Zeotap Assignment
Task-3

Objective: Segmenting customers based on their transaction behavior using some cluster techniques, and also utilizing both profile information (from Customers.csv) and transaction information (from Transactions.csv).

Number of Clusters Formed: 3

Cluster Technique Used: K-Means

Clustering Insights:
This analysis by clustering shows distinct customer segments based on their transaction behaviors. The segments represent customers with a high purchasing history and frequent or occasional customers**.**

Further, It can be observed through the scatter plot that the performance improves by using 2 principal components for separating the 3 clusters, although, there is still some overlapping but it is separated enough, to be used for making some data-driven decisions.

Clustering Evaluation Metrics:
DB-Index: The DB index value is 0.7888, which is between 0.5 to 1, that indicates, the clusters are relatively well-separated, but there are some minor overlaps.

I used one more Evaluation metric which is Silhouette Score, and for this clustering, I got a value of 0.4239. A score closer to 1 indicates well-defined clusters, while a score closer to -1 indicates poor clustering, here the score suggests that clustering is decent, but still have some room for improvement.

Conclusion: From the clustering, it is evident that the segmentation is useful for identifying different customer groups based on transaction history, although future improvements can be made by experimenting with more features, different algorithms, or fine-tuning the number of clusters