***BTech-IT, Sem-7, Term work, DAIE, IT704***

**Dharmsinh Desai University, Nadiad**

**Department of Information Technology**

**DAIE, IT704**

**B.Tech. IT, Sem: VII**

**Submitted By**

**Roll No:** IT076

**Name:** Dishant Modh

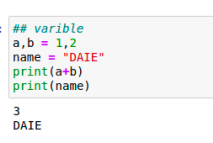
**Experiment – 01 (Introduction to Python Programming)**

**Aim:** Introduction to python programming for Data Mining **Tools/Apparatus:** Visual Studio Code, Jupyter Notebook Extension

**Procedure:**

**1. Python Programming Language Introduction:**

**1.1. Variables:**

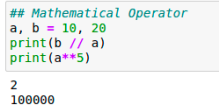
****

**1.2 Logical Operator**

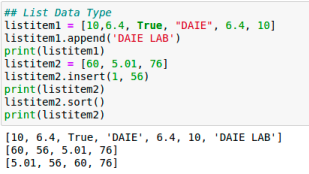
****

1 | P a g e

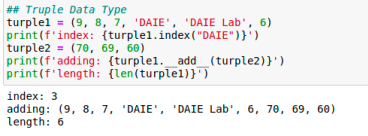
**1.3. Mathematical Operators:**

**1.4 List Data Type**

***BTech-IT, Sem-7, Term work, DAIE, IT704***

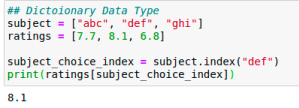
******

**1.5. Tuple Data Type:**

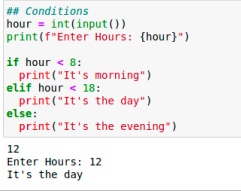
****2 | P a g e

**1.6 Dictionary Data Type**

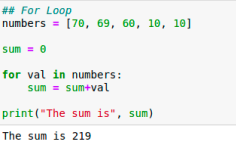
***BTech-IT, Sem-7, Term work, DAIE, IT704***

******

**1.7. Conditions:**

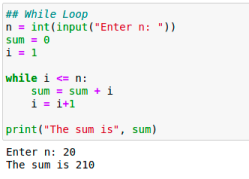
****

**1.8 For Loop**

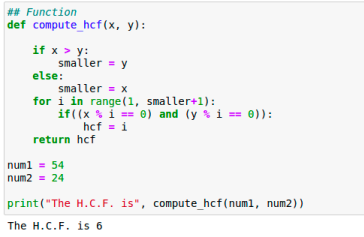
****

3 | P a g e

**1.9 While Loop**

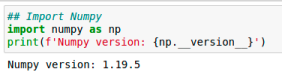
**1.10. Functions:**

***BTech-IT, Sem-7, Term work, DAIE, IT704***

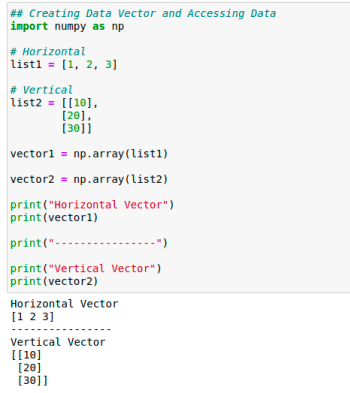
******4 | P a g e

***BTech-IT, Sem-7, Term work, DAIE, IT704***

**2. Handling multi-dimensional data and element-wise operators using Numpy: 2.1. Import Numpy:**

****

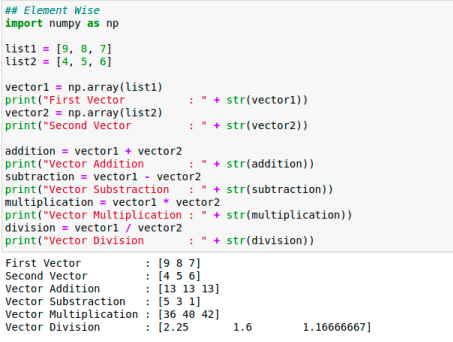
**2.2 Creating Data Vector**

****

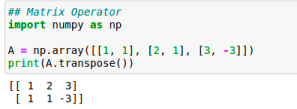
5 | P a g e

**2.3. Element Wise Operators:**

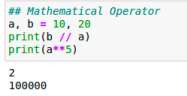
***BTech-IT, Sem-7, Term work, DAIE, IT704***

******

**2.4 Matrix Operation**

****

**2.5. Mathematical Operators:**

****

6 | P a g e

***BTech-IT, Sem-7, Term work, DAIE, IT704***

**Experiment – 02 (Data pre-processing)**

**Aim:** To perform data pre-processing using pandas package in python

**Tools/Apparatus:** Visual Studio Code, Jupyter Notebook Extension

**Procedure:**

**1 Import Pandas**

****

**2 Create Data Series**

****7 | P a g e

**3. Create Data Frame:**

***BTech-IT, Sem-7, Term work, DAIE, IT704***

******

**4. Reading Data from CSV file:**

****

**5. Reading inbuilt Data Sets:**

****

8 | P a g e

***BTech-IT, Sem-7, Term work, DAIE, IT704***

******9 | P a g e

**6. Performing Mathematical Operation:**

***BTech-IT, Sem-7, Term work, DAIE, IT704***

******

**7. Handling Missing Values:**

****10 | P a g e

***BTech-IT, Sem-7, Term work, DAIE, IT704***

******

**8. Data Normalization:**

****11 | P a g e

**8.1 Min- Max Normalization**

***BTech-IT, Sem-7, Term work, DAIE, IT704***

******12 | P a g e

**8.2. Z-score Normalization:**

***BTech-IT, Sem-7, Term work, DAIE, IT704***

******13 | P a g e

***BTech-IT, Sem-7, Term work, DAIE, IT704***

**Experiment – 03 (Statistical Analysis and Visualization of Data)**

**Aim:** To perform statistical analysis and visualization of data using pandas and matplotlib in python. (mean, mod, median, standard deviation, variance, correlation, bar chart, line chart, pie graph, histogram, box plot, scatter plot)

**Tools/Apparatus:** Visual Studio Code, Jupyter Notebook Extension

**Procedure:**

**1.1 Import Numpy**

****

**1.2. Mean, Median, Mode, Standard Deviation and Variance of Data: **

**1.3. Correlation Coefficent and Percentile Data:**

****

14 | P a g e

**2.1. Import Matplotlib for PyPlot: 2.2 Line Chart**

***BTech-IT, Sem-7, Term work, DAIE, IT704***

******15 | P a g e

**2.3. Scatter Plot:**

***BTech-IT, Sem-7, Term work, DAIE, IT704***

****2.4 Bar Chart**

****

16 | P a g e

***BTech-IT, Sem-7, Term work, DAIE, IT704***

**2.5. Pie Chart:**

**2.6 Box Plot**

****

17 | P a g e

***BTech-IT, Sem-7, Term work, DAIE, IT704***

**2.7. Sub Plot:**

****18 | P a g e

***BTech-IT, Sem-7, Term work, DAIE, IT704***

**2.8 Plotting in Pandas**

****19 | P a g e

***BTech-IT, Sem-7, Term work, DAIE, IT704***

******20 | P a g e