Introduction to the Movie Dataset

This movie dataset consists of **45,000 movies** released on or before July 2017. It is a **sample** from *TMDb & GroupLens* reported movies. **Source:** *Kaggle*

We have a set of visualizations about this data in 8 slides. The dataset includes several key attributes that are essential for analysis:

- **Genres**: Categorizes films into various genres like action, drama, comedy, etc.
- **Release Years**: Tracks when each movie was released, allowing for trend analysis over time.
- **Box Office Earnings**: Provides financial data to assess the commercial success of films.

These features are crucial for identifying trends and making informed conclusions for our study.

Today...

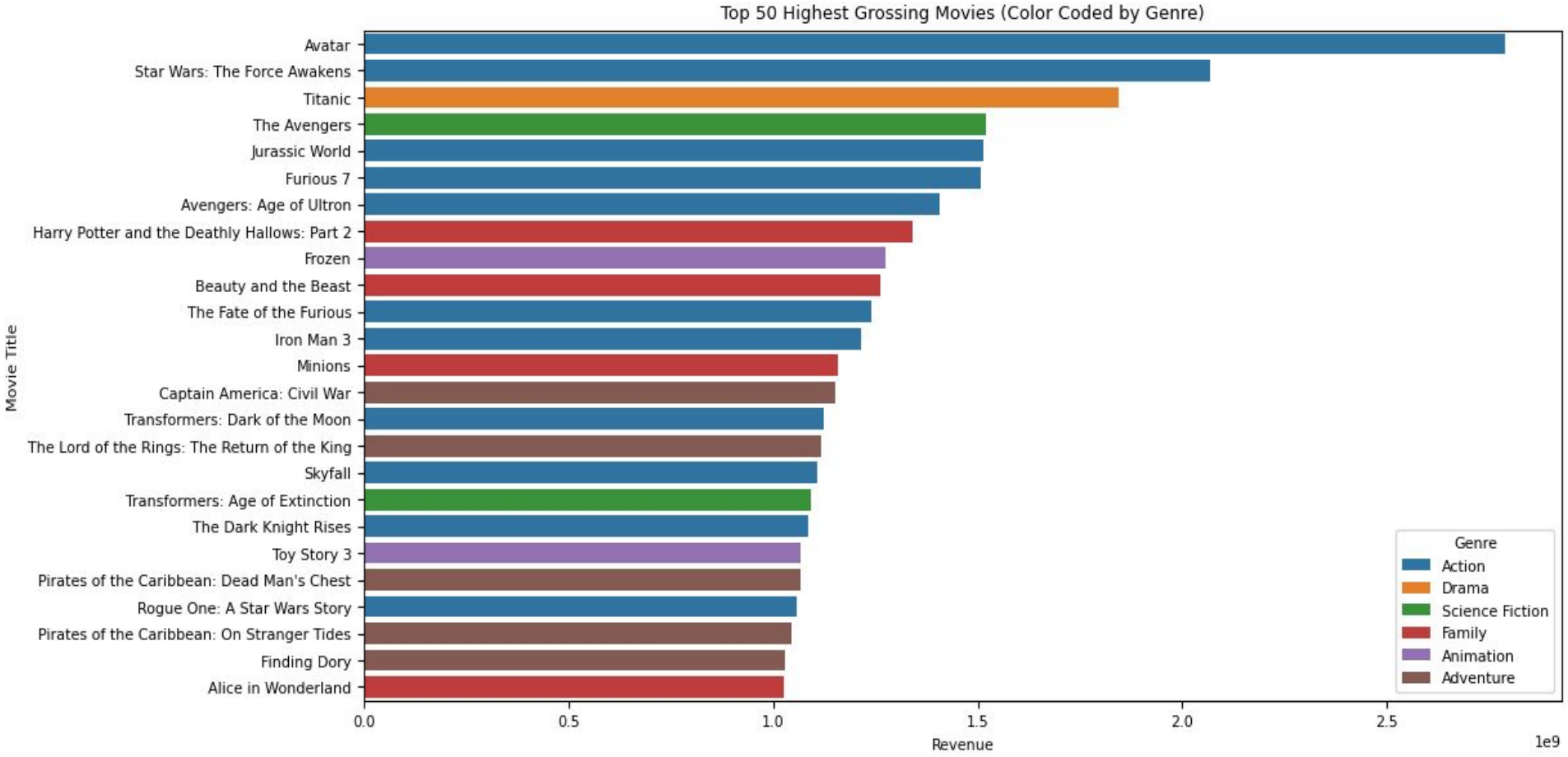
We want you to identify interesting trends and patterns in movies using visualizations of this dataset.

