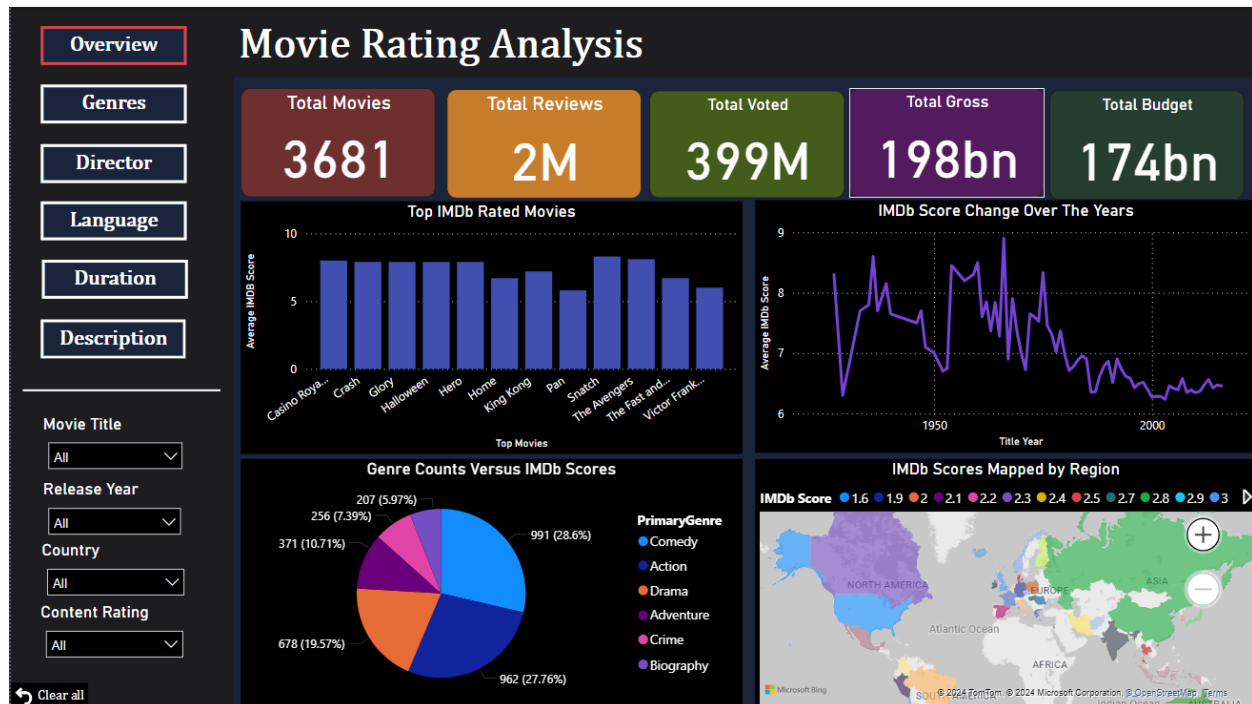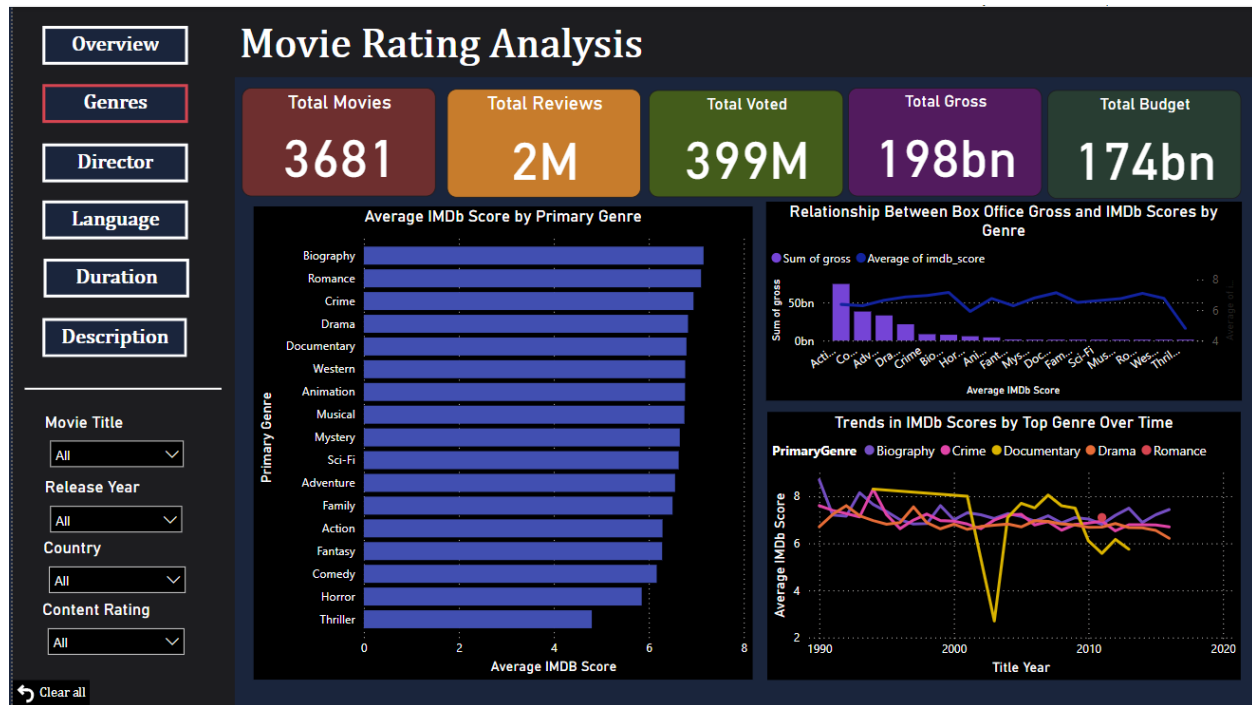