



**Pimpri Chinchwad Education Trust's
Pimpri Chinchwad College of Engineering
(An Autonomous Institute Affiliated to SPPU,Pune)**

List of Assignments

Department: Computer Engg.

Academic Year: 2024- 2025

Semester: I

Laboratory: Data Exploration and Visualization (BCE23VS03)

| Assignment No | | Assignment List |
|---------------|---|--|
| 1 | a | Create a dictionary to store student names and their corresponding scores in a test. Write functions to add new students, update scores, delete students, and find the student with the highest score. |
| | b | Using python perform List Operations and list slicing in detail with necessary programs. |
| | c | Perform matrix addition subtraction and multiplication using NumPy library |
| 2 | a | Write a Python program to create a dataframe by importing CSV, JSON files and perform various operations. |
| | b | Perform web scraping for any website (using scrapy/beautiful soap/selenium) |
| | c | Write a Python program to import any dataset from UCI/Kaggle, perform data cleaning and remove the outliers. |
| 3 | a | Download or create any dataset and perform Exploratory data analysis for understanding the data and finding the hidden patterns and perform the following tasks <ul style="list-style-type: none">● Print summary statistics and write the observations if any● Univariate analysis and write the observations if any (Example datasets: automobile/ car dataset , Loan prediction data set or any other) |
| | b | Data Visualization with Matplotlib and Seaborn Dataset: Use the iris dataset or any available datasets. Perform the following tasks: <ul style="list-style-type: none">● Data Loading and Summary |

| | | |
|---|---|--|
| | | <ul style="list-style-type: none"> • Visualizations with Matplotlib • Visualizations with Seaborn |
| | c | <p>Exploratory Data Analysis using any datasets from UCI/Kaggle and Perform the following tasks:</p> <ul style="list-style-type: none"> • Data Loading and Summary Statistics • Data Distribution Analysis • Bivariate / multivariate analysis • Correlation analysis and write observation |
| 4 | a | <p>Download suitable Housing data from kaggle/ UCI / any other source, Predict House prices using Linear Regression and Evaluate the model using RMSE or any other suitable metric. Example Dataset: (https://www.kaggle.com/datasets/yasserh/housing-prices-dataset) (Note: Use sci-kit learn Library)</p> |
| | b | <p>Download Diabetic data from kaggle/ UCI / any other source, Build classification model for diabetic prediction and Evaluate the model using accuracy, precision, Recall etc.,. Example Dataset: (https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database) (Note: Use sci-kit learn Library)</p> |
| | c | <p>Hypothesis test: Assume our business has two units that make pizzas. Check if there is any significant difference in the average diameter of pizzas between the two making units.</p> |
| 5 | | Mini Project |

Let's create some dummy age data for the population of voters in the entire country and a sample of voters in Maharashtra and test whether the average age of voters in Maharashtra differs from the population