

# “Customer Segmentation”

## Explaining K-Means Clustering in Simple Terms

I performed the steps of using K-Means Clustering to group customers based on how they spend.

**Gathering Data:** First, we collect customer information and their transaction records. This includes:

1. Customer Profiles: Basic details like name and region.
2. Transactions: How much they buy and how often.

This helps us understand our customers better.

**Understanding Customer Spending:** We focus on:

1. Total Money Spent: How much a customer has spent overall.
2. Total Number of Purchases: How many times they've bought something.

This gives us insight into their shopping habits.

**Combining Data:** Next, we combine customer details with their transaction info. This way, we have a complete picture of each customer's spending history and number of purchases in one place.

**Focusing on Key Details:** For clustering, we only consider:

1. TotalSpent (money spent) and
2. TransactionCount (number of purchases).

These details are crucial for grouping customers.

**Normalizing Data:** We scale the data so that spending doesn't overshadow the number of purchases. This ensures that the program can fairly compare customers, whether they spend a lot or a little.

**Grouping with K-Means:** Now, we use K-Means to group customers based on their spending habits and frequency. For example, with 4 clusters, we might get:

Group 1: Big spenders who shop often.

Group 2: Small spenders who rarely shop.

Group 3: Moderate spenders with steady activity.

Group 4: Irregular buyers who spend a lot but shop rarely.

K-Means finds the best way to group customers.

**Measuring Group Quality (DBI Index):** After forming groups, we check how well they are separated. The Davies-Bouldin Index (DBI) helps us here:

1. Small DBI value: Groups are distinct.
2. High DBI value: Groups overlap, making clustering less useful.

**Visualizing Clusters:** Finally, we visualize the results. We plot customers on a graph, with each dot representing a customer. Dots of the same color belong to the same cluster. This helps us see which groups have the most customers or the highest spending.

### Insights:

1. Number of Clusters: Tells us how many customer groups there are.
2. Type of Customers in Each Cluster: For example, high spenders, frequent shoppers, low spenders, or rare high spenders.
3. DBI Value: Shows how clear the groups are (lower is better).

### Why is This Useful?

Businesses can tailor marketing and offers to specific groups and identify high-value customers and focus on them.

Understanding these groups helps in making better business decisions.

### Output:

