

```
In [1]: import pandas as pd
import numpy as np
```

```
In [2]: # Reading dataset
ds=pd.read_csv(r"wholesale customers data.csv")
ds.head()
```

```
Out[2]:
```

	Channel	Region	Fresh	Milk	Grocery	Frozen	Detergents_Paper	Delicassen
0	2	3	12669	9656	7561	214	2674	1338
1	2	3	7057	9810	9568	1762	3293	1776
2	2	3	6353	8808	7684	2405	3516	7844
3	1	3	13265	1196	4221	6404	507	1788
4	2	3	22615	5410	7198	3915	1777	5185

```
In [3]: # Checking null values
ds.isnull().sum()
```

```
Out[3]: Channel      0
Region      0
Fresh      0
Milk      0
Grocery      0
Frozen      0
Detergents_Paper  0
Delicassen    0
dtype: int64
```

```
In [4]: x=ds.iloc[:,2:8].values
```

```
In [5]: from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
x = sc.fit_transform(x)
```

KMeans Clustering

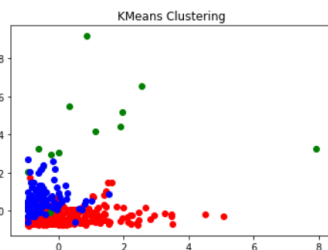
```
In [6]: from sklearn.cluster import KMeans
kmeans=KMeans(n_clusters=3)
kmeans.fit(x)
```

```
Out[6]: KMeans(algorithm='auto', copy_x=True, init='k-means++', max_iter=300,
n_clusters=3, n_init=10, n_jobs=None, precompute_distances='auto',
random_state=None, tol=0.0001, verbose=0)
```

```
In [7]: y_pred = kmeans.predict(x)
```

```
In [8]: import matplotlib.pyplot as plt

plt.title('KMeans Clustering')
plt.scatter(x[y_pred == 0,0],x[y_pred == 0,1],color = 'red')
plt.scatter(x[y_pred == 1,0],x[y_pred == 1,1],color = 'green')
plt.scatter(x[y_pred == 2,0],x[y_pred == 2,1],color = 'blue')
plt.show()
```



Agglomerative Clustering

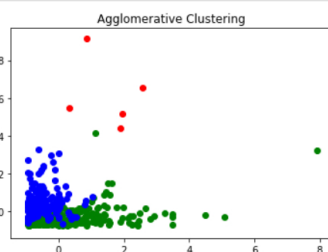
```
In [9]: from sklearn.cluster import AgglomerativeClustering
agc = AgglomerativeClustering(n_clusters=3,affinity = 'euclidean',linkage = 'ward')
```

```
In [10]: y_agc_pred = agc.fit_predict(x)
```

```
In [12]: plt.title('Agglomerative Clustering')

plt.scatter(x[y_agc_pred == 0,0],x[y_agc_pred ==0,1],color= 'red')
plt.scatter(x[y_agc_pred == 1,0],x[y_agc_pred ==1,1],color= 'green')
plt.scatter(x[y_agc_pred == 2,0],x[y_agc_pred ==2,1],color= 'blue')

plt.show()
```



Done

