Customer Segmentation using K-Means Clustering

Presented to:
Md. Mynoddin
Assistant Professor, Department of CSE, RMSTU

Presented By:
Md. Hasibul Islam (2018-15-30)
Afrida Tasnim Niha (2018-15-25)

INTRODUCTION

1. Businesses need data-driven strategies to categorize customers effectively.
2. K-Means Clustering groups customers based on Annual Income and Spending Score.
3. Enables personalized marketing and improves customer retention.

Future Work

1. Integrate **Age** and **Gender** for multi-dimensional segmentation.
2. Test **DBSCAN** or **Hierarchical Clustering** for comparison.
3. Develop a dashboard for real-time cluster monitoring.

Limitations

1. **Feature Scaling Missing**
   Income and spending score not normalized.
2. **Spherical Clusters**
   K-Means struggles with complex shapes
3. **Static Analysis**
   No real-time updates or demographic integration.

INTRODUCTION

1. **Targeted Marketing**
   Identify distinct customer groups for marketing.
2. **Reduce Manual Effort**
   Reduce manual effort in customer analysis.
3. **Data-Driven Insights**
   Improve business strategies using data-driven insights.
4. **Enhance Satisfaction**
   Enhance customer satisfaction through tailored offers.

Motivation

1. **Time-Consuming**
   Manual segmentation is time-consuming and subjective.
2. **Growing Demand**
   Growing demand for automated solutions in retail.
3. **Hidden Patterns**
   Leverage machine learning to uncover customer behavior.

Tools & Technologies

Python
Programming Language

Libraries
Pandas, Scikit-learn, Matplotlib, Seaborn

Platform
Google Colab

Dataset Collection

- **Source** : **Kaggle**
- **Dataset:** **Mall Customer Information**
- **Features:** **200 rows, 5 Columns**

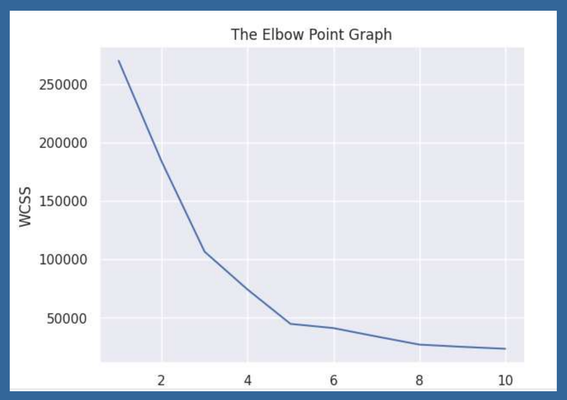
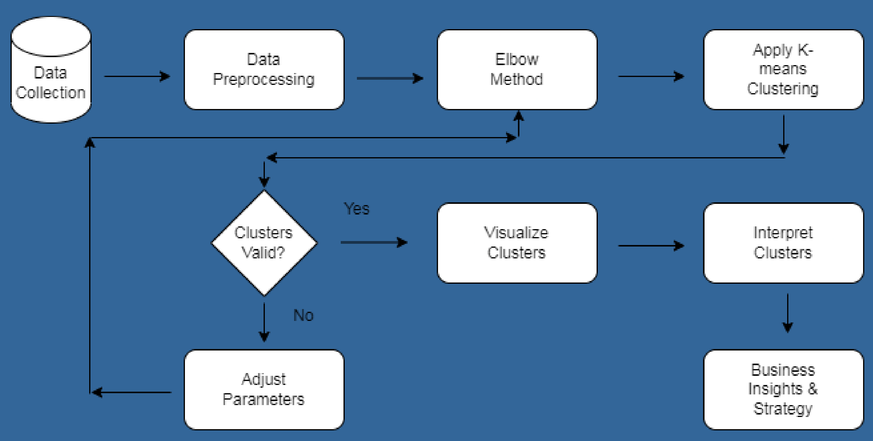


Figure 1. Elbow Point Graph



Block Diagram: Customer Segmentation Using K-Means Clustering

Methodology

1. **Data Preprocessing**
   Check for null values.
   Select features.
2. **Elbow Method**
   Determine optimal clusters (k=5).
3. **Clustering**
   Apply K-Means algorithm.
4. **Visualization**
   Scatter plot of clusters with centroids.
5. **Interpretation**
   Label clusters
   (e.g., "High Income, Low Spenders").

Results

1. Identified 5 distinct customer groups.
2. Enabled targeted marketing strategies.
3. Improved customer retention plans.

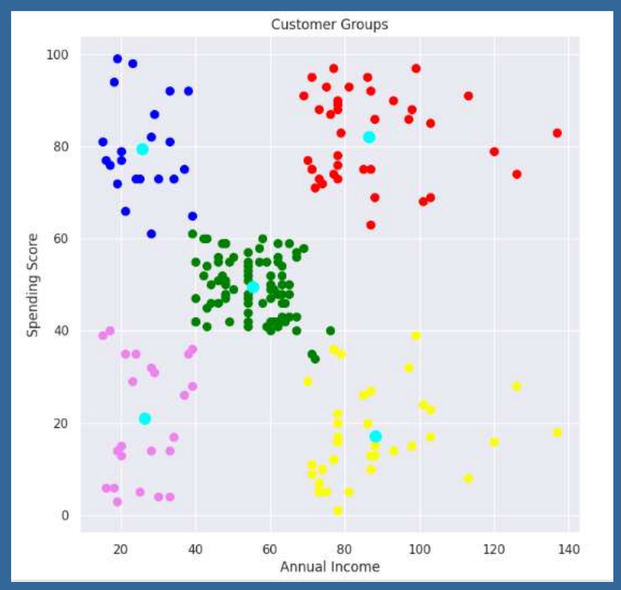


Figure 2. Cluster Visualization

Conclusion

1. K=5 clusters effectively segment customers into actionable groups.
2. High-income low-spenders require targeted promotions.
3. Model provides a foundation for advanced retail analytics.