

BIDV Project



Team:
Shreya (19csu299)
Tuhini (19csu416)

Topic: IMDB Ratings



ACKNOWLEDGEMENT

It is our pleasure to be indebted to various people, who directly or indirectly contributed to the development of this work and who influenced our thinking, behavior, and acts during study.

We express our sincere gratitude to Miss Akansha (Faculty), for providing an opportunity to learn and implement the Concepts of Tableau in different ways and on different Topics. Thanks for helping us on each step. We are thankful for her support, cooperation, and motivation provided to us during the learning lessons and theoretical parts of the machine learning and artificial intelligence.

We would also like to thank the almighty and our friends with whom we shared our day-to-day experience and received lots of suggestions that improved our quality of work.

Shreya (19csu299)

Tuhini (19csu416)

CONTENT

| | |
|-------------------------------|----|
| ACKNOWLEDGEMENT | 2 |
| INTRODUCTION | 4 |
| ABOUT PROJECT | 4 |
| ABOUT DATASET..... | 5 |
| PROJECT LINK | 6 |
| ANALYSIS AND PREDICTION | 7 |
| CONCLUSION | 21 |
| BIBLIOGRAPHY | 21 |

INTRODUCTION

IMDb is an online database of information related to films, television programs, home videos, video games, and streaming content online – including cast, production crew and personal biographies, plot summaries, trivia, ratings, and fan and critical reviews. Most data in the database is provided by volunteer contributors. The site enables registered users to submit new material and edits to existing entries. Users with a proven track record of submitting data are given instant approval for additions or corrections to cast, credits, and other demographics of media product and personalities. However, image, name, character name, plot summaries, and title changes are supposedly screened before publication, and usually take between 24 and 72 hours to appear. Users are also invited to rate any film or TV series on a scale of 1 to 10

OUR PROJECT

Our Project is based on IMDB scores, ratings, Facebook likes of movie, director, actors, genre In this project we are building Dashboard and Stories to explore IMDb(Internet Movie Database) Top 5000 movies. We had built an interactive dashboard with the help of global filters and dashboard actions so that we can explore data and answer some own questions. With the help of Story we are trying to answer a few of the common questions that can be thought by anyone.

ABOUT DATASET

Our dataset contains 1 csv file named as movie_metadata containing string, int and float datatypes and are totally different from one another. Dataset has been collected / downloaded from Kaggle.

More about dataset:

Contains data in 29 sections i.e.

- a. color
- b. Director_Name
- c. num_critic_for_reviews
- d. duration
- e. cast_total_facebook_likes
- f. director_facebook_likes
- g. actor_3_facebook_likes
- h. actor_2_facebook_likes
- i. actor_1_facebook_likes
- j. gross
- k. genres
- l. actor_1_name
- m. movie_title
- n. num_voted_users
- o. actor_3_name
- p. plot_keywords
- q. facenumber_in_poster
- r. movie_imdb_link
- s. num_user_for_reviews
- t. language
- u. country
- v. content_rating
- w. budget
- x. title_year
- y. actor_2_facebook_likes
- z. imdb_score
- aa. aspect_ratio
- bb. movie_facebook_likes

