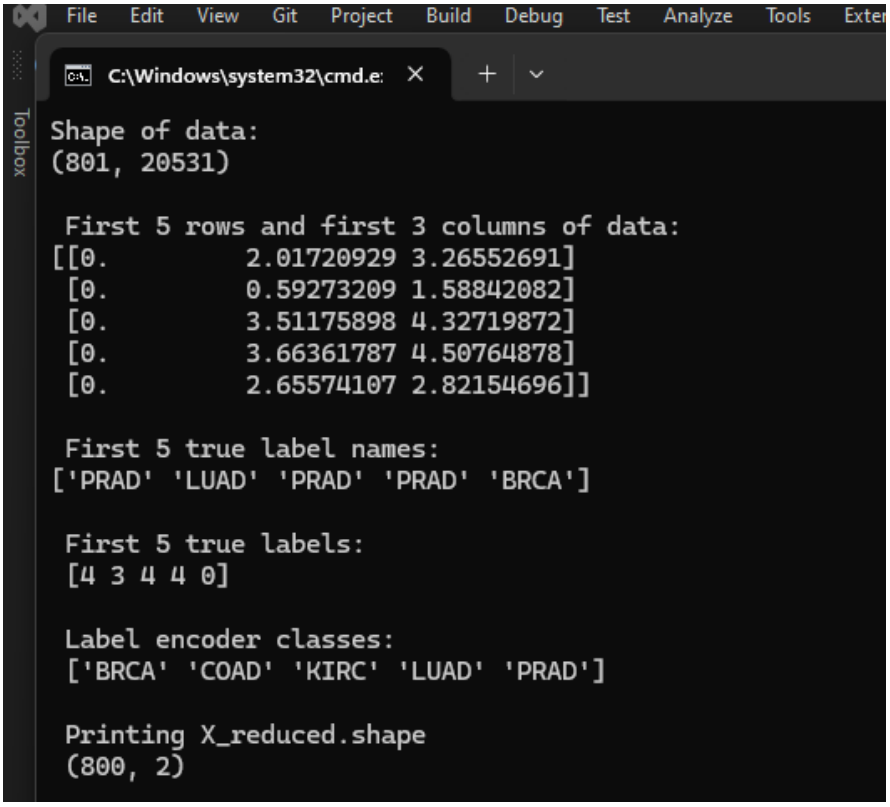
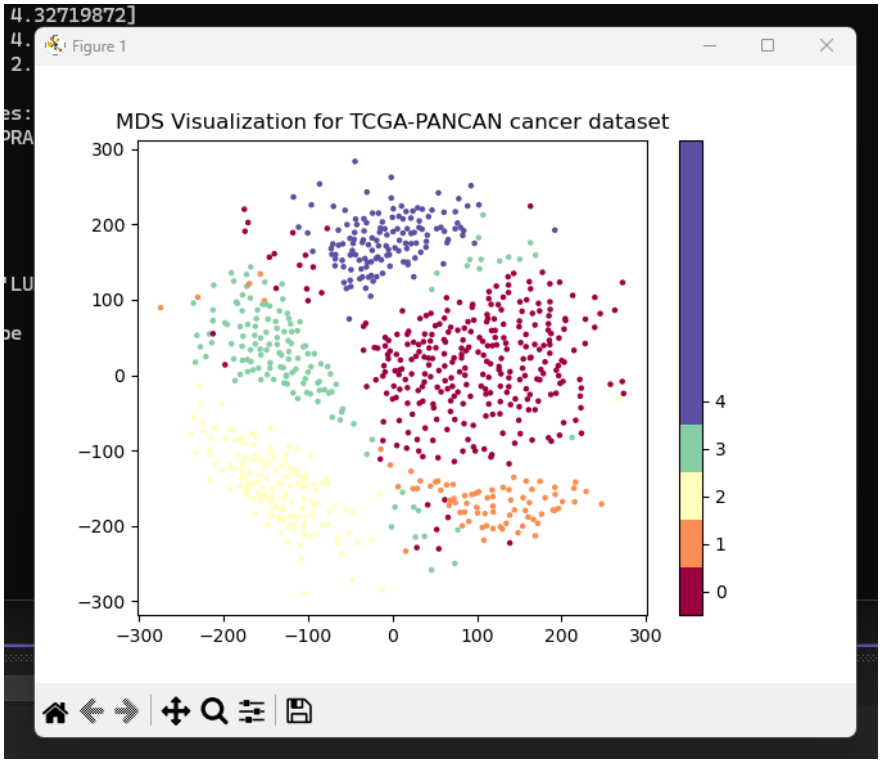