# Data Mining and Business Intelligence
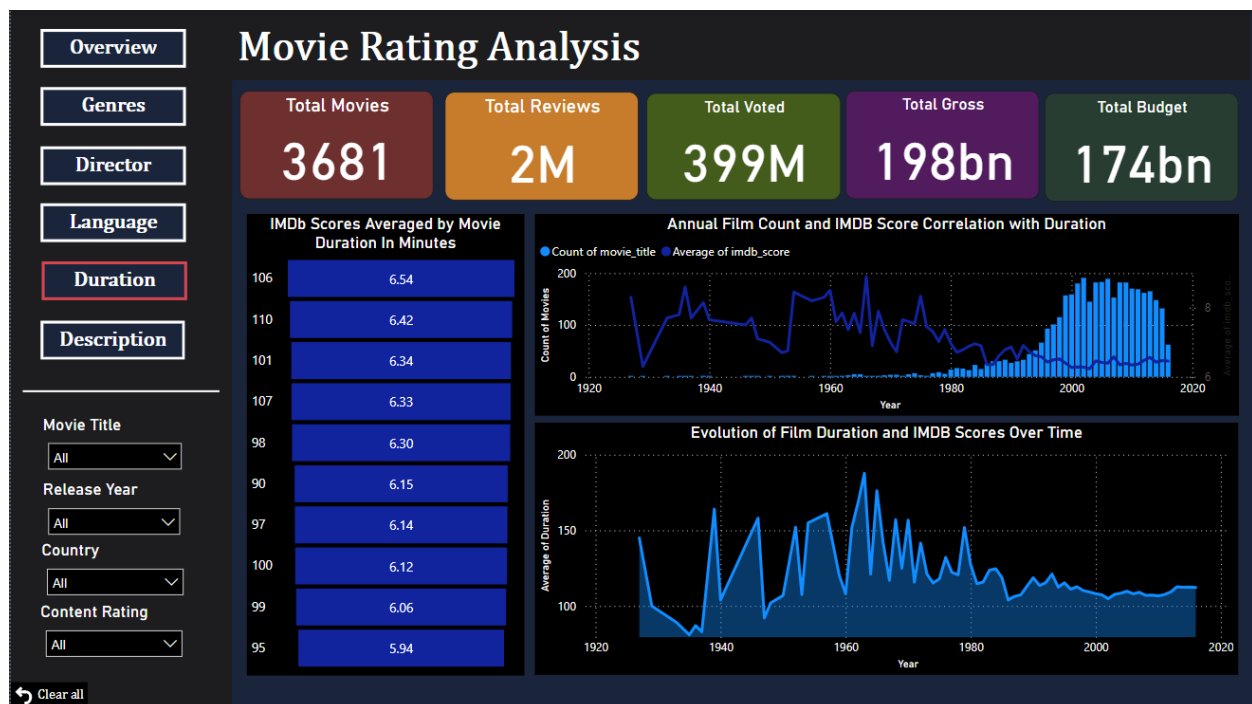