

Experiment 10 Aim: Implementation of DBSCAN clustering. Objectives: To Study DBSCAN clustering To form clusters using DBSCAN clustering algorithm Course Outcomes: CO4, CO5

```
In [29]: import matplotlib.pyplot as plt
import pandas as pd
from sklearn.cluster import DBSCAN
```

```
In [30]: df=pd.read_csv("Exp-10.csv")
```

```
In [31]: df.head()
```

```
Out[31]:
```

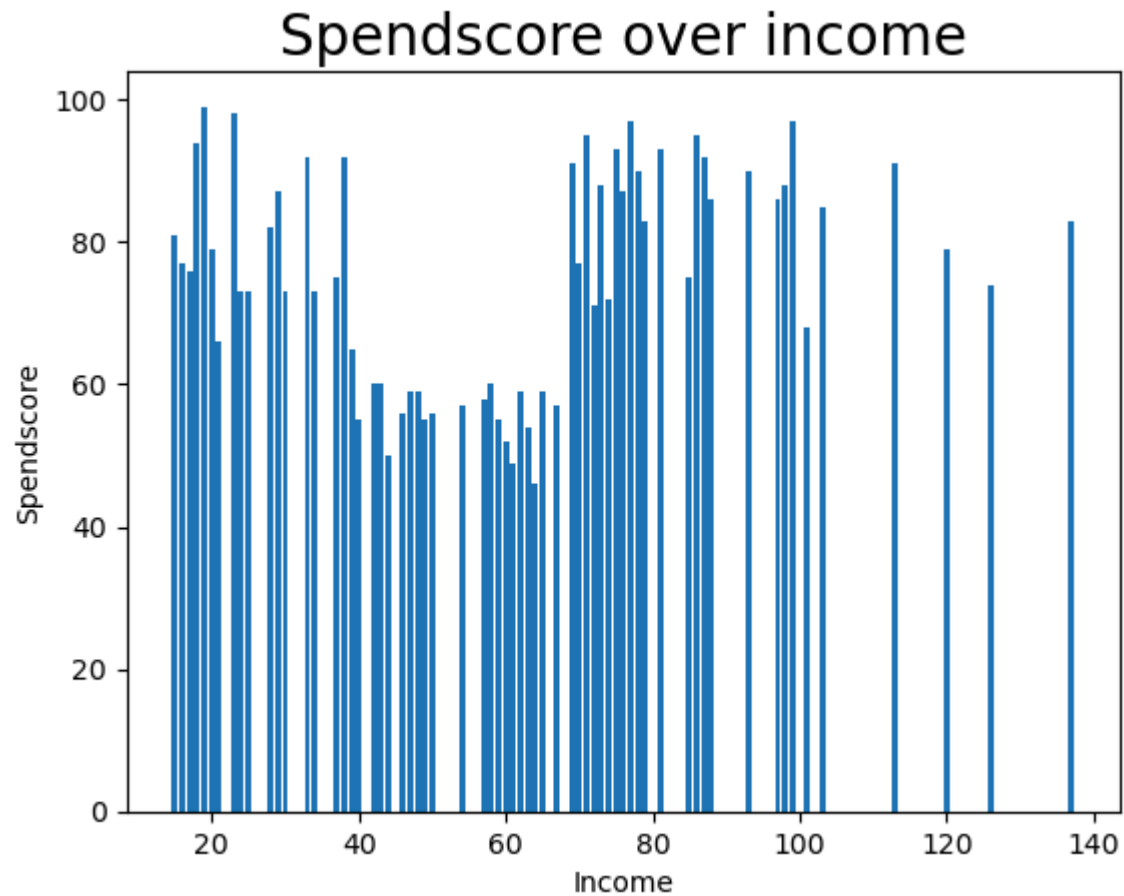
	CustomerID	Gender	Age	Annual Income (k\$)	Spending Score (1-100)
0	1	Male	19	15	39
1	2	Male	21	15	81
2	3	Female	20	16	6
3	4	Female	23	16	77
4	5	Female	31	17	40

```
In [32]: df=df.drop(['CustomerID'],axis=1)
```

```
In [33]: df.rename(columns={'Annual Income (k$)':'Income','Spending Score (1-100)':'SpendScore'},inplace=True)
```

```
In [34]: plt.bar(df['Income'],df['SpendScore'])
plt.title('Spendscore over income',fontsize=20)
plt.xlabel('Income')
plt.ylabel('Spendscore')
```

```
Out[34]: Text(0, 0.5, 'Spendscore')
```



```
In [35]: x=df.iloc[:,[2,3]].values
```

```
In [36]: x.shape
```

```
Out[36]: (200, 2)
```

```
In [37]: db=DBSCAN(eps=3,min_samples=4,metric='euclidean')
```

```
In [38]: model=db.fit(x)
```

```
In [39]: label=model.labels_
```

In [40]: label

```
Out[40]: array([-1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1,
        -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1,
        -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1,
        -1, -1, -1, -1, -1,  0,  0,  0,  0, -1, -1,  0, -1,  0, -1,  0,  0,
        -1,  0, -1, -1,  0, -1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,
         1,  1,  1, -1,  2,  1,  2,  2,  2,  2,  2,  2,  2,  2,  2,  2,
         2,  2,  2,  2,  2,  2,  2,  2,  2,  2,  2,  2,  3,  2,
         3,  3, -1,  3, -1, -1,  4, -1, -1, -1,  4,  5,  4, -1,  4,  5, -1,
         5,  4, -1,  4,  5, -1, -1,  6, -1, -1, -1,  7, -1,  6, -1,  6, -1,
         7, -1,  6, -1,  7, -1,  7, -1, -1, -1, -1, -1, -1, -1, -1, -1,
         8, -1,  8, -1,  8, -1,  8, -1, -1, -1, -1, -1, -1, -1, -1, -1,
        -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1], dtype=int64)
```

