```
# DB-scan clustering
```

In [9]:

```
from sklearn.cluster import DBSCAN
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
```

In [4]:

```
dataset = pd.read_csv('/Users/myyntiimac/Desktop/Mall_Customers.csv')
X = dataset.iloc[:, [3, 4]].values
```

In [26]:

```
from sklearn.preprocessing import StandardScaler
X = StandardScaler().fit_transform(X)
```

In [27]:

```
# Perform DBSCAN clustering
dbscan = DBSCAN(eps=0.3, min_samples=5)
labels = dbscan.fit_predict(X)
```

In [28]:

```
# Get the number of unique clusters excluding noise points
unique_labels = np.unique(labels)
unique_labels = unique_labels[unique_labels != -1]
num_clusters = len(unique_labels)
num_clusters
```

Out[28]:

```
7
```

In [31]:

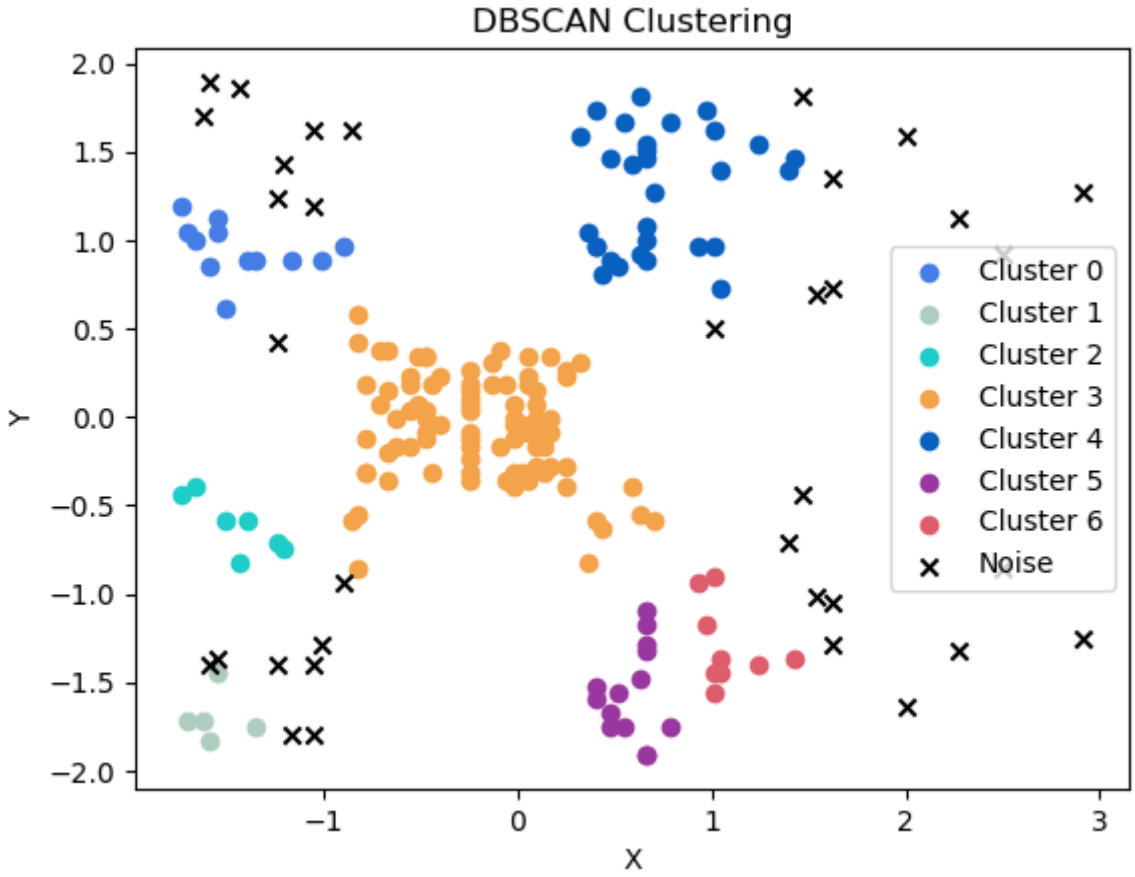
```
# Generate random colors for the clusters
colors = np.random.rand(num_clusters, 3)

# Plot the clusters with random colors
for label, color in zip(unique_labels, colors):
    cluster_data = X [labels == label]
    plt.scatter(cluster_data[:, 0], cluster_data[:, 1], c=[color], label=f'Clust

# Plot noise points as black X's
noise_data = X [labels == -1]
plt.scatter(noise_data[:, 0], noise_data[:, 1], marker='x', color='black', label

# Set plot title and labels
plt.title('DBSCAN Clustering')
plt.xlabel('X')
plt.ylabel('Y')
plt.legend()

# Show the plot
plt.show()
```



In [39]:

```
df1=pd.read_csv("/Users/myyntiimac/Desktop/new_dataframe.csv")
df1.head()
```

Out[39]:

	CustomerID	Genre	Age	Annual Income (k$)	Spending Score (1-100)	Cluster	agguluramative
0	1	Male	19	15	39	2	4
1	2	Male	21	15	81	3	3
2	3	Female	20	16	6	2	4
3	4	Female	23	16	77	3	3
4	5	Female	31	17	40	2	4

In [40]:

```
df1["DBC"] = labels
```

In [41]:

```
df1.head()
```

Out[41]:

	CustomerID	Genre	Age	Annual Income (k$)	Spending Score (1-100)	Cluster	agguluramative	DBC
0	1	Male	19	15	39	2	4	2
1	2	Male	21	15	81	3	3	0
2	3	Female	20	16	6	2	4	1
3	4	Female	23	16	77	3	3	0
4	5	Female	31	17	40	2	4	2

In [46]:

```
df1.to_csv('all_cluster.csv', index=False)
```

In [47]:

```
import os

# Get the current working directory
current_dir = os.getcwd()

# Specify the filename
filename = 'all_cluster.csv'

# Combine the directory and filename to get the full file path
file_path = os.path.join(current_dir, filename)

# Print the file path
print("File saved at:", file_path)
```

```
File saved at: /Users/myyntiimac/all_cluster.csv
```

In [ ]:

```

```