In [1]:

```
from sklearn.cluster import KMeans
import pandas as pd
from sklearn.preprocessing import MinMaxScaler
from matplotlib import pyplot as plt
%matplotlib inline
```

In [2]:

```
df=pd.read_csv(r"C:\Users\Shubham\Desktop\Machine learning\EastWestAirlines.csv")
```

In [3]:

```
df.head()
```

Out[3]:

	ID#	Balance	Qual_miles	cc1_miles	cc2_miles	cc3_miles	Bonus_miles	Bonus_trans	Flight_miles_12mo	Flight_trans_12	Days_si
0	1	28143	0	1	1	1	174	1	0	0	
1	2	19244	0	1	1	1	215	2	0	0	
2	3	41354	0	1	1	1	4123	4	0	0	
3	4	14776	0	1	1	1	500	1	0	0	
4	5	97752	0	4	1	1	43300	26	2077	4	

In [13]:

```
df1=df.drop('ID#',axis=1)
df1
```

Out[13]:

	Balance	Qual_miles	cc1_miles	cc2_miles	cc3_miles	Bonus_miles	Bonus_trans	Flight_miles_12mo	Flight_trans_12	Days_sin
0	28143	0	1	1	1	174	1	0	0	
1	19244	0	1	1	1	215	2	0	0	
2	41354	0	1	1	1	4123	4	0	0	
3	14776	0	1	1	1	500	1	0	0	
4	97752	0	4	1	1	43300	26	2077	4	
5	16420	0	1	1	1	0	0	0	0	
6	84914	0	3	1	1	27482	25	0	0	
7	20856	0	1	1	1	5250	4	250	1	
8	443003	0	3	2	1	1753	43	3850	12	
9	104860	0	3	1	1	28426	28	1150	3	
10	40091	0	2	1	1	7278	10	0	0	
11	96522	0	5	1	1	61105	19	0	0	
12	43382	0	2	1	1	11150	20	0	0	
13	43097	0	1	1	1	3258	6	0	0	
14	17648	0	1	1	1	0	0	0	0	
15	28495	0	4	1	1	49442	15	0	0	
16	51890	0	4	1	1	48963	16	0	0	
17	13958	0	1	1	1	4291	5	0	0	
18	91473	0	3	1	1	27408	17	0	0	
19	23354	0	3	1	1	10447	5	0	0	
20	120576	0	5	1	1	58831	23	250	2	

	Balance	Qual_miles	cc1_miles	cc2_miles	cc3_miles	Bonus_miles	Bonus_trans	Flight_miles_12mo	Flight_trans_12	Days_sin
21	185681	2024	1	1	1	13300	16	1800	9	
22	20584	0	1	1	1	3450	11	3450	11	
23	66275	0	1	1	1	2533	11	150	1	
24	205651	500	1	1	1	4025	21	700	4	
25	20726	0	1	1	1	1375	4	0	0	
26	18521	0	1	1	1	1227	2	1227	2	
27	8828	0	1	1	1	0	0	0	0	
28	59763	0	3	1	1	33772	20	100	1	
29	19221	0	1	1	1	4655	8	500	1	
...	...	...	...	...	...	...	...	...	...	
3969	12532	0	1	1	1	2069	9	0	0	
3970	52584	0	1	1	1	2500	1	0	0	
3971	20954	0	1	1	1	375	3	0	0	
3972	35185	0	2	1	1	5957	7	0	0	
3973	72297	0	3	1	1	16241	16	0	0	
3974	58387	0	3	1	1	14581	39	0	0	
3975	9128	0	1	1	1	0	0	0	0	
3976	37520	0	1	1	1	19924	6	0	0	
3977	9399	0	1	1	1	2125	10	0	0	
3978	10071	0	2	1	1	27701	16	0	0	
3979	57793	0	3	1	1	20959	15	1198	3	
3980	28867	0	3	1	1	19169	28	0	0	
3981	1010	0	1	1	1	0	0	0	0	
3982	11463	0	1	1	1	339	4	0	0	
3983	26173	0	1	1	1	305	1	0	0	
3984	404	0	1	1	1	550	3	0	0	
3985	59017	0	4	1	1	34746	25	0	0	
3986	34235	0	1	1	1	18910	7	250	1	
3987	11933	0	1	1	1	249	3	79	1	
3988	5000	0	1	1	1	2125	3	0	0	
3989	2622	0	1	1	1	1625	6	0	0	
3990	11310	0	1	1	1	5021	2	0	0	
3991	39142	0	3	1	1	14981	28	0	0	
3992	11181	0	1	1	1	929	12	0	0	
3993	3974	0	1	1	1	365	3	0	0	
3994	18476	0	1	1	1	8525	4	200	1	
3995	64385	0	1	1	1	981	5	0	0	
3996	73597	0	3	1	1	25447	8	0	0	
3997	54899	0	1	1	1	500	1	500	1	
3998	3016	0	1	1	1	0	0	0	0	

