- There were no missing values in the dataset.
- For clustering, only the "Annual Income" and "Spending Score" features were considered.
- The features were standardized using the StandardScaler from scikit-learn to ensure equal importance during clustering.

- The Within-Cluster Sum of Squares (WCSS) method was used to determine the optimal number of clusters.
- The elbow plot (WCSS vs. number of clusters) was created, and the optimal number of clusters was identified as 5 based on the "elbow" in the plot.

- The K-Means algorithm was applied to the scaled "Annual Income" and "Spending Score" features with the number of clusters set to 5.
- The algorithm assigned each data point to one of the 5 clusters based on its proximity to the cluster centroids.

- The clusters were visualized by plotting the data points on a scatter plot, with the x-axis representing "Annual Income" and the y-axis representing "Spending Score".
- The cluster centroids were also plotted on the same scatter plot.

- Cluster 4: This cluster consists of customers with relatively high spending scores but low annual income.
- Cluster 2: This cluster consists of customers with high annual income but low spending scores.
- Cluster 5: This cluster consists of customers with both high annual income and high spending scores.
- Clusters 1 and 3: These clusters represent customers with moderate annual income and spending scores.

- **Insights**:

- The K-Means clustering algorithm effectively identified distinct groups of customers based on their annual income and spending behavior.
- The cluster analysis revealed different customer segments, which could be valuable for targeted marketing strategies or personalized offerings.
- Cluster 4 (high spenders with low income) and Cluster 5 (high spenders with high income) could be the most profitable segments for the mall, as they represent customers with high spending tendencies.
- Cluster 2 (high income, low spenders) could be a potential target for promotional campaigns or offers to encourage increased spending.
- The clustering results can help the mall management understand their customer base better and tailor their marketing efforts accordingly.