**Relatório de Análise de Dados das Empresas do Ambition Box**

Projeto por Victor Matheus Araújo Oliveira

Data de Conclusão: 22/02/2024

Contato:

* Email: victormatheus@alu.ufc.br
* LinkedIn: www.linkedin.com/in/victor-mattheus
* Github: https://github.com/VictorMatheus

**Visão Geral**

Este projeto consiste em uma análise de dados das empresas listadas no Ambition Box, uma plataforma que fornece insights sobre avaliações de empresas, salários, entrevistas e mais. O conjunto de dados contém informações sobre 10.000 empresas, incluindo seus nomes, descrições, avaliações dos usuários, áreas altamente e criticamente avaliadas, número total de avaliações, salários médios, número de entrevistas realizadas, número total de vagas disponíveis e informações sobre benefícios oferecidos.

**Objetivos**

* Explorar e entender as características das empresas listadas no Ambition Box.
* Identificar padrões e insights úteis relacionados às avaliações dos usuários e outras métricas fornecidas.
* Fornecer uma análise abrangente para auxiliar na tomada de decisões relacionadas a carreira e pesquisa de mercado.

Fortaleza - CE

Fevereiro - 2024

**Metodologia**

* Aquisição de Dados: O conjunto de dados foi obtido do Kaggle, uma plataforma online para compartilhamento de conjuntos de dados.
* Limpeza e Preparação de Dados: Os dados foram limpos e preparados para análise, incluindo tratamento de valores ausentes e formatação de colunas.
* Análise Descritiva Inicial: Foram realizadas análises descritivas iniciais para entender a distribuição das variáveis e identificar padrões.
* Visualização de Dados: Diversas visualizações foram criadas para representar graficamente os dados e facilitar a interpretação.
* Análise de Correlação: Foi realizada uma análise de correlação entre as variáveis para identificar possíveis relações entre elas.
* Identificação das Empresas com Melhores e Piores Avaliações: Foram identificadas as empresas com as melhores e piores avaliações médias e comparadas suas características.

**Sobre o Conjunto de Dados**

O conjunto de dados utilizado neste projeto é proveniente do Ambition Box, uma plataforma online que oferece insights valiosos sobre empresas, incluindo avaliações de usuários, salários, entrevistas e mais. Composto por informações detalhadas de 10.000 empresas, o conjunto de dados abrange uma ampla gama de métricas, como avaliações gerais, áreas altamente e criticamente avaliadas, número total de revisões, médias salariais, entrevistas realizadas, oportunidades de emprego e benefícios oferecidos. Esses dados proporcionam uma visão abrangente do panorama empresarial, possibilitando análises detalhadas e a identificação de tendências significativas no mercado de trabalho.O conjunto de dados inclui as seguintes colunas:

1. **Company\_name:** O nome da empresa listada na Ambition Box.
2. **Descrição:** uma breve descrição da empresa, geralmente obtida de seu perfil ou site.
3. **Avaliações:** A classificação geral da empresa fornecida pelos usuários do Ambition Box.
4. **Highly\_rated\_for:** Áreas ou aspectos da empresa que são bem avaliados pelos usuários.
5. **Critical\_rated\_for:** Áreas ou aspectos da empresa que são avaliados criticamente pelos usuários.
6. **Total\_reviews:** O número total de avaliações enviadas para a empresa no Ambition Box.
7. **Avg\_salary:** O salário médio informado para a empresa, se disponível.
8. **Interviews\_taken:** Quantidade de entrevistas realizadas pela empresa, se disponível.
9. **Total\_jobs\_available:** O número total de cargos disponíveis na empresa.
10. **Total\_benefits:** Informações sobre benefícios oferecidos pela empresa, caso existam.

