BCSE209P

Machine Learning Lab

L45+L46

Evaluation Lab – 01

**SET I**

Answer the following questions.

1) Import the necessary packages /modules and read the data set

2) Drop the **column ‘Unnamed: 0’** and **rename** the columns as given below [hint: rename] (2 Marks).

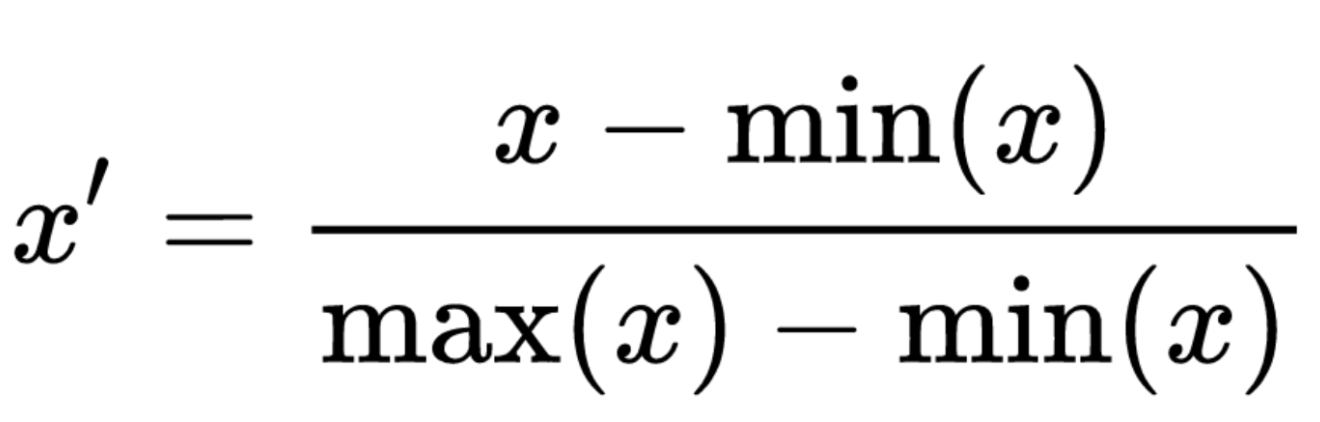
|  |  |
| --- | --- |
| **Given** | **Rename** |
| ride.alc.driver | drunken |
| age4 | age |
| smoke | smoked |
| DriverLicense | driver\_license |
| In\_Month | Income |

3) Fill in the **missing values** using the proper strategy for the columns identified from question number 2. (2 Marks)

4) Visualize the “**female**” using **pie chart** with labels. (2 Marks)

5) Draw the drunken driver details in bar chart [**x axis: No and Yes** **& Y axis: Count**]. (2 Marks)

6) Apply **normalization** on the column “Income” and add the normalized values into the new column called “IncomeV1” (2 Marks).



Note: TypeError: cannot convert the series to <class 'int'> which means you need to convert the obhect types into int

Example :

You need to use astype if df['age'] is object dtype.

df['intage'] = df['age'].astype(int)