In [41]: *#identifying the points which makes up our core points*

```
sample_cores=np.zeros_like(label,dtype=bool)
```

```
sample_cores[db.core_sample_indices_]=True
```

```
#Calculating the number of clusters
```

```
n_clusters=len(set(label))- (1 if -1 in label else 0)
```

```
print('No of clusters:',n_clusters)
```

```
No of clusters: 9
```

In [42]: y\_means = db.fit\_predict(x)

```
plt.figure(figsize=(7,5))
```

```
plt.scatter(x[y_means == 0, 0], x[y_means == 0, 1], s = 50, c = 'pink')
```

```
plt.scatter(x[y_means == 1, 0], x[y_means == 1, 1], s = 50, c = 'yellow')
```

```
plt.scatter(x[y_means == 2, 0], x[y_means == 2, 1], s = 50, c = 'cyan')
```

```
plt.scatter(x[y_means == 3, 0], x[y_means == 3, 1], s = 50, c = 'magenta')
```

```
plt.scatter(x[y_means == 4, 0], x[y_means == 4, 1], s = 50, c = 'orange')
```

```
plt.scatter(x[y_means == 5, 0], x[y_means == 5, 1], s = 50, c = 'blue')
```

```
plt.scatter(x[y_means == 6, 0], x[y_means == 6, 1], s = 50, c = 'red')
```

```
plt.scatter(x[y_means == 7, 0], x[y_means == 7, 1], s = 50, c = 'black')
```

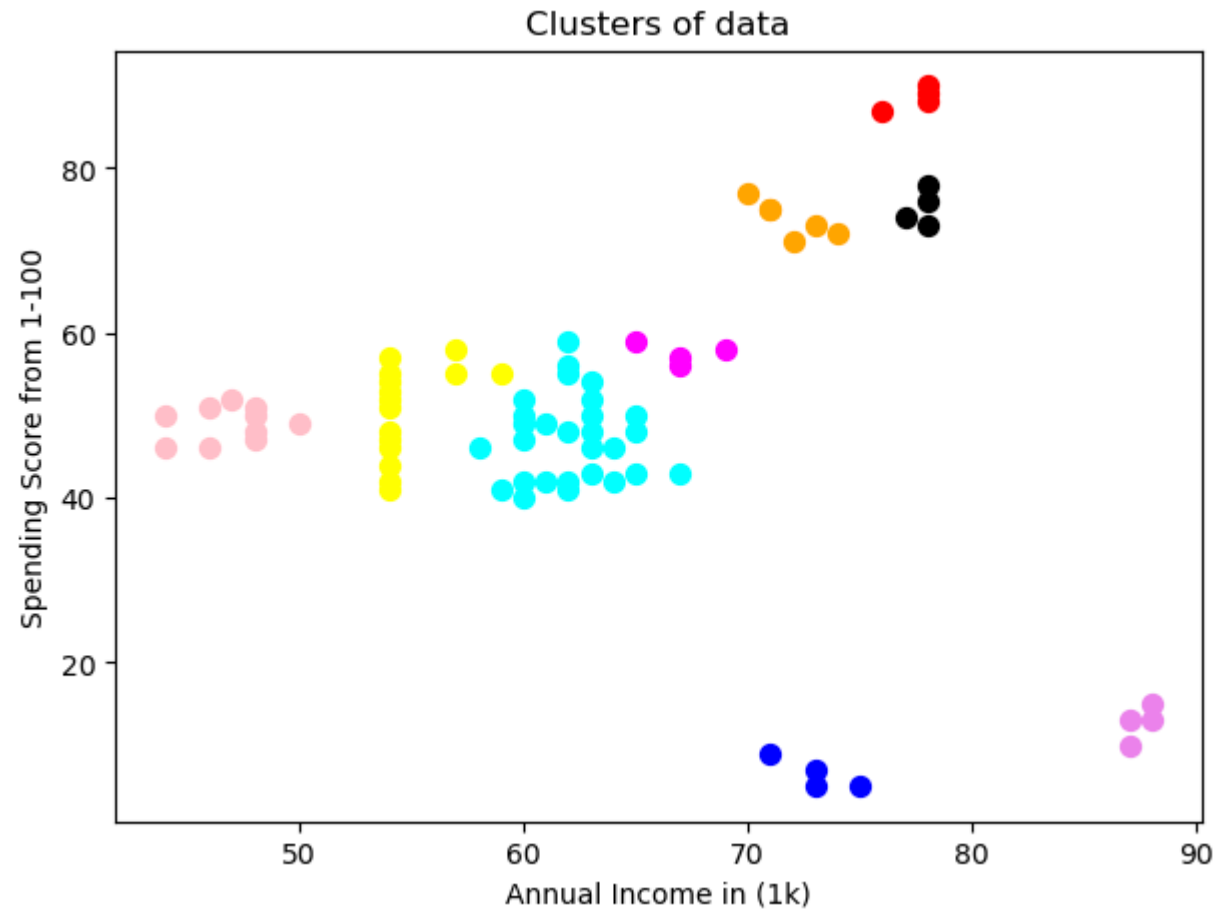
```
plt.scatter(x[y_means == 8, 0], x[y_means == 8, 1], s = 50, c = 'violet')
```

```
plt.xlabel('Annual Income in (1k)')
```

```
plt.ylabel('Spending Score from 1-100')
```

```
plt.title('Clusters of data')
```

```
plt.show()
```



Learning Outcomes: I learned about DBSCAN clustering algorithm and how to form the clusters.

Result/discussion: I have successfully implemented DBSCAN clustering algorithm and formed clusters.

In [ ]: