e0123016 raagavendran sundar

```python
import pandas as pd

pd.set_option('display.max_rows', 500)
pd.set_option('display.max_columns', 500)
pd.set_option('display.width', 1000)
```

```python
url = "imports-85.csv"
col_names=['symboling','normalized-losses','fuel-type','aspiration','num-of-doors',
                          'length','width','height','curb-weight','engine
                          'fuel-system','bore','stroke','compression-rati

df_car = pd.read_csv(url,sep=',',names = col_names ,na_values="?",  header=None)
```

```python
df_car.shape
```

```python
df_car.head(5)
```

```python
df_car.info()
```

```python
df_car.columns
```

```python
df_car = df_car[[ 'aspiration', 'num-of-doors',  'drive-wheels',  'num-of-cylinders
```

```python
df_car.head()
```

```python
df_car.info()
```

```python
df_car['num-of-doors'].value_counts()
```

```python
door_mapper = {"two": 2,
               "four": 4}
```

```python
df_car['doors'] = df_car["num-of-doors"].replace(door_mapper)
```

```python
df_car.head()
```

```python
df_car['num-of-cylinders'].value_counts()
```

```python
cylinder_mapper = {"two":2,
                   "three":3,
                   "four":4,
                   "five":5,
                   "six":6,
                   "eight":8,
                   "twelve":12}
```

```python
df_car['cylinders'] = df_car['num-of-cylinders'].replace(cylinder_mapper)
```

```python
df_car.head()
```

```python
df_car['drive-wheels'].value_counts()
```

```python
df_car = pd.get_dummies(df_car,columns=['drive-wheels'])
```

```python
df_car.head()
```

```python
df_car['aspiration'].value_counts()
```

```python
df_car = pd.get_dummies(df_car,columns=['aspiration'], drop_first=True)
```

```python
df_car.head()
```

In [1]:
```python
import pandas as pd
url = "imports-85.csv"
col_names = [
  'symboling', 'normalized-losses', 'make', 'fuel-type', 'aspiration',
  'num-of-doors', 'body-style', 'drive-wheels', 'engine-location',
  'wheel-base', 'length', 'width', 'height', 'curb-weight',
  'engine-type', 'num-of-cylinders', 'engine-size', 'fuel-system',
  'bore', 'stroke', 'compression-ratio', 'horsepower', 'peak-rpm',
  'city-mpg', 'highway-mpg', 'price'
]
df_car = pd.read_csv(url, sep=',', names=col_names, na_values="?", header=None)
df_car.columns = df_car.columns.str.strip()
df_car_subset = df_car[[
  'aspiration', 'num-of-doors', 'drive-wheels', 'num-of-cylinders',
  'engine-location', 'fuel-type'
]].copy()
print(df_car_subset.head())
```

```
/home/ec2-user/anaconda3/envs/python3/lib/python3.10/site-packages/pandas/core/com
putation/expressions.py:21: UserWarning: Pandas requires version '2.8.4' or newer
of 'numexpr' (version '2.7.3' currently installed).
  from pandas.core.computation.check import NUMEXPR_INSTALLED
  aspiration num-of-doors drive-wheels  num-of-cylinders  engine-location  \
0        two   convertible        front               130             88.6
1        two   convertible        front               130             88.6
2        two     hatchback        front               152             94.5
3       four         sedan        front               109             99.8
4       four         sedan        front               136             99.4

  fuel-type
0       std
1       std
2       std
3       std
4       std
```

In [5]:
```python
df_encoded = pd.get_dummies(df_car_subset, columns=['engine-location', 'fuel-type',
# Display the final result
print("Encoded Dataset (No Filtering):")
print(df_encoded.head())
```

```
Encoded Dataset (No Filtering):
  aspiration num-of-doors drive-wheels  num-of-cylinders  \
0        two  convertible        front               130
1        two  convertible        front               130
2        two    hatchback        front               152
3       four        sedan        front               109
4       four        sedan        front               136

   engine-location_86.6  engine-location_88.4  engine-location_88.6  \
0                 False                 False                  True
1                 False                 False                  True
2                 False                 False                 False
3                 False                 False                 False
4                 False                 False                 False

   engine-location_89.5  engine-location_91.3  engine-location_93.0  ...  \
0                 False                 False                 False  ...
1                 False                 False                 False  ...
2                 False                 False                 False  ...
3                 False                 False                 False  ...
4                 False                 False                 False  ...

   engine-location_106.7  engine-location_107.9  engine-location_108.0  \
0                  False                  False                  False
1                  False                  False                  False
2                  False                  False                  False
3                  False                  False                  False
4                  False                  False                  False

   engine-location_109.1  engine-location_110.0  engine-location_112.0  \
0                  False                  False                  False
1                  False                  False                  False
2                  False                  False                  False
3                  False                  False                  False
4                  False                  False                  False

   engine-location_113.0  engine-location_114.2  engine-location_115.6  \
0                  False                  False                  False
1                  False                  False                  False
2                  False                  False                  False
3                  False                  False                  False
4                  False                  False                  False

   engine-location_120.9
0                  False
1                  False
2                  False
3                  False
4                  False

[5 rows x 112 columns]
```