| movie_metadata (1) | | | | | | | | | | | | | | | | | | | | | | | | | | |
|--------------------|-------------|----------|----------|----------|----------|-------------|----------|----------|----------|-----------------------|-----------------------|-----------|-------------|-------------|-----------|--------------|--|-----------|----------|----------|-----------|----------|------------|-------|--|--|
| color | director | num_crit | duration | director | factor_3 | factor_2 | factor_1 | gross | genres | actor_1 | movie_title | num_votes | cast_total | actor_3 | num_votes | plot_key | movie_im | num_votes | language | country | content_r | budget | title_year | actor | | |
| Color | James Car | 723 | 178 | 0 | 855 | Joel Davic | 1000 | 7.61E+08 | Action | Ac | CH Poun Avatar | 886204 | 4834 | Wes Studi | 0 | avatar | http://www | 3054 | English | USA | PG-13 | 2.37E+08 | 2009 | | | |
| Color | Gore Vert | 302 | 169 | 563 | 1000 | Orlando B | 40000 | 3.09E+08 | Action | Ac | Johnny De Pirates of | 471220 | 48350 | Jack Davel | 0 | goddies | http://www | 1238 | English | USA | PG-13 | 8E+08 | 2007 | | | |
| Color | Sam Menk | 602 | 148 | 0 | 161 | Rory Kinni | 11000 | 2E+08 | Action | Ac | Christoph Spectra | 275868 | 11700 | Stephanie | 1 | bomb leg | http://www | 994 | English | UK | PG-13 | 2.45E+08 | 2015 | | | |
| Color | Christoph | 813 | 164 | 22000 | 23000 | Christian f | 27000 | 4.48E+08 | Action | Th | Tom Hard-The Dark | 1144337 | 106759 | Joseph Gc | 0 | deception | http://www | 2701 | English | USA | PG-13 | 2.5E+08 | 2012 | | | |
| Color | Doug Walker | | | 131 | | Rob Walki | 131 | | Document | Doug Wall Star Wars: | | 8 | 143 | | 0 | | http://www.imdb.com/title/tt5289954/?ref=fn_tt_1 | | | | | | | | | |
| Color | Andrew St | 462 | 132 | 475 | 530 | Samantha | 640 | 73058679 | Action | Ac | Daryl Sabi John Cart | 212204 | 1873 | Polly Wall | 1 | alien | http://www | 738 | English | USA | PG-13 | 2.64E+08 | 2012 | | | |
| Color | Sam Raim | 392 | 156 | 0 | 4000 | James Fra | 24000 | 3.37E+08 | Action | Ac | J.K. Simmi Spider-Me | 383056 | 46055 | Kirsten Du | 0 | sandman | http://www | 1902 | English | USA | PG-13 | 2.58E+08 | 2007 | | | |
| Color | Nathan Gr | 324 | 100 | 15 | 284 | Donna Mu | 799 | 2.01E+08 | Adventur | Brad Garri Tangled | 294810 | 2036 | M.C. Gaini | 1 | 17th cent | http://www | 387 | English | USA | PG | 2.6E+08 | 2010 | | | | |
| Color | Joss Whee | 635 | 141 | 0 | 19000 | Robert Do | 26000 | 4.59E+08 | Action | Ac | Chris Hem Avengers: | 462669 | 92000 | Scarlett Jc | 4 | artificial i | http://www | 1117 | English | USA | PG-13 | 2.5E+08 | 2015 | | | |
| Color | David Yate | 375 | 153 | 282 | 10000 | Daniel Rai | 25000 | 3.02E+08 | Adventur | Alan Ricki Harry Pott | 321795 | 58753 | Rupert Gri | 3 | blood | http://www | 973 | English | UK | PG | 2.5E+08 | 2009 | | | | |
| Color | Zack Snyd | 673 | 183 | 0 | 2000 | Lauren Co | 15000 | 3.3E+08 | Action | Ac | Henry Cav Batman v | 371639 | 24450 | Alan D. Pu | 0 | based on | http://www | 3018 | English | USA | PG-13 | 2.5E+08 | 2016 | | | |
| Color | Bryan Sing | 434 | 169 | 0 | 903 | Marlon Br | 18000 | 2E+08 | Action | Ac | Kevin Spa Supermar | 240396 | 29991 | Frank Lan | 0 | crystal | http://www | 2367 | English | USA | PG-13 | 2.09E+08 | 2006 | | | |
| Color | Marc Forst | 409 | 106 | 395 | 393 | Mathieu F | 451 | 1.68E+08 | Action | Ac | Giancarlo Quantum | 330784 | 2023 | Rory Kinni | 1 | action her | http://www | 1243 | English | UK | PG-13 | 2E+08 | 2008 | | | |
| Color | Gore Veri | 313 | 151 | 563 | 1000 | Orlando B | 40000 | 4.23E+08 | Action | Ac | Johnny De Pirates of | 522040 | 48486 | Jack Davel | 2 | box office | http://www | 1832 | English | USA | PG-13 | 2.25E+08 | 2006 | | | |
| Color | Gore Veri | 450 | 150 | 563 | 1000 | Ruth Willi | | | | | | 181792 | 45757 | Tom Willi | 1 | horse | http://www | 711 | English | USA | PG-13 | 2.15E+08 | 2013 | | | |
| Color | Zack Snyd | 733 | 143 | 0 | 748 | Christoph | 15000 | 2.91E+08 | Action | Ac | Henry Cav Man of St | 548573 | 20495 | Harry Leni | 0 | based on | http://www | 2536 | English | USA | PG-13 | 2.25E+08 | 2013 | | | |
| Color | Andrew A. | 258 | 150 | 80 | 201 | Pierfranc | 22000 | 1.42E+08 | Action | Ac | Peter Dinl The Chor | 149922 | 22697 | Damla In | 4 | brother bi | http://www | 438 | English | USA | PG | 2.25E+08 | 2008 | | | |
| Color | Joss Whee | 703 | 173 | 0 | 19000 | Robert Do | 26000 | 6.23E+08 | Action | Ac | Chris Hem The Aven | 995415 | 87697 | Scarlett Jc | 3 | alien inva | http://www | 1722 | English | USA | PG-13 | 2.2E+08 | 2012 | | | |
| Color | Rob Marsh | 448 | 136 | 252 | 1000 | Sam Claffi | 40000 | 2.41E+08 | Action | Ac | Johnny De Pirates of | 370704 | 54083 | Stephen C | 4 | blackbear | http://www | 484 | English | USA | PG-13 | 2.5E+08 | 2011 | | | |
| Color | Barry Sonr | 451 | 106 | 188 | 718 | Michael St | 10000 | 1.79E+08 | Action | Ac | Will Smith Men in Bli | 268154 | 12572 | Nicole Sch | 1 | alien | http://www | 341 | English | USA | PG-13 | 2.25E+08 | 2012 | | | |
| Color | Peter Jack | 422 | 164 | 0 | 773 | Adam Bro | 5000 | 2.55E+08 | Adventur | Aidan Turi The Hobbi | 354228 | 9152 | James Nei | 0 | army | http://www | 802 | English | New Zeal | PG-13 | 2.5E+08 | 2014 | | | | |
| Color | Marc Web | 559 | 153 | 464 | 963 | Andrew G | 15000 | 2.62E+08 | Action | Ac | Emma Sto The Amaz | 451803 | 28489 | Chris Zylki | 0 | lizard | http://www | 1225 | English | USA | PG-13 | 2.3E+08 | 2012 | | | |
| Color | Ridley Soc | 343 | 156 | 0 | 738 | William H | 891 | 1.05E+08 | Action | Ac | Mark Add Robin Hoc | 211765 | 3244 | Scott Grirr | 0 | 1190s | http://www | 546 | English | USA | PG-13 | 2E+08 | 2010 | | | |
| Color | Peter Jack | 509 | 186 | 0 | 773 | Adam Bro | 5000 | 2.58E+08 | Adventur | Aidan Turi The Hobbi | 483540 | 9152 | James Nei | 6 | dwarf | http://www | 951 | English | USA | PG-13 | 2.25E+08 | 2013 | | | | |
| Color | Chris Weir | 251 | 113 | 129 | 1000 | Eva Green | 16000 | 70083519 | Adventur | Christoph The Golde | 149019 | 24106 | Kristin Soc | 2 | children | http://www | 666 | English | USA | PG-13 | 1.8E+08 | 2007 | | | | |
| Color | Peter Jack | 446 | 201 | 0 | 84 | Thomas Ki | 6000 | 2.18E+08 | Action | Ac | Naomi Wi King Kong | 316018 | 7123 | Evan Parki | 0 | animal na | http://www | 2618 | English | New Zeal | PG-13 | 2.07E+08 | 2005 | | | |

Project links

Story1-

https://public.tableau.com/app/profile/tuhini.dhankhar/viz/imdbproject_16378261647560/Story1

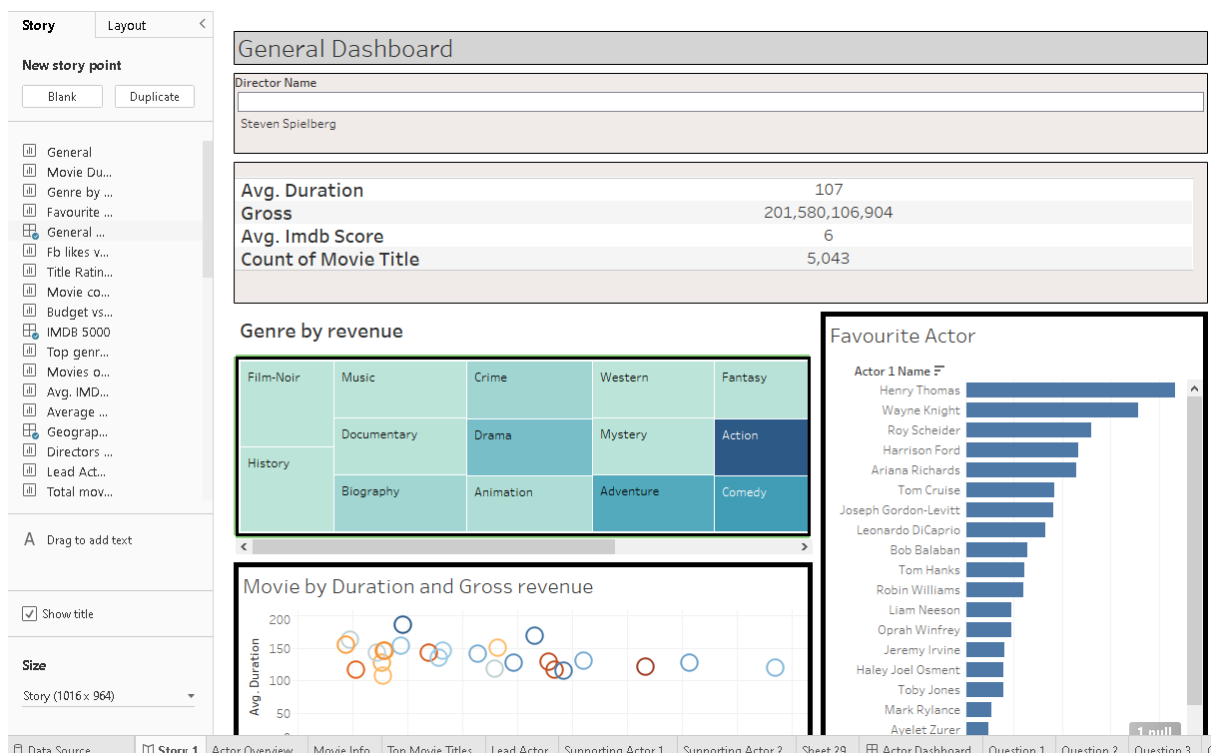
Story2 -

<https://public.tableau.com/app/profile/tuhini.dhankhar/viz/imdbquestions/Story2>

