OPTIMIZING CUSTOMER EXPERIENCE FOR EVERYTHING PLUS

A Segmentation Analysis of Household E-commerce

Analyst : Team Practicum Batch 1

July 2023

OUTLINES

- Project Description
- Objectives
- Methods
- Findings
- Customer Segmentation
- Recommendations

PROJECT DESCRIPTION

This project focuses on creating **customer segmentation based on the purchase history data of Everything Plus**, an online store specializing in household goods.

By understanding customers' distinct purchasing behaviours and preferences, Everything Plus aims to tailor their marketing strategies, improve customer satisfaction, and drive higher conversion rates.

Objective

Identify the customer's behaviour trend and potential features that can be used to build customer segmentation

Create customer segmentation that can be used to enhance the personalization of offers for different user groups.

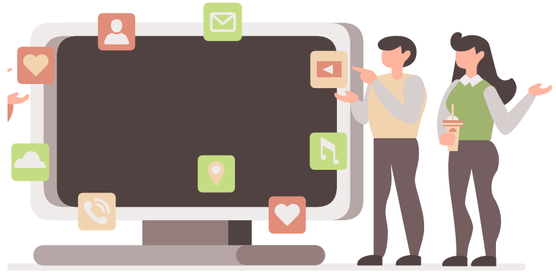
METHOD

Data Preprocessing

- The goal was to have a clean and reliable dataset
- Ensuring its accuracy and consistency
- Involving removing duplicates, handling missing values, and addressing any data quality issues

Exploratory Data Analysis

- Seasonal analysis
- Analyzing customers' spending behaviours



Feature Engineering

- We derived additional meaningful features from the existing data to enhance our analysis

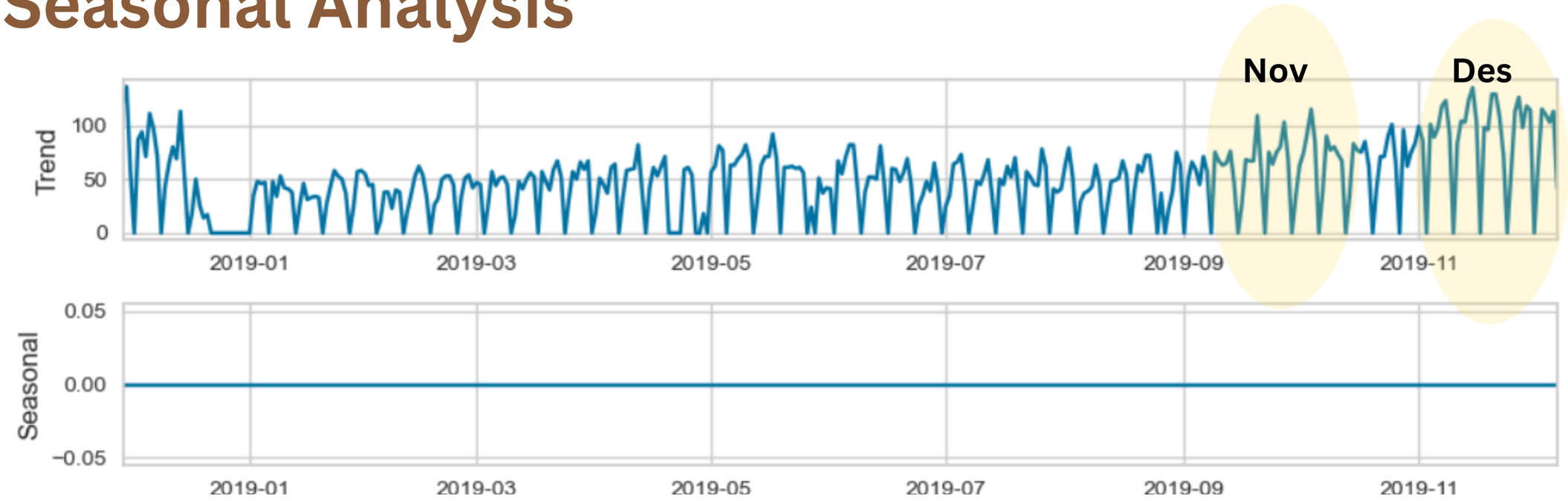
Segmentation

- RFM Segmentation
- Combination of RFM and KMeans Segmentation
- KMeans Segmentation

data exploration

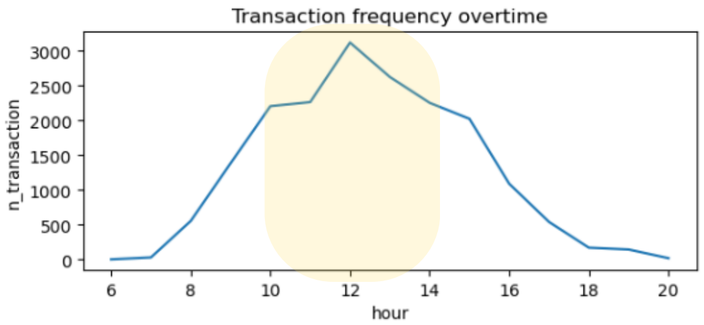
FINDINGS

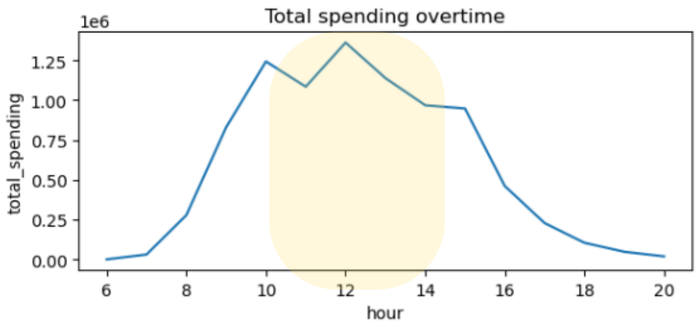
FINDINGS
B Seasonal Analysis

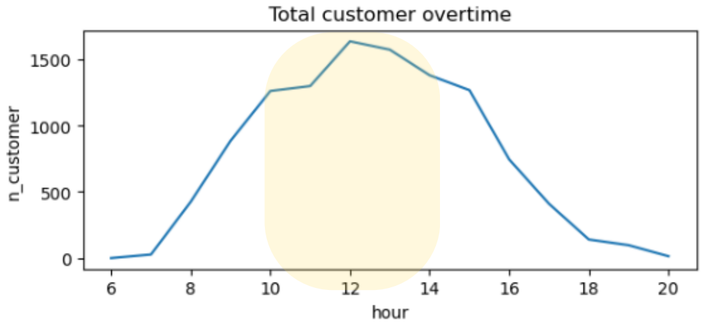


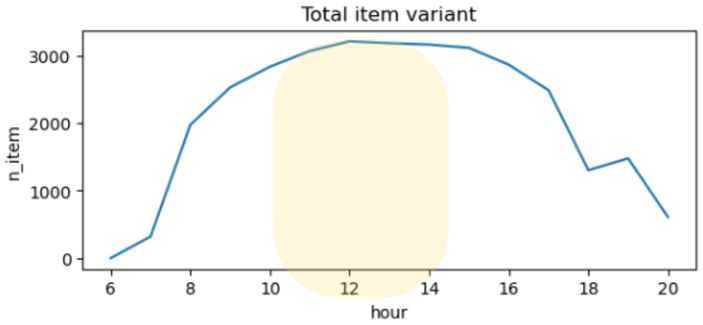
- **Seasonality does not heavily influence customer behaviors**
- **There's only a minor increase in various metrics during the fourth quarter (Nov-Des).** These metrics include *total transactions, unique customers, revenue,* and *the number of purchased item variants*. The increase is observed both on a monthly and daily basis. Seasonality does not heavily influence custom

FINDINGS
B Seasonal Analysis









- **The peak hours** are observed before and after lunch, specifically between **10:00 AM and 2:00 PM.**
- During this time period, customers tend to **engage in more transactions** and **purchase a larger variety of items,** resulting in **higher overall spending.**
- The rush hour showcases heightened customer engagement and presents a valuable opportunity to capture increased sales and revenue.

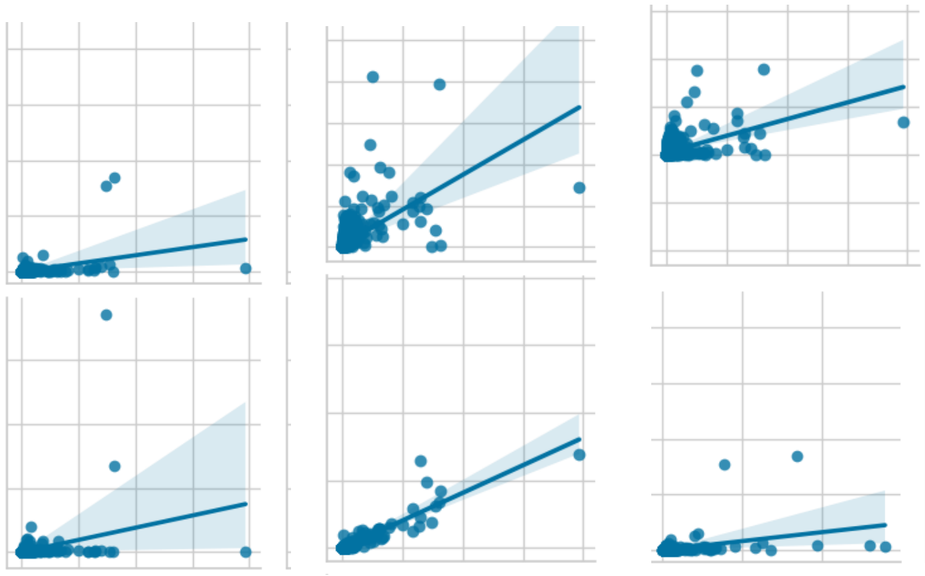
A

FINDINGS
Customers' Spending

If *transaction freq/ revenue/ item variants/ quantity* **increase ...**

then *transaction freq/ revenue/ item variants/ quantity* ***tend to*** **increase**

All variables in the dataset display positive correlations, indicating a consistent trend where an increase in one variable corresponds to an increase in the other variables.



