Phase 1 Project

Movie Analysis

Summary

- With Microsoft already being a leader in the technology space, the organization would like to embark on a new venture-- this time onto the big screen. This project focuses on gathering pertinent data to create an analysis on which direction this tech giant should focus on to augment success for their first major film.

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

Business Problem

- It's not about reinventing the wheel here. Who can we model to replicate success in the first production.
- Are online movie database ratings relevant to potential gross earning success.
- What direction should the movie studio take to replicate such optimal gross earnings.

Data

- API

  - The Movie DB, we attempted to to do an API call, which was successful. This API was our initial dataset, however we decided to move towards a different direction.

- Web Scraping

  - IMDB's Top 1000 movies list the following characteristics were used: **movie, year, time_minute, genre, imdb_rating, metascore, vote, and gross_earning**.
  - Box Office Mojo's Top 1000 movies list: **title, year, worldwide_lifetime_gross, domestic_lifetime_gross, domestic_%, foreign_lifetime_gross, and foreign_%.**

Methods

- Finding the most successful movies from a global and domestic perspective.

- Finding the features of the top worldwide grossing movies and creating a dataframe that shows similar character attributes.

- Finding relationships based on character attributes.

- Coming up with inferences based on visualizations and selecting characteristics within the dataset.

Top 10 Worldwide Lifetime Movie Grossing
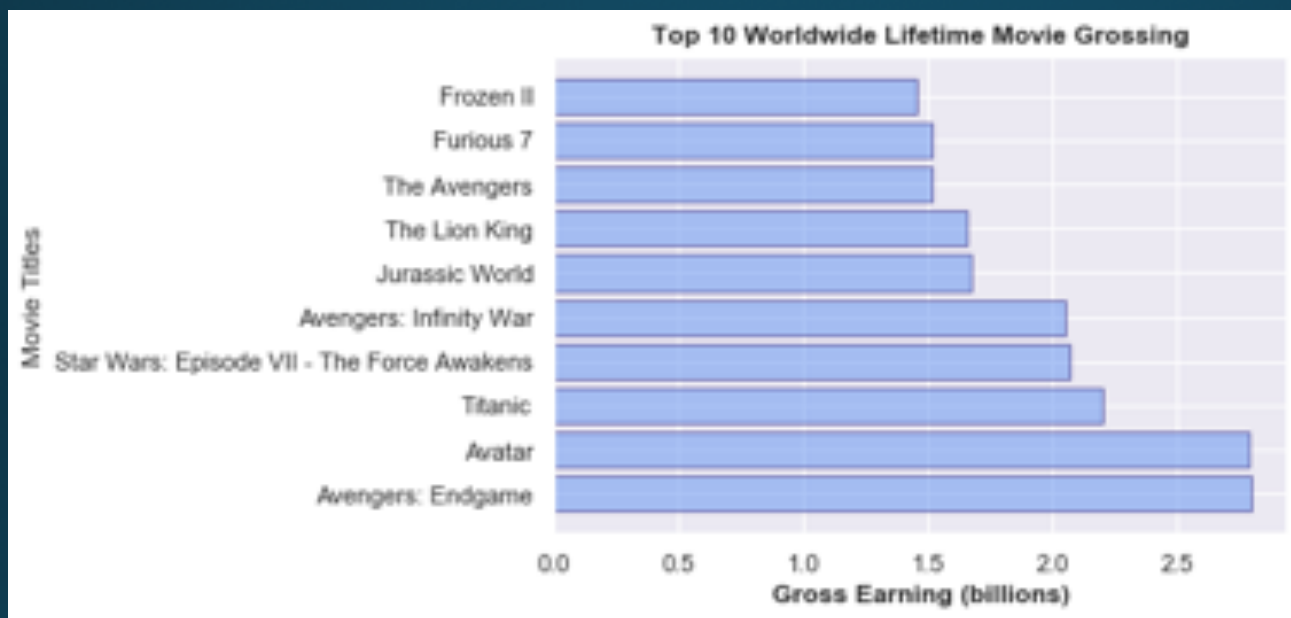
The ideal Movie

	Unnamed: 0	movie	year	time_minute	genre	imdb_rating	metascore	vote	gross_earning
19	22	Avengers: Endgame	2019	181	Action, Adventure, Drama	8.4	78.0	778254	858.37