ANALYSIS AND PREDICTIONS

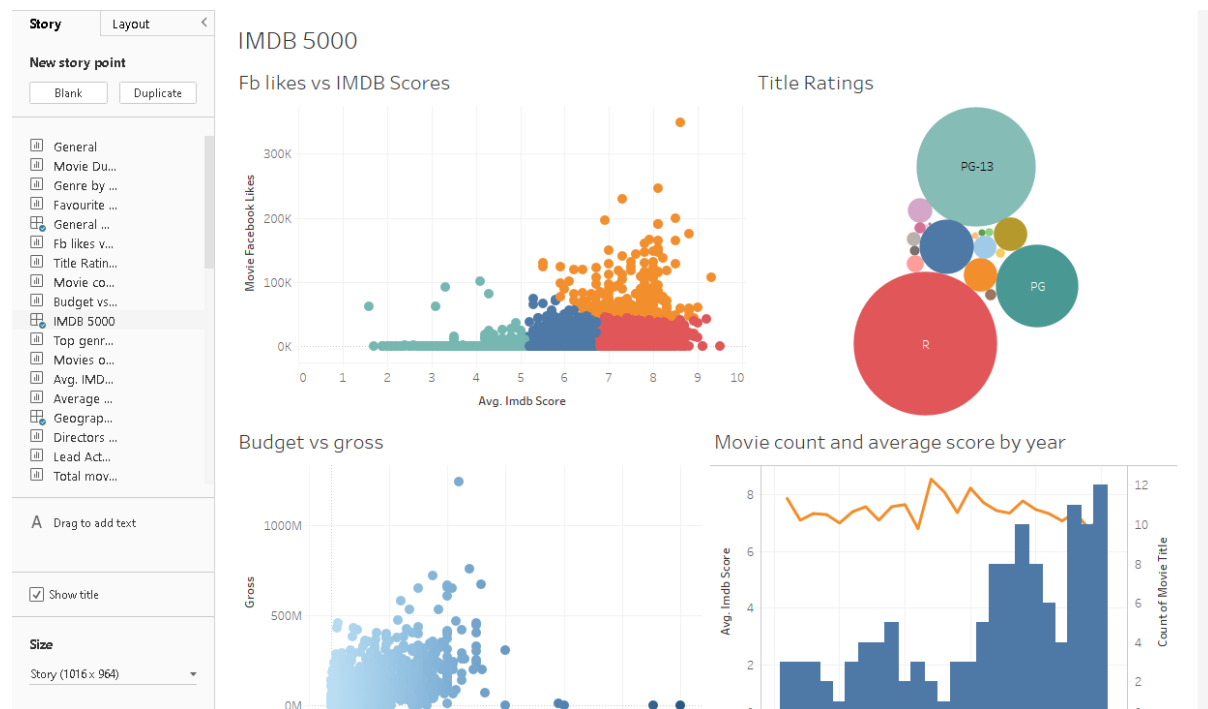
Story1

General Analysis



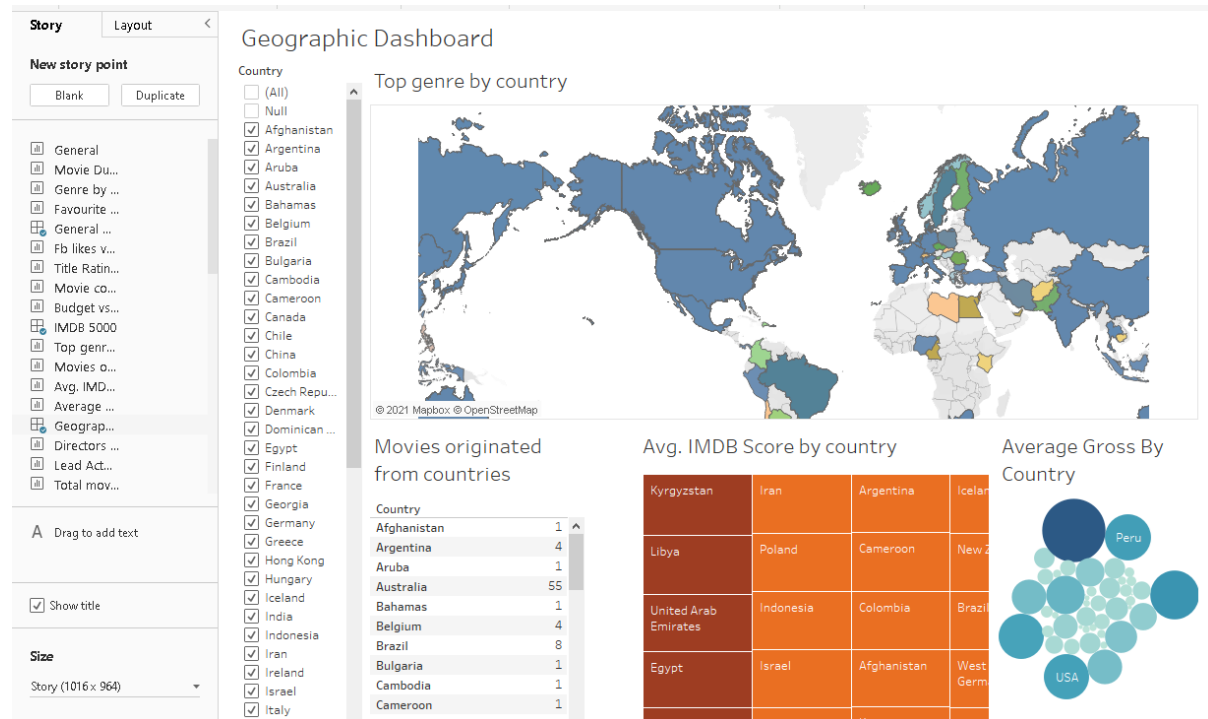
- This is the general analysis of the dataset where we are finding the avg duration of the movies their gross and total no. of movie by the Director chosen
- Genre by revenue determines the revenue generated by the different genres
- Movie by Duration and gross revenue determines the duration and the gross revenue relation with each other
- Favourite actor based on the gross