We want you to discuss:

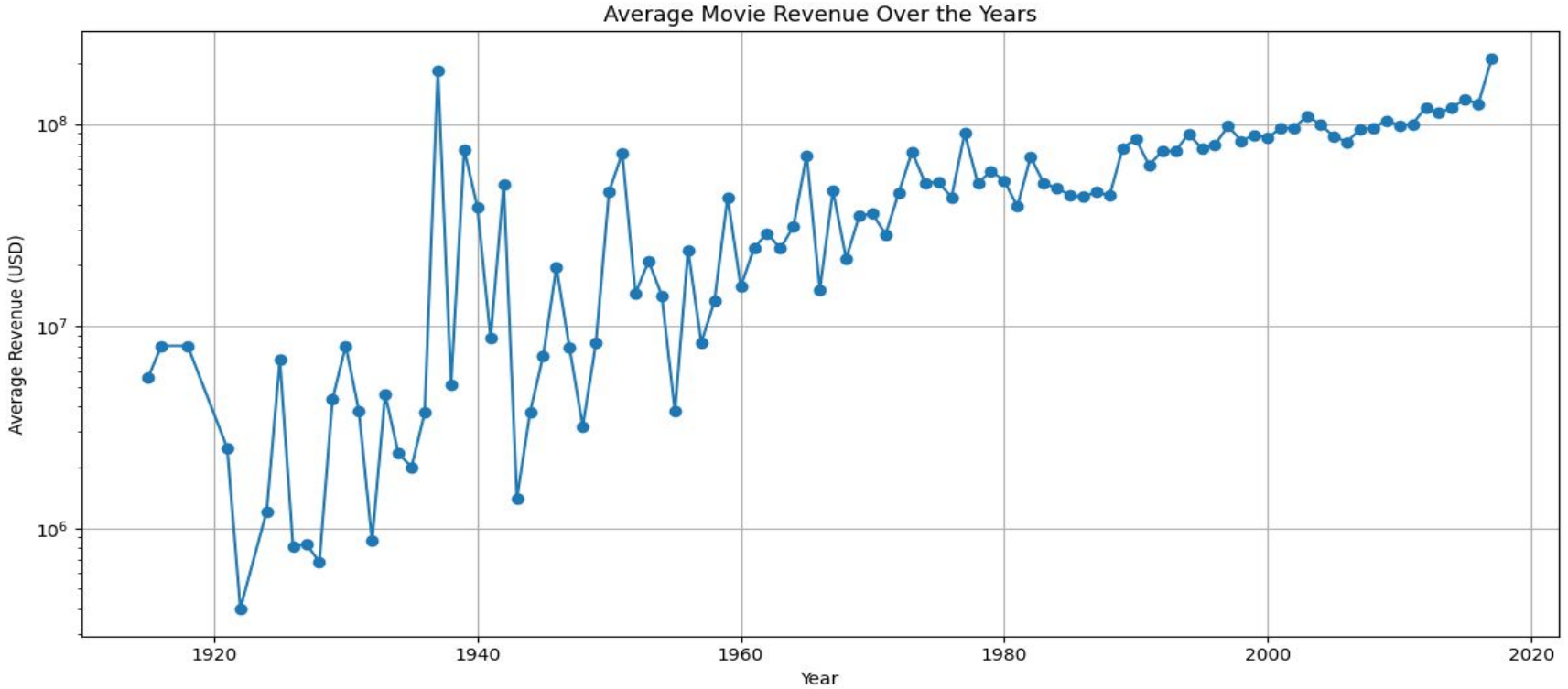
- What you think visualization shows,
- What you learn from this visualization about the movies,
- What you think the visualization means, and
- What are the things that surprise you about these visualizations?