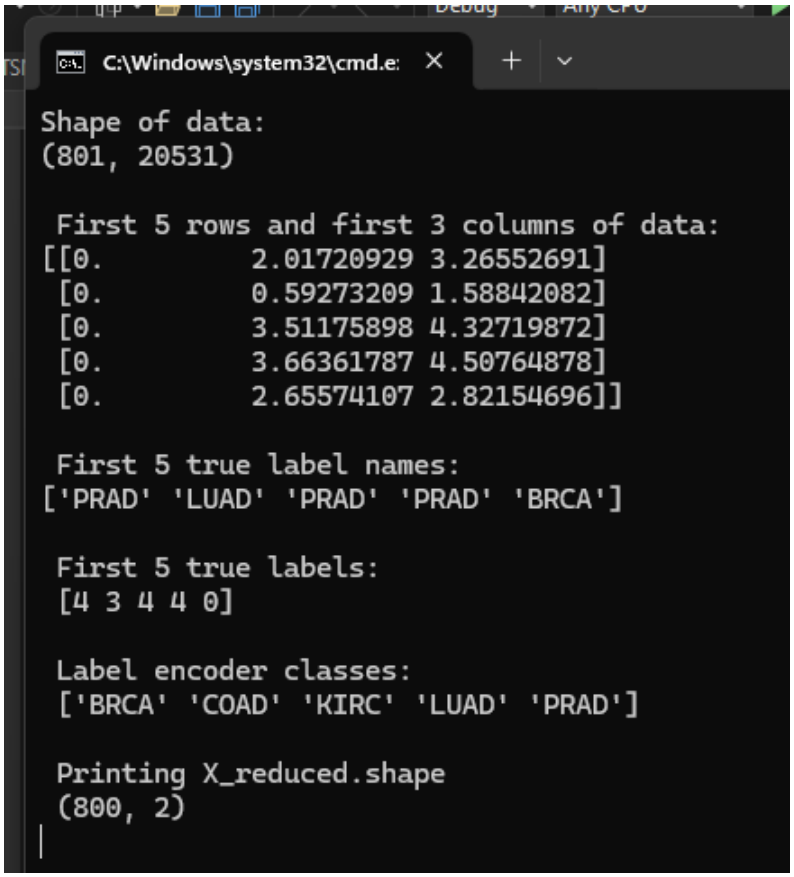
ASSIGNMENT 08

