Data Set

TMDb movie data (cleaned from original data on Kaggle)

Data Cleaning and Wrangling

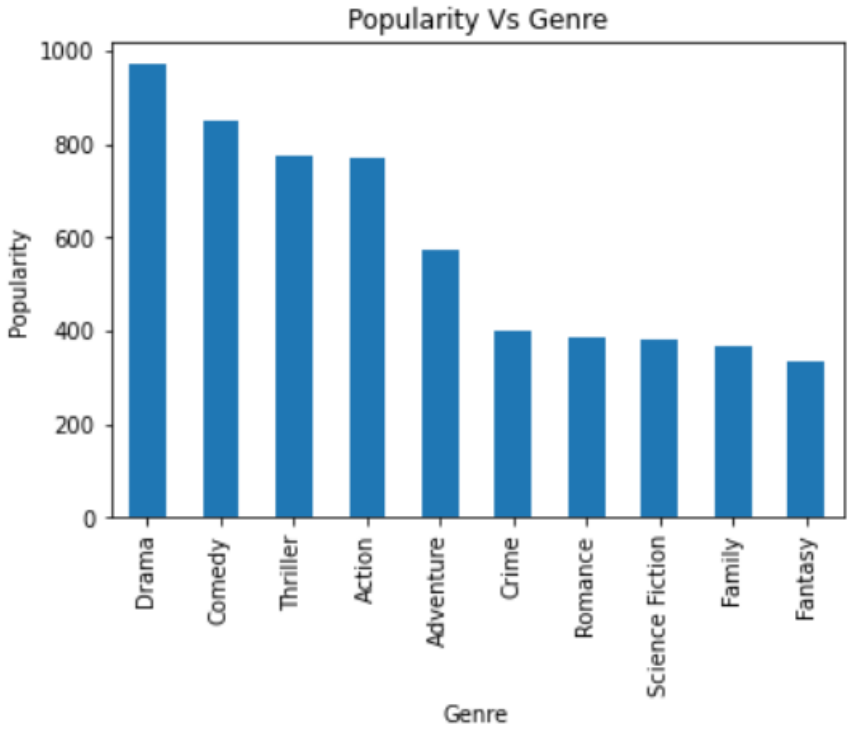
1) Checking for NAN values then dropping them because we can't fill ids or directors with mean or so

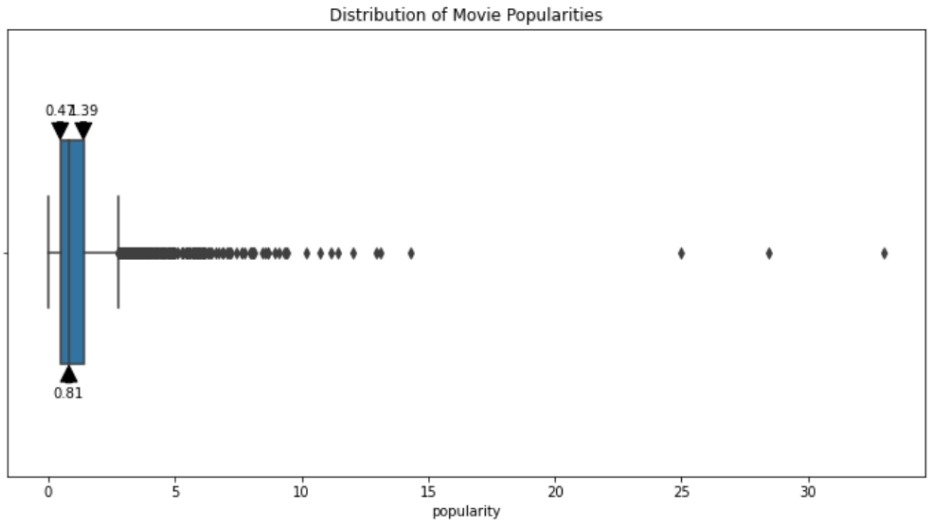
2) Dropping the columns that won't be needed like(Homepage, tagline, etc..)

3) removing the data that has zero revenue or zero budget

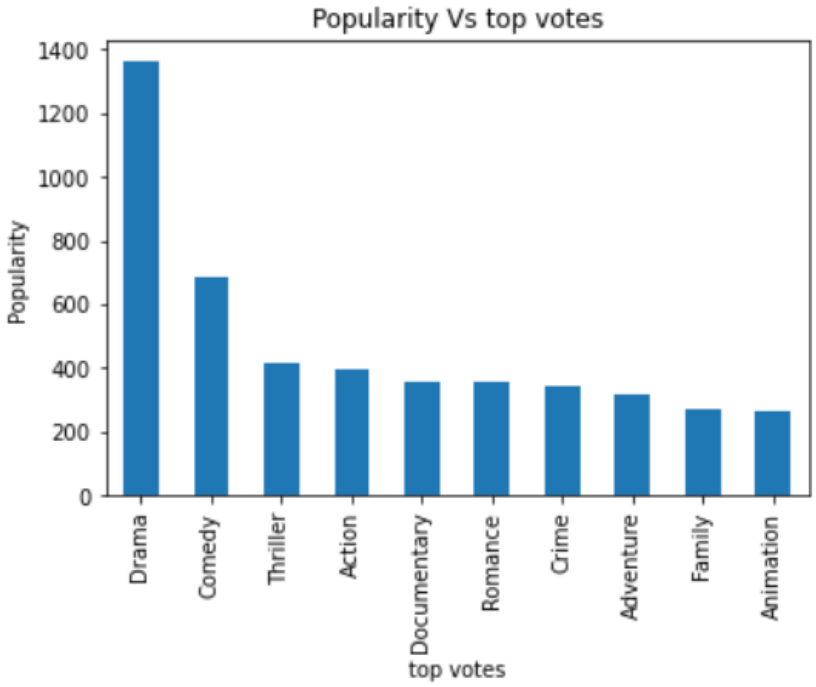
4) Splitting The Genres into different rows

