Aim :

   To perform Feature Generation techniques in the given data set

Program Code :

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
df=pd.read_excel("/content/Workshop_feature Engg.csv.xlsx")
df1 = df.copy()
df1.head()
df1.info()
df1.isnull().sum()
from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()
df1['fueltype'] = le.fit_transform(df1['fueltype'])
sns.set(style ="darkgrid")
sns.countplot(df1['fueltype'])
from sklearn.preprocessing import OneHotEncoder
enc = OneHotEncoder()
df2 = df.copy()
enc = enc.fit_transform(df2[['CarName']]).toarray()
encoded_colm = pd.DataFrame(enc)
df2 = pd.concat([df2, encoded_colm], axis=1)
df2 = df.drop(['CarName'], axis=1)
df2.head(10)
df3 = df.copy()
df3 = pd.get_dummies(df3, prefix=['aspiration'], columns=['aspiration'])
df3.head(5)
df4 = df.copy()
df4 = pd.get_dummies(df4, prefix=['doornumber'], columns=['doornumber'])
df4.head()
df5 = df.copy()
df5 = pd.get_dummies(df5, prefix=['carbody'], columns=['carbody'])
df5.head()
df6 = df.copy()
df6 = pd.get_dummies(df6, prefix=['drivewheel'], columns=['drivewheel'])
df6.head()
df7 = df.copy()
df7 = pd.get_dummies(df7, prefix=['enginelocation'], columns=['enginelocation'])
df7.head()
df8 = df.copy()
df8 = pd.get_dummies(df8, prefix=['enginetype'], columns=['enginetype'])
df8.head()
df9 = df.copy()
df9 = pd.get_dummies(df9, prefix=['cylindernumber'], columns=['cylindernumber'])
df9.head()
```

Output :

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
df=pd.read_excel("/content/Workshop_feature Engg.csv.xlsx")
df1 = df.copy()
df1.head()
```

	car_ID	CarName	fueltype	aspiration	doornumber	carbody	drivewheel	enginelocation	wheelbase	carlength	carwidth	carheight	curbweight	enginetype	cylindernumber	enginesize	fuelsystem
0	1	alfa-romero giulia	gas	std	two	convertible	rwd	front	88.6	168.8	64.1	48.8	2548	dohc	four	130	mpfi
1	2	alfa-romero stelvio	gas	std	two	convertible	rwd	front	88.6	168.8	64.1	48.8	2548	dohc	four	130	mpfi
2	28	dodge coronet custom	gas	turbo	two	sedan	fwd	front	93.7	157.3	63.8	50.6	2191	ohc	four	98	mpfi
3	29	dodge dart custom	gas	std	four	wagon	fwd	front	103.3	174.6	64.6	59.8	2535	ohc	four	122	2bbl
4	30	dodge coronet custom (sw)	gas	turbo	two	hatchback	fwd	front	95.9	173.2	66.3	50.2	2811	ohc	four	156	mfi

```
df1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 19 entries, 0 to 18
Data columns (total 17 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   car_ID          19 non-null     int64
 1   CarName         19 non-null     object
 2   fueltype        19 non-null     object
 3   aspiration      19 non-null     object
 4   doornumber      19 non-null     object
 5   carbody         19 non-null     object
 6   drivewheel      19 non-null     object
 7   enginelocation  19 non-null     object
 8   wheelbase       19 non-null     float64
 9   carlength       19 non-null     float64
 10  carwidth        19 non-null     float64
 11  carheight       19 non-null     float64
 12  curbweight      19 non-null     int64
 13  enginetype      19 non-null     object
 14  cylindernumber  19 non-null     object
 15  enginesize      19 non-null     int64
 16  fuelsystem      19 non-null     object
dtypes: float64(4), int64(3), object(10)
memory usage: 2.6+ KB
```

```
df1.isnull().sum()
```

```
car_ID            0
CarName           0
fueltype          0
aspiration        0
doornumber        0
carbody           0
drivewheel        0
enginelocation    0
wheelbase         0
carlength         0
carwidth          0
carheight         0
curbweight        0
enginetype        0
cylindernumber    0
enginesize        0
fuelsystem        0
dtype: int64
```

```python
from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()
df1['fueltype'] = le.fit_transform(df1['fueltype'])
sns.set(style ="darkgrid")
sns.countplot(df1['fueltype'])
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:43: FutureWarning: Pass the following variable as a k
  FutureWarning
<matplotlib.axes._subplots.AxesSubplot at 0x7f04400f6cd0>
```
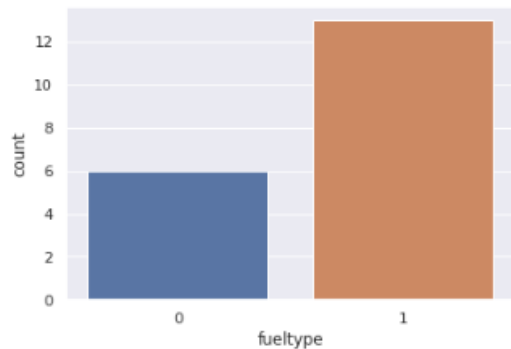


```python
from sklearn.preprocessing import OneHotEncoder
enc = OneHotEncoder()
df2 = df.copy()
enc = enc.fit_transform(df2[['CarName']]).toarray()
encoded_colm = pd.DataFrame(enc)
```

```python
df2 = pd.concat([df2, encoded_colm], axis=1)
df2 = df.drop(['CarName'], axis=1)
df2.head(10)
```