FINDINGS
A Customers' Spending

Customers with **longer purchase recency ...**

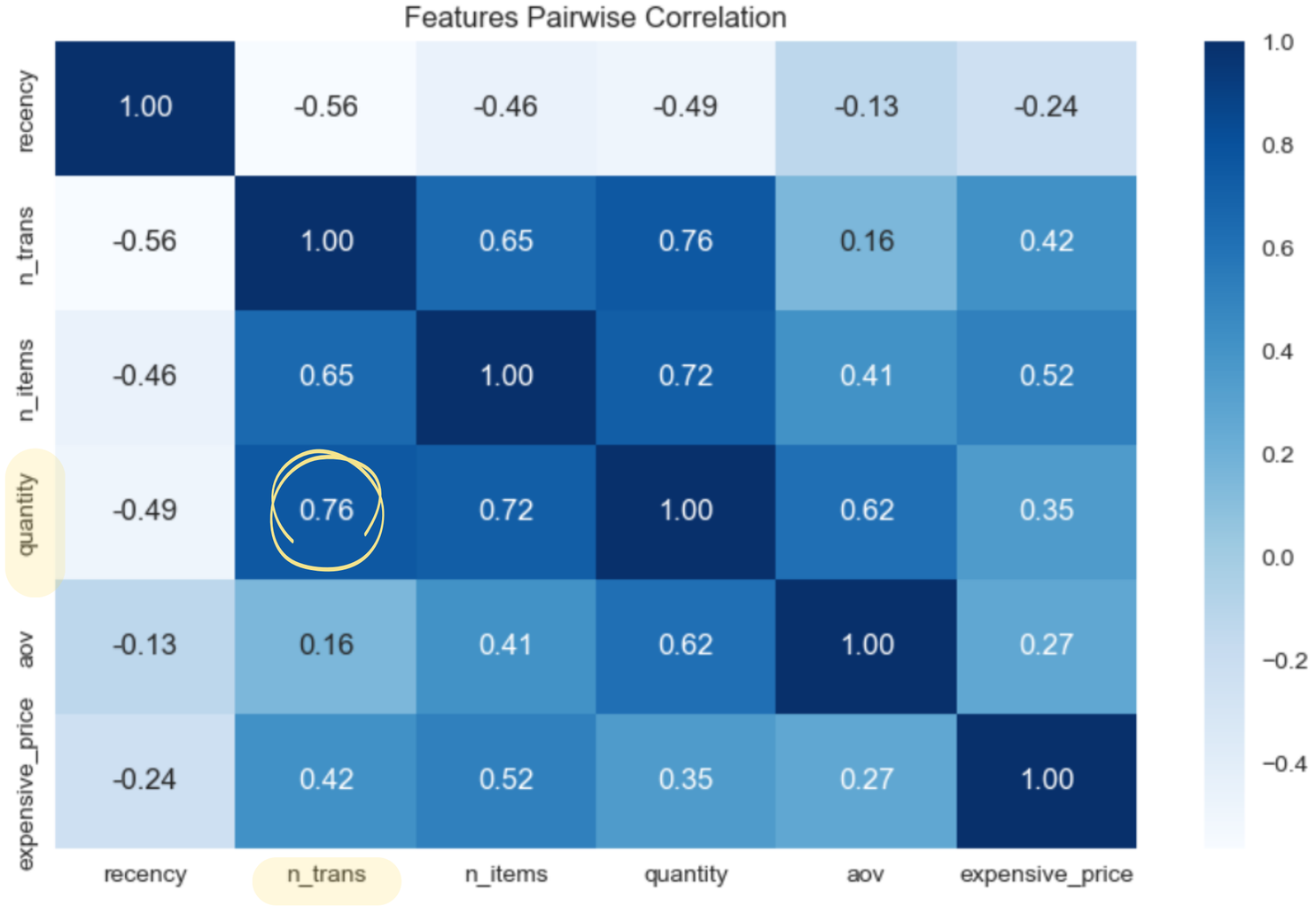
tend to have lower *transaction freq/ revenue/ item variants/ quantity*

- In contrast, the **variable *recency*** stands out with its **negative correlation to all other variables.**
- Customers who have **not purchased recently tend to exhibit lower transaction frequency, spend less, purchase fewer item variants,** and have lower overall engagement with the company.
- Understanding this negative correlation can help identify customers who may require re-engagement strategies to encourage repeat purchases and increase their overall involvement with the business.

FINDINGS
A Customers' Spending

- The closer the absolute value is to 1 or -1, the stronger the relationship between the variables
- A stronger correlation implies that **changes in one variable are more likely to be accompanied by corresponding changes in the other variable.**
- For example:

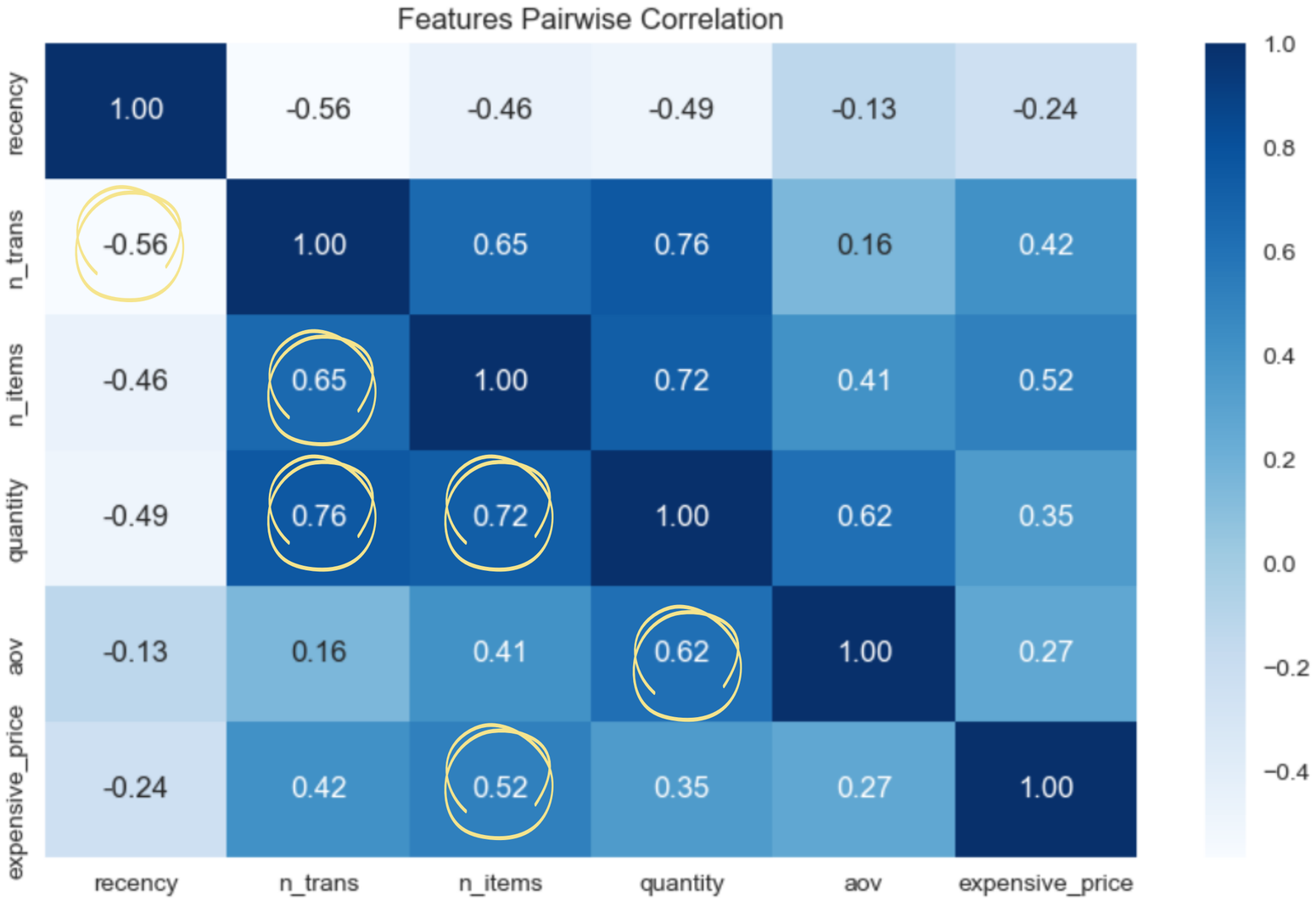
Our data analysis shows a clear trend indicating that ***customers who make frequent transactions also tend to purchase larger quantities.*** *In other words, as the frequency of transactions increases, so does the quantity of items purchased.*