Ratings and scores Analysis



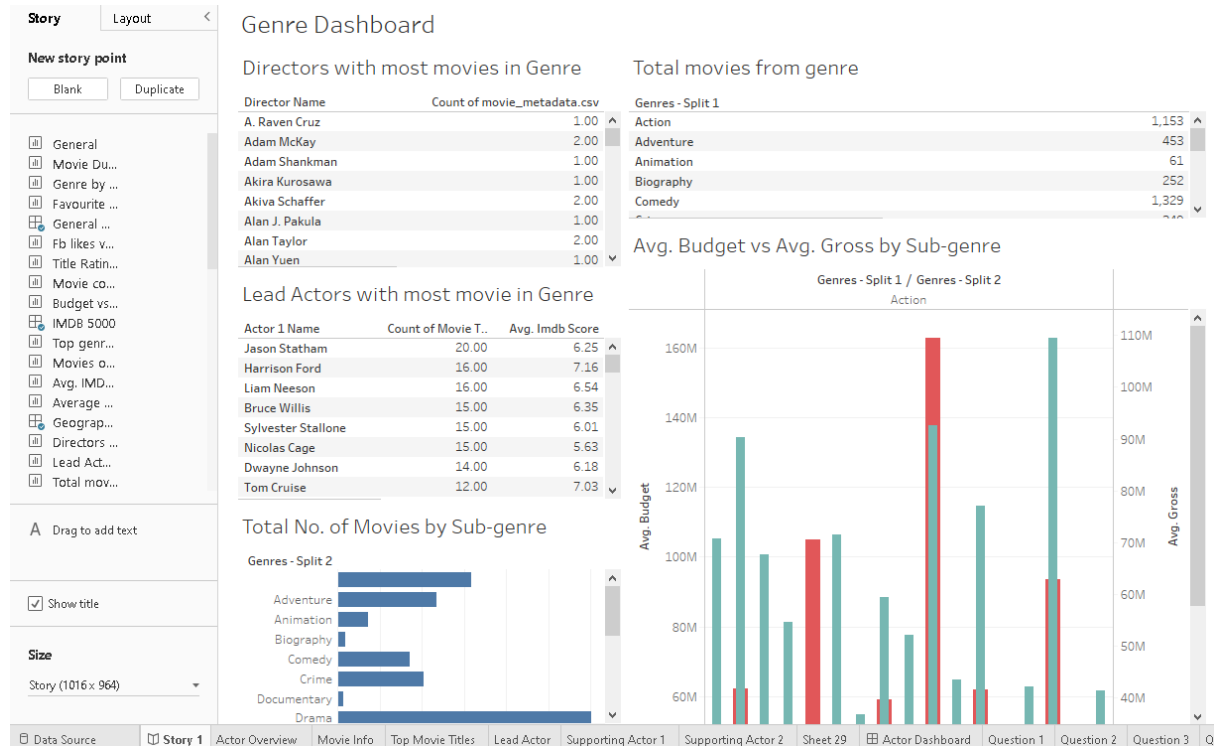
- In this dashboard we are we are analysing the ratings scores. In Fb likes vs IMDB score we are determining
- Title rating determines the content rating which is the most preferable or popular among the watchers
- Budget vs gross determines the relationship between the budget and gross of the movie with Avengers having low budget and highest gross and kites having highest budget and low gross .
- Movie count and average score by year conveys the count of the movie in the specific year and their average IMDB score. As the count of the movie increases their average IMDB ratings drops

Movies By Country Analysis



- This is general analysis of the movies, their ratings and gross based on the country
- Top genre by country determines the top genre of the country and the gross of that genre in that. Maximum countries have action as their top genre like Canada, Russia, India and gross of 718479434, 7268659, 9016849 respectively
- Movies originated from country direct the number of movies originated in the country. USA having the highest count of the originated movie and India having 34 movies.
- Avg. IMDB score tells the average IMDB score of the movie originated in that country with Egypt, United Arab Emirates etc. as their highest and New line, Peru etc. as their lowest.
- Average gross by country has New Zealand as their highest average gross country followed by Peru, Taiwan, USA

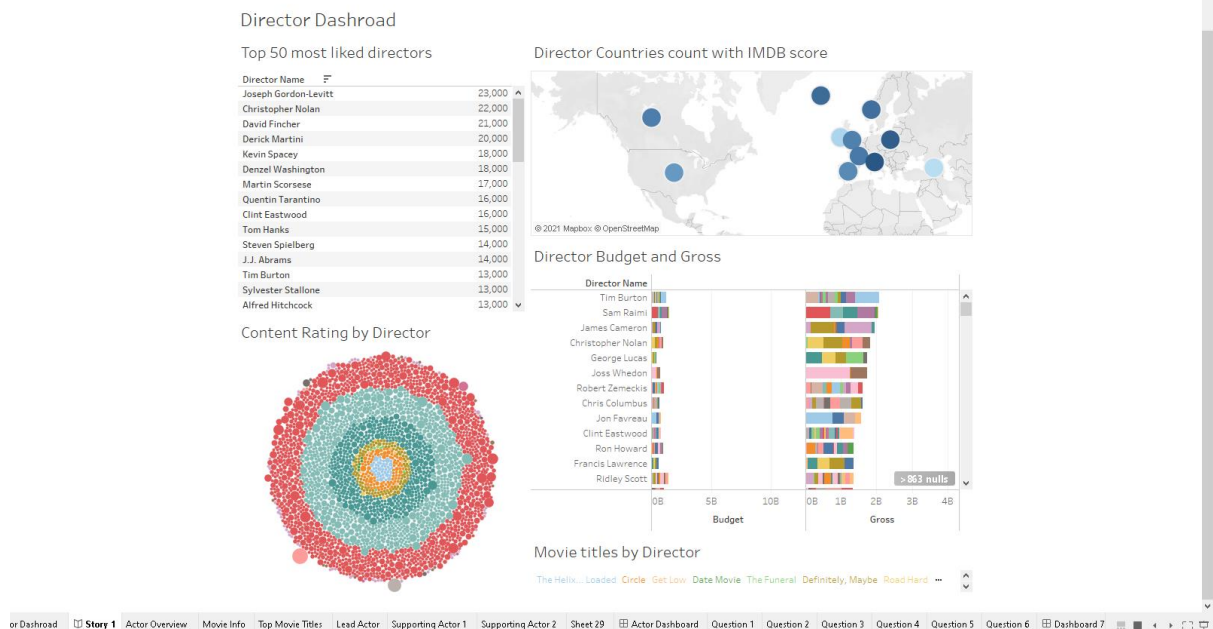
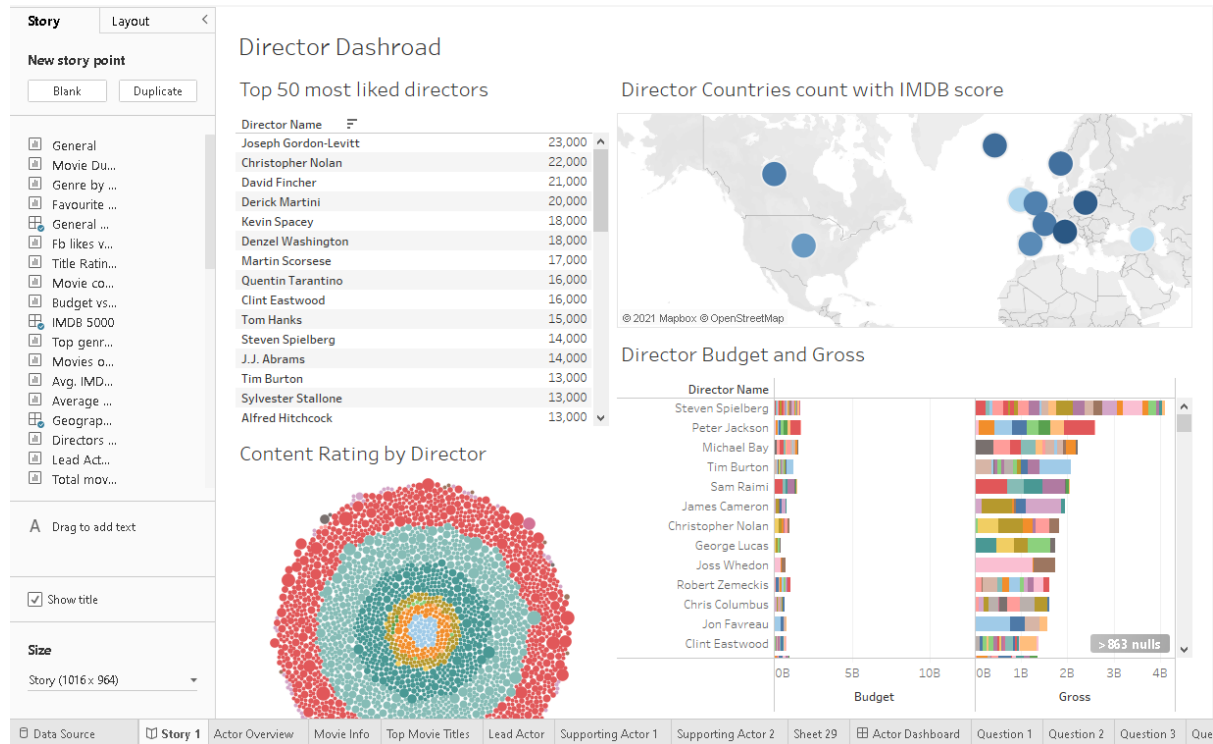
Movie Genres Analysis



