# Task 1: Exploring Multiple Categorical Variables with Crosstab

Objective:

The goal of this task was to explore how multiple categorical variables relate to each other using a contingency table. For that, I used the `pandas.crosstab()` function in Python, which is great for summarizing data in a clear, tabular format.

Dataset:

I created a custom dataset that simulates user activity on an online learning platform. It includes the following columns:

- User\_Gender: Male or Female

- Course\_Category: Type of course enrolled in — Programming, Marketing, or Design

- Device\_Used: Device used to access the course — Laptop, Mobile, or Tablet

- Payment\_Status: Whether the user’s payment was successful or failed

Analysis:

To analyze the data, I used a crosstab with:  
- User\_Gender as the index  
- A combination of Course\_Category and Payment\_Status as columns  
This helped show how payment outcomes vary across genders and course types. I also included total counts using the `margins=True` parameter to make the table easier to interpret.

Tools Used:

- Python (with pandas library)

- Google Colab (for writing and running the code)

Files Included:

- Task1\_MultipleCategoricalVariables.ipynb: The full code, including dataset creation and crosstab

- task1\_crosstab\_output.csv: The final contingency table in CSV format

- Task1\_Report.docx: This report file

Summary:

This task was a good exercise in understanding how to compare multiple categorical features in a dataset. It’s a useful step when preparing for deeper data analysis or building machine learning models.