Task 3: Customer Segmentation / Clustering

- The Davies-Bouldin Index is a metric used to evaluate the quality of clustering. It measures how well-separated and compact the clusters are.
- A **lower DBI** value indicates that the clusters are more compact (i.e., the points within each cluster are close to each other) and well-separated (i.e., the clusters are far apart from each other).
- A **higher DBI** suggests that the clusters are either overlapping or not well-separated.

DBI of 0.972 is relatively **low** and generally indicates that your clustering result is of **good quality**. Specifically:

- The clusters are reasonably compact, meaning that customers within the same cluster are fairly close to each other.
- The clusters are sufficiently **separated**, which suggests there is a clear distinction between the groups formed by the clustering algorithm.

Cluster Characteristics:

- Since we have seven clusters, each cluster likely represents a **group of customers** with specific shared characteristics (e.g., purchasing patterns, spending behavior, demographics).
- The clusters could potentially be analyzed further to **identify customer segments**, such as:
    - High-spending customers.
    - Frequent buyers.
    - Low engagement customers.
    - Customers from different regions or demographic backgrounds.

Evaluating Cluster Quality (Compactness and Separation):

- **Compactness**:
    - **Compactness** refers to how closely related the data points within a cluster are. A low DBI means that the customers in the same cluster are relatively close to each other in terms of their features (e.g., spending behavior, transaction frequency).
- **Separation**:
    - **Separation** measures how far apart the clusters are from each other. A low DBI indicates that the clusters are sufficiently spread out, making it easy to distinguish between different customer segments.

Conclusion:

- The **DBI value of 0.972** indicates that your **seven clusters** are well-formed, with reasonably well-separated and compact groups of customers.
- The **number of clusters** seems appropriate, and the clustering algorithm appears to have captured distinct patterns in customer behaviour.