

A NASA Space Shuttle is shown in orbit above Earth's cloud-covered surface. The shuttle is oriented diagonally from the top left towards the bottom right. The background is the deep black of space, and a solid red rectangle is visible in the top right corner.

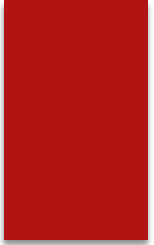
# Anomaly Detection with NASA shuttle data

ASHISH PIYA

# Objective

- ▶ Detect anomalous data among the Shuttle Dataset provided by NASA

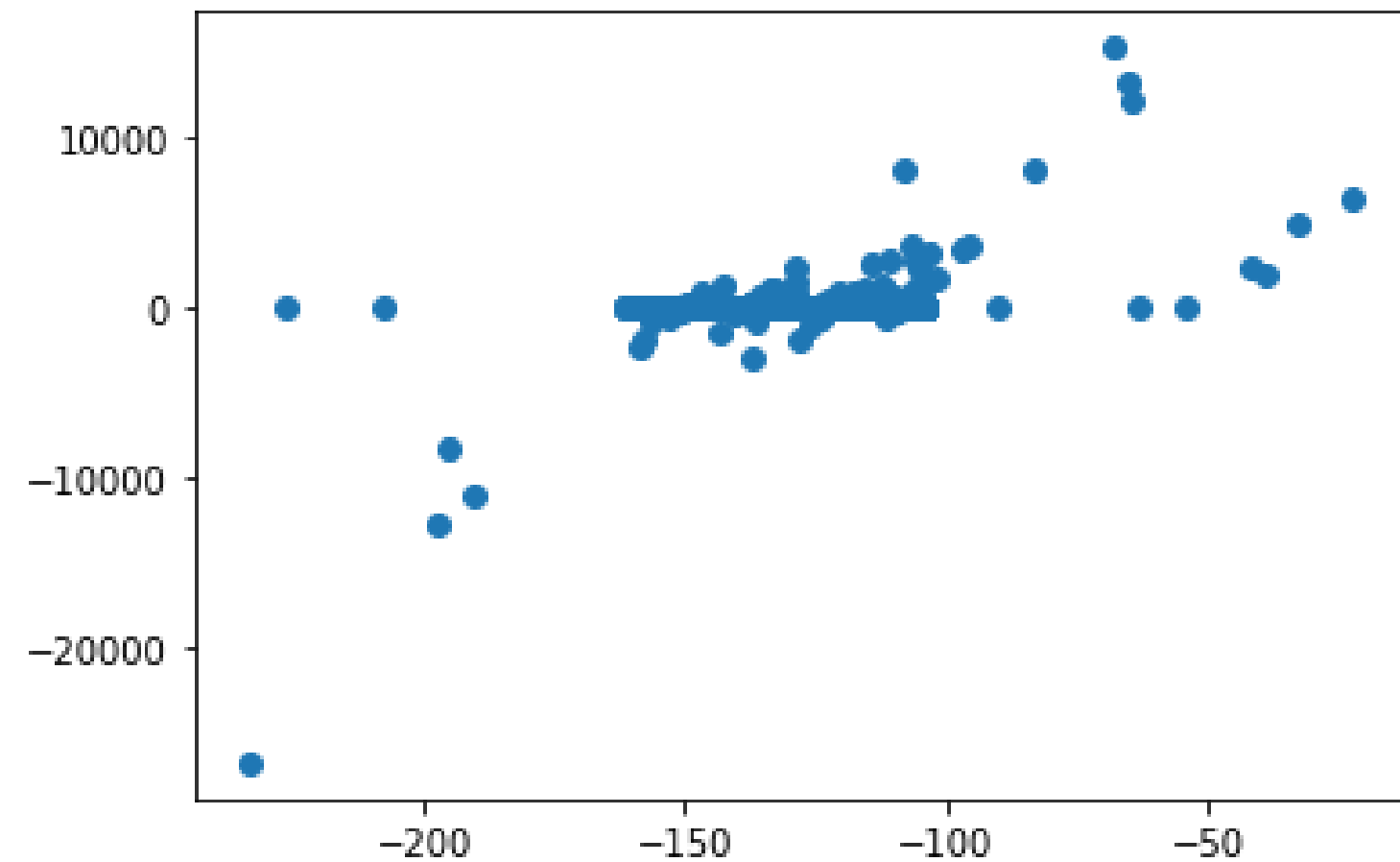
- ▶ 9 Undisclosed category
- ▶ Around 49000 rows of data



