_2_ciclos               |     | -0.14                     |
| Inovação x Dias.Ausências                                      |     | -0.14                     |
| Indice._de_Pontualidade x Banco_de_horas                       |     | -0.14                     |
| Entradas_nao_negociadas_ano<br>dias_nao_util_trab_ano          | x   | -0.14                     |
| Media_Clima_Relacionamento_2_ciclos x folgas                   |     | -0.14                     |



| <b>Par de Variáveis</b>                                         | <b>Coefficiente de Correlação</b> |
|-----------------------------------------------------------------|-----------------------------------|
| tempo_banco x Qtde.Cert..Avançadas                              | -0.14                             |
| Resultado_Eixo_Desempenho x Dias.Ausências                      | -0.14                             |
| tempo_banco x Colaboração                                       | -0.15                             |
| Media_Clima_Relacionamento_2_ciclos x qtde_horas_treina_sem     | -0.15                             |
| Resultado_Geral_9box x faltas_banco                             | -0.15                             |
| faltas_banco x Otimiza_Processo_Trabalho                        | -0.15                             |
| faltas_banco x Adaptabilidade                                   | -0.15                             |
| Total_Cert_internas x Salario                                   | -0.15                             |
| tempo_banco x Resultado_Competencia                             | -0.15                             |
| Media_Pesq_Clima_Estilo_Gestão_2_ciclos x qtde_horas_treina_sem | -0.15                             |
| faltas_banco x Autonomia                                        | -0.15                             |
| faltas_banco x Eficiência                                       | -0.16                             |
| Media_Clima_Geral_2_ciclos x folgas                             | -0.16                             |
| Media_Pesq_Clima_Valorizacao_2_ciclos x folgas                  | -0.16                             |
| Resultado_Competencia x faltas_banco                            | -0.16                             |
| Eficiência x Dias.Ausências                                     | -0.16                             |
| Dias.Ausências x Otimiza_Processo_Trabalho                      | -0.16                             |
| Salario x Qtde.Cert..Básicas                                    | -0.16                             |
| Entradas_nao_negociadas_ano x Idade                             | -0.16                             |
| faltas_banco x ab_acordo_atual                                  | -0.16                             |
| Resultado_Geral_9box x tempo_banco                              | -0.17                             |
| folgas x Media_Clima_Auto_Inov_2_ciclos                         | -0.17                             |
| tempo_banco x Adaptabilidade                                    | -0.17                             |
| nr_depend_creche_novo x Idade                                   | -0.17                             |
| Resultado_Eixo_Desempenho x Idade                               | -0.18                             |
| Idade x Media_GPF_Processos_Int._3_ciclos                       | -0.19                             |
| Idade x ab_acordo_atual                                         | -0.20                             |
| Colaboração x Idade                                             | -0.21                             |
| Idade x Otimiza_Processo_Trabalho                               | -0.21                             |
| Idade x Media_GPF_Clientes_3_ciclos                             | -0.22                             |
| Idade x Media_GPF_Aprendizado_3_ciclos                          | -0.23                             |
| Idade x Qtde.Cert..Intermediárias                               | -0.24                             |
| Resultado_Competencia x Idade                                   | -0.25                             |

| Par de Variáveis                                                      | Coefficiente de Correlação |
|-----------------------------------------------------------------------|----------------------------|
| Qtde.Cert..Avançadas x Idade                                          | -0.26                      |
| Idade x Eficiência                                                    | -0.26                      |
| Autonomia x Idade                                                     | -0.26                      |
| Inovação x Idade                                                      | -0.27                      |
| Idade x Total_Cert_internas                                           | -0.27                      |
| Resultado_Geral_9box x Idade                                          | -0.29                      |
| tempo_banco x ab_acordo_atual                                         | -0.30                      |
| Adaptabilidade x Idade                                                | -0.30                      |
| dias_nao_util_trab_ano x Me-<br>dia_Pesq_Clima_Valorizacao_2_ciclos   | -0.36                      |
| dias_nao_util_trab_ano x Me-<br>dia_Pesq_Clima_Estilo_Gestão_2_ciclos | -0.37                      |
| Media_Clima_Auto_Inov_2_ciclos x<br>dias_nao_util_trab_ano            | -0.40                      |
| ab_acordo_atual x anuenio                                             | -0.40                      |
| dias_nao_util_trab_ano x Me-<br>dia_Clima_Geral_2_ciclos              | -0.41                      |
| dias_nao_util_trab_ano x Me-<br>dia_Clima_Relacionamento_2_ciclos     | -0.44                      |
| Media_Geral_Placar_GDP_3_ciclos x Horas.FIP                           | -0.61                      |

