**Assignment 1**

**Aim:** Preprocess data Using Python.

**Objective:**

* To clean Data and make it noise free
* Prepare Data for Analysis

**Problem Statement:**

**Case Study:**

Download different dataset from kaggle, UCI like diamonds.csv, adult.csv and WorldCupMatches.csv.

Use Python Libraries like pandas, numpy, sklearn etc.

* **Perform different task like basic data view operations.**
* **Perform different data preprocessing tasks like**
* **Handle Missing Values**
* **Tuple duplication**
* **Data Integration**
* **Normalization**
* **Correlation Analysis**

**Theory:**

**Explain the following points with standard function used in python:**

* **Data Preprocessing**
* **Data Cleaning**
* **Data Integration**
* **Data Redundancy**
* **Data Transformation**

**Input:** Dataset

**Output:** Cleaned, integrated, transformed dataset.

**Platform:** Windows

**Conclusion:** Thus, we have learned different preprocessing techniques using python on .csv file.

**FAQ’s:**

1. What are different data types in data mining?
2. What are the applications of correlational analysis?
3. What is noise in data?