Movie Genre Analysis dashboard has Director with most movie in genre, Total movies from genre, lead actors with most in genre, total no. of movies by sub- genre and avg. budget vs Avg. Gross by sub-genre

- Directors with movies in genre show Michael Bay as the Director of the highest count of Action genre.
- Lead actor with most movie in genre show the count of the movie and their average IMDB rating. In this we have Jason Statham as the highest count of the movie for the action genre
- A movie can have multiple genres, for the next we are determining the sub-genre that has highest count of the movie and drama in the subgenre category has the highest.
- Budget vs gross by sub-genre determines the relationship between the budget and gross of the genres of the movie. Action and Family is the least popular among the watchers as difference between the budget and the gross is the largest and it has quite a high budget and action and mystery is the most popular among the watchers

Director Stats

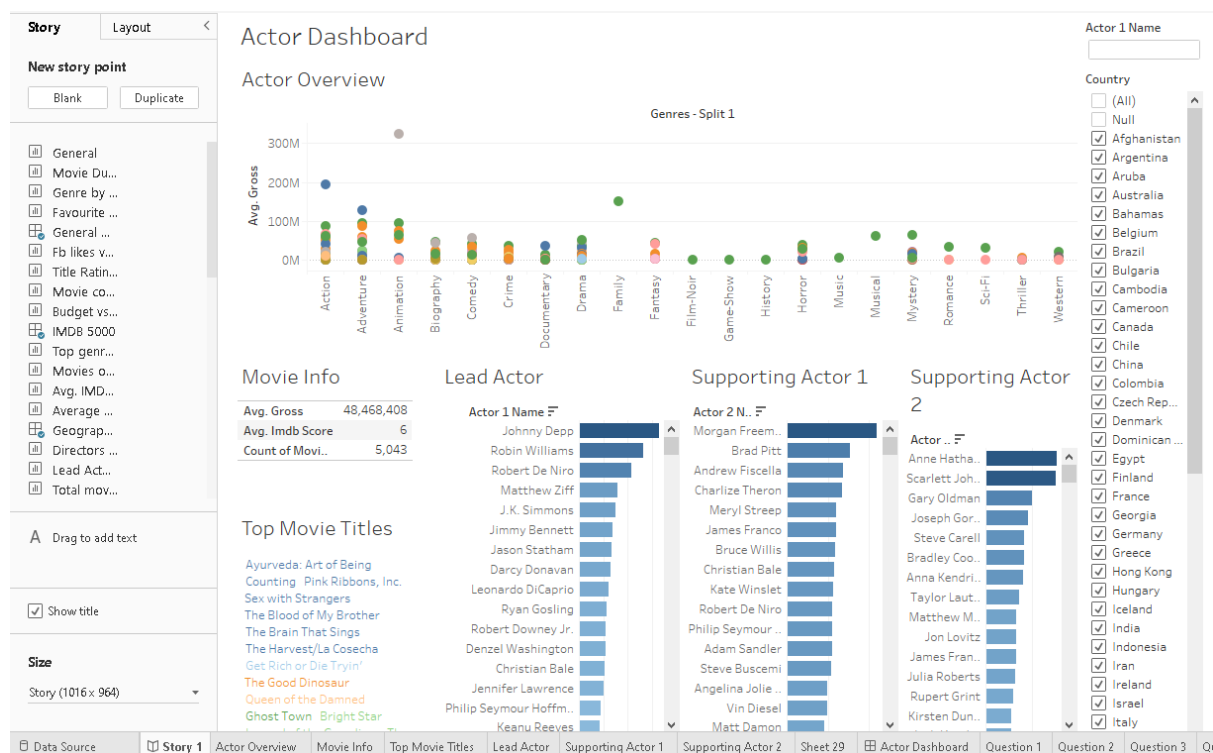


Director stats show the top 50 most liked directors, content rating by director, Director budget and gross

- In top 50 most liked directors we can see Josoph Gordon leviit as the most liked director based on the Facebook likes

- In Director count with IMDB score we can see the directors from different countries and their average IMDB score. Italy has the highest followed by Poland
- Content Rating by director shows us the content popular among the directors i.e R , Director, and their movie count
- Movie titles by director is a word cloud of the movies directed by the director.
- Director budget and gross depicts the movies directed by the director their budget and their gross.

Actor Analysis

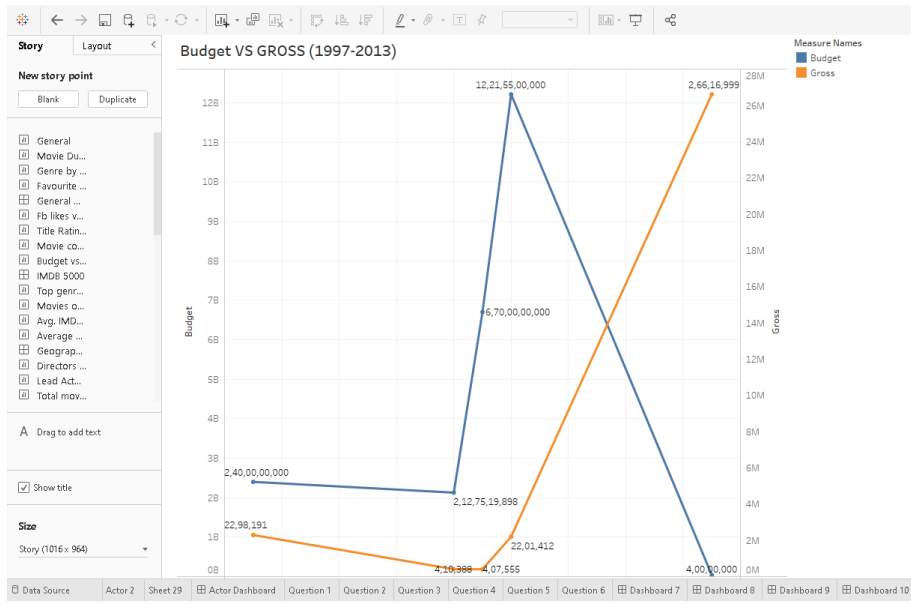


Actor Analysis dashboard has the actor overview, movie info, lead actor , supporting actor 1 and 2, top movie titles

- Actor overview depicts the country, genre and gross with animation being the most famous in Japan.
- Top Lead actor based on the Facebook likes is Johnny Depp, for supporting 1 it is Morgan Freeman, supporting 2 is Anne Hathaway, A movie with all these 3 actors will be super hit movie.
- Next we have is the top movie titles word cloud with their actor names.

