

# MICROSOFT MOVIE STUDIO DATA ANALYSIS PROJECT

By Abel Odhiambo



# **There is Always a Light at The End of A Tunnel**





# Objectives

- ❖ Find the top movie genre
- ❖ Find the most popular movie genre
- ❖ Find the reception a movie gets worldwide and locally
- ❖ Find the average runtime for a movie and the runtime trend
- ❖ Find gross distribution locally and worldwide
- ❖ Find which is the best movie



# 1. Business Understanding

In my project Microsoft wants to start a movie studio. It wants to ensure that it comes up with a competitive and profitable movie studio.

It also wants to ensure the studio standards will match the worlds best movie studio



## 2.Data Understanding

In this project the data that would be need are;

- Dataset that has movie genres
- Dataset that has movie budget and gross profit locally and worldwide
- Dataset that has runtime distribution and years for the same movies
- Dataset that has movie ratings



## 3.Data Preparation

For high accuracy for the project , I chose 4 datasets out of the 5 given data set.

### 3.1: Data Cleaning

- Some data with null values especially data with amounts were replaced by 0 while categorical data which were important were replaced with object values.
- Data with missing key values were dropped to enable more accuracy for the analysis.
- Data with combined date and time were split according to their respective years
- Some calculation columns were added eg column for calculating percentage gross profit and the date that were split



## 4.Data Analysis

### 4.1: Top movie genre

One data set from the selected 4 had the movies categorized into their respective genres.

I used the movies average ratings and then grouped them by their genres

The top 5 movie genres were;

1. Comedy, Documentary, Fantasy
2. Documentary, Family, Musical
3. History, Family , Music
4. History, Sport
5. Game-show

## 4.2: Most popular genres

I used the same data set in 4.1 to calculate the most popular movie genre

I used the votes for the movies as people viewed then added them and grouped the results by genre

The top 5 most popular movies are;

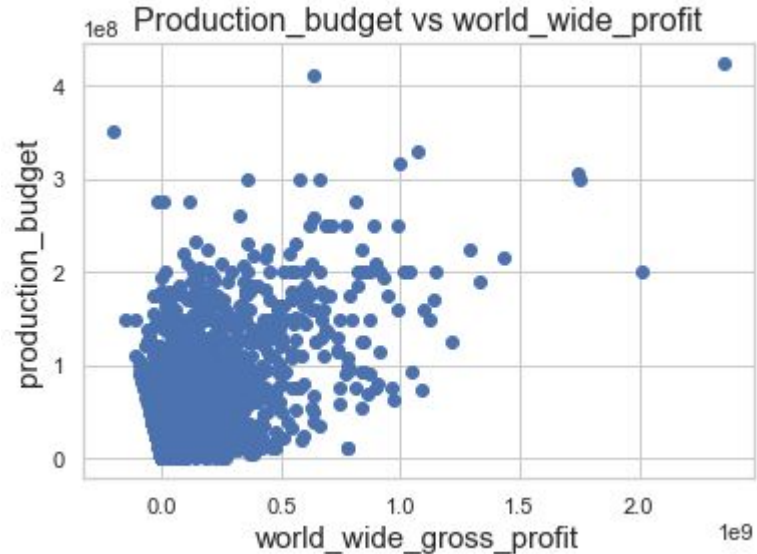
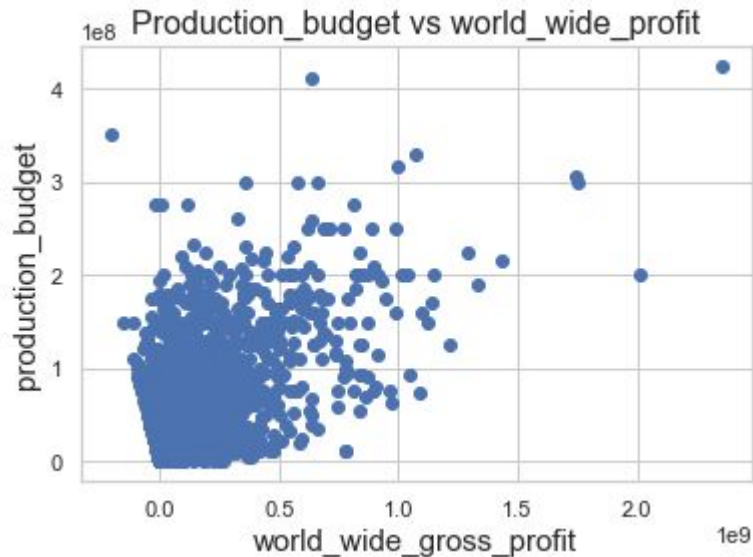
1. Action, Adventure, Sci-Fi
2. Action, Adventure, Fantasy
3. Adventure, Animation, Comedy
4. Drama
5. Comedy, Drama, Romance



