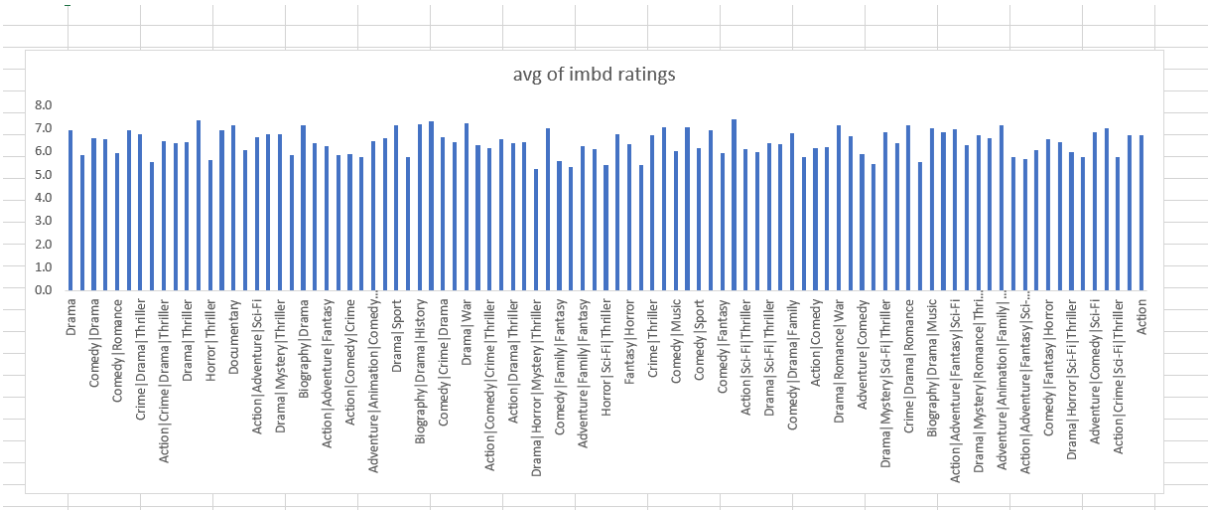
IMDB Movie Analysis

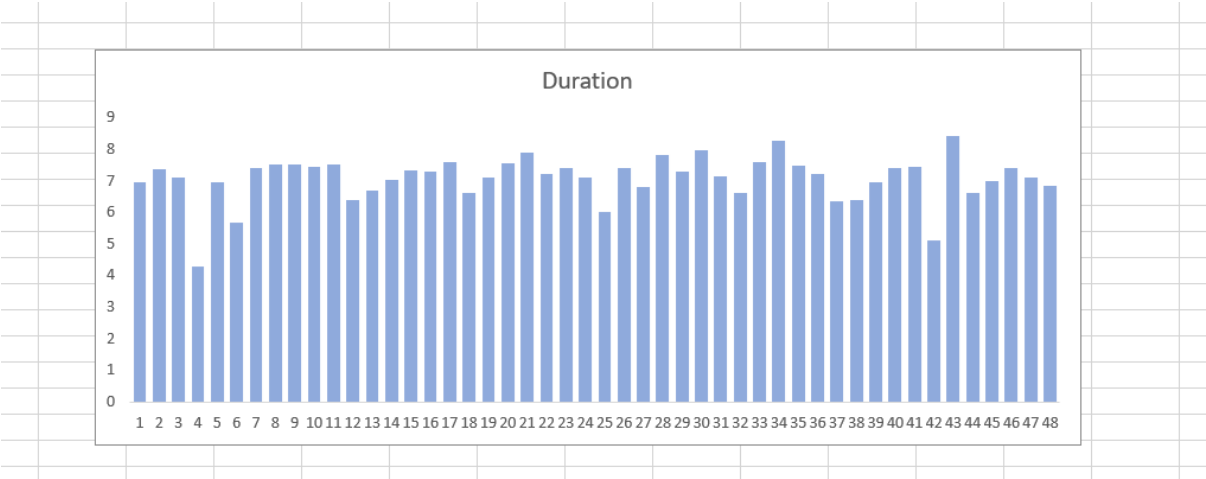
1. Determine the most common genres of movies in the dataset. Then, for each genre, calculate descriptive statistics (mean, median, mode, range, variance, standard deviation) of the IMDB scores.

