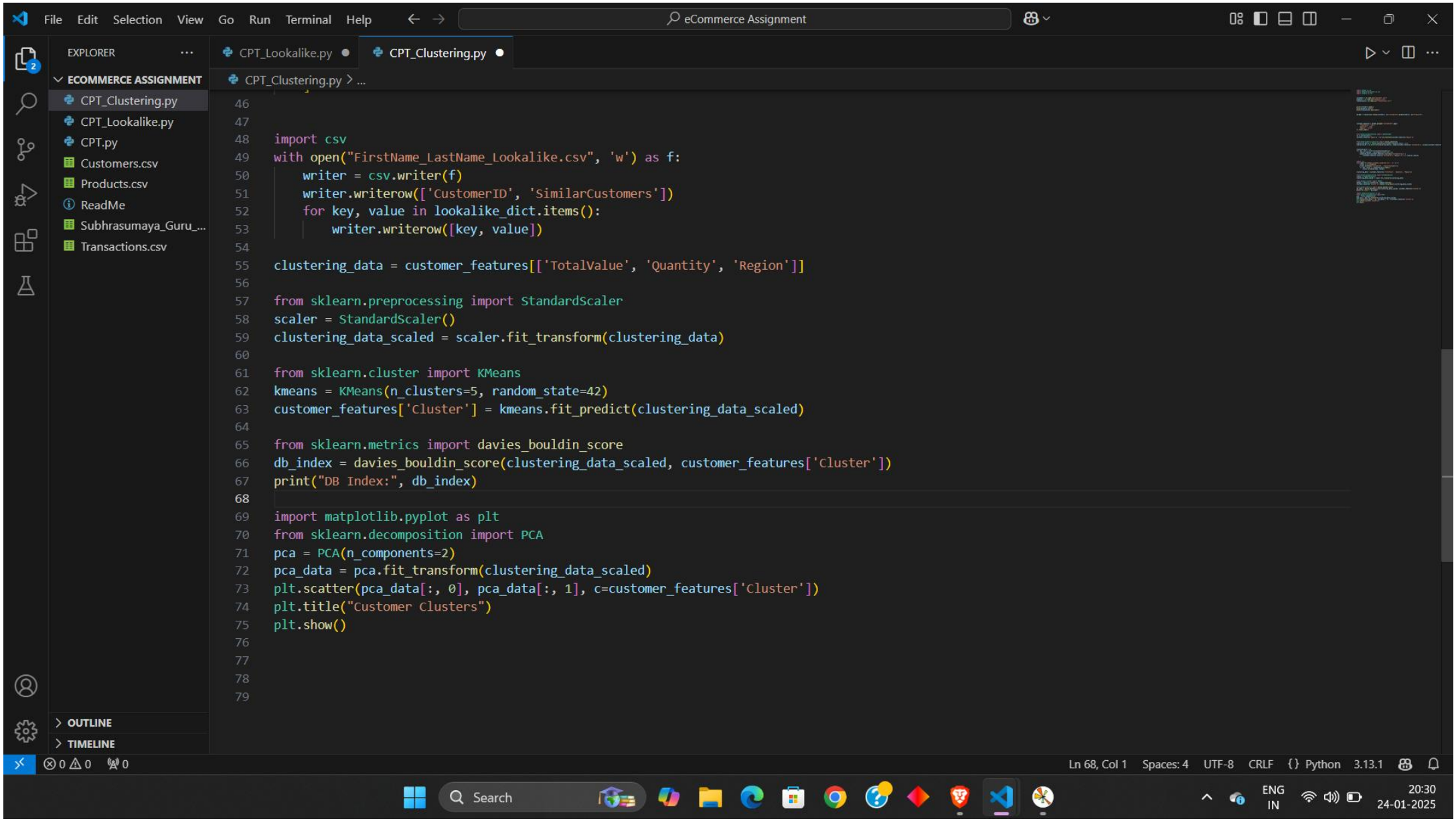
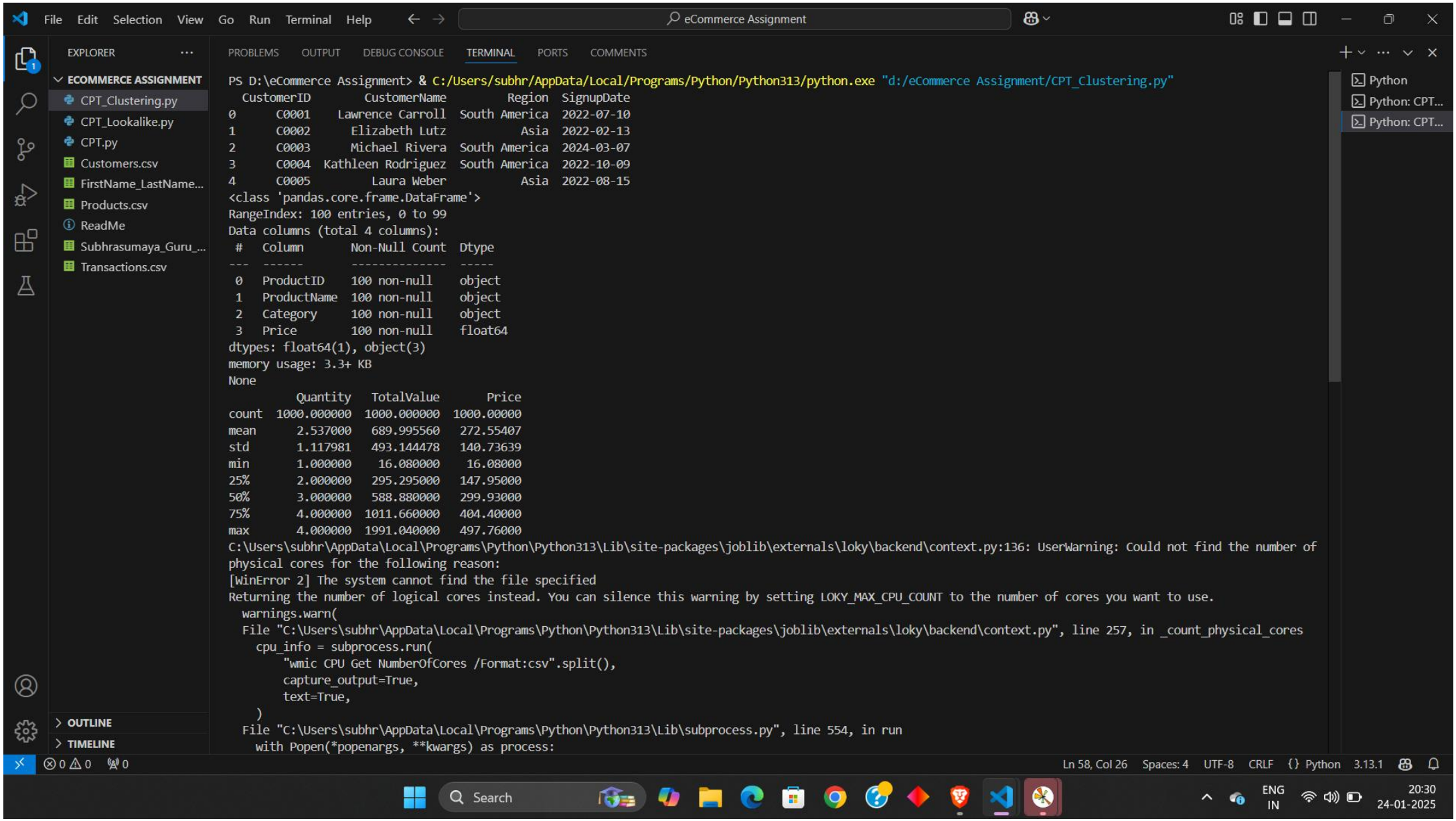
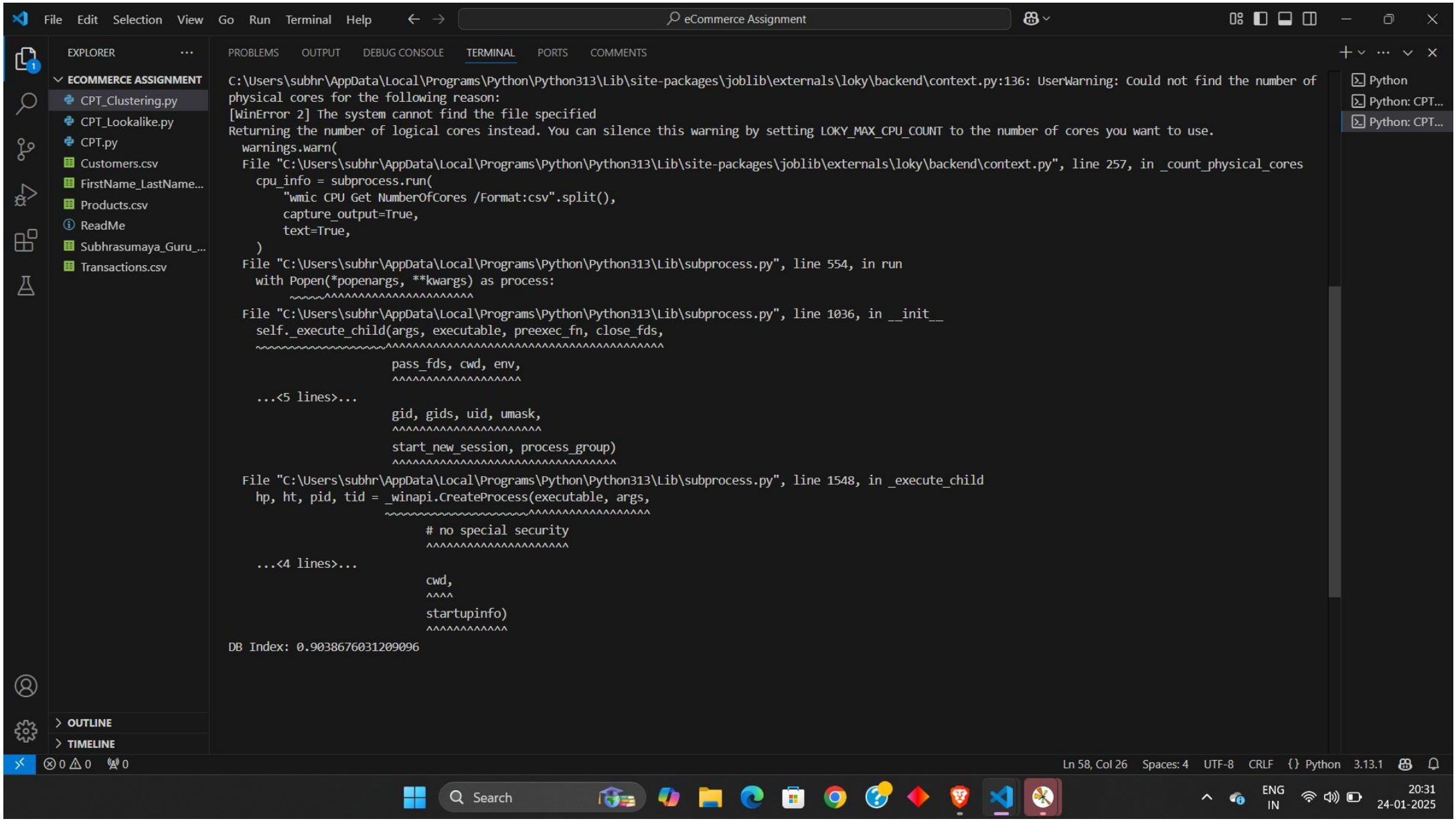

File Edit Selection View Go Run Terminal Help
eCommerce Assignment
EXPLORER
ECOMMERCE ASSIGNMENT
CPT_Clustering.py
CPT_Lookalike.py
CPT.py
Customers.csv
Products.csv
ReadMe
Subhrasumaya_Guru_...
Transactions.csv
OUTLINE
TIMELINE
CPT_Lookalike.py
CPT_Clustering.py
CPT_Clustering.py > ...
46
47
48 import csv
49 with open("FirstName_LastName_Lookalike.csv", 'w') as f:
50 writer = csv.writer(f)
51 writer.writerow(['CustomerID', 'SimilarCustomers'])
