

A cartoon illustration of an elderly woman with short, curly hair, wearing glasses and a red headband. She is sitting on a purple sofa, looking at a laptop screen. The background features stylized purple and white clouds and four red plus signs.

# Lola and Lolo's New Superpower: Navigating online shopping with ease

**Presented by: DataBayani**

Hector Banda, Job Chua, Miggy Miguel, Joshua Pecson, Kimberlee Say



In the competitive world of credit cards, understanding customer behavior is key. At **Adobo Advantage Cards (AAC)**, we've collected a comprehensive dataset of credit card transactions for RFM analysis. This dataset, drawn from our data warehouse, includes transaction details, customer demographics, and amounts spent in various categories.

Our challenge to you is to use this data to perform RFM analysis. By segmenting customers, uncover valuable insights, discover spending patterns, and help us make informed decisions. Can you turn this data into actionable strategies that enhance customer satisfaction and drive business growth?

# The Challenge

# Adobo Advantage Cards Dataset

## Demographics

53-97

Age

59

PH Cities

3:50

Female:Male



# Adobo Advantage Cards Dataset

## Transactions

89

Acct. No.

100,000

Transactions

2020-2021

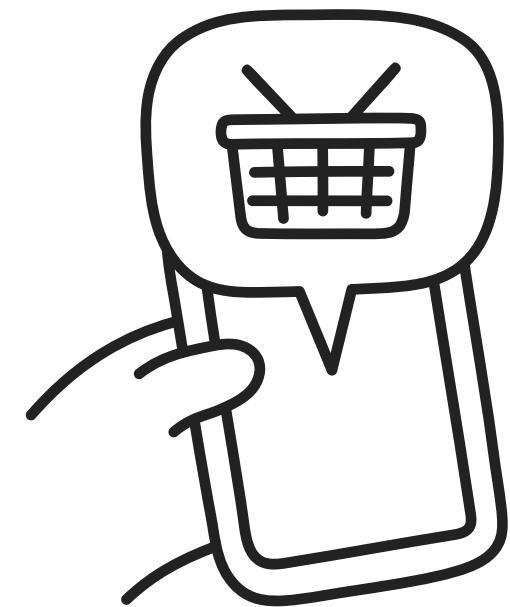
Transaction Date

\$7,029,204

Trans. Amount

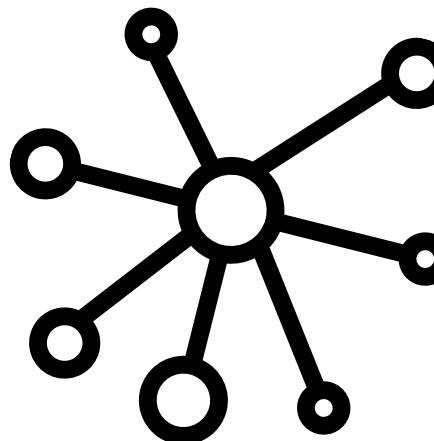


# RESEARCH OBJECTIVES



1

How do online and offline credit card purchases compare in the dataset spanning?



2

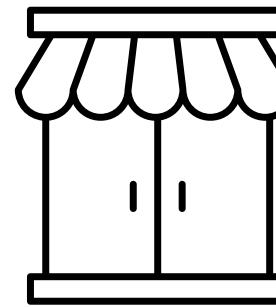
What trends emerge from the transaction data?



3

What strategies can be devised to increase online transactions?

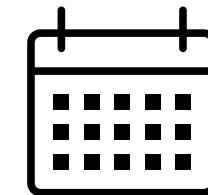
# Scope and Limitations



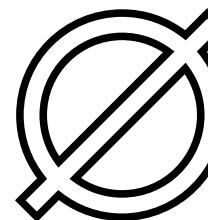
The dataset primarily covers shopping categories with "**Net**" and "**POS**" transactions.



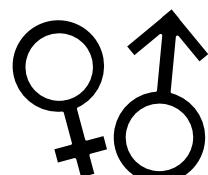
Measurement by number of transactions and amount.



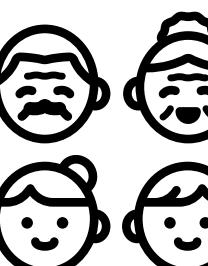
Data spans 2020-2021.



Contains null values for "jobs" and "categories."



