

Final project in Module 1

Guofa Shou

Self-paced

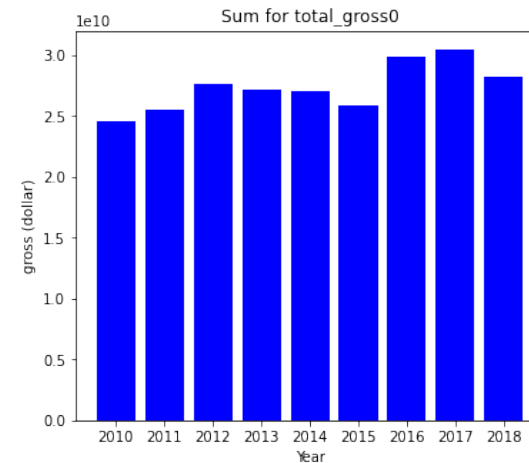
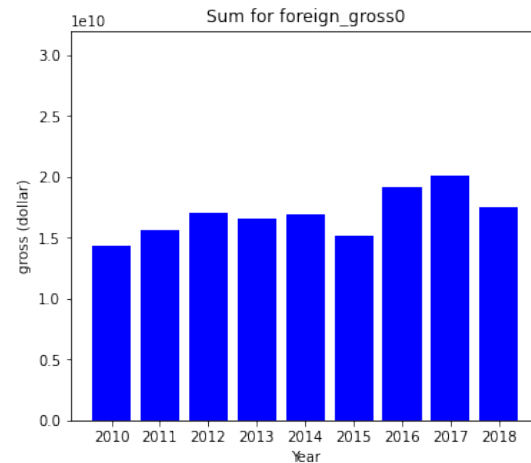
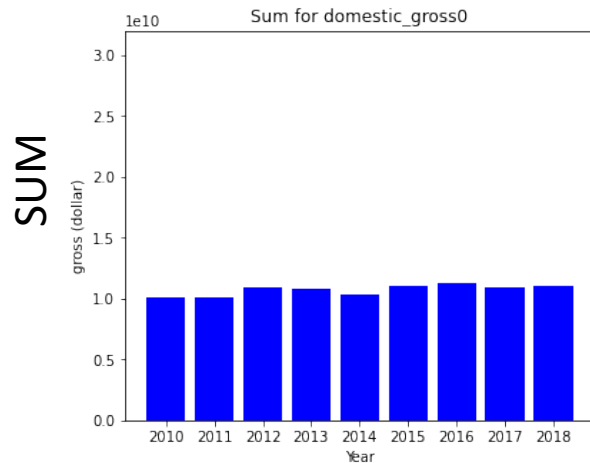
Outline

- The aim of the project:
 - To provide suggestions regarding what type of films should be created in terms of box office in the newly found Microsoft studio
- Three questions are addressed sequentially by data analyses to fulfill the aim:
 - 1) What's the general performance of movie industry in terms of box office for recent years?
 - 2) What types of films have large number of box office?
 - 3) What types of films have good ratings?
- Conclusion

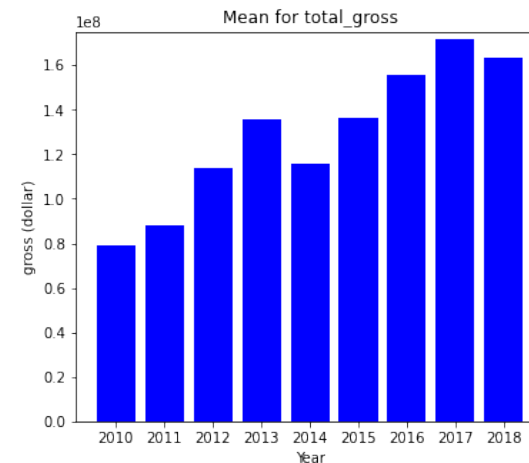
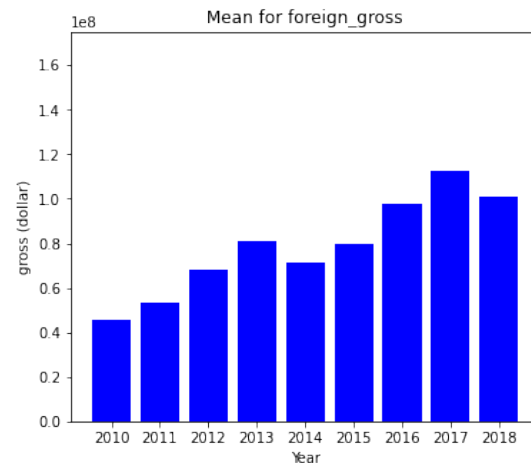
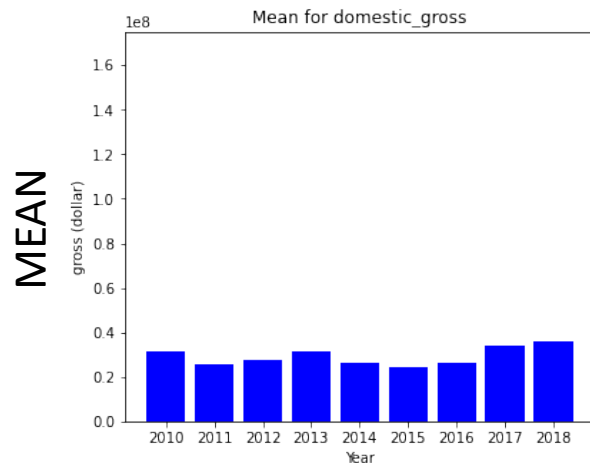
Q1: What's the general performance of movie industry in terms of box office for recent years?