## B Lista de Atributos

- **Dados pessoais:**

- Idade
- Gênero
- Estado civil
- Raça
- Quantidade de dependentes
- UF de localização

- **Dados de carreiras:**

- Cargo atual
- Banco de horas

- Dias de absenteísmo
- Tempo de empresa
- Férias disponíveis
- Saldo de folgas
- Dias de ausências
- Anuênio
- Abonos acordo anterior
- Abonos acordo atual
- Total de abonos
- Faltas totais
- Faltas por categoria
- Índice de pontualidade
- Salário
- Número de entradas não negociadas
- Número de prorrogações não autorizadas
- Dias não úteis de trabalho
- Exame periódico
- Opção por aposentadoria privada
- Carga horária
- Turno

• **Dados de competências:**

- Autonomia
- Adaptabilidade
- Inovação
- Colaboração
- Resultado no eixo de competências
- Otimização de processos de trabalho
- Eficiência
- Resultado no eixo de desempenho
- Resultado 9box
- Média da avaliação financeira em 3 ciclos

- Média da avaliação de clientes em 3 ciclos
  - Média da avaliação de aprendizado em 3 ciclos
  - Média da avaliação de processos em 3 ciclos
  - Média da avaliação socioambiental em 3 ciclos
  - Média da avaliação geral em 3 ciclos
  - Média da pesquisa de clima sobre autonomia e inovação em 2 ciclos
  - Média da pesquisa de clima sobre relacionamento em 2 ciclos
  - Média da pesquisa de clima sobre estilo de gestão em 2 ciclos
  - Média da pesquisa de clima sobre valorização em 2 ciclos
  - Média geral da pesquisa de clima em 2 ciclos
- **Dados de formação:**
    - Grau de instrução
    - Bolsa de estudo
    - Quantidade de certificações básicas
    - Quantidade de certificações intermediárias
    - Quantidade de certificações avançadas
    - Total de certificações internas
    - Horas de treinamento no semestre
    - Conhecimentos
    - Habilidades
    - Ferramentas

## C Lista de Atributos

- **Dados Pessoais:**
  - Idade
  - Genero
  - PPP
  - PCD
  - UF de prova
  - Classificacao

- **Anos de experiência profissional:** (Mapeados juntos a especialistas da organização, com base nos principais conhecimentos utilizados)
  - Agile Coach
  - AI Ops
  - AI Tools
  - Analytics
  - Android
  - Ansible
  - Application Server
  - Inteligencia Artificial
  - Assembly
  - AWS
  - Azure
  - Banco de dados distribuídos
  - C/C ++
  - CI/CD
  - Cobit
  - Cobol
  - Computing Vision
  - CX
  - Engenharia de Dados
  - Ciência de Dados
  - Design
  - Docker
  - FinOps
  - Flutter
  - Framework Angular
  - Framework React
  - Go
  - Google Cloud
  - IOS

- ISO 20000
- IT Infraesture
- ITL/ITSM
- IISM
- Java
- Java Script
- JCL
- Kanban
- Kubernetes
- Mainframe CICS/Batch
- Mainframe DB2
- Mainframe Hardware
- Mainframe Software
- Middleware
- Monitoramento
- Natural
- Net
- NLP/LLMs
- Node JS
- Observabilidade
- OpenStack
- Operational System
- Others
- Publics Clouds
- Pupper
- Python
- Quarkus
- React Native
- Research
- Scala
- Scrum

- Signal Processement
- Spring Boot
- SRE
- Storage
- Terraform
- UX
- Virtual Server
- Web Developer
- Writing

- **Áreas de maior interesse de trabalho:**

- Agile Coach
- Cibersegurança
- Desenvolvedor Cloud Java/Java Script
- Desenvolvedor Mainframe/Cobol
- Desenvolvedor Mobile
- Desenvolvedor Web
- Especialista em UX
- Governança de TI
- Infraestutura de TI
- Inteligencia Artificial/Ciência de dados