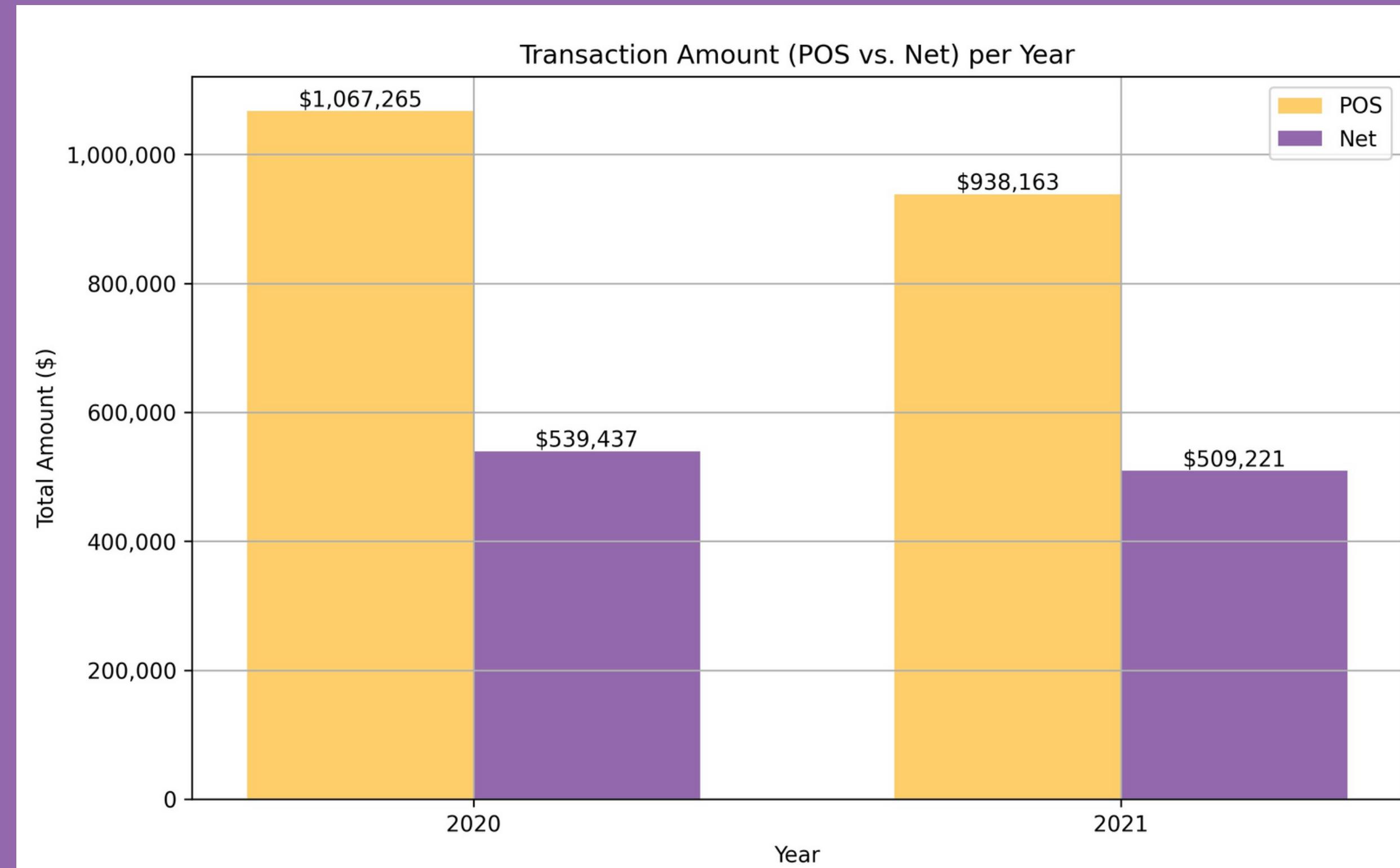
Dataset exhibits a significant gender imbalance with a higher number of male samples.



Dataset is constricted to the elderly population (53-97 years old).