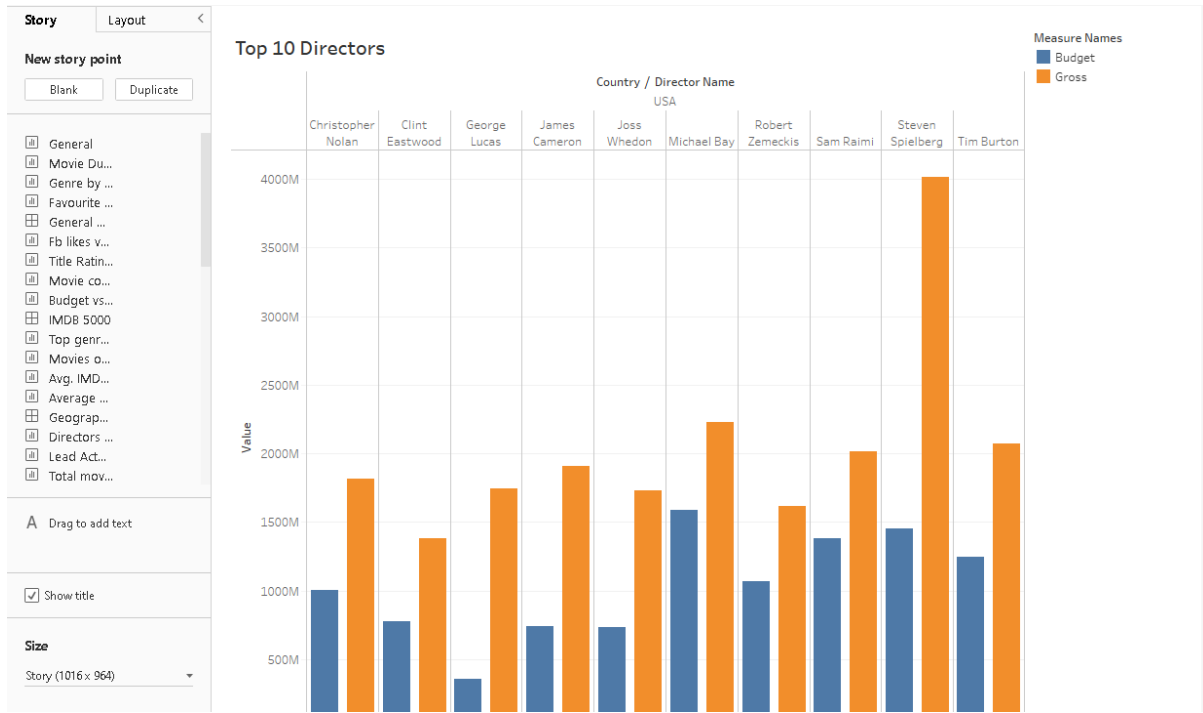
Story2 (Questions)

What is the growth of movie industry?



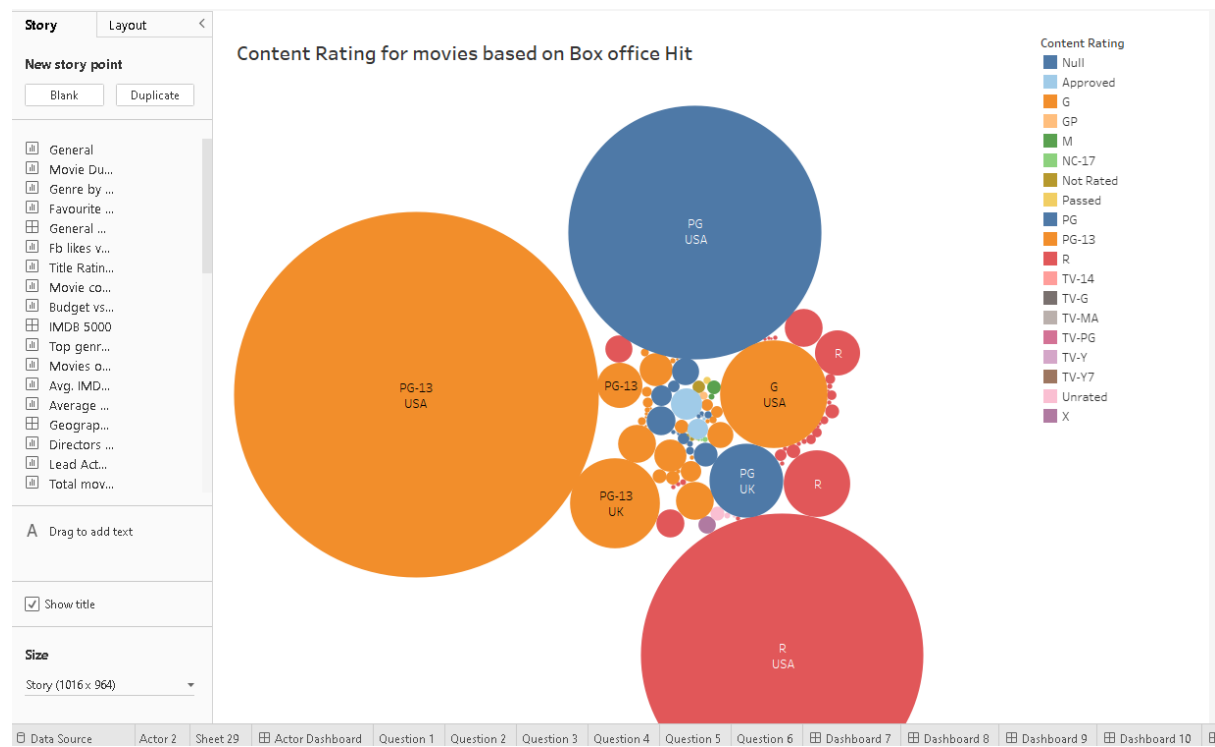
- Here we have plotted graph between Budget Vs Gross across different timeline ranging from 1997 to 2013.
- It is Depicted that after 2005 movie industry have increased the trajectory curve of making good profit over the movies.
- It acts as a suitable indicator that movie industry is growing faster and making better movies.

Who are the top directors and why?



