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

[In 2]:
```
df = pd.read_csv("C:\MACHINE LEARNING\ly-master\Mll_bmaara\Exercise\Mall_Customers.csv")
```

[In 3]:
```
df
```

[Out 3]:

	CustomerID	Gender	Age	Annual Income (k$)	Spending Score (1-100)
0	1	Male	19	15	39
1	2	Male	21	15	81
2	3	Female	20	16	6
3	4	Female	23	16	77
4	5	Female	31	17	40
...	...	...	...	...	...
195	196	Female	35	120	79
196	197	Female	45	126	28
197	198	Male	32	126	74
198	199	Male	32	137	18
199	200	Male	30	137	83

200 rows × 5 columns

[In 4]:
```
df.describe()
```

[Out 4]:

	CustomerID	Age	Annual Income (k$)	Spending Score (1-100)
count	200.000000	200.000000	200.000000	200.000000
mean	100.500000	38.850000	60.560000	50.200000
std	57.879185	13.969007	26.264721	25.823522
min	1.000000	18.000000	15.000000	1.000000
25%	50.750000	28.750000	41.500000	34.750000
50%	100.500000	36.000000	61.500000	50.000000
75%	150.250000	49.000000	78.000000	73.000000
max	200.000000	70.000000	137.000000	99.000000

[In 5]:
```
df.dtypes
```

[Out 5]:
```
CustomerID                   int64
Gender                      object
Age                          int64
Annual Income (k$)           int64
Spending Score (1-100)       int64
dtype: object
```

[In 6]:
```
df.shape
```

[Out 6]:
```
(200, 5)
```

[In 7]:
```
df.isnull().sum()
```

[Out 7]:
```
CustomerID                   0
Gender                       0
Age                          0
Annual Income (k$)           0
Spending Score (1-100)       0
dtype: int64
```

[In 8]:
```
df.drop(["CustomerID","Gender"], axis = 1, inplace = True)
df
```

[Out 8]:

	Age	Annual Income (k$)	Spending Score (1-100)
0	19	15	39
1	21	15	81
2	20	16	6
3	23	16	77
4	31	17	40
...	...	...	...
195	35	120	79
196	45	126	28
197	32	126	74
198	32	137	18
199	30	137	83

200 rows × 3 columns

[In 9]:
```
# Data scale from 0 - 1

scaler = MinMaxScaler()
scaler.fit(df["Annual Income (k$)"])
df["Annual Income (k$)"] = scaler.transform(df["Annual Income (k$)"])

scaler.fit(df["Age"])
df["Age"] = scaler.transform(df["Age"])

scaler.fit(df["Spending Score (1-100)"])
df["Spending Score (1-100)"] = scaler.transform(df["Spending Score (1-100)"])

df
```

[Out 9]:

	Age	Annual Income (k$)	Spending Score (1-100)
0	0.019231	0.000000	0.387755
1	0.057692	0.000000	0.816327
2	0.038462	0.008197	0.051020
3	0.096154	0.008197	0.775510
4	0.250000	0.016393	0.397959
...	...	...	...
195	0.326923	0.860656	0.795918
196	0.519231	0.909836	0.275510
197	0.269231	0.909836	0.744898
198	0.269231	1.000000	0.173469
199	0.230769	1.000000	0.836735

200 rows × 3 columns

[In 10]:
```
#KMeans clustor model

km = KMeans(n_clusters=4)
y_predicted = km.fit_predict(df)
y_predicted
```

[Out 10]:
```
array([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       
```