A Movie Analysis for Microsoft

By Benjamin Dean

Introduction

This project :

- provides a general analysis of movie data
- focuses on the utilization of data science tools in the Python programming environment
- answers a business problem with the data science process

Outline

- Business Problem
- Data Exploration
- Methods
- Results
- Conclusions

Business Problem

- Microsoft Corporation is forming their own movie studio and seeks advice
- This project explores and begins to answer the following questions:
  - What defines a successful movie?
  - Which genres produce the most profit?
  - Which directors bring the most to a movie in terms of profit and ratings
  - What is the highest grossing production budget to begin with

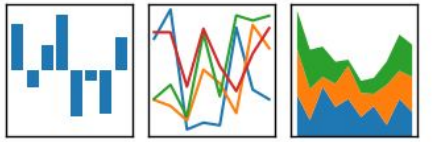
Data

- This project uses six datasets from the websites IMDB.com, TheMovieDb.org and the-numbers.com
- Once datasets were merged they included 1213 observations from the decade 2010-2019
- Data included production budgets, worldwide grosses, director name, movie titles, etc.

- Datasets were imported, cleaned and merged using Python's Pandas and Numpy libraries

pandas

$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$



NumPy

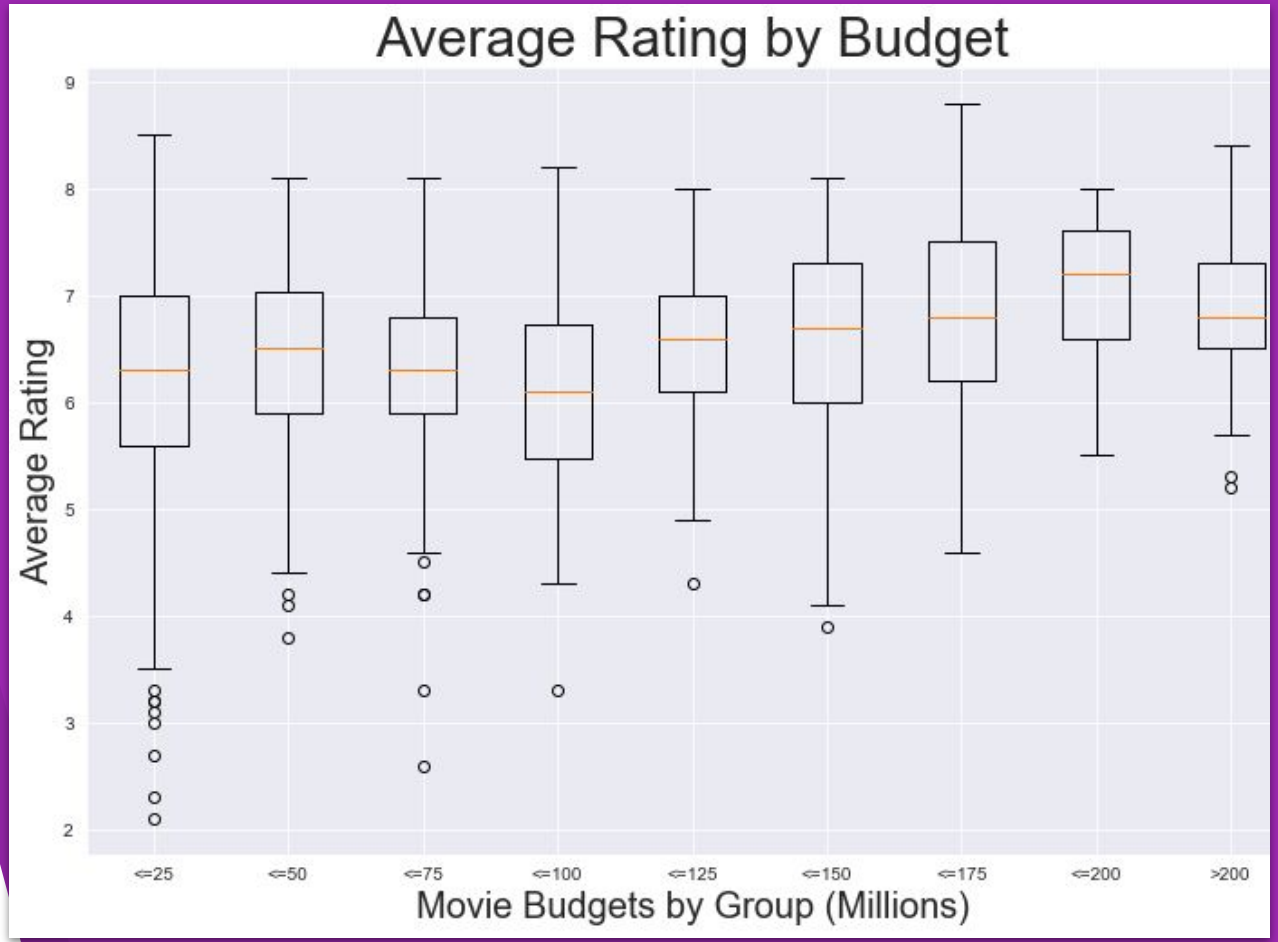
- Visualizations were made using the Matplotlib and Seaborn libraries

matplotlib

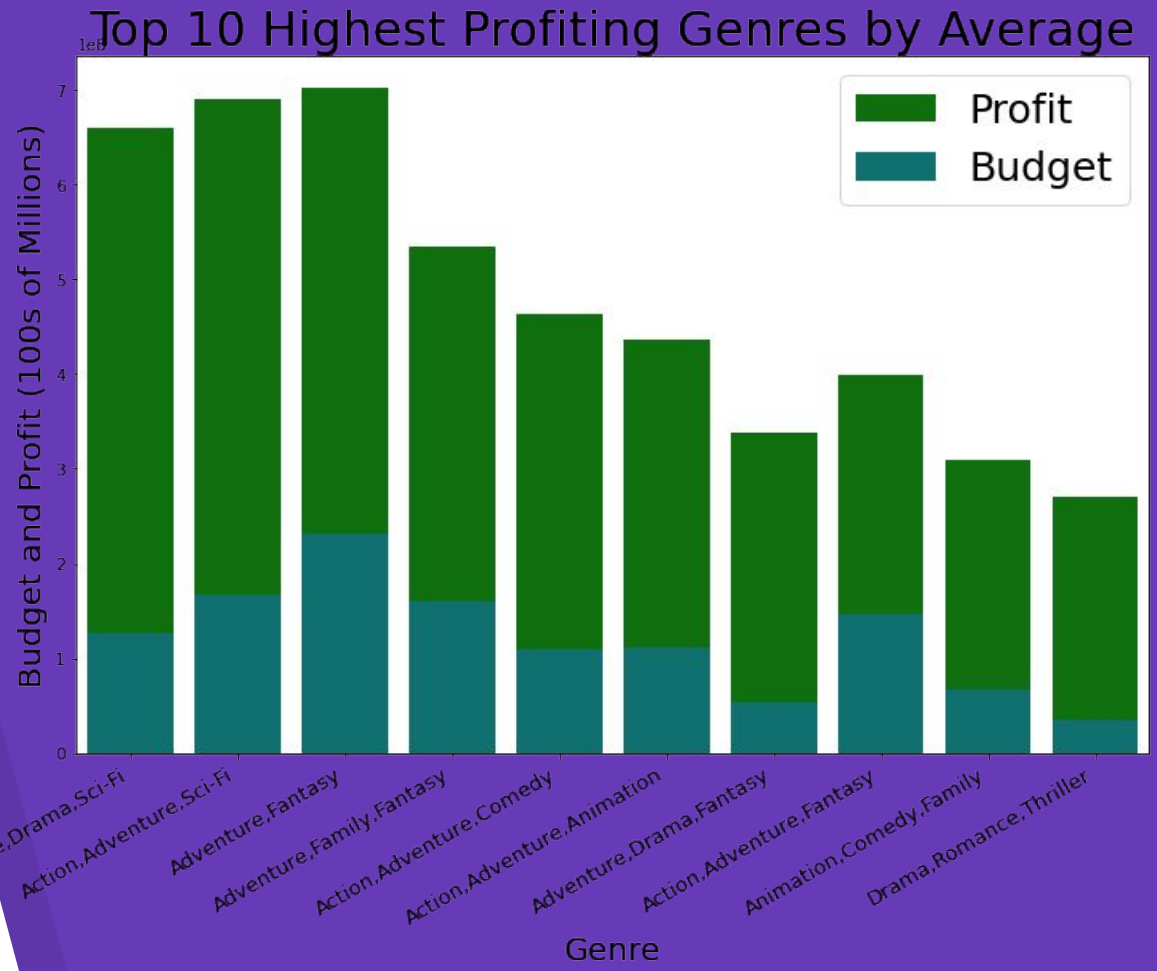
seaborn

Results

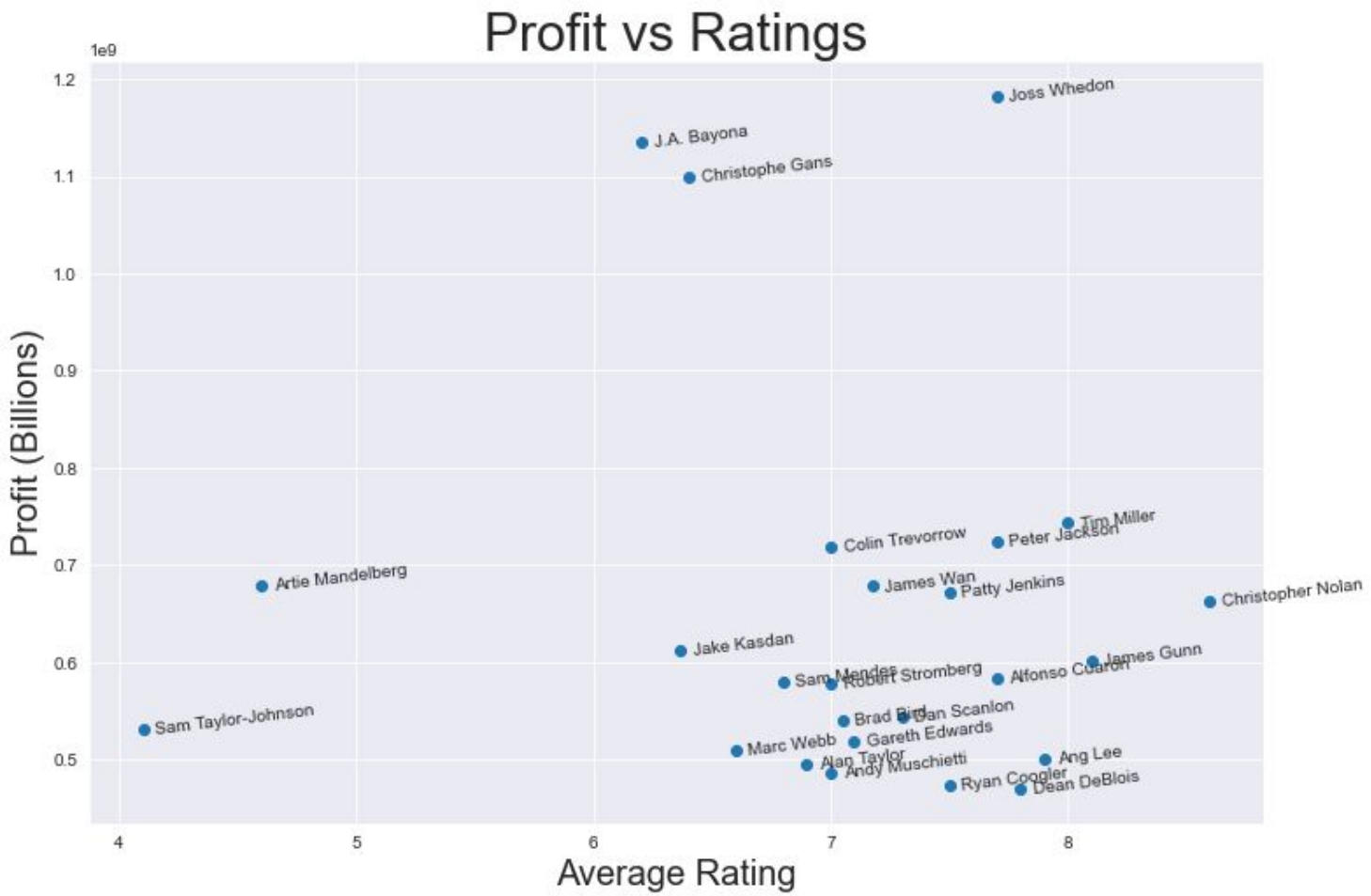
- Here we see that the variance generally decreases as budget increases
- Analysis suggests Microsoft choose a budget between 176mm and 200mm



Results


Top 10 Highest Profiting Genres by Average
1e8
Budget and Profit (100s of Millions)
0
1
2
3
4
5
6
7
Profit
Budget
Adventure,Drama,Sci-Fi
Action,Adventure,Sci-Fi
Adventure,Fantasy
Adventure,Family,Fantasy
Action,Adventure,Comedy
Action,Adventure,Animation
Adventure,Drama,Fantasy
Action,Adventure,Fantasy
Animation,Comedy,Family
Drama,Romance,Thriller
Genre

Results


Profit vs Ratings
1e9
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5
Profit (Billions)
4
5
6
7
8
Average Rating
Joss Whedon
J.A. Bayona
Christophe Gans
Tim Miller
Colin Trevorrow
Peter Jackson
Artie Mandelberg
James Wan
Patty Jenkins
Christopher Nolan
Jake Kasdan
James Gunn
Sam Mendes
Robert Stromberg
Alfonso Cuaron
Sam Taylor-Johnson
Brad Bird
Dan Scanlon
Gareth Edwards
Marc Webb
Alan Taylor
Andy Muschietti
Ang Lee
Ryan Coogler
Dean DeBlois

Conclusions

Analysis suggests Microsoft take the following actions:

Begin with a production budget within the range of $176mm - $200mm

Focus on the genres: Action, Adventure, Fantasy, and Science-Fiction

Hire directors Joss Whedon, Christopher Nolan, Tim Miller and Peter Jackson

Future Analysis

- Employ webscraping to incorporate more data
- Explore other datasets that contain information of staff of movies

Thank you
for reading

GitHub: Benjamin2817
Email: benjamin.a.dean.17@gmail