	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	y
0	50	21	77	0	28	0	27	48	22	1
1	53	0	82	0	52	-5	29	30	2	0
2	37	0	76	0	28	18	40	48	8	0
3	37	0	79	0	34	-26	43	46	2	0
4	85	0	88	-4	6	1	3	83	80	1

# Unsupervised Clustering

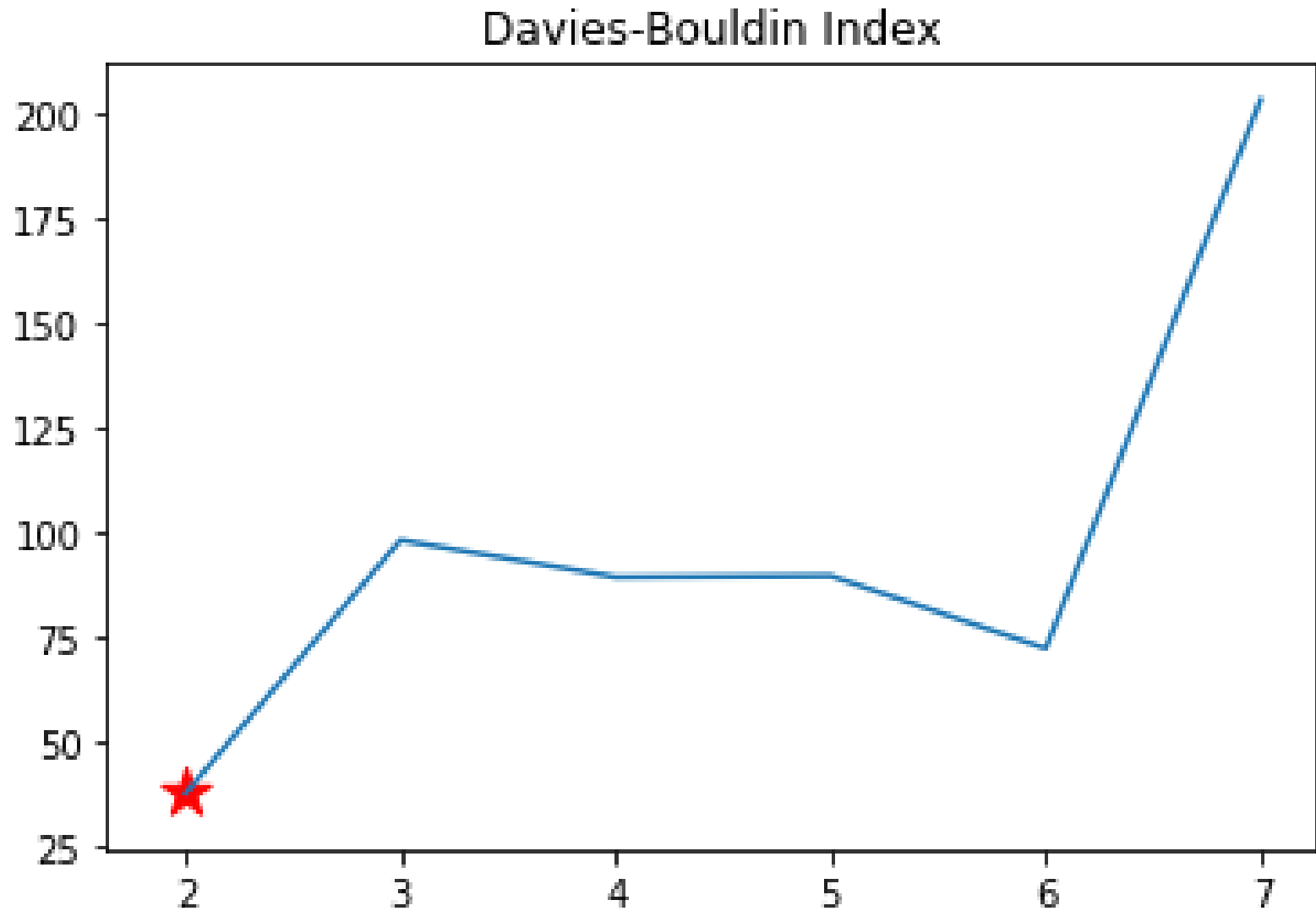
- ▶ K-means clustering
- ▶ Gaussian Mixture Model clustering
- ▶ DBSCAN