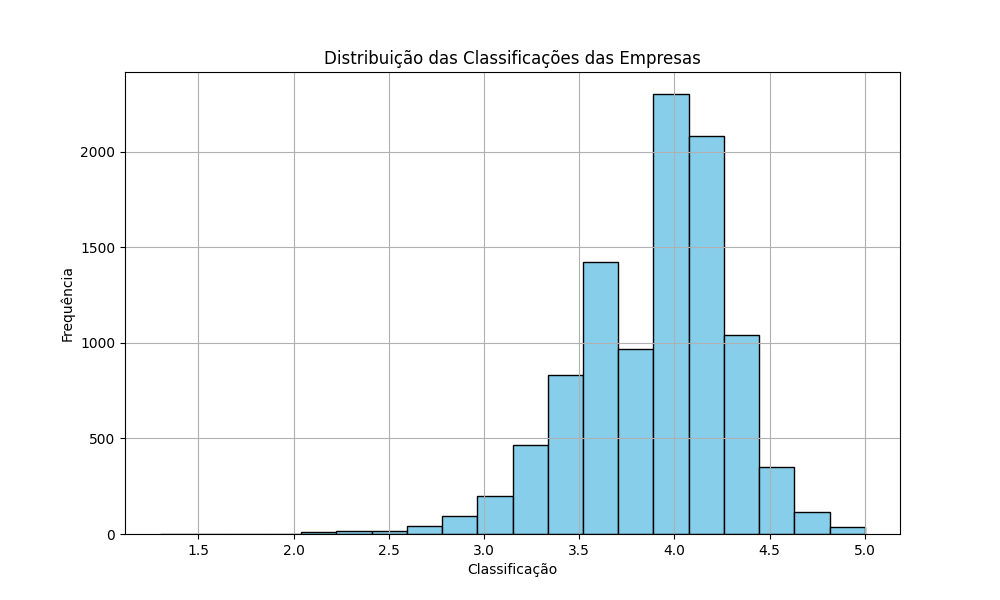
O conjunto de dados está disponível no site Kaggle pelo link a seguir: <https://www.kaggle.com/datasets/vedantkhapekar/top-10000-companies-dataset/data>

**Desafios e Resultados**

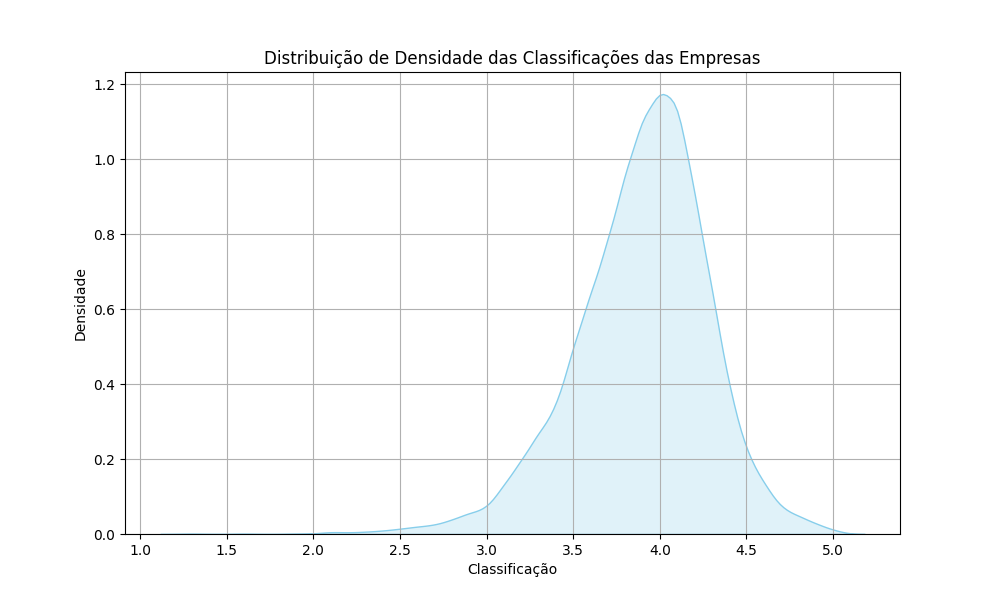
Os resultados apresentados neste relatório foram obtidos por meio da análise de dados utilizando códigos em Python. Todos os códigos utilizados estão disponíveis no arquivo ambition\_box\_analysis.py, que acompanha este relatório. A análise foi conduzida com base em 6 desafios propostos e resolvidos pelo autor deste relatório.

Desafio 1: Qual é a distribuição das classificações (ratings) das empresas? Existe algum padrão notável?

