

MOVIE DATA ANALYSIS

BRIAN KIMATHI M

machingabrian@gmail.com

The following libraries were used in the analysis of the data and also in the data visualizations

```
import sqlite3
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

Data

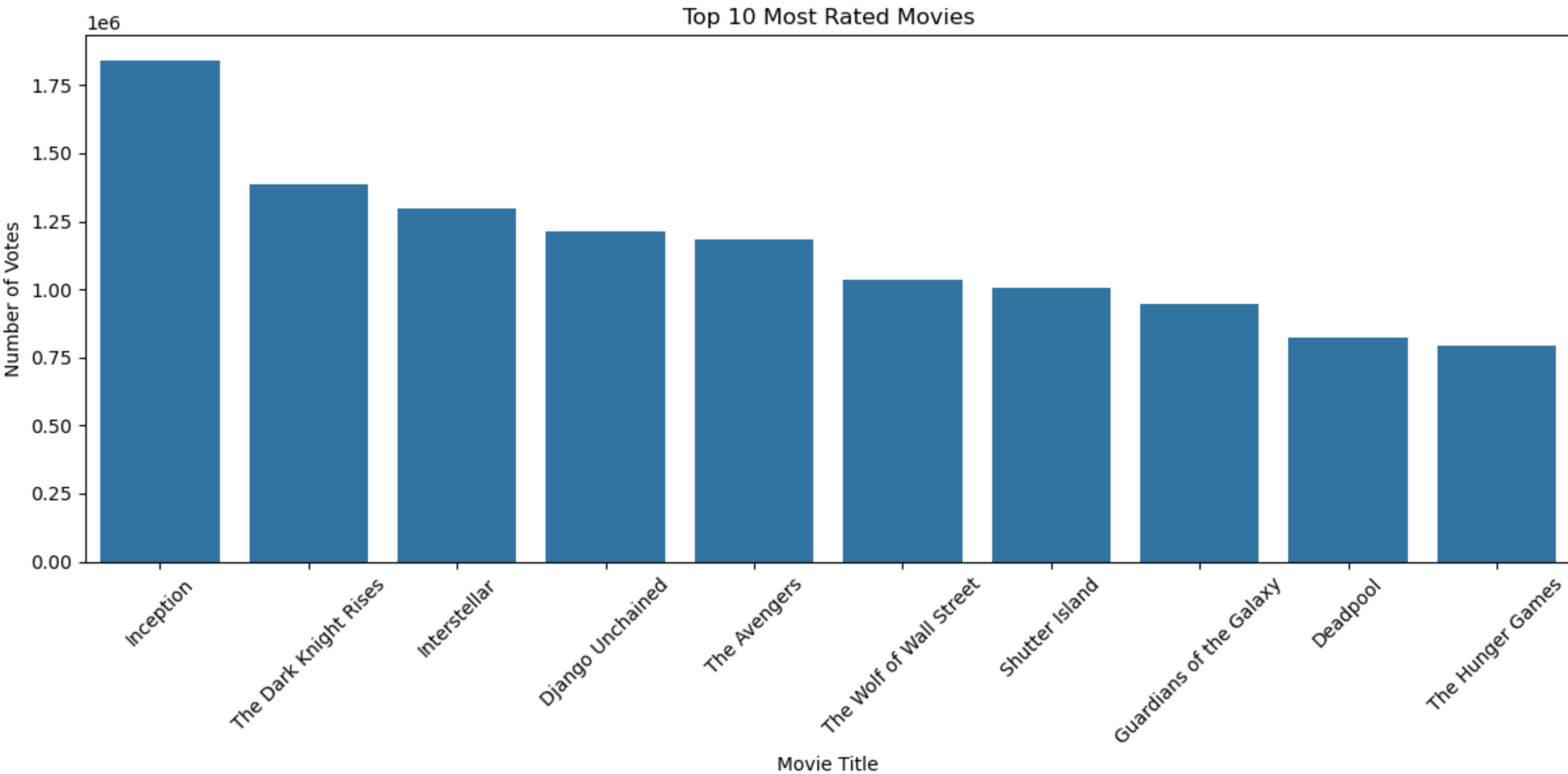
- This project implored the use of ;
- Movie_basics dataset