At the end, we will ask you what you learned from these visualizations and if they were helpful in understanding the data about movies.

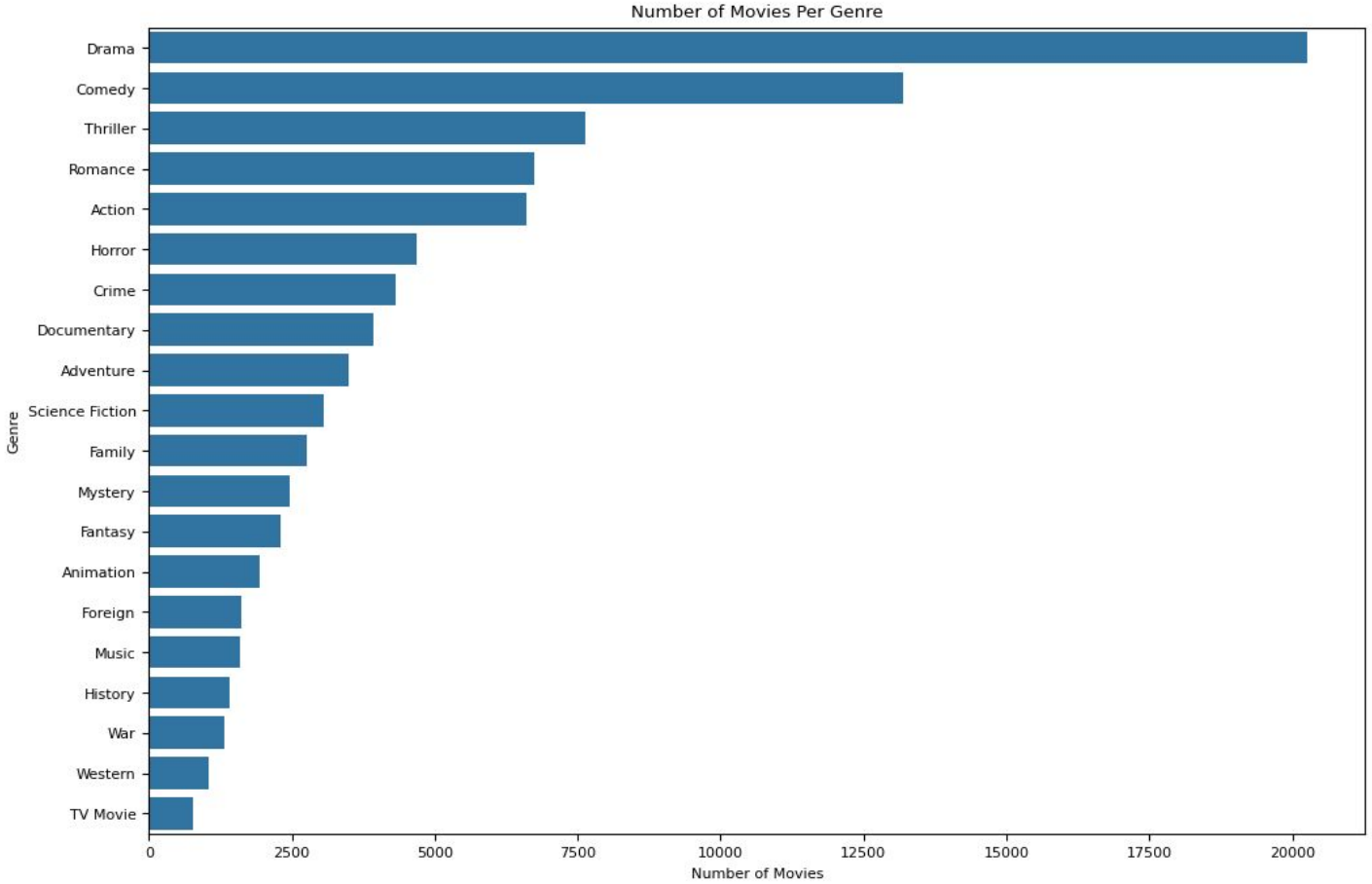
Highest Grossing Movies



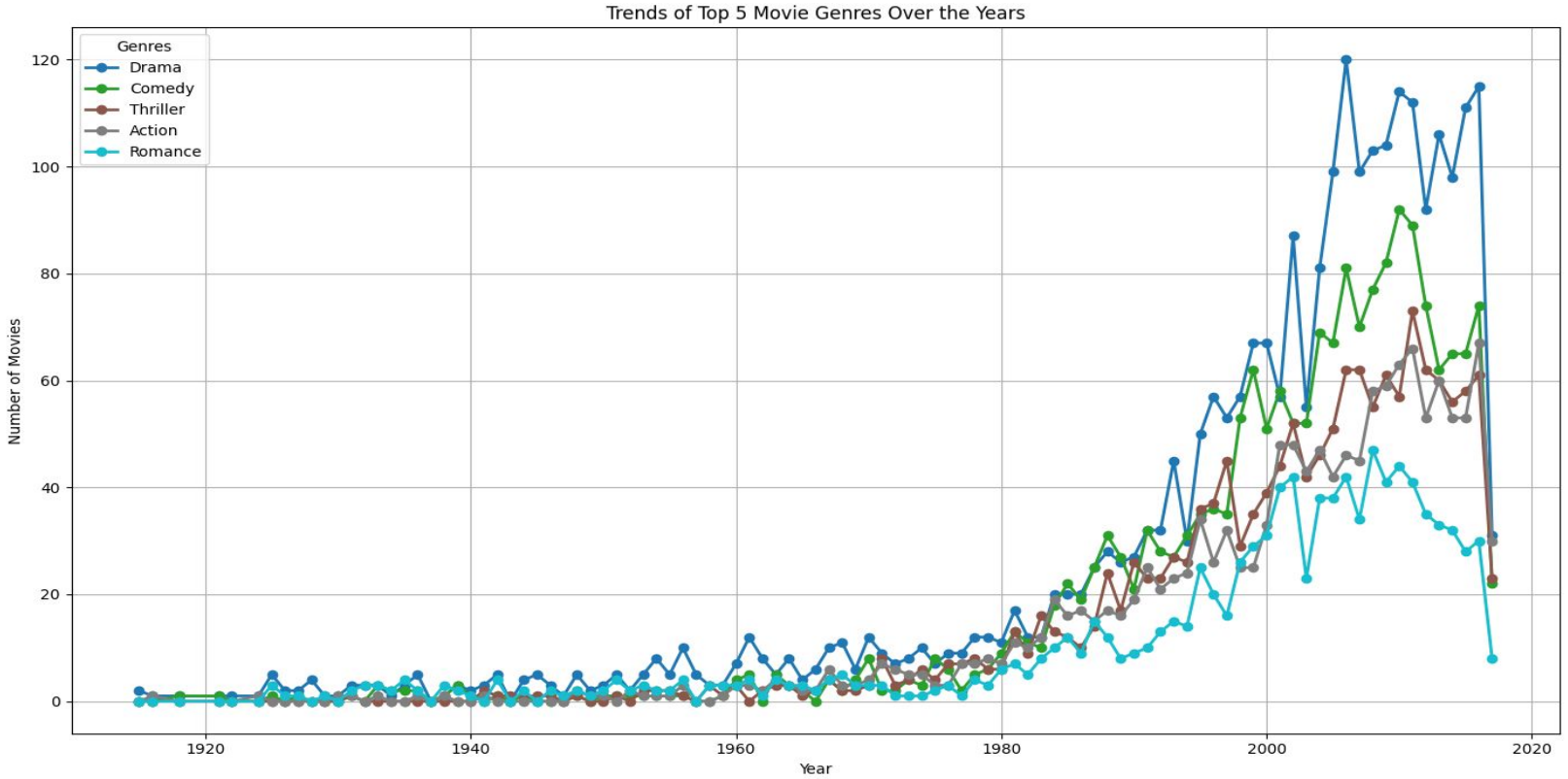
Average Movie Revenue Over The Years