Distribuição das Classificações das Empresas:



Distribuição de Densidade das Classificações das Empresas:



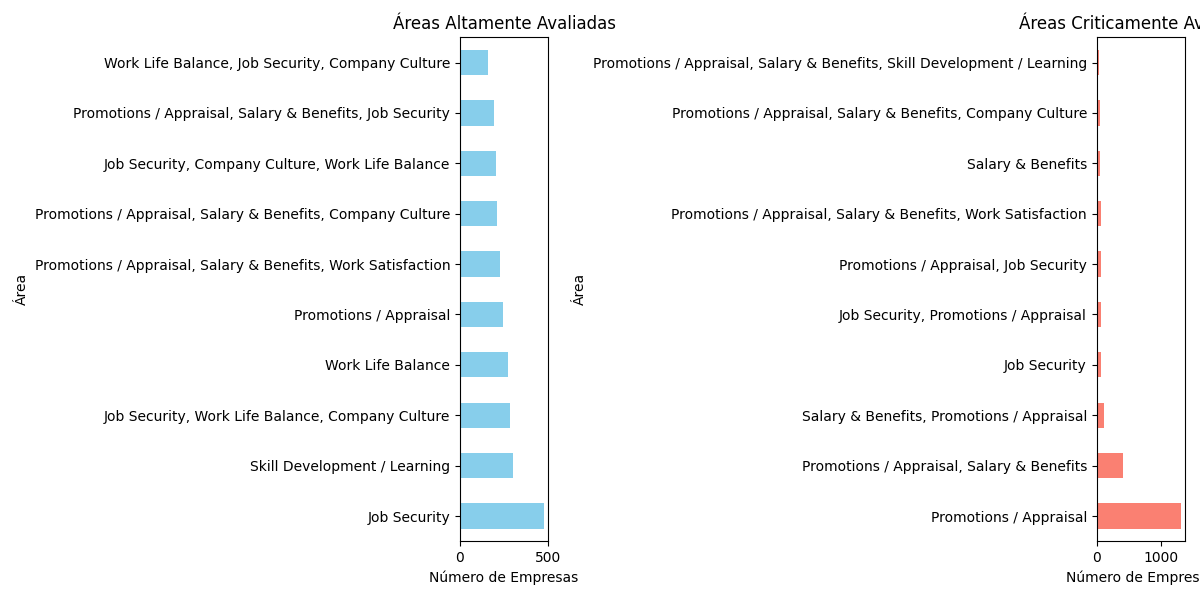
Concentração em Faixas Específicas: Há uma concentração de classificações entre 3,5 e 3,7, com uma frequência média de aproximadamente 1000 a 1500. Isso sugere que muitas empresas são avaliadas nessa faixa, indicando uma certa neutralidade ou moderada satisfação dos usuários com essas empresas. As classificações entre 4,0 e 4,2 também apresentam uma concentração significativa, com uma frequência média de 2100 a 2600. Isso pode indicar que muitas empresas são bem avaliadas nessa faixa, sugerindo uma alta satisfação dos usuários com essas empresas. As classificações mais altas, como 4,4, 4,5 e 5, mostram uma frequência média relativamente menor. Isso pode indicar que menos empresas recebem classificações excepcionalmente altas, o que pode ser atribuído a uma variedade de fatores, como altos padrões de qualidade ou uma pequena amostra de empresas excepcionais.

Distribuição Assimétrica: A distribuição das classificações parece ser assimétrica, com uma maior frequência de empresas recebendo classificações na faixa média (3,5 a 4,2) e uma diminuição na frequência à medida que as classificações aumentam para faixas mais altas.

Essa análise proporciona uma compreensão valiosa da distribuição das classificações das empresas e pode nos ajudar a identificar áreas de foco para análises mais detalhadas, como investigar as razões por trás das classificações mais baixas ou examinar o que diferencia as empresas altamente avaliadas das outras.

Desafio 2: Quais são as áreas ou aspectos mais frequentemente destacados como altamente avaliados (Highly\_rated\_for) e criticamente avaliados (Critically\_rated\_for)?

Áreas altamente avaliadas e Áreas criticamente avaliadas:



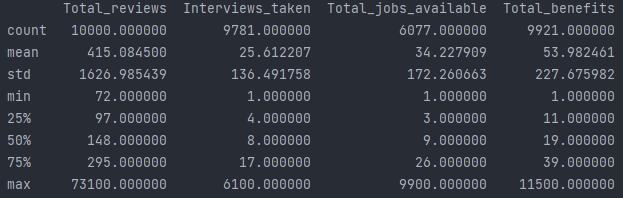
Áreas Altamente Avaliadas: Job security (segurança no emprego) emerge como a área mais bem avaliada na maioria das empresas, com quase 500 empresas destacando-a como altamente avaliada. Isso sugere que a segurança no emprego é uma prioridade ou ponto forte comum entre essas empresas.

