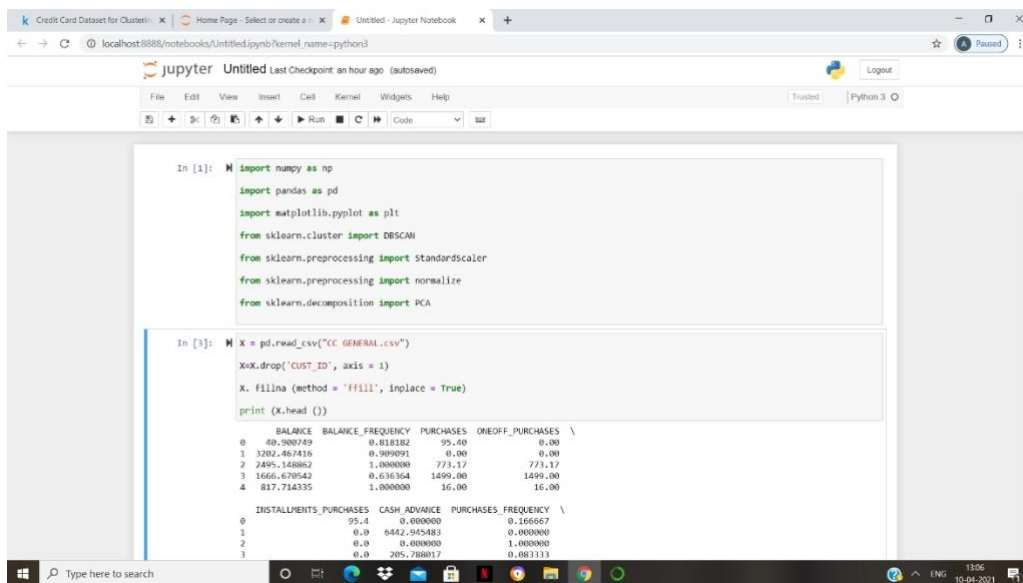


# MODEL LAB-DWDM

-R.SRIJILL 18BCS034

1)BALANCE-NUMERIC, 2)BALANCE FREQUENCY-NUMERIC ,3)PURCHASES-NUMERIC,  
4)ONEOFF\_PURCHASES-NUMERIC, 5)INSTALLMENTS\_PURCHASES-NUMERIC, 6)CASH ADVANCE-  
NUMERIC,7)PURCHASES FREQUENCY-NUMERIC



The screenshot shows a Jupyter Notebook with two code cells. The first cell imports necessary libraries: numpy, pandas, matplotlib.pyplot, DBSCAN from sklearn.cluster, StandardScaler from sklearn.preprocessing, and PCA from sklearn.decomposition. The second cell loads a CSV file named 'CC\_GENERAL.csv', drops the 'CUST\_ID' column, fills missing values with 'f', and prints the first five rows of the dataset. The output shows a table with columns: BALANCE, BALANCE\_FREQUENCY, PURCHASES, ONEOFF\_PURCHASES, INSTALLMENTS\_PURCHASES, CASH\_ADVANCE, and PURCHASES\_FREQUENCY.

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.cluster import DBSCAN
from sklearn.preprocessing import StandardScaler
from sklearn.preprocessing import normalize
from sklearn.decomposition import PCA

In [3]: X = pd.read_csv('CC_GENERAL.csv')
X=X.drop('CUST_ID', axis = 1)
X.fillna(method = 'f', inplace = True)
print (X.head ())
```

	BALANCE	BALANCE_FREQUENCY	PURCHASES	ONEOFF_PURCHASES	INSTALLMENTS_PURCHASES	CASH_ADVANCE	PURCHASES_FREQUENCY
0	40.900749	0.818182	95.60	0.00	0	95.4	0.166667
1	3202.467416	0.909091	0.00	0.00	1	0.0	0.000000
2	2495.148862	1.000000	773.17	773.17	0	0.0	0.000000
3	1666.679542	0.636364	1499.00	1499.00	0	0.0	0.000000
4	817.714335	1.000000	16.00	16.00	3	205.788017	0.666667

Credit Card Dataset for Clusterin x Home Page - Select or create a x Untitled - Jupyter Notebook x +

localhost:8888/notebooks/Untitled.ipynb?kernel\_name=python3

jupyter Untitled Last Checkpoint: an hour ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