- Movie_ratings dataset
- Numbers dataset

Data Cleaning and Data exploration

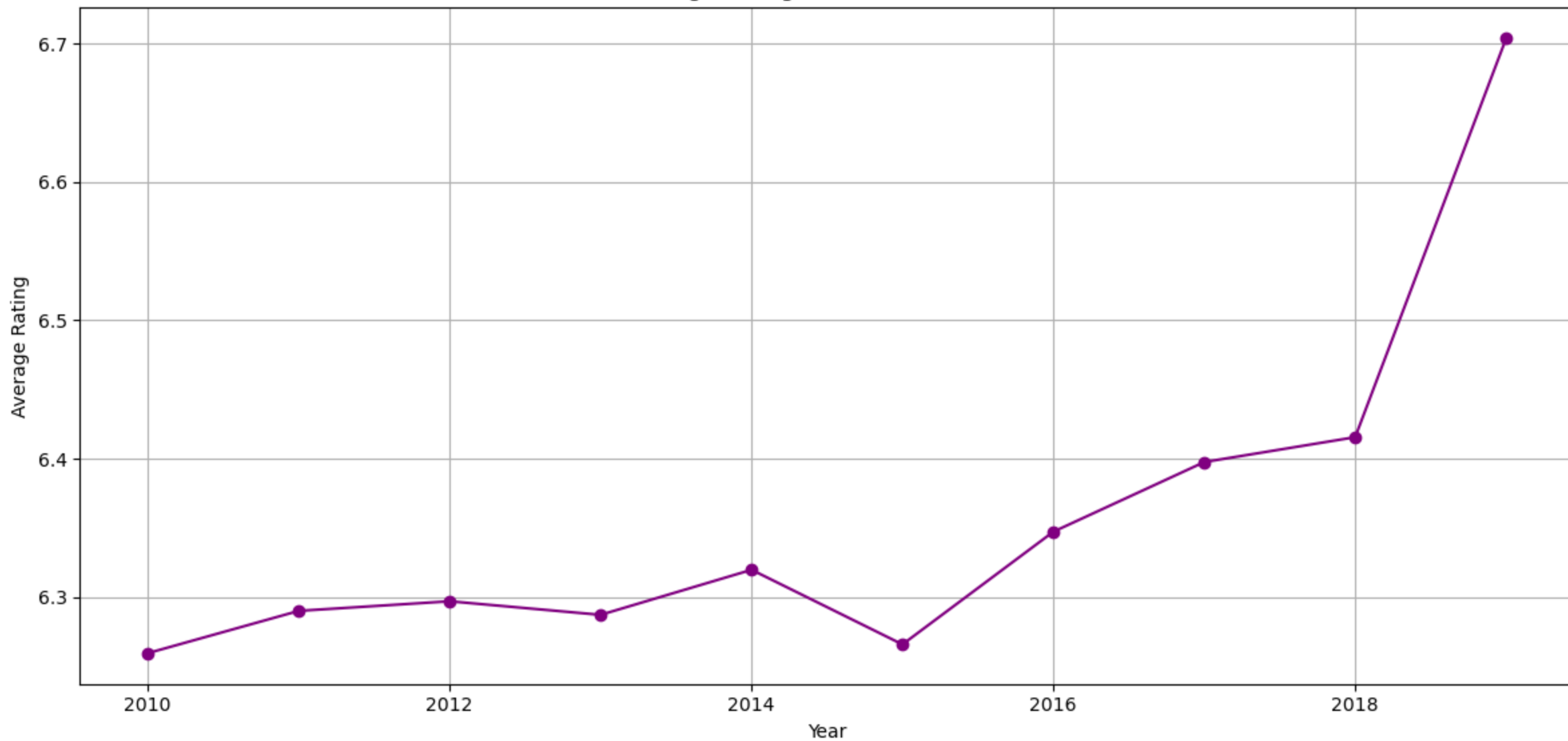
- The 3 data sets underwent data cleaning process;
- Checking and dropping of null values
- Checking and dropping of duplicates(none were found)
- Data standardization(making all column names lowercase)
