

...



...



...

E-Commerce RFM Clustering

Satryo Sunu Prabowo

Table of Contents

01

...

Background

Background of the analysis and reason we started working on this project

02

...

Initial Analysis

Relevant findings from exploratory data analysis

03

...

Cluster Analysis

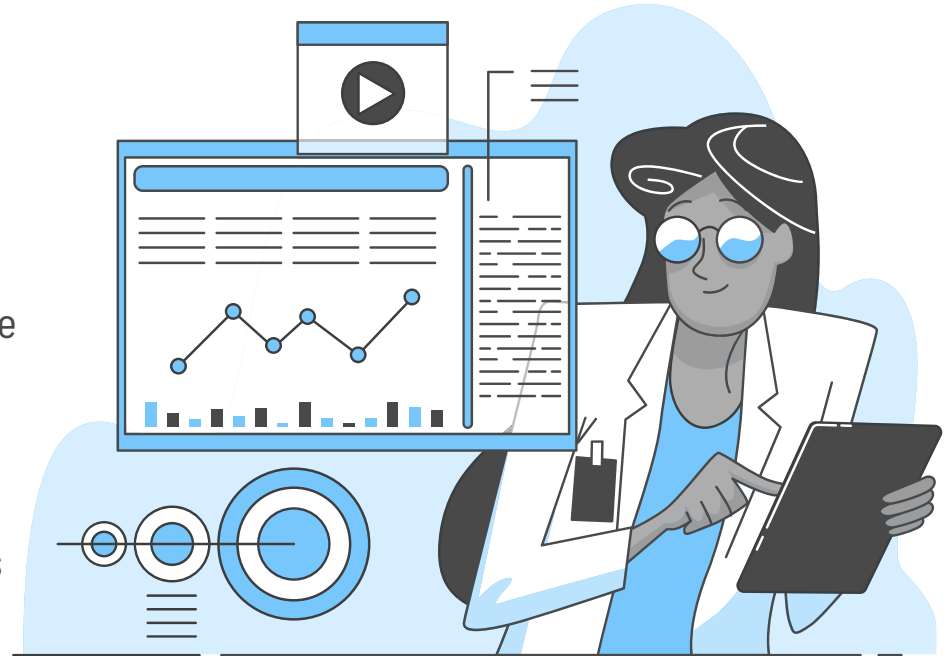
Identify cluster and determine the cluster based on the user behavior

04

...

Recommendations

Provides actions for business to take





01

Background



Company Background

Established Online Retailer (Founded 2007)

This London-based e-commerce company has been selling gifts and homewares for adults and children since 2007.

Global Customer Base

Their customer reach extends beyond London, attracting direct-to-consumer sales from a worldwide audience.

Dual Sales Channel

While the primary focus is direct-to-consumer sales, the company also caters to B2B sales with small businesses purchasing in bulk for resale through retail channels.



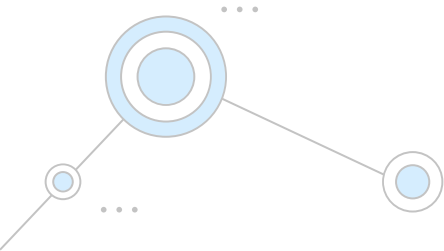
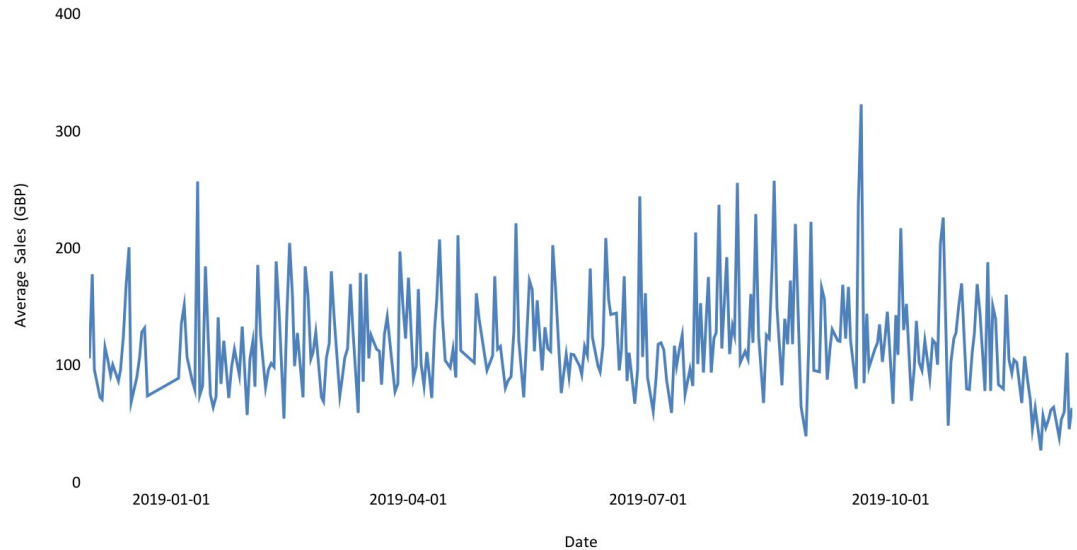
Overall sales trend were declining towards end of 2019.

While the trend appears relatively flat for most months, the sharp rise and subsequent fall towards the end of 2019 suggest **potential seasonal factors** or **promotional activities** that may have influenced consumer demand during that period.

There is a significant decline in sales performance towards the end of 2019. In November 2019, there was a **sharp drop of 22.77%**, and in December 2019, **sales declined even further by a staggering 36.13%**.

Average Sales Performance

From December 2018 - December 2019





Project Background

Our revenue declined on December 2019 to 36,13%, therefore we aim to increase declining revenue from -36,13% to 20% in Q1 2020.

This project analyzes our **sales data** from November 2018 - November 2019 to identify **high-potential customer segments** to **tailor marketing plan** based on their profile, potentially mitigates the decline back to 20% for Q1 2020 .

Business Overview

Dec 2018 - Dec 2019

Total Transactions

19,787

Revenue Earned

£61.1M

Total Items Sold

5,4M

4,717

Total Customer

3,752

Total Unique
Products

Methodology



Business Understanding

Establish the goal and scope of the analysis



Exploratory Data Analysis

Utilize Python for exploratory data analysis to understand data pattern



Data Cleaning

Utilize Python for cleaning; drop missing value, and exclude cancelled transactions