# Referências

- [1] Andersen, M. K. (2017). Human capital analytics: the winding road. *Journal of Organizational Effectiveness: People and Performance*, 4(2). 2, 3
- [2] Angrave, D., Charlwood, A., Kirkpatrick, I., Lawrence, M., and Stuart, M. (2016). Hr and analytics: why hr is set to fail the big data challenge. *Human Resource Management Journal*, 26(1):1–11. 17, 24
- [3] Bazerman, M. H. and Moore, D. A. (2012). *Judgment in Managerial Decision Making*. Wiley. 14
- [4] Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer. 29, 31
- [5] Boudreau, J. W. and Ramstad, P. M. (2007). *Beyond HR: The New Science of Human Capital*. Harvard Business Review Press. 14, 16
- [6] Bresciani, S., Ferraris, A., and Del Giudice, M. (2021). The management of organizational resources as antecedent to the innovation paradox: The role of big data analytics. *Technological Forecasting and Social Change*, 171:120986. 2
- [7] C., S. M. (2007). The past, present, and future of dynamic performance research. research in personel and human resources management. 60
- [8] Cao, M., Zhao, S., and Xu, Y. (2021). How hr systems are implemented matters: high-performance work systems and employees’ thriving at work. *Asia Pacific Journal of Human Resources*. 3
- [9] Cascio, W. F. and Boudreau, J. W. (2011). *Investing in people: Financial impact of human resource initiatives*. FT Press. 29
- [10] Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., and Wirth, R. (2000). Crisp-dm 1.0: Step-by-step data mining guide. Technical report, SPSS Inc. 7, 8
- [11] Chen, H., Chiang, R. H., and Storey, V. C. (2014). Business intelligence and analytics: From big data to big impact. *MIS quarterly*, pages 1165–1188. 30
- [12] Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press. 39
- [13] Cropanzano, R., Bowen, D. E., and Gilliland, S. W. (2007). The management of organizational justice. *Academy of Management Perspectives*, 21(4):34–48. 15



- [14] Davenport, T. H. and Harris, J. G. (2007). *Competing on Analytics: The New Science of Winning*. Harvard Business Review Press. 8, 18
- [15] Davenport, T. H., Harris, J. G., and Morison, R. (2010). The new world of business analytics. *International Institute for Analytics*. 8
- [16] Dineen, B. R., . W. (2012). Screening-oriented recruitment messages: Antecedents and relationships with applicant pool quality. pages 343–360. 50
- [17] Edwards, M. R. and Edwards, K. (2019). *Predictive HR Analytics: Mastering the HR Metric*. Kogan Page, London. 8, 23
- [18] Falletta, S. (2014). In search of hr intelligence: Evidence-based hr analytics practices in high performing companies. *People & Strategy*, 36(4):28–37. 18, 22
- [19] Farber, L. (2009). *Estatística Aplicada*. Pearson Education do Brasil, São Paulo, 4<sup>a</sup> edição edition. 39, 40
- [20] Ferguson, R. (2019). *Practical Data Science with Python*. Apress. 30
- [21] Fitz-enz, J. (2010). *The New HR Analytics: Predicting the Economic Value of Your Company's Human Capital Investments*. AMACOM. 16
- [22] Fitz-enz, J. and Mattox II, J. R. (2014). *Predictive Analytics for Human Resources*. Wiley & SAS Business Series. John Wiley & Sons, Hoboken, New Jersey. 3
- [23] Freedman, D. A. (2009). *Statistical Models: Theory and Practice*. Cambridge University Press. 39
- [24] Gartner (2013). The importance of 'big data': A definition. *Gartner IT Glossary*. 8
- [25] Goleman, D. (2000). *Working with Emotional Intelligence*. Bantam Books. 13
- [26] Guthrie, J. P. and Datta, D. K. (1998). Corporate strategy, executive selection, and firm performance. *Human Resource Management*, 37(2):101–115. Available at: [https://doi.org/10.1002/\(SICI\)1099-050X\(199822\)37:2%3C101::AID-HRM2%3E3.0.CO;2-%23](https://doi.org/10.1002/(SICI)1099-050X(199822)37:2%3C101::AID-HRM2%3E3.0.CO;2-%23). 50
- [27] Harris, J. G., Craig, E., and Light, D. A. (2011). Talent and analytics: new approaches, higher roi. *Journal of Business Strategy*, 32(6):4–13. 4, 16
- [28] Hastie, R. and Dawes, R. M. (2010). *Rational Choice in an Uncertain World: The Psychology of Judgment and Decision Making*. SAGE Publications. 14
- [29] Hausknecht, J. P., Rodda, J., and Howard, M. J. (2009). Targeted employee retention: Performance-based and job-related differences in reported reasons for staying. *Human Resource Management*, 48(2):269–288. 13
- [30] Highhouse, S. (2008). Stubborn reliance on intuition and subjectivity in employee selection. *Industrial and Organizational Psychology*, 1(3):333–342. 13