Áreas Criticamente Avaliadas: Promotions / appraisal (promoções / avaliação de desempenho) é identificada como a área mais criticamente avaliada na maior parte das empresas, com aproximadamente 1400 empresas destacando-a como criticamente avaliada. Isso indica que muitas empresas enfrentam desafios ou questões significativas relacionadas a promoções e avaliação de desempenho, o que pode afetar negativamente a satisfação dos funcionários.

Essas descobertas podem fornecer insights importantes para empresas que buscam melhorar sua reputação e satisfação dos funcionários, destacando áreas de força e áreas que precisam de atenção e melhorias.

Desafio 3: Qual é a distribuição dos valores nas colunas Total\_reviews, Interviews\_taken, Total\_jobs\_available e Total\_benefits?

Resumo estatístico das colunas:



Total\_reviews:

* O número médio de revisões (Total\_reviews) é de aproximadamente 415, com uma grande variação devido ao desvio padrão elevado (1626.99).
* A mediana está em torno de 148, o que indica que metade das empresas tem menos de 148 revisões, enquanto a outra metade tem mais.
* A quantidade mínima de revisões é 72 e a máxima é 73100.

Interviews\_taken:

* A média de entrevistas realizadas (Interviews\_taken) é de cerca de 25, com um desvio padrão considerável (136.49).
* A mediana está em torno de 8, o que significa que metade das empresas realizou menos de 8 entrevistas, enquanto a outra metade realizou mais.
* O número mínimo de entrevistas realizadas é 1 e o máximo é 6100.

Total\_jobs\_available:

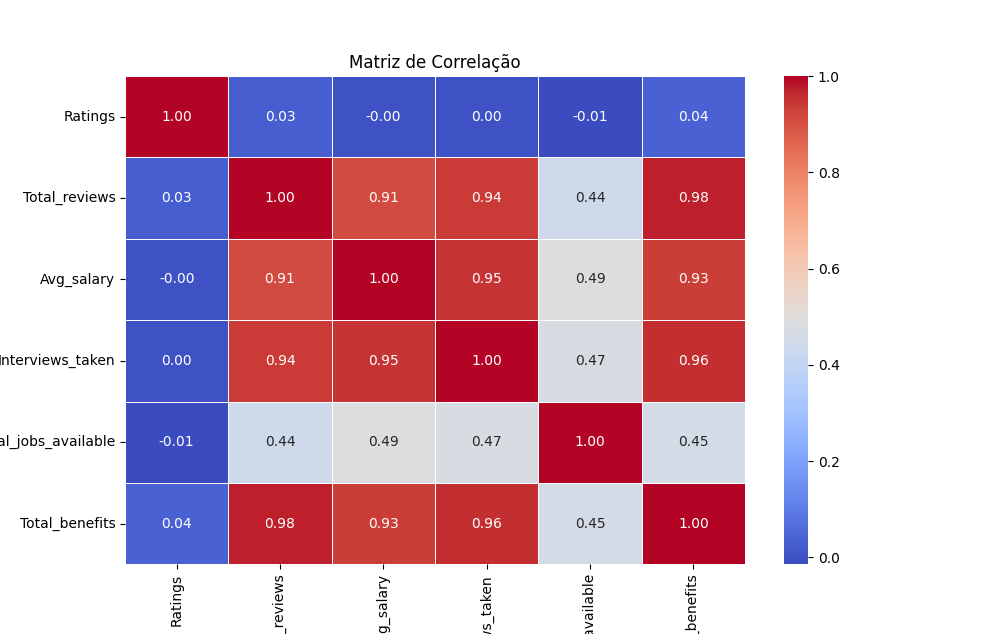
* A média de vagas disponíveis (Total\_jobs\_available) é de aproximadamente 34, com uma grande variação (desvio padrão de 172.26).
* A mediana está em torno de 9, o que indica que metade das empresas tem menos de 9 vagas disponíveis, enquanto a outra metade tem mais.
* O número mínimo de vagas disponíveis é 1 e o máximo é 9900.

Total\_benefits:

* A média de benefícios oferecidos (Total\_benefits) é de cerca de 54, com um desvio padrão considerável (227.68).
* A mediana está em torno de 19, o que significa que metade das empresas oferece menos de 19 benefícios, enquanto a outra metade oferece mais.
* O número mínimo de benefícios é 1 e o máximo é 11500.

