

Microsoft Movie Studio Analysis

Deztany Jackson
Data Science Flex
Phase 1 Project
July, 2022

Summary

This analysis directs Microsoft's potential studio head with actionable insights for their movie studio development. The analysis observed and extracted patterns from data of some of the "best" rated movies shown in the theatres in 2010s.

- Action, Adventure and Sci-Fi genres have some of the biggest budgets, top directors as well as engaged viewers
- The most common genre is Drama with varying budgets and many directors to choose from
- Optimize initial resources by not investing in Family, Sports or News genres.

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

Business Problem

- Understanding best box office films
- Characterize attributes that affect films
- Recommendations on movie films to make

Ref: [Phase 1 Project Description] (https://learning.flatironschool.com/courses/4963/pages/phase-1-project-description?module_item_id=370765), 2022

Data (Primary)

Attribute	Description
Movie Rating (Average)*	Target variable describing a movie's average rating (from average ratings that range from "1" to "10").
Number of Votes	Describes the number of votes the average is calculated from (ranging from "5" to "1.84 Mil")
Movie Genre	Describes 25 unique different types of movies (e.g. Action, Comedy, War). Movies can be classified into multiple types.
Movie Director	Describes the director who directed a specific movie
Movie Budget	Describes the movie budget allowance from 5782 movies. Range from 1E6 to 300E6 Million Dollars

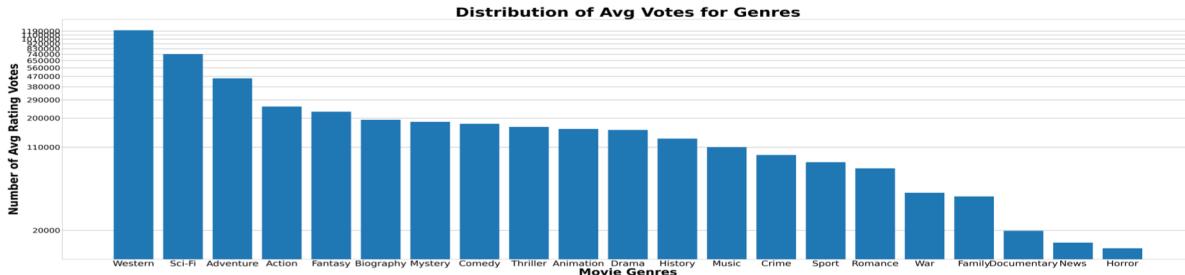
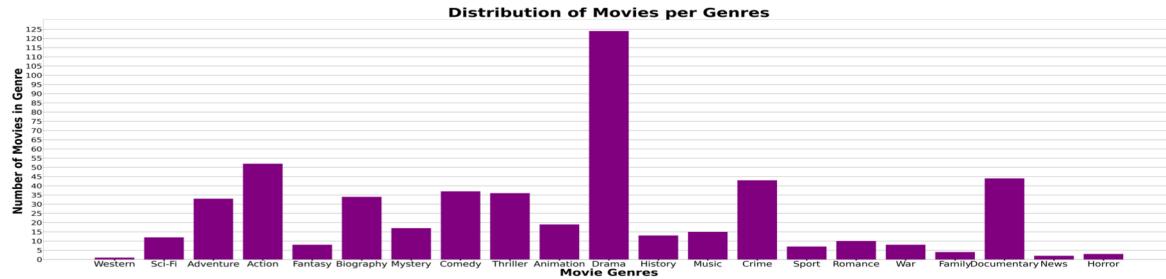
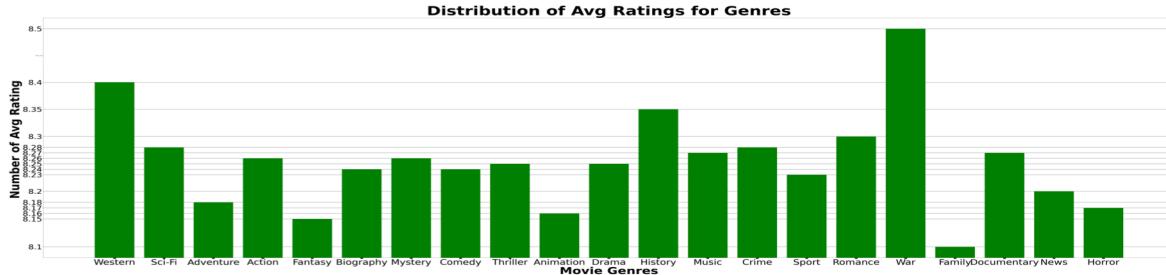
- Data taken from multiple data sources (i.e. IMBD, The-Numbers). These sources included limited overlapping data.

