

# Customer Segmentation FOR TEQUILA SALES ENHANCEMENT

---

By Christian Cusson



# INTRODUCTION



# DISCLAIMER

For this project, I made sure to protect the privacy of Siempre Tequila's customers and the company's intellectual property. I've either removed or made anonymous any personal information in the dataset we analyzed. Additionally, some of the data has been either changed or randomized to ensure further protection of Siempre's IP.

# TECHNOLOGIES USED



## JUPYTER NOTEBOOK

Interactive computing and development environment.

## PYTHON

Main programming language used for backend calculations and data processing.

## SCIKIT-LEARN

Python library used for machine learning models and data preprocessing.

## TABLEAU

Tool for data visualization.

# Data UTILIZATION

---

01

OnOff Premises

02

Classes of Trade

03

Sales Data for  
each SKU (4.5L  
Cases)

04

Total Cases Sold

05

Did Buys

06

# Purchases

07

# Months  
Purchased

08

Rate of Sale: Units  
/ Month

09

Ave Unit Price

# Project STEPS

```
# Columns that need encoding
categorical_features = ['OnOff Premises', 'Classes of Trade']
numerical_features = columns_to_keep[2:]

# Defining a transformer for scaling and encoding
preprocessor = ColumnTransformer(
    transformers=[
        ('num', StandardScaler(), columns_to_keep[2:]),
        ('cat', OneHotEncoder(), categorical_features)
    ])

# Applying the transformations
data_preprocessed = preprocessor.fit_transform(data_behavioral)

optimal_clusters = 5
kmeans = KMeans(n_clusters=optimal_clusters, init='k-means++', max_iter=300, n_init=10, random_state=0)
cluster_labels = kmeans.fit_predict(features)

# Add the cluster labels to the DataFrame
data_preprocessed_df['Cluster'] = cluster_labels
```

## EDA, Data Cleaning & Feature Selection

- Conducted exploratory data analysis to understand data distributions.
- Handled missing values and data inconsistencies.
- Managed outliers to prevent skewing the analysis.
- Chose relevant features for clustering.

## Encoding, Scaling & K-Means Clustering

- Prepared data through encoding and scaling.
- Used StandardScaler() for numerical data and OneHotEncoder() for categorical data.
- Performed K-Means Clustering to segment the customer base into distinct groups.

## Cluster Analysis & Profiles

- Analyzed the cluster means and created detailed profiles based on segment characteristics.



# Cluster PROFILES

0

## Low Volume Clients

These clients typically buy less and often only during specific times like special events or when there are discounts. They're not frequent buyers and usually make smaller purchases compared to others.

3

## Retail and Hospitality Chains

Characterized by extremely high purchase volumes, these clients consistently demand products for business operations and corporate events. Includes large chain corporations, hospitality chains, and institutional clients, each with diverse and specific needs.

1

## Niche Hospitality Venues

Comprised of boutique bars, especially those with a focus on Tequila or Mexican themes, and other specialized venues. This group tends to prefer quality over quantity, buying more regularly than the first group.

2

## Volume Purchasers

Made up of wholesalers, large retail outlets, and organizers of major events who need reliable bulk purchasing options. This cluster shows consistent, high-volume purchasing habits.