A

FINDINGS

Customers' Spending



Check Other variables with high correlation!

result

CUSTOMER SEGMENTATION

RFM Segmentation

RFM segments the customers based on **Recency, Frequency,** and **Monetary Value**, which are three key factors that help us understand customer behaviour

Recency

the time since a customer's most recent purchase (*days)

Customers who have made recent purchases are often more engaged and responsive to marketing efforts

Frequency

the number of purchases made by a customer within a specific timeframe (*one year)

Customers with higher purchase frequency may be considered more loyal or valuable to the business.

Monetary Value

the total amount of money spent by a customer

Customers with higher monetary value demonstrate higher spending potential or greater investment in the business.

*unit in our analysis

RFM Segmentation

We assigned a scoring system ranging from 1 to 5 for each RFM component.
This scoring approach allows us to quantify and rank the importance of each factor in determining customer segmentation.

Recency Score

	days
1	373 - 263
2	264 - 142
3	143 - 50
4	49 - 17
5	< 17

Frequency Score

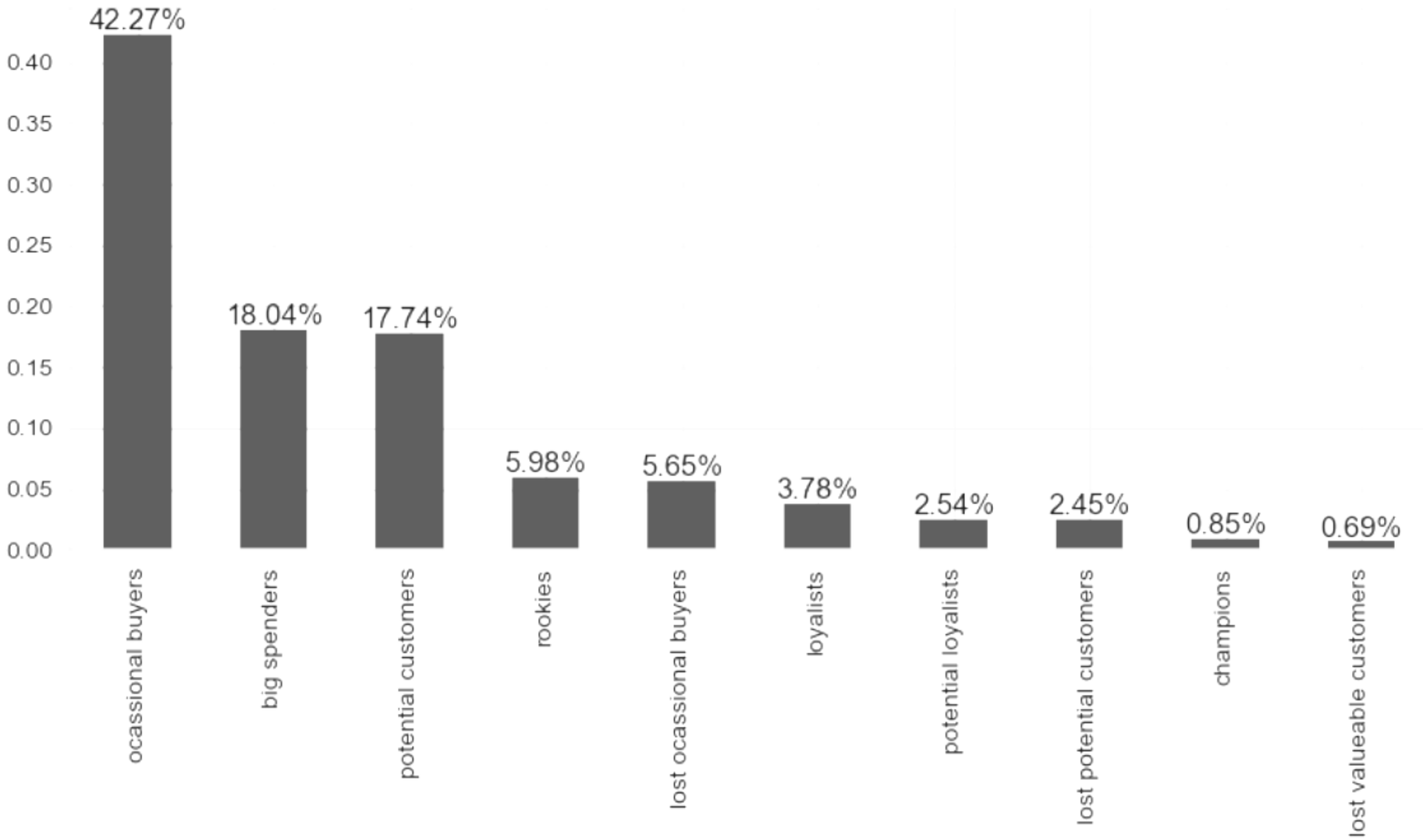
	times
1	1 - 2
2	3 - 5
3	6-13
4	14 - 30
5	> 30

