import pandas as pd

import numpy as np

dataset=pd.read\_csv("Cars.csv")

dataFrame1=pd.DataFrame(dataset)

#take care of categorical values

modeOfCylinders=dataFrame1['Cylinders'].mode()

dataFrame1['Cylinders'].fillna(modeOfCylinders[0], inplace= True)

modeOfUsedNew=dataFrame1['UsedNew'].mode()

dataFrame1['UsedNew'].fillna(modeOfUsedNew[0], inplace= True)

#take care of numerical values

from sklearn.impute import SimpleImputer

imputer = SimpleImputer(missing\_values = np.nan, strategy = "mean")

imputer = imputer.fit(dataFrame1.iloc[:,4:7])

dataFrame1.iloc[:,4:7] = imputer.transform(dataFrame1.iloc[:,4:7])

#handeling outliers

Q1 = dataFrame1.Price.quantile(0.25)

Q3 = dataFrame1.Price.quantile(0.75)

Q2 = dataFrame1.Price.quantile(0.50)

IQR = Q3 - Q1

lower\_limit = Q1 - 1.5\*IQR

upper\_limit = Q3 + 1.5\*IQR

df\_no\_outlier = dataFrame1[(dataFrame1.Price>lower\_limit)&(dataFrame1.Price<upper\_limit)]

# convert dataset to object

dfIntoObject = df\_no\_outlier.iloc[:,:].values

#Convert Categorical columns to Numerical columns

from sklearn.compose import ColumnTransformer

from sklearn.preprocessing import OneHotEncoder

columnTransformer = ColumnTransformer([('encoder', OneHotEncoder(), [0])],remainder='passthrough')

dfIntoObject = np.array(columnTransformer.fit\_transform(dfIntoObject[:,0:7]), dtype = np.float64)