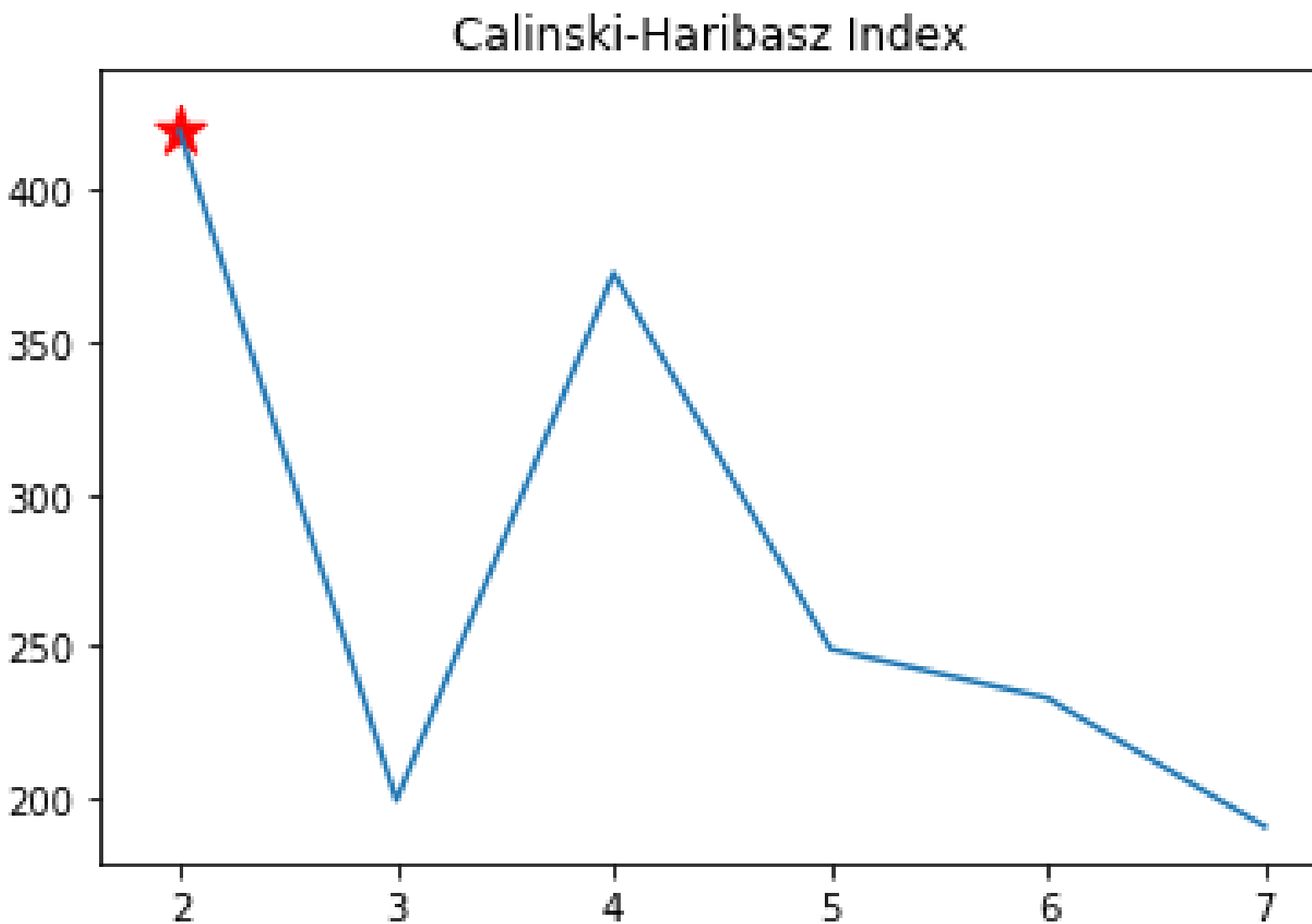


Dimensionality  
Reduction for  
Visualization

## Davies-Bouldin Index

Suggested  
number of  
Cluster 2



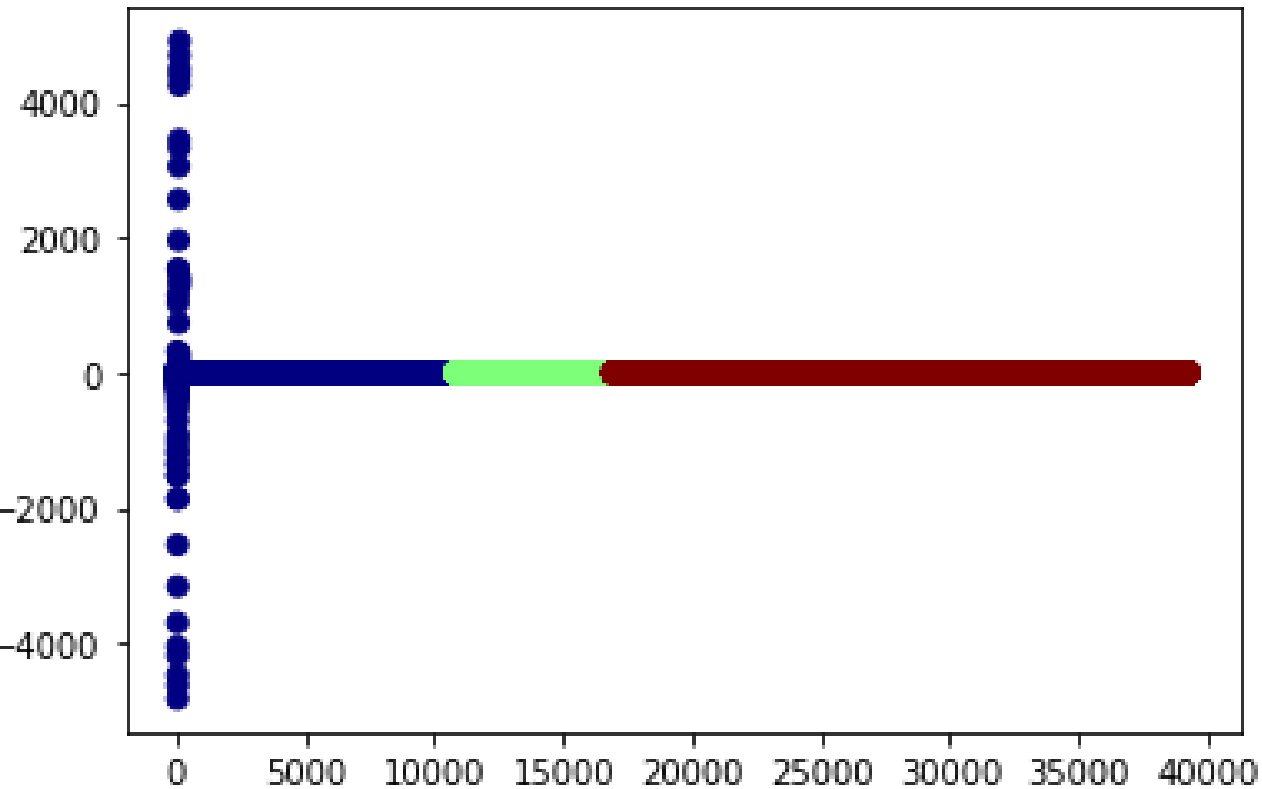


Calinski-Haribasz Index

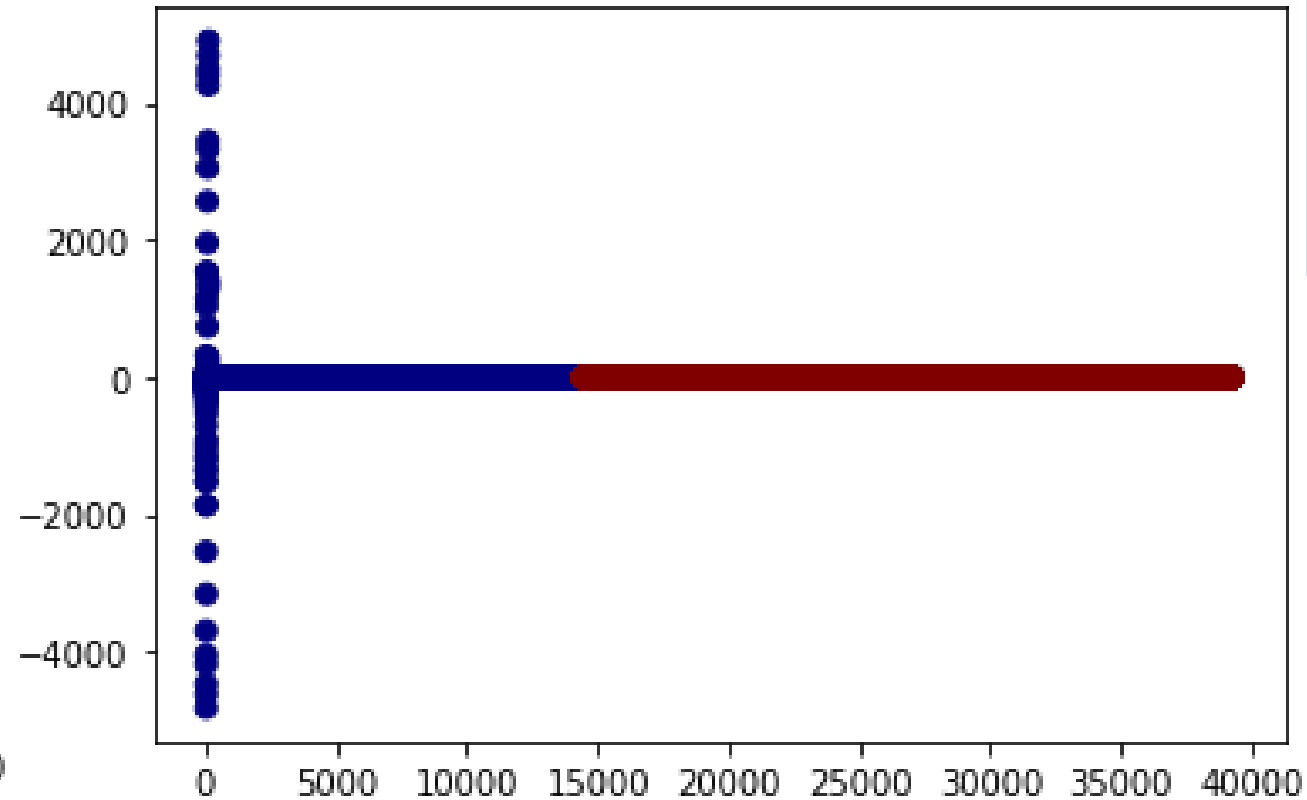
Suggested  
number of  
Cluster 2

# GMM Silhouette Plot

$k = 3$



$k = 2$





# Performance Metric

- ▶ 80/20 Train-Test Dataset split

- ▶ Train Dataset

Accuracy: 0.9335	F1_score: 0.96
------------------	----------------

Precision: 0.999	Recall: 0.9289
------------------	----------------

- ▶ Test Dataset

Accuracy: 0.9316	F1_score: 0.9619
------------------	------------------

Precision: 1.0	Recall: 0.9266
----------------	----------------



# Question?