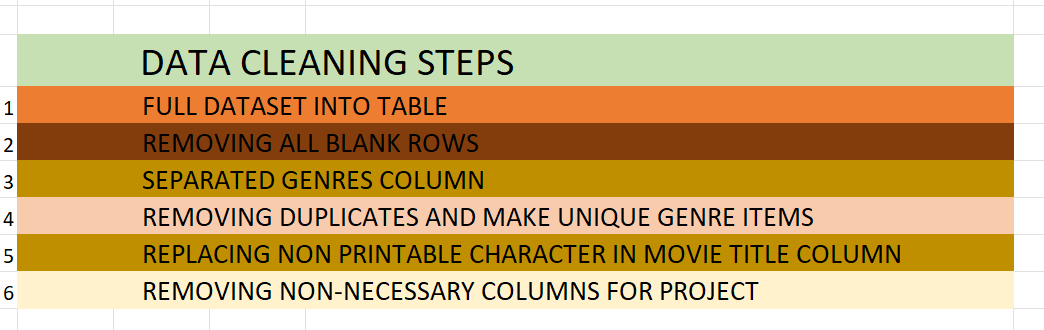
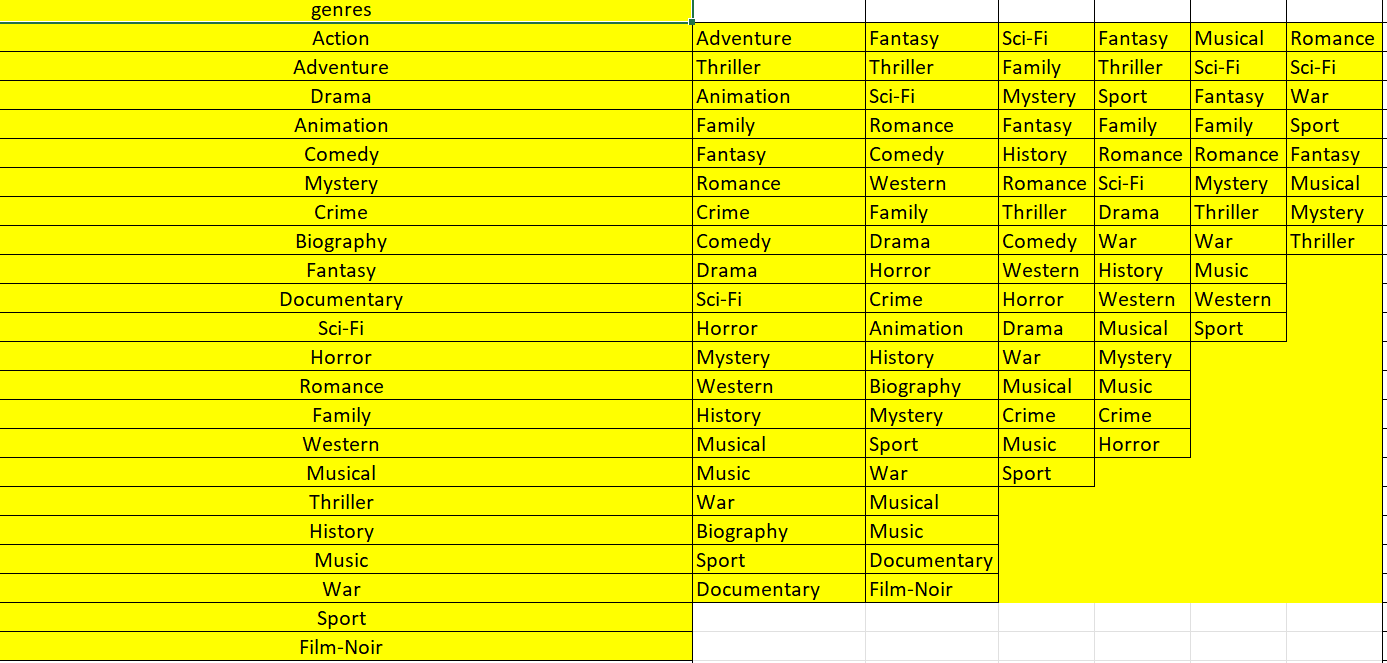
**: -IMDB MOVIE ANALYSIS: -**

**PROJECT DESCRIPTION:** The project was related to IMDB Movies. The objective of this project is to conduct a comprehensive data analysis of the IMDB Movie datasets. The analysis mainly focuses on several key aspects such as Movie genres, Duration, Language, Directors and Budget and its impact on IMDB scores and financial success of the movies.

**APPROACH:** Following steps are being followed

**DATA CLEANING:**





**DATA ANALYSIS:** Using Microsoft Excel various analysis on the detests are performed to get the output of the tasks given. This includes data manipulation to extract information, calculating descriptive statistic and visualizing relationship between variables

REPORT GENERATION**:** Reports are generated based on the result of analysis which includes insights from each analysis tasks, supported by visualizations and statistical summaries

**TECH-STACK USED:** Software: Microsoft Excel 2021

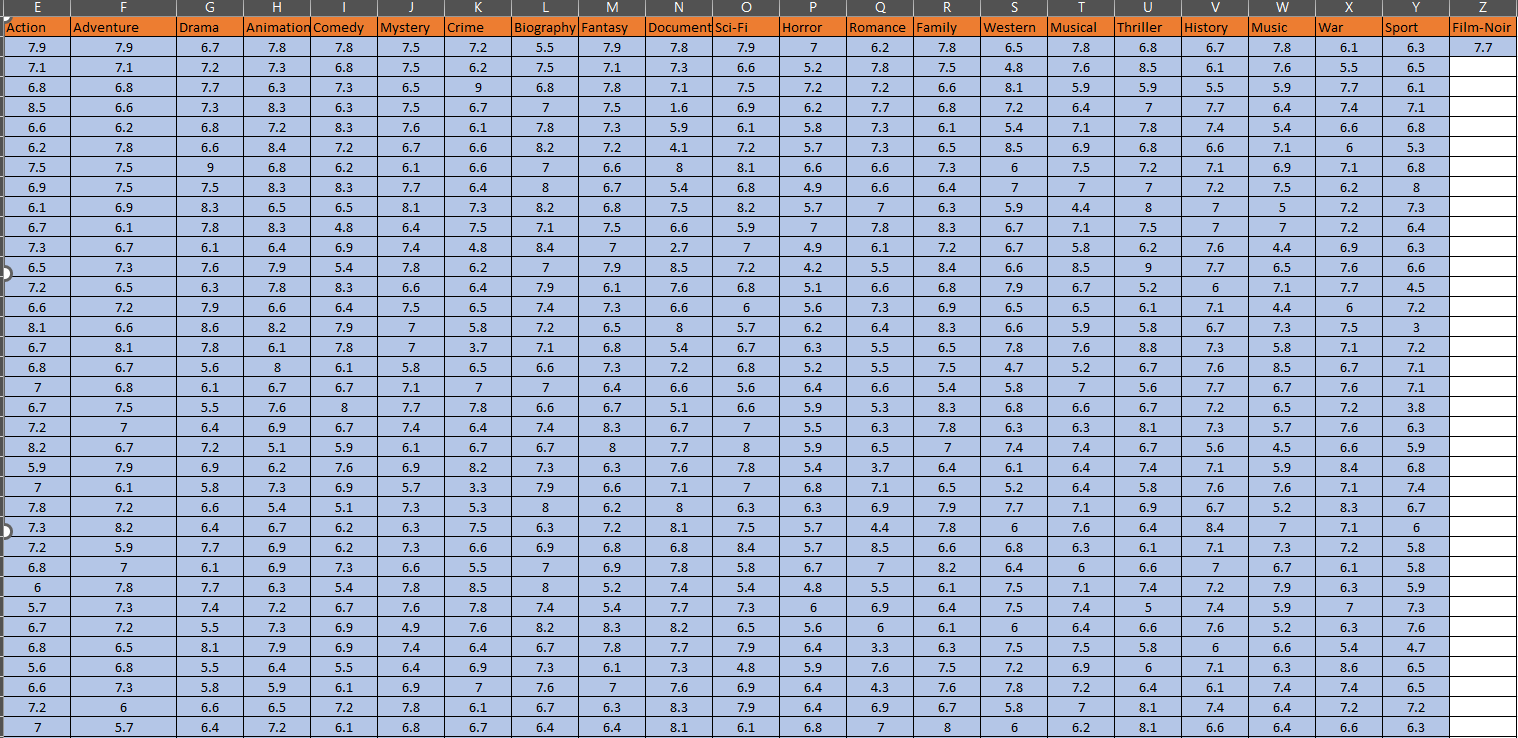
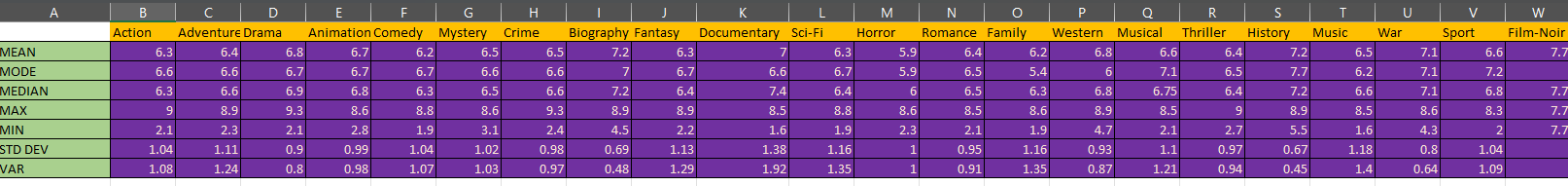
Purpose: Excel was chosen for its extensive data analysis functions, including pivot tables, charts, and statistical functions and visualization of the IMDB data

**INSIGHTS:** various key insights were uncovered through the data analytics process. Meaningful trends and patterns were observed in the data

* Movie genre Analysis
* Movie Duration Analysis
* Language Analysis
* Director Analysis
* Budget Analysis

1. **Movie Genre Analysis:** Analyse the distribution of movie genres and their impact on the IMDB score.

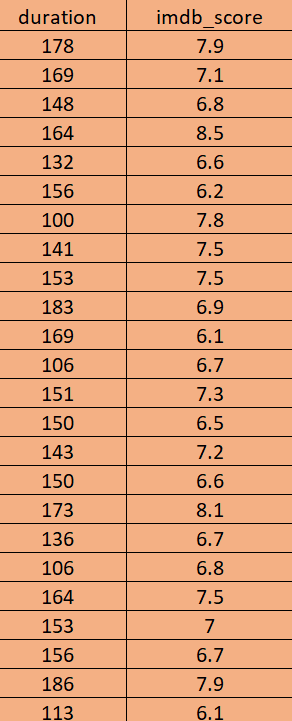
**Task:** Determine the most common genres of movies in the dataset. Then, for each genre, calculate descriptive statistics (mean, median, mode, range, variance, standard deviation) of the IMDB scores.



**B. Movie Duration Analysis:**Analyse the distribution of movie durations and its impact on the IMDB score.

**Task:** Analyse the distribution of movie durations and identify the relationship between movie duration and IMDB score.

|  |  |
| --- | --- |
| MOVIE DURATION |  |
| MEAN | 110.26 |
| MEDIA | 106.00 |
| STD | 22.65 |



**C. Language Analysis:**Situation: Examine the distribution of movies based on their language.

**Task:**Determine the most common languages used in movies and analyze their impact on the IMDB score using descriptive statistics.

|  |  |
| --- | --- |
| Top 5 most common languages | |
| **Language** | **No of Movies** |
| English | 3598 |
| French | 34 |
| Spanish | 23 |
| Mandarin | 15 |
| Japanese | 10 |

|  |  |
| --- | --- |
| **Language** | **Average of imdb\_score** |
| Aboriginal | 7.0 |
| Arabic | 7.2 |
| Aramaic | 7.1 |
| Bosnian | 4.3 |
| Cantonese | 7.3 |
| Czech | 7.4 |
| Danish | 7.9 |
| Dari | 7.5 |
| Dutch | 7.6 |
| English | 6.4 |
| Filipino | 6.7 |
| French | 7.4 |
| German | 7.8 |
| Hebrew | 8.0 |
| Hindi | 7.2 |
| Hungarian | 7.1 |
| Indonesian | 7.9 |
| Italian | 7.2 |
| Japanese | 7.7 |
| Kazakh | 6.0 |
| Korean | 7.7 |
| Mandarin | 7.1 |
| Maya | 7.8 |
| Mongolian | 7.3 |
| None | 8.5 |
| Norwegian | 7.2 |
| Persian | 8.1 |
| Portuguese | 7.8 |
| Romanian | 7.9 |
| Russian | 6.5 |
| Spanish | 7.1 |
| Thai | 6.6 |
| Vietnamese | 7.4 |
| Zulu | 7.3 |

