

Moringa School Data Science part Time project

Prepared By :Purity Riungu

Date; 16/04/2023

Project Overview;

Project Title ;Microsoft Movie studio

- Project Microsoft Movie Studio was undertaken to conduct data analysis and advice the management of Microsoft of the best Movie studio to venture in.
- The goal of the project was to deliver a high-quality, user-friendly application that would improve our client's productivity and efficiency. To achieve this, the project was broken down into several key tasks, including requirements gathering, design, development, testing, and deployment.
- The project scope was defined in the project scope statement below, which included a detailed description of the project's objectives. The scope was managed throughout the project lifecycle using project management tools and techniques, such as work breakdown structures, data analysis visualization and project status reports. This helped to ensure that the project stayed on track, and any changes to the scope were carefully evaluated.

The scope of the Project;

The project was carried out as per the scope below

Business Understanding

Data understanding and Preparation

Data Analysis

Data Cleaning

Data Visualization

Conclusion and Reporting to management

Business Understanding



The Business Questions that will be tackled in the project will be;

- What are the most successful genres of movies?
- What are the key characteristics of successful movies? For instance the enthralling storyline, the Characters/cast.
- How can Microsoft new movie studio position it self to be come the best?

Data understanding and Preparation

In this project, I will work with various datasets which was collected from various locations. Some data are compressed CSV,TSV and IMDB. The data to be used is contained in files below;

- im.db.zip
- bom.movie_gross.csv.gz

The data that was extracted and used in the project was;

- Bom.movie_gross.csv.gz
- Rt.movie_info.tsv.gz
- Rt.reviews.tsv.gz
- Tmdb.movies.csv.gz
- tn.movie_budget.csv.gz

Data Analysis and cleaning

- The first step was to read the data from various files that were provided. For instance
- Bom.movie_gross.csv.gz
- Rt.movie_info.tsv.gz
- Rt.reviews.tsv.gz
- Tmdb.movies.csv.gz
- tn.movie_budget.csv.gz

After reading the data, analysis of the data was conducted this was through;