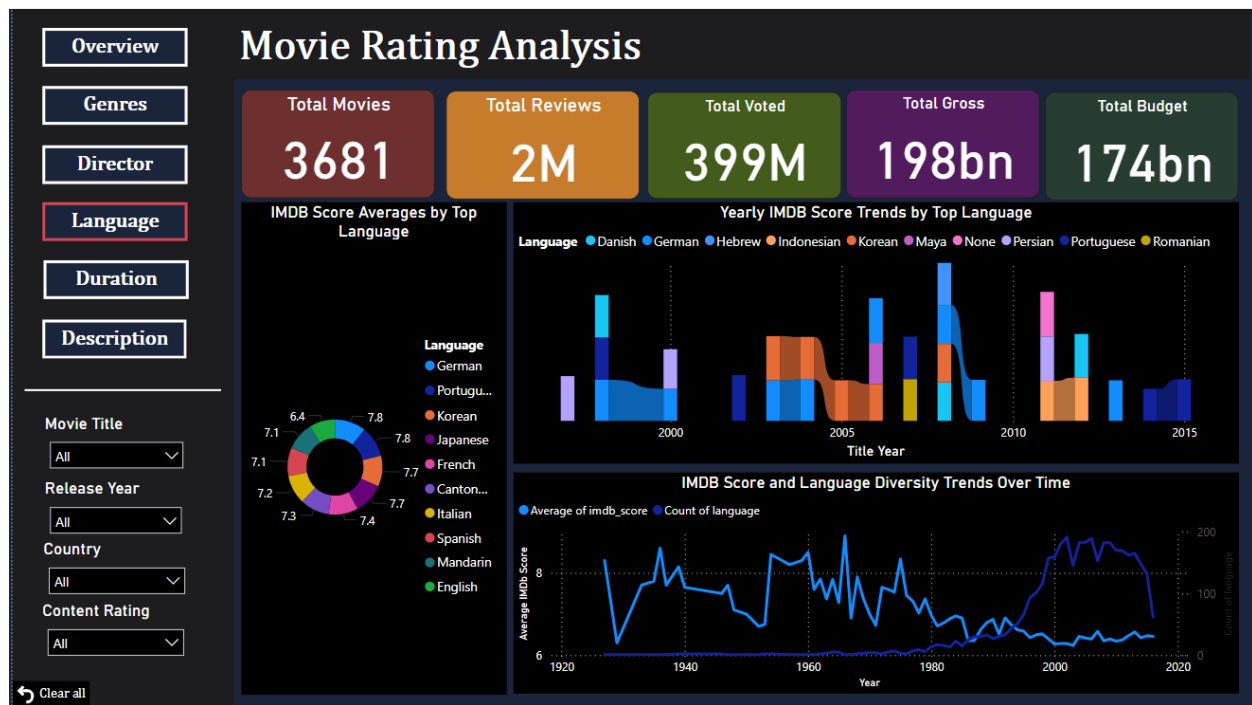
Final Project

