**Customer Segmentation Using Clustering**

### ****Goal****

The goal of this task was to segment customers of a shopping mall based on their demographic and spending patterns, using unsupervised learning. This helps businesses understand customer groups and tailor marketing strategies.

### ****Dataset****

* Source: Mall Customer Segmentation Data (Kaggle).
* Features used:
  + Age
  + Annual Income (k$)
  + Spending Score (1–100)
* Gender was mapped to numeric values (0 = Male, 1 = Female) but not used in clustering.

### ****Exploratory Data Analysis (EDA)****

* **Age Distribution**: Most customers are between 20–50 years old.
* **Annual Income Distribution**: Income is fairly spread, with peaks around 40–70k.
* **Spending Score Distribution**: Customers are evenly distributed across low and high spending.
* **Correlation Heatmap**: Weak correlations were found among age, income, and spending score, confirming the need for clustering.

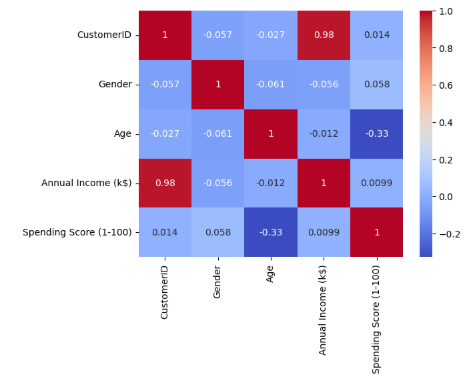


Figure Correlation Heatmap

### ****Feature Scaling****

* StandardScaler was applied to ensure all features (Age, Annual Income, Spending Score) had equal weight in clustering.

### ****Clustering****

* **Elbow Method**: The plot of inertia suggested an optimal number of clusters around **k = 5**.
* **K-Means with k=5** was applied to the scaled features.
* **Silhouette Score**: The clustering quality was around 0.55 (moderate but reasonable for real-world customer data).

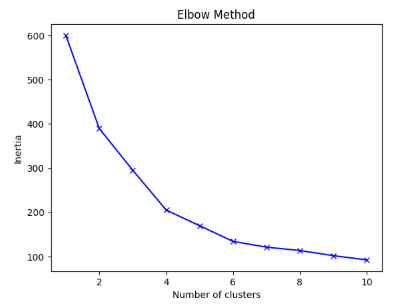


Figure Elbow method

### ****Cluster Insights****

From the scatterplot of **Annual Income vs Spending Score**, clusters represent distinct customer groups:

1. **Low Income – Low Spending**: Price-sensitive customers, likely occasional buyers.
2. **High Income – High Spending**: Premium customers; ideal for loyalty programs.
3. **Low Income – High Spending**: Enthusiastic but budget-limited shoppers.
4. **Mid Income – Moderate Spending**: Average customers, potential to upsell.
5. **High Income – Low Spending**: Wealthy but less engaged shoppers.

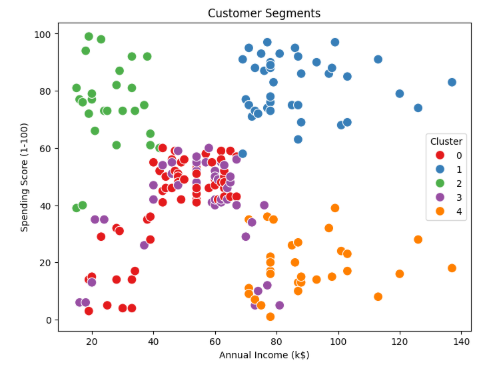


Figure Clusters

### ****Business Insights****

* Cluster 0 represents **budget-conscious customers** who may respond well to discounts.
* Clusters 1 & 2 (high spending) can be targeted for **exclusive offers & loyalty programs**.
* Cluster 3 forms the **steady middle group**, contributing consistently to sales.
* Cluster 4 customers have high income but low spending, suggesting untapped potential — targeted advertising or personalized offers might engage them more.

### ****Conclusion****

Clustering successfully divided mall customers into **five distinct groups**. These insights can help businesses **personalize marketing**, **allocate resources effectively**, and **improve customer retention**.