Which genres are most popular?

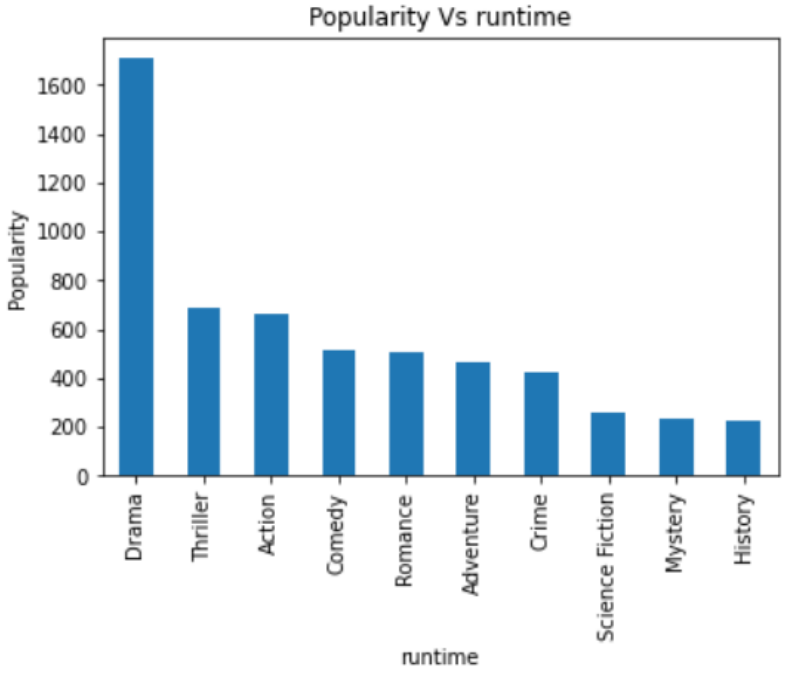




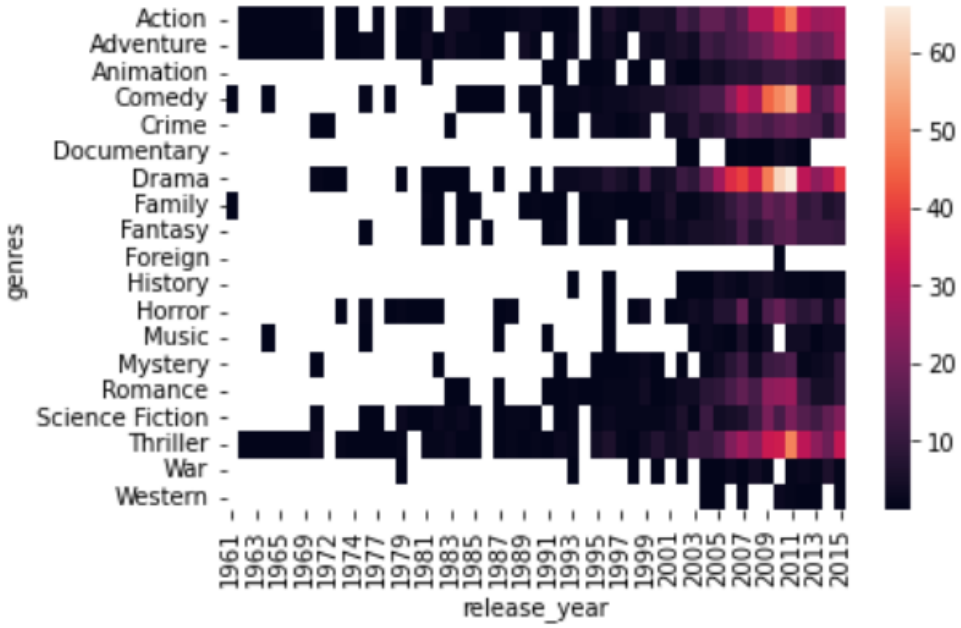
Which genres have the highest votes?



Which genres have the longest runtime?