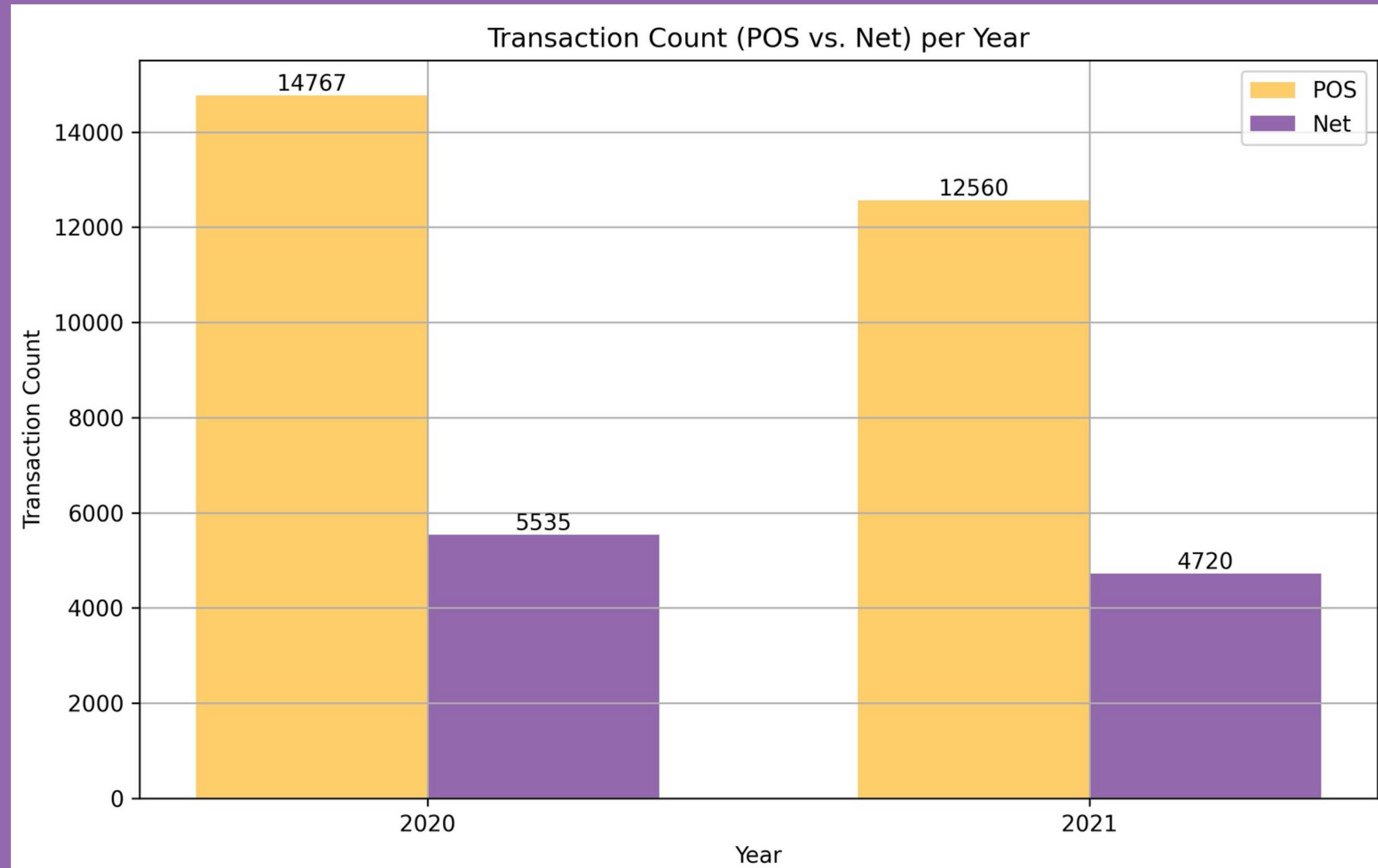
# Comparison of Transaction Amount

- From a high-level perspective, there is a significant difference between the amount that cardholders spend onsite (POS) versus online (Net)



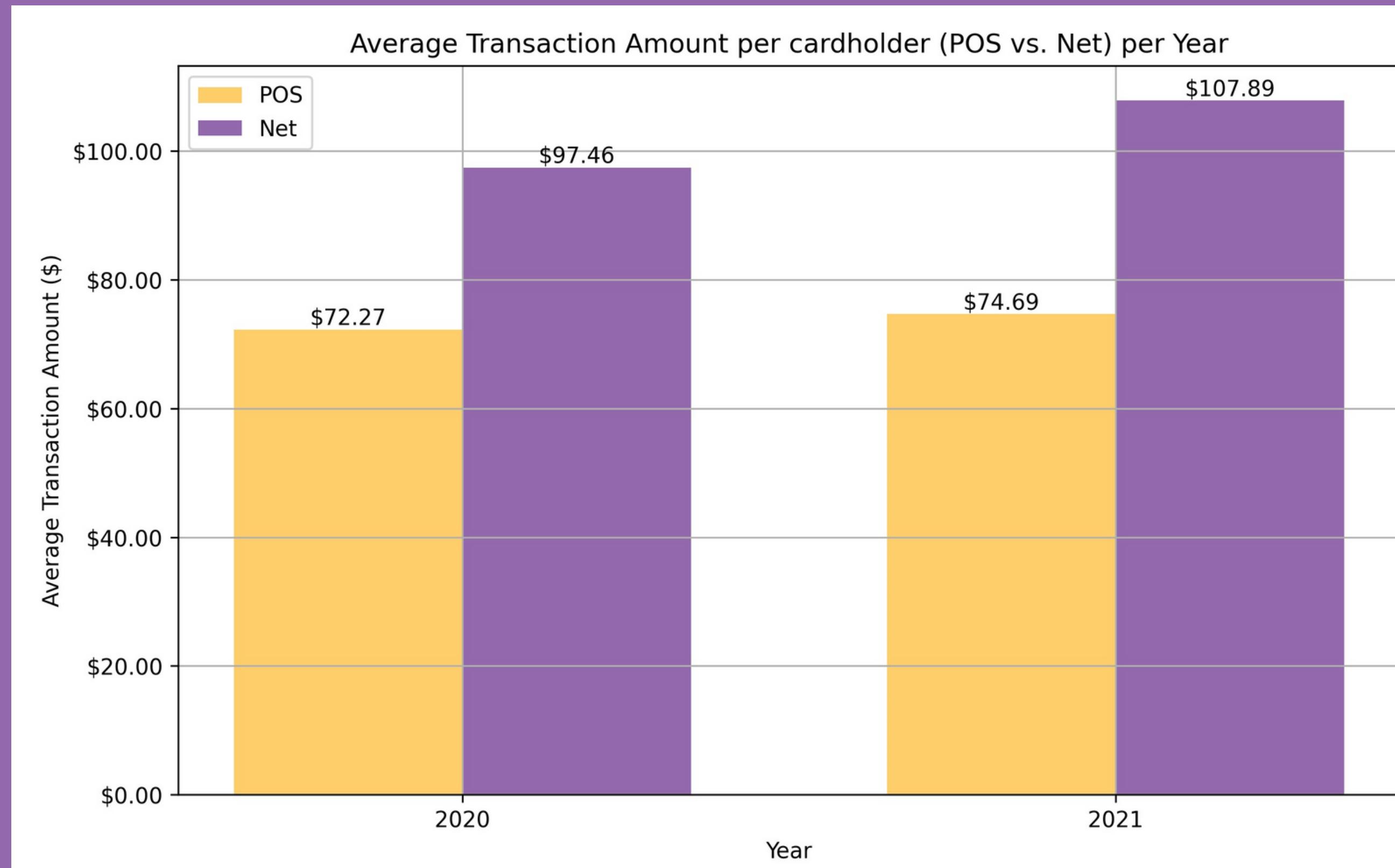
# Comparison of Transaction Count

- Using the transaction count as another metric, there is also a significant difference between onsite (POS) versus online (Net) transactions



# Comparison of Average Transaction Amount

- Cardholders, on average, spend more online (Net) versus on-site (POS)



