

# CUSTOMER SEGMENTATION with RFM

September 2024



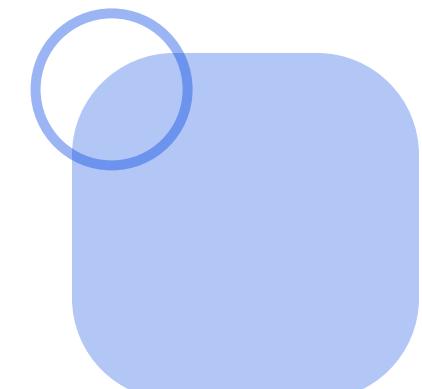
dev.abdullah.khaled@gmail.com

+201557504902





# CONTENT



- |           |  |           |  |
|-----------|--|-----------|--|
| <b>01</b> | Overview                                     | <b>07</b> | Hierarchical Clustering                          |
| <b>02</b> | Business Goal                                | <b>08</b> | RFM with cluster id and<br>cluster labels        |
| <b>03</b> | The Dataset                                  | <b>09</b> | Box plot to visualize<br>Hierarchical Clustering |
| <b>04</b> | The RFM                                      | <b>10</b> | Final Analysis                                   |
| <b>05</b> | The RFM with cluster id                      |           |  |
| <b>06</b> | Box plots to visualize<br>k-means clustering |           |  |
- 

# OVERVIEW



Online retail is a transnational data set which contains all the transactions occurring between 01/12/2010 and 09/12/2011 for a UK-based and registered non-store online retail. The company mainly sells unique all-occasion gifts. Many customers of the company are wholesalers.

# BUSINESS GOAL



We aim to segment the Customers based on RFM so that the company can target its customers efficiently.

# THE DATASET

---

	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country
0	536365	85123A	WHITE HANGING HEART T-LIGHT HOLDER	6	01-12-2010 08:26	2.55	17850.0	United Kingdom
1	536365	71053	WHITE METAL LANTERN	6	01-12-2010 08:26	3.39	17850.0	United Kingdom
2	536365	84406B	CREAM CUPID HEARTS COAT HANGER	8	01-12-2010 08:26	2.75	17850.0	United Kingdom
3	536365	84029G	KNITTED UNION FLAG HOT WATER BOTTLE	6	01-12-2010 08:26	3.39	17850.0	United Kingdom
4	536365	84029E	RED WOOLLY HOTTIE WHITE HEART.	6	01-12-2010 08:26	3.39	17850.0	United Kingdom

# THE RFM

---

	CustomerID	Amount	Frequency	Recency
0	12346.0	0.00	2	325
1	12347.0	4310.00	182	1
2	12348.0	1797.24	31	74
3	12349.0	1757.55	73	18
4	12350.0	334.40	17	309

