Assignment Lab 3

Analyzing the Relationship Between Smoking and Tea Consumption
Using Chi-Square Test in R

Objective

The objective of this assignment is to analyze the relationship between **smoking status** and **tea consumption** using a **Chi-Square Test** and visualize the data using **bar plots** in **R**.

Dataset

The dataset used in this assignment is **"NutIntake.csv"**, which contains information about individuals' **smoking habits (smoke)** and **tea consumption (tea)** that we want to focused.

Tasks

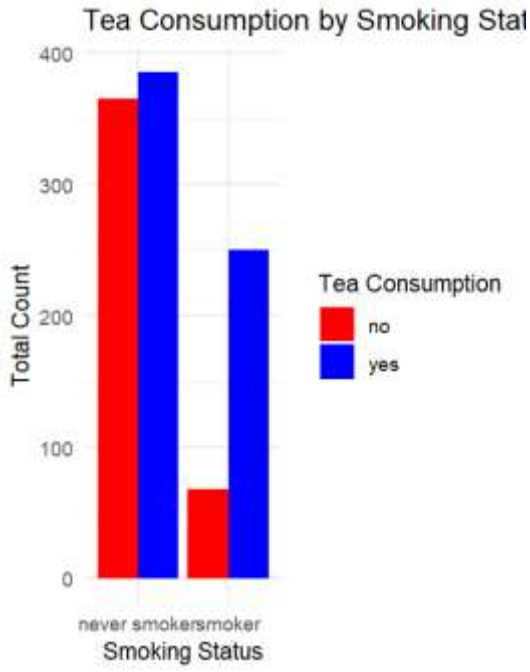
Task1 : Answer what is your null hypothesis and alternative hypothesis

Task 2: Load the Dataset

Write R code to **read** the dataset and check its structure.

Task 3: Make a barplot

Hint! in this case; summarise(total.number = n(), .groups = "drop") would be helpful



Task 4: Conduct Chi-Square Test

Hint! structure of xtabs

- If you specify a **response variable**, xtabs() will **sum** that variable within the contingency table. => xtabs(response_variable ~ predictor1 + predictor2, data = dataset)
- If **no response variable is specified**, xtabs() **counts the number of occurrences** of each combination of predictor variables. This means R will **count** how many times

each combination of smoke and tea appears in the dataset. => xtabs(~ predictor1 + predictor2, data = dataset)

You will get Nut.mat look like this:

```
                    tea
smoke                no yes
  never smoker 365 385
  smoker        68 250
```

Then, perform Chi-square and show the result.

Task 6: Answer that your Chi-square result is rejected or accept your null hypothesis

Submit via Mycourse including 2 files (.R and . html)

The file name is

- **lab3_studentId1_ studentId2_ studentId3.R**
- **lab3_studentId1_ studentId2_ studentId3.html**