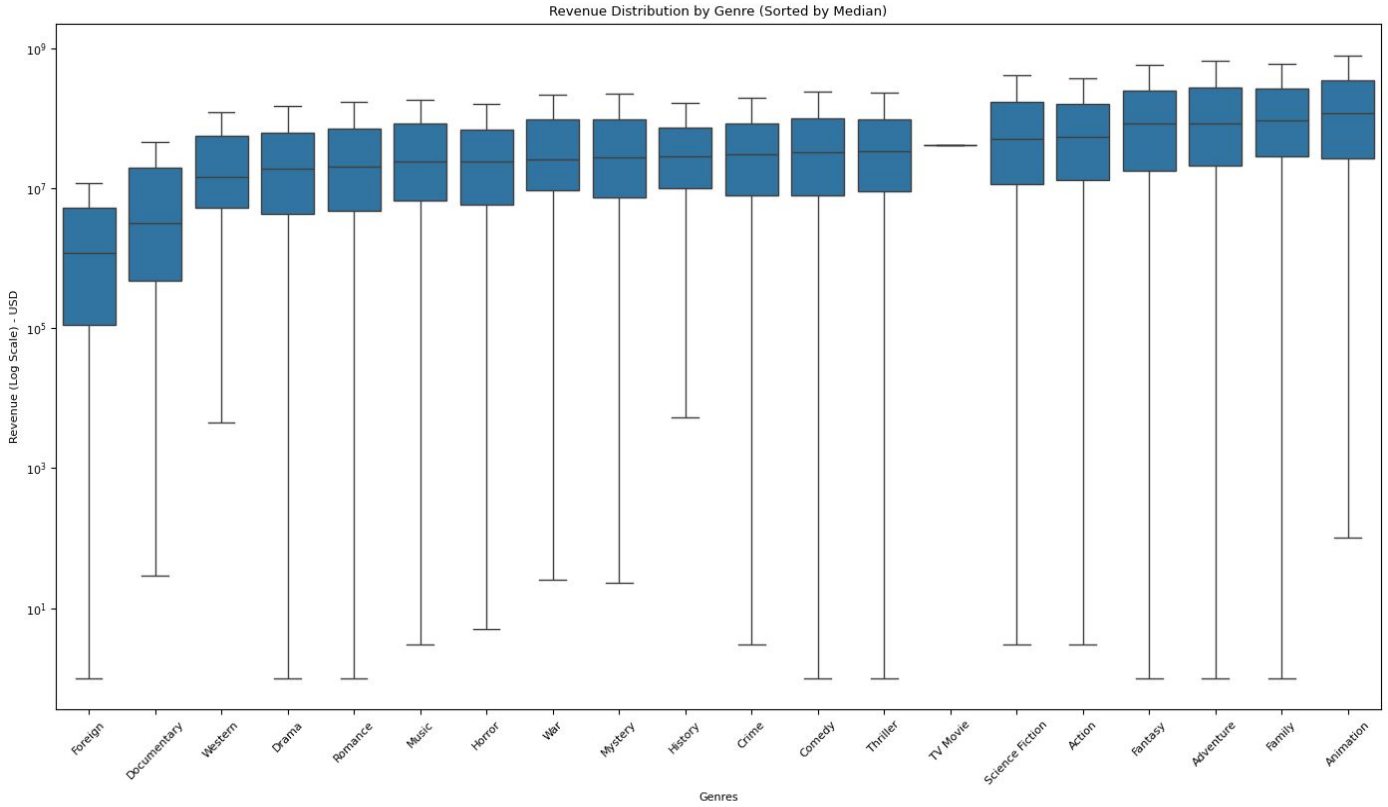
Number Of Movies Per Genre

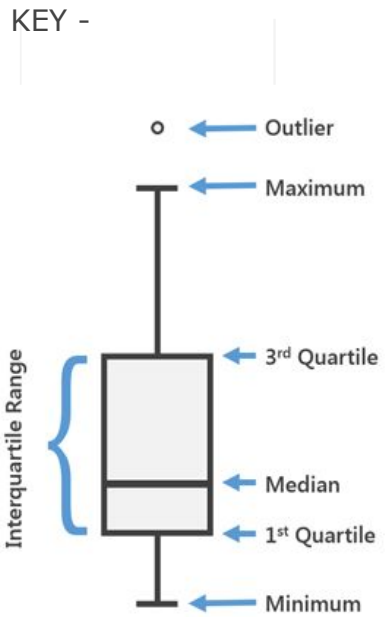