    - Since the dataset contained a large volume of information, I used a **Pivot Table** to organize and analyze the data.
    - Within the Pivot Table, I selected **Genre** as the category and applied the **Count** function to calculate the number of movies per genre.
    - Alternatively, the data could also be counted using the =COUNT() function.
    - I then calculated key statistical measures such as **Average IMDB Rating**, **Median**, **Maximum**, **Minimum**, **Variance**, and **Standard Deviation** using built-in Excel functions.
    - Finally, I represented the results visually by plotting a graph to show how **movie genre impacts IMDB ratings**.



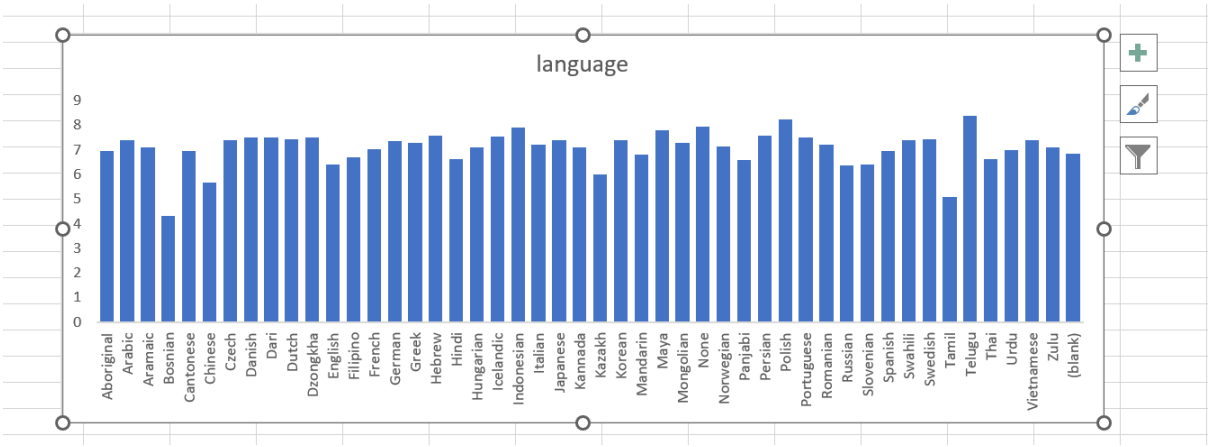
2. Analyze the distribution of movie durations and identify the relationship between movie duration and IMDB score.

    - I used the **Pivot Table** function to simplify the analysis.
    - I selected **Duration** as the variable and calculated the **average IMDB ratings** with respect to it.
    - To enhance understanding, I represented the results in a **pictorial form using a column chart**.



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

- I used the **Pivot Table** function to simplify the analysis.
- I selected **language** as the variable and calculated the **average IMDB ratings** with respect to it.
- To enhance understanding, I represented the results in a **pictorial form using a column chart**.



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

- I used the **Pivot Table** to calculate the average IMDB rating for each director.
- Then, I applied the =PERCENTILE() function to determine the **top 5% of directors** based on their ratings.
- From this analysis, it was observed that any director with an **average IMDB rating above 7.78** falls into the **top 5% category**.

5. Analyze the correlation between movie budgets and gross earnings, and identify the movies with the highest profit margin.

- This analysis could not be fully achieved using a Pivot Table, as the values were too large to process directly. Therefore, I created a **new sheet** and copied the **Gross** and **Budget** columns into it.
- I calculated the **Margin** value (Gross – Budget) to assess profitability.
- I then used the **MAX() function** to find the highest margin.
- To study the relationship between **Gross** and **Budget**, I applied the **CORREL() function**, which resulted in a correlation value of **0.10218**.
- Since the value is very close to 0, it indicates that **Gross and Budget are not strongly correlated**, meaning a higher budget does not necessarily guarantee higher gross earnings.