- Data standardization(stripping of any whitespaces)

Visualizations

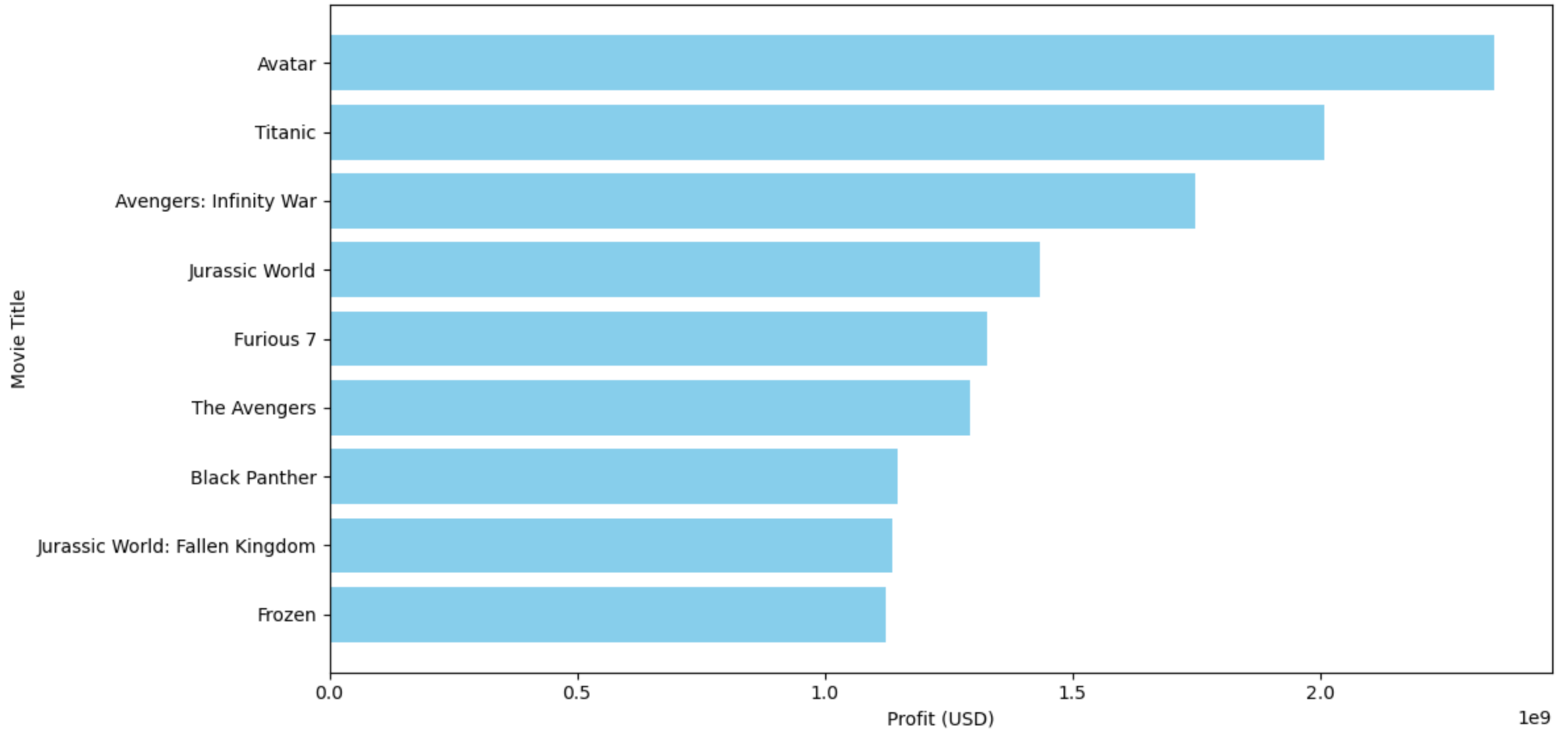
- After data was cleaned and ready to use I created visualization to help better come up with answers for this movie project
- The following viaualization were the results;

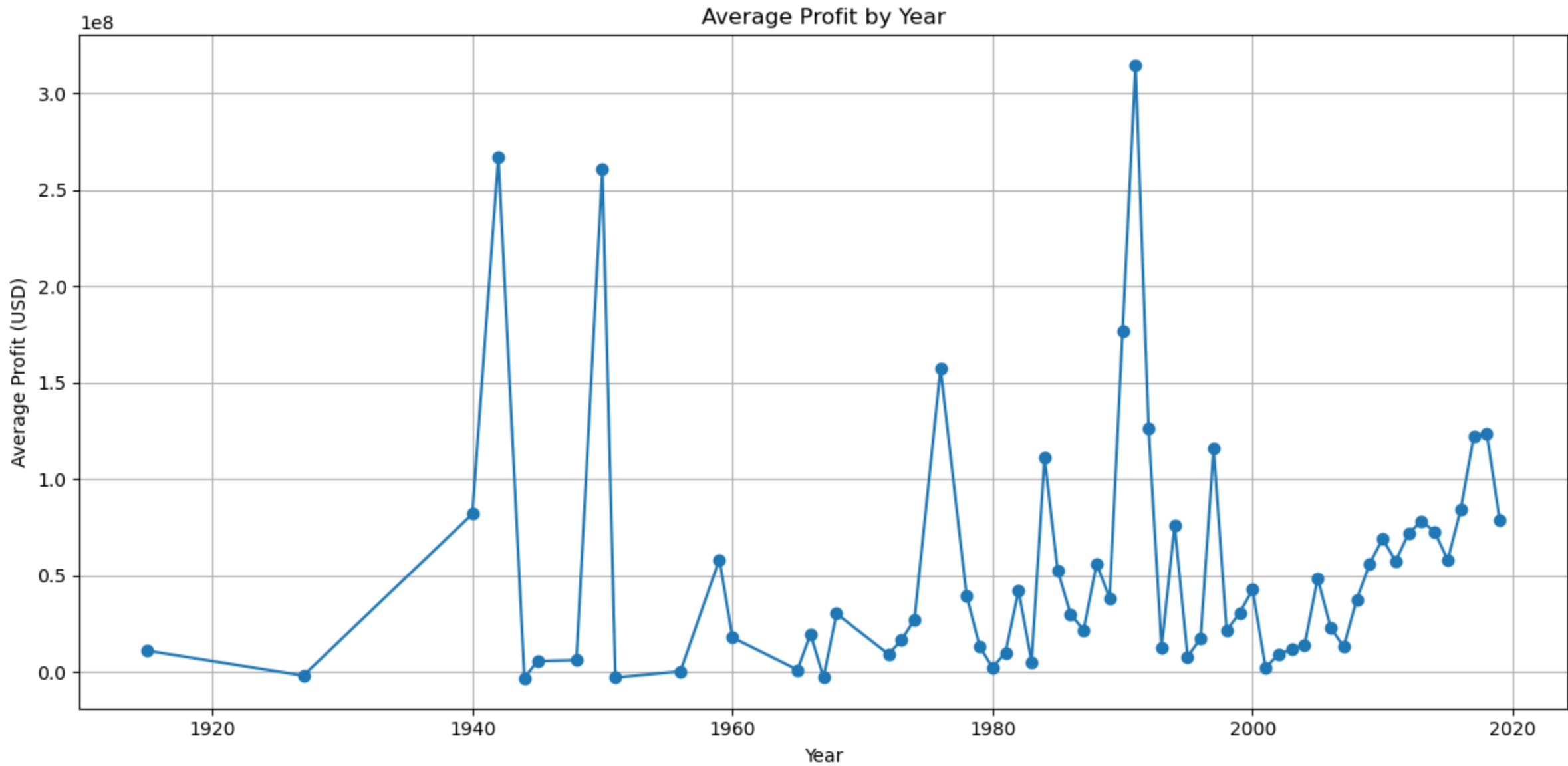


Average Rating of Movies Over Time

