DS321 – Phase 2 Project Assignment

Objective

In this phase you will expand your project into the foundations of a full research paper. You must:

1. Write a clear **Introduction** explaining the research problem and motivation.
2. Conduct a **Related Work** review based on prior studies.
3. Perform **Exploratory Data Analysis (EDA)** using descriptive statistics and visualizations.

Use the **IEEE template uploaded on e-Learning** for all formatting and citations.

Tasks and Deliverables

1 – Introduction

Write 3–4 academic paragraphs covering:

- The research problem and its importance.
- Real-world relevance of the dataset.
- Key challenges or knowledge gaps.
- Your main research goal or hypothesis.

Use a scientific tone, not bullet points.

2 – Related Work

Prepare a short literature review (15-20 references minimum):

- Summarize previous approaches and findings.
- Identify what limitations or open gaps remain.
- Cite sources in **IEEE style** (numbered [1], [2], …).
- List full references at the end of the paper.

3 – Dataset Description and Statistical Analysis

Provide quantitative information about your dataset:

- Number of samples, features, variable types.
- Summary statistics (mean, median, std dev, min, max).

Then visualize key variables with Python (**pandas**, **matplotlib** or **seaborn**):

Plot	Purpose
Histogram	Show data distribution
Scatter plot	Show correlation between variables
Box plot	Detect outliers

Each figure must include **title, axis labels, and caption**.

Submission Requirements

- **IEEE-formatted PDF** (exported from the template).
- **Python code file** (.ipynb or .py) used for the visualizations.
- Submit both files on **e-Learning** before the deadline.

Notes

- Work must be original; plagiarism = zero grade, AI generated text= zero grade.
- Use **the same dataset you selected in Phase 1**.