# THE RFM WITH CLUSTER ID

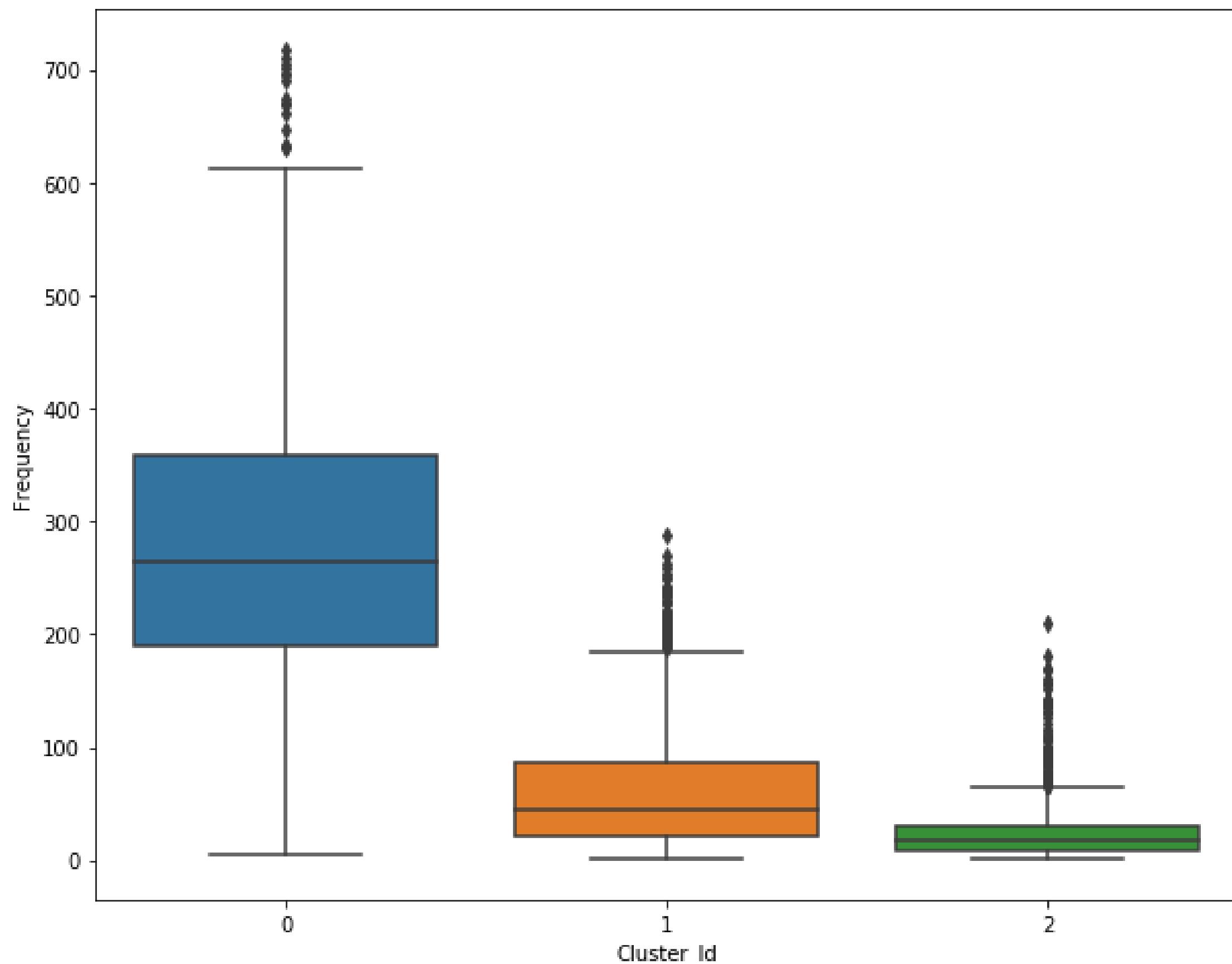
---

T

	CustomerID	Amount	Frequency	Recency	Cluster_Id
0	12346.0	0.00	2	325	2
1	12347.0	4310.00	182	1	0
2	12348.0	1797.24	31	74	1
3	12349.0	1757.55	73	18	1
4	12350.0	334.40	17	309	2

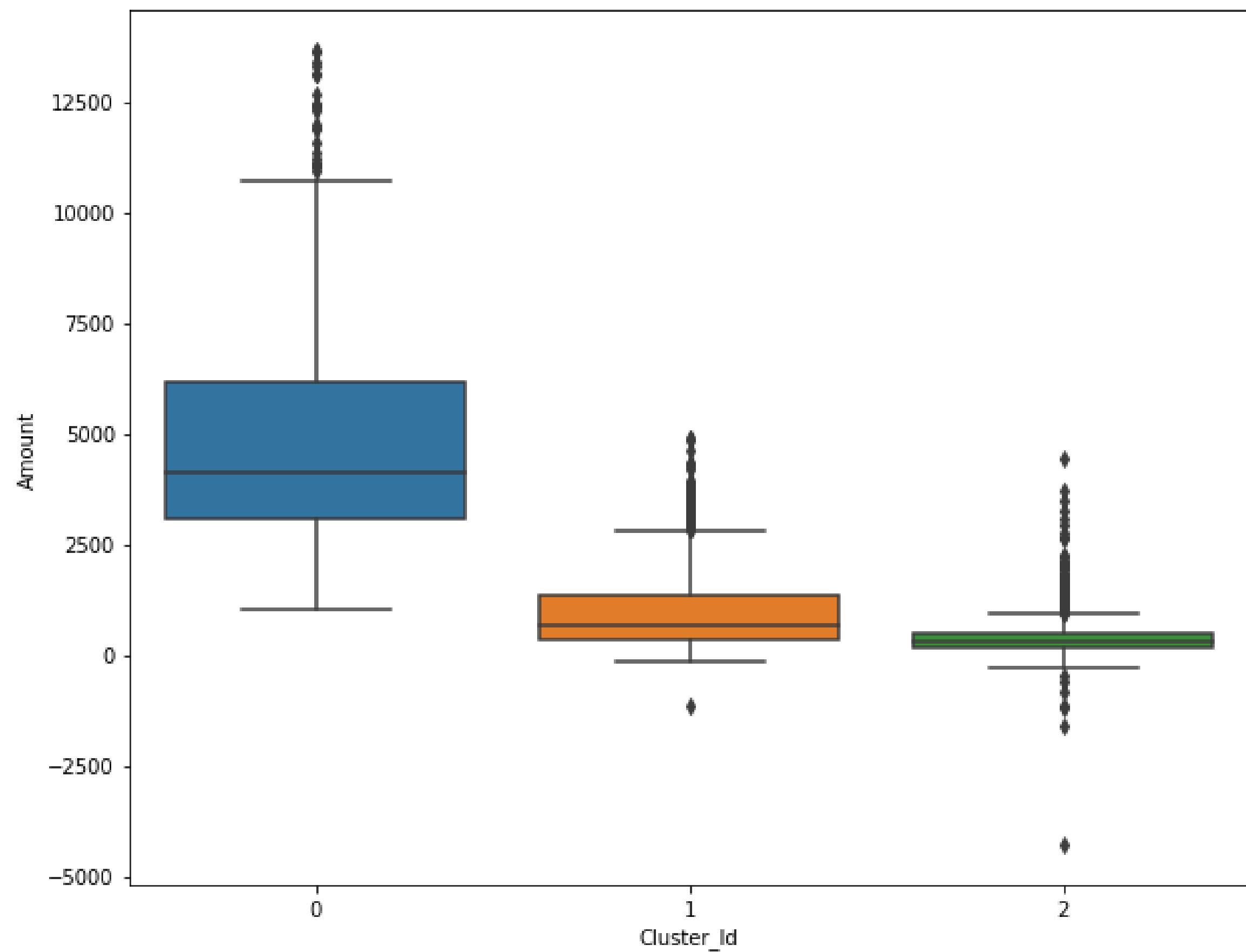
## K-means Clustering

BOX PLOT TO  
VISUALIZE  
CLUSTER ID  
VS  
FREQUENCY



## K-means Clustering

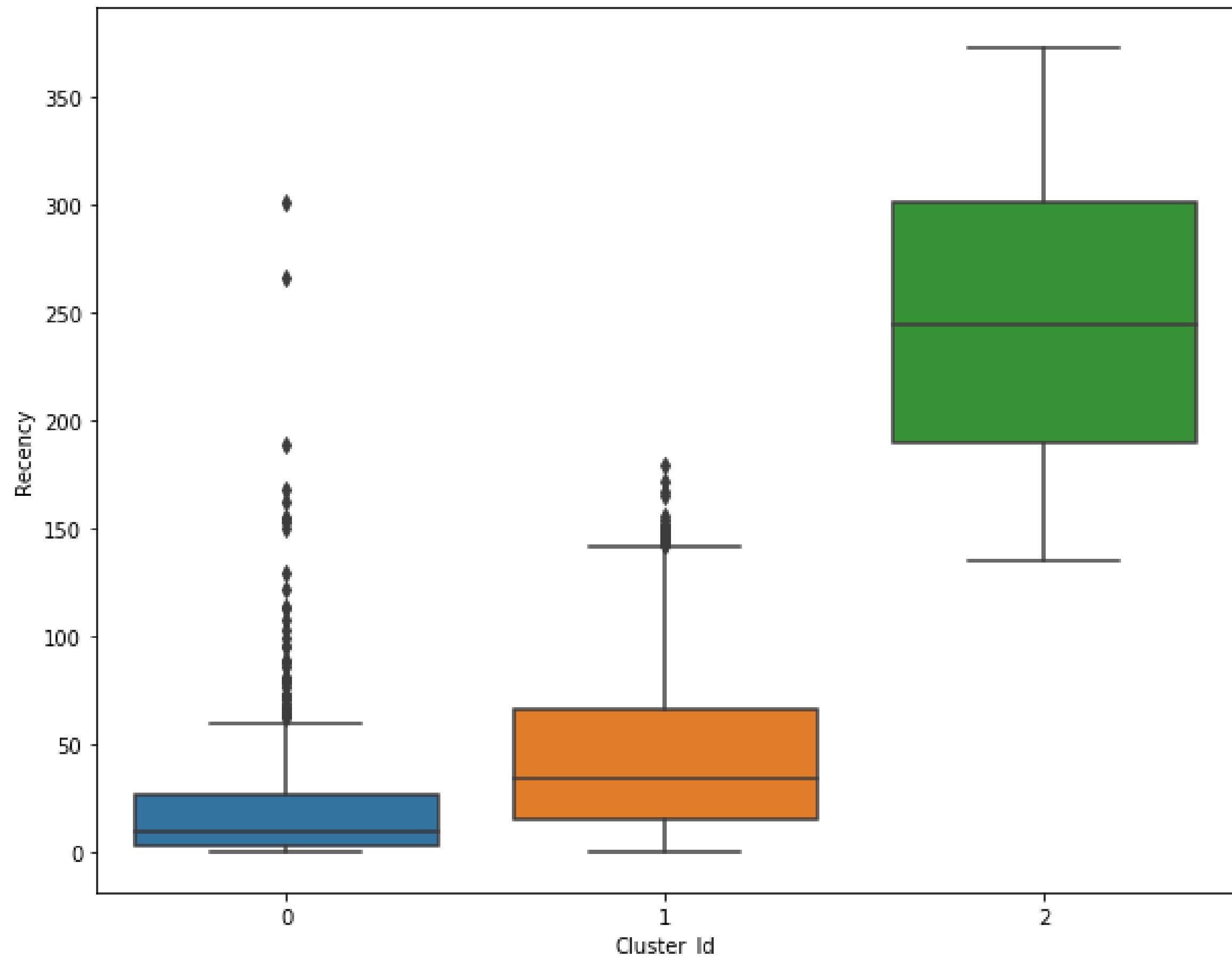
BOX PLOT TO  
VISUALIZE  
CLUSTER ID  
VS  
AMOUNT



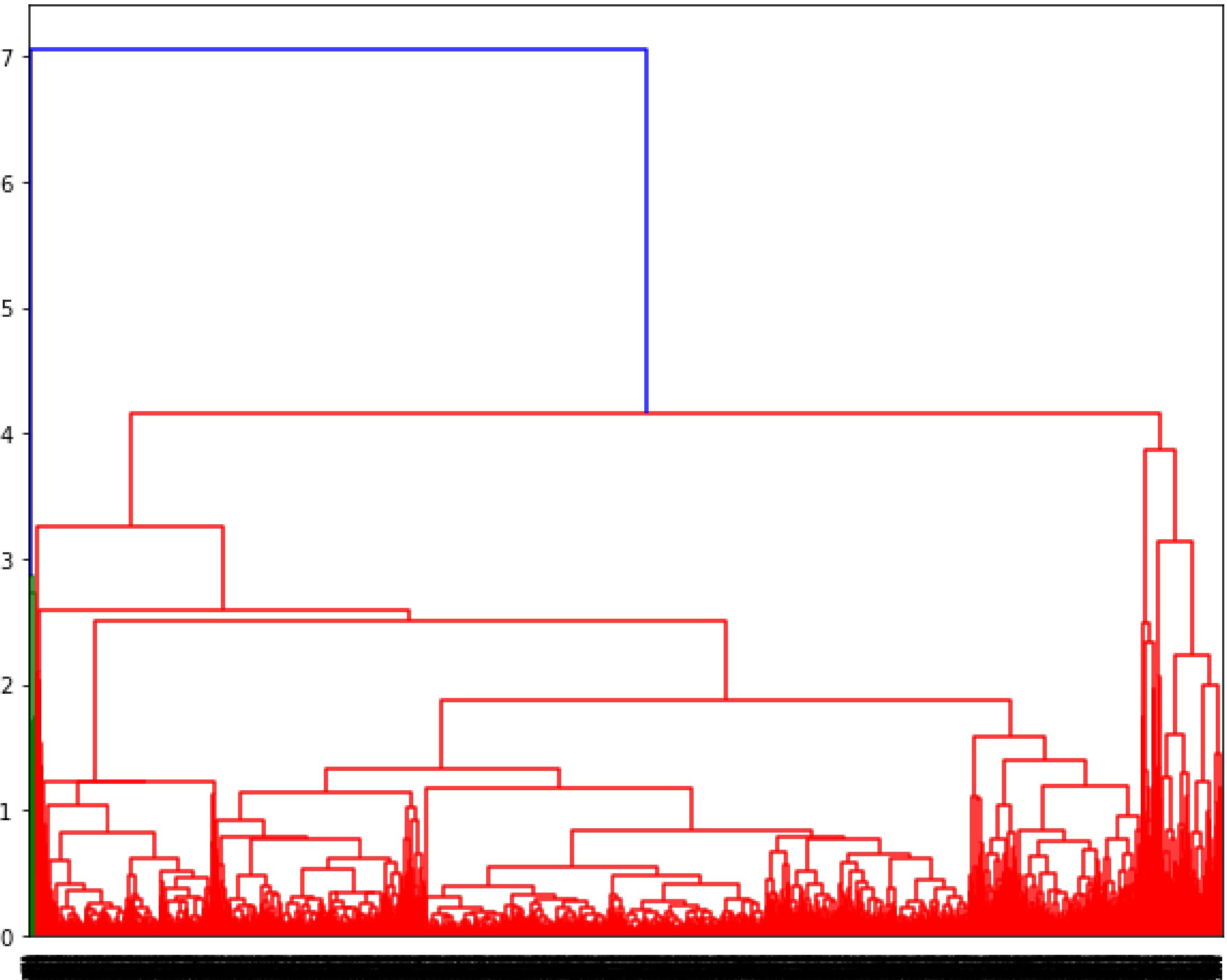
## K-means Clustering

BOX PLOT TO  
VISUALIZE  
CLUSTER ID  
VS  
RECENCY

T



# HIERARCHICAL CLUSTERING (AVERAGE LINKAGE)



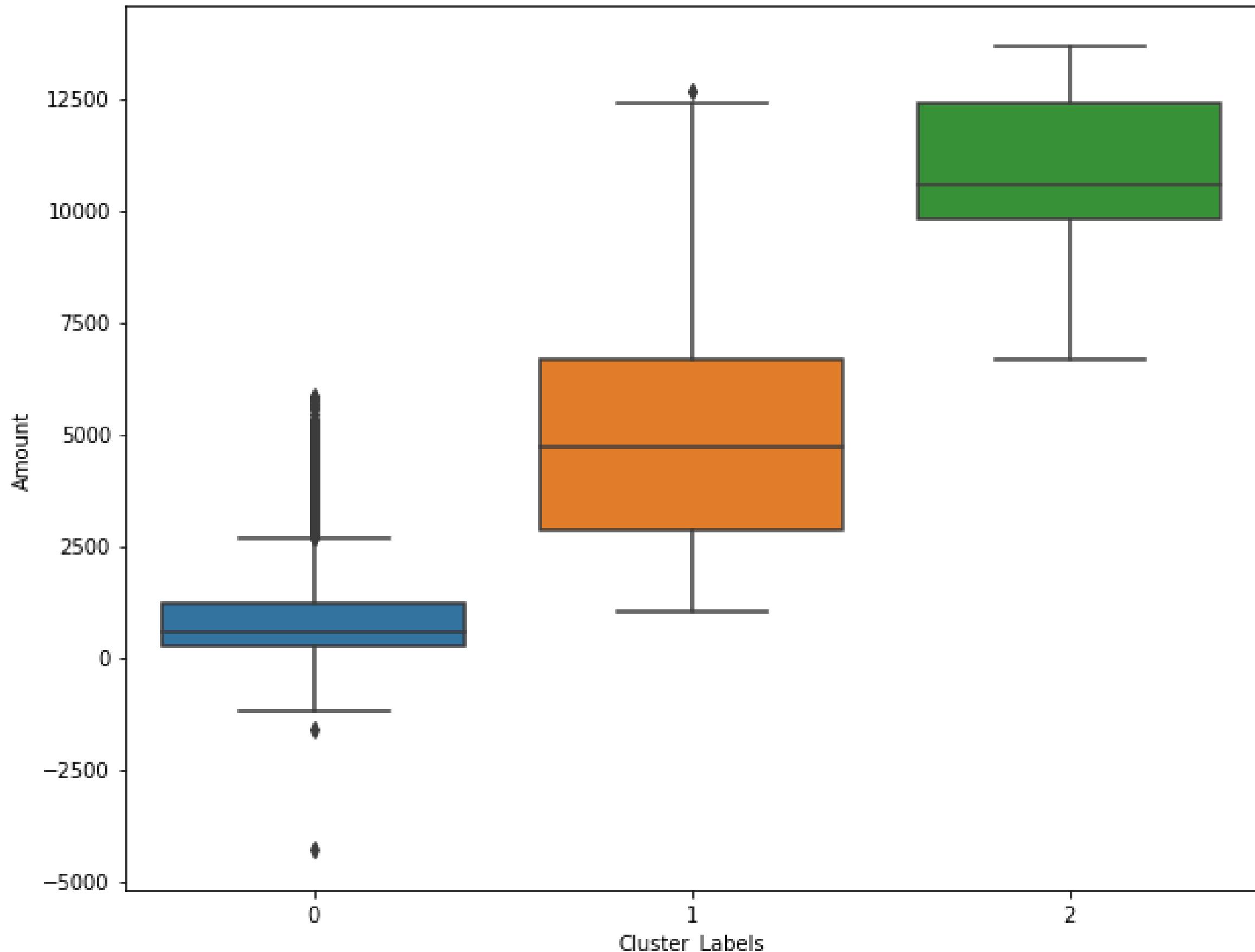
# RFM WITH CLUSTER ID AND CLUSTER LABELS

	CustomerID	Amount	Frequency	Recency	Cluster_Id	Cluster_Labels
0	12346.0	0.00	2	325	2	0
1	12347.0	4310.00	182	1	0	0
2	12348.0	1797.24	31	74	1	0
3	12349.0	1757.55	73	18	1	0
4	12350.0	334.40	17	309	2	0

## Hierarchical Clustering (Average linkage)

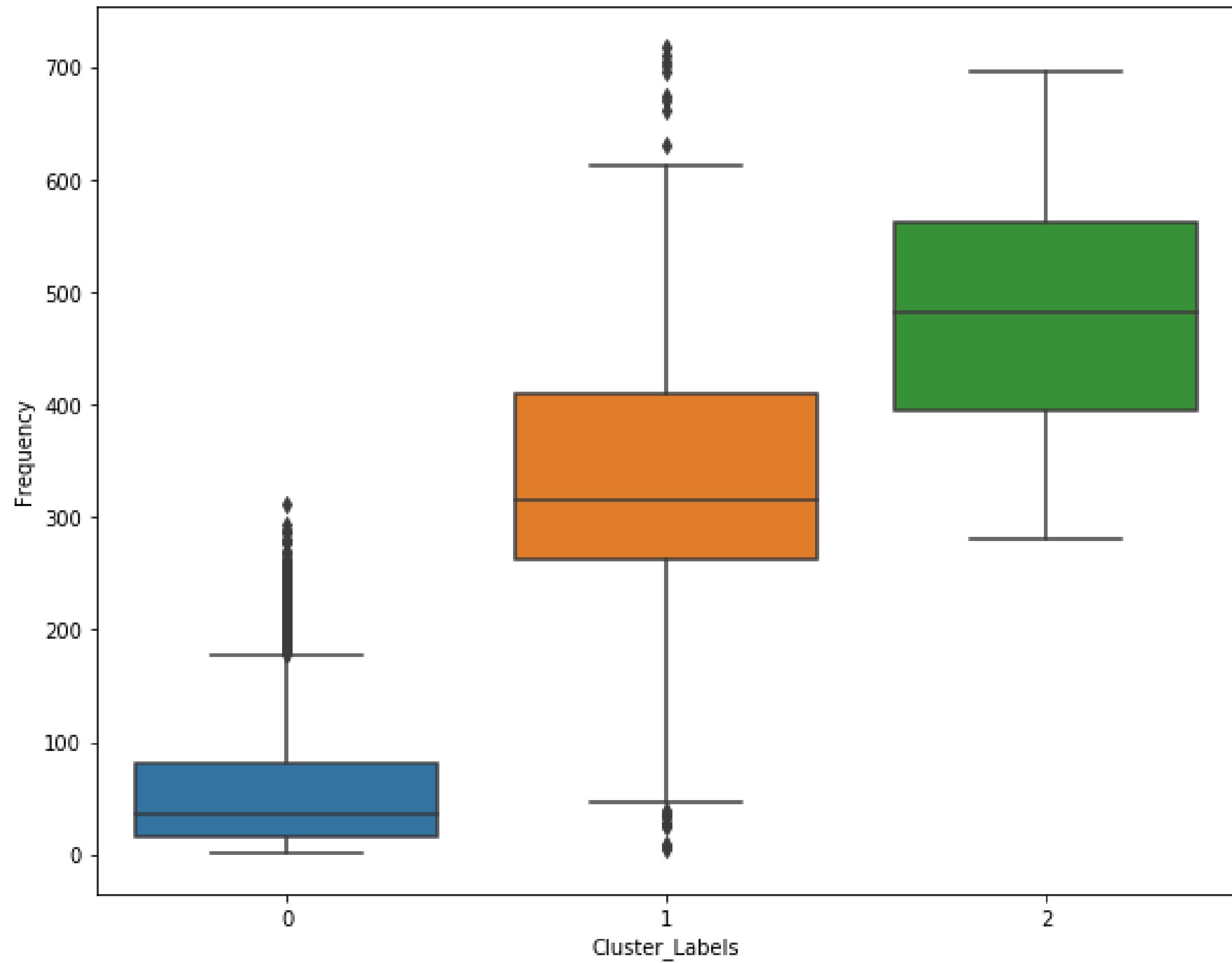
BOX PLOT TO  
VISUALIZE  
CLUSTER  
LABELS

VS  
AMOUNT



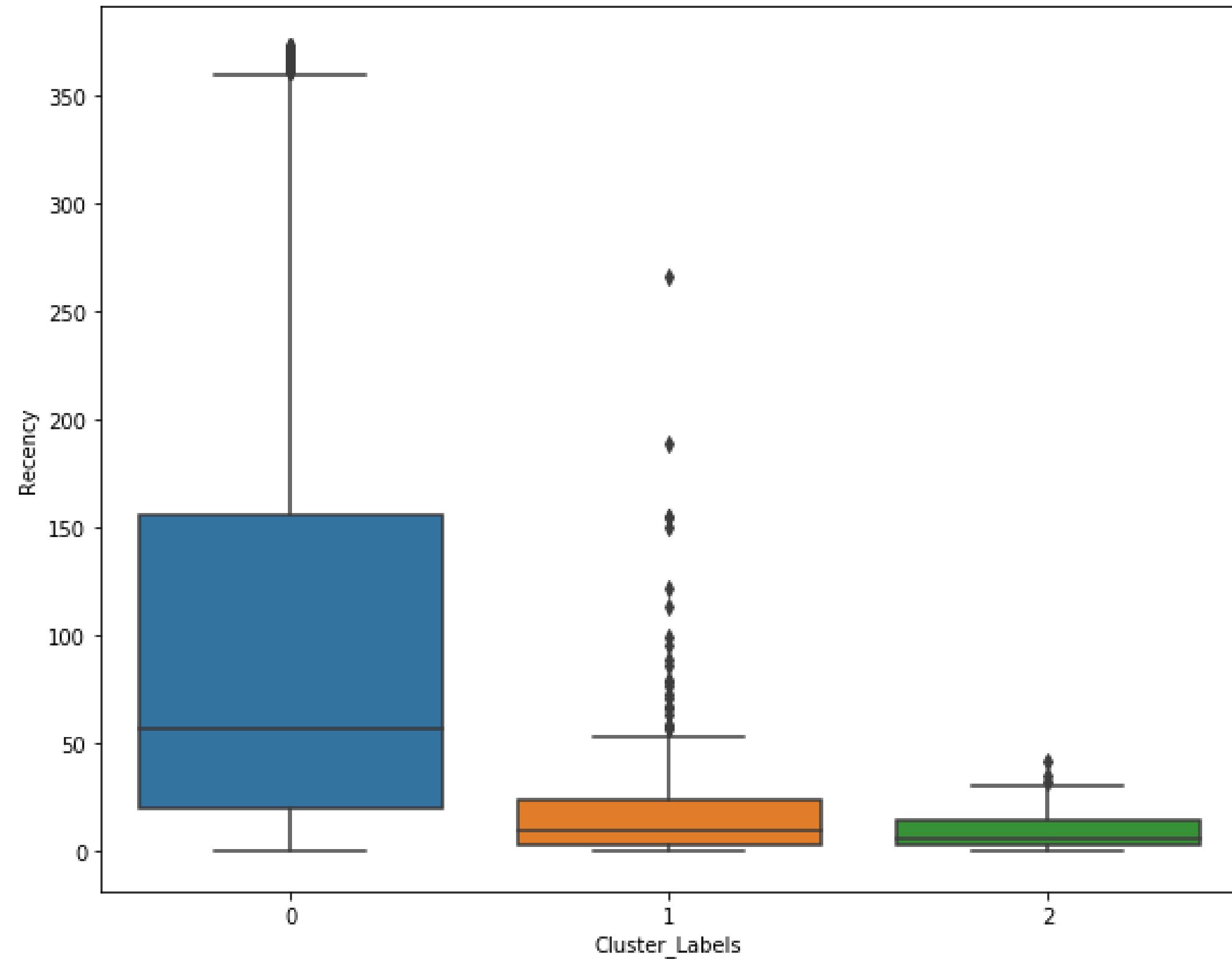
## Hierarchical Clustering (Average linkage)

BOX PLOT TO  
VISUALIZE  
CLUSTER  
LABELS  
VS  
FREQUENCY



## Hierarchical Clustering (Average linkage)

BOX PLOT TO  
VISUALIZE  
CLUSTER  
LABELS  
VS  
RECENCY



# FINAL ANALYSIS



## INFERENCE:



### K-Means Clustering with 3 Cluster Ids

- Customers with Cluster Id 1 are the customers with high amount of transactions as compared to other customers.
- Customers with Cluster Id 1 are frequent buyers.
- Customers with Cluster Id 2 are not recent buyers and hence least of importance from business point of view.

### Hierarchical Clustering with 3 Cluster Labels

- Customers with Cluster\_Labels 2 are the customers with high amount of transactions as compared to other customers.
- Customers with Cluster\_Labels 2 are frequent buyers.
- Customers with Cluster\_Labels 0 are not recent buyers and hence least of importance from business point of view.



# THANK YOU

FOR YOUR ATTENTION



dev.abdullah.khaled@gmail.com

+201557504902