# METHODOLOGY

**Data  
preprocessing**

data cleaning

feature engineering

**Exploratory  
data analysis**

demographic check

spend analysis

**RFM  
analysis**

RFM score computation

customer categorization

**K-means  
clustering**

data scaling & transformation

cluster optimization

**Model  
interpretation**

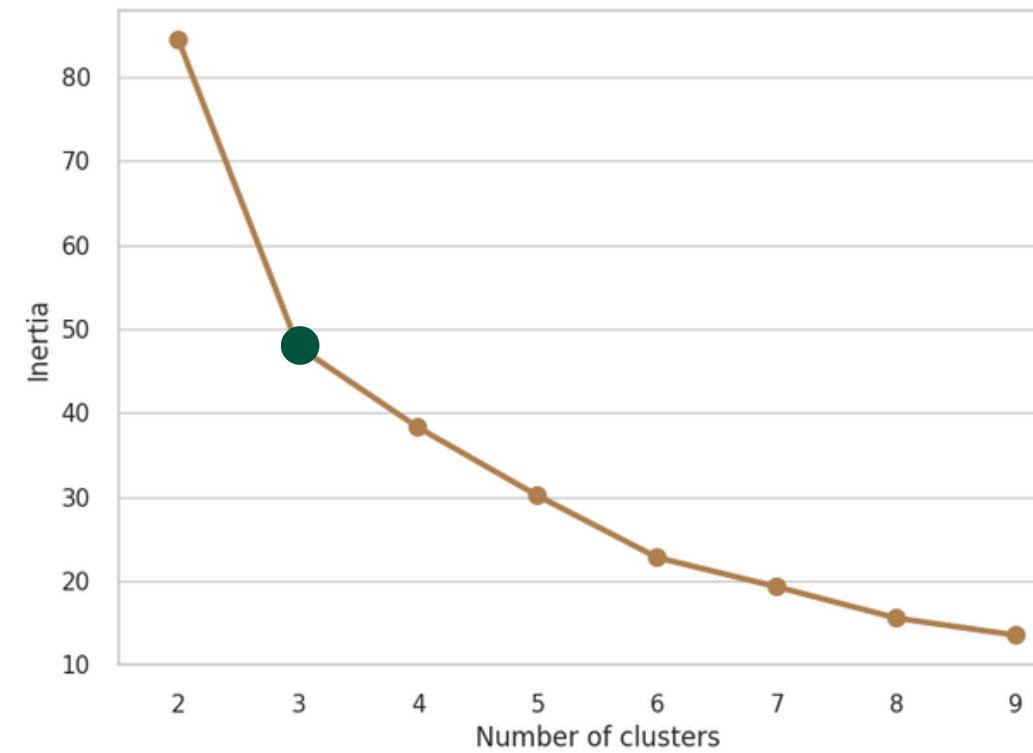