- [31] Karimi, M. and Viliyani, K. S. (2024). Employee turnover analysis using machine learning algorithms. *arXiv preprint arXiv:2402.03905*. 23
- [32] Keeney, R. L. and Raiffa, H. (1993). *Decisions with Multiple Objectives: Preferences and Value Trade-offs*. Cambridge University Press. 11
- [33] Kesner, I. F. and Sebor, T. C. (1994). Executive succession: Past, present, future. *Journal of Management*, 20(2):327–372. Available at: [https://psycnet.apa.org/doi/10.1016/0149-2063\(94\)90019-1](https://psycnet.apa.org/doi/10.1016/0149-2063(94)90019-1). 50
- [34] Kroenke, D. M. (2012). *Database Concepts*. Pearson Education. 29
- [35] Król, K. and Zdonek, D. (2020). Analytics maturity models: An overview. *Information*, 11(3):142. 5
- [36] Kuhn, H. W. (1955). The hungarian method for the assignment problem. *Naval Research Logistics (NRL)*, 2(1-2):83–97. 29, 31
- [37] Levenson, A. (2018). Using workforce analytics to improve strategy execution. *Human Resource Management*, 57(3):685–700. 17
- [38] Marler, J. H. and Boudreau, J. W. (2017). An evidence-based review of hr analytics. *The International Journal of Human Resource Management*, 28(1):3–26. 18
- [39] Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill. 30
- [40] Pearson, K. (1895). Notes on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, 58:240–242. 39
- [41] Peeters, T., Paauwe, J., and Van De Voorde, K. (2020). People analytics effectiveness: developing a framework. *Journal of Organizational Effectiveness: People and Performance*, 7(2):203–219. 24
- [42] Peres, A. R. and Laurindo, F. J. B. (2022). Uma proposta de framework para a execução e estruturação do people analytics. In *XLII Encontro Nacional de Engenharia de Produção*. ABEPRO. 26
- [43] Pessach, D. et al. (2020). Employees recruitment: A prescriptive analytics approach via machine learning and mathematical programming. *Elsevier Journal*, 134. 50, 60
- [44] Polzer, J. T. (2023). The rise of people analytics and the future of organizational research. *Research in Organizational Behavior*, 42:100181. 23
- [45] Provost, F. and Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big data*, 1(1):51–59. 2
- [46] Rasmussen, T. and Ulrich, D. (2015). Learning from practice: how hr analytics avoids being a management fad. *Organizational Dynamics*, 44(3):236–242. 2, 24
- [47] Rothwell, W. J. (2010). *Effective Succession Planning: Ensuring Leadership Continuity and Building Talent from Within*. AMACOM. 13

- [48] Santos, Q. (2018). *Gestão de pessoas: do processo intuitivo ao People Analytics*. Einstein (SP). 3, 60
- [49] Savage, L. J. (1954). *The Foundations of Statistics*. Dover Publications. 12
- [50] Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157):1124–1131. 12
- [51] Van den Heuvel, S. and Bondarouk, T. (2017). The rise (and fall?) of hr analytics: A study into the future application, value, structure, and system support. *Journal of Organizational Effectiveness: People and Performance*, 4(2):157–178. 21
- [52] Van Rossum, G. and Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace. 29
- [53] Von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press. 11
- [54] West, M. (2019). *People Analytics For Dummies*. John Wiley & Sons, Hoboken, New Jersey. 19, 23, 51, 61
- [55] White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., and Schmidt, D. C. (2023). A prompt pattern catalog to enhance prompt engineering with chatgpt. *ArXiv*, abs/2302.11382. 67
- [56] Wixom, B. H. and Watson, H. J. (2010). The bi-based organization. *International Journal of Business Intelligence Research (IJBIR)*, 1(1):13–28. 29
- [57] Yinusa, A. and Faezipour, M. (2023). Optimizing healthcare delivery: A model for staffing, patient assignment, and resource allocation. *Applied System Innovation*, 6(5):78. 20