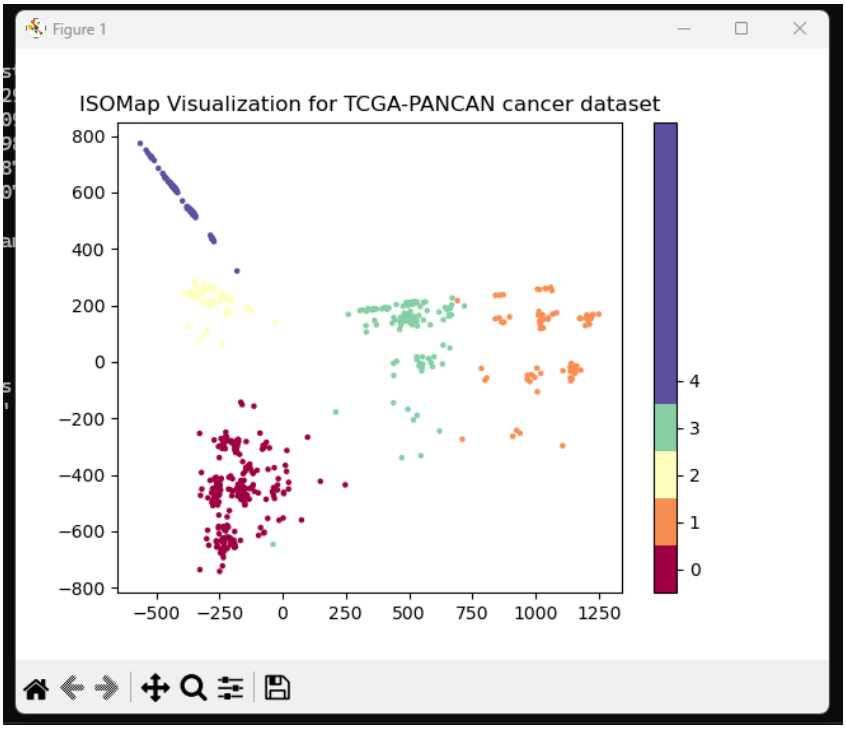
MDS OUTPUT





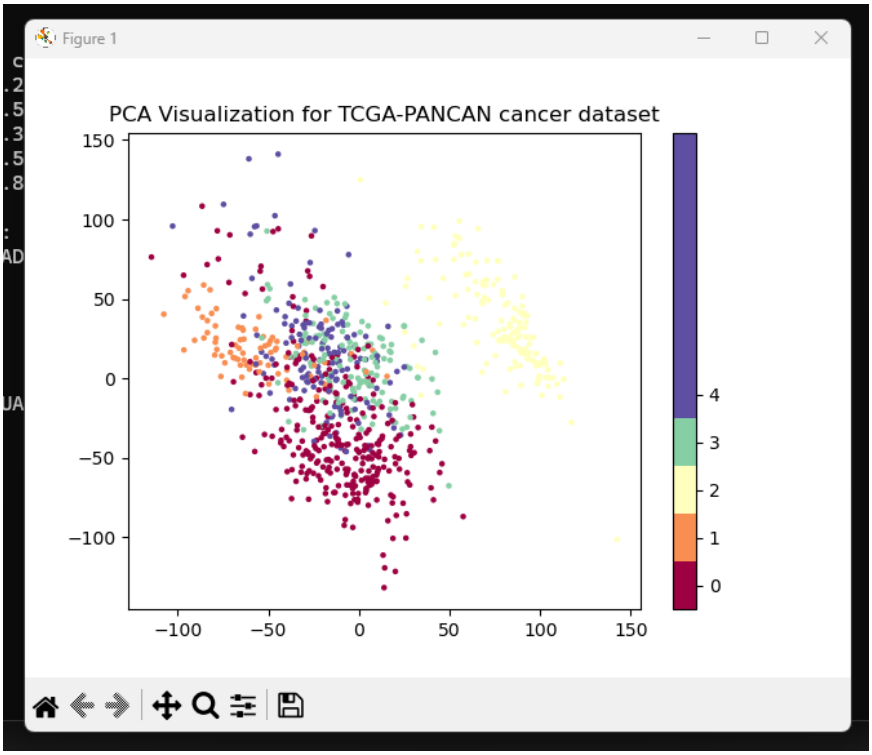
ISOMAP OUTPUT





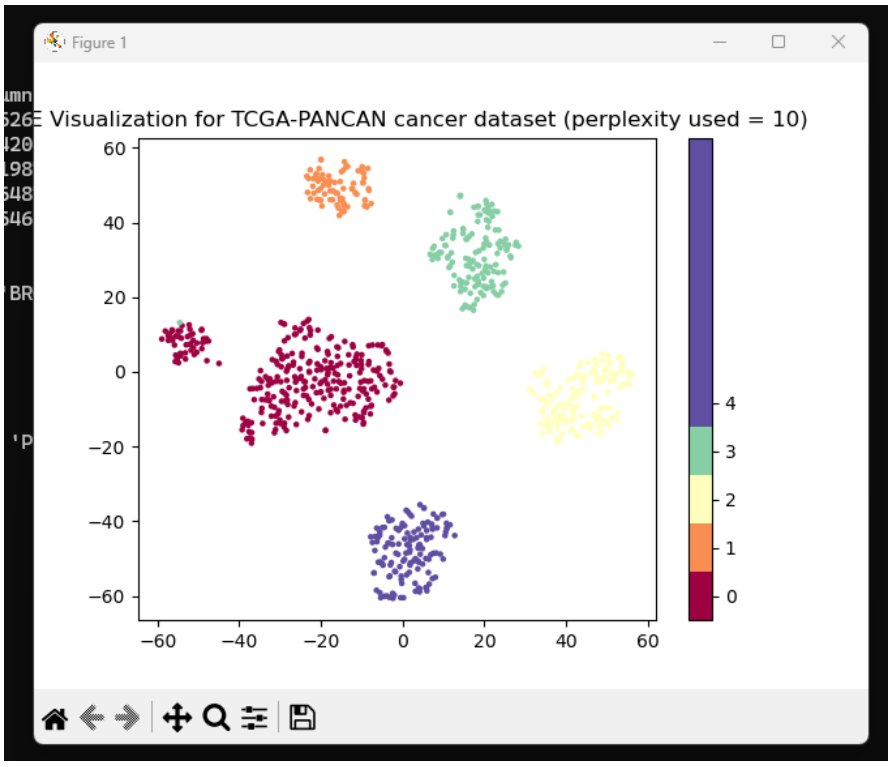
PCA OUTPUT

```
C:\Windows\system32\cmd.e:
Shape of data:
(801, 20531)

 First 5 rows and first 3 columns of data:
[[0.         2.01720929 3.26552691]
 [0.         0.59273209 1.58842082]
 [0.         3.51175898 4.32719872]
 [0.         3.66361787 4.50764878]
 [0.         2.65574107 2.82154696]]

 First 5 true label names:
['PRAD' 'LUAD' 'PRAD' 'PRAD' 'BRCA']

 First 5 true labels:
 [4 3 4 4 0]

 Label encoder classes:
 ['BRCA' 'COAD' 'KIRC' 'LUAD' 'PRAD']

 Printing pca_result.shape
 (800, 2)
Press any key to continue . . .
```



TSNE OUTPUT

```
Shape of data:
(801, 20531)

 First 5 rows and first 3 columns of data:
[[0.         2.01720929 3.26552691]
 [0.         0.59273209 1.58842082]
 [0.         3.51175898 4.32719872]
 [0.         3.66361787 4.50764878]
 [0.         2.65574107 2.82154696]]

 First 5 true label names:
['PRAD' 'LUAD' 'PRAD' 'PRAD' 'BRCA']

 First 5 true labels:
 [4 3 4 4 0]

 Label encoder classes:
 ['BRCA' 'COAD' 'KIRC' 'LUAD' 'PRAD']

 Printing tsne_result.shapee
 (800, 2)
```