- Data sources: bom.movie_gross.csv and tn.movie_budgets.csv
- Two summarized figures:
 - 1) Gross values: sum and mean values of the domestic, foreign, and total gross for all types of films in the years between 2010 and 2018
 - 2) Profits: Sum and mean values of the worldwide and domestic profit

1) Gross values: sum and mean of selected films between 2010-2018



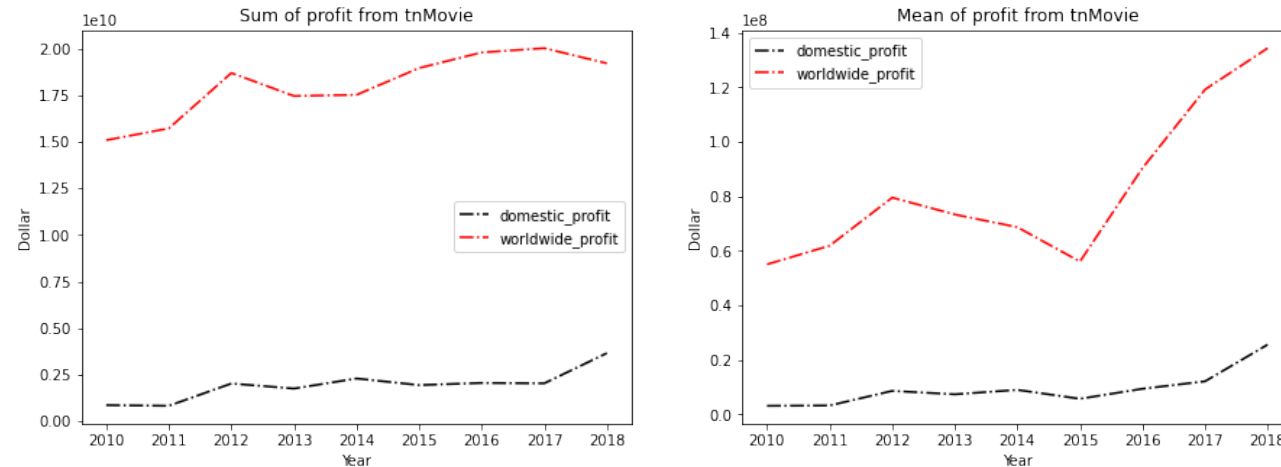
- From sum: domestic gross is stable, but foreign gross has an increased trend. Therefore, the worldwide gross has an increased trend



- From mean: similar patterns as for sum values, i.e., domestic market is stable, but foreign gross has an increased trend, which contributes to the increased trend in worldwide

