Clustering Results Report

The K-Means clustering algorithm successfully split the dataset into **four different clusters**. These clusters are meaningful groups of customers according to their total spending and the quantity of purchases, providing significant insight into the customer's behavior.

Two crucial metrics were used to evaluate the quality of clustering. The calculated ***Davies-Bouldin Index*** is ***0.710***, meaning that the clusters are compact and well-separated. The lower value of the Davies-Bouldin Index indicates that the algorithm for clustering customers groups similar customers well and has distinct boundaries between the clusters. Such a value will show a high degree of intra-cluster cohesion as well as good inter-cluster separation.

The ***Silhouette Score*** was computed to be ***0.454***, which indicates that the clusters are moderately well-defined. The score does indicate that most customers are well-classified to their respective clusters; however, it also indicates slight overlap between a few clusters, and this can be minimized by considering more features such as product preference, purchase frequency, or seasonal trends, for further fine-tuning of the clustering process.

The overall clustering results can be considered a good basis for customer segmentation. The resultant segmentation can then be used in creating customized marketing strategies for each cluster. For example, loyalty programs and premium offerings can be given to high-spending customers, while low-spending customers can be offered discounts or bundled promotions to increase their engagement. This reveals the possibility of optimizing business strategies with the help of customer segmentation.

References

McKinney, W. (2017). *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython* (2nd ed.). O'Reilly Media.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830. Retrieved from https://scikit-learn.org/stable/

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, 9(3), 90–95. doi:10.1109/MCSE.2007.55