Customer Segmentation Report

Varun Goyal

January 27, 2025

Introduction

This report details the customer segmentation analysis conducted using KMeans clustering algorithm. The analysis utilized three key features: DaysSinceSignup , Total Spent, Total Quantity. The goal was to identify distinct customer segments to enhance targeted marketing strategies and improve customer engagement.

KMeans Clustering

Methodology

The K-Means algorithm is a popular unsupervised learning method used for clustering data into k distinct groups. The Elbow method was used to determine the optimal number of clusters, which was found to be 4.

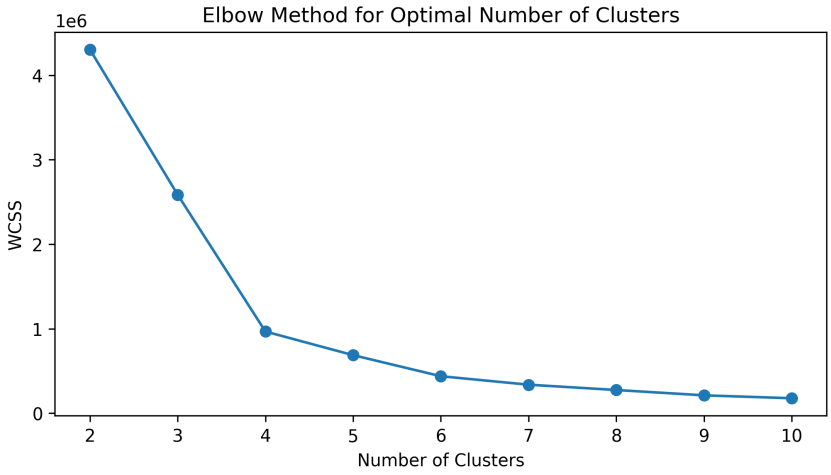

Figure 1: Elbow Method for Determining Optimal Clusters

Clustering Metrics

The clustering performance was evaluated using the Davies-Bouldin Index and Silhouette Score:

- **Davies-Bouldin Index:** 0.4553932001330353
- **Silhouette Score:** 0.622603236902826

These metrics suggest moderate cluster separation, indicating potential overlap between clusters.

Visualization

The 3D visualization below illustrates the customer segmentation achieved with KMeans, highlighting the distribution of customers across the identified clusters.

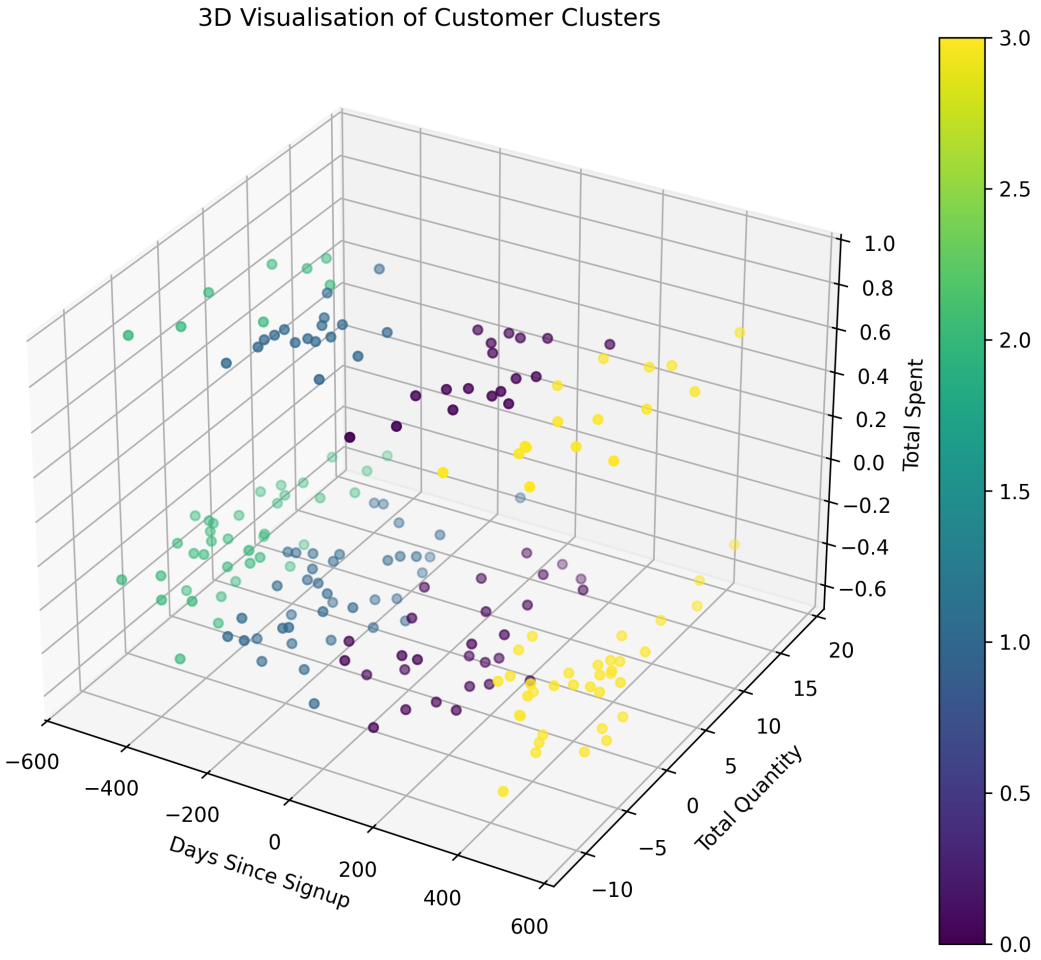

Figure 2: 3D Customer Segmentation with KMeans