2) Profit values: sum and mean profit values of selected films between 2010-2018

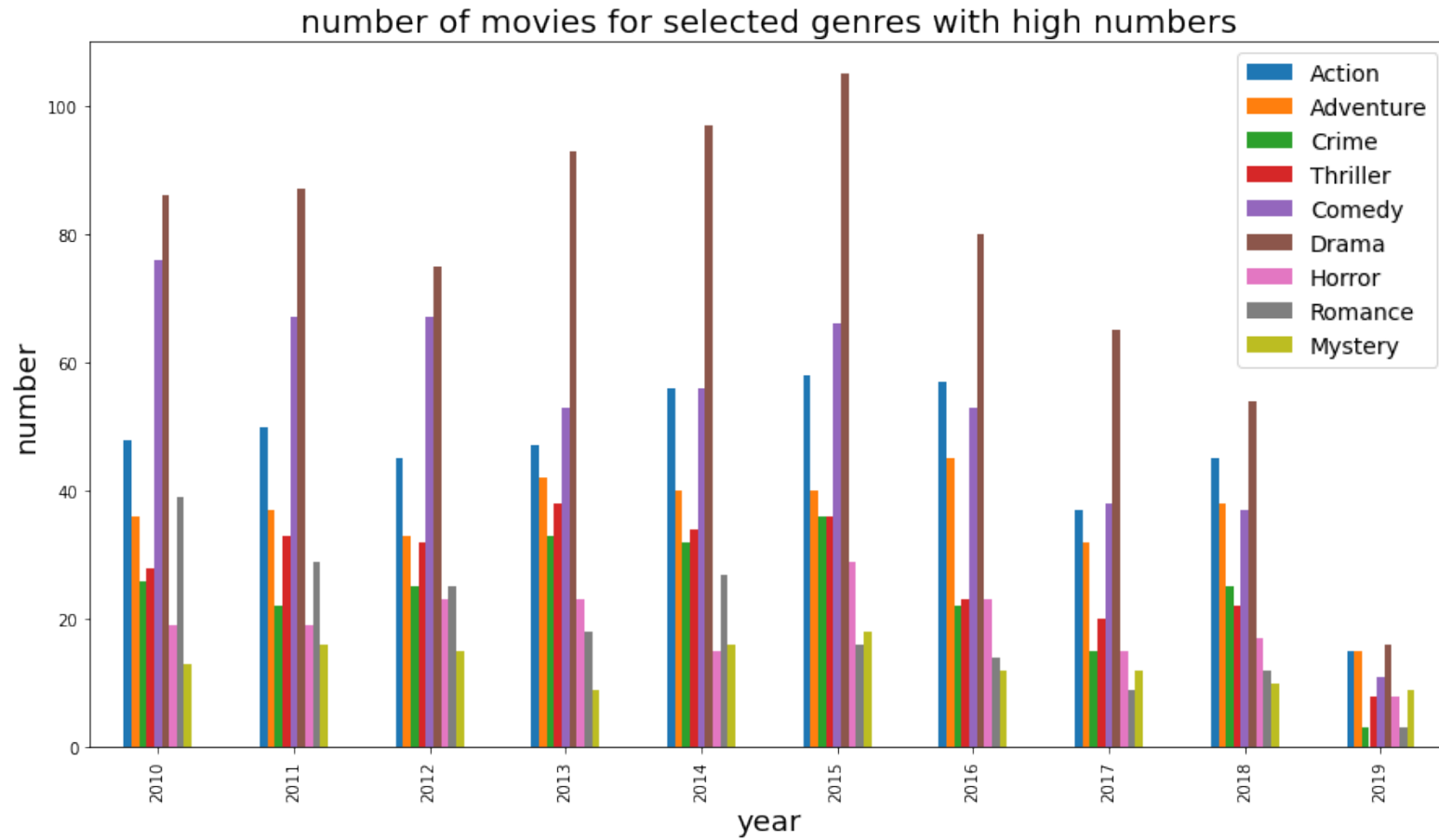
Sum and mean profit values



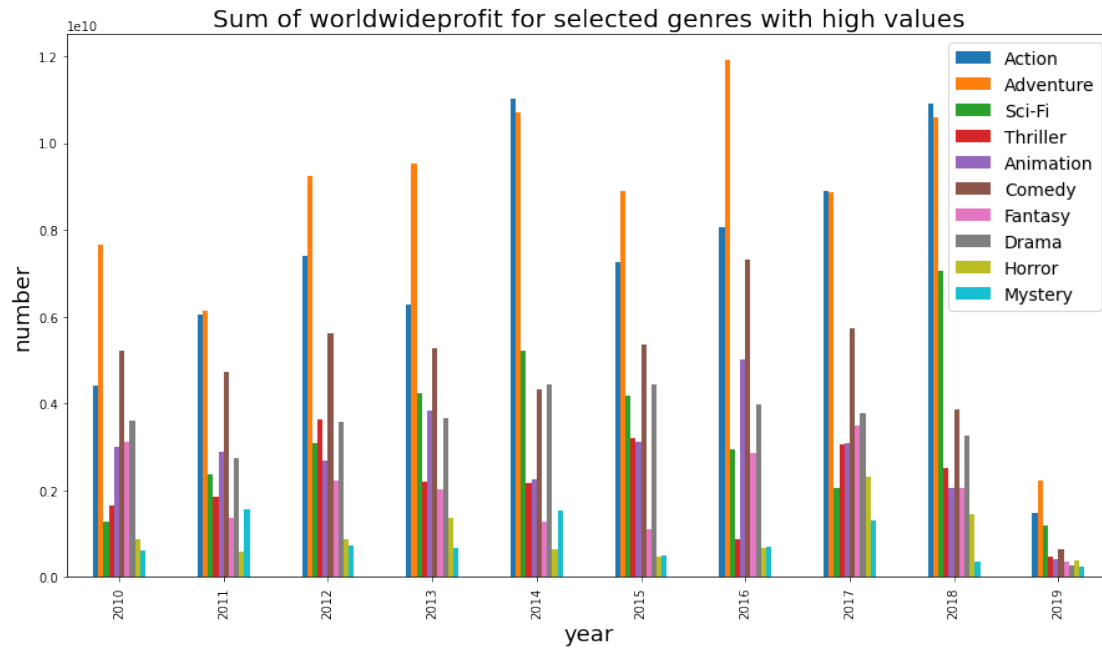
- Both sum and mean profit values indicate an increased trend for both domestic and worldwide profits
 - There is a trough in 2015
-
- Answer to Q1: The film industry is still developing very well, and lots of profits could be achieved, especially from foreign market

Q2: What types of films have large number of box office?

- Data sources: imdbTitleBasics and tnMovie_Budgets
- Three summarized figures:
 - 1) Number of movies for different types
 - 2) Sum of worldwide profit for selected genres with high values
 - 3) Mean of worldwide profit per film for selected genres

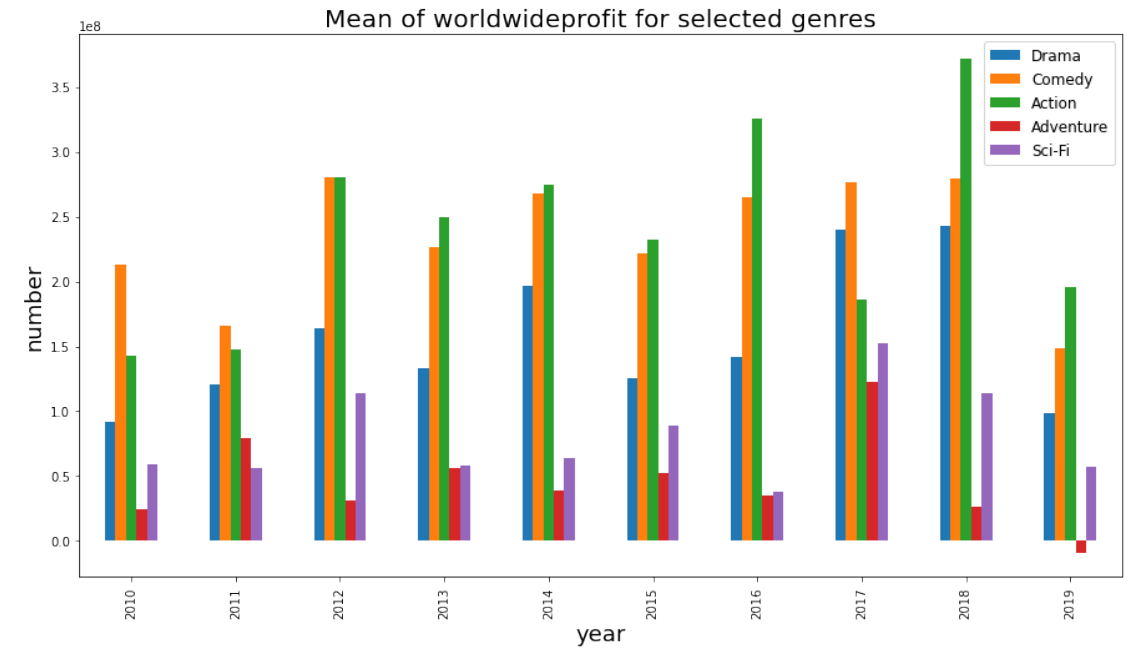


Drama, Comedy, Action, Adventure are the top four categories in terms of number of movies released



Adventure, Action, Comedy, and Sci-Fi, are the top 4 categories in terms of sum of worldwide profit

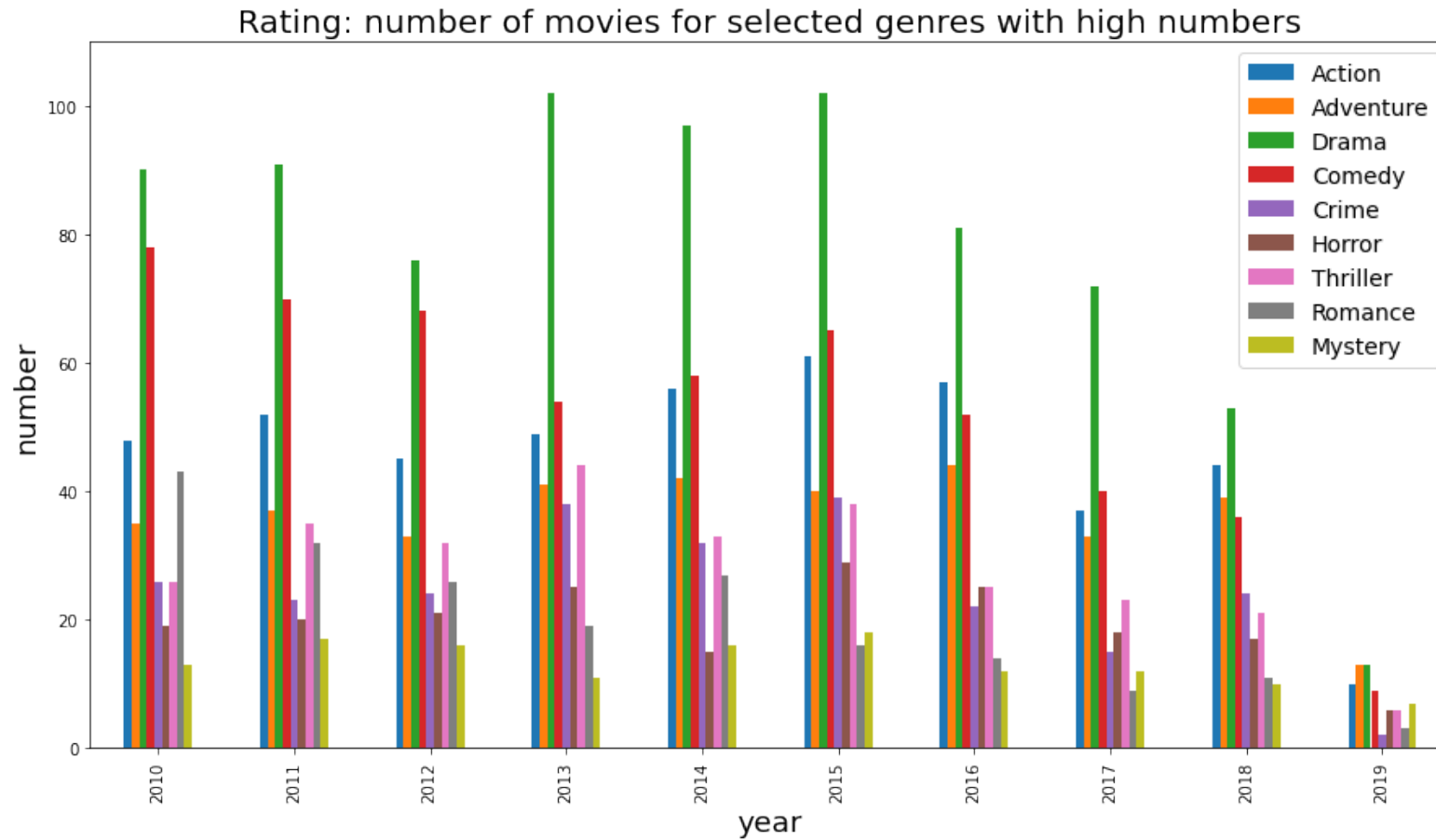
- Answer to Q2: Action and comedy are two types of films that have high box office in recent years



