



Assignment 2 – Descriptive Statistics

This assignment is based on the **Climate Dataset**. The goal is to apply the skills learned during the second Excel lab session to compute summary statistics, visualize distributions, and interpret patterns.

Dataset Information

The dataset contains climate records with the following variables:

- **Year:** Numerical (2000–2025)
- **Month:** Numerical (1–12)
- **Season:** Categorical (Winter, Spring, Summer, Autumn)
- **Region:** Categorical (Africa, Asia, Europe, North America, South America, Oceania)
- **Average_Temperature_C:** Numerical (–10 to 40 °C)
- **Rainfall_mm:** Numerical (0–500 mm)
- **Humidity_percent:** Numerical (30–90%)
- **Wind_Speed_kmh:** Numerical (0–50 km/h)
- **Solar_Radiation_Wpm2:** Numerical (100–800 W/m²)

Objectives

By completing this assignment, you will demonstrate your ability to:

- Compute summary statistics (mean, median, variance, standard deviation)
- Create and interpret frequency tables
- Visualize data using histograms and box plots
- Detect patterns, skewness, and outliers
- Write simple, meaningful interpretations

Tasks

1. Summary Statistics

- For the following variables: Average_Temperature_C, Rainfall_mm, Humidity_percent, Wind_Speed_kmh
- Calculate Mean, Median, Mode

- Calculate Minimum, Maximum, Range
- Calculate Variance and Standard Deviation

2. Frequency Tables

- Create a frequency table for Rainfall_mm grouped into 5 categories (e.g., Very Low, Low, Moderate, High, Very High).
- Create a frequency table for Wind_Speed_kmh with suitable ranges.

3. Visualizations

- Build a histogram for Rainfall_mm and comment on its shape (normal, skewed).
- Build a box plot for Average_Temperature_C by Season (to compare temperature variation across seasons).

4. Interpretation

Write 3–4 bullet-point insights, for example:

- Which variable shows the highest variability?
- Are there outliers in Rainfall or Temperature?
- Does any region/season show unusual patterns?

Submission Guidelines

Submit a single Excel file with the following sheets:

- **Sheet 1:** Cleaned dataset
- **Sheet 2:** Summary statistics
- **Sheet 3:** Frequency tables
- **Sheet 4:** Histograms
- **Sheet 5:** Box plots
- **Sheet 6:** Insights (typed in cells)

Submission Deadline: Sunday 28th September 2025 11:59pm GMT.

NOTE: Please save your assignment with your full name (e.g., *John Doe Assignment 1*) and send it to the following email addresses:

ruchinbox@gmail.com, tijanijumoke410@gmail.com & charlesopondo1376@gmail.com