Average Gross Earning for IMDB Ratings

Relationship Between Voters to Gross Earning
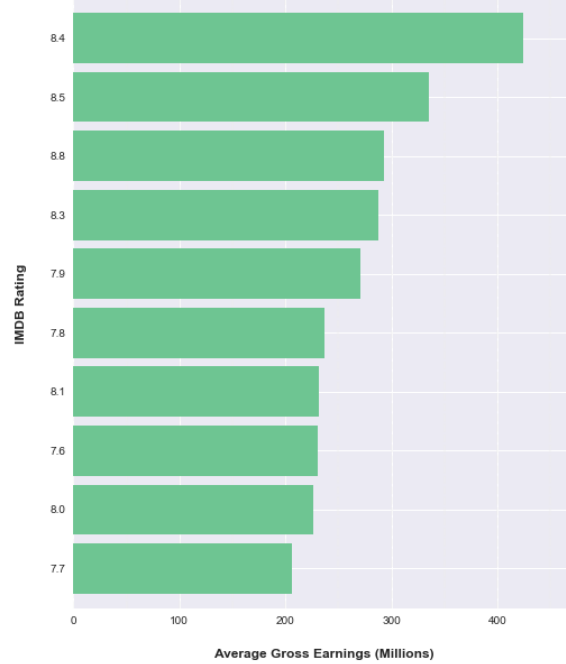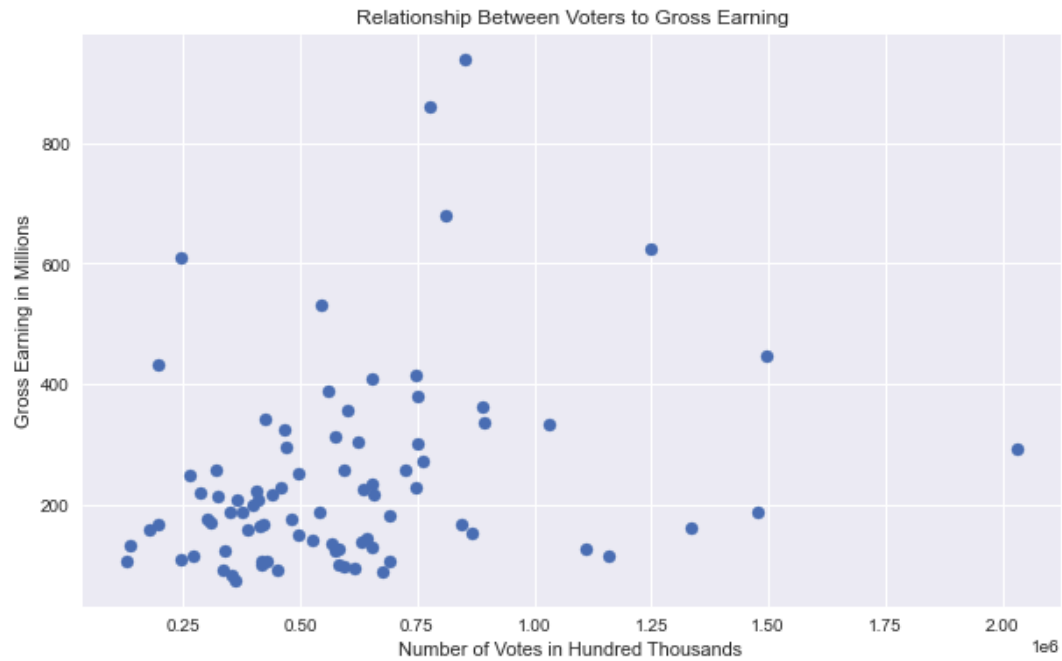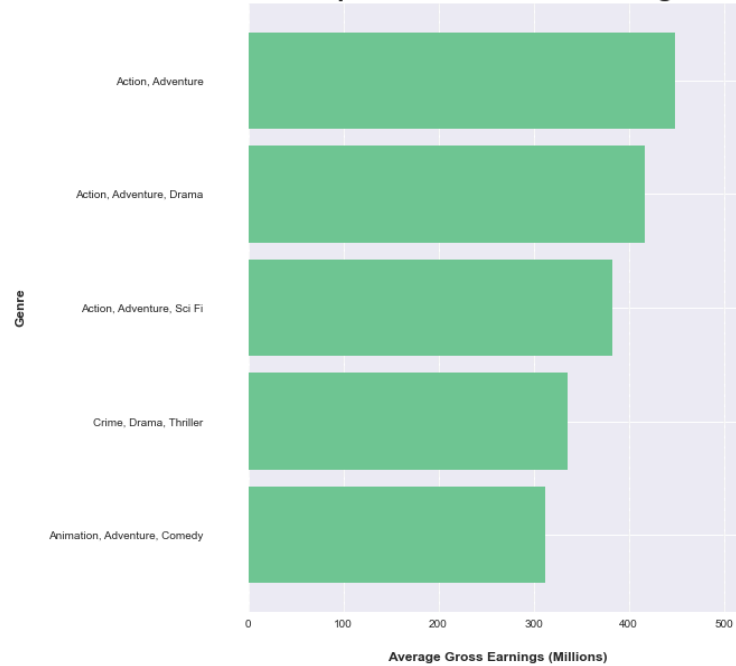
Average IMDB Rating for Genre