significance testing

demographics deep-dive



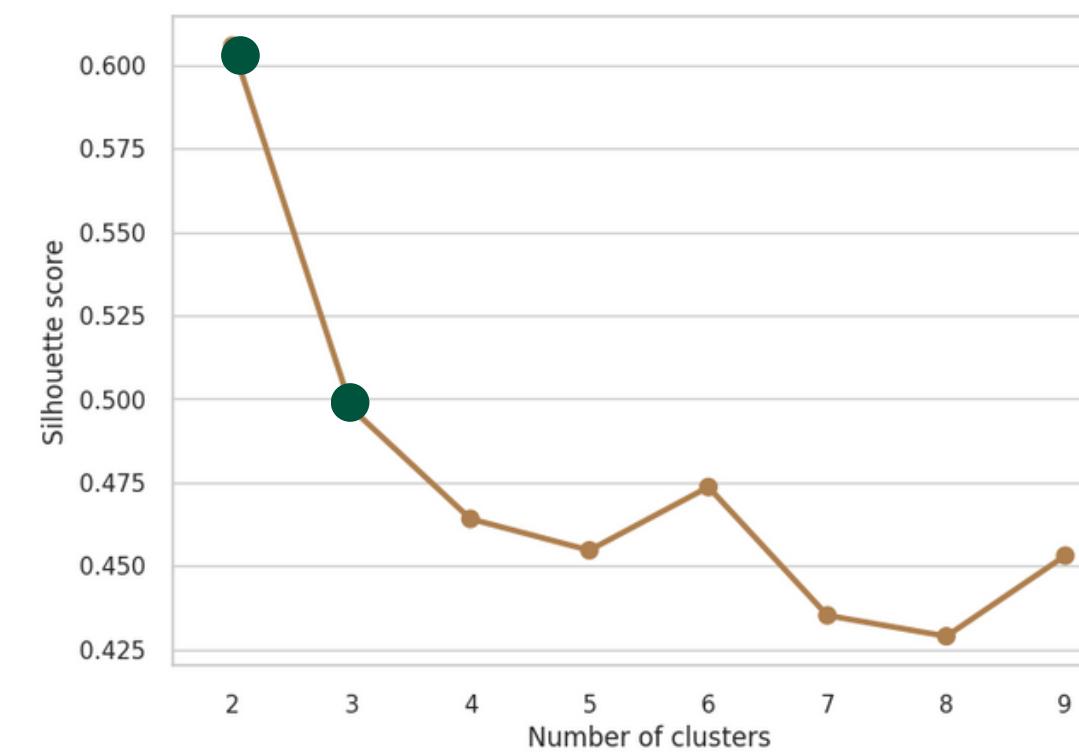
# + Cluster optimization: $k = 3$

**Elbow method**



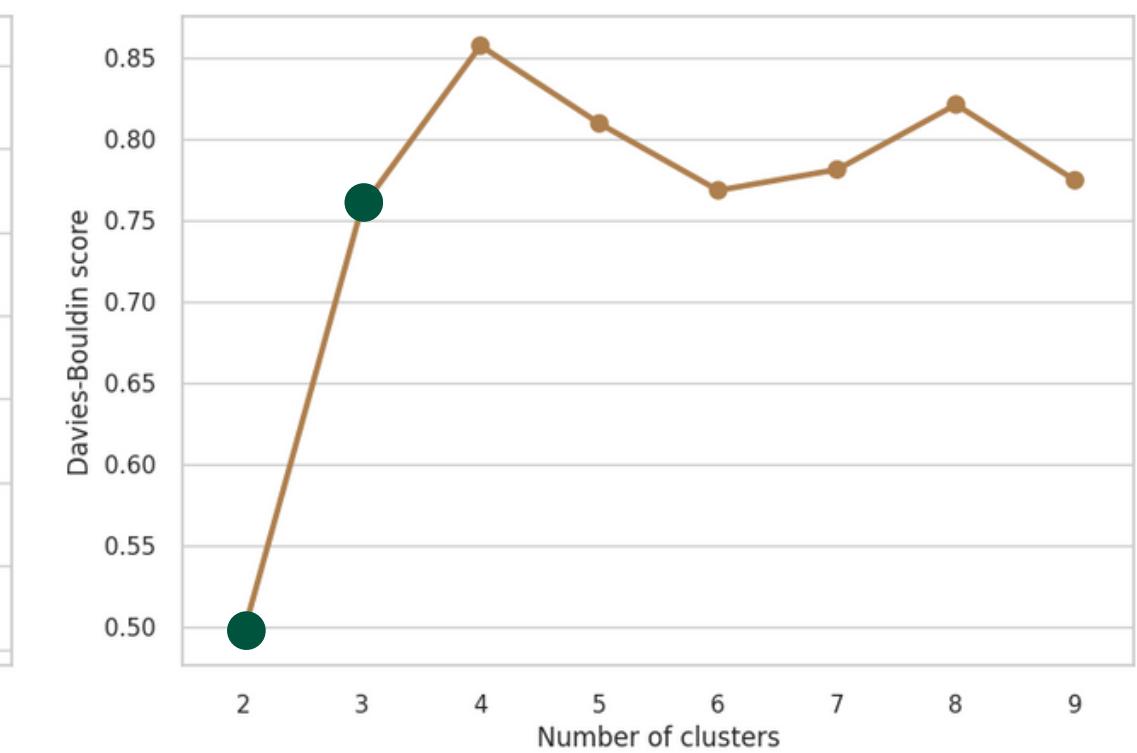
$k = 3$

**Silhouette score**



$k = 2$  or  $k = 3$

**Davies-Bouldin score**

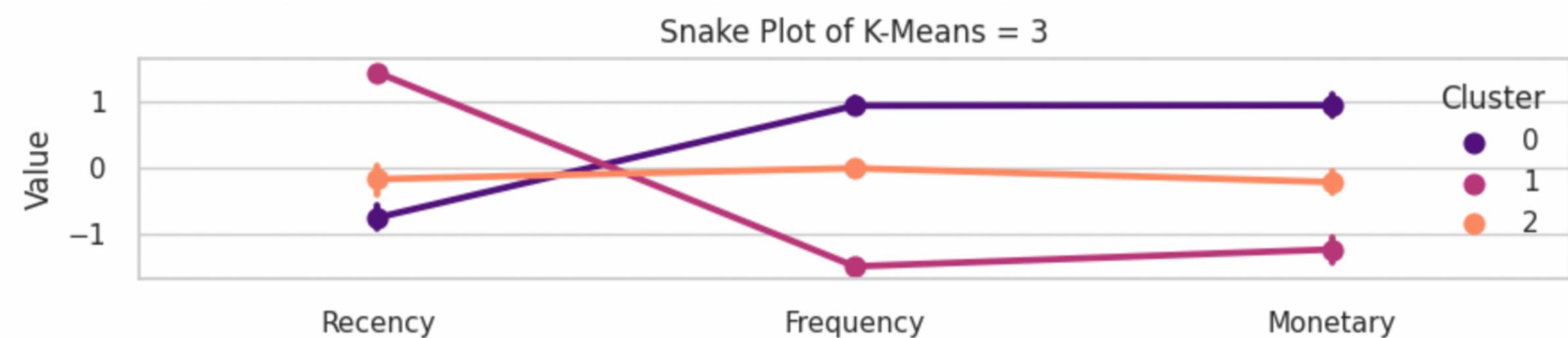
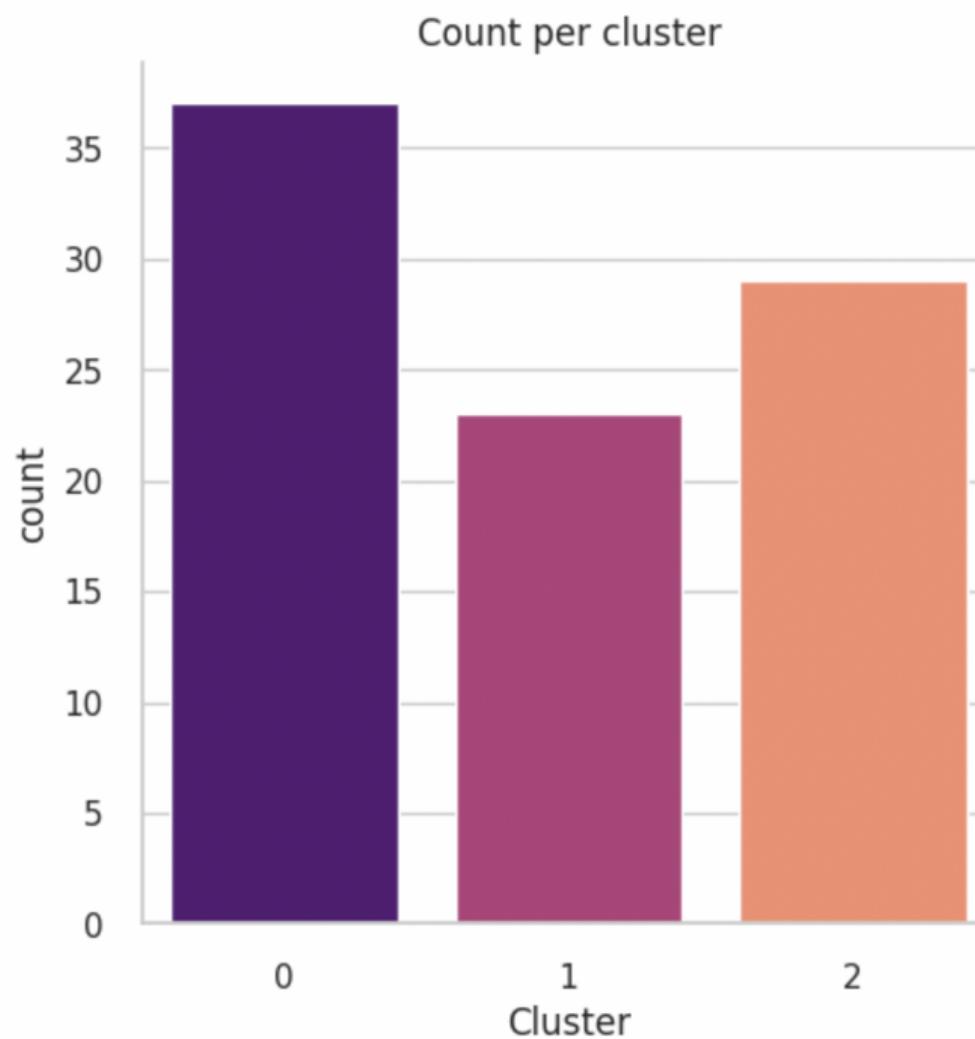