Top Five Movie Genres Over The Years

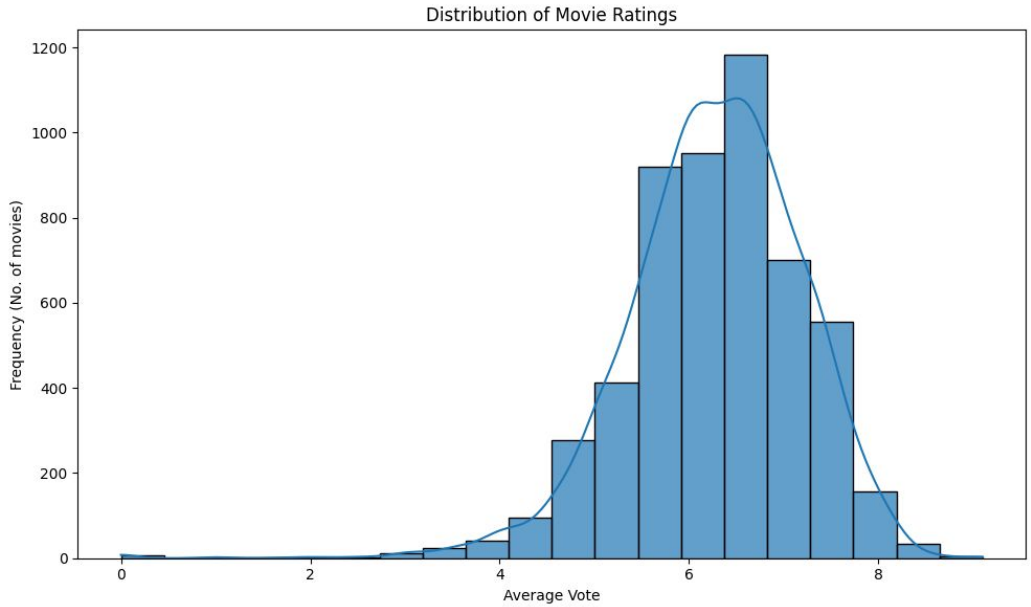


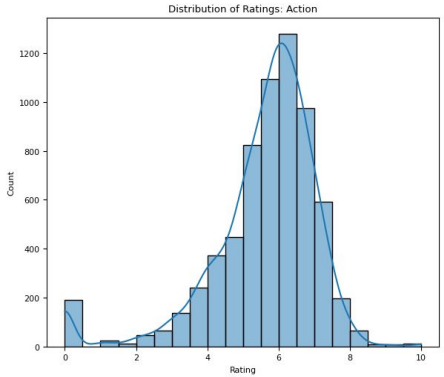
Budget by Genre (Box plot)

