**Project 2 - Team 4**

**Project Name:** Movie Analysis

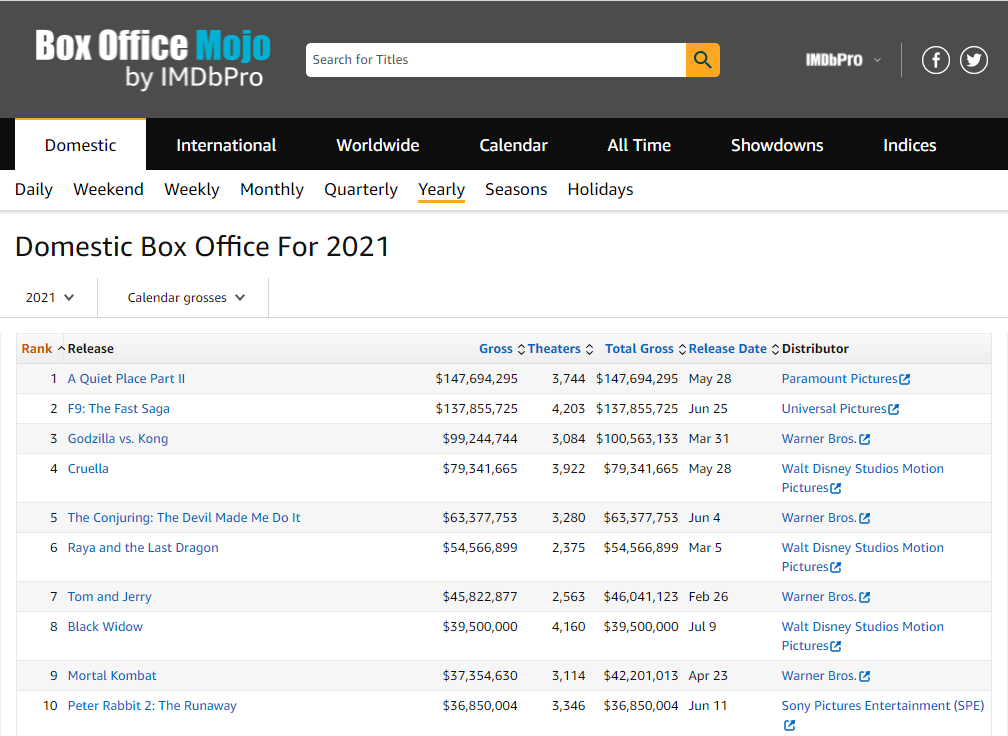
**Team Members:** Taylor Sperry, Harsha Vinoy, Alex Gainer, and Connor Scherer

**GitHub Repo:** https://github.com/tsperr/Project-2

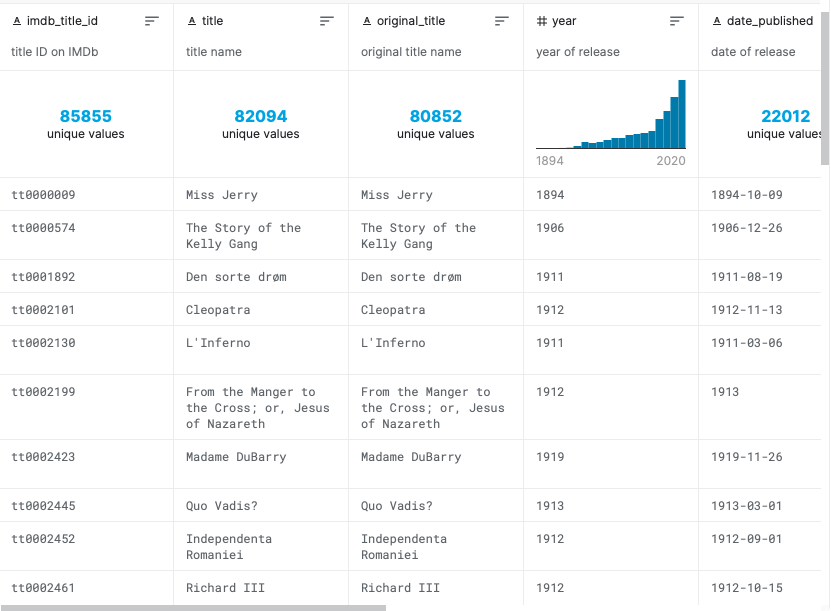
**Project Description:** We will be looking at past movie data to understand different patterns and trends among key variables. Using the Kaggle dataset on movies from 2000 to current, we will analyze the top 10 movies with the highest revenue internationally, identify any correlation between the budget of the movie and how much revenue it generated, look at the production countries to determine patterns and understand any trends based on the movie genres and the movie ratings. Finally, we will look at the current top 10 US movies based on IMDB charts.

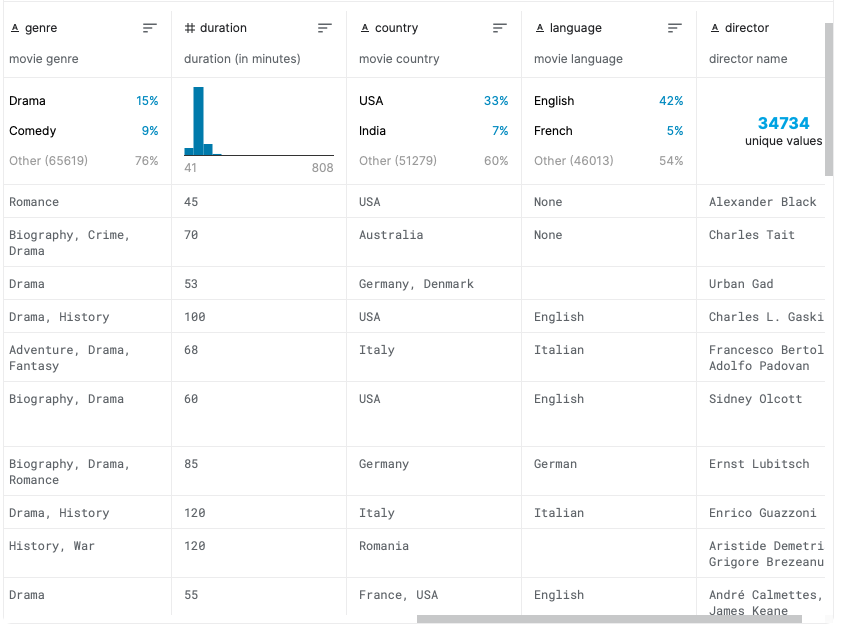
**Sources:**

* <https://www.boxofficemojo.com/year/2021/?grossesOption=calendarGrosses>
  + Current movies in theaters, shows ranking, gross revenue
  + Create visualization to show top 10 movies in that day
  + Web scraping to pull the name of the movie and total gross income
  + Screenshot:



* <https://www.omdbapi.com/>
  + API - movies data
  + Bring in IMDB rating for each movie
* <https://www.kaggle.com/stefanoleone992/imdb-extensive-dataset>
  + Movie dataset from IMdb
  + Columns:
    - title ID on IMDb
    - Movie title
    - Year of release (2000 to present)
    - Genre
    - Movie country
    - Budget
    - International Gross income
    - Production company
  + Screenshots





**Data Flow Structure**:

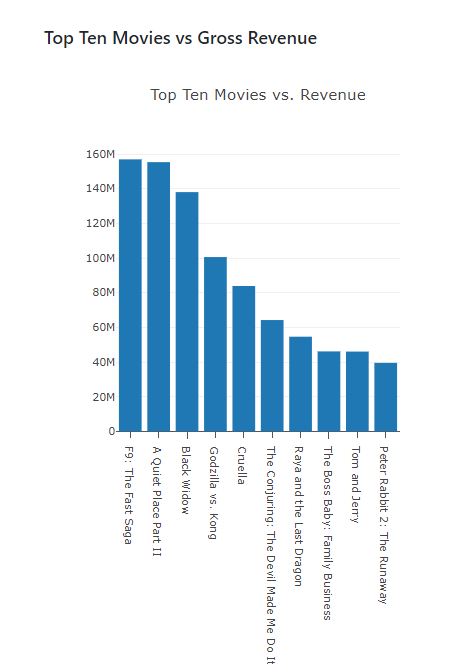
* Jupyter notebook - clean, webscrape data
* Postgres - store data
* Flask
* Html - webpage design
* JS - graphs (d3.json--url from flask)

**Chart Visualizations:**

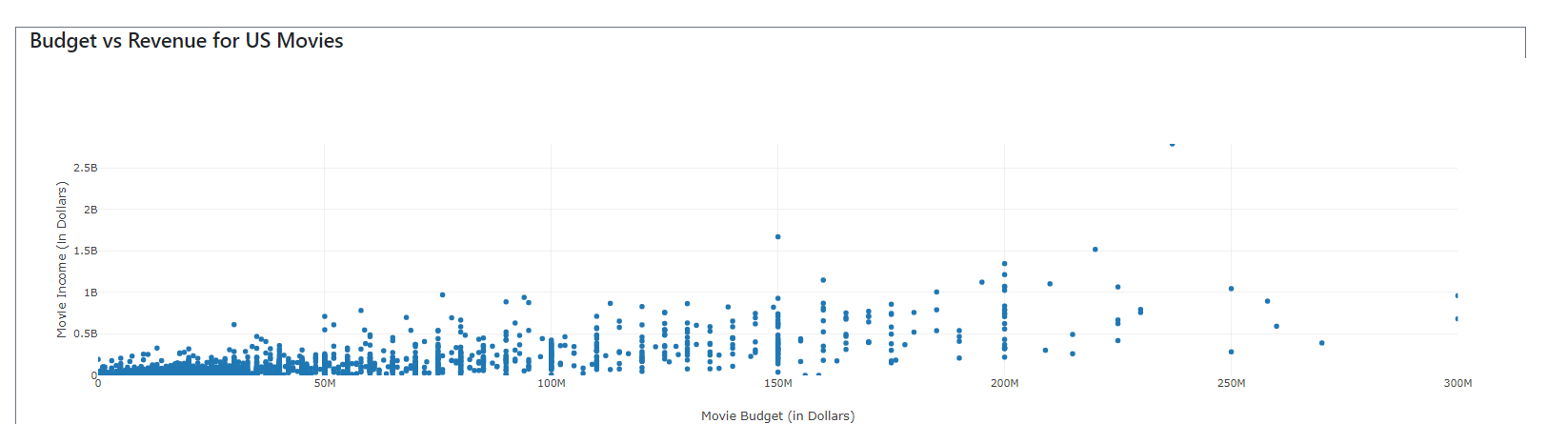
1. Production Country Visualization:



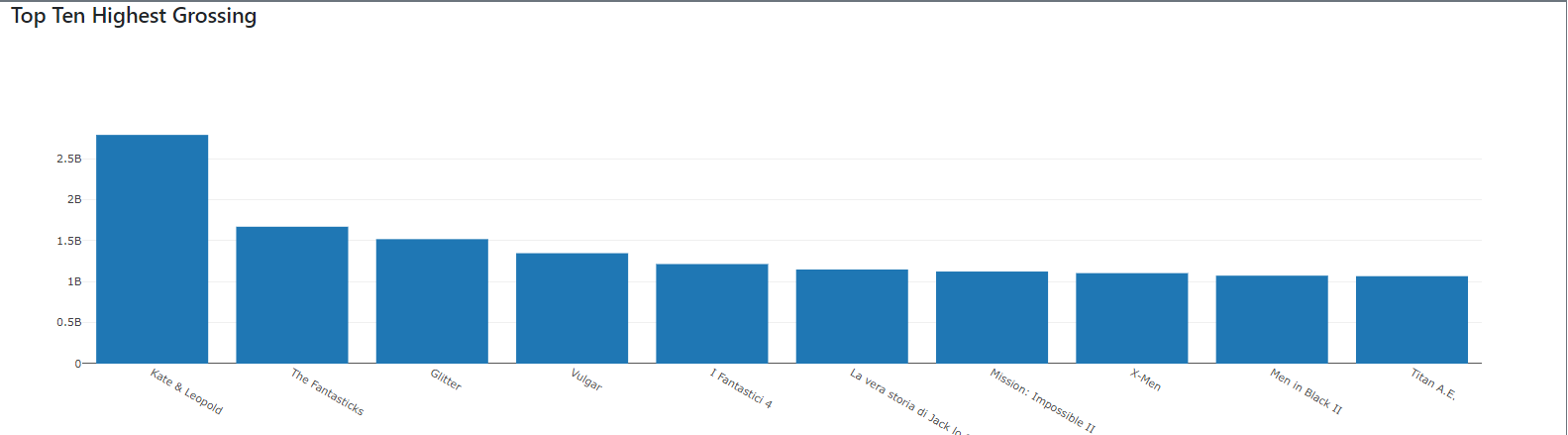
1. Current US Top 10 Movies Visualization



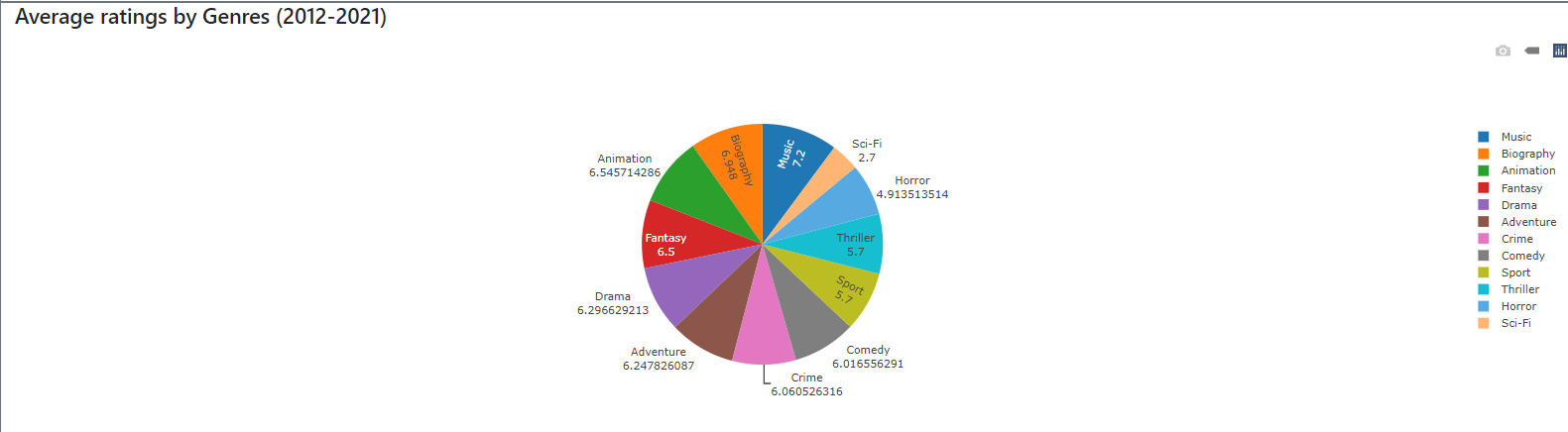
1. Budget vs Gross Income Visualization



1. Top 10 movies with Highest Revenue



1. Movie genres and Rating Visualization



**Webpage design Sketch:**