### 4.3: Reception a movie gets worldwide and locally

From the chosen datasets 1 data set had movies production cost and the gross the movies made both locally and internationally

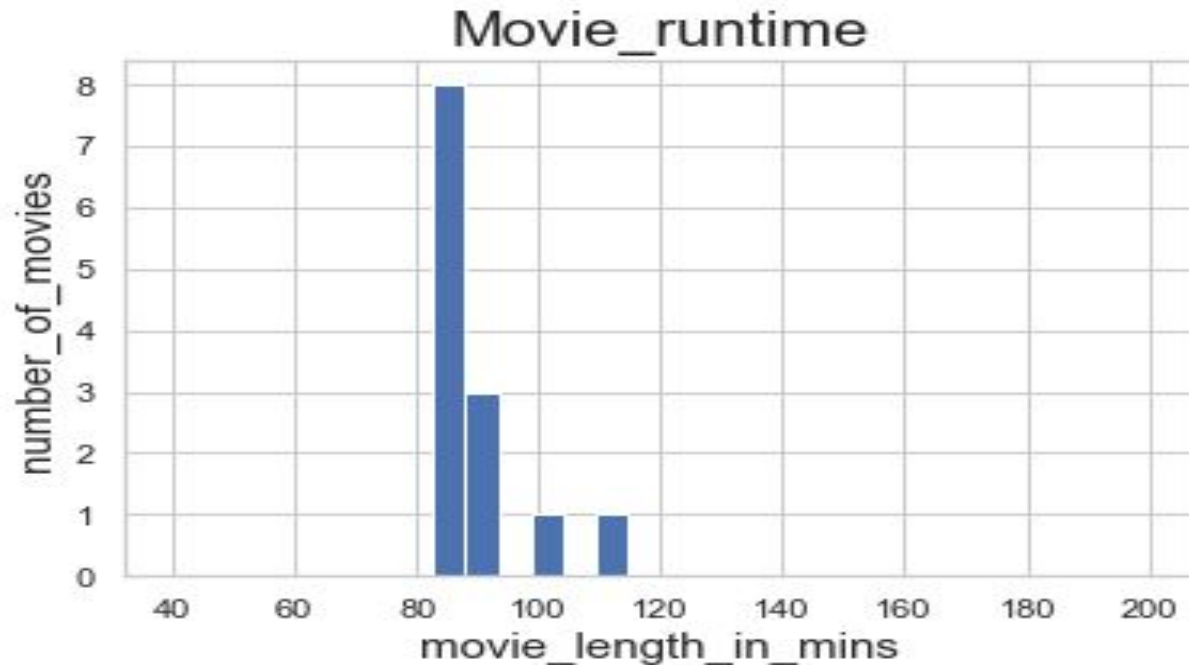
From the below diagram we can see that that movies with high budget do sell well worldwide compared to locally gross



#### 4.4: Average runtime for a movie

From the chosen data set, 1 dataset was showing the runtime for a movie.

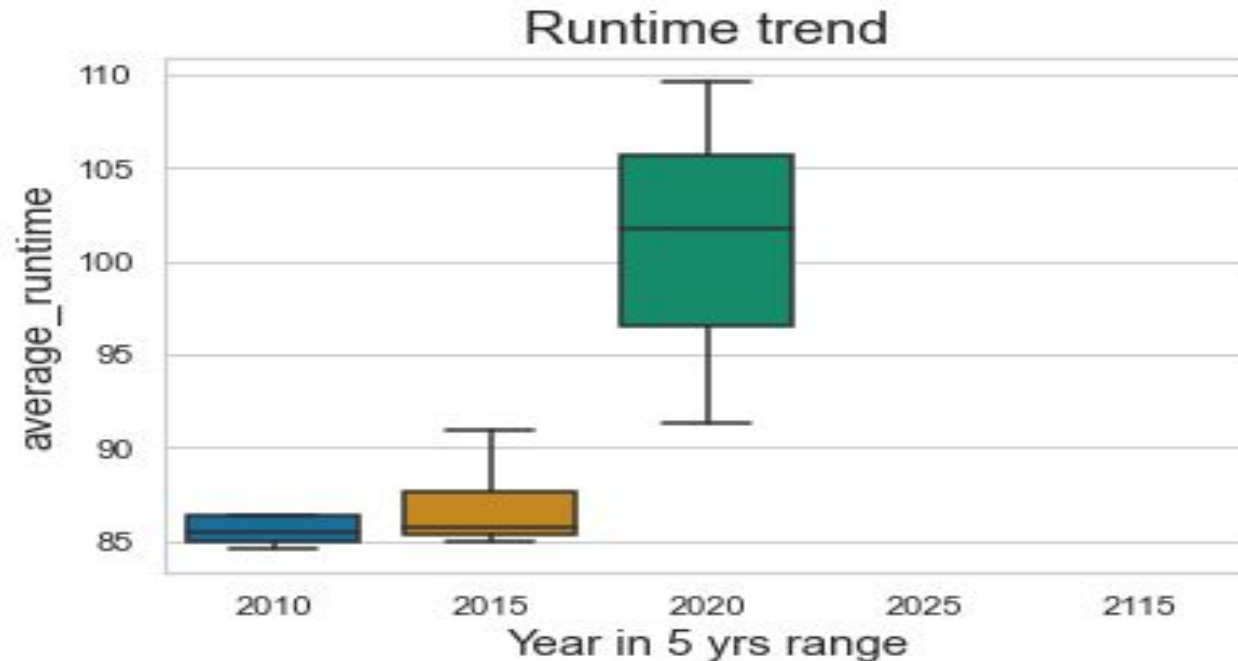
As shown below i first analysed the distribution of time in movies where i found that most movies runtime range between 80 mins and 90 mins.