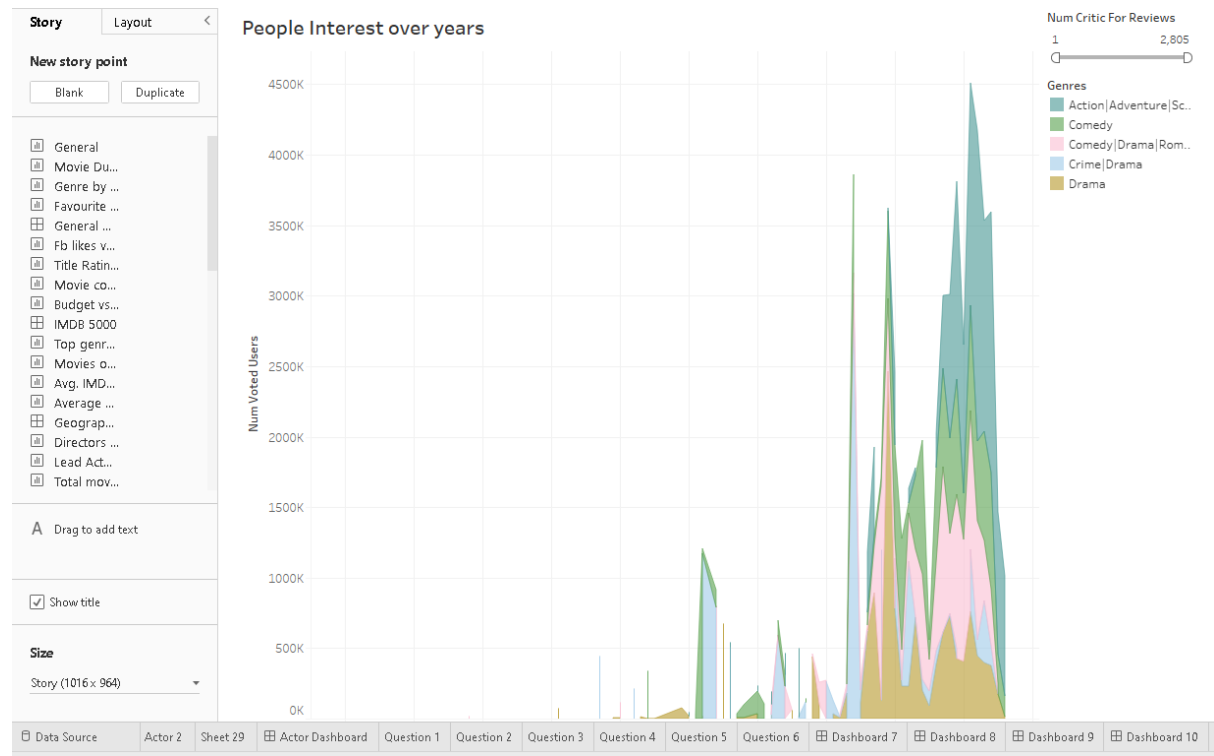
- Here we have listed top 10 directors of all time based upon their success over their movies and making it a box office hit.
- From the above story line, we can see that Steven Spielberg is the legendary person when compared to all the directors. He has made a huge box office collection of 4,014,061,704.00 over his movies where the budget of movies costs 1,452,900,870.00.
- Next to Steven Spielberg is Michael Bay his movies made a box office collection of 2,231,242,537.00 where the actual budget over the movie's costs 1,587,000,000.00.
- Next person who comes to the list is Tim Burton made a box office collection of 2,071,275,480.00 over his movies where his actual budget over his movie's costs only 1,247,000,000.00.
- From the above story line, we can view the top 10 director's gross vs budget

What is the role of movie rating on its success?



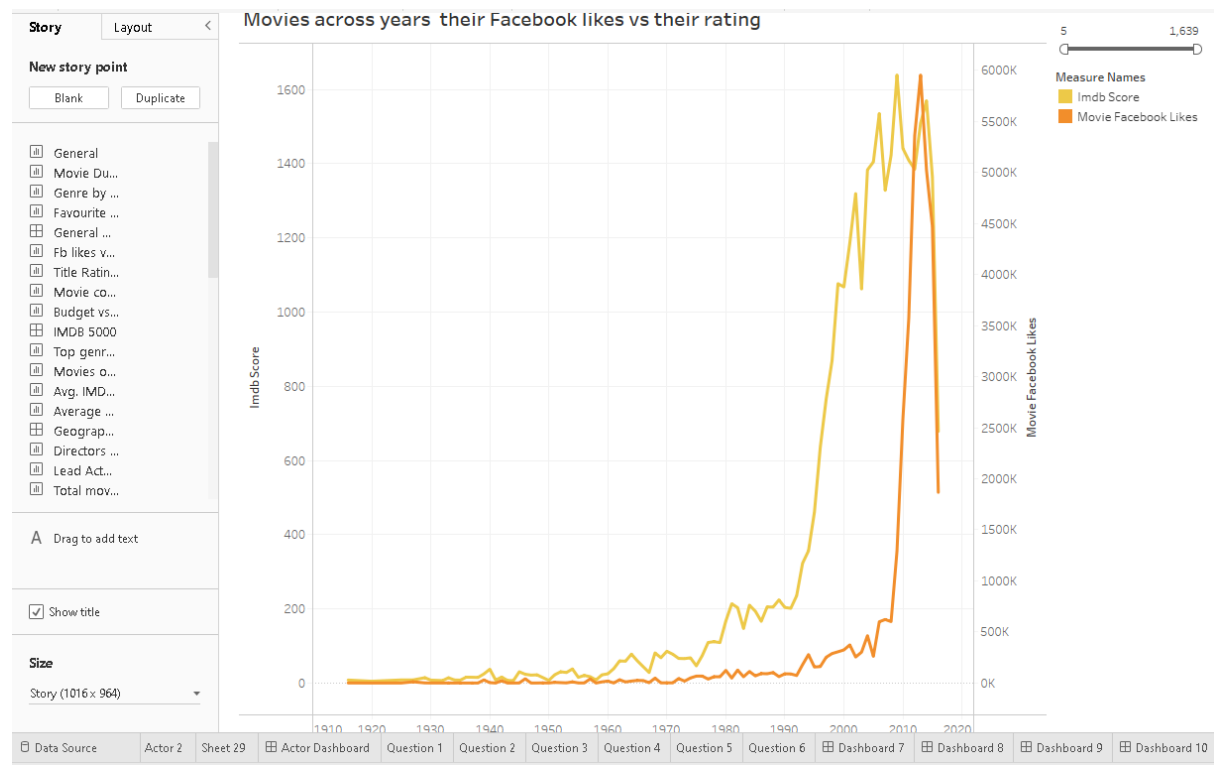
- From the above circle graph, we can see that PG-13 is the content rating for the top grossing movies from USA.
- Next to that we have content rating R that stands second and comes from USA followed by PG.
- PG-13 comes from UK directed top box office hit films.
- From the above circle, we are able to say that USA stands a pioneer in the movie business, and they stand tall and firm with different variety of movies, towards attracting the world audience than any other countries in the world.

How has the viewer's interest changed over the years?



