# **HVA24: Mini Hackathon 1.0 - Data Science Track**

## Rules

* The Mini Hackathon is an individual event. Please refrain from discussing or collaborating with others.
* The goal is to challenge yourself in Data Processing, Data Analysis, and Data Visualization. External help, including tools like ChatGPT, is NOT allowed. The objective is to test your own skills and abilities.
* The event starts now, and submissions are due before Monday, 26th August 2024, at 10 AM.

## Problem Statement

### Overview:

In this Data Science challenge, participants will explore a [dataset from Kaggle](https://www.kaggle.com/datasets/harshitshankhdhar/imdb-dataset-of-top-1000-movies-and-tv-shows) featuring the top 1000 movies and TV shows from IMDB. The focus of this challenge is to utilize your skills in SQL and Power BI to process, analyze, and visually present meaningful insights derived from the dataset.

### Task:

Your task is to dive into the IMDB dataset and uncover interesting insights about the movies and TV shows. The emphasis should be on data processing, analysis, and visualization using SQL and Power BI.

### Suggested Areas for Analysis:

* **Director's Impact on Earnings:** Analyze how movies directed by different directors perform in terms of gross earnings. Are there any noticeable trends or patterns?
* **Genre Popularity Over the Years:** Investigate how different genres have evolved in popularity over time.
* **Correlation Between IMDB Ratings and Commercial Success:** Explore if there's a relationship between a movie's IMDB rating and its box office earnings.
* **Impact of Movie Length on Ratings or Earnings:** Examine whether the duration of a movie influences its ratings or financial success.
* **Actor Influence on Movie Success:** Analyze how the presence of certain actors correlates with a movie's performance in terms of ratings and earnings.
* **Release Date Analysis:** Study the impact of a movie's release date (e.g., month, season) on its success.

These are suggested areas for analysis. Feel free to explore other aspects of the data that you find interesting or relevant.

### Tools to Use:

* **Data Processing and Analysis:** Focus on using SQL for querying and data manipulation.
* **Visualization:** Use Power BI to create compelling visualizations that effectively communicate your findings.

## Submission

Aim to complete all suggested analyses. However, if you're unable to finish everything, submit whatever you've accomplished within the time frame.

Submit the following via the provided [Google Form](https://forms.gle/Fcq3KPX59oyrDVpd6):

* Github Link: A link to your GitHub repository containing all your SQL queries and Power BI files.
* A Loom video:
  + Demonstrating the workflow of your project.
  + Explaining your analysis and visualizations.
* (Optional but encouraged) A link to any deployed visualizations or dashboards.

# 