# Third Year B. Tech (EL & CE)

**Semester: VI Subject:** Data Science for Engineering

**Name: Shreerang Mhatre Class: TY**

**Roll No: 52 Batch: A2**

# Experiment No: 03

**Name of the Experiment**: **Data Pre-processing.**

**Performed on: 18/03/2024**

**Submitted on: 18/04/2024**

**Problem Statement:**

Load

Data and perform Data Pre-processing.

Input: df = pd.read\_csv

('https://raw.githubusercontent.com/selva86/datasets/master/Cars93\_miss.csv&#39;)

1. Read a csv file to create a data frame and print top records.

ii. Check if there are any missing values in the data.

iii. Drop null values / Impute the missing values with mean / median.

iv. Import ‘crim’ and ‘medv’ columns of the Boston Housing dataset as a dataframe and ge the nrows, ncolumns, datatype, summary stats of each column of a dataframe.

v. Which manufacturer, model and type has thehighest Price?

vi. How to create one-hot encodings of a categorical variable.



