1. Data cleaning- This was done with the aim of achieving;

- The accurate data to use

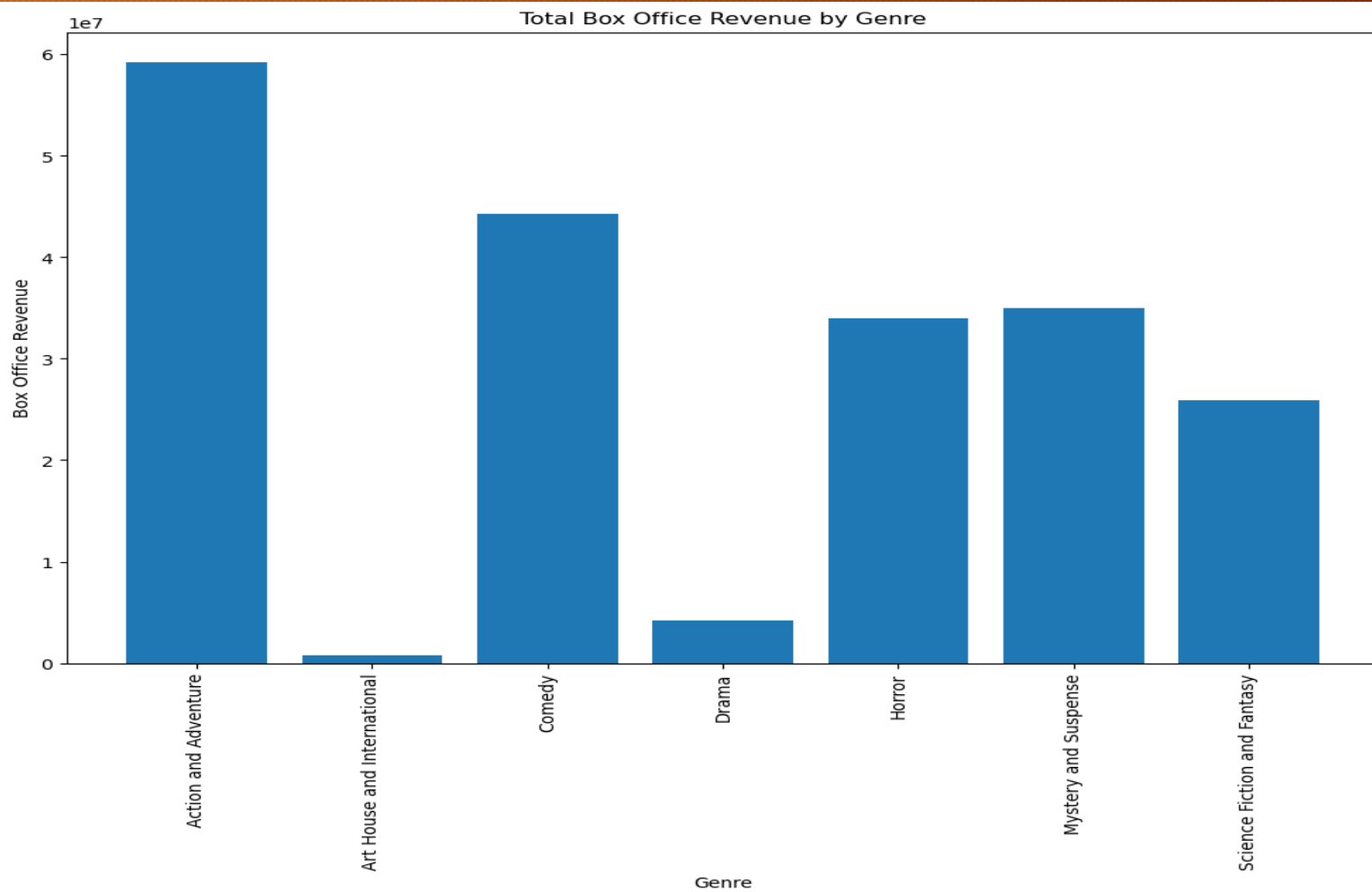
- Increased efficiency

- Better Data quality

Data Analysis and Cleaning

2. Merging different data sets with common columns that are important in the project.
3. Removing duplicates, removing columns that were not necessary the quality data for use.

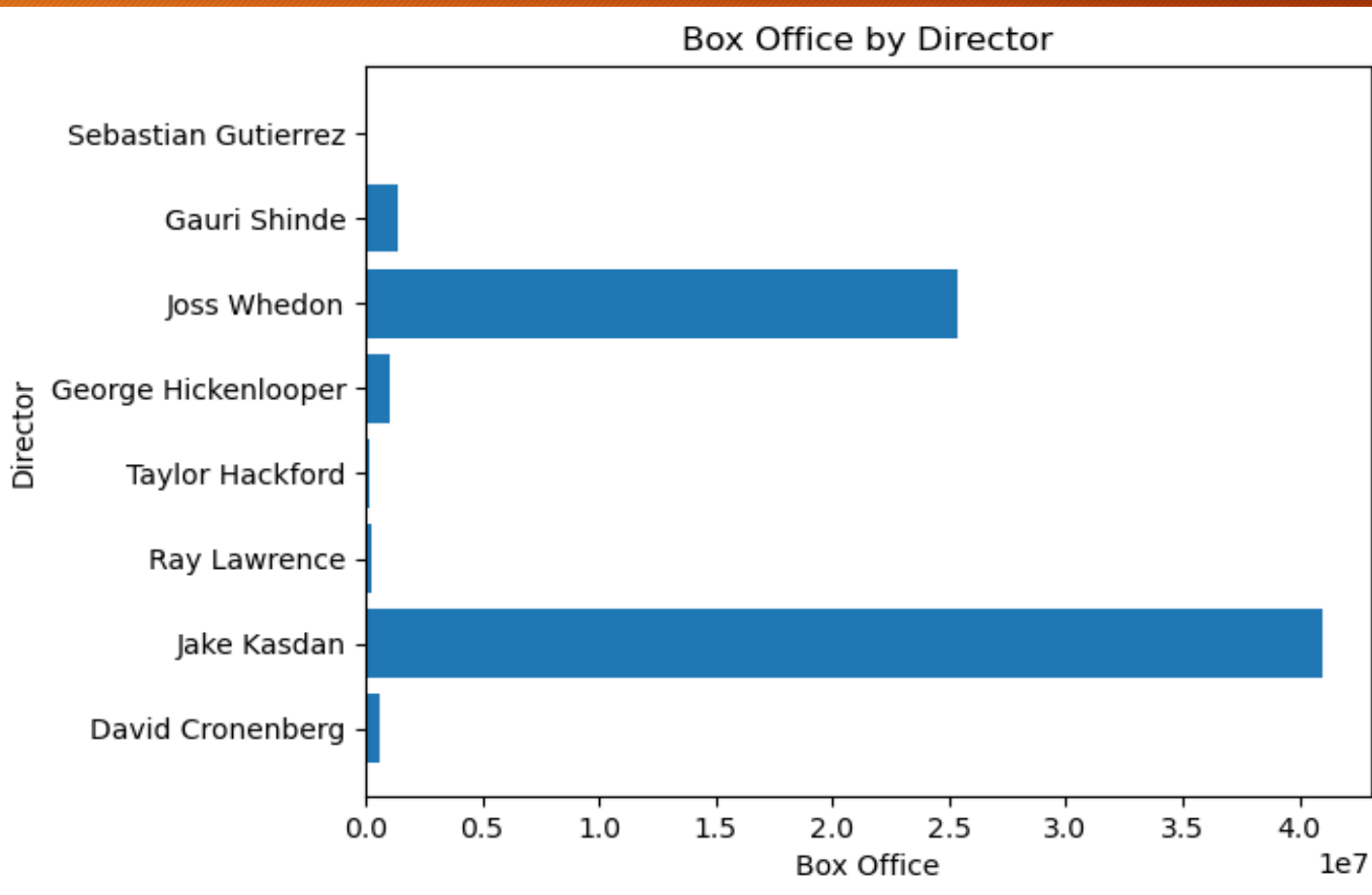
Data Visualization



The Graph below was used in describing the sales of different genres of movies. The Action and Adventure movies made the highest sales for the period under review. The Comedy genres were the second best selling genre.

Therefore Microsoft should consider venturing in the genres which made more sales to maximize their returns.

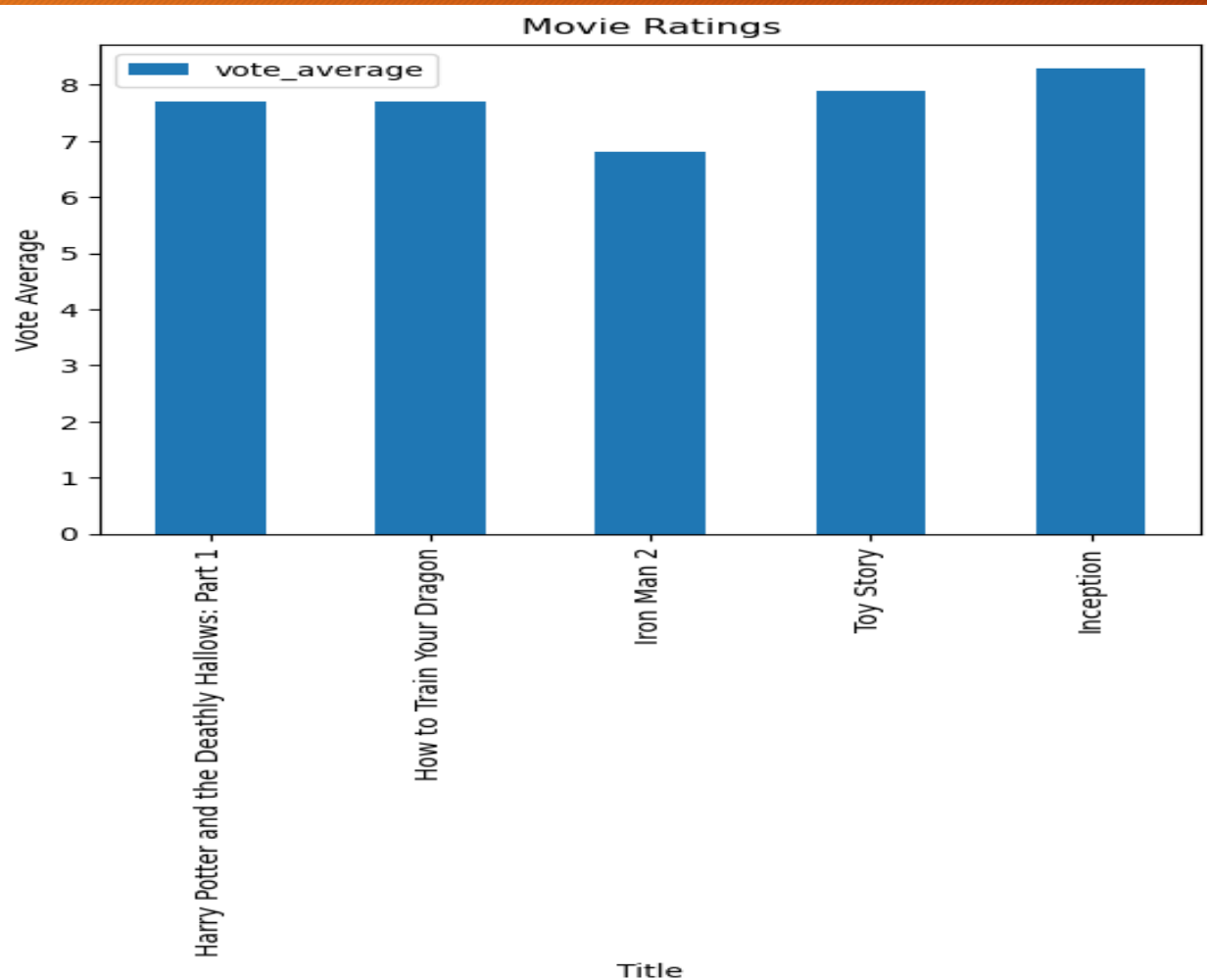
Data Visualization; Cont



From the data analysis, there was an indication of the various directors and how the sales made from the movies they directed. One of the factors to consider is the capacity and the skills of the directors who will direct the movies.