Methods

- Method
 - Analysis Tools: SQL & Python
 - Data Prep:
Filtered- Minimum 8/10 Average Rating & Minimum Average (5330)Rating Votes
Removed missing data
Combined datasets (from different sources) which limited data analysis amount
 - Data Analysis:
Focused on assessing patterns around rating and genre
Rating patterns support further exploration

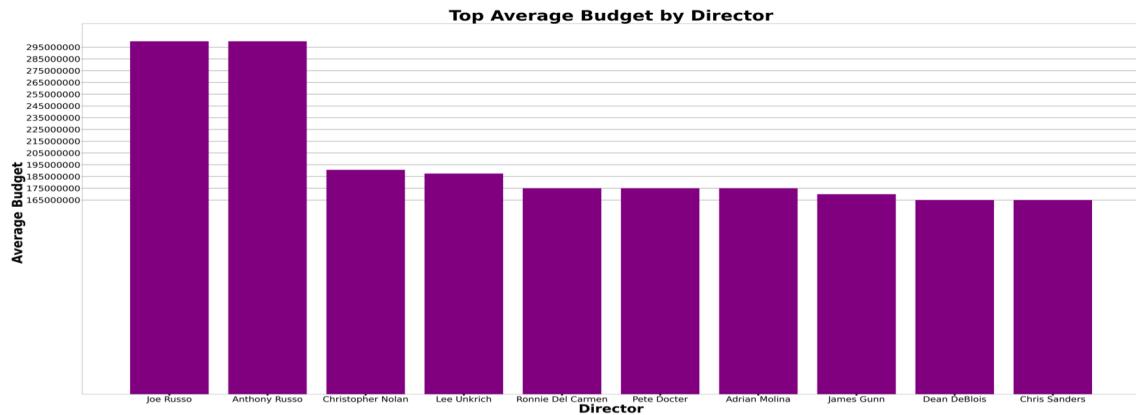
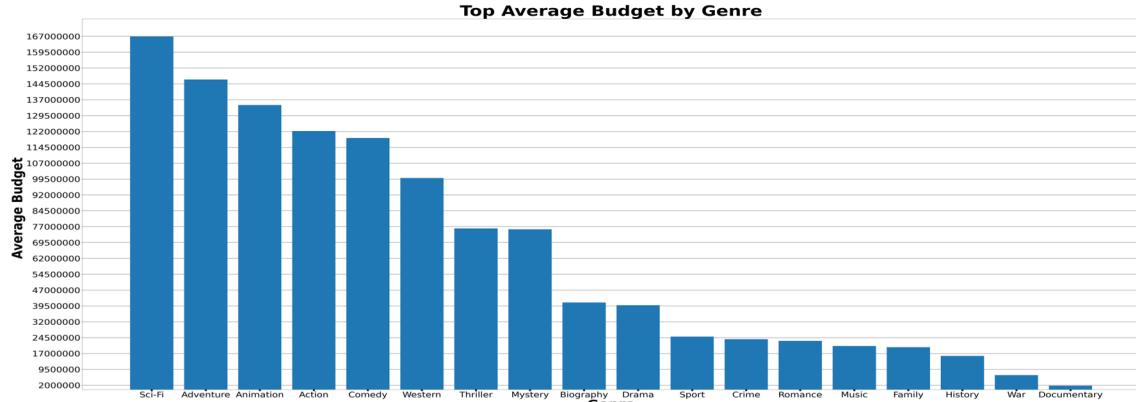
Results: Genre Focused

- Most movies made: Drama
- Favorable values (i.e. ratings, movies made and avg votes): Action and Adventure