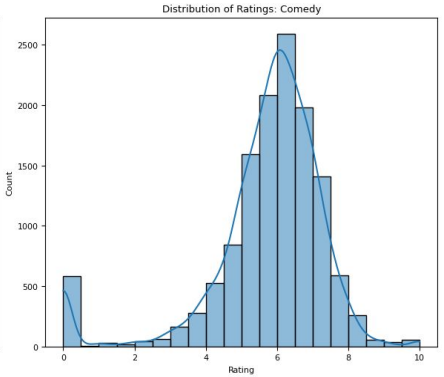


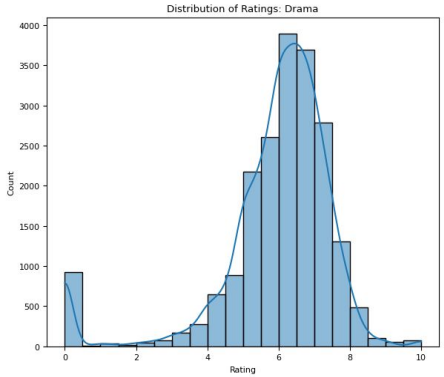


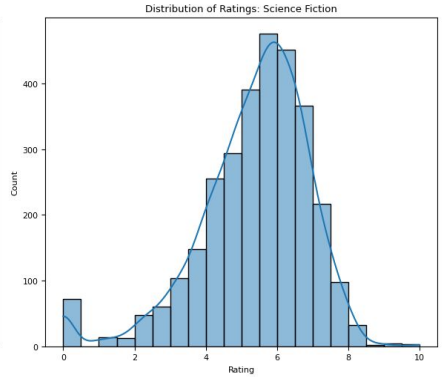
Distribution Of Movie Ratings



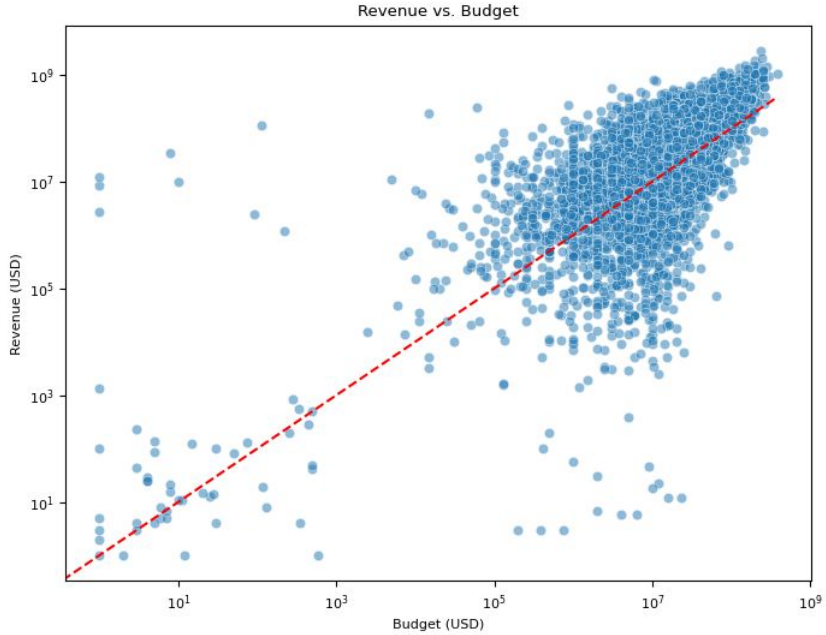


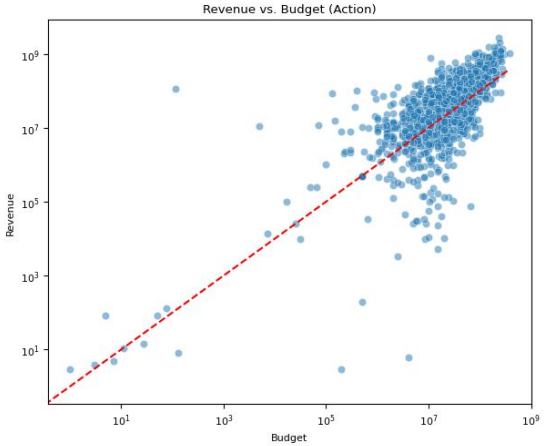


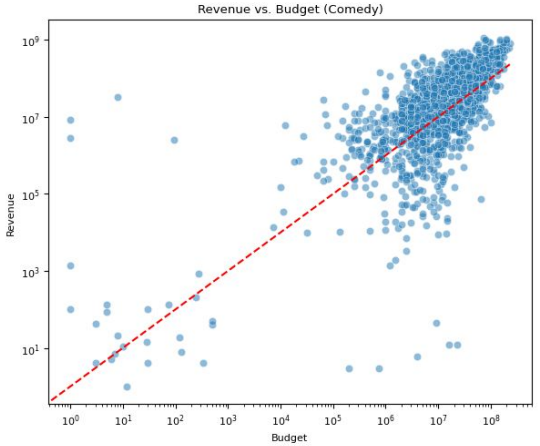


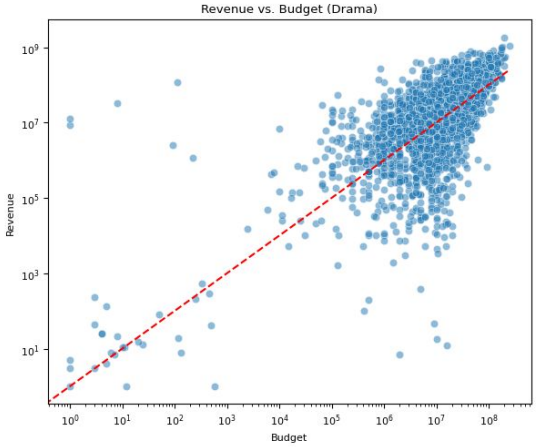


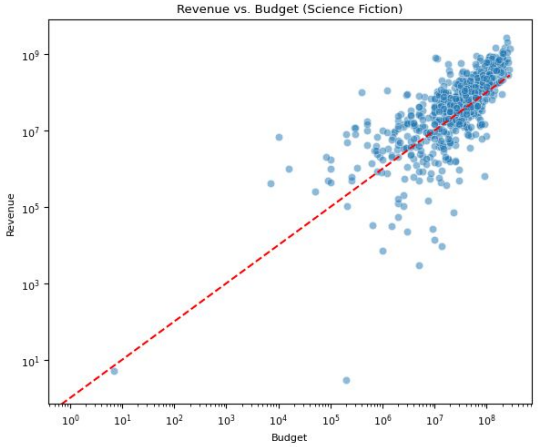
Scatter Plot: Revenue vs. Budget





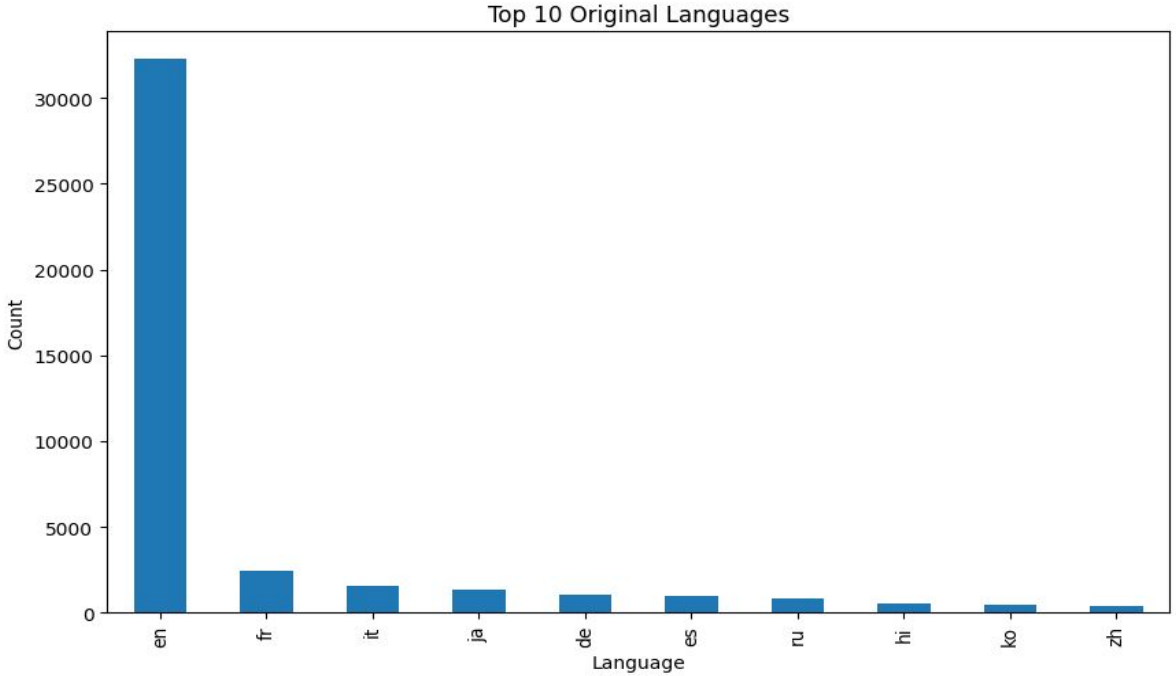






Original Language Distribution

- **en**-english
- **fr**-french
- **it**-italian
- **ja**-japanese
- **de**-german
- **es**-spanish
- **ru**-russian
- **hi**-hindi
- **ko**-korean
- **zh**-chinese



Reflection!

What have you learned from these visualizations and if they were helpful in understanding the data about movies?

And what visualizations you found were the most interesting to you?