MOVIE INDUSTRY ANALYSIS

By Nancy Maina

OBJECTIVES

➢ **Project Overview**

➢ **Business understanding**

➢ **Data understanding**

➢ **Data Analysis & Visualizations**

➢ **Recommendation & Conclusion**

PROJECT OVERVIEW

- This organization had decided to join in the fun by creating new movie studio

- This project explores the various kinds of films and their performance at the box office.

- Findings translated into actionable insights

BUSINESS UNDERSTANDING

- **Objective:** Establish a new movie studio, determine genres of films to produce based on current box office trends.

- **Business Goals:** Maximize box office revenues, establish brand recognition, and cater to diverse audience preferences.

- **Key Questions:**

a) Which genres are currently performing well at the box office?

b) How do factors like budget and directors impact box office success?

c) Which studios will we be competing with?
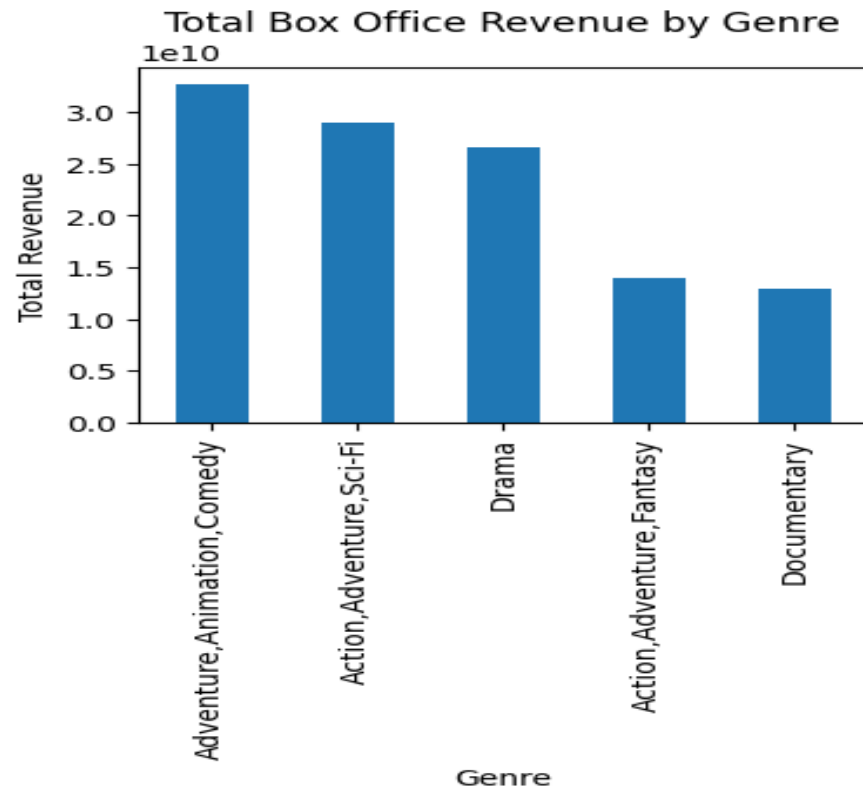
DATA UNDERSTANDING

Data Sources:

- IMDB SQLite Database (im.db): Contains movie metadata including titles, genres, ratings. Box Office Mojo (bom.movie_gross.csv.gz): Provides box office gross revenue data.

- Explore data to understand distributions, correlations, and completeness.

- Identify relevant variables such as genre, budget, ratings, release dates.

DATA ANALYSIS PROCESS

- **Relevant libraries imported:** for data analysis including Pandas, Matplotlib, and Seaborn, Numpy and SciPy for easy data manipulation.

- **Selected the relevant dataset and data frames**: This project will use IMDB database as it has large datasets for analysis(more than 150,000 movies) using three primary columns selected including genres, primary_name(directors) and primary title. The project will also use the bom.movie_gross.csv data sets to calculate the box office revenue per genre.

- **Data Cleaning:** There were a few columns dropped entirely since the majority of the data was missing, and for the purpose of this project, they were basically useless.

- **Merging Data Frames:** Using the Pandas merge method, several data frames were merged based on either similar column values

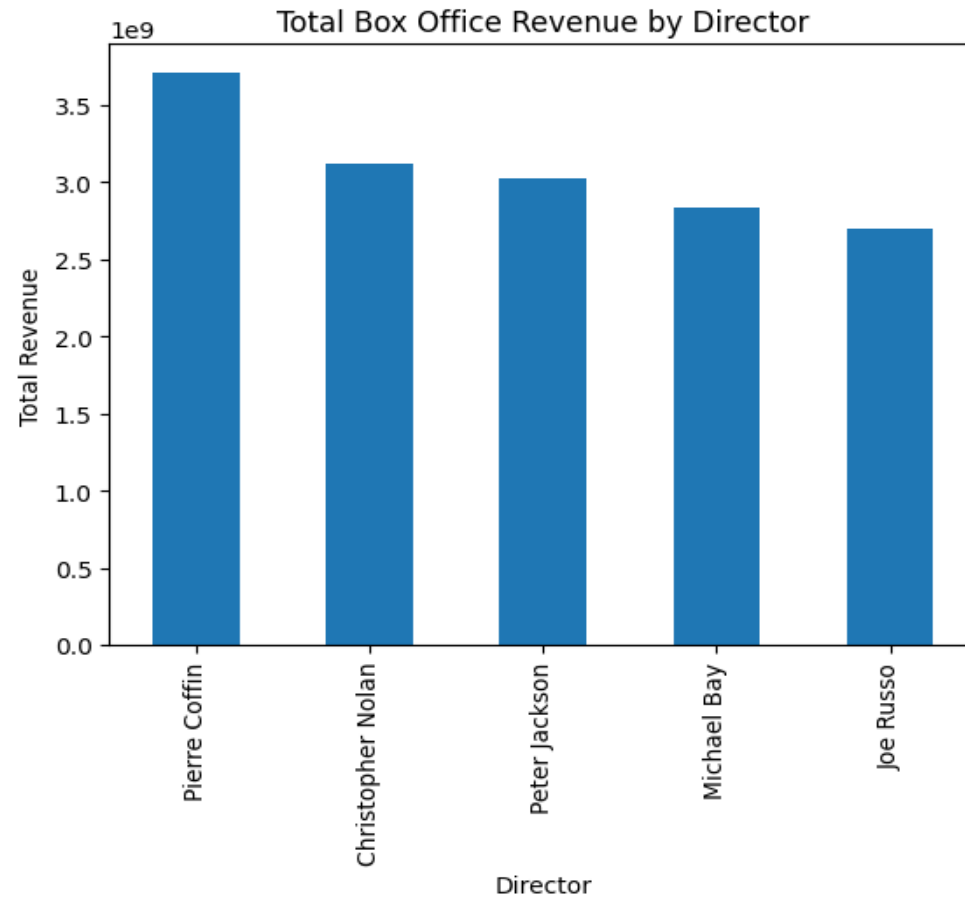- **Visualizing the data:** Used to answer the three questions under business understanding.