	car_ID	fueltype	aspiration	doornumber	carbody	drivewheel	enginelocation	wheelbase	carlength	carwidth	carheight	curbweight	enginetype	cylindernumber	enginesize	fuelsystem
0	1	gas	std	two	convertible	rwd	front	88.6	168.8	64.1	48.8	2548	dohc	four	130	mpfi
1	2	gas	std	two	convertible	rwd	front	88.6	168.8	64.1	48.8	2548	dohc	four	130	mpfi
2	28	gas	turbo	two	sedan	fwd	front	93.7	157.3	63.8	50.6	2191	ohc	four	98	mpfi
3	29	gas	std	four	wagon	fwd	front	103.3	174.6	64.6	59.8	2535	ohc	four	122	2bbl
4	30	gas	turbo	two	hatchback	fwd	front	95.9	173.2	66.3	50.2	2811	ohc	four	156	mfi
5	35	gas	std	two	hatchback	fwd	front	93.7	150.0	64.0	52.6	1956	ohc	four	92	1bbl
6	61	gas	std	four	sedan	fwd	front	98.8	177.8	66.5	55.5	2410	ohc	four	122	2bbl
7	62	gas	std	two	hatchback	fwd	front	98.8	177.8	66.5	53.7	2385	ohc	four	122	2bbl
8	63	gas	std	four	sedan	fwd	front	98.8	177.8	66.5	55.5	2410	ohc	four	122	2bbl
9	64	diesel	std	four	sedan	fwd	front	98.8	177.8	66.5	55.5	2443	ohc	four	122	idi

```python
df3 = df.copy()
df3 = pd.get_dummies(df3, prefix=['aspiration'], columns=['aspiration'])
df3.head(5)
```

	car_ID	CarName	fueltype	doornumber	carbody	drivewheel	enginelocation	wheelbase	carlength	carwidth	carheight	curbweight	enginetype	cylindernumber	enginesize	fuelsystem	aspiration_std	aspiration_turbo
0	1	alfa-romero giulia	gas	two	convertible	rwd	front	88.6	168.8	64.1	48.8	2548	dohc	four	130	mpfi	1	0
1	2	alfa-romero stelvio	gas	two	convertible	rwd	front	88.6	168.8	64.1	48.8	2548	dohc	four	130	mpfi	1	0
2	28	dodge coronet custom	gas	two	sedan	fwd	front	93.7	157.3	63.8	50.6	2191	ohc	four	98	mpfi	0	1
3	29	dodge dart custom	gas	four	wagon	fwd	front	103.3	174.6	64.6	59.8	2535	ohc	four	122	2bbl	1	0
4	30	dodge coronet custom (sw)	gas	two	hatchback	fwd	front	95.9	173.2	66.3	50.2	2811	ohc	four	156	mfi	0	1

```python
df4 = df.copy()
df4 = pd.get_dummies(df4, prefix=['doornumber'], columns=['doornumber'])
df4.head()
```