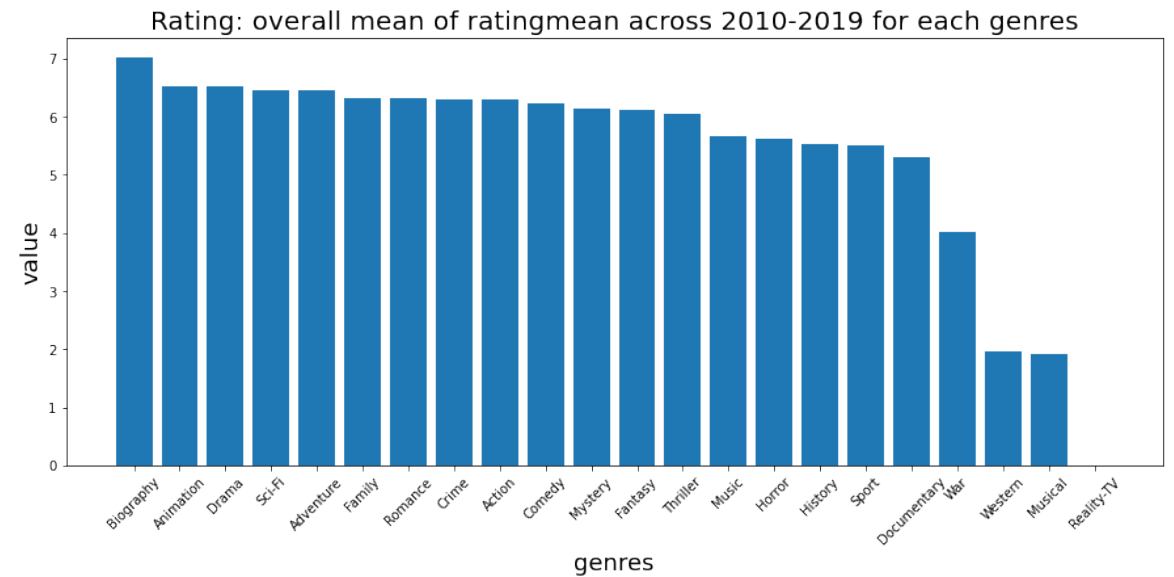
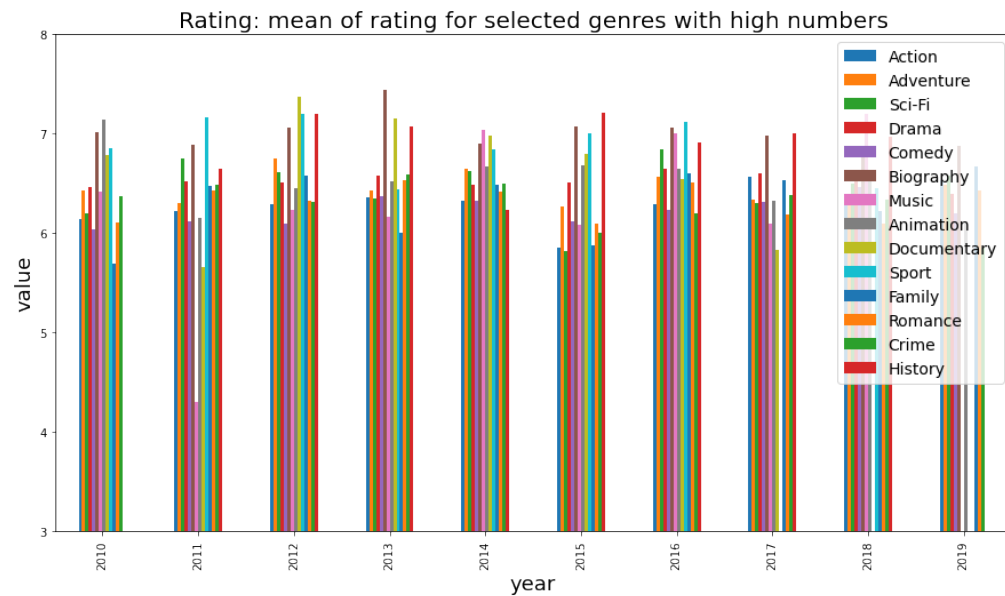
Action, comedy and Drama have higher mean values of worldwide profit

Q3: What types of films have good ratings?

- Data sources: imdbTitleRatings and the one generated in Q2
- Three summarized figures:
 - 1) Number of movies in the selected genres with high ratings
 - 2) Mean of rating for selected genres with high ratings
 - 3) Overall mean of mean rating for all genres across 2010 to 2019



Drama, Comedy, Action, Adventure are the top four categories in terms of number of movies



Answer to Q3: Biography, Animation, Drama, Sci-Fi, and Adventure are top five for the ratings

Conclusion

- Based on rating and box office data, adventure and action should be the focus in the newly found Microsoft studio