Customer Segmentation

1. **Objective**:

   The goal is to group customers into meaningful categories (or clusters) based on their behavior—such as how much they've spent and how long they've been with the business. This helps in understanding customer patterns and creating targeted strategies.

2. **Datasets**:

   a. **Customer Data**: Contains basic customer details like signup dates.
   b. **Transaction Data**: Contains details of purchases made by customers, including the amounts spent.

3. **Step 1 - Calculate Total Spending**:

   a. Combine the customer and transaction data using their unique IDs.
   b. For each customer, calculate the total amount they've spent by summing up their transactions.
   c. Fill in zero for customers with no recorded transactions.

4. **Step 2 - Calculate Customer Tenure**:

   a. Calculate the number of days a customer has been with the business by subtracting their signup date from today's date.
   b. This gives a measure of how long they've been loyal to your business.

5. **Step 3 - Standardize the Data**:

   a. Spending amounts and tenure values can vary widely (e.g., some customers spend $500 while others spend $10).
   b. Standardize the data so all values fall within a similar range, making it easier to compare and analyze.

6. **Step 4 - Apply K-Means Clustering**:

    a. Use the K-Means algorithm to group customers into four clusters based on their standardized spending and tenure data.
    b. Each cluster represents a group of customers with similar behaviors, such as high spenders, loyal customers, or new customers.

7. **Step 5 - Evaluate Clusters**:

    a. Calculate the **Davies-Bouldin Index** to check the quality of the clusters.
    b. A lower Davies-Bouldin score means the clusters are more distinct and meaningful.
    c. Here Davies-Bouldin Index: 0.7947787492. This is a relatively low value, which suggests that the clusters are likely well-separated and compact.

8. **Step 6 - Visualize the Clusters**:

    a. Create a scatter plot where each point represents a customer.
    b. The x-axis shows standardized spending, the y-axis shows standardized tenure, and the colors represents the clusters.
    c. This visualization helps identify patterns, like which group includes loyal customers or low spenders.

9. **Insights and Applications**:

    a. Use the clusters to design business strategies. For example:
        i. Reward high-spending, loyal customers with special offers.
        ii. Encourage new or low-spending customers to engage more with discounts or promotions.
    b. The visualization helps make data-driven decisions to improve customer engagement and business performance.

Actionable Insights:

- **Cluster 0 (Purple)**:
  - These are potentially **high-value customers**. Focus on retaining them through loyalty programs or exclusive offers.
- **Cluster 1 (Green)**:
  - These are **loyal, medium-spending customers**. Engage them further to increase their spending (e.g., by upselling or cross-selling).
- **Cluster 2 (Yellow)**:
  - These could be **low-priority customers**. Re-engage them through targeted promotions to increase activity.
- **Cluster 3 (Blue)**:
  - These are **new customers**. Create onboarding campaigns to nurture their loyalty and encourage spending.