$k = 2$  or  $k = 3$

We select  $k = 3$  as it will allow for more demographic-specific insights, as we will see.

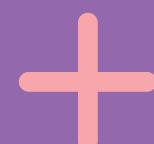
# First look at cluster behavior: $k=3$

The sample is nearly evenly distributed amongst the clusters.



Distinct behaviors between clusters can be observed.

# One-way ANOVA test



## Hypotheses

H0 Means across clusters are equal.

H1 At least two means are not equal.

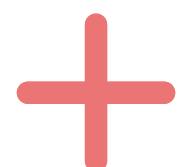
Take  $p < 0.05$

# Results



All metrics are significantly different among the clusters.

	p-value	Conclusion
Recency	6.68e-18	Significantly different
Frequency	1.27e-31	Significantly different
Monetary	1.71e-22	Significantly different



# RESULTS & RECOMMENDATIONS



# CLUSTER

Cluster

Count

Recency

Frequency

Monetary

**Digital Lolo-di**  
Champions

37

⬇️ 1 day

⬆️ 516

⬆️ \$ 37,140

**M.I.A**  
Lost

23

⬆️ 278 days

⬇️ 4

⬇️ \$ 1,696

**Budget Beshies**  
Promising, Price Sensitive

29

⬇️ 3 days

⬆️ 225

⬆️ \$ 16,260

# PERSONA



68  
years old

DIGITAL  
LOLO-DI

Engineer

Affluent

Tech Savvy

Romualdo

Ligaya

M.I.A

Librarian

Online

Non-essentialist

Low-Tech User

69  
years old



Vicente

BUDGET  
BESHIE

Product  
Designer

Comfortable

Moderately  
Techie



68  
years old

## Online shopping level: Expert 💯

Implement a loyalty program and offer retention incentives like NAFFL and low-interest rates to keep them engaged.

+  
+  
+  
**Digital Lolo-dis**



+  
+  
+  
**M.I.A**



Akin ka na lang, pls? 😊👉👈

Build an online community, provide educational content, introduce grocery-focused rewards, and carry out re-targeting and re-activation efforts.

Scroll, shop, repeat. 🛍

Provide dedicated senior support, exclusive discounts/promos, and employ strategies like bundling, cross-selling, and upselling.



# CONCLUSION

## Objective 1

How do online and offline credit card purchases compare in the dataset spanning?

Onsite customers shops more frequently but online customers spends more

## Objective 2

What trends emerge from the transaction data?

Digital Lolo-di



Budget Beskie

## Objective 3

What strategies can be devised to increase online transactions?

Online shopping level: Expert

Scroll, shop, repeat.