3999 rows × 11 columns

In [14]:

```
scaler=MinMaxScaler()
scaler.fit(df1)
df1=scaler.transform(df1)
```

In [15]:

```
df1
```

Out[15]:

```
array([[0.01650773, 0.        , 0.        , ..., 0.        , 0.84374246,
        0.        ],
       [0.01128788, 0.        , 0.        , ..., 0.        , 0.83988425,
        0.        ],
       [0.02425685, 0.        , 0.        , ..., 0.        , 0.84784181,
        0.        ],
       ...,
       [0.0431695 , 0.        , 0.5       , ..., 0.        , 0.16879672,
        1.        ],
       [0.03220189, 0.        , 0.        , ..., 0.01886792, 0.16867615,
        0.        ],
       [0.00176908, 0.        , 0.        , ..., 0.        , 0.16831444,
        0.        ]])
```

In [16]:

```
k_rng=range(1,10)
sse=[]
for k in k_rng:
    km=KMeans(n_clusters=k)
    km.fit(df1)
    sse.append(km.inertia_)
```

In [17]:

```
sse
```

Out[17]:

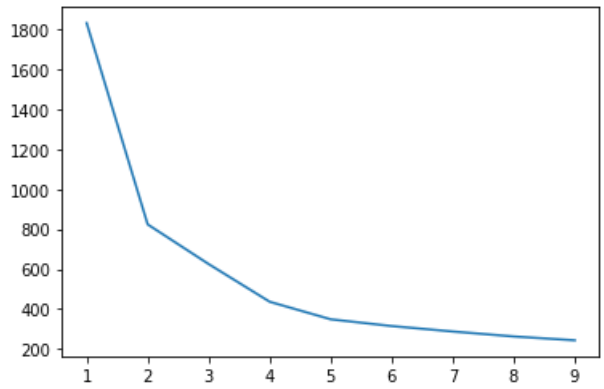
```
[1830.7932128584112,
 823.6756984125207,
 625.168488157072,
 436.7088576193262,
 348.9433217254137,
 315.15077900014745,
 287.9277021471628,
 263.048492559619,
 243.93051296436536]
```

In [18]:

```
plt.plot(k_rng,sse)
```

Out[18]:

```
[<matplotlib.lines.Line2D at 0x1487baae550>]
```



In [19]:

```
km=KMeans(n_clusters=4)
y_pred=km.fit_predict(df1)
```

In [20]:

```
y_pred
```

Out[20]:

```
array([0, 0, 0, ..., 3, 0, 0])
```

In [21]:

```
df['clusters']=y_pred
```

In [22]:

```
df
```

Out[22]:

	ID#	Balance	Qual_miles	cc1_miles	cc2_miles	cc3_miles	Bonus_miles	Bonus_trans	Flight_miles_12mo	Flight_trans_12	Da
0	1	28143	0	1	1	1	174	1	0	0	
1	2	19244	0	1	1	1	215	2	0	0	
2	3	41354	0	1	1	1	4123	4	0	0	
3	4	14776	0	1	1	1	500	1	0	0	
4	5	97752	0	4	1	1	43300	26	2077	4	
5	6	16420	0	1	1	1	0	0	0	0	
6	7	84914	0	3	1	1	27482	25	0	0	
7	8	20856	0	1	1	1	5250	4	250	1	
8	9	443003	0	3	2	1	1753	43	3850	12	
9	10	104860	0	3	1	1	28426	28	1150	3	
10	11	40091	0	2	1	1	7278	10	0	0	
11	12	96522	0	5	1	1	61105	19	0	0	
12	13	43382	0	2	1	1	11150	20	0	0	
13	14	43097	0	1	1	1	3258	6	0	0	
14	15	17648	0	1	1	1	0	0	0	0	
15	16	28495	0	4	1	1	49442	15	0	0	
16	17	51890	0	4	1	1	48963	16	0	0	
17	18	13958	0	1	1	1	4291	5	0	0	
18	19	91473	0	3	1	1	27408	17	0	0	
19	20	23354	0	3	1	1	10447	5	0	0	
20	21	120576	0	5	1	1	58831	23	250	2	
21	22	185681	2024	1	1	1	13300	16	1800	9	
22	23	20584	0	1	1	1	3450	11	3450	11	
23	24	66275	0	1	1	1	2533	11	150	1	
24	25	205651	500	1	1	1	4025	21	700	4	
25	26	20726	0	1	1	1	1375	4	0	0	
26	27	18521	0	1	1	1	1227	2	1227	2	
27	28	8828	0	1	1	1	0	0	0	0	
28	29	59763	0	3	1	1	33772	20	100	1	
29	30	19221	0	1	1	1	4655	8	500	1	
...	...	...	...	...	...	...	...	...	...	...	
3969	3992	12532	0	1	1	1	2069	9	0	0	
3970	3993	52584	0	1	1	1	2500	1	0	0	
3971	3994	20954	0	1	1	1	375	3	0	0	
3972	3995	35185	0	2	1	1	5957	7	0	0	
3973	3996	72297	0	3	1	1	16241	16	0	0	
3974	3997	58387	0	3	1	1	14581	39	0	0	

	ID#	Balance	Qual_miles	cc1_miles	cc2_miles	cc3_miles	Bonus_miles	Bonus_trans	Flight_miles_12mo	Flight_trans_12	Da
3975	3998	21128	0	1	1	1	0	0	0	0	
3976	3999	37520	0	1	1	1	19924	6	0	0	
3977	4000	9399	0	1	1	1	2125	10	0	0	
3978	4001	10071	0	2	1	1	27701	16	0	0	
3979	4002	57793	0	3	1	1	20959	15	1198	3	
3980	4003	28867	0	3	1	1	19169	28	0	0	
3981	4004	1010	0	1	1	1	0	0	0	0	
3982	4005	11463	0	1	1	1	339	4	0	0	
3983	4006	26173	0	1	1	1	305	1	0	0	
3984	4007	404	0	1	1	1	550	3	0	0	
3985	4008	59017	0	4	1	1	34746	25	0	0	
3986	4009	34235	0	1	1	1	18910	7	250	1	
3987	4010	11933	0	1	1	1	249	3	79	1	
3988	4011	5000	0	1	1	1	2125	3	0	0	
3989	4012	2622	0	1	1	1	1625	6	0	0	
3990	4013	11310	0	1	1	1	5021	2	0	0	
3991	4014	39142	0	3	1	1	14981	28	0	0	
3992	4015	11181	0	1	1	1	929	12	0	0	
3993	4016	3974	0	1	1	1	365	3	0	0	
3994	4017	18476	0	1	1	1	8525	4	200	1	
3995	4018	64385	0	1	1	1	981	5	0	0	
3996	4019	73597	0	3	1	1	25447	8	0	0	
3997	4020	54899	0	1	1	1	500	1	500	1	
3998	4021	3016	0	1	1	1	0	0	0	0	

3999 rows × 13 columns

In [24]:

```
---------------------------------------------------------------------------
AttributeError                            Traceback (most recent call last)
<ipython-input-24-8c9895330210> in <module>
----> 1 KMeans.cluster_centers_

AttributeError: type object 'KMeans' has no attribute 'cluster_centers_'
```

In [ ]: