Movie Data Analysis

Helping guide business ventures through exploratory data analysis.

By: Andrew Reusche

Summary

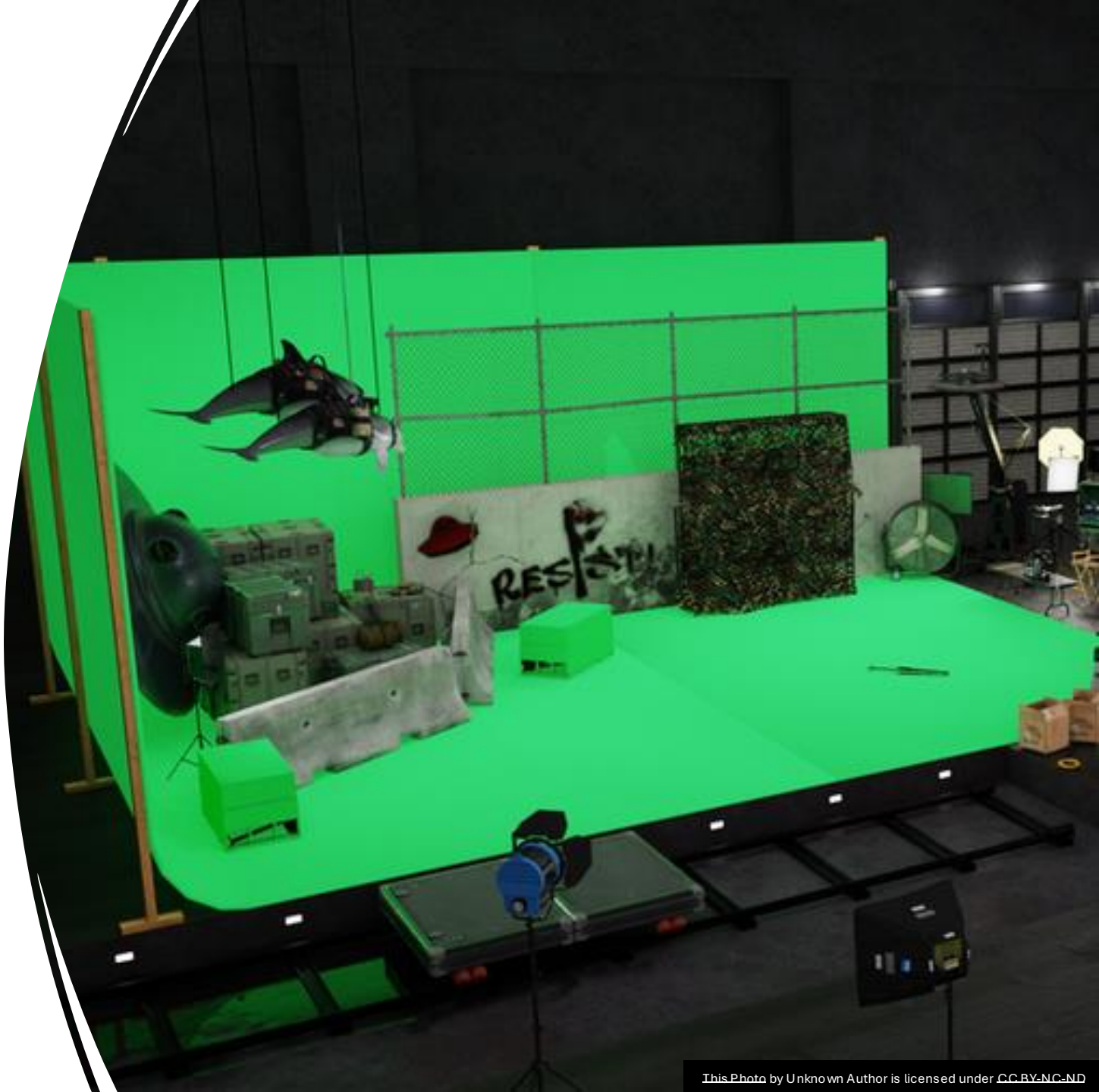
- This project uses data cleaning, feature engineering, and descriptive data analysis of movie datasets from IMDB.com and The-Numbers.com to help our new movie studio director make informed decisions about the direction this new business venture needs to go to succeed in the modern-day market.

Outline

- **Business Problem:** Help movie studio decide what type of film to create.
- **Data:** Movie data from IMDB.com and The-Numbers.com
- **Methods:** Data Cleaning, Feature Engineering, and Analysis
- **Results:** Production Budget Level, Movie Genre, Movie Release Date
- **Conclusions and Next Steps:** Analysis Conclusions and possible future analyses.

