**Machine Learning Lab – Eval-01**

**Date: 14-Feb-2023 Max. Marks: 15 Duration: 60 Minutes**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**Note: The built in python functions describe() and info() should not be used for any of the questions to solve.**

**Q1)** Using python code write the following data to a csv file:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **REGNO** | **NAME** | **SUB1** | **SUB2** | **SUB3** |
| 2001 | ABC | 25 | 35 | 45 |
| 2002 | DEF | 35 | 45 | 55 |
| 2003 | GHI | 48 | 76 | 85 |
| 2004 | JKL | 37 | 54 | 74 |
| 2005 | MNO | 48 | 78 | 98 |

Perform the following operations on the above data:

(i) For each subject display the minimum, maximum and mean values

(ii) Determine if a student has failed or passed in a subject. The pass marks in each subject is m40.

**Q2)** For the given dataset titled “**abalone.csv**” perform the following operations on the data set and display the appropriate output:

**Use Pandas library to read the csv file and display the following statistics:**

(a) The dimensions of the entire data set

(b) The data type of each column

(c) Five Number summary of every numeric column

(d) Convert the non-numeric column “**sex**” into a numeric column with the appropriate value

assigned.