

MovieLens Data Exploration

Project Data Description:

MovieLens data sets were collected by the GroupLens Research Project at the University of Minnesota.

Datasets: Download from Olympus.

Domain:

Entertainment and Internet

Context:

The GroupLens Research Project is a research group in the Department of Computer Science and Engineering at the University of Minnesota. The data is widely used for collaborative filtering and other filtering solutions. However, we will be using this data to act as a means to demonstrate our skill in using Python to “play” with data.

Datasets Information:

- **Data.csv:** It contains information of ratings given by the users to a particular movie.
Columns: user id, movie id, rating, timestamp
- **item.csv:** File contains information related to the movies and its genre.
Columns: movie id, movie title, release date, unknown, Action, Adventure, Animation, Children's, Comedy, Crime, Documentary, Drama, Fantasy, Film-Noir, Horror, Musical, Mystery, Romance, Sci-Fi, Thriller, War, Western
- **user.csv:** It contains information of the users who have rated the movies
Columns: user id, age, gender, occupation, zip code

Objective:

To implement the techniques learnt as a part of the course.

Learning Outcomes:

- Exploratory Data Analysis
- Visualization using Python
- Pandas – groupby, merging

Tasks and steps:

Please refer the Jupyter notebook

References:

- <https://movielens.org/>