## Project 1: Exploratory Data Analysis (EDA) and Visualization

### Description

This project involves conducting an exploratory data analysis (EDA) on the `iris` dataset and creating various visualizations to understand the dataset's structure and relationships.

### Dataset

- \*\*iris\*\*: Built-in dataset in R containing measurements for iris flowers.

### R Code

1. \*\*Data Cleaning\*\*: Load and clean the dataset.

2. \*\*Histograms\*\*: Create histograms of numerical variables.

3. \*\*Boxplots\*\*: Create boxplots to compare distributions by category.

4. \*\*Scatterplots\*\*: Create scatterplots to explore relationships between variables.

5. \*\*Save Plots\*\*: Save visualizations as PNG files.

### How to Run

1. Install and load the required packages.

2. Copy the R code into a Jupyter Notebook cell and run it.

3. View and interpret the generated plots.

**Part 2:**

### R Code

1. \*\*Data Preprocessing\*\*: Load and preprocess the dataset.

2. \*\*Summary Statistics\*\*: Compute summary statistics grouped by species.

3. \*\*Correlation Matrix\*\*: Compute and visualize the correlation matrix.

4. \*\*Pair Plots\*\*: Generate pair plots using `GGally`.

5. \*\*Advanced Visualizations\*\*:

- Boxplot with jitter.

- Density plot.

- Feature engineering and visualization.

6. \*\*Save Plots\*\*: Save visualizations as PNG files.