

Data Engineering Project

End-to-End Data Modelling ETL and Visualization

Phase-3

Purpose of Study and Analysis Questionnaires

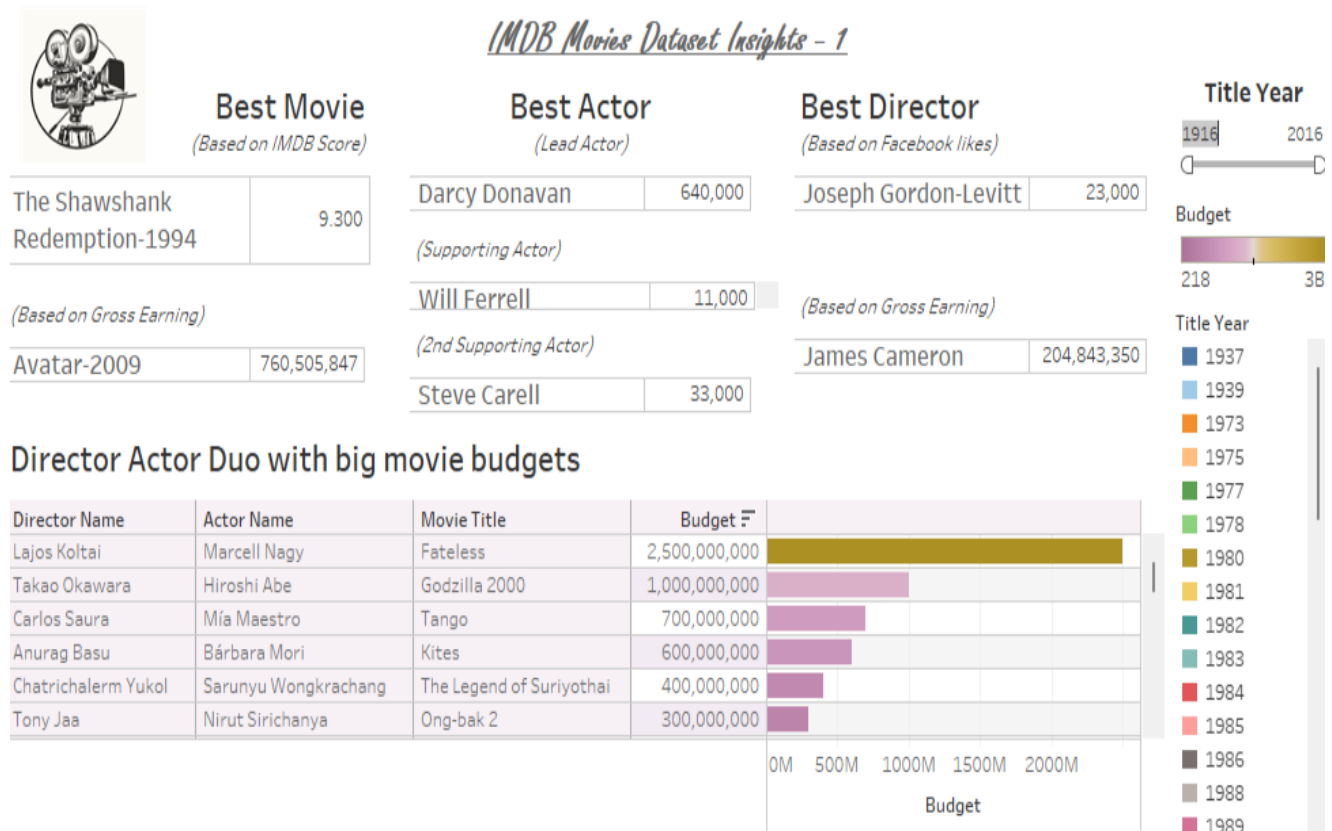
As part of phase-3, two dashboards are created using Tableau Desktop. IMDB movies dataset contains data for 4,794 movies that to for period for 100 years (from 1916 to 2016). Through dashboards, following questions are tried to be answered:

- Which movie is best for the given period?
 - On basis of best IMDB score.
 - On basis of gross earning of the movie.
- Which actor is most followed by audience on Facebook?
 - Most followed Leading actor.
 - Most followed Supporting actor.
 - Most followed 2nd Supporting actor.
- Who is best Director?
 - Based on followers on Facebook.
 - Based on gross earning of the movie directed.
- List of director-actor duo with big movie budgets.
- Year wise popular movie-plots with most popular shown first.
- What is IMDB score distribution?
- Movie duration in minutes over the years.
- Language wise distribution of number of movies.
- List of actors most followed on Facebook with their count and comparison.
- Country-wise movies gross collection.
- Most liked genre over the years.

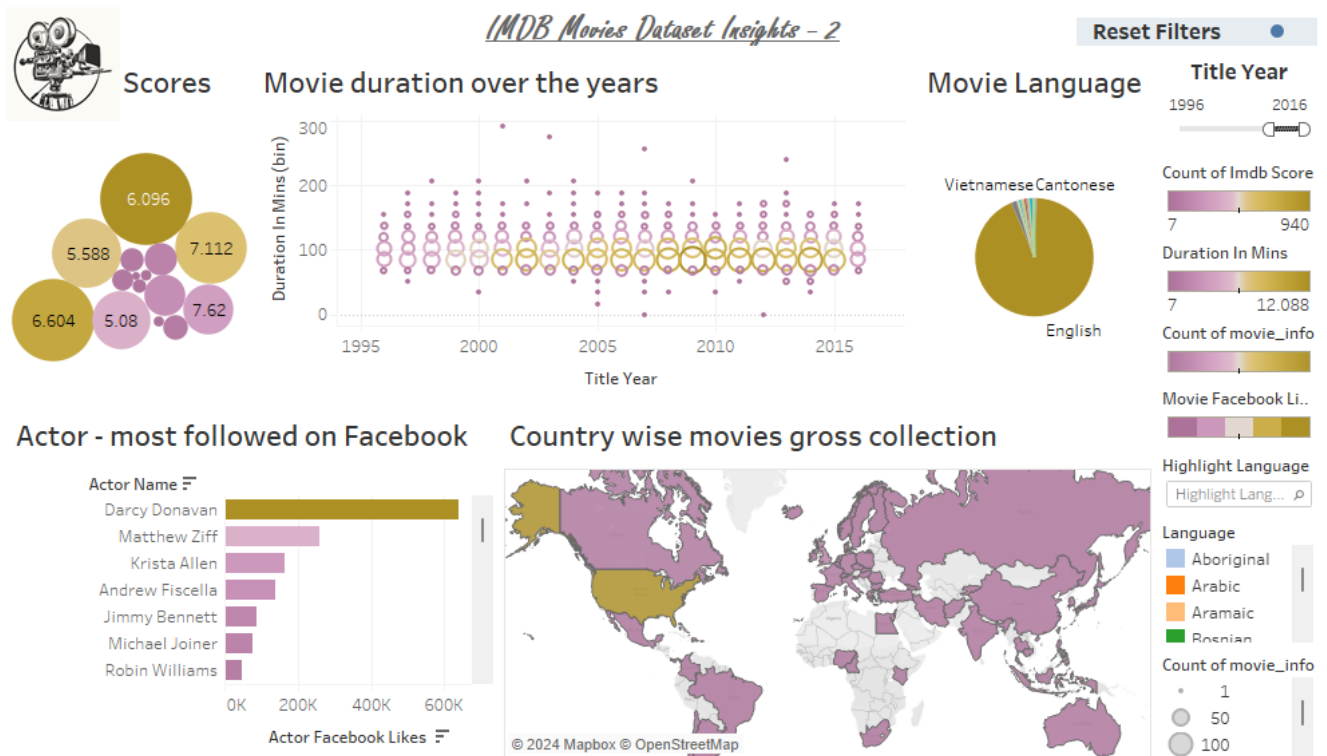
All the plots/charts are interlinked through filter-action and there is flexibility to change the range of years for which analysis is being performed.

Screenshots of final dashboards

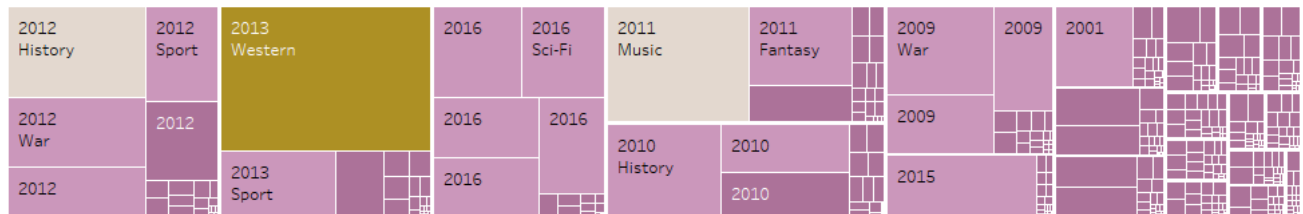
Dashboard IMDB Movies Dataset Insights - 1



Dashboard IMDB Movies Dataset Insights - 2



Most like genre over years



Deliverable 3 Screenshots:

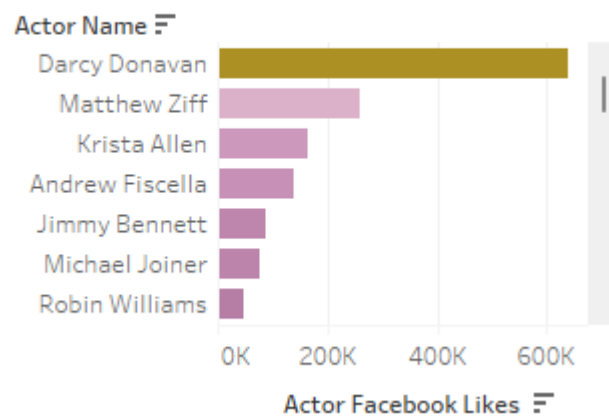
- Dashboard should have 3+ graphs
 - One pie chart

