### **Basic Programming Assignment: Clinical Data Analysis [20mins]**

#### **Objective:**

Students can learn the basics of clinical data analysis using a real-world dataset. Students will practice data cleaning, exploratory data analysis (EDA), and basic statistical analysis techniques.

#### **Dataset:**

A dataset representing clinical trial data has been provided. The dataset includes patient demographics, clinical measurements, and treatment outcomes.

Download Dataset [clinical\_trial\_data](https://docs.google.com/spreadsheets/d/1KVVF1eAq24KHIfZLyf_rszfuSQ0ULlm51FlV9t9Py28/edit?gid=0#gid=0)

#### **Assignment Instructions:**

1. **Data Cleaning:**
   * Load the dataset into a Pandas DataFrame.
   * Check for missing values and handle them appropriately (e.g., impute with mean/median, remove rows, etc.).
   * Ensure that data types for each column are correct (e.g., numerical columns are of type float/int, categorical columns are of type object).
2. **Exploratory Data Analysis:**
   * Generate summary statistics for numerical columns (mean, median, standard deviation, etc.).
   * Create visualizations to understand the distribution of key variables (e.g., histograms, box plots).
   * Explore relationships between variables using scatter plots and correlation matrices.
3. **Basic Statistical Analysis:**
   * Compare the mean age of patients between two treatment groups using a t-test.
   * Analyze the distribution of treatment outcomes and calculate the proportion of positive outcomes for each treatment group.
   * Perform a chi-square test to determine if there is a significant association between treatment group and outcome.
4. **Documentation and Reporting:**
   * Include your code with comments explaining each step.
   * ***Write a brief report summarizing your findings from the EDA and statistical analysis. Include the visualizations and statistical results in your report.***