# IMDB TOP 1000 Analysis

Zaid Al Otel, Hubert Magdziak, Mateusz Pliszka, Michał Woźniak

#### Why this topic?

- We are interested in both movies and data analysis
- So we can go through the EDA process using Python
- To have a powerful tool to find the best films
- To have a chance to use Google's BigQuery
- To satisfy our curiosity
- To excel in data visualization

#### Our approach

- 1. Data preparation in python wrangling and feature engineering
- Loading data into BigQuery
- 3. Connecting to BigQuery in Python
- 4. Data analysis
- 5. Creating an interactive report

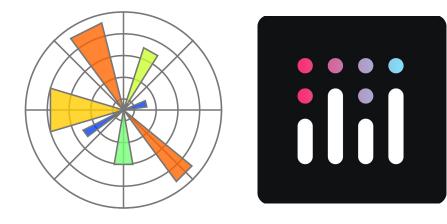


### Examples of problems we will to answer

What are the outliers in IMDB rating and meta rating relation?	IQR rule Z-score rule
Do better-rated movies generate more revenue?	Correlation testing
How has the movie duration change over time and genres?	Graphical analysis
Which actors appear most often in movies?	Graphical analysis

## Our final goal

A public interactive report for easy browsing of our findings.



## Thank you for your attention