### 4.4.1: The runtime trend

From the below chart we can also seeing as yeah moves the runtime of the movie also moves.

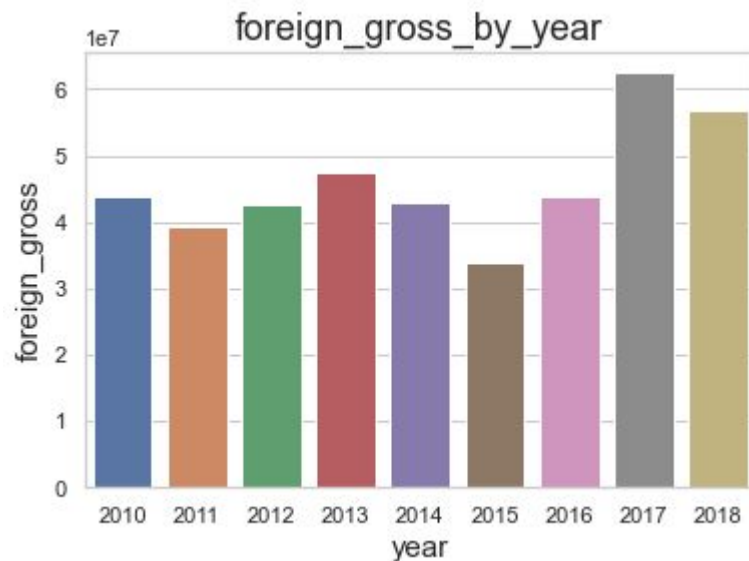
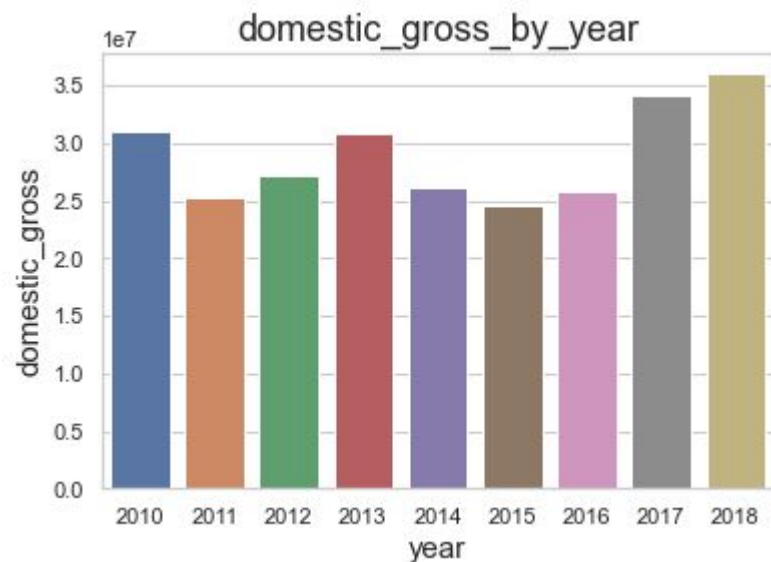
This might indicate that popularity in long running movies like movies with episodes.



## 4.5: Gross distribution locally and worldwide

Two datasets had columns showing the gross for a movie locally and world\_wide

I came up with the two bar charts to compare the distribution of gross locally and worldwide



## 4.6: The most popular movie

3 of the data sets had data indicating movies and their respective ratings

Through my analysis I found that the top 5 most popular movies;

1. Avengers:Infinity war
2. John Wick
3. Spider-Man: into the spider-verse
4. The Hobbit The Battle of The Five Armies
5. The Avengers

The reason for this analysis is that you can get some ideas or features such as picture quality, actors, movie set up e.t.c that Microsoft can adapt to come up with the best competitive movies.



## **5. Summary**

From my analysis I have provided Microsoft with the answers to all of my objectives.

Microsoft adapting the solutions will result to a highly profitable competitive movie studio



## **6.Recommendation**

I recommend Microsoft to adapt the top movie genre to the new movie studio, average movie runtime and when they are making movies they should target the worldwide audience



THANK YOU!!!