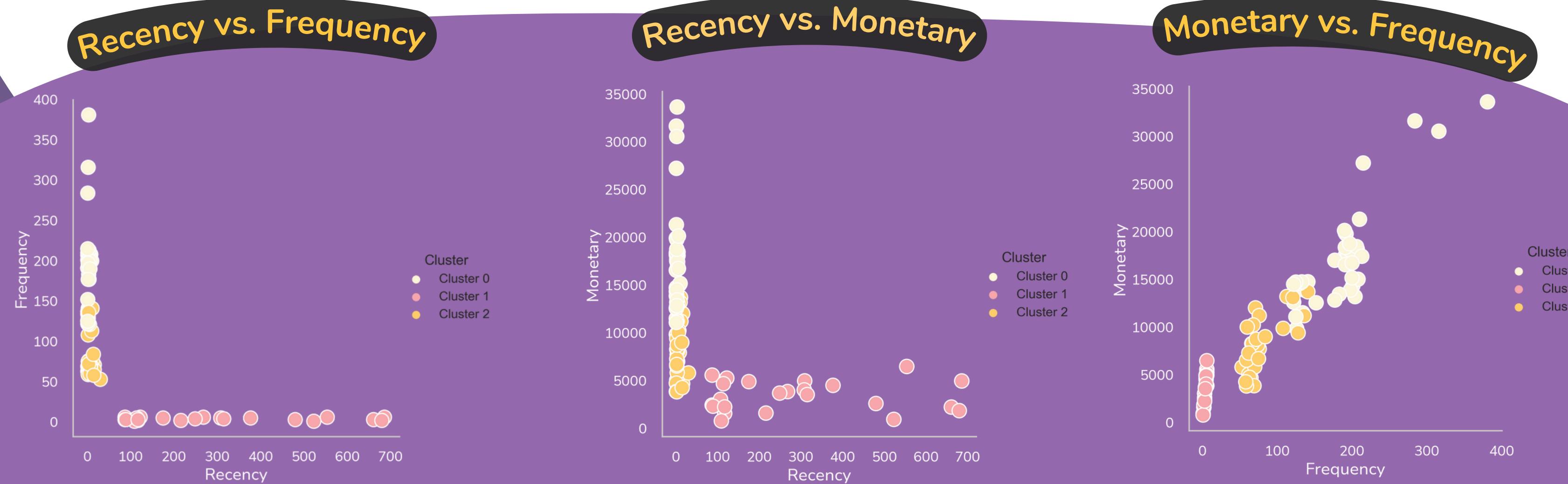
Akin ka na lang, pls?

A stylized illustration of two people, a man and a woman, shaking hands over a purple globe. The man on the left has short, wavy hair and is wearing a black long-sleeved shirt. The woman on the right has curly hair and is wearing a purple top with a white collar. They are both smiling. The globe is purple with white continents and oceans. There are small yellow plus signs at the top and bottom center of the globe.

**Thank you for  
listening!**

Don't hesitate to ask any questions!