Monetary Score

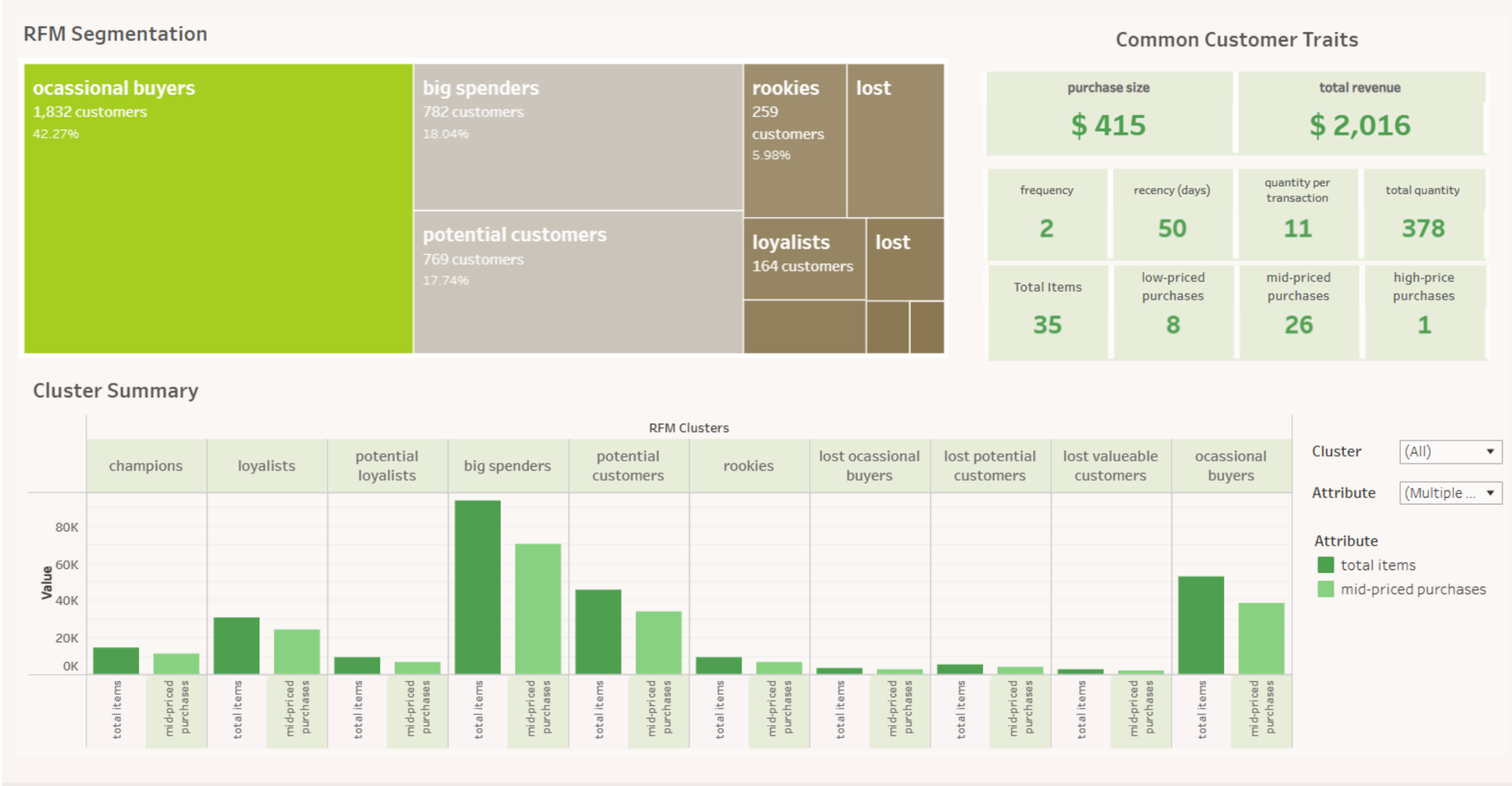
	€
1	3.75 - 304
2	305 - 662
3	143 - 1,631
4	1,632 - 5,735
5	> 5,735

Note: The range values for each score can be modified according to specific needs.

