

▼ Suvodeep Das

Section-B

Roll- 90

▼ IMPORTING PANDAS

```
import pandas as pd
from google.colab import drive
```

▼ EXTRACTING DATA FROM CSV FILE

```
drive.mount('/content/drive')
filename = '/content/drive/My Drive/datasets/auto-mpg.csv'
df = pd.read_csv(filename)
df.head()
```

🔗 Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount(

| | mpg | cylinders | displacement | horsepower | weight | acceleration | model year | origin | name |
|---|------|-----------|--------------|------------|--------|--------------|---------------|--------|------------------------|
| 0 | 18.0 | 8 | 307.0 | 130.0 | 3504 | 12.0 | 70 | 1 | chevro chevi mal |
| 1 | 15.0 | 8 | 350.0 | 165.0 | 3693 | 11.5 | 70 | 1 | bu skyl 3 |

▼ PRINTING ROWS AND COLUMNS

```
df.shape
```

```
(398, 9)
```

```
len(df)
```

```
398
```

▼ DATA TYPE OF FILENAME

```
type(filename)

str
```

▼ SEEING THE NULL VALUES

```
df.isnull()
```

```
df.isnull().sum()
```

```
mpg          0
cylinders    0
displacement 0
horsepower   6
weight       0
acceleration 0
model year   0
origin       0
car name     0
dtype: int64
```

▼ OUTLIERS REMOVAL

```
def remove_outlier(ds,col):  
    quart1 = ds[col].quantile(0.25)  
    quart3= ds[col].quantile(0.75)  
    IQR = quart3 - quart1 #Inter quartile range  
    low_val = quart1 - 1.5*IQR  
    high_val = quart3 + 1.5*IQR  
    ds = ds.loc[(ds[col] > low_val) & (ds[col] < high_val)]  
    return ds  
data = remove_outlier(df, "mpg")  
nrowcount = data.shape[0]  
print(nrowcount)
```

397

✓ 0s completed at 1:41 PM

