

I have analyzed the TMDb movies database. This data set contains information about 10,000 movies collected from The Movie Database (TMDb), including user ratings and revenue.

### **Questions:-**

Q1) Show the names top 10 highest budget movies and plot the bar graph original title vs budget?

Q2) Show the name of top 10 popular movies, plot the bar graph between original title and their popularity?

Q3) Which director has done most movies. Show the top 10 director in terms of number of movies on bar graph?

Q4) Show the relation between budget vs revenue. Analyze the graph and give a conclusion on high budget movies vs high revenue?

Q5) Show the popularity of movie according to budget. Analyze the graph and give conclusion on the relation between high budget movie and their popularity?

Q6) Group the revenue by month and show that which months have high revenues.? Which will help producer to decide the release date of movies. Plot the graph which will help producer easily to choose his release month

### **Description about investigation**

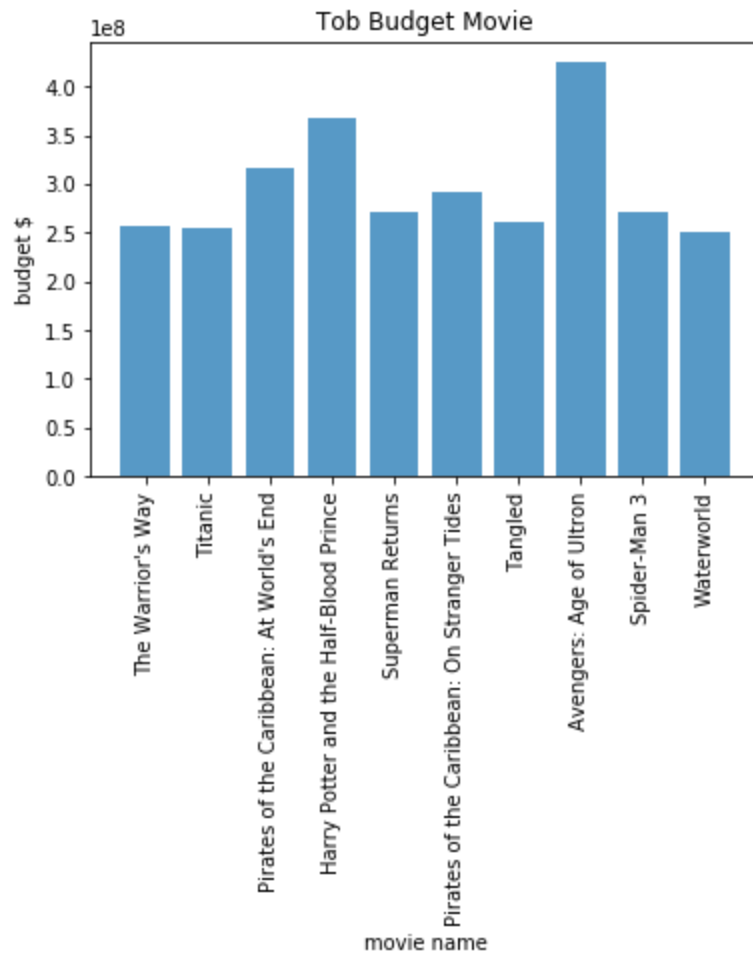
- I have collected the data from kaggle and I have tried to analyze some of the open question.
- I have investigated the relation between revenue and budget. I have also analyzed which is the best release month for high revenues.
- I have analyzed the popularity of movies and number of movies done by directors.
- I have tried to analyze which is good release month in terms of revenue