Percentage of Customer Segments



RFM Segmentation

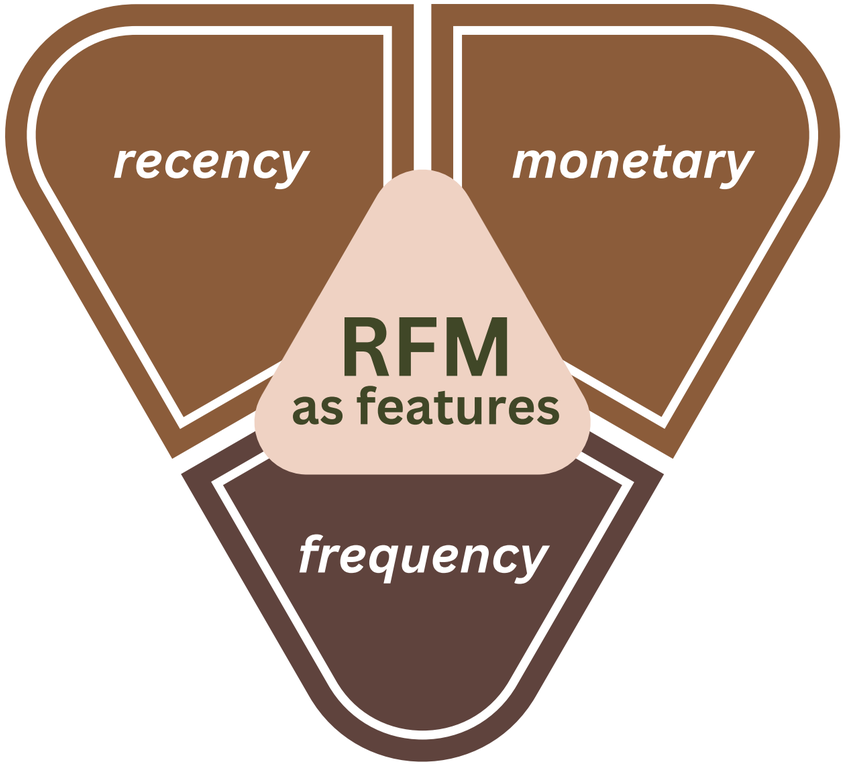


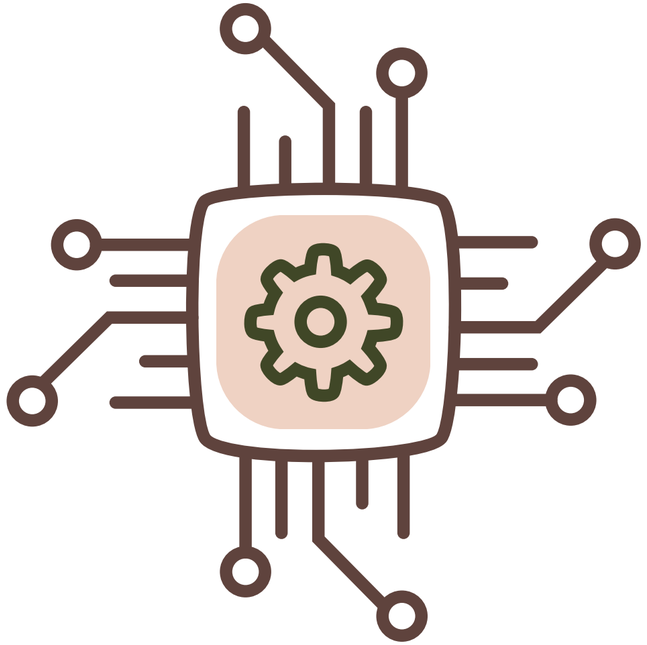
Explore the characteristics of the segments by clicking on the link below!

bit.ly/segmentation_result

RFM + K-Means Segmentation

- RFM + KMeans method combines the RFM analysis with the KMeans clustering algorithm.
- Instead of scoring for *recency, frequency,* and *monetary value,* we utilize standardized values as the main features of the KMeans model.
- This allows us to save time from the feature engineering process





clustering with K-Means algorithm

General Customers
(70.62%)

avg. recency **43 days**
recency **0-163 days**

avg. transaction frequency **3-4**

avg. monetary value **€ 1,343**

Lost Cheap Customers
(24.57%)

avg. recency **248 days**
recency **143-373 days**

avg. transaction frequency **1-2**

avg. monetary value **€ 474**

Potential Customers
(4.50%)

avg. recency **15 days**
recency **0-372 days**

avg. transaction frequency **22**

avg. monetary value **€ 13,470**

Loyalists
(0.16%)

avg. recency **1 day**
recency **0-3 days**

avg. transaction frequency **127**

avg. monetary value **€ 50,117**

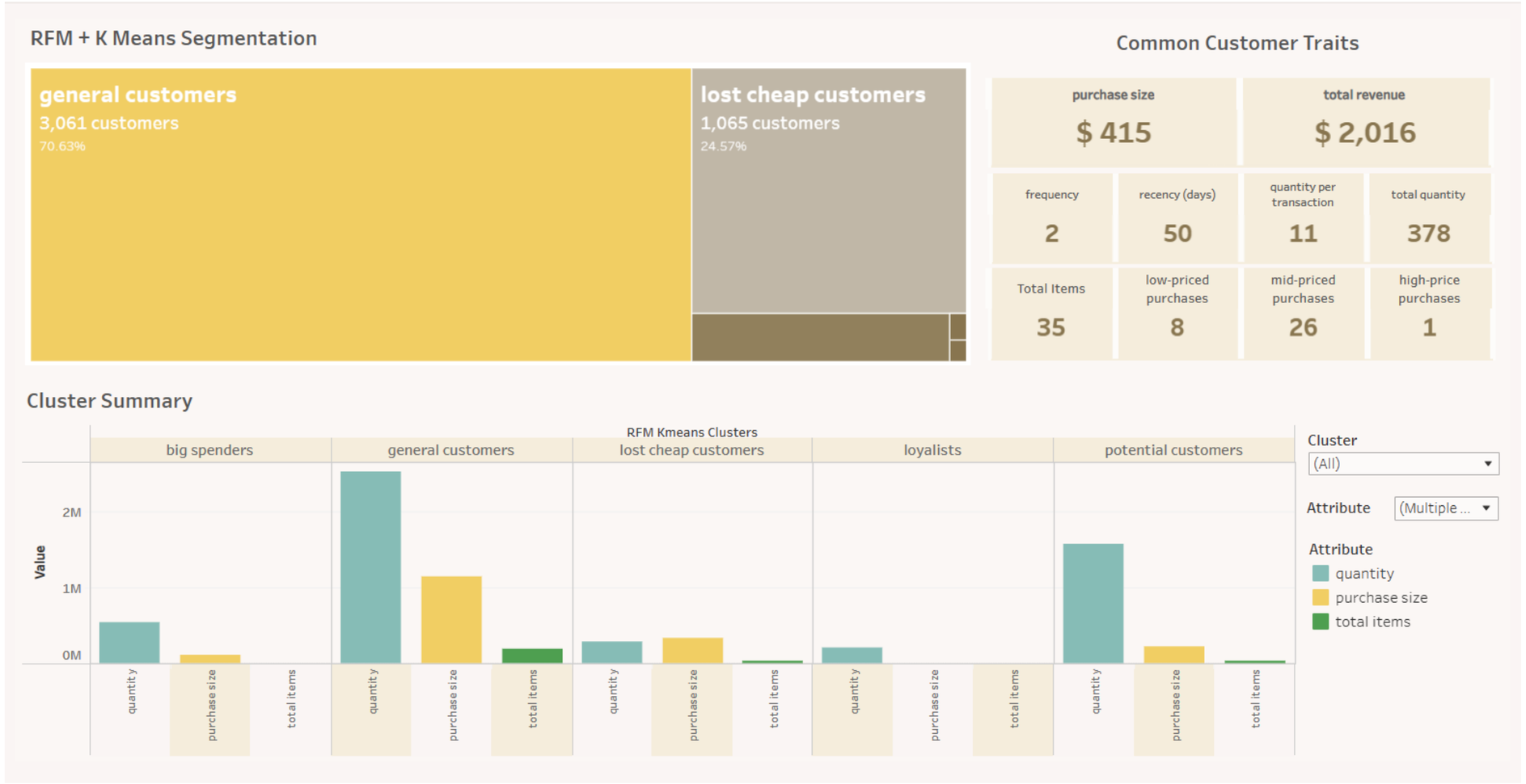
Big Spenders
(0.14%)

