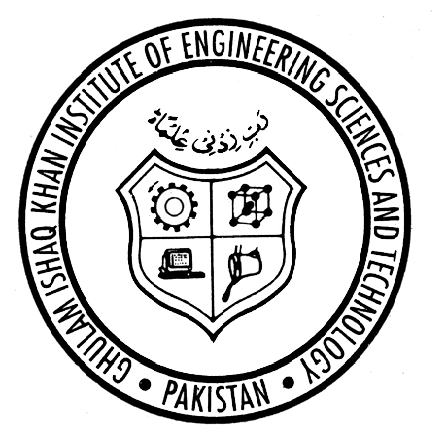
***Ghulam Ishaq Khan Institute of Engineering Sciences and Technology (GIKI)***



***Faculty of Engineering Sciences (FES)***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ES 111** | **Spring 2025** | **Course Project** | **Weight 7%** | **Due: 09 May 2025** |

**CLO 5 🡪 PLO 2: Demonstrate ability to implement learned concepts in a modern tool (such as R, Python).**

# Objective

The students will learn how to extract different types of information from a data set and how to represent that information through different tools. The students will learn to do this using a software tool.

# Process

You are to demonstrate your learning of the software tool through the following steps:

* Find a dataset of your own choice from the internet. The dataset should be large enough for processing but not too large (between 1,000 data points to 10,000 data points, for example).
* Write a code for finding the average and variance of the dataset.
* Write a code to process the data in such a way that it can be represented using a histogram, as well as a pie chart. Frequency distribution of the data set should be clearly mentioned, also through code. An explanation should be included in the report as to how the distribution was attained.
* Write a code to find the average and variance of the dataset using the frequency distribution found in the previous step. Are the values the same as found previously? Explain.
* Using only 80% of the dataset, find a 95% confidence interval for the mean and variance. Now use this data to predict a 95% tolerance interval. Use the remaining 20% data to validate your calculations. This step is to be done through calculations as well as coding.
* Based on your knowledge of statistics, form a hypothesis about the data set. Now write a code to validate your hypothesis. Give an analysis of your findings.

# Solution Required

**Note-1**  Make one report containing main body and annexures (any part which is required as annexure should not be part of the main body).

**Note-2** You can work in group of 2 to 3. You need to identify the team lead. Only team lead will submit the report, other member(s) will not submit any thing. The report is to be written as per the provided template and submitted in pdf format. Late submission will not be accepted. Registration numbers of all team members to be included in the code (as comments) as well as the top of the report. Report should mention the team lead.

# Rubrics

**Report 7%**

**Format 2%**

**Results and Discussion 3%**

**Codes 2%**