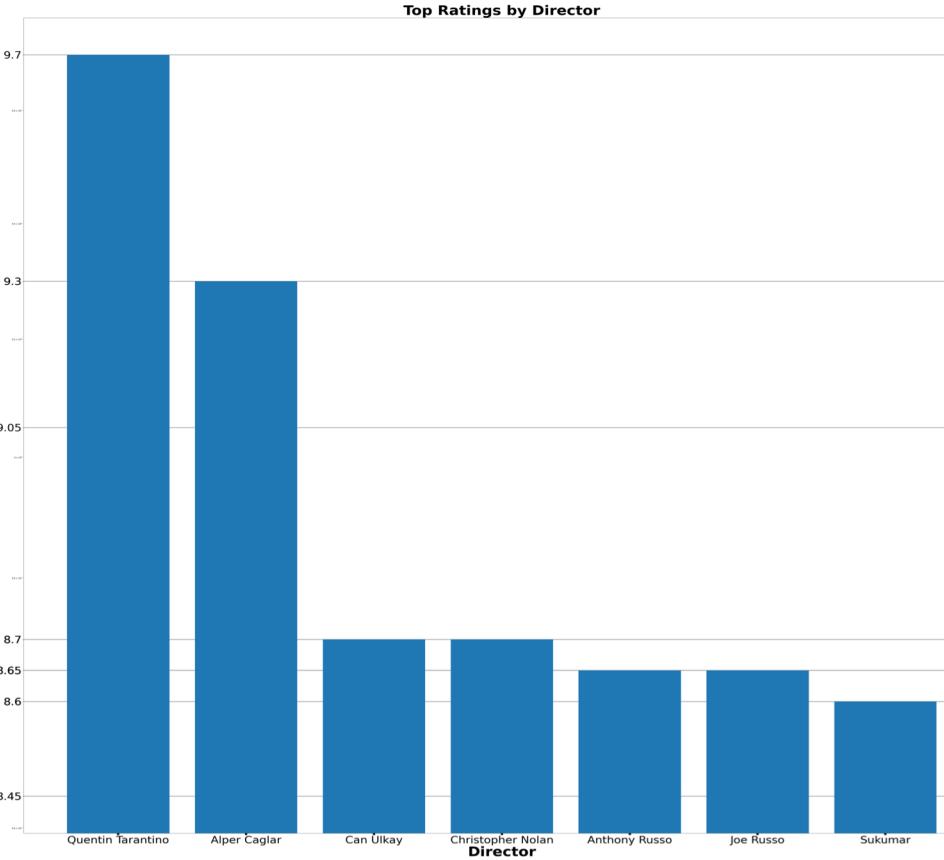
Results : Budget Focused

- High budgeted movie genres were the Sci-Fi, Action, Animation and Adventure
- The Russo brothers and Christopher directed the top budgeted movies



Results: Director Focused

- The high budgeted movie genres: Sci-Fi, Action, Animation and Adventure
- Top budgeted movies: The Russo brothers and Christopher Nolan



Director	genres	Avg_Votes	Avg_Rating
Alper Caglar	Action	100568.0	9.3
	Drama	59006.0	8.6
	War	100568.0	9.3
Anthony Russo	Action	556031.0	8.6
	Adventure	556031.0	8.6
	Sci-Fi	556031.0	8.6
Can Ülkay	Drama	17763.0	8.4
	History	26743.0	8.7
	War	26743.0	8.7
Christopher Nolan	Action	1614418.0	8.6
	Adventure	1570200.0	8.7
	Drama	1299334.0	8.6
	Sci-Fi	1570200.0	8.7
	Thriller	1387769.0	8.4
Joe Russo	Action	556031.0	8.6
	Adventure	556031.0	8.6
	Sci-Fi	556031.0	8.6
Lee Unkrich	Adventure	479706.0	8.3
	Animation	479706.0	8.3
	Comedy	479706.0	8.3
Quentin Tarantino	Comedy	5600.0	9.7
	Drama	608503.0	9.0
	Western	1211405.0	8.4
Sukumar	Action	28266.0	8.4
	Drama	15407.0	8.6

Conclusions

- Invest in at least one of high budget films in the category of "Action, Adventure or Sci-Fi".
- Invest in multiple movies in "Drama, Documentary and/or Crime". If one genre, choose "Drama".
These types of movies are popular and lower cost risk.
- For the initial start of the movie studio, DO NOT invest resources in genres focused on "Family, Sports or News".

Future improvements:

- Add more parameters to analysis (e.g. region, language).
- Do more detailed work into the genres (causality and correlation)
- Analyze the combinations of genres with respect to the target variable and independent variables.
- Research more on Microsoft's current technical ecosystem and their holistic vision as a company.

Thank You!

Email: dmvinedata@gmail.com

GitHub: [@dmvinedata](https://github.com/dmvinedata)

LinkedIn: [linkedin.com/in/deztanyjackson/](https://www.linkedin.com/in/deztanyjackson/)