avg. recency **7 days**
recency **0-24 days**

avg. transaction frequency **42**

avg. monetary value **€ 190,463**

RFM + KMeans Segmentation



Explore the characteristics of the segments by clicking on the link below!

bit.ly/segmentation_result

KMeans Segmentation

- We applied the KMeans algorithm to cluster customers based on their purchasing behaviour.
- The KMeans segmentation incorporates features engineered from the customer data and utilizes a standardized approach for effective clustering.

General Customers

(62.57%)

- avg. recency **46 days**
- recency **0-169 days**
- avg. frequency **3**
- avg. monetary value **€ 1,082**
- avg. number of item variants **45**
- avg. purchase size **€ 347**
- avg. quantity **685**
- avg. expensive items **1**

Lost Cheap Customers

(24.18%)

- avg. recency **250 days**
- recency **141-373 days**
- avg. frequency **1-2**
- avg. monetary value **€ 543**
- avg. number of item variants **23**
- avg. purchase size **€ 337**
- avg. quantity **276**
- avg. expensive items **0-1**

Potential Customers

(12.85%)

- avg. recency **20 days**
- recency **0-179 days**
- avg. frequency **12**
- avg. monetary value **€ 6,123**
- avg. number of item variants **188**
- avg. purchase size **€ 556**
- avg. quantity **3,529**
- avg. expensive items **5**

Loyalists

(0.35%)

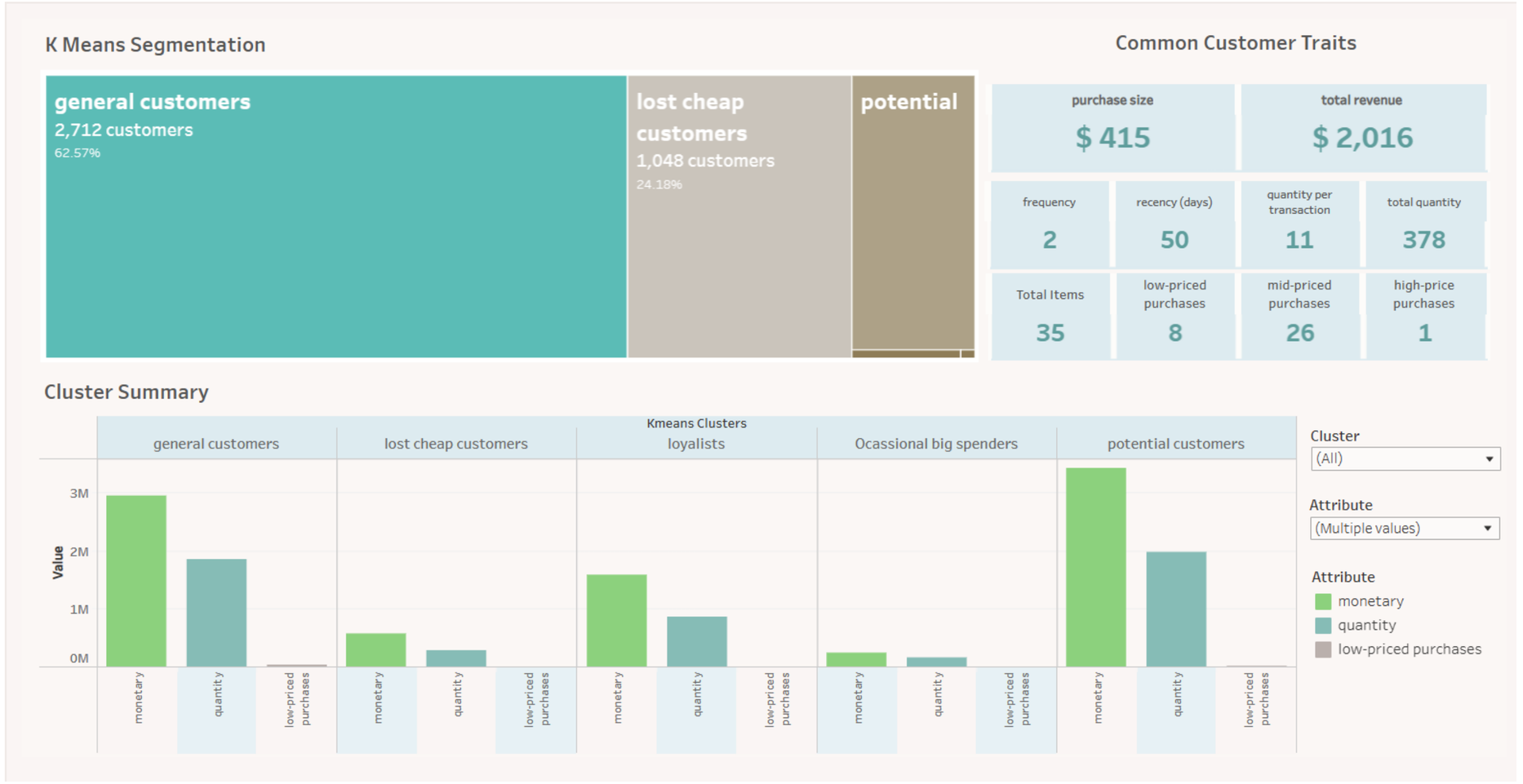
- avg. recency **4 days**
- recency **0-24 days**
- avg. frequency **80**
- avg. monetary value **€ 104,988**
- avg. number of item variants **789**
- avg. purchase size **€ 2,128**
- avg. quantity **58,131**
- avg. expensive items **19**

Occasional Big Spenders

(0.046%)

- avg. recency **162 days**
- recency **0-325 days**
- avg. frequency **1-2**
- avg. monetary value **€ 122,828**
- avg. number of item variants **2**
- avg. purchase size **€ 80,709**
- avg. quantity **77,606**
- avg. expensive items **0**

KMeans Segmentation



Explore the characteristics of the segments by clicking on the link below!

bit.ly/segmentation_result

RECOMMENDATIONS STRATEGY

Recency-based promotions

- Offer them exclusive discounts or special offers to customers who have made a purchase recently but have not made a purchase in a specific time frame

Loyalty rewards

- Create a loyalty program where customers earn points or rewards based on the number of items variant purchased and the recency of their purchases.

Personalized Recommendations

- Leverage the knowledge of customers' past spending and purchase quantities to provide personalized product recommendations

Targeted Promotions

- Send personalized emails/notifications to customers based on their purchase history
- Tailor the content, offers, and recommendations in the emails to make them relevant and engaging to each customer segment.

Limited-time offers

- Offer limited-time discounts or exclusive access to new products to create a sense of urgency and encourage them to make a purchase.

Re-Engagement Campaigns

- Target customers who have not made a purchase for an extended period. Send them targeted campaigns with exclusive offers or rewards to re-engage them and bring them back to your store.

IMPROVEMENT AREA

Considering Special Offers

Any promo/offers/additional fee/postage may significantly affect the customer's preferences to buy a product. It's better to conduct a more advanced analysis of this

Further Analysis for Price Preference

Check whether customers who buy premium/low price will prefer a similar range price for future transaction

ML Clustering

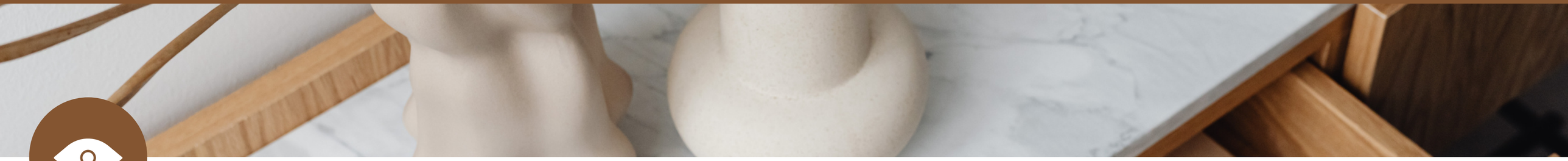
- Evaluate the best features to be passed for customer segmentation. apply standardization
- Try to use other methods to choose the optimum number of clusters (e.g., elbow method, etc
- Iterate and fine-tune

Product Category Analysis

need to be carried out to get more accurate customer segmentation

Customers' Feedback

Collect feedback from customers in different segments to understand more their preferences, satisfaction levels, and areas for improvement.(e.g, rating)



REFERENCES

- Niggl, D. (2022). *Customer Segmentation using K-Means Clustering*. https://pub.towardsai.net/customer-segmentation-using-k-means-clustering-c0c6307ec3f7
- Keenan, M. (2022). *21 Proven Ecommerce Marketing Strategies to Try*. https://www.shopify.com/blog/ecommerce-marketing#16
- *Bilcius, K. The Complete Guide on Behavioral Segmentation in Marketing*. https://www.verfacto.com/blog/customer-segmentation/behavioral-segmentation-marketing/
- mdozmorov.github. (2017). *Chi-square test, Fisher's Exacttest, McNemar's Test*. https://mdozmorov.github.io/BIOS567.2017/presentations/08_FuncEnrich/08b_enrichment_stats.pdf
- Meyer-Waarden, Lars; Benavent, Christophe; Castéran, Herbert. (2013). *The effects of purchase orientations on perceived loyalty programmes' benefits and loyalty*. **International Journal of Retail & Distribution Management**, Vol.41.Issue.3: 201-225. http://search.proquest.com/docview/1315333092? accountid=34643.

MEET THE TEAM

Anand Pandey
Team Leader

Ica Candra
Junior Analyst

Jonathan Chan
Junior Analyst

Zamani
Junior Analyst

THANK YOU

Team Practicum Batch 1
July 2023