52 for key, value in lookalike_dict.items():
53 writer.writerow([key, value])
54
55 clustering_data = customer_features[['TotalValue', 'Quantity', 'Region']]
56
57 from sklearn.preprocessing import StandardScaler
58 scaler = StandardScaler()
59 clustering_data_scaled = scaler.fit_transform(clustering_data)
60
61 from sklearn.cluster import KMeans
62 kmeans = KMeans(n_clusters=5, random_state=42)
63 customer_features['Cluster'] = kmeans.fit_predict(clustering_data_scaled)
64
65 from sklearn.metrics import davies_bouldin_score
66 db_index = davies_bouldin_score(clustering_data_scaled, customer_features['Cluster'])
67 print("DB Index:", db_index)
68
69 import matplotlib.pyplot as plt
70 from sklearn.decomposition import PCA
71 pca = PCA(n_components=2)
72 pca_data = pca.fit_transform(clustering_data_scaled)
73 plt.scatter(pca_data[:, 0], pca_data[:, 1], c=customer_features['Cluster'])
74 plt.title("Customer Clusters")
75 plt.show()
76
77
78
79
Ln 68, Col 1 Spaces: 4 UTF-8 CRLF Python 3.13.1
Search
ENG
IN
20:30
24-01-2025


File Edit Selection View Go Run Terminal Help
eCommerce Assignment
EXPLORER
ECOMMERCE ASSIGNMENT
CPT_Clustering.py
CPT_Lookalike.py
CPT.py
Customers.csv
FirstName_LastName...
