**Code:**

**#Step 1: Loading and Preparing the Data**

# Load necessary libraries library (ggplot2)

# Load the dataset

superstore\_data <- read.csv("D:/MSc DS/Semester 1/Retail Market Analysis/Practical/superstore\_data.csv")

**#Step2: Design experiment and load the necessary data.**

# Scenario: An e-commerce platform wants to enhance the checkout process to reduce cart abandonment.

# Problem Statement: Assess the impact of recent purchases ('Recency'), in-store purchases ('NumStorePurchases'), and web purchases ('NumWebPurchases') on customer engagement ('NumWebVisitsMonth').

selected\_data <- superstore\_data[, c("Id","Year\_Birth", "Marital\_Status", "Education", "Dt\_Customer", "Recency", "NumStorePurchases", "NumWebPurchases", "NumWebVisitsMonth")]

selected\_data <- unique(selected\_data) **#Step3: Implement Randomization technique** # Set a seed for reproducibility (optional) set.seed(123)

# Randomly assign treatment and control groups based on specified proportions selected\_data$treatment\_group <- ifelse(runif(nrow(selected\_data)) <= 0.7, "Treatment", "Control") # Split the dataset into treatment and control groups

treatment\_data <- selected\_data[selected\_data$treatment\_group == "Treatment", ] control\_data <- selected\_data[selected\_data$treatment\_group == "Control", ] **#Step4: Implement Simple Random Sample splitting technique**

# Define the size of the sample you want to extract (e.g., 70% of the data) sample\_size <- floor(0.7 \* nrow(selected\_data))

# Perform simple random sampling to select a sample from the dataset

sampled\_data <- selected\_data[sample(1:nrow(selected\_data), size = sample\_size, replace = FALSE), ] # Check the dimensions of the sampled data

dim(sampled\_data)

### Interpretation and Implications:

The following are the Sample Sizes:

* Control Data: 666 observations. * Sampled Data: 568 observations. * Selected Data: 2240 observations.

* Superstore Data: 2240 observations (initial dataset). * Treatment Data: 1574 observations.

* Sample Size: 1568 (seems to be a specified or calculated value).

**Randomization Outcome:** The randomization procedure allocated 666 observations to the control group and 1574 observations to the treatment group. This random assignment helps avoid bias when analyzing the impact of different factors on customer behavior.

**Simple Random Sample:** A simple random sample extracted 568 observations from the initial dataset (superstore data) for analysis. This sample represents a subset of the entire dataset and allows for analysis while retaining the diversity of observations.

**Sample Size:** The specified sample size for analysis appears to be 1568, which is close to the combined size of the control and treatment groups. It might represent the desired sample size for the experiment or analysis.

**Conclusion:** The randomization and sampling procedures have successfully divided the dataset into control and treatment groups and extracted a random sample for analysis.

**Design & Conduct experiments for Marketing Campaigns**