

T5 Bootcamp

Data scientists

Project Proposal

Movie Dataset

- **Question/Problem statement :**

What's the factors that effect the movie rating? And how we can predict that movie will be trend?

- **Data Description :**

Main Data source is IMDb <https://www.imdb.com/> 'Top 1000" (Internet movie database)

- **Dataset Size:** 1,000 rows and 8 features contains (Name, Production Year, Movie Type, Rate the Site, Website visitors Vote, Duration, Us Gross, Meta Score.
- **Data Type:** Integer, Date, String.

- **Tools :**

- **Programs :**

We will use Python ,(HTML,CSS) optional.

- **Libraries :**

Pandas, matplotlib, seaborn, numpy, datetime, BeautifulSoup, requests.

- **Functions :**

We will use several functions that we define in the project.

- **Plots :**

We will use one of them (lineplot, Scatter, bar, hits, pie, circle).

- **MVP Goal:**

The goal of this project is to identify the most effective factors on movie rating and build a model to predict the rating based on that factors.