

# Project Phase 1

By Stephen Ndirangu

# Business understanding

- ▶ Microsoft sees all the big companies creating original video content and they want to get in on the fun. They have decided to create a new movie studio, but they don't know anything about creating movies. ¶
- ▶ This analysis is to help reveal insights for Microsoft's new movie studio. The point of interest is which movies are doing well currently in the box office.
- ▶ This analysis will focus on ratings and revenue of the movies as the relevant information

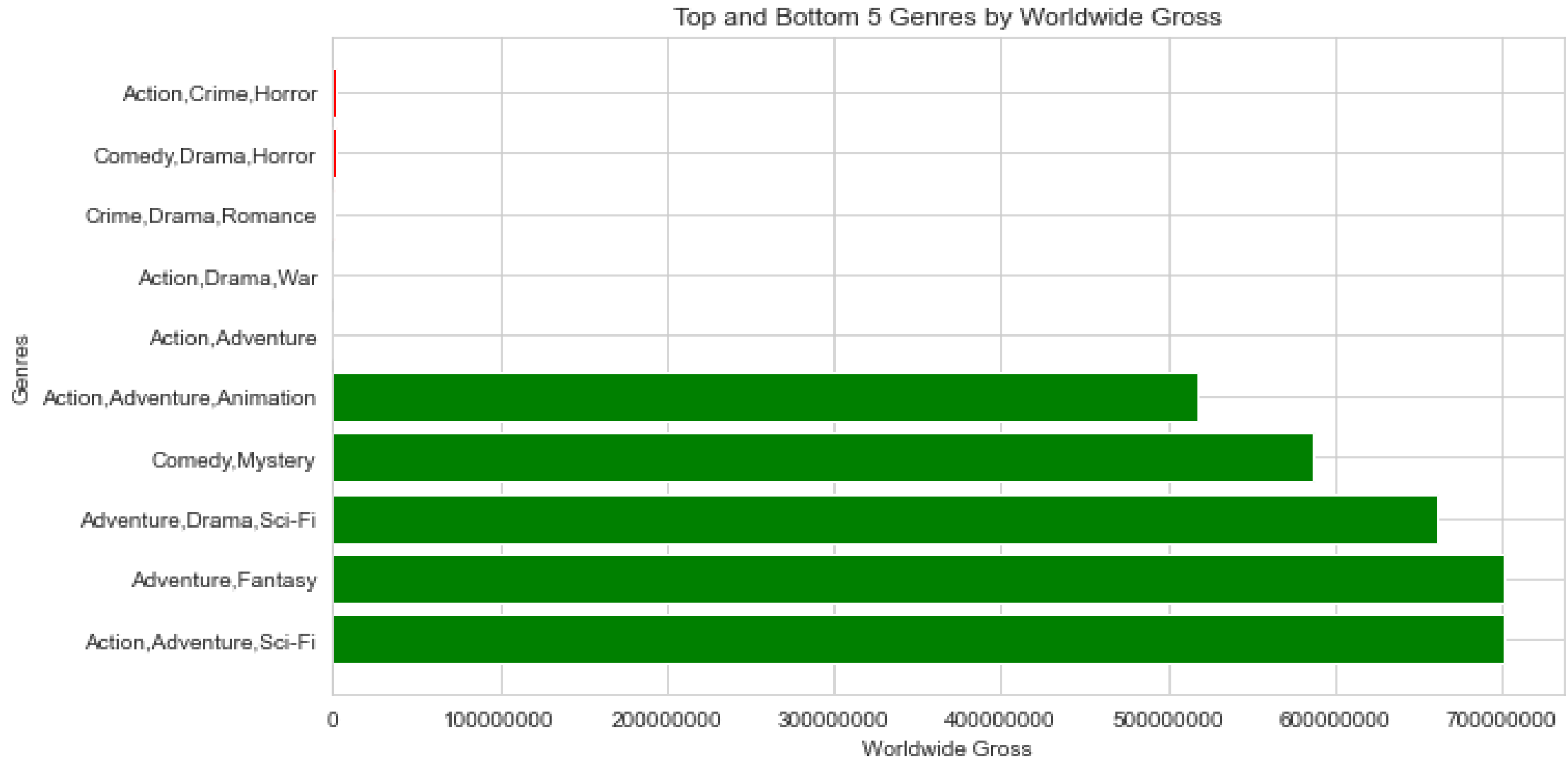
# Data used:

- ▶ Box office Mojo dataset
- ▶ IMDB SQL database
- ▶ The numbers dataset
- ▶ They are trustworthy as they are legitimate sites for getting reviews for movies

