Unit 3.1Graded Assignment

Instructions:

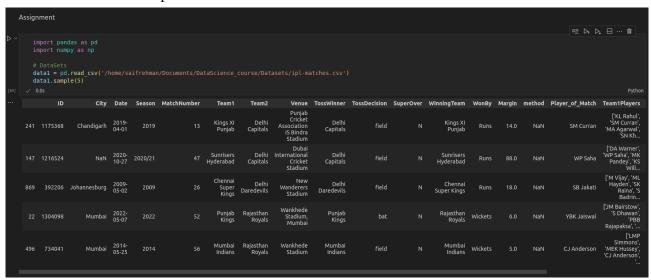
Implement a label encoder for categorical data using pure Python, Pandas and NumPy

Submitted by:

- 1. Ali Nasir (2303.KHI.DEG.012)
- 2. Saif ur Rehman (2303.KHI.DEG.007)

Solution:

1. First we have import our dataset.



2. We have implemented function in two use cases if the run by default it will labelize all the object based column. In second case we will labialize in the desired column. We have saved all the str type data frame a list alled var, which will labelize all the categorical data. The def labelencoder function states that make copy of our data frame in encoded_dataframe variable → used for loop pn dataframe columns → using if condition to check the data type = "O" → using for loop to encode the columns → categories to set the unique values separate → using for loop to emurate the the values assign unique value to a label → The map() function is used to apply the label encoding dictionary (encode) to each value in the column. The astype() method is used to convert the resulting column to the int16 data type for memory efficiency → finally return encoded dataframe

3. When we run the function we can see that it has labialize all the categorical data

