Balaji Vijayaraj

v22balvi@du.se

Lab – 3 Report

**TASK – 1**

**Introduction:**

The purpose of this report is to discover potential new locations in Sweden that could be suitable for IKEA department stores using k-means clustering. We have been provided with a dataset, ikea\_kommun\_data.txt, which contains essential features for many of Sweden's municipalities, such as population density, average income, and unemployment rate. By applying k-means clustering, we aim to group these municipalities based on their similarities in terms of these features, which will allow us to identify potential locations where IKEA stores could be successful. The report will discuss the data preprocessing steps, the selection of the optimal number of clusters, and the interpretation of the clustering results. Finally, we will provide recommendations for potential new IKEA store locations based on the analysis.

**Method:**

1. Data cleaning: The dataset was checked for data types, and categorical data/textual was excluded to ensure only numerical data was used for clustering.
2. PCA: Principal component analysis (PCA) was performed for dimensionality reduction, and the data was normalized/scaled to prepare it for clustering. PCA was chosen to explain 90-95% of the total variance in the data.
3. Elbow method: The Elbow method was applied to select the optimum number of clusters, which turned out to be 3.
4. K-means clustering: K-means clustering was applied using the results from PCA to the induction model.
5. Cluster labeling: A cluster label column was appended to the original dataset.
6. Has IKEA column: A "has\_ikea" column was appended to the original dataset.
7. Cluster exploration: The clusters were explored and found to be separated into small, medium, and large municipalities.
8. Analysis: The clusters were analyzed based on population, sales index, revenue, etc. Municipalities where IKEA already exists or is close to existing IKEA were excluded from the analysis.

**Result:**

The selection of new stores was based on two criteria, namely population and sales index. After analysis and evaluation, the top three candidates that emerged were Borås, Eskilstuna, and Södertälje. These municipalities were identified as potential locations for new IKEA stores.

Note: Cities located in close proximity to existing IKEA stores are excluded from the analysis.