### **Data Wrangling**

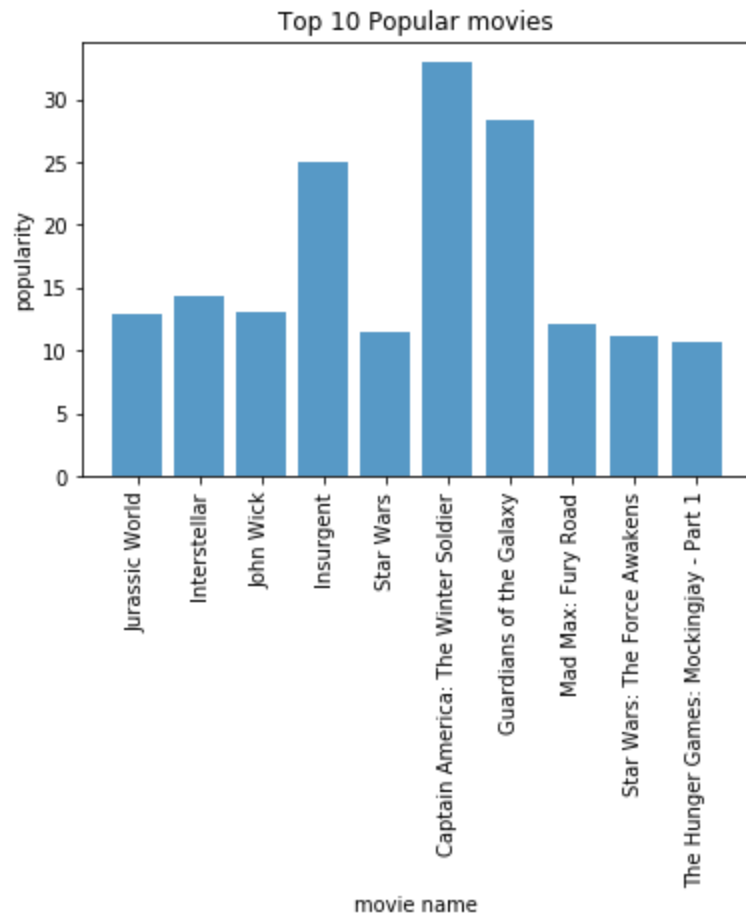
- In analyzing the release month vs revenue. I have first group the revenue by month. Then I have plotted the graph.
- In analyzing the top 10 high budget or revenue. First I have fetched the top movies based on budget or revenue respectively and then plotted the graph.

### **Summary and plot communicating your final result**

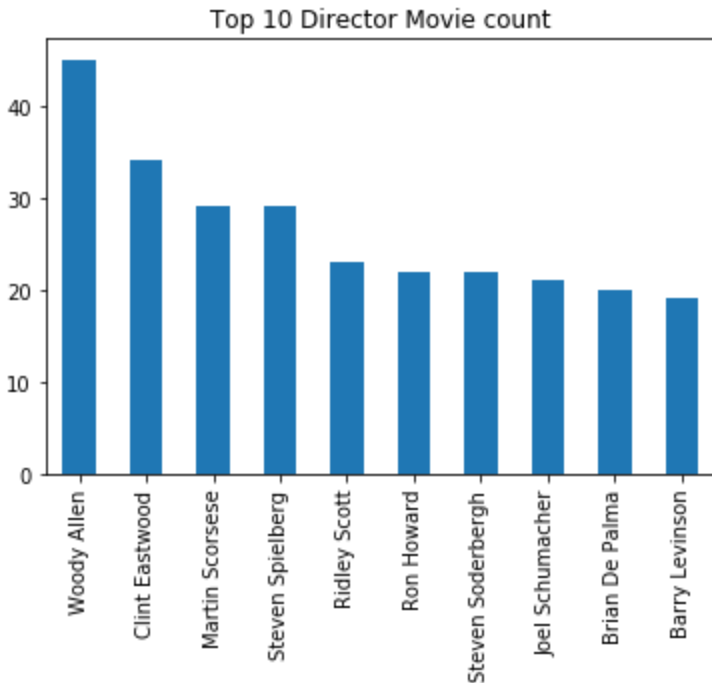
- Avenger-Age of Ultron is very high budget movie and second is Harry potter and half deadly prince.



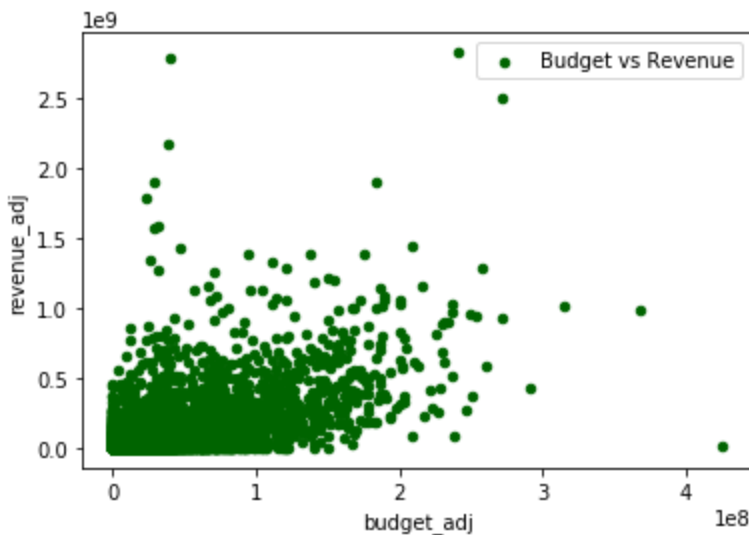
- Most popular movie is Captain America: The winter soldier and second most popular is guardian of galaxy.