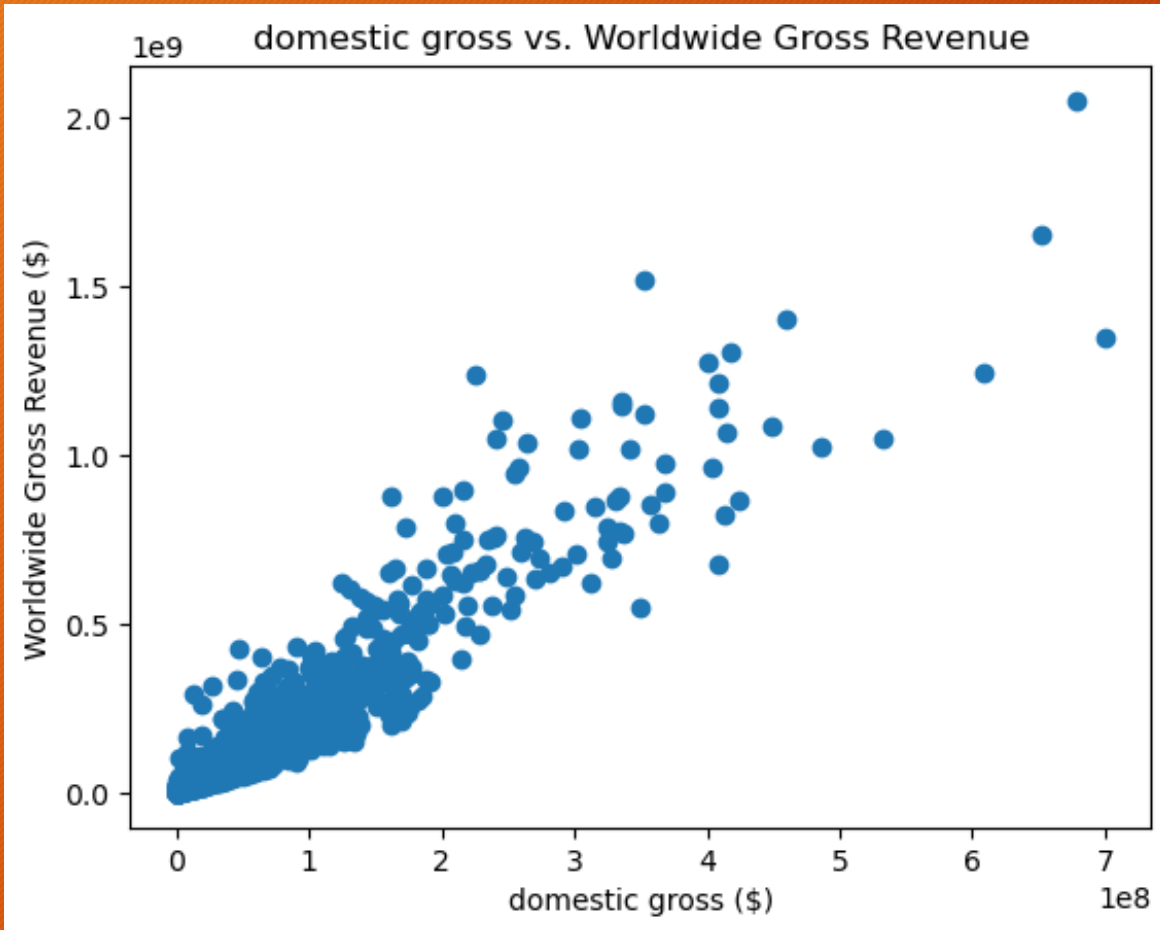
From the sample taken, movies directed by Jake Kasdan and Joss Whedon made the most sales. Therefore Microsoft should consider the directors.

Data Visualization; Cont



The Rating for different movies in the industry should be a factor to consider by Microsoft.

Data Visualization; Cont



The graph below was used to show the relationship between the Domestic and worldwide gross revenues. The two have a positive correlation, this means the movies can perform well both locally and internationally

Conclusion.

In quest of answering the questions below;

- What are the most successful genres of movies?
- What are the key characteristics of successful movies? For instance the enthralling storyline, the Characters/cast.
- How can Microsoft new movie studio position it self to be come the best?

We advice the management to consider the factors below in deciding on the movie shop;

1. The Genres of the movie. The management should consider historical trends of genres which made more sales to maximize their returns
2. The director of the movies they will be producing
3. The audience of the movies
4. The reception of the previous genres across the globe.

Recommendations;

I Recommend that the management should consider the factors below as they venture in the movie shop

1. The Genres of the movie.
2. The Director of the movies they will be producing
3. The audience of the movies
4. The reception of the previous genres across the globe.
5. The profitability of different genres.