Assignment #2: EDA

1. EDA and Statistical Graphics

Exploratory Data Analysis (EDA) uses statistical graphics to identify patterns, trends, and outliers. The primary goal is to uncover hidden insights. While no specific EDA tasks are assigned, your analysis report should include:

- **Distribution of single variables** (with visualizations).
- **Relationships between two variables**.
- **Pairwise comparisons**.

Note: The approach differs for **categorical vs. continuous variables**. Always consider variable types when analyzing relationships.

2. Guidelines for Visual Representation

Visuals (tables and figures) must include:

Tables: Required components of a tabular representation

- **Caption** (clear description).
- **Header** (column labels).

Figures: required graphical components of a figure

- **Title**.
- **Axis labels**.
- **Caption**.
- **Legend** (if applicable).
- **Annotations** (where needed).

Each table/figure must be briefly interpreted based on observed patterns.

A Brief Report on EDA

The assignment is based on the single data file you created in week #3. Write a short report to summarize the above EDA. The required components are

Project Title

- Provide a practical, descriptive, and concise title for the project. **Note:** *You should not simply use Exploratory Data Analysis as your project title since it is too general and does not contain any information related to the data set and potential research questions based on the data set.*

Objective Statement & Background

- Clearly state the objectives of this EDA (Thinking about the critical information you need to know to formulate meaningful analytic questions and perform analysis).
- Describe the **data sources** and their characteristics.

Exploratory Data Analysis (EDA): Formulate some analytic questions you want to address. These questions will guide your EDA.

- Focus on **meaningful results only**. That is,
- Each subsection should cover **one analysis** with visuals.
- Include **narrative interpretations** of patterns.

Results & Discussion

- Summarize key findings.
- Make **inferential statements** based on statistical insights.
- Propose **new research questions** if new patterns emerge.
- **Visuals should support arguments** (avoid raw screenshots; generate directly from R).
- **Interpret all tables/figures** included.

- **Prioritize interactive graphics** where possible.