# VISUALIZING CLUSTERS

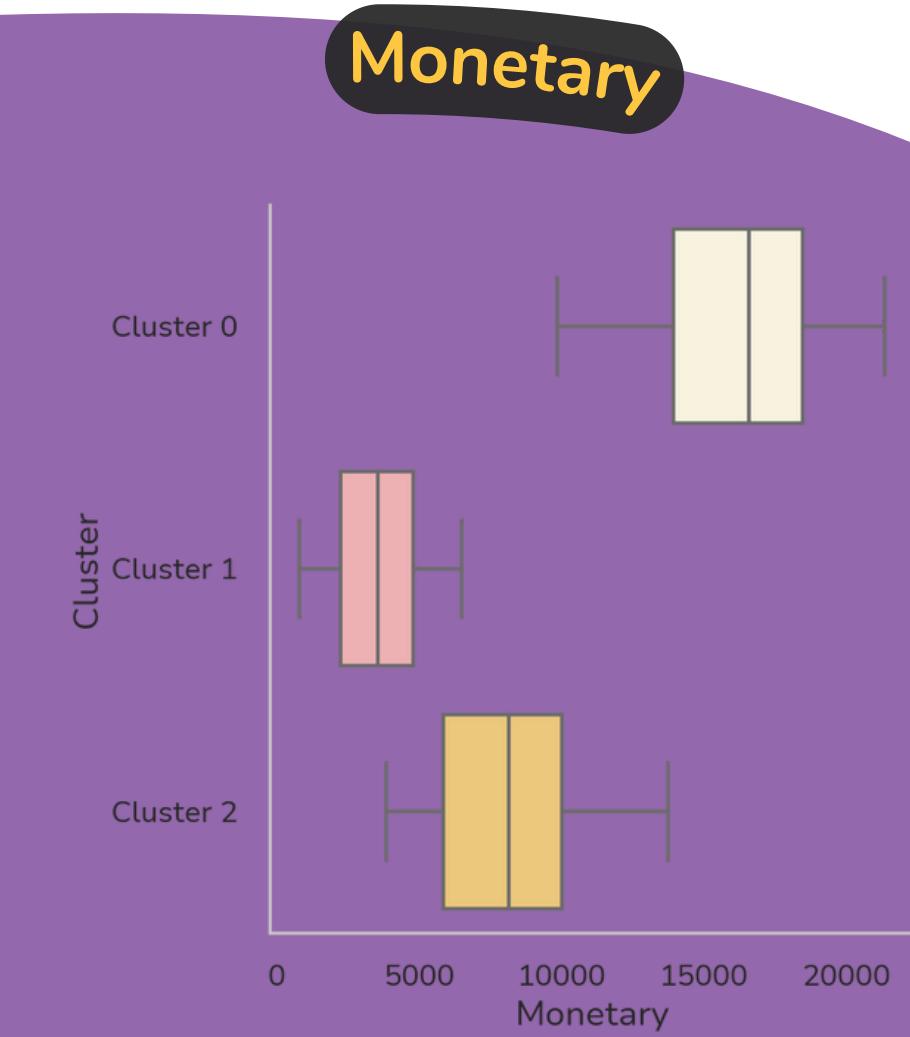
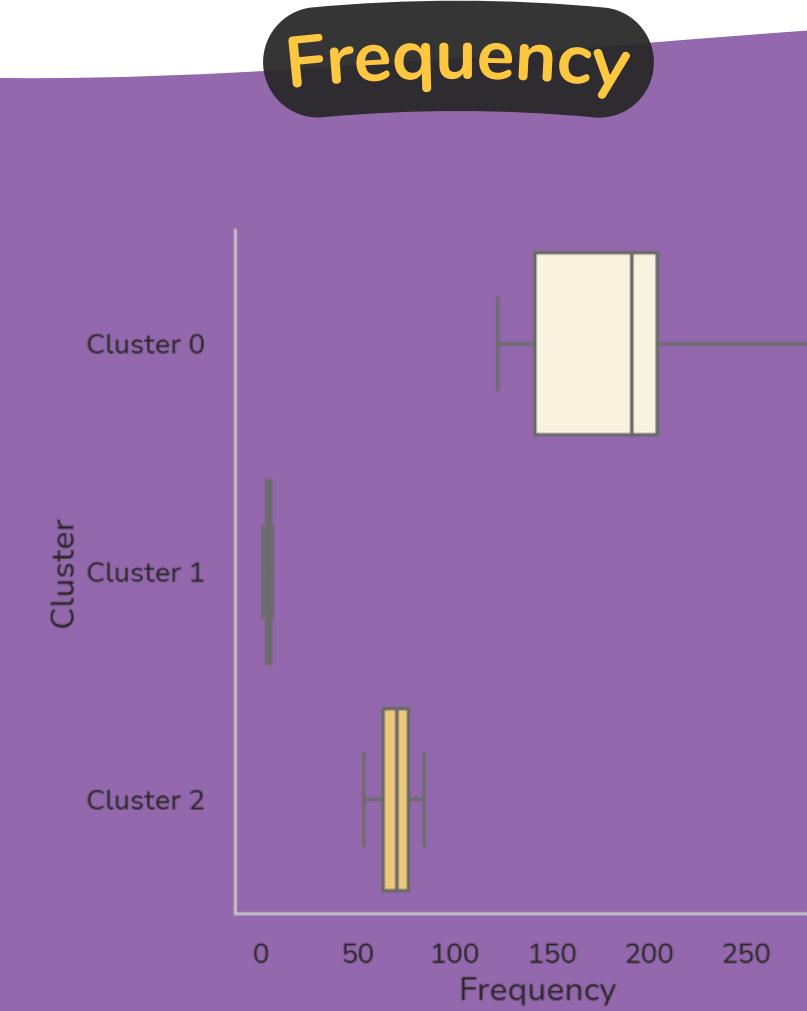
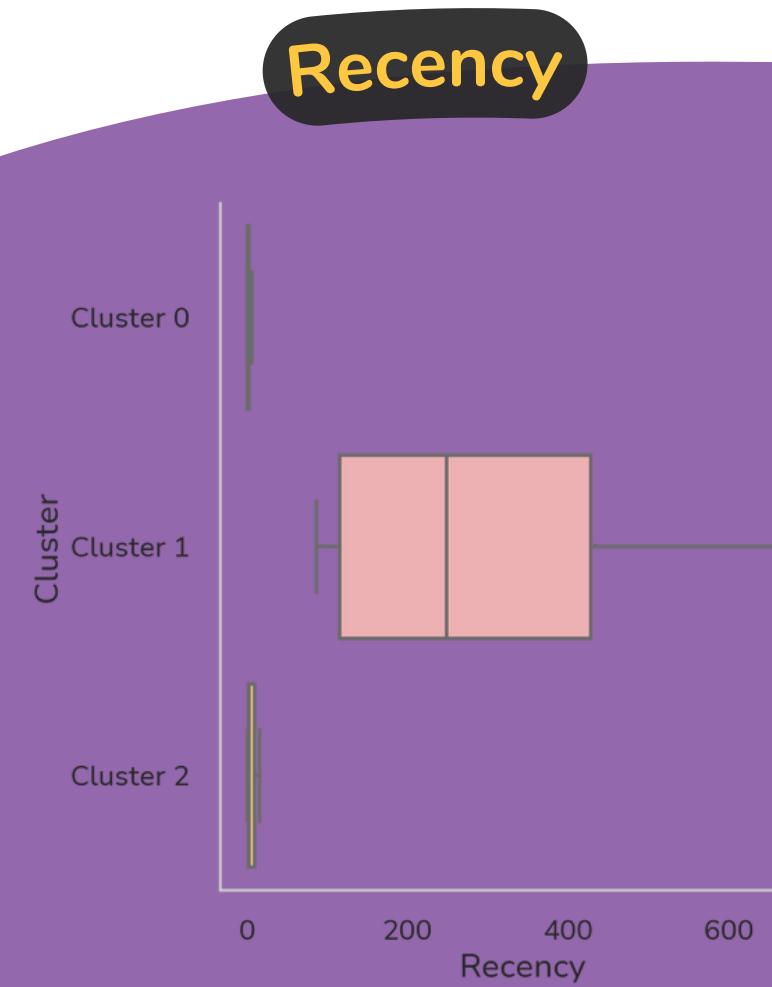


- Both Cluster 0 and 2 have low recency
- Cluster 0 has a higher frequency vs. Cluster 2
- Cluster 1 has higher and more spread out recency and lowest frequency

- Similar behavior as first graph.
- Cluster 0 has higher monetary vs. Cluster 2
- Cluster 1 has the worst performance

- Cluster 0 has the highest frequency and monetary
- Cluster 2 has a lower frequency and monetary vs. Cluster 0
- Cluster 1 has the lowest frequency and monetart

# BOXPLOT



The three clusters are statistically distinct for RFM