E-commerce Customer Segmentation

Abstract:

A key challenge for e-commerce businesses is to analyze the trend in the market to increase their sales. The trend can be easily observed if the companies can group the customers; based on their activity on the e-commerce site. This grouping can be done by applying different criteria like previous orders, mostly searched brands and so on.

Problem Statement:

Given the e-commerce data, use k-means clustering algorithm to cluster customers with similar interest.

Dataset Information:

The data was collected from a well known e-commerce website over a period of time based on the customer's search profile.

Variable Description:

Column	Description
Cust_ID	Unique numbering for customers
Gender	Gender of the customer
Orders	Number of orders placed by each customer in the past

Remaining 35 features (brands) contains the number of times customers have searched them

Scope:

Problem Statement – K means

- Analyzing the existing customer data and getting valuable insights about the purchase pattern
- Data pre-processing including missing value treatment
- Segmenting customer based on the optimum number of clusters ('k') with the help of silhouette score

Learning Outcome:

The students will get a better understanding of how the variables are linked to each other and will be able to apply cluster analysis to business problem such as customer segmentation.