Model : K-means Clustering

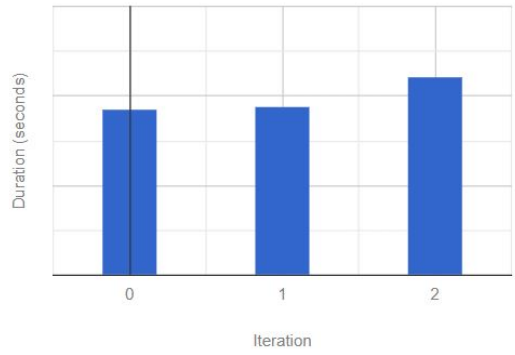
#NUMBER_CLUSTERS = 3

```
CREATE MODEL`bads7105-313104.Supermarketdata.Supermarketdata_CLUSTERS`
OPTIONS( MODEL_TYPE='KMEANS',NUM_CLUSTERS=3,KMEANS_INIT_METHOD='RANDOM')
AS SELECT*FROM`bads7105-313104.Supermarketdata.Supermarketdata`
```

Result :

Iteration	Training Data Loss	Duration (seconds)	Cluster Centroid Id	Cluster Radius
2	14.9072	11.07	1	3.65257595
			2	4.14393402
			3	3.79291357
1	15.0042	9.41	1	3.6669992
			2	4.13359775
			3	3.79386618
0	24.8039	9.26	1	4.89444928
			2	5.13461195
			3	4.86181473

Duration (seconds)



Loss

Metrics

Davies–Bouldin index	4.1016
Mean squared distance	14.9072

Numeric features

This table shows the centroid value for each feature. Use the select menu to view more numeric features.

Selected Features

BASKET_ID, QUANTITY, SHOP_DATE, SHOP_HOUR, SHOP_WEEK, SHOP_WEEKD...

Centroid Id	Count	BASKET_ID	QUANTITY	SHOP_DATE
1	269,736	0.0000	1.2444	
2	280,069	0.0000	1.8880	
3	406,769	0.0000	1.3994	

Categorical features

Each chart below shows the category value distribution for a particular feature. Use the select menu to view more categorical features.

Selected Features
10 options selected