

DATAMININGPROJECT

Master in Data Science and Advanced Analytics

NOVA Information Management School

Universidade Nova de Lisboa

Strategical Guideline Proposal: A Cluster Analysis Approach

Prepared for client:

ABCDEeats Inc.

Company: Group 99

Consultants:

Hugo Fonseca, 20240520

Ricardo Pereira, 20240745

Jéssica Vicente, 20230744

Fall/Spring Semester 2024-2025

1. INTRODUCTION

The main objective of our application is to facilitate customer segmentation and exploratory data analysis. The application aims to help users uncover meaningful patterns and insights from the diverse customer dataset, enabling data-drive decision-making processes. With an intuitive interface, users can explore and visualize in an easier way data distributions and analyze cluster characteristics without requiring programming skills.

2. Main Features

This application is divided into four main areas:

- **Cluster Exploration:**
 - Detailed exploration of customer clusters, including size, summary statistics, and key characteristics.
 - Visualizations such as bar charts to analyze average metrics for selected clusters and feature groups.
- **Filter Clusters:**
 - Interactive filtering of clusters based on attribute ranges.
 - Visualization of cluster composition through pie charts.
- **Compare Clusters:**
 - Side-by-side comparison of selected clusters using feature distributions.
 - Box plots to highlight variations across clusters for specific features.

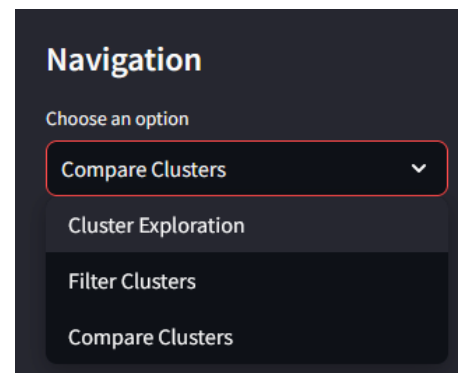


Figure 1-Navigation

3. RESULTS

The application provides users with the ability to explore EDA results and derive actionable insights from customer segmentation. By visualizing and comparing clusters, users can identify patterns such as:

- Customer groups with similar purchasing behavior or preferences.
- Outlier clusters representing unique or underserved customer segments.
- Key factors influencing customer behaviors across clusters.

Cluster Characteristics:

- **Cluster 0:** Represents low-value customers with sporadic purchasing behavior. These customers exhibit low frequency and monetary value.
- **Cluster 1:** High-value customers with frequent and high-spending patterns. These customers also show extended engagement over time.
- **Cluster 2:** Casual shoppers with moderate engagement and spending levels.
- **Cluster 3:** Represents customers with specific niche behaviors, such as consistent engagement in particular timeframes or product categories.

Cluster Exploration feature:

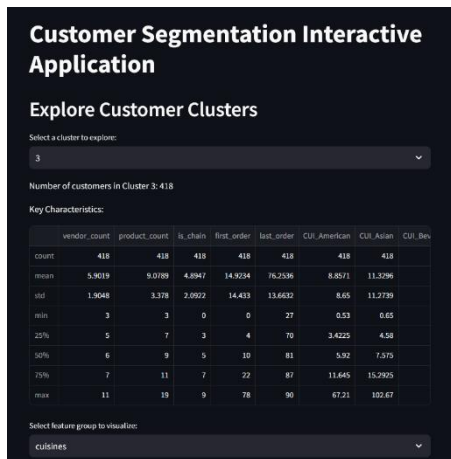


Figure 2- Table of summary statistics



Figure 3- Bar char of average metrics

Filter Clusters Feature:

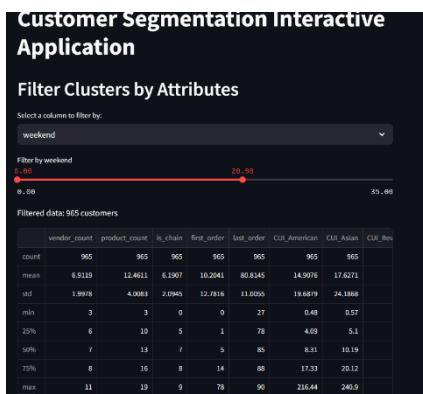


Figure 4- Filtered summary table

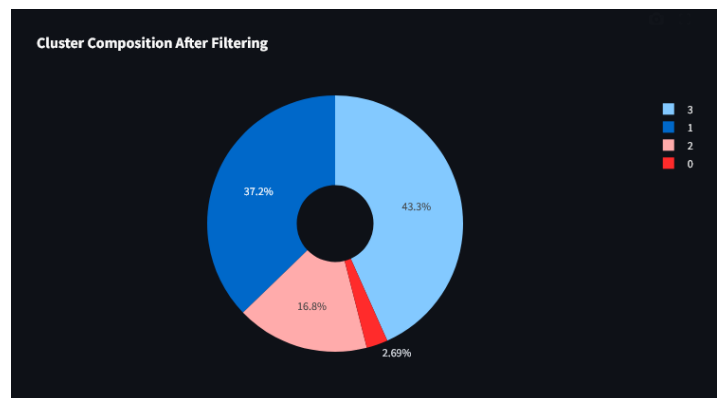


Figure 5- Pie char of cluster composition by feature

Compare Clusters feature:

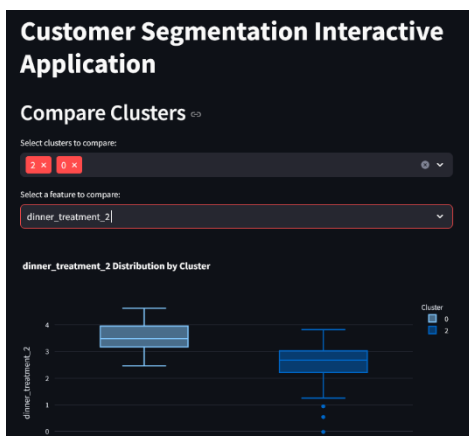


Figure 2- Box plot of different clusters by feature