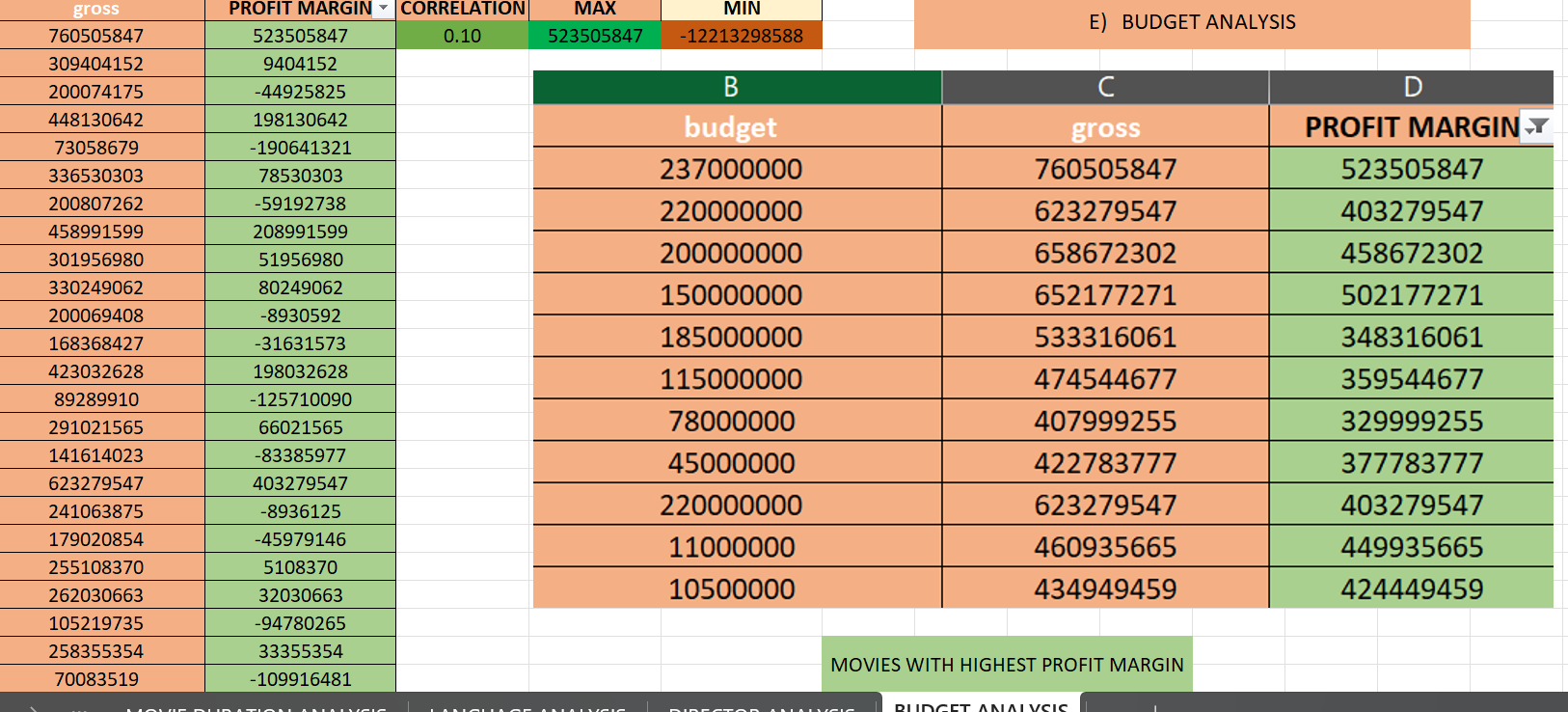
**D.** Director Analysis:Influence of directors on movie ratings.

Task: Identify the top directors based on their average IMDB score and analyse their contribution to the success of movies using percentile calculations.

|  |  |  |
| --- | --- | --- |
| **director\_name** | **Average of imdb\_score** | percentaile |
| Akira Kurosawa | 8.7 | 99.40% |
| Tony Kaye | 8.6 | 99.20% |
| Charles Chaplin | 8.6 | 99.20% |
| Ron Fricke | 8.5 | 98.70% |
| Majid Majidi | 8.5 | 98.70% |
| Damien Chazelle | 8.5 | 98.70% |
| Alfred Hitchcock | 8.5 | 98.70% |
| Sergio Leone | 8.4 | 98.70% |
| Christopher Nolan | 8.4 | 98.70% |
| Richard Marquand | 8.4 | 98.30% |

**E. Budget Analysis:** Explore the relationship between movie budgets and their financial success.

Task: Analyse the correlation between movie budgets and gross earnings, and identify the movies with the highest profit margin.



**EXCEL REFER LINK: -** [**IMDB -PROJECT 5.xlsx**](IMDB%20-PROJECT%205.xlsx)