Products.csv
ReadMe
Subhrasumaya_Guru_...
Transactions.csv
OUTLINE
TIMELINE
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS COMMENTS
PS D:\eCommerce Assignment> & C:/Users/subhr/AppData/Local/Programs/Python/Python313/python.exe "d:/eCommerce Assignment/CPT_Clustering.py"
CustomerID CustomerName Region SignupDate
0 C0001 Lawrence Carroll South America 2022-07-10
1 C0002 Elizabeth Lutz Asia 2022-02-13
2 C0003 Michael Rivera South America 2024-03-07
3 C0004 Kathleen Rodriguez South America 2022-10-09
4 C0005 Laura Weber Asia 2022-08-15
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100 entries, 0 to 99
Data columns (total 4 columns):
Column Non-Null Count Dtype
0 ProductID 100 non-null object
1 ProductName 100 non-null object
2 Category 100 non-null object
3 Price 100 non-null float64
dtypes: float64(1), object(3)
memory usage: 3.3+ KB
None
Quantity TotalValue Price
count 1000.000000 1000.000000 1000.00000
mean 2.537000 689.995560 272.55407
std 1.117981 493.144478 140.73639
min 1.000000 16.080000 16.08000
25% 2.000000 295.295000 147.95000
50% 3.000000 588.880000 299.93000
75% 4.000000 1011.660000 404.40000
max 4.000000 1991.040000 497.76000
C:\Users\subhr\AppData\Local\Programs\Python\Python313\Lib\site-packages\joblib\externals\loky\backend\context.py:136: UserWarning: Could not find the number of physical cores for the following reason:
[WinError 2] The system cannot find the file specified
Returning the number of logical cores instead. You can silence this warning by setting LOKY_MAX_CPU_COUNT to the number of cores you want to use.
warnings.warn(
File "C:\Users\subhr\AppData\Local\Programs\Python\Python313\Lib\site-packages\joblib\externals\loky\backend\context.py", line 257, in _count_physical_cores
cpu_info = subprocess.run(
"wmic CPU Get NumberOfCores /Format:csv".split(),
capture_output=True,
text=True,
)
File "C:\Users\subhr\AppData\Local\Programs\Python\Python313\Lib\subprocess.py", line 554, in run
with Popen(*popenargs, **kwargs) as process:
Python
Python: CPT...
Python: CPT...
Ln 58, Col 26 Spaces: 4 UTF-8 CRLF Python 3.13.1
Search
ENG IN
20:30 24-01-2025


File Edit Selection View Go Run Terminal Help
eCommerce Assignment
EXPLORER
ECOMMERCE ASSIGNMENT
CPT_Clustering.py
CPT_Lookalike.py
CPT.py
Customers.csv
FirstName_LastName...
Products.csv
ReadMe
Subhrasumaya_Guru_...
Transactions.csv
OUTLINE
TIMELINE
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS COMMENTS
C:\Users\subhr\AppData\Local\Programs\Python\Python313\Lib\site-packages\joblib\externals\loky\backend\context.py:136: UserWarning: Could not find the number of physical cores for the following reason:
[WinError 2] The system cannot find the file specified
Returning the number of logical cores instead. You can silence this warning by setting LOKY_MAX_CPU_COUNT to the number of cores you want to use.
warnings.warn(
File "C:\Users\subhr\AppData\Local\Programs\Python\Python313\Lib\site-packages\joblib\externals\loky\backend\context.py", line 257, in _count_physical_cores
cpu_info = subprocess.run(
"wmic CPU Get NumberOfCores /Format:csv".split(),
capture_output=True,
text=True,
)
File "C:\Users\subhr\AppData\Local\Programs\Python\Python313\Lib\subprocess.py", line 554, in run
with Popen(*popenargs, **kwargs) as process:
File "C:\Users\subhr\AppData\Local\Programs\Python\Python313\Lib\subprocess.py", line 1036, in __init__
self._execute_child(args, executable, preexec_fn, close_fds,
pass_fds, cwd, env,
...<5 lines>...
gid, gids, uid, umask,
start_new_session, process_group)
File "C:\Users\subhr\AppData\Local\Programs\Python\Python313\Lib\subprocess.py", line 1548, in _execute_child
hp, ht, pid, tid = _winapi.CreateProcess(executable, args,
no special security
...<4 lines>...
cwd,
startupinfo)
DB Index: 0.9038676031209096
Python
Python: CPT...