# Methods used

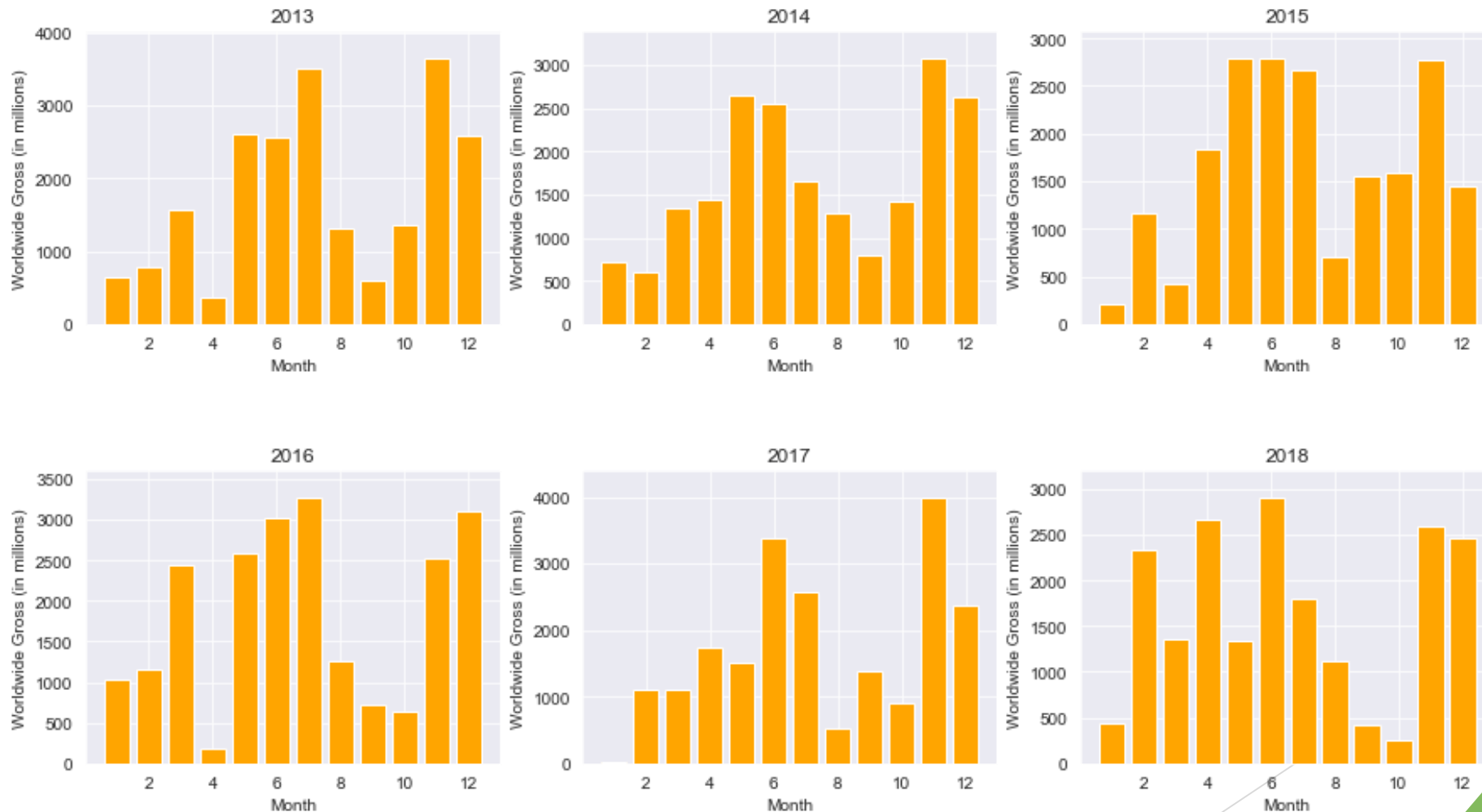
- ▶ Data cleaning, choosing to drop Nan columns if many, eg Runtime Minutes had however, kept some and renamed to “Missing” instead for Genres
- ▶ Dropped Entire Column “Foreign gross” in the second dataset, as the 3<sup>rd</sup> dataset would have worldwide gross

# Top and bottom 5 Genres based on Worldwide Gross



# Monthly gross per month in last 6 years

Monthly Worldwide Gross for Last 6 Years



# Budget vs Worldwide Gross



# Conclusion

## Budget

- ▶ The production budget has very strong positive correlation with the domestic and worldwide grosses

## Average rating

- ▶ -movies rated higher did much better in the box office, both for domestic and worldwide audiences, and in fact, movies rated lower than 5 did very poor
- ▶ -curiously, the rating isn't really affected by increased budget, probably due to the fact that many things go into a production
- ▶ -it would seem the average rating has a weak positive correlation with both domestic and worldwide gross,

## Genre

- ▶ Generally genres don't really affect a rating, but the top 5 that make money worldwide are anything of the "adventure genre" involving sci-fi, animation and comedy
- ▶ ### Runtime
- ▶ -Rating has a strong positive correlation with the runtime of the movie, but closer investigation shows the distribution is clustered within a value range of 80 to 140 minutes

## Month of release

- ▶ -it would seem from the per-year and mean of 5 years bar plots, that in the middle of the year and in November, these are the best times to release movies



# Recommendations to Microsoft

- ▶ 1) Microsoft should invest in the Genres of Adventure, with sci-fi, comedy, and/or animation, with main projects being of the genre "Adventure, Drama, Sci-fi" as it gets high reviews as well as high grossing worldwide
- ▶ 2) A high budget in these genres will give better returns worldwide, somewhere above the 100 million mark
- ▶ 3) A runtime of Between 80 minutes and 140 minutes is the most consistent at good ratings
- ▶ 4) Releasing between months 4-6, and at the tail end of the year might also give good worldwide grosses...probably because these are the months in which holidays occur eg easter, christmas