	car_ID	CarName	fueltype	aspiration	carbody	drivewheel	enginelocation	wheelbase	carlength	carwidth	carheight	curbweight	enginetype	cylindernumber	enginesize	fuelsystem	doornumber_four	doornumber_two
0	1	alfa-romero giulia	gas	std	convertible	rwd	front	88.6	168.8	64.1	48.8	2548	dohc	four	130	mpfi	0	1
1	2	alfa-romero stelvio	gas	std	convertible	rwd	front	88.6	168.8	64.1	48.8	2548	dohc	four	130	mpfi	0	1
2	28	dodge coronet custom	gas	turbo	sedan	fwd	front	93.7	157.3	63.8	50.6	2191	ohc	four	98	mpfi	0	1
3	29	dodge dart custom	gas	std	wagon	fwd	front	103.3	174.6	64.6	59.8	2535	ohc	four	122	2bbl	1	0
4	30	dodge coronet custom (sw)	gas	turbo	hatchback	fwd	front	95.9	173.2	66.3	50.2	2811	ohc	four	156	mfi	0	1

```python
df5 = df.copy()
df5 = pd.get_dummies(df5, prefix=['carbody'], columns=['carbody'])
df5.head()
```

	car_ID	CarName	fueltype	aspiration	doornumber	drivewheel	enginelocation	wheelbase	carlength	carwidth	...	curbweight	enginetype	cylindernumber	enginesize	fuelsystem	carbody_convertible	carbody_hardtop	carbody_hatchback	carbody_sedan	carbody_wagon
0	1	alfa-romero giulia	gas	std	two	rwd	front	88.6	168.8	64.1	...	2548	dohc	four	130	mpfi	1	0	0	0	0
1	2	alfa-romero stelvio	gas	std	two	rwd	front	88.6	168.8	64.1	...	2548	dohc	four	130	mpfi	1	0	0	0	0
2	28	dodge coronet custom	gas	turbo	two	fwd	front	93.7	157.3	63.8	...	2191	ohc	four	98	mpfi	0	0	0	1	0
3	29	dodge dart custom	gas	std	four	fwd	front	103.3	174.6	64.6	...	2535	ohc	four	122	2bbl	0	0	0	0	1
4	30	dodge coronet custom (sw)	gas	turbo	two	fwd	front	95.9	173.2	66.3	...	2811	ohc	four	156	mfi	0	0	1	0	0

5 rows × 21 columns

```python
df6 = df.copy()
df6 = pd.get_dummies(df6, prefix=['drivewheel'], columns=['drivewheel'])
df6.head()
```

	car_ID	CarName	fueltype	aspiration	doornumber	carbody	enginelocation	wheelbase	carlength	carwidth	carheight	curbweight	enginetype	cylindernumber	enginesize	fuelsystem	drivewheel_fwd	drivewheel_rwd
0	1	alfa-romero giulia	gas	std	two	convertible	front	88.6	168.8	64.1	48.8	2548	dohc	four	130	mpfi	0	1
1	2	alfa-romero stelvio	gas	std	two	convertible	front	88.6	168.8	64.1	48.8	2548	dohc	four	130	mpfi	0	1
2	28	dodge coronet custom	gas	turbo	two	sedan	front	93.7	157.3	63.8	50.6	2191	ohc	four	98	mpfi	1	0
3	29	dodge dart custom	gas	std	four	wagon	front	103.3	174.6	64.6	59.8	2535	ohc	four	122	2bbl	1	0
4	30	dodge coronet custom (sw)	gas	turbo	two	hatchback	front	95.9	173.2	66.3	50.2	2811	ohc	four	156	mfi	1	0

```python
df7 = df.copy()
df7 = pd.get_dummies(df7, prefix=['enginelocation'], columns=['enginelocation'])
df7.head()
```

	car_ID	CarName	fueltype	aspiration	doornumber	carbody	drivewheel	wheelbase	carlength	carwidth	carheight	curbweight	enginetype	cylindernumber	enginesize	fuelsystem	enginelocation_front
0	1	alfa-romero giulia	gas	std	two	convertible	rwd	88.6	168.8	64.1	48.8	2548	dohc	four	130	mpfi	1
1	2	alfa-romero stelvio	gas	std	two	convertible	rwd	88.6	168.8	64.1	48.8	2548	dohc	four	130	mpfi	1
2	28	dodge coronet custom	gas	turbo	two	sedan	fwd	93.7	157.3	63.8	50.6	2191	ohc	four	98	mpfi	1
3	29	dodge dart custom	gas	std	four	wagon	fwd	103.3	174.6	64.6	59.8	2535	ohc	four	122	2bbl	1
4	30	dodge coronet custom (sw)	gas	turbo	two	hatchback	fwd	95.9	173.2	66.3	50.2	2811	ohc	four	156	mfi	1

```python
df8 = df.copy()
df8 = pd.get_dummies(df8, prefix=['enginetype'], columns=['enginetype'])
df8.head()
```

	car_ID	CarName	fueltype	aspiration	doornumber	carbody	drivewheel	enginelocation	wheelbase	carlength	carwidth	carheight	curbweight	cylindernumber	enginesize	fuelsystem	enginetype_dohc	enginetype_ohc	enginetype_ohcv
0	1	alfa-romero giulia	gas	std	two	convertible	rwd	front	88.6	168.8	64.1	48.8	2548	four	130	mpfi	1	0	0
1	2	alfa-romero stelvio	gas	std	two	convertible	rwd	front	88.6	168.8	64.1	48.8	2548	four	130	mpfi	1	0	0
2	28	dodge coronet custom	gas	turbo	two	sedan	fwd	front	93.7	157.3	63.8	50.6	2191	four	98	mpfi	0	1	0
3	29	dodge dart custom	gas	std	four	wagon	fwd	front	103.3	174.6	64.6	59.8	2535	four	122	2bbl	0	1	0
4	30	dodge coronet custom (sw)	gas	turbo	two	hatchback	fwd	front	95.9	173.2	66.3	50.2	2811	four	156	mfi	0	1	0

```python
df9 = df.copy()
df9 = pd.get_dummies(df9, prefix=['cylindernumber'], columns=['cylindernumber'])
df9.head()
```

	car_ID	CarName	fueltype	aspiration	doornumber	carbody	drivewheel	enginelocation	wheelbase	carlength	carwidth	carheight	curbweight	enginetype	enginesize	fuelsystem	cylindernumber_eight	cylindernumber_five	cylindernumber_four
0	1	alfa-romero giulia	gas	std	two	convertible	rwd	front	88.6	168.8	64.1	48.8	2548	dohc	130	mpfi	0	0	1
1	2	alfa-romero stelvio	gas	std	two	convertible	rwd	front	88.6	168.8	64.1	48.8	2548	dohc	130	mpfi	0	0	1
2	28	dodge coronet custom	gas	turbo	two	sedan	fwd	front	93.7	157.3	63.8	50.6	2191	ohc	98	mpfi	0	0	1
3	29	dodge dart custom	gas	std	four	wagon	fwd	front	103.3	174.6	64.6	59.8	2535	ohc	122	2bbl	0	0	1
4	30	dodge coronet custom (sw)	gas	turbo	two	hatchback	fwd	front	95.9	173.2	66.3	50.2	2811	ohc	156	mfi	0	0	1