Python: CPT...
Ln 58, Col 26 Spaces: 4 UTF-8 CRLF Python 3.13.1
Search
ENG
IN
20:31
24-01-2025

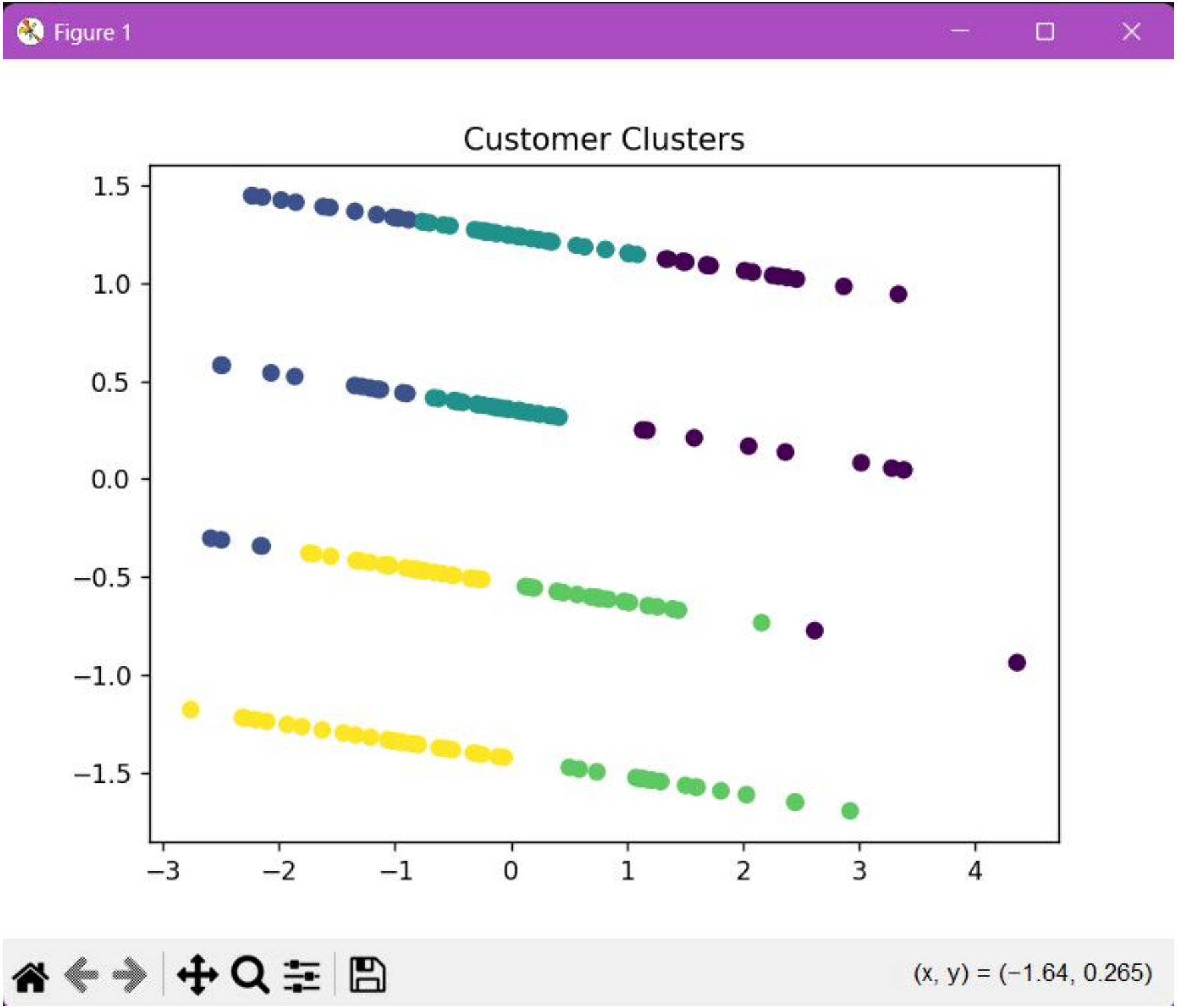


Figure 1
Customer Clusters
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−3
−2
−1
0
1
2
3
4
(x, y) = (−1.64, 0.265)