

IMDb - Análise Exploratória de Dados (EDA)

O presente relatório tem como objetivo realizar uma análise exploratória dos dados cinematográficos. A análise visa entender as principais características dos filmes, explorar relações entre variáveis e extrair insights que possam orientar decisões estratégicas na produção de novos filmes.

O relatório EDA realizado no Notebook descreve todas as análises iniciais para o resumo deste documento.

Link Notebook:

https://github.com/Raul-Lemelle/lighthouse_desafio_ciencia_dados/blob/main/notebooks/eda.ipynb

Métodos Utilizados

Para realizar a análise exploratória dos dados, foram utilizadas técnicas estatísticas e visualizações gráficas.

- Visão Geral dos Dados: Database do tipo “csv”, conforme descrito abaixo:

desafio_indicium_imdb.csv

Colunas:

- Series_Title => Nome do filme
 - Released_Year => Ano de lançamento
 - Certificate => Classificação etária
 - Runtime => Tempo de duração
 - Genre => Gênero
 - IMDB_Rating => Nota do IMDB
 - Overview => Overview do filme
 - Meta_score => Média ponderada de todas as críticas
 - Director => Diretor
 - Star1 => Ator/atriz #1
 - Star2 => Ator/atriz #2
 - Star3 => Ator/atriz #3
 - Star4 => Ator/atriz #4
 - No_of_Votes => Número de votos
 - Gross => Faturamento
-
- A Análise Exploratória de Dados foi conduzida utilizando a linguagem de programação **Python**, utilizando bibliotecas como **Pandas**, **NumPy** e **Matplotlib** para análise e visualização dos dados.
 - Identificação de outliers e tratamento de dados ausentes quando necessário.



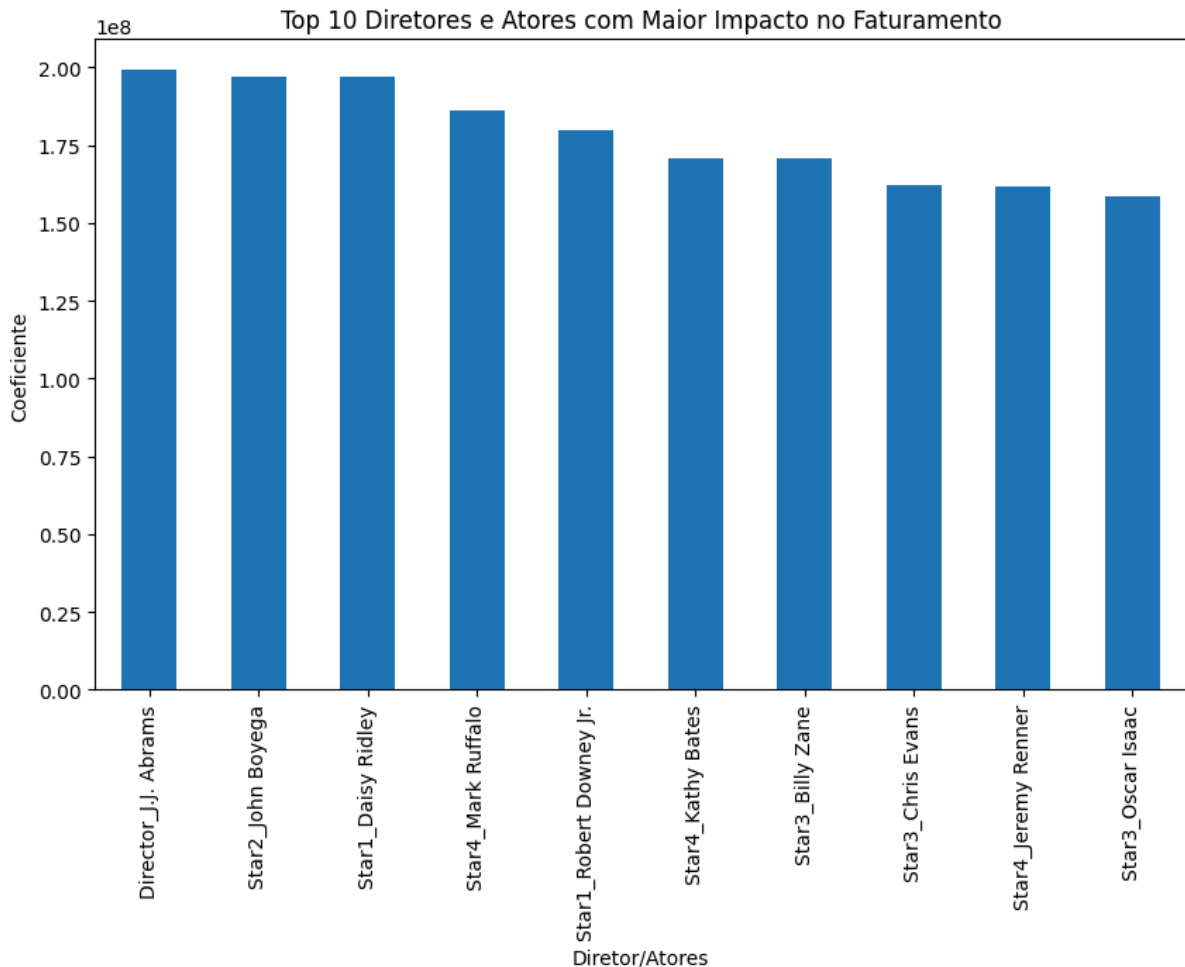
IMDb - Análise Exploratória de Dados (EDA)

- **Fatores de Faturamento:**

Link Notebook:

https://github.com/Raul-Lemelle/lighthouse_desafio_ciencia_dados/blob/main/notebooks/analysis_regressao_linear_gross.ipynb

Foi realizado um modelo de Regressão Linear para inferir dos coeficientes quais diretores e atores têm maior impacto no faturamento. Diretores e atores com coeficientes maiores têm um impacto mais significativo.



Os coeficientes listados mostram quanto cada diretor e ator específico contribui para o faturamento previsto do filme, mantendo todas as outras variáveis constantes. Alguns insights específicos incluem:

- Director_J.J. Abrams: Este diretor tem o maior coeficiente positivo, indicando que filmes dirigidos por J.J. Abrams tendem a ter um impacto significativo positivo no faturamento.
- Star2_John Boyega e Star1_Daisy Ridley: Os atores John Boyega e Daisy Ridley também têm coeficientes positivos elevados, sugerindo que filmes estrelados por eles estão associados a um aumento significativo no faturamento.

IMDb - Análise Exploratória de Dados (EDA)

- Star4_Mark Ruffalo, Star1_Robert Downey Jr., Star4_Kathy Bates, Star3_Billy Zane, Star3_Chris Evans, Star4_Jeremy Renner, Star3_Oscar Isaac: Outros atores também mostram coeficientes positivos, indicando que suas presenças nos filmes estão correlacionadas com um faturamento mais alto.

- **Previsão da Nota do IMDb**

Relatório Previsão da Nota do IMDb

Link:

https://github.com/Raul-Lemelle/lighthouse_desafio_ciencia_dados/blob/main/reports/relatorio_previsao_nota_imdb.pdf

IMDb - Relatório Final das Análises

Link:

https://github.com/Raul-Lemelle/lighthouse_desafio_ciencia_dados/blob/main/reports/imdb_relatorio_final_das_analises.pdf