Movie Language



- One bar graph

Actor - most followed on Facebook



- One table

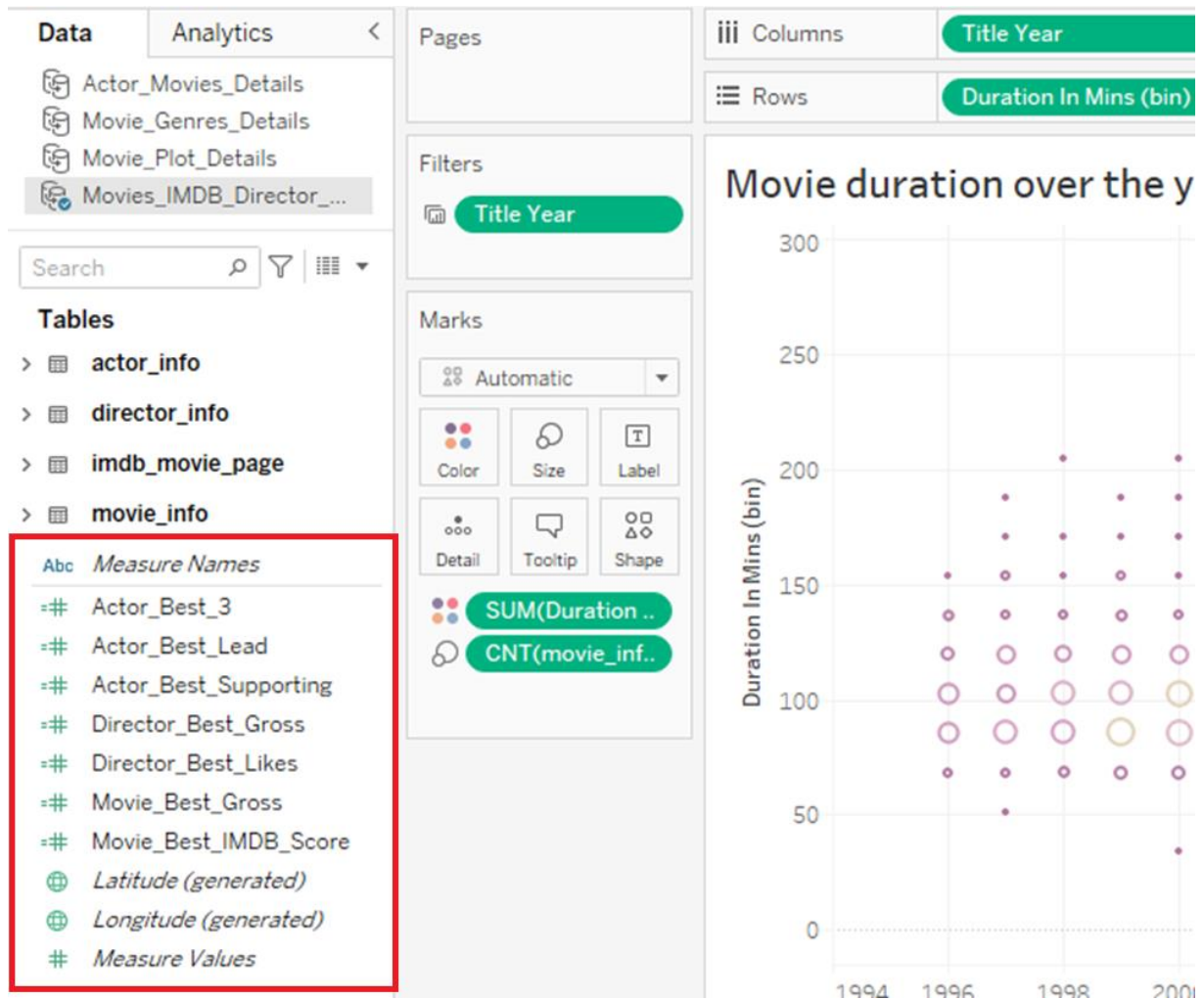
Director Actor Duo with big movie budgets

Director Name	Actor Name	Movie Title	Budget ₹
Lajos Koltai	Marcell Nagy	Fateless	2,500,000,000
Katsuhiro Ôtomo	William Hootkins	Steamboy	2,127,519,898
Takao Okawara	Hiroshi Abe	Godzilla 2000	1,000,000,000
Carlos Saura	Mía Maestro	Tango	700,000,000
Anurag Basu	Bárbara Mori	Kites	600,000,000
Chatrichalerm Yukol	Sarunyu Wongkrachang	The Legend of Suriyothai	400,000,000

