Data Science - IMDB Top Movies and TV Shows

Rules

- HVA Mini Hackathon is an individual event. Please refrain from discussing with friends or others.
- The aim is to practice Data Processing, Data Analysis, and Data Visualisation.
- The event **starts now** and submissions are due before 8-Jan-2023, Monday at 10 AM.

Problem Statement

Overview

In this Data Science challenge for Mini Hackathon 2.0, participants will utilize a dataset from Kaggle featuring the top 1000 movies and TV shows from IMDB. The objective is to process, analyze, and visually present insights derived from the dataset.

Task

Your task is to explore the IMDB dataset and uncover interesting insights about the movies and TV shows. Focus on data processing, analysis, and visualization to convey your findings effectively.

Suggested Areas for Analysis:

1. **Director's Impact on Gross Earnings:** Analyze how movies directed by different directors perform in terms of gross earnings. Are there any noticeable trends or patterns?
2. **Other Potential Analyses:**
    - Genre popularity over the years.
    - Correlation between IMDB ratings and commercial success.
    - Impact of movie length on ratings or earnings.

Tools to Use

- You may use SQL, Excel, Power BI, or Python (Pandas) for processing and analyzing the data.

- For visualization, you may use any tool or library within the mentioned technologies (e.g., Excel charts, Power BI, Python's Matplotlib/Seaborn).

Submission

Try to complete all the features listed. But in case you are not able to complete all the features, you can submit whatever portion of the project you complete within the given time.

You can submit the following links in the Google Form:

- GitHub repository link for all your code.
- A Loom video:
    - Showcasing the workflow of the project, and
    - Explaining the code
- Preferably, a link to the deployed code on Vercel or equivalent (optional but encouraged).