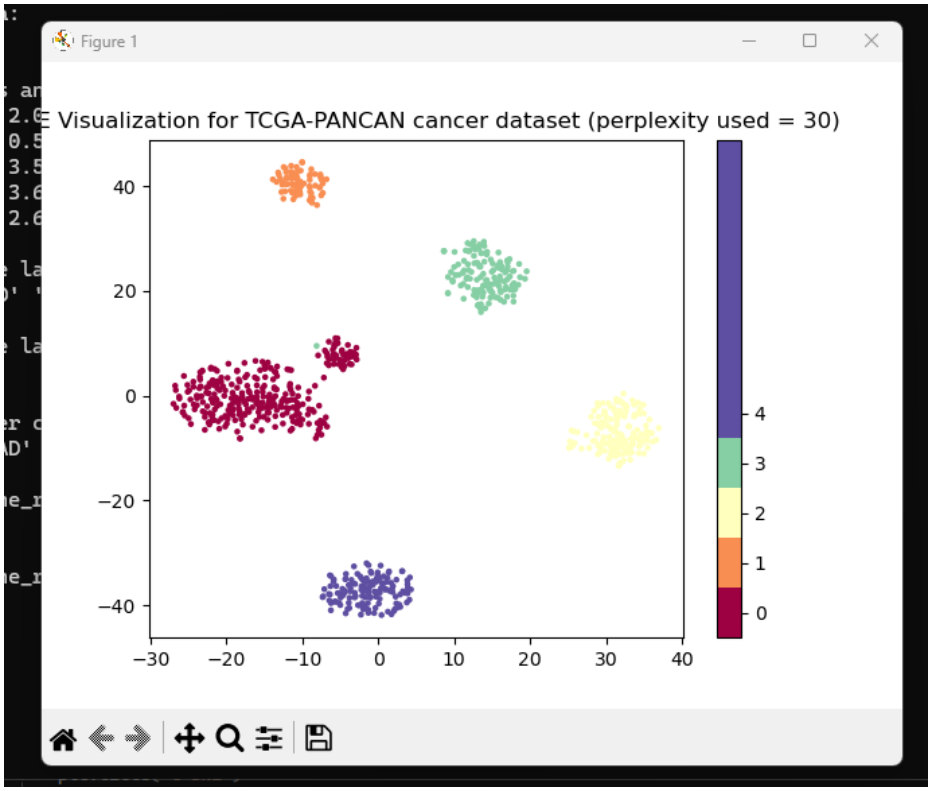


Figure 1
Visualization for TCGA-PANCAN cancer dataset (perplexity used = 30)
40
20
0
−20
−40
−30
−20
−10
0
10
20
30
40
4
3
2
1
0

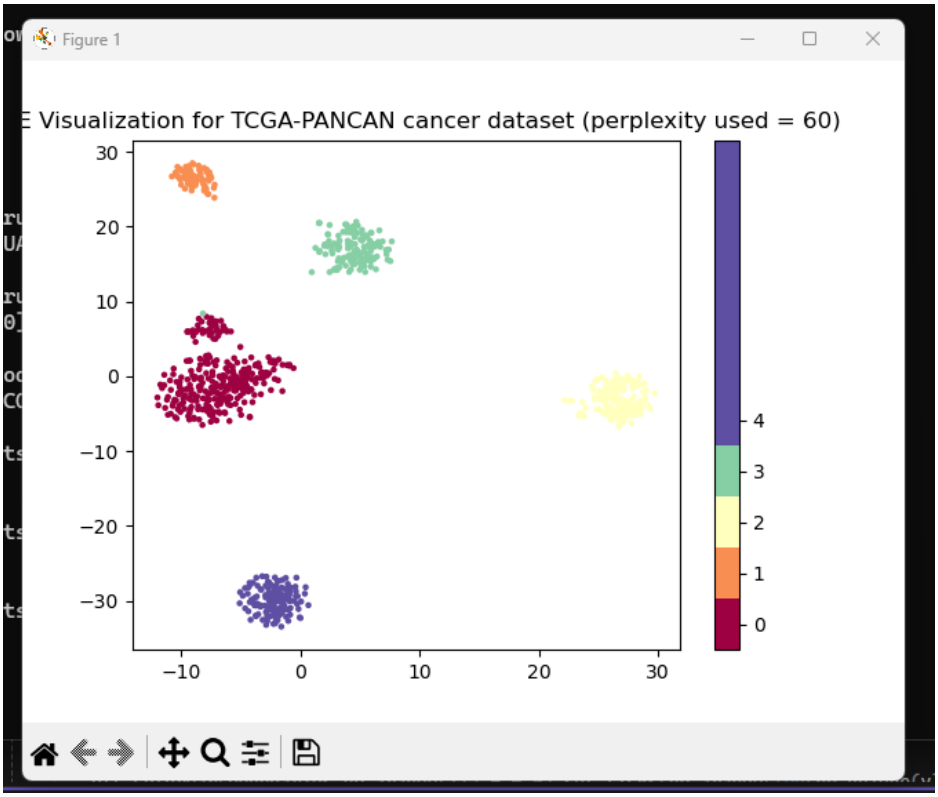


Figure 1
Visualization for TCGA-PANCAN cancer dataset (perplexity used = 60)
30
20
10
0
−10
−20
−30
−10
0
10
20
30
4
3
2
1
0

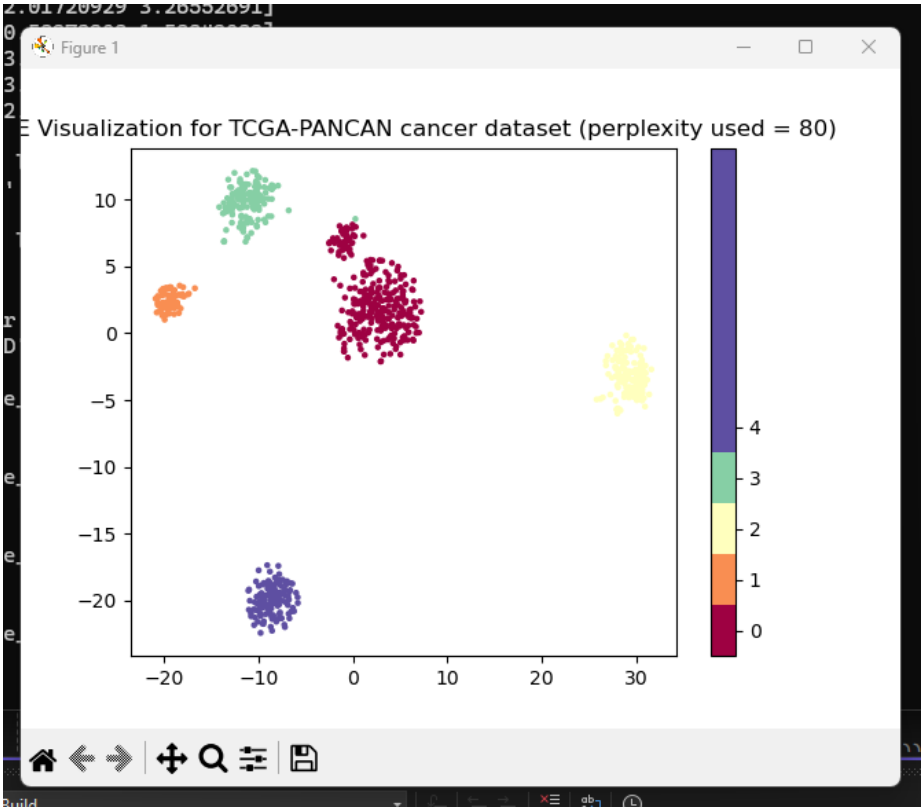


Figure 1
Visualization for TCGA-PANCAN cancer dataset (perplexity used = 80)
10
5
0
−5
−10
−15
−20
−20
−10
0
10
20
30
4
3
2
1
0

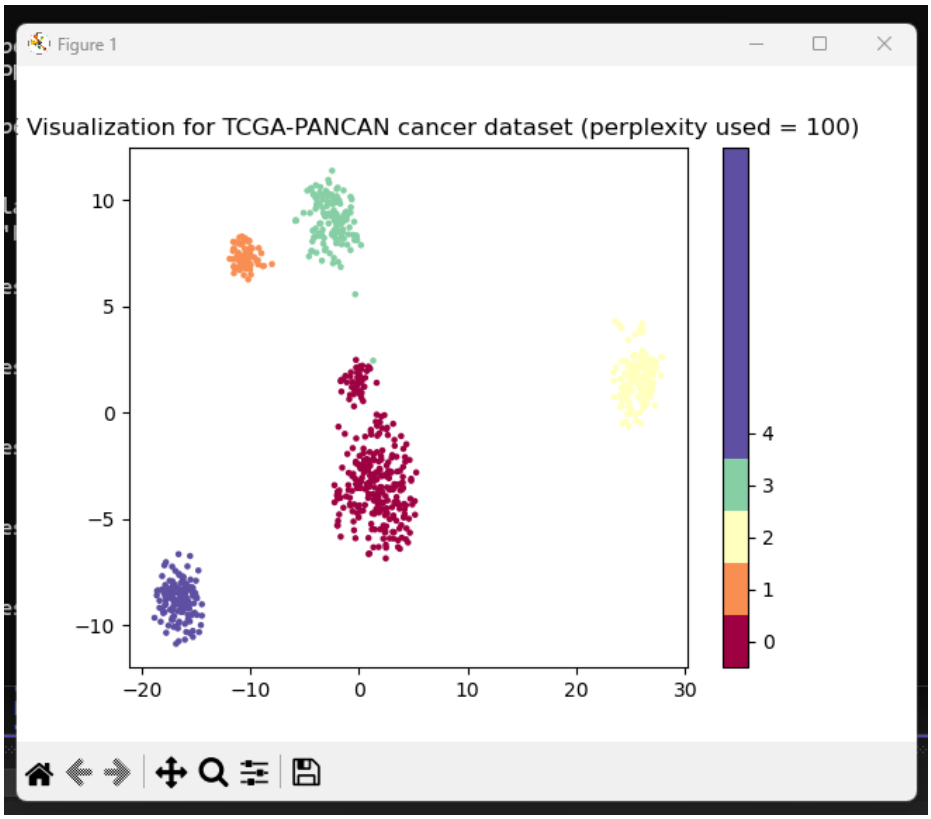
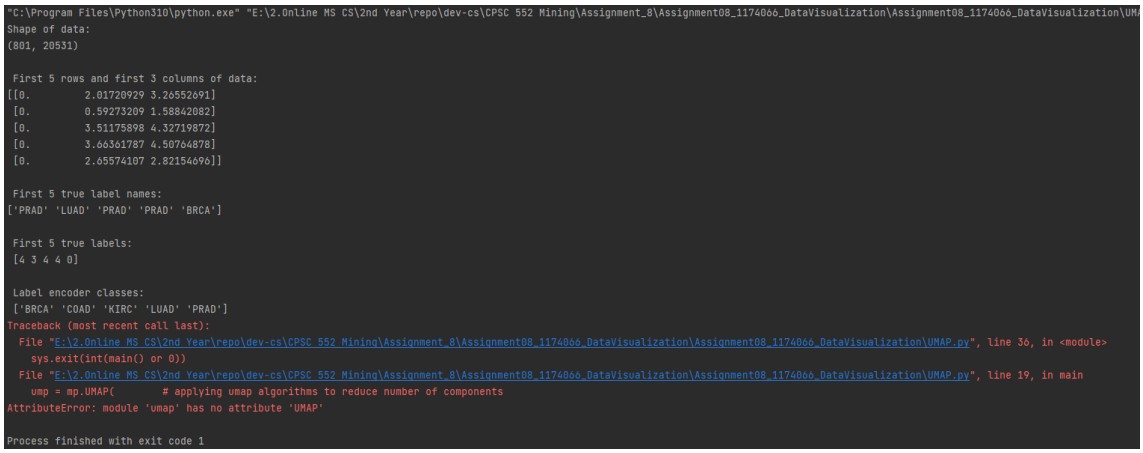


