MovieLens Project

By Shanti Jogi

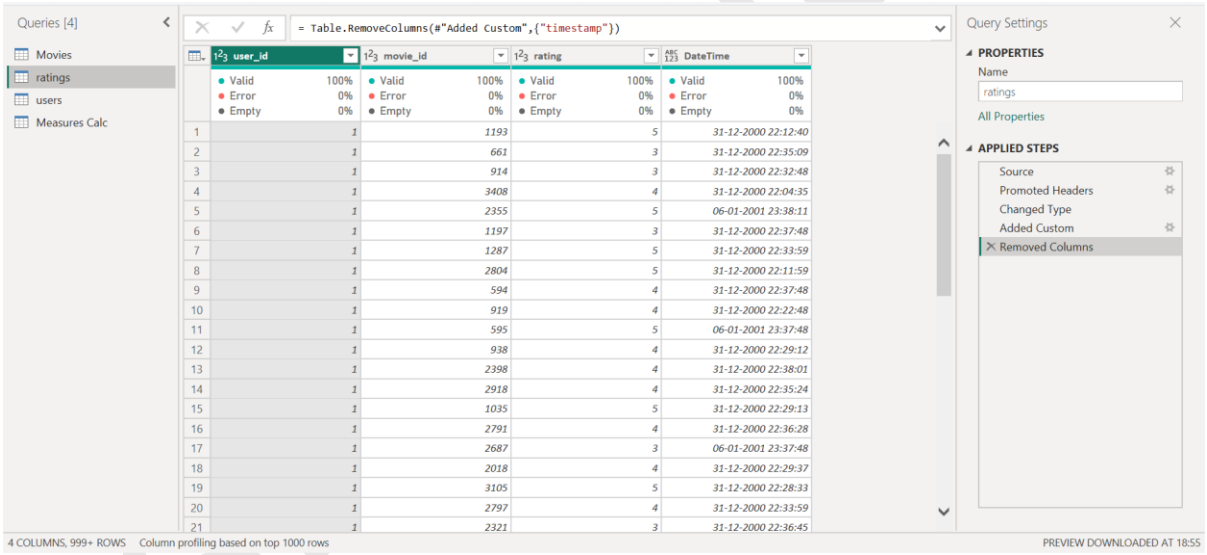
Project Overview

The MovieLens project was aimed at analysing movie ratings data using Microsoft Power BI. The dataset included information about **movies**, **users**, and **ratings**. The goal was to clean, transform, and visualize the data to uncover insights about user preferences and rating trends.

Data Cleaning and Preparation

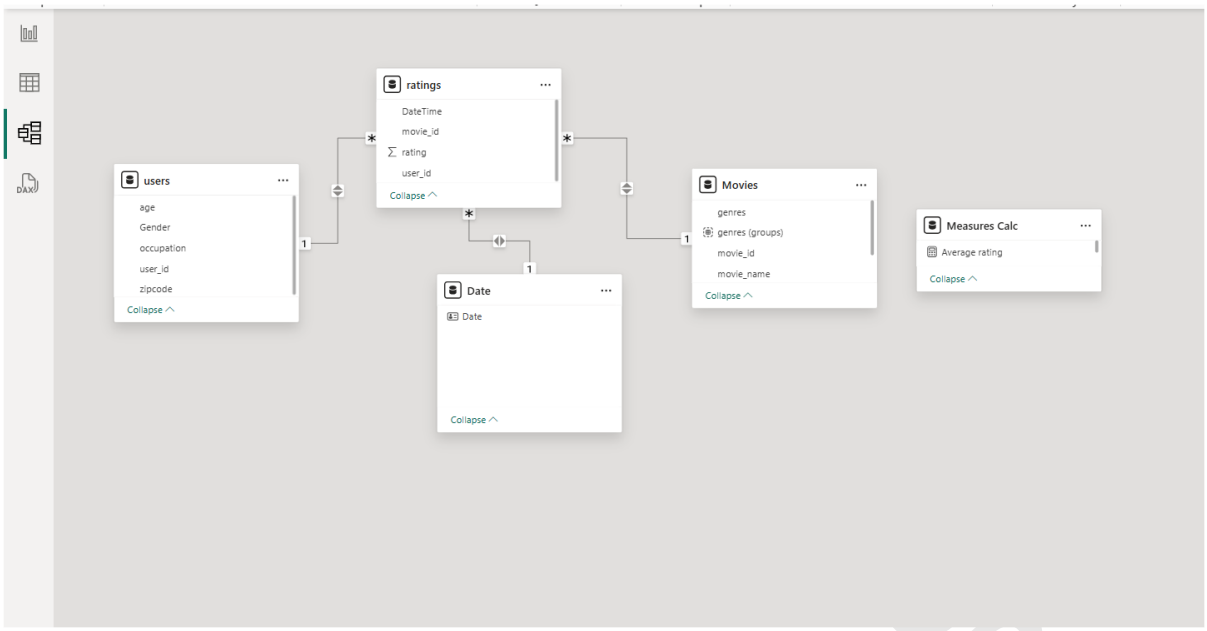
Performed in **Power Query**:

- Imported multiple datasets: movies, ratings, and users.
- Split the genres column using a delimiter (,) to handle multiple genres.
- Converted the rating timestamp (in seconds) to a standard **date and time** format.
- Removed unnecessary columns and renamed fields for clarity.



Data Modelling

- Built **entity relationships** between:
  - movies and ratings (via movieId)
  - ratings and users (via userId)
- Created a **Master Date Table** using DAX.
- Linked the Date Table with the Ratings table using the rating_date.
- Marked the Date Table as a **Date Table** to enable time intelligence functions.



Report View Customization

- Grouped columns into folders in the Report View for better organization.
- Created calculated columns and measures using DAX.

Visuals and Dashboard Elements

- **Cards** for key performance indicators (KPIs) like:
  - Total number of ratings
  - Number of users
  - Average rating
- **Bar/Column Charts** to show:
  - Top-rated movies
  - Most active users
- **Line Chart** to display rating trend.
- **Stacked Column Chart** to compare genre preferences by gender.
- **Slicers** for interactive filtering by genre, gender, and date range.

Analysis Highlights

- **Genre Preferences**: Notable variation in genre preference across gender.
- **Rating Trends**: Average ratings fluctuated across time, genres and the volume.
- **Top Movies**: Identified highest-rated and most-rated movies.
- **User Activity**: Highlighted the most active users based on the number of rated movies.

Tools Used

- Microsoft Power BI
- Power Query Editor
- DAX (Data Analysis Expressions)
- Data Modelling (Relationships, Date Table)

Outcome

- Developed an interactive Power BI dashboard.
- Cleaned and transformed MovieLens data for accurate analysis.
- Applied best practices in data modelling and visualization.
- Delivered meaningful insights about user behaviour and genre trends.