DATA ANALYSIS & VISUALIZATIONS: BEST PERFORMING GENRE

Total Box Office Revenue by Genre

Top 5 Genres performing well currently based on revenue generated by Genre
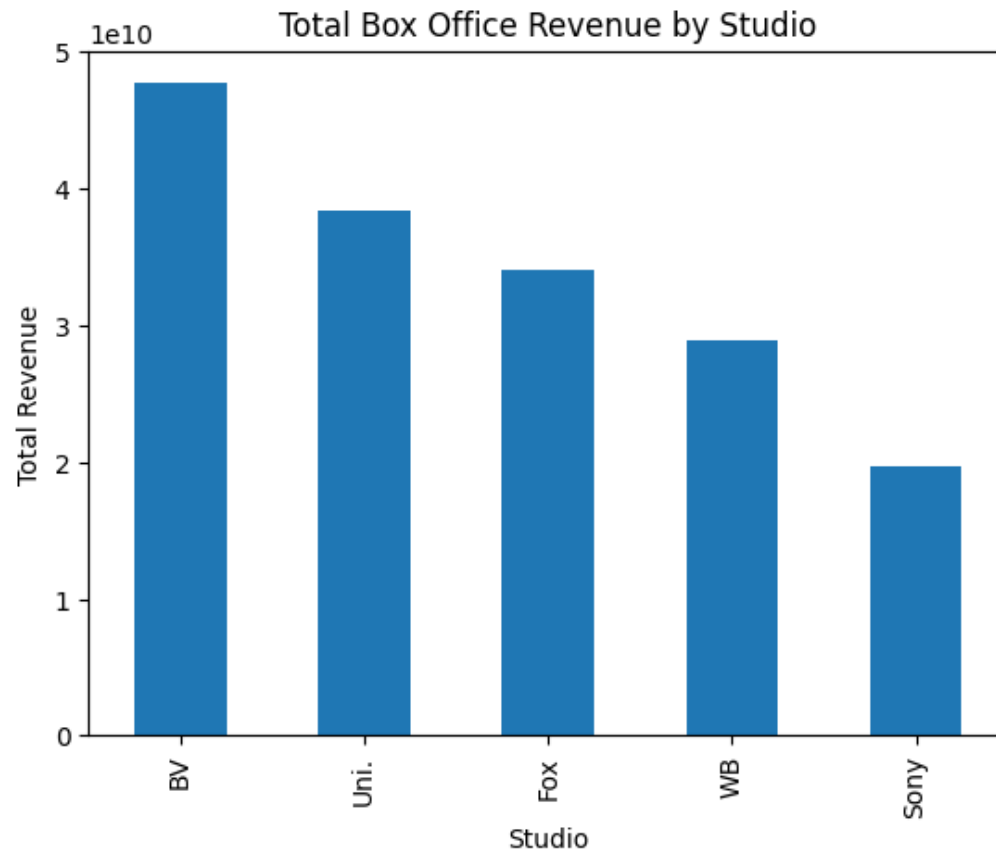
1. Adventure
2. Action
3. Animation
4. Comedy
5. Drama

DATA ANALYSIS & VISUALIZATIONS: BEST PERFORMING DIRECTORS

Total Box Office Revenue by Director

- **Top 5 Directors performing well currently based on revenue generated in their films**

1. **Pierre Coffin**
2. **Christopher Nolan**
3. **Peter Jackson**
4. **Michael Bay**
5. **Joe Russo**

DATA ANALYSIS & VISUALIZATIONS: BEST PERFORMING STUDIO

CONCLUSION & RECOMMENDATIONS

❑ **Focus on producing Action and Adventure movies:** These genres consistently generate the highest box office revenues.

❑ **Collaborate with successful directors:** Partnering with experienced top 5 directors like **Pierre Coffin, Christopher Nolan, Peter Jackson, Michael Bay, and Joe Russo** can significantly increase the chances of success.

❑ **Consider benchmarking top 5 studios:** The following studios tend to perform exceptionally well including **BV, Uni., Fox, WB, and Sony,** to learn ways they strategically can maximize revenue potential.

REFERENCES

- https://www.boxofficemojo.com), IMDB (https://www.imdb.com/interfaces/),
- The Movie Database (https://www.themoviedb.org/?language=en-US).