Cluster Characteristics Report Summary

Objective:

The project aims to segment retail stores into meaningful groups based on their sales performance and operational metrics using the K-Means clustering algorithm.

By analyzing various store attributes such as annual sales, customer satisfaction, employee count, and store size, we identify distinct clusters representing different performance levels among retail stores

```
cluster_summary = data.groupby('Cluster').mean()
print(cluster_summary)
```

```
                StoreID  StoreLocation    AnnualSales  NumberOfEmployees  \
Cluster
0         1213.576410       1.939487  679552.986738          21.273846
1         2436.005055       1.620051  478776.417801          28.528222
2          835.953174       1.877211  478605.305498          31.680541

          MonthlyFootTraffic  CustomerSatisfactionScore  \
Cluster
0                5492.493508                   5.678636
1                5575.667986                   5.424987
2                5244.211525                   5.243011

          AverageTransactionValue  OperatingHours   StoreSize  ProductVariety
Cluster
0                      306.485981        0.625641  564.901132      142.318974
1                      254.789598        0.882056  562.195289      103.026116
2                      205.798943        1.142560  529.262970       86.335068
```

Overview of Clusters:

Cluster 0 - High-performing stores:

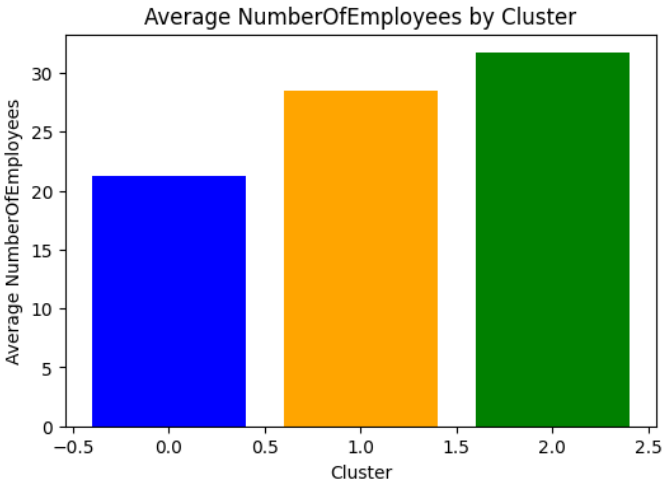
- Highest Annual Sales (~6.8 Lakhs avg)
- Moderate Number of Employees (~21)
- Good Customer Satisfaction (~5.67)
- Highest Average Transaction Value (~306)
- Larger Store Size and Product Variety
- Moderate Operating Hours

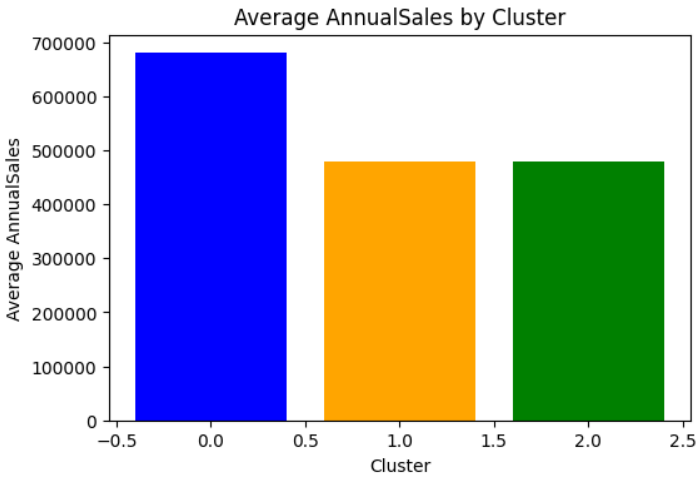
Cluster 1 - Moderate-performing stores:

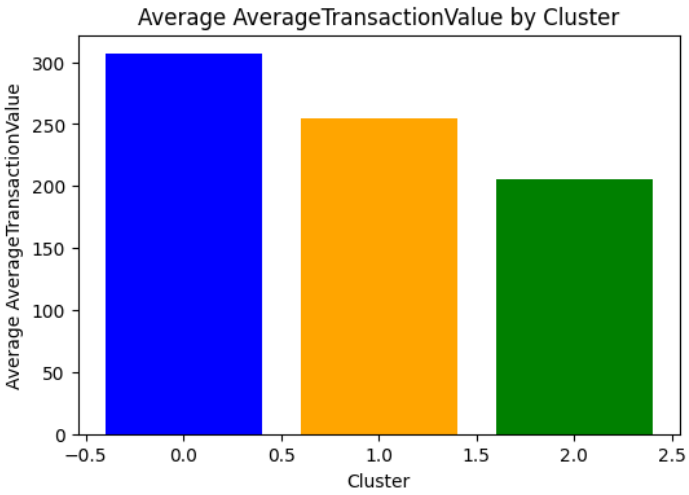
- Medium Annual Sales (~4.7 Lakhs)
- Slightly higher Employee Count (~28)
- Moderate Customer Satisfaction (~5.42)
- Average Transaction Value (~254)
- Balanced Foot Traffic
- Slightly higher Operating Hours

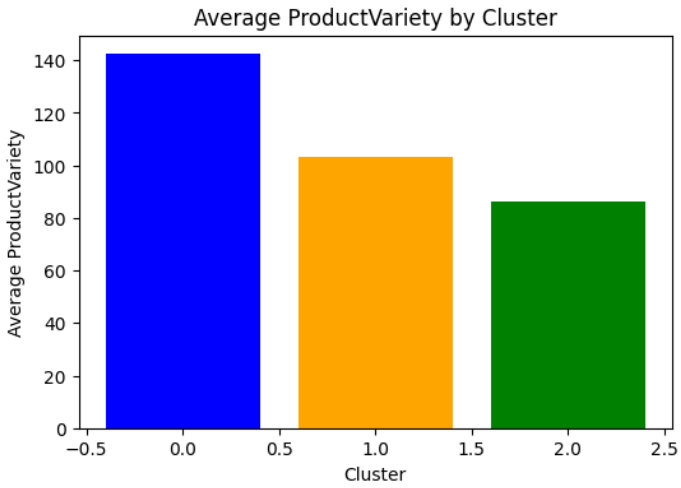
Cluster 2 - Low-performing stores:

- Lowest Annual Sales (~4.7 Lakhs)
- Highest Employee Count (~31)
- Lowest Customer Satisfaction (~5.24)
- Smallest Store Size and Product Variety
- Longest Operating Hours
- Lower Transaction Value









Interpretation:

- Cluster 0 represents top-performing stores that generate strong sales and maintain good customer satisfaction with efficient staffing and balanced operations.
- Cluster 1 consists of average stores that perform moderately across all metrics, showing potential for improvement in sales and satisfaction.
- Cluster 2 contains underperforming stores that, despite higher employee counts and longer operating hours, still show low sales and satisfaction, suggesting inefficiency or operational challenges.

Business Insights:

1. Focus on Cluster 2 to identify inefficiencies and review management, marketing, or staff productivity.

2. Leverage Cluster 0 stores as benchmarks for best practices.

3. Support Cluster 1 with moderate-level sales strategies to push them toward Cluster 0 performance.

4. Optimize store operations and product variety based on patterns observed in Cluster 0.

Conclusion:

The K-Means clustering approach successfully grouped retail stores into three distinct segments based on performance.

This analysis provides actionable insights for management to make data-driven decisions, focusing efforts on low-performing stores and replicating the success of top-performing ones.