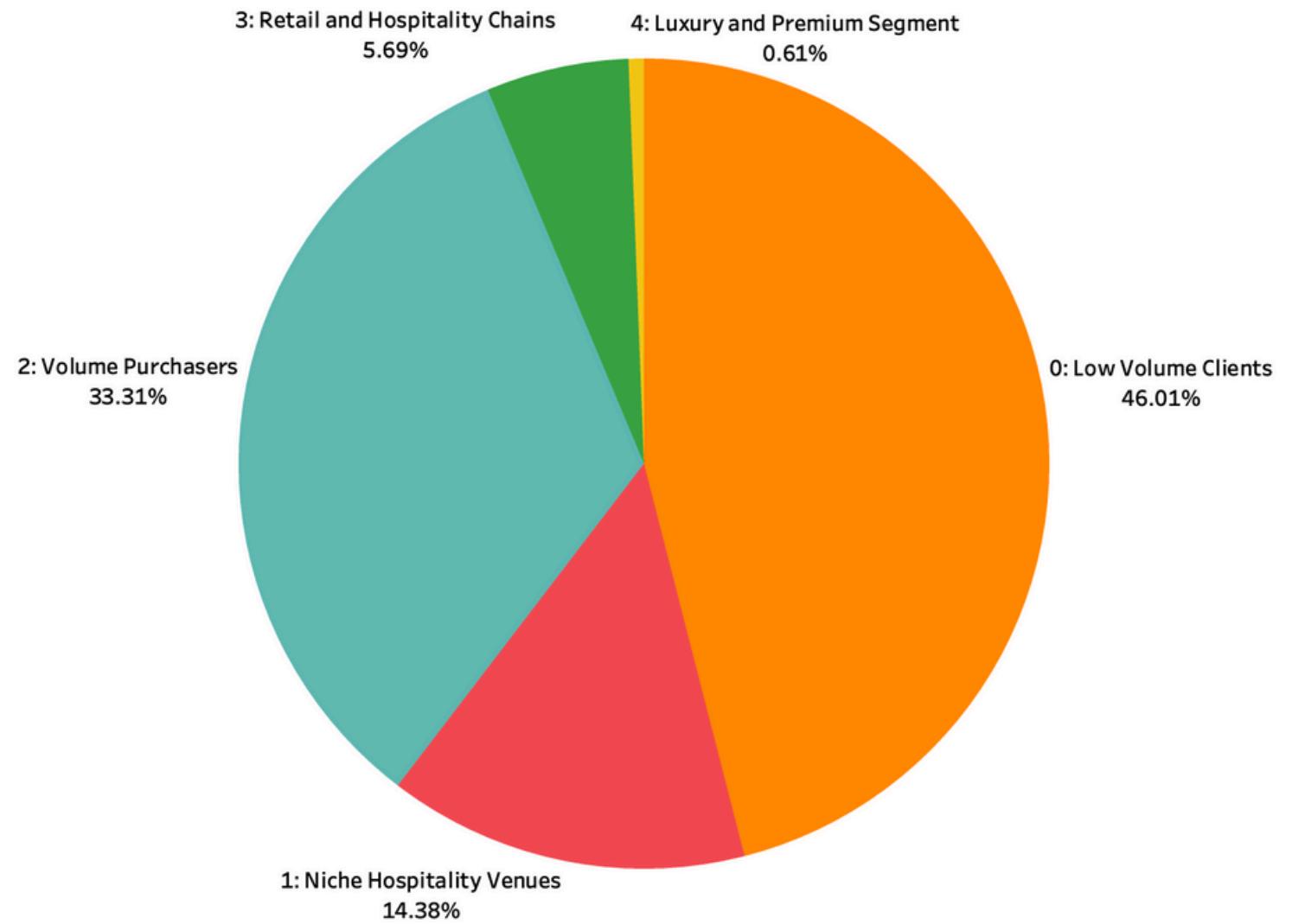
4

## Luxury and Premium Segment

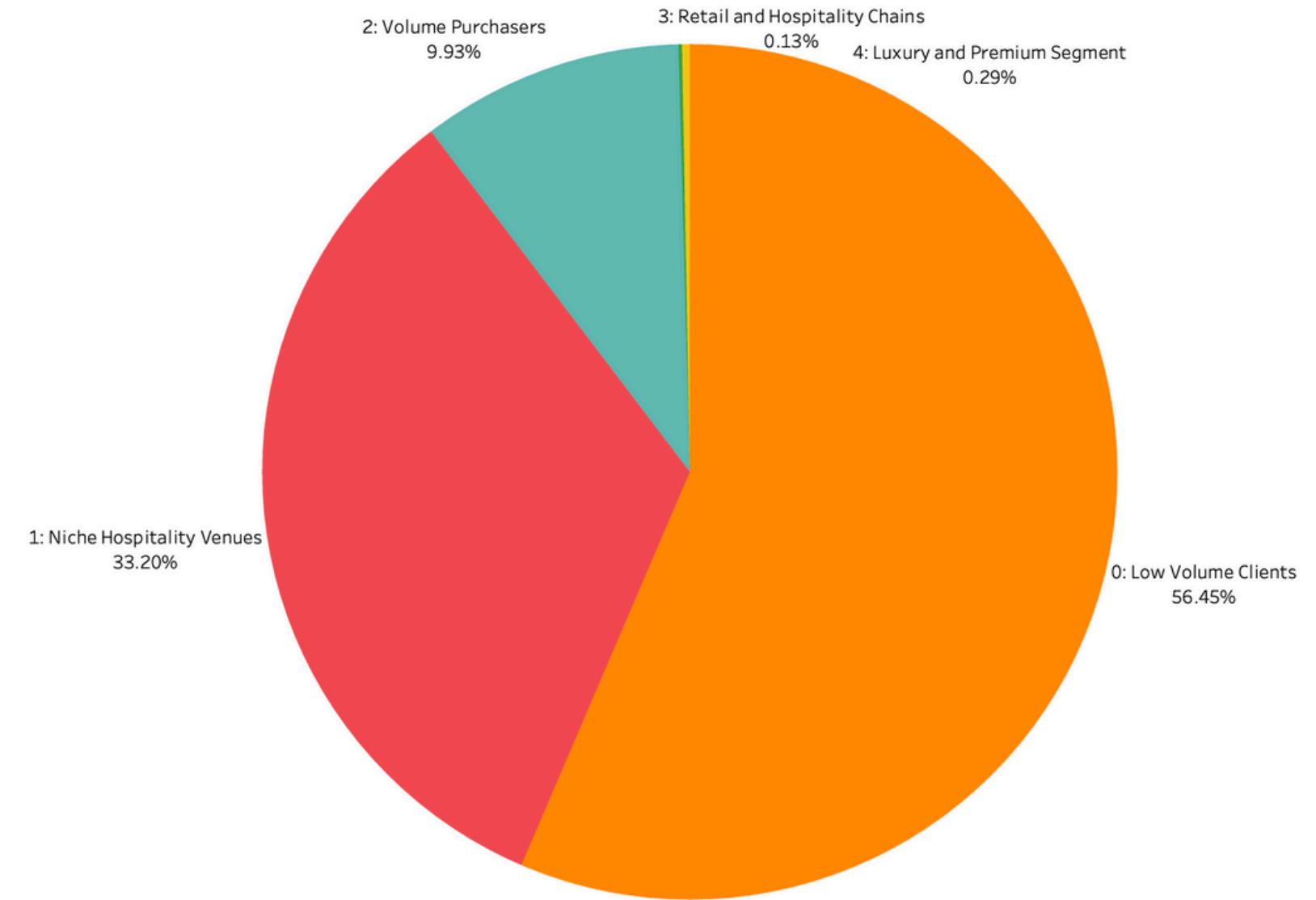
This cluster shows a strong preference for premium products, with a significant focus on quality and exclusivity. Composed of luxury hotels, upscale restaurants, and exclusive clubs.