Movie Rating Analysis – Based on IMDB Rating

Faiyaz Sattar. Mohammad Abbas Yunus, Pui Hei Wong

Movie Rating Analysis

Overview · **Genres** · **Director** · **Language** · **Duration** · **Description**

Total Movies	Total Reviews	Total Voted	Total Gross	Total Budget
3681	2M	399M	198bn	174bn

Average IMDb Score by Primary Genre

Primary Genre (top to bottom): Biography, Romance, Crime, Drama, Documentary, Western, Animation, Musical, Mystery, Sci-Fi, Adventure, Family, Action, Fantasy, Comedy, Horror, Thriller

X-axis: Average IMDB Score (0 to 8)

Relationship Between Box Office Gross and IMDb Scores by Genre

● Sum of gross   ● Average of imdb_score

Y-axis: Sum of gross (0bn, 50bn)

X-axis genres: Acti.., Co.., Adv.., Dra.., Crime, Bio.., Hor.., Ani.., Fant.., Mys.., Doc.., Fam.., Sci-Fi, Mus.., Ro.., Wes.., Thri..

X-axis label: Average IMDb Score

Trends in IMDb Scores by Top Genre Over Time

PrimaryGenre ● Biography ● Crime ● Documentary ● Drama ● Romance

Y-axis: Average IMDb Score (2 to 8)

X-axis: Title Year (1990, 2000, 2010, 2020)

Movie Rating Analysis

Overview · **Genres** · **Director** · **Language** · **Duration** · **Description**

Total Movies	Total Reviews	Total Voted	Total Gross	Total Budget
3681	2M	399M	198bn	174bn

IMDB Score Averages by Top Director

Director	Score
Peter Jackson	7.68
Robert Zemeckis	7.31
Woody Allen	7.00
Tim Burton	6.93
Martin Scorsese	7.68
Clint Eastwood	7.21
Spike Lee	6.73
Steven Sod...	6.71
Steven Spielberg	7.54
Ridley Scott	7.07

Box Office Earnings Distribution by Top Director

Values shown: 2bn, 4bn, 2bn, 3bn, 2bn, 2bn, 2bn

Director Name
● Steven Spielberg
● Peter Jackson
● Michael Bay
● Tim Burton
● Sam Raimi
● James Cameron
● Christopher Nolan

Movie Count by Top Director

Y-axis: Count of Movies (0 to 20)

X-axis (Director Name): Steven Spielberg, Clint Eastwood, Woody Allen, Ridley Scott, Martin Scorsese, Steven Soderbe..., Tim Burton, Renny Harlin, Spike Lee, Barry Levinson, John Carpenter, Michael Bay, Oliver Stone, Robert Rodriguez, Robert Zemeckis, Ron Howard

Movie Rating Analysis

Overview

Genres

Director

Language

Duration

Description

Movie Title
All

Release Year
All

Country
All

Content Rating
All

↺ Clear all

Total Movies	Total Reviews	Total Voted	Total Gross	Total Budget
3681	2M	399M	198bn	174bn

IMDB Score Averages by Top Language

Language
- German
- Portugu...
- Korean
- Japanese
- French
- Canton...
- Italian
- Spanish
- Mandarin
- English

6.4   7.8
7.1   7.8
7.1   7.7
7.2   7.7
7.3   7.4

Yearly IMDB Score Trends by Top Language

Language ● Danish ● German ● Hebrew ● Indonesian ● Korean ● Maya ● None ● Persian ● Portuguese ● Romanian

2000   2005   2010   2015

Title Year

IMDB Score and Language Diversity Trends Over Time

● Average of imdb_score ● Count of language

Average IMDB Score   Count of language

8

6

1920   1940   1960   1980   2000   2020

Year

Movie Rating Analysis

Overview

Genres

Director

Language

Duration

Description

Movie Title
All

Release Year
All

Country
All

Content Rating
All

↺ Clear all

Total Movies	Total Reviews	Total Voted	Total Gross	Total Budget
3681	2M	399M	198bn	174bn

IMDb Scores Averaged by Movie Duration In Minutes

Duration	Score
106	6.54
110	6.42
101	6.34
107	6.33
98	6.30
90	6.15
97	6.14
100	6.12
99	6.06
95	5.94

Annual Film Count and IMDB Score Correlation with Duration

● Count of movie_title ● Average of imdb_score

Count of Movies   Average of imdb_score

200

100

0

1920   1940   1960   1980   2000   2020

Year

Evolution of Film Duration and IMDB Scores Over Time

Average of Duration

200

150

100

1920   1940   1960   1980   2000   2020

Year

Result Interpretation

Main Insights:

Duration (0.37 correlation): Longer movies tend to have higher IMDb scores, suggesting that films with substantial content may be rated more favorably.

Gross (0.22 correlation) shows a modest positive correlation, indicating that movies which perform better financially tend to have higher scores, though the correlation is not very strong.

Title year has a slight negative correlation (-0.14) with IMDb scores, indicating that more recent movies might have slightly lower scores on average compared to older movies.

Thriller stands out as the lowest with an average IMDb score of 4.80, indicating that it might be more niche with a potentially smaller but dedicated audience base.

IMDb Rating by Genre Analysis: After our analyzation, we found the top genre that has the highest IMDb rating is Biography even though it is not the highest grossing genre. This suggests that the movie genre plays crucial role in getting IMDb rating even if the movies of that genre does not do well in box office. It shows potential growth of movie production of that specific genre.

Director Influence on Rating: Our analysis shows that Martin Scorsese and Peter Jackson both are the highest rated director in IMDb. Even though Steven Spielberg has produced more movies and is the highest grossing director. This shows how movie rating can be influenced just by director name and fame. This suggests movies with those directors can get higher rating in IMDb.