0M 500M 1000M 1500M 2000M

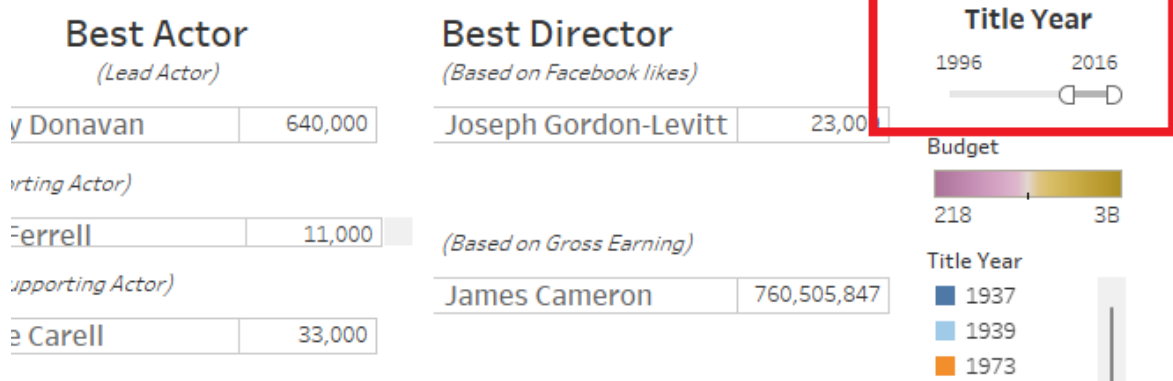
Budget

- Show examples of creating calculated fields



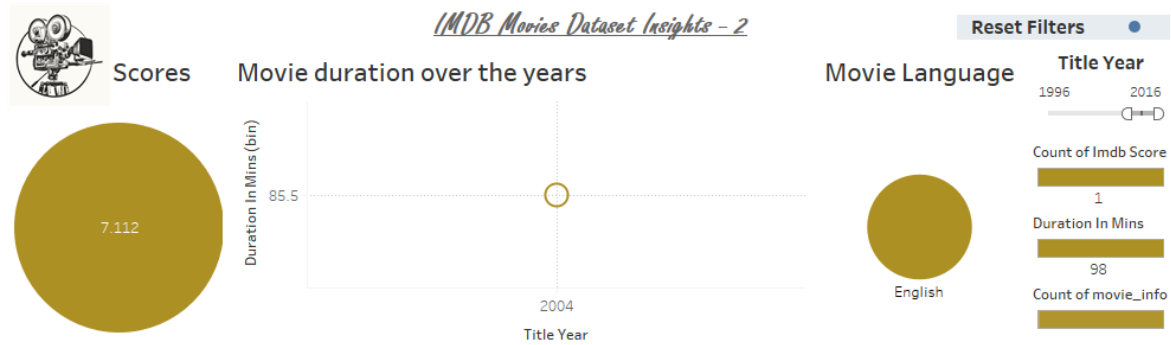
- Show examples of creating filters

IMDB Movies Dataset Insights - 1

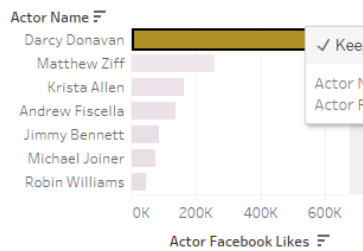


- Show examples of using filter actions based on selection on graph

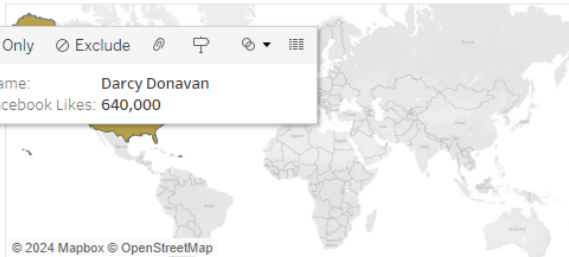
On selecting Actor "Darcy Donavan" all other graphs have filtered based on actor's name.



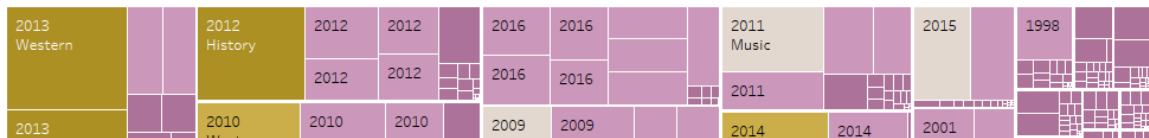
Actor - most followed on Facebook



Country wise movies gross collection



Most like genre over years



• Show examples of using formatting

Although all the components are formatted to choose particular colour theme, fonts, and order, as example of using formatting, title is shared in screenshot:

