# Code Link and Description

The complete implementation of the project can be accessed through the following Google Colab notebook:

https://colab.research.google.com/drive/1J7FfNrdpi5bUC5j4JSvZGHMf4RQpo_QR?usp=sharing

This notebook includes all steps necessary to reproduce the analysis described in the project report. It is written in Python and utilizes libraries such as `pandas`, `matplotlib`, `seaborn`, and `scipy` for data processing, statistical testing, and visualization.

## What the Code Does

1. **Data Import and Cleaning**
    - Loads Netflix viewing history and menstruation tracking data from CSV files.
    - Standardizes column names and converts date fields into proper date time format.
    - Labels each viewing based on whether it occurred during a menstruation period.
2. **Genre Analysis**
    - Separates and normalizes genre data.
    - Calculates frequency distributions of genres during and outside menstruation periods.
3. **Hypothesis Testing**
    - Applies a chi-square test of independence to examine whether genre preferences differ between menstruation and non-menstruation periods.
    - The test results indicate a statistically significant difference (p-value $< 0.001$).
4. **Data Visualization**
    - Generates multiple plots to support the analysis:
        - A bar chart comparing genre preferences
        - A line chart showing monthly viewing trends
        - A boxplot displaying daily viewing counts based on menstruation status
5. **Output**
    - All generated visualizations are saved into a `plots/` folder for documentation and reporting purposes.