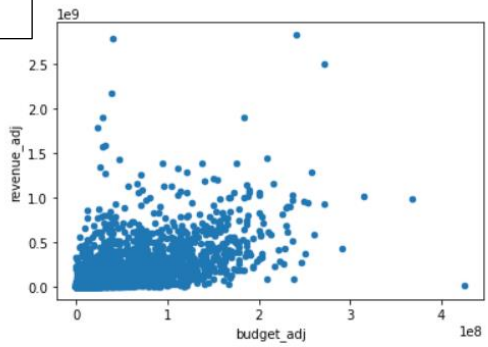
- ***Now You Can see that the popularity increases over the years***



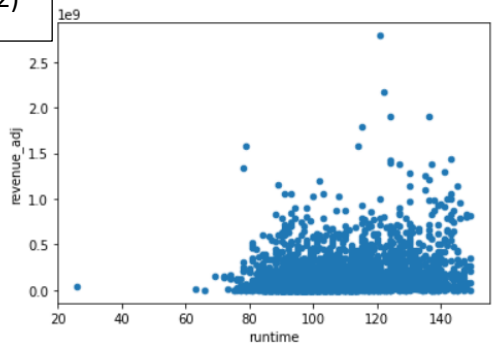
- ***The most popular genres are Drama, Comedy and Thriller based on the heat***

- ***What kinds of properties are associated with movies that have high revenues?***
    1. ***Relation between Budget and Revenue***
    2. ***Runtime and Revenue***
    3. ***Popularity and Revenue***
    4. ***Vote Average and Revenue***

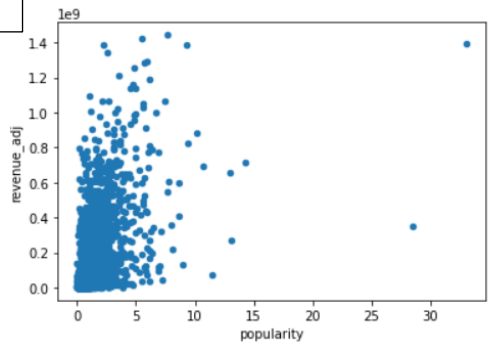
1)



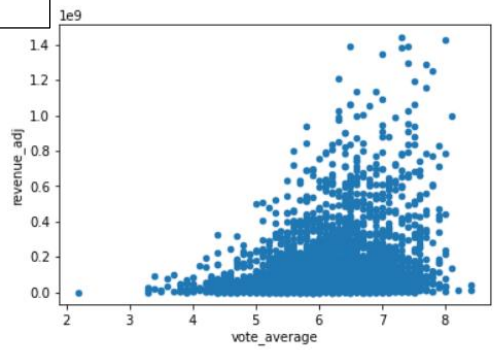
2)

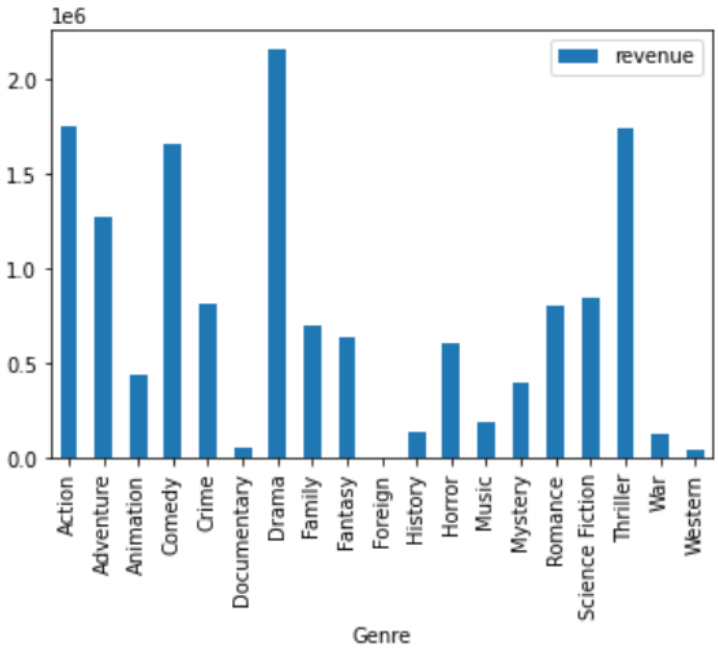


3)



4)





The Revenue Vs Movie Genres

Conclusions

(1): *Finally, It is obvious that the most popular genres are comedy and Drama*

(2): *The most voted and longest runtime genres are Drama and Thriller*

(3): *There is a positive correlation between runtime and revenue of the movie*

(4): *There is a positive correlation between vote average and revenue of the movie*

(5): *The Highest Revenue per Genre is for Drama*

Recommendation:

We could make a detailed analysis for the most popular (repeated) actor of every popular cast we can gather more data about their social media accounts, their last work, Their salary, etc..

Also to get an accurate correlation for the revenue we should get the exact budget for every movie including(salaries, promotions, production needs, etc..)

References

https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.loc.html

https://stackoverflow.com/questions/45926230/how-to-calculate-1st-and-3rd-quartiles

https://www.youtube.com/watch?v=0U9cs2V-Mqc