Business Problem

- **New Business Venture:** Movie Production Studio
- **Venture Challenge:** Current limited knowledge of the movie market
- **Value Added to Venture:** Through Exploratory Data Analysis (EDA), provide new venture decision-makers with information they can use to help select what type of movies should be made.



Data Used & Data Limitations

Data Used

- Movie data from IMDM.com and The-Numbers.com
- 139,457 different movies released from 1915-2020
- Movie details such as earnings, movie staff, production budget, rating, and more

Data Limitations

- Limited information on movies
- It can only highlight characteristics successful and unsuccessful movies had, instead of claiming cause.

Data Analysis Methods

- **Data Cleaning:** Filtered out all movies not in the recent market and all data not immediately relevant to impactful decisions our stakeholders could make.
- **Feature Engineering:** Modify filtered data to extract more meaningful information, that can be used to analyze current movies.
- **Data Analysis:** Analyze the newly engineered data to determine what specific characteristics modern successful and unsuccessful movies have.

Movie Level of Success in The Domestic Market

Investment Loss: Made less money than was invested in production.

Profit Made: More money was earned than was invested in the production budget.

Investment Doubled: At least twice as much money was earned as was invested in the production budget.

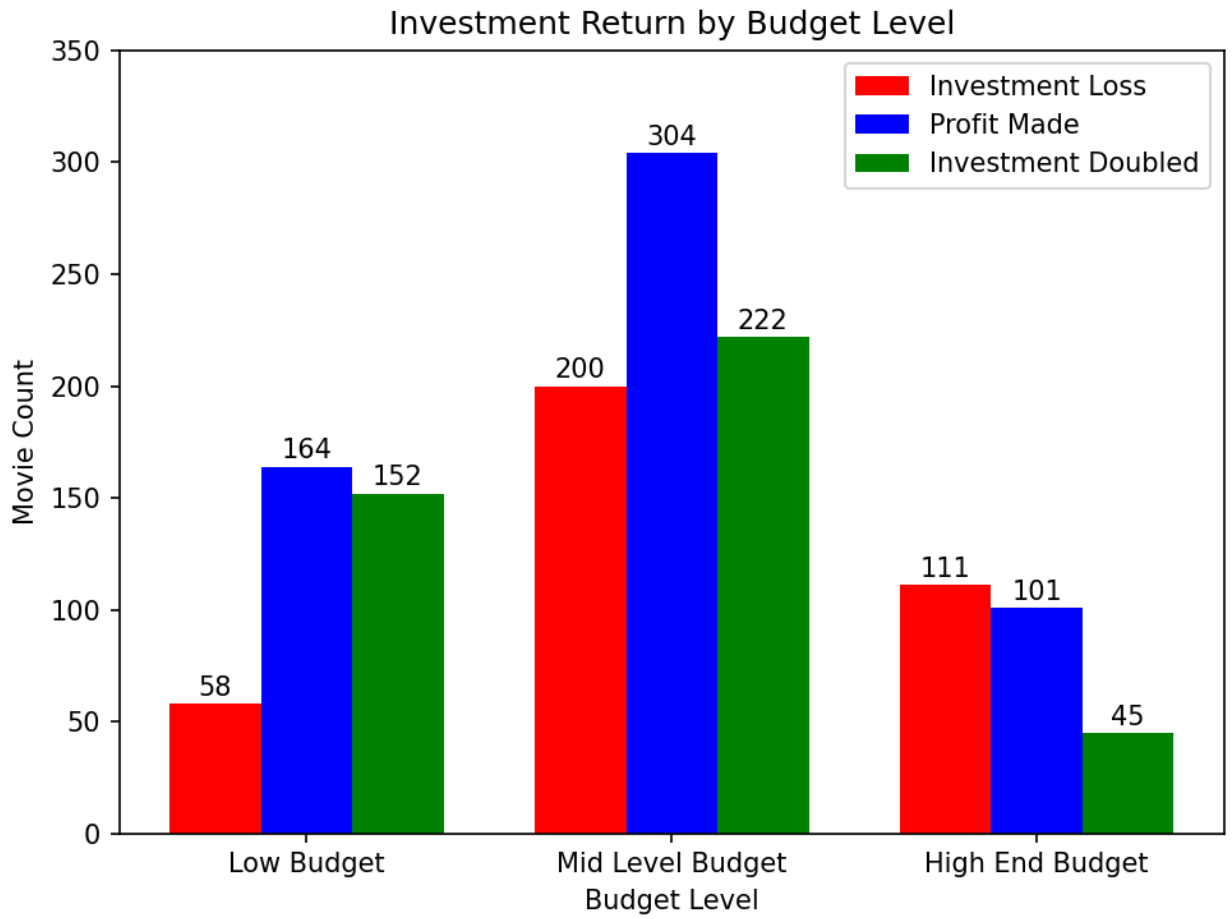
Results: Production Budget Level

Low Budget: Less than $5,000,000

Mid-Level Budget: Between $5,000,000 and $50,000,000

High End Budget: Over $50,000,000

Low Budgets resulted in the lowest frequency of investment losses and the highest frequency of investments being doubled.



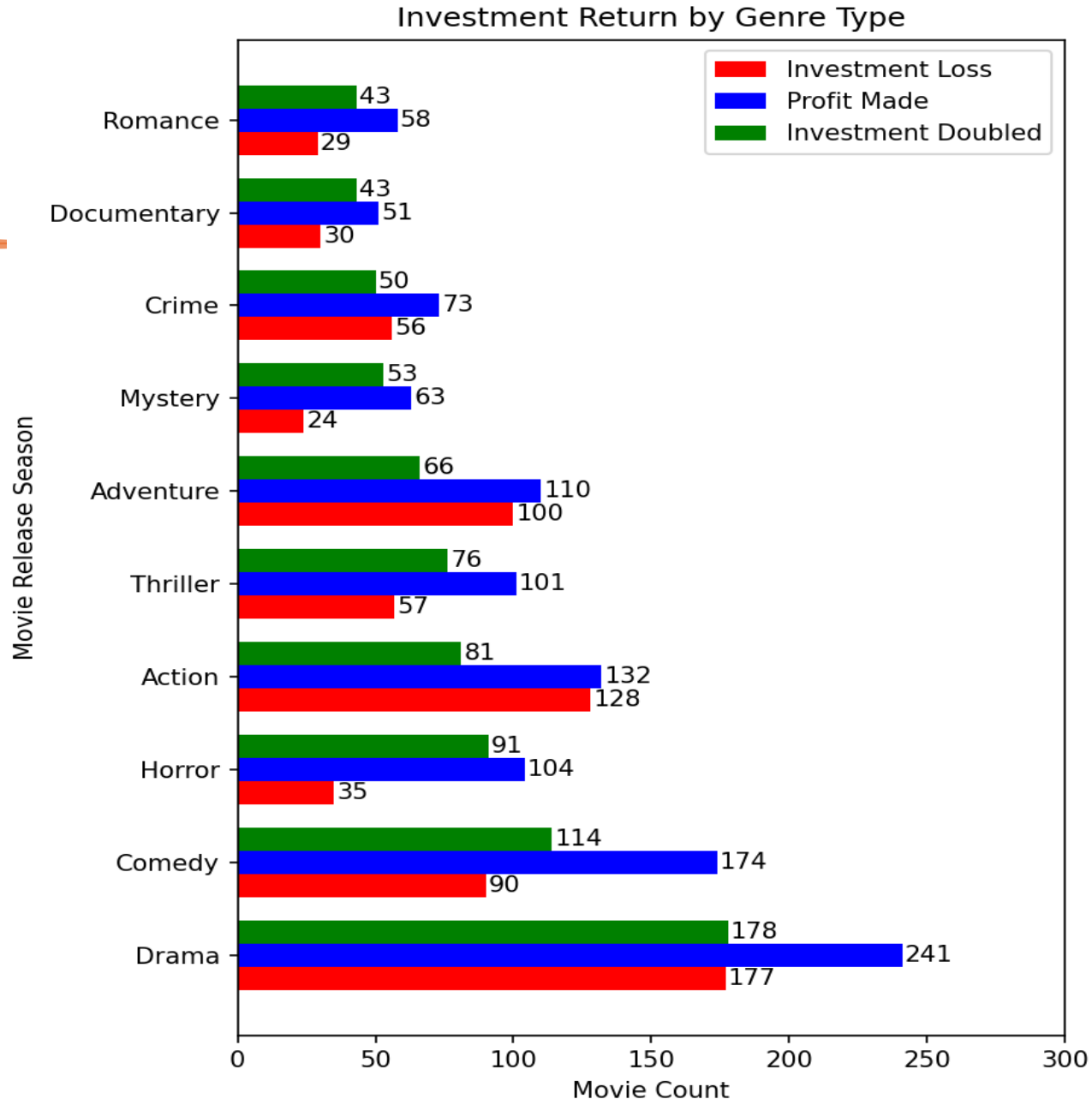
Results: Movie Genre

Most Created: Drama

Worst Performer: Action/ Adventure

Best Performer: Horror

Horror movies resulted in the lowest frequency of investment losses and the highest frequency of investments being doubled.



Results: Movie Release Date

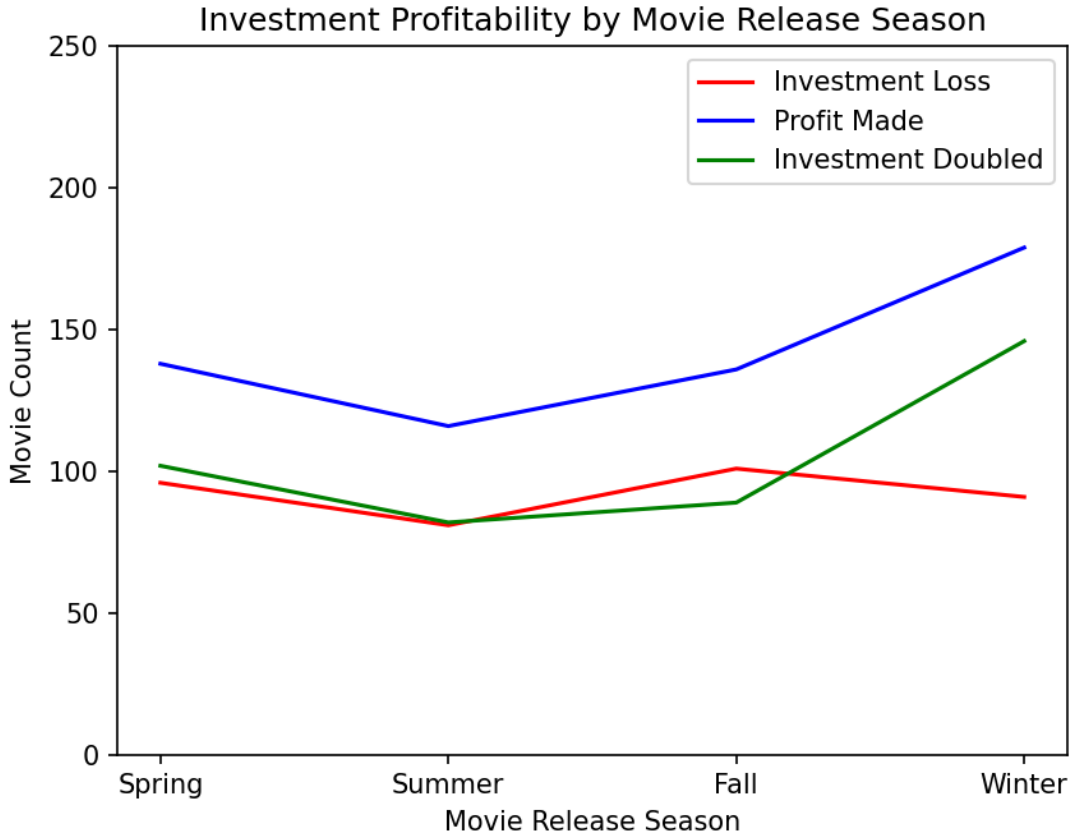
Spring: March, April, May

Summer: June, July, August

Fall: September, October, November

Winter: December, January, February

The Winter Season resulted in the biggest deviation from investment losses, with a significant increase in the frequency of both movie profits and investments being doubled.



Conclusions

- Start out with a smaller production budget under $5,000,000.
- Make a movie in the horror genre.
- Release the movie during one of the winter months (December, January, February).

Following these suggestions could allow us to break into the market at a lower price point and with minimized risks. Additionally, once we have a solid footing in the market and have gained experience and recognition we could expand out to global markets, more lavish productions, and other genres.

Next Steps

- (Potential Further Analyses)

- Actor Analysis: Given more data about different actors, this analysis could help narrow down who the current popular actors are, or the actors that cause the biggest spike in ticket sales when put in a movie.
- Marketing Technique Analysis: Given more data about different marketing techniques, this could shed light on what the most effective current strategies are, and what it would take to roll them out.
- Global Market Analysis: Given more data about the global market this could help predict what types of movies and attributes of movies would be successful in other countries around the world, increasing our market share.

Thank you!

- Email: A.J.Reusche@gmail.com
- Github: https://github.com/AndrewReusche
- LinkedIn: https://www.linkedin.com/in/andrew-reusche-1397bb311/