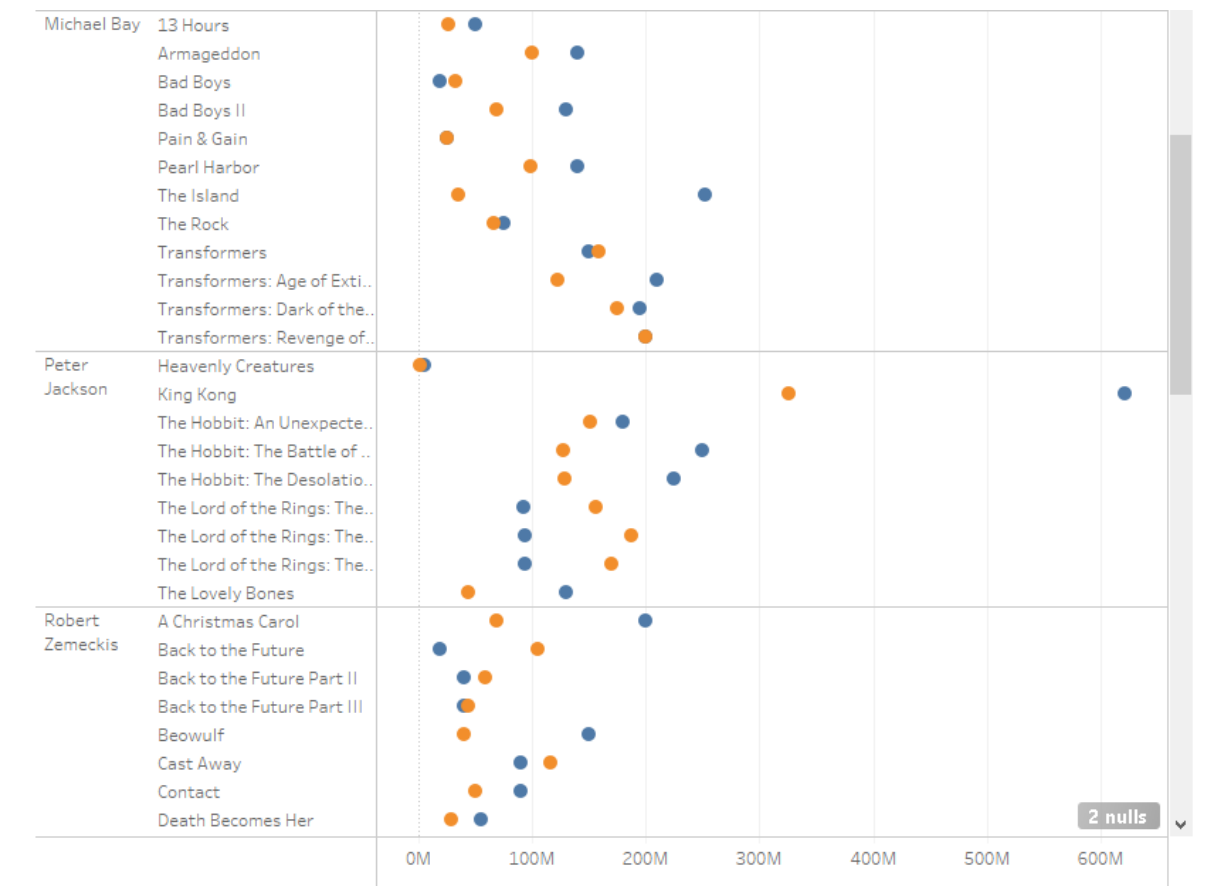
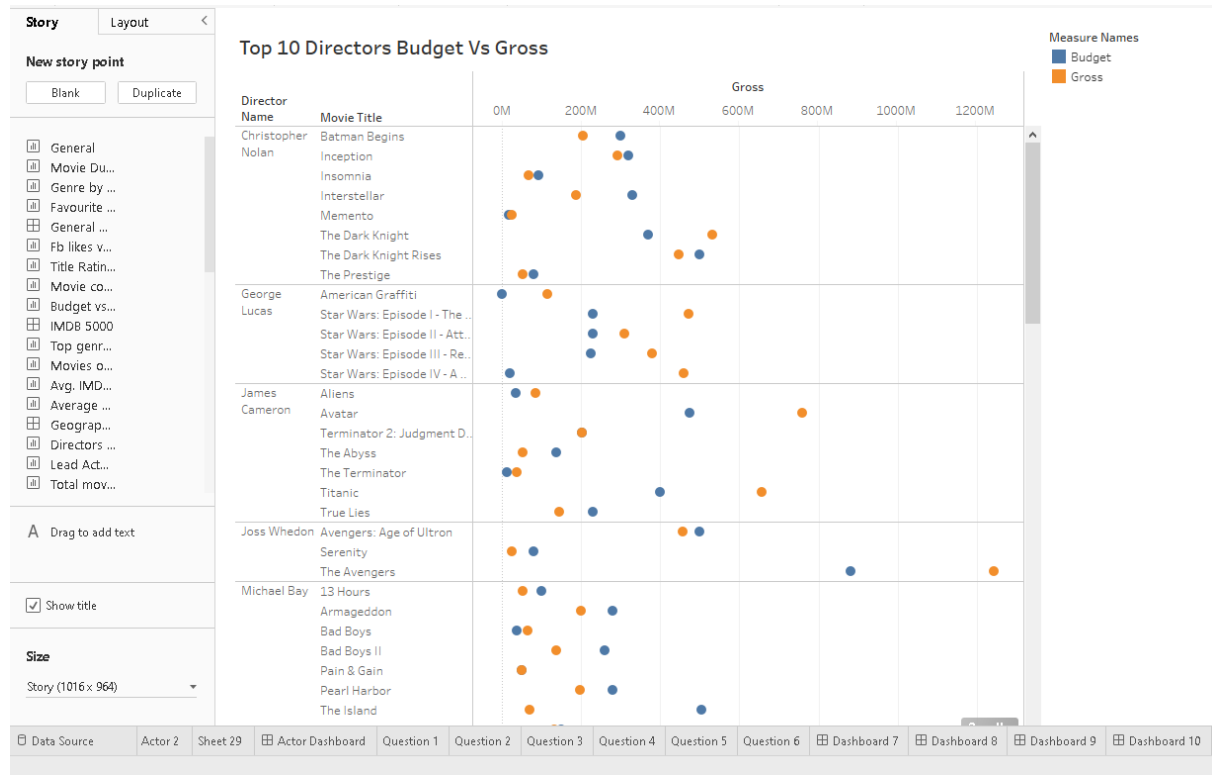
- From the above chart, we can say that initially people are interested in Drama that is during the period of 1950.
- Then comes comedy, which had a serious hype and interest during 1970.
- We can see that Crime/Drama have made a great deal during 1995.
- Then comes the one we love Action/Adventure/Sci-Fi ruled the movie industry starting from 2005 and till now.

What is the effect of Facebook on imdb score and movie?



- IMDB Rating towards the movies has increased from the initial 1915 and till now.
- Here we can see from the chart that IMDB Rating has increased from year to year , since the movie industry are making better movies every year and understanding the people interests.
- After the evolution of Facebook in 2005, people used to rank the movies in the FB by liking the movies.
- The Facebook rating of movies have increased over years and people are interested in liking the movies that are impressive. This liking system acts as a medium to say that whether a movie is better or good or average or flop.

Who are the ultimate earners and stars of box office?





- Christopher Nolan, George Lucas, James Cameron, Joss Wheldon, Michael Bay, Peter Jackson, Robert Zemeckis, Sam Raimi, Steven Spielberg, Tim Burton are the top 10 directors in the movie history.
- Christopher Nolan has directed 8 massive hit movies. It includes Interstellar, The Dark Knight Rises, Inception, and so on. He made a significant impact towards handling the movie with creativity and perfectionism which gives any of the audience a goose bump.
- Then comes George Lucas the director of Star Wars series. Star Wars 1 made a box office hit 474,544,677 whereas the actual budget is 115,000,000.
- Next person to the list is James Cameron known for the outstanding film 'Avatar', 'Titanic', 'Terminator Series'. He stands one among the top 10 directors.
- Peter Jackson movies include 'Lord of the ring', 'Hobbit series' have made huge, massive hit in Hollywood.
- We have mentioned the following directors as the top directors based upon their Box office collection vs budget for the movies they directed.

CONCLUSION

- In this project we have seen the Top movie titles, leading actor , supporting 1 and 2 , their count of movies, average IMDB score
- Top directors their budget and gross, Top movie titles and the country they are famous
- Content ratings popular among the watchers, top directors of that field and it's role in movie success
- Effect of Facebook on imdb score and movie
- How the viewer's interest changed over the years

BIBLIOGRAPHY

- www.youtube.com
- www.kaggle.com
- www.google.com
- www.github.com
- Faculties

BIDV Project

IMDB Ratings