Figure 1
Visualization for TCGA-PANCAN cancer dataset (perplexity used = 100)
10
5
0
−5
−10
−20
−10
0
10
20
30
4
3
2
1
0

UMAP OUTPUT



```
"C:\Program Files\Python310\python.exe" "E:\2.Online MS CS\2nd Year\repo\dev-cs\CPSC 552 Mining\Assignment_8\Assignment08_1174066_DataVisualization\Assignment08_1174066_DataVisualization\UMA
Shape of data:
(801, 20531)

 First 5 rows and first 3 columns of data:
[[0.         2.01720929 3.26552691]
 [0.         0.59273209 1.58842082]
 [0.         3.51175898 4.32719872]
 [0.         3.66361787 4.50764878]
 [0.         2.65574107 2.82154696]]

 First 5 true label names:
['PRAD' 'LUAD' 'PRAD' 'PRAD' 'BRCA']

 First 5 true labels:
 [4 3 4 4 0]

 Label encoder classes:
 ['BRCA' 'COAD' 'KIRC' 'LUAD' 'PRAD']
Traceback (most recent call last):
  File "E:\2.Online MS CS\2nd Year\repo\dev-cs\CPSC 552 Mining\Assignment_8\Assignment08_1174066_DataVisualization\Assignment08_1174066_DataVisualization\UMAP.py", line 36, in <module>
    sys.exit(int(main() or 0))
  File "E:\2.Online MS CS\2nd Year\repo\dev-cs\CPSC 552 Mining\Assignment_8\Assignment08_1174066_DataVisualization\Assignment08_1174066_DataVisualization\UMAP.py", line 19, in main
    ump = mp.UMAP(        # applying umap algorithms to reduce number of components
AttributeError: module 'umap' has no attribute 'UMAP'

Process finished with exit code 1
```

CONCLUSION

From the above outputs TSNE and UMAP have shown good visualization of clusters by displaying a clear separation between them compared to other techniques such as PCA, ISOMAPS and MDS.

For this specific dataset, TSNE performed very well even for different numbers of perplexities.