```

4 817.714335 1.000000 16.00 16.00
INSTALLMENTS_PURCHASES CASH_ADVANCE PURCHASES_FREQUENCY \
0 95.4 0.000000 0.166667
1 0.0 6442.945483 0.000000
2 0.0 0.000000 1.000000
3 0.0 205.788017 0.083333
4 0.0 0.000000 0.083333
ONEOFF_PURCHASES_FREQUENCY PURCHASES_INSTALLMENTS_FREQUENCY \
0 0.000000 0.083333
1 0.000000 0.000000
2 1.000000 0.000000
3 0.083333 0.000000
4 0.083333 0.000000
CASH_ADVANCE_FREQUENCY CASH_ADVANCE_TRX PURCHASES_TRX CREDIT_LIMIT \
0 0.000000 0 2 1000.0
1 0.250000 4 0 7000.0
2 0.000000 0 12 7500.0
3 0.083333 1 1 7500.0
4 0.000000 0 1 1200.0
PAYMENTS MINIMUM_PAYMENTS PRC_FULL_PAYMENT TENURE
0 201.802084 139.509787 0.000000 12
1 4103.032597 1072.340217 0.222222 12
2 622.066742 627.284787 0.000000 12
3 0.000000 627.284787 0.000000 12
4 678.334763 244.791237 0.000000 12

```

```

In [6]: scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)
X_normalized = normalize(X_scaled)
X_normalized = pd.DataFrame(X_normalized)

```

Type here to search

Credit Card Dataset for Clusterin x Home Page - Select or create a x Untitled - Jupyter Notebook x +

localhost:8888/notebooks/Untitled.ipynb?kernel\_name=python3

jupyter Untitled Last Checkpoint: an hour ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

```

In [12]: pca = PCA(n_components = 2)
X_principal = pca.fit_transform(X_normalized)
X_principal = pd.DataFrame(X_principal)
X_principal.columns = ['p1', 'p2']
print(X_principal.head())

```

```

      p1      p2
0 -0.489949 -0.679975
1 -0.519098  0.544832
2  0.330633  0.268877
3 -0.481656 -0.097606
4 -0.563512 -0.482506

```

```

In [16]: db_default = DBSCAN(eps = 0.0975, min_samples = 3).fit(X_principal)
labels = db_default.labels_

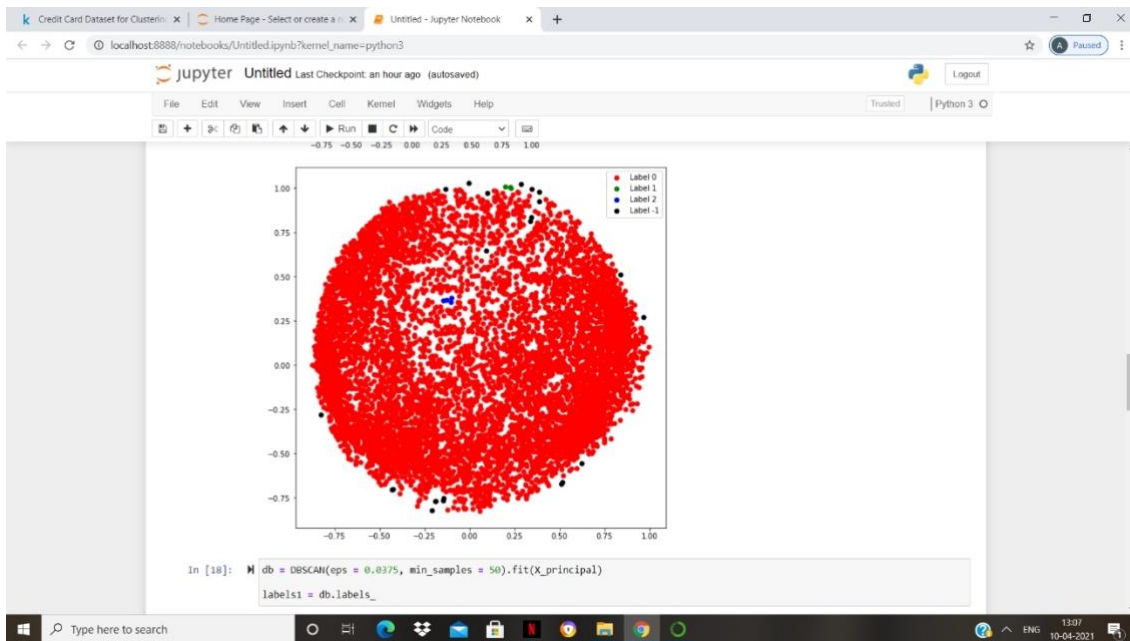
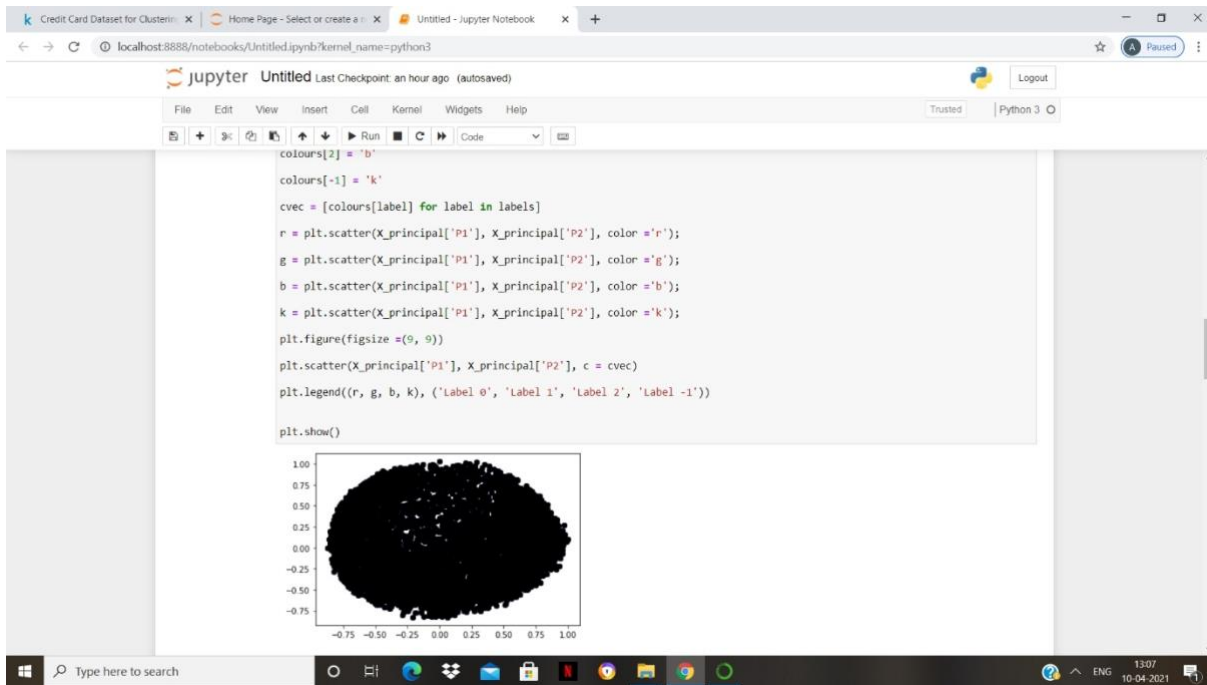
```

```

In [17]: colours = {}
colours[0] = 'r'
colours[1] = 'g'
colours[2] = 'b'
colours[-1] = 'k'
cvec = [colours[label] for label in labels]
r = plt.scatter(X_principal['p1'], X_principal['p2'], color = 'r');
g = plt.scatter(X_principal['p1'], X_principal['p2'], color = 'g');

```

Type here to search



Credit Card Dataset for Cluster... X Home Page - Select or create a... X Untitled - Jupyter Notebook X +

localhost8888/notebooks/Untitled.ipynb?kernel\_name=python3

jupyter Untitled Last Checkpoint: an hour ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

In [20]:

```
colours1 = {}  
colours1[0] = 'r'  
colours1[1] = 'g'  
colours1[2] = 'b'  
colours1[3] = 'c'  
colours1[4] = 'y'  
colours1[5] = 'm'  
colours1[-1] = 'k'  
  
cvec = [colours1[label] for label in labels]  
colors = ['r', 'g', 'b', 'c', 'y', 'm', 'k']  
  
r = plt.scatter(  
    X_principal['P1'], X_principal['P2'], marker = 'o', color = colors[0])  
g = plt.scatter(  
    X_principal['P1'], X_principal['P2'], marker = 'o', color = colors[1])  
b = plt.scatter(  
    X_principal['P1'], X_principal['P2'], marker = 'o', color = colors[2])  
c = plt.scatter(  

```

Type here to search

1507 10-04-2021

Jupyter Notebook interface showing a code cell with the following Python code:

```

c = plt.scatter(
    X_principal['P1'], X_principal['P2'], marker='o', color = colors[3])
y = plt.scatter(
    X_principal['P1'], X_principal['P2'], marker='o', color = colors[4])
m = plt.scatter(
    X_principal['P1'], X_principal['P2'], marker='o', color = colors[5])
k = plt.scatter(
    X_principal['P1'], X_principal['P2'], marker='o', color = colors[6])

plt.figure(figsize=(9, 9))
plt.scatter(X_principal['P1'], X_principal['P2'], c = cvec)
plt.legend((r, g, b, c, y, m, k),
           ('Label 0', 'Label 1', 'Label 2', 'Label 3', 'Label 4',
            'Label 5', 'Label -1'),
           scatterpoints = 1,
           loc = 'upper left',
           ncol = 3,
           fontsize = 8)
plt.show()

```

The notebook is titled "Untitled" and shows the last checkpoint as "an hour ago (autosaved)". The interface includes a menu bar (File, Edit, View, Insert, Cell, Kernel, Widgets, Help) and a toolbar with icons for file operations, running, and saving. The status bar at the bottom indicates the kernel is "Python 3".

