

Mechatronics Engineering and Automation Program

CSE473: Computational Intelligence

Lab Assignment #02



Assignment: Analyzing a Dataset with Pandas

Objective:

You will use the pandas library to load a dataset, explore its structure, clean the data, and perform some basic analysis to gain insights.

Dataset:

Use the [Iris dataset](#). The Iris dataset contains measurements of 150 iris flowers from three different species: Iris-setosa, Iris-versicolour, and Iris-virginica. Each flower is described by four features: sepal length, sepal width, petal length, and petal width.

Tasks:

1. Load the dataset:

- Download the Iris dataset from the provided link.
- Load the dataset into a pandas DataFrame.

2. Explore the dataset:

- Display the first 10 rows of the DataFrame.
- Show a summary of the dataset, including the data types and missing values.
- Display basic statistics for each numerical column.

3. Data Cleaning:

- Check for and handle any missing values.
- Ensure all data types are appropriate for analysis.

4. Data Analysis:

- Calculate the mean and median for each numerical column.
- Find the correlation between the numerical features.
- Group the data by species and calculate the mean for each feature within each species.
- Plot histograms for each numerical feature.

Submission:

- Submit your Python script (.py file) containing the implementation of the tasks above.
 - Ensure your script is well-commented and includes any necessary explanations of your approach and results.
-