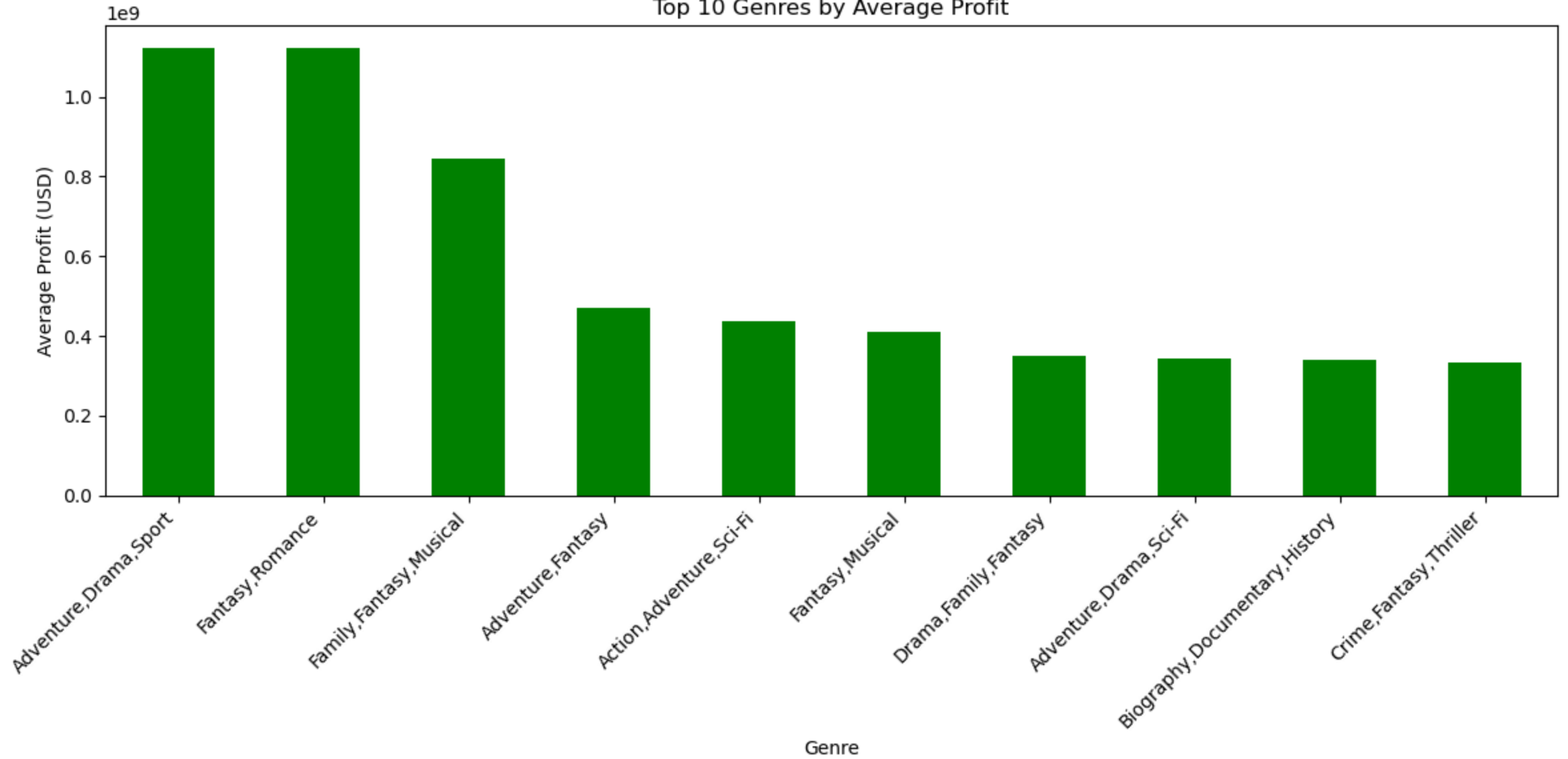


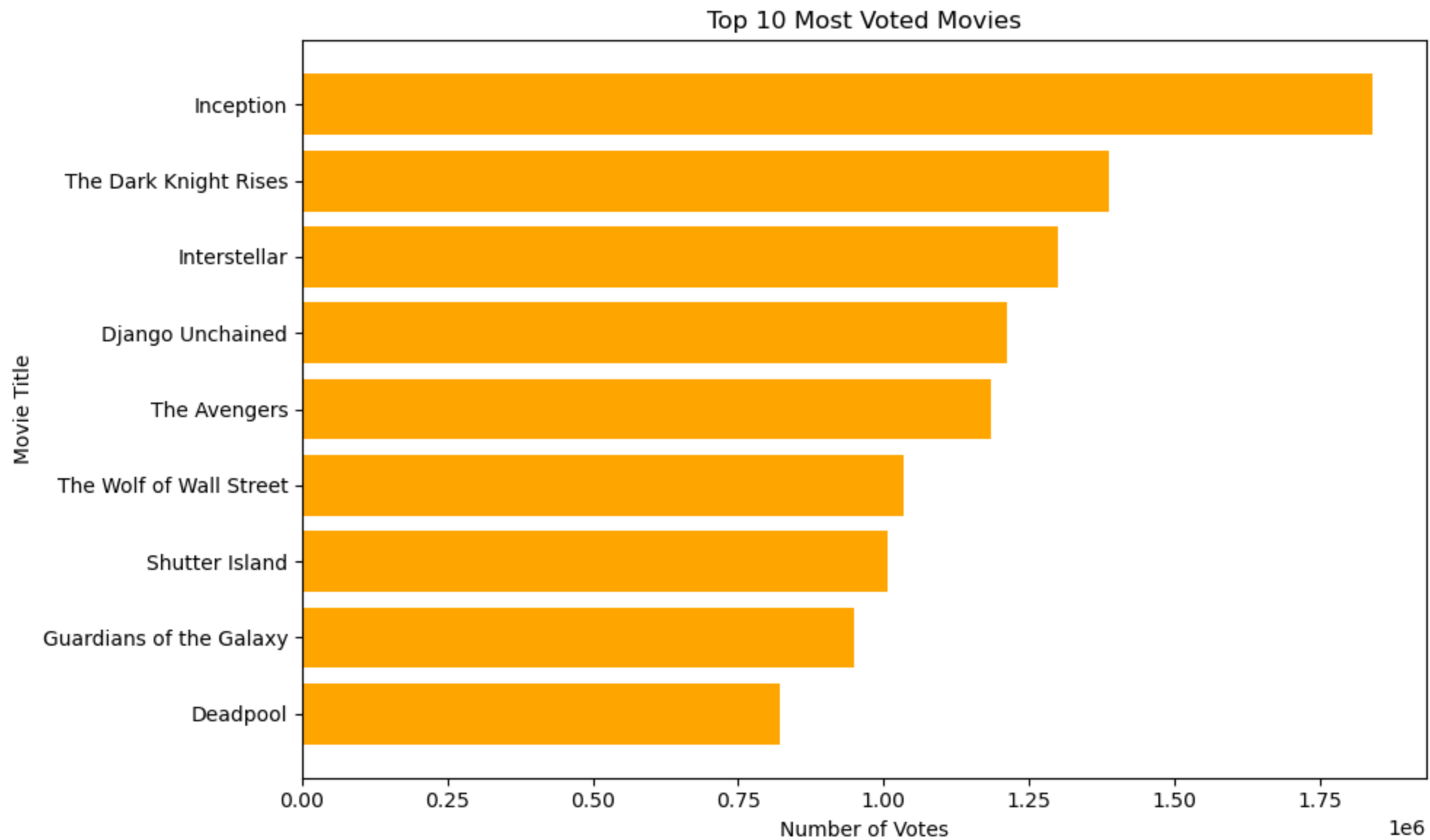
Top 10 Most Profitable Movies





Top 10 Genres by Average Profit





Business recommendations

- 1. Focus your studio on creating original films with the built-in potential to become the next major sci-fi or fantasy franchise.
- 2. Focus on creating a diversified portfolio of films,(i.e a balance of lower-budget, lower-risk projects with a few high-concept blockbusters) to lower/reduce volatility while still pursuing high-profit opportunities.
- 3. Focus on producing films that combine multiple genres, specifically targeting high-profit blends such as "Adventure, Drama, Sport" or "Fantasy, Romance," to maximize your average return
- **other recommendations**
 - 1. Focus on producing unique, thought-provoking films that are designed to be talked about and debated.(it helps to have a highly engaged audience.)
 - 2. Prioritize investing in high-quality storytelling and production.(i.e films that receive strong critical and audience ratings.)

