Task 3: Customer Segmentation / Clustering

1. Objective:
   Group customers into distinct segments based on their profile information (region, signup year) and transaction behavior (total purchase value, quantity purchased).

2. Clustering Process:

- We used KMeans clustering, a common method to group similar data points into clusters.
- Features like region, total spending, signup year, and total quantity purchased were selected.
- Data was standardized (scaled) to ensure fair comparison across different units (e.g., dollars vs. years).

3. Cluster Evaluation:

- The Davies-Bouldin Index (DB Index) was calculated to measure cluster quality. A lower DB Index means better clusters.
- Silhouette Score was also used to confirm that the clusters are well-separated and compact.

4. Visualization:

- A scatter plot shows how customers are grouped into clusters, with each cluster having its own color for easy identification.
- This makes it easy to interpret the clustering results and understand the customer segments.

5. Outcome:

- The number of clusters, the DB Index value, and the visual representation provide insights into how customers are segmented and how distinct these groups are.