# KEY INSIGHTS

% of Total Sales by Cluster

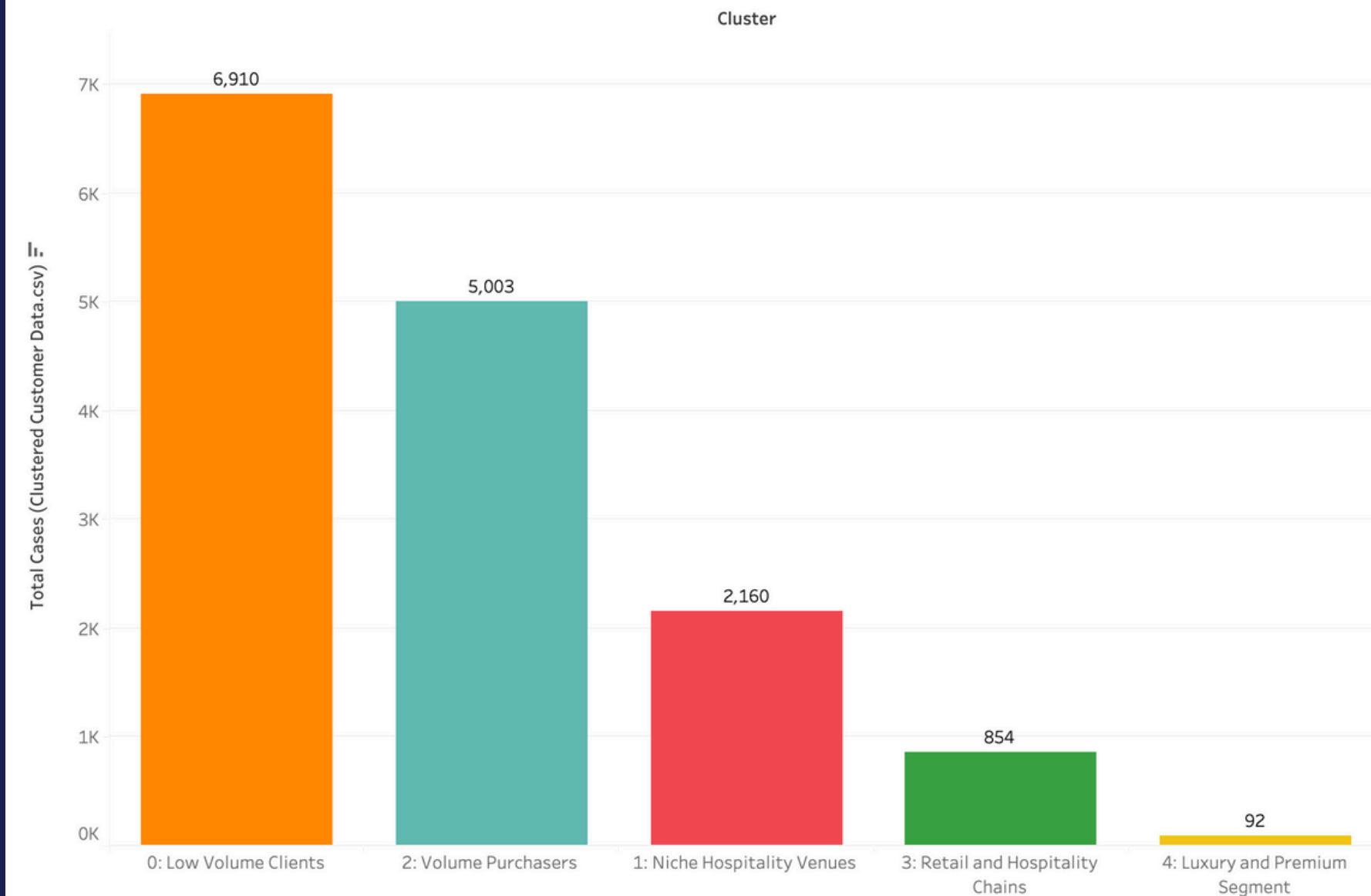


% Of Customers by Cluster

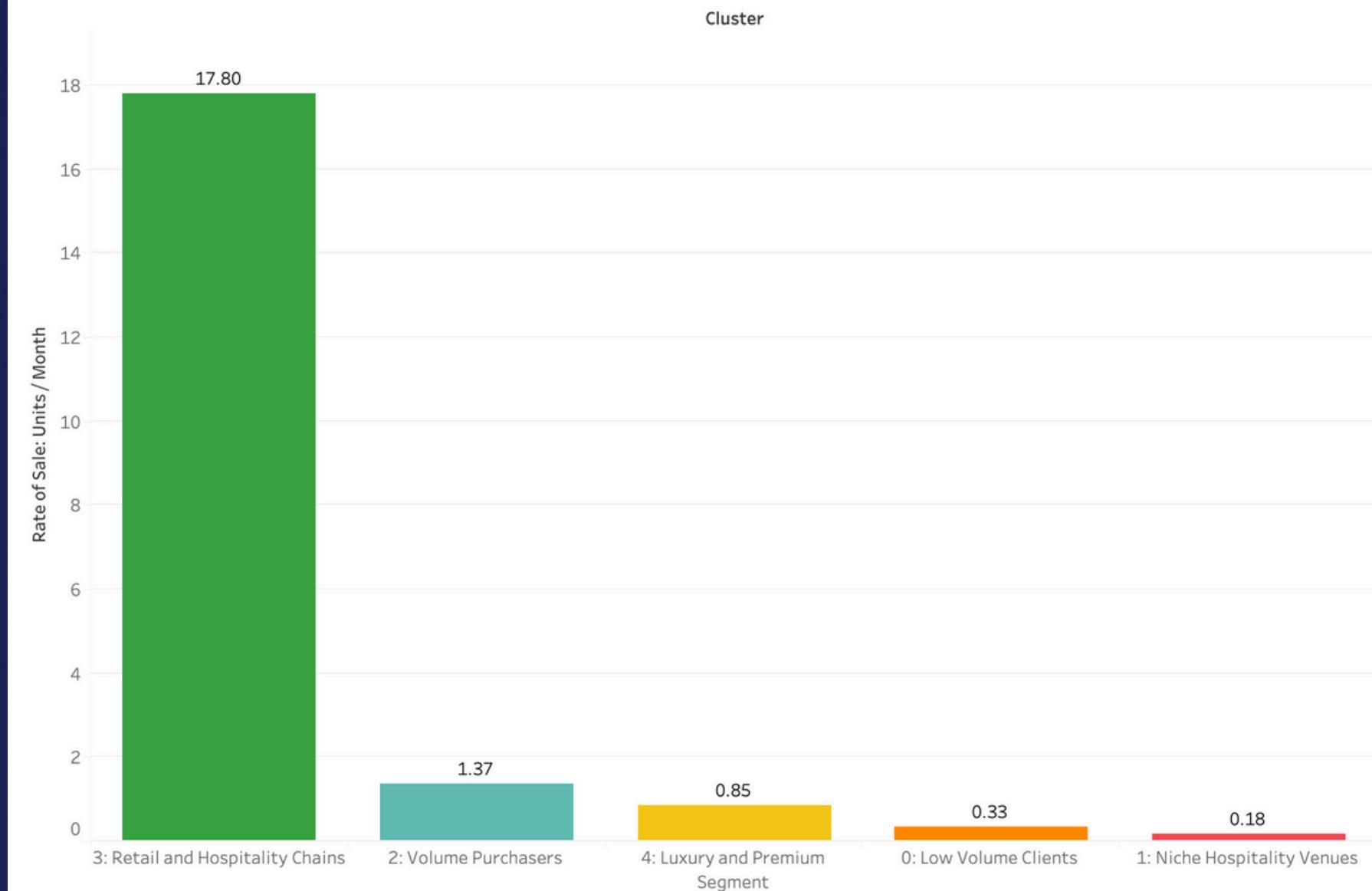


# KEY INSIGHTS

Total Case Sales by Cluster



AVG Rate of Sale by Cluster



## **Challenges FACED**

---

- Data Gathering & Preprocessing Complexity.
- Determining the Optimal Number of Clusters.
- Integrating Domain Knowledge.

## **Future DIRECTION**

---

- Monitor and adapt segmentation strategies as needed.
- Explore additional analytics techniques for deeper insights.
- Implement insights across marketing and sales strategies.

# CHRISTIAN CUSSON

<https://www.linkedin.com/in/christian-cusson-7b982a91/>

[chris.r.cusson@gmail.com](mailto:chris.r.cusson@gmail.com)

+1 819-210-9626

Ottawa, Ontario