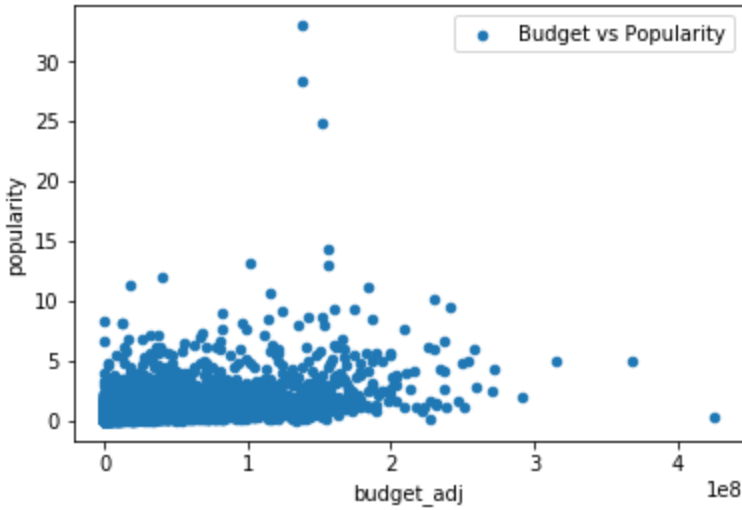
- Woody-Allen has done 40plus movies,which is the highest movie done by any director.Second is Clint Eastwood .He has done 30 plus movies



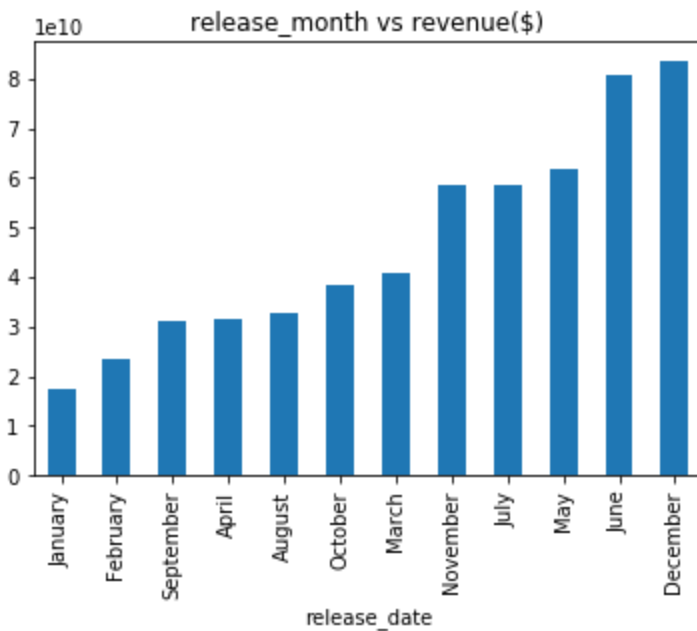
- We can see from the graph most of movies has good revenue according to budget. Some of the movies which has high budget. But it has less revenue. Some of the movies having low budget and has very high revenue.



- Some of the average budget movies has very high popularity. Very few of the very high budget has less popularity.



- December and June has very high revenue. December is holiday time. June is summer vacation time. I think most of the high budget movie will try to release during this June or December



#### Limitation:-

- Some of high budget movie have low popularity. We can't exactly what was the reason behind it. It could be because of new Actor/Actress or new Director. It could be due to bad script

- Some of the medium budget movies have high popularity. We can't predict exact reason. It could be due to various reasons like famous Actor/Actress or famous director or may be holiday season or could be autobiographical movie or real event movies or any historical movies. It could be due to very good script. There could be many reasons for popularity of movies.