Desafio 4: Analisar as correlações com matriz de correlação

Matriz de correlação:



Correlações:

1. Total\_reviews e Avg\_salary: Correlação de 0.91

* Existe uma correlação positiva forte (0.91) entre o número total de avaliações e o salário médio. Isso sugere que as empresas que oferecem salários mais altos tendem a receber um maior número de avaliações.

1. Total\_reviews e Interviews\_taken: Correlação de 0.94

* Há uma correlação positiva muito forte (0.94) entre o número total de avaliações e o número de entrevistas realizadas pelas empresas. Isso pode indicar que as empresas que conduzem mais entrevistas geralmente recebem mais avaliações de seus funcionários ou candidatos.

1. Total\_reviews e Total\_benefits: Correlação de 0.98

* Existe uma correlação positiva extremamente forte (0.98) entre o número total de avaliações e o número total de benefícios oferecidos pelas empresas. Isso sugere que empresas que oferecem mais benefícios tendem a receber mais avaliações.

1. Avg\_salary e Interviews\_taken: Correlação de 0.95

* Há uma correlação positiva muito forte (0.95) entre o salário médio e o número de entrevistas realizadas pelas empresas. Isso pode indicar que empresas que oferecem salários mais altos têm uma demanda maior por seus cargos, resultando em um maior número de entrevistas.

1. Avg\_salary e Total\_benefits: Correlação de 0.93

* Existe uma correlação positiva forte (0.93) entre o salário médio e o número total de benefícios oferecidos pelas empresas. Isso sugere que empresas que oferecem salários mais altos também tendem a fornecer mais benefícios.

1. Interviews\_taken e Total\_benefits: Correlação de 0.96

Existe uma correlação positiva muito forte (0.96) entre o número de entrevistas realizadas e o número total de benefícios oferecidos pelas empresas. Isso sugere que empresas que conduzem mais entrevistas também tendem a oferecer mais benefícios.

Essas correlações sugerem uma forte associação entre o número total de revisões, o salário médio, o número de entrevistas realizadas e o total de benefícios oferecidos pelas empresas. É importante notar que correlação não implica causalidade, ou seja, apenas porque duas variáveis estão correlacionadas, não significa necessariamente que uma cause a outra.

Desafio 5: Quais são as empresas com as melhores e piores avaliações médias?

Com base na análise das avaliações médias das empresas os resultados obtidos foram esses:

Empresas com as melhores avaliações médias:

* Salaryfy - Avaliação média de 5.0
* Stirring Minds - Avaliação média de 5.0
* IBC Techno - Avaliação média de 5.0
* Royal Migration Solutions - Avaliação média de 5.0
* Lavender Technology - Avaliação média de 5.0

Empresas com as piores avaliações médias:

* TalentServe - Avaliação média de 2.1
* Artefact Projects - Avaliação média de 2.0
* Miric Biotech - Avaliação média de 1.9
* Ascentrik Research Services - Avaliação média de 1.6
* DM Consultants - Avaliação média de 1.3

Desafio 6: Existe uma diferença significativa entre o número total de avaliações para as empresas mais bem avaliadas e as menos avaliadas?

* Média do número total de avaliações para as empresas mais bem avaliadas: 259.0
* Média do número total de avaliações para as empresas menos avaliadas: 117.8

O resultado indica que a média do número total de avaliações para as empresas mais bem avaliadas é de aproximadamente 259, enquanto a média do número total de avaliações para as empresas menos avaliadas é de cerca de 117.8. Isso sugere que há uma diferença significativa no número total de avaliações entre as empresas mais bem avaliadas e as menos avaliadas. As empresas mais bem avaliadas tendem a ter um número maior de avaliações, indicando possivelmente uma maior visibilidade e interação por parte dos usuários em relação a essas empresas.

**Conclusão**

Esta análise fornece insights valiosos sobre as empresas listadas no Ambition Box, destacando suas características, pontos fortes e áreas para melhoria. Os resultados obtidos podem ser úteis para profissionais em busca de oportunidades de carreira, empresas em busca de informações sobre concorrentes e pesquisadores interessados em explorar tendências no mercado de trabalho.

**Contato**

Para mais informações ou colaborações, entre em contato:

* Email: victormatheus@alu.ufc.br
* LinkedIn: www.linkedin.com/in/victor-mattheus
* Github: https://github.com/VictorMatheus