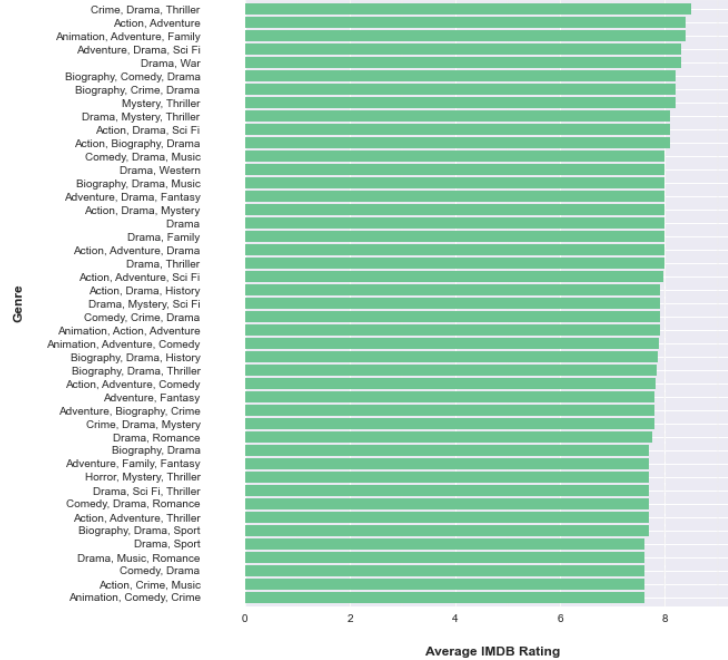| Genre	
Crime, Drama, Thriller	
Action, Adventure	
Animation, Adventure, Family	
Adventure, Drama, Sci Fi	
Drama, War	
Biography, Comedy, Drama	
Biography, Crime, Drama	
Mystery, Thriller	
Drama, Mystery, Thriller	
Action, Drama, Sci Fi	
Action, Biography, Drama	
Comedy, Drama, Music	
Drama, Western	
Biography, Drama, Music	
Adventure, Drama, Fantasy	
Action, Drama, Mystery	
Drama	
Drama, Family	
Action, Adventure, Drama	
Drama, Thriller	
Action, Adventure, Sci Fi	
Action, Drama, History	
Drama, Mystery, Sci Fi	
Comedy, Crime, Drama	
Animation, Action, Adventure	
Animation, Adventure, Comedy	
Biography, Drama, History	
Biography, Drama, Thriller	
Action, Adventure, Comedy	
Adventure, Fantasy	
Adventure, Biography, Crime	
Crime, Drama, Mystery	
Drama, Romance	
Biography, Drama	
Adventure, Family, Fantasy	
Horror, Mystery, Thriller	
Drama, Sci Fi, Thriller	
Comedy, Drama, Romance	
Action, Adventure, Thriller	
Biography, Drama, Sport	
Drama, Sport	
Drama, Music, Romance	
Comedy, Drama	
Action, Crime, Music	
Animation, Comedy, Crime	

Average IMDB Rating

Conclusions

- Who can we mimic?

  - Modeling the success of **Avengers: Endgame** would be the the most ideal route in in terms of box office success and rating.

- What about the ratings?

  - We did find that ratings to gross earning do share a relationship, but the number of votes do not necessarily contribute to a movies overall lifetime gross earnings

- How are we going to mimic success.

  - We found that a specific genres (Action and Adventure) are both profitable and popular, with an aim for an IMDB rating over 7.8

Next Steps

- Further considerations in optimizing data and models.
    - Financial conversations (budget and potential expenses)
    - Casting recommendations
    - Director recommendations
    - Box office release locations (foreign vs domestic)
    - Projecting earnings
    - See next for preview for Relationship Between Movie Runtime to Gross Earning

Relationship Between Movie Runtime to Gross Earning
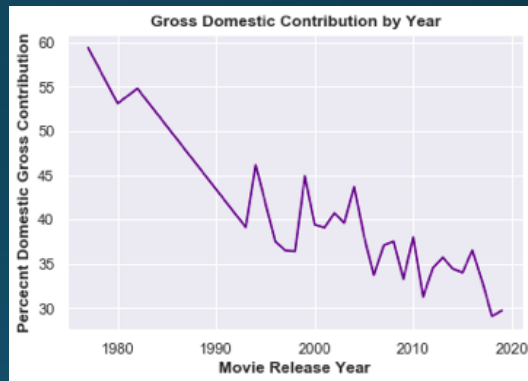
Thank You!

Christian Corrales
Ijeoma Akamnonu

GitHub:
https://github.com/ccorrales1228/Phase_1_Project