```python
In [53]: import pandas as pd
         import numpy as np
         import matplotlib.pyplot as plt
         import seaborn as snb
         from sklearn.cluster import    KMeans
         from scipy.spatial.distance import cdist
```

```python
In [75]: var=pd.read_csv('C://Users/Gopi/Desktop/machine learning/csv files/cars.csv')
         var.head(3)
```

Out[75]:

	HP	MPG	VOL	SP	WT
0	49	53.700681	89	104.185353	28.762059
1	55	50.013401	92	105.461264	30.466833
2	55	50.013401	92	105.461264	30.193597

```python
In [76]: var1 =(var-var.min())/(var.max()-  var.min())
         var1.head(3)
```

Out[76]:

	HP	MPG	VOL	SP	WT
0	0.000000	1.000000	0.354545	0.065975	0.349986
1	0.021978	0.911362	0.381818	0.084193	0.395709
2	0.021978	0.911362	0.381818	0.084193	0.388381

```python
In [84]: model=KMeans(n_clusters=5)
         abc = model.fit_predict(var1)
         print(abc)
```

```
[4 4 4 4 4 4 4 0 0 4 0 0 4 0 3 4 4 4 4 0 4 3 4 3 4 3 0 3 3 4 3 3 3 3 3 3 3
 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 1 1 3 3 3 3 1 3 1 1 1 0 1 1 1 1 2 1 1
 1 1 2 1 2 1 1]
```

```python
In [82]: temp = model.predict(var1)
         print(temp)
```

```
[0 0 0 0 0 0 0 4 4 0 0 4 3 0 0 0 0 4 0 3 3 0 3 0 3 0 3 4 3 3 0 3 3 3 3 3 3 3
 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 1 1 3 3 3 3 1 3 1 1 1 1 4 1 1 1 1 2 1 1 1
 1 1 2 1 2 1 1]
```

```python
In [85]: var['clusters']=pd.DataFrame(temp)
         print(var.head(3))
```

```
   HP        MPG  VOL          SP         WT  clusters
0  49  53.700681   89  104.185353  28.762059         4
1  55  50.013401   92  105.461264  30.466833         4
2  55  50.013401   92  105.461264  30.193597         4
```

```python
In [80]: var['clusters'].value_counts()
```

```
Out[80]: 3    36
         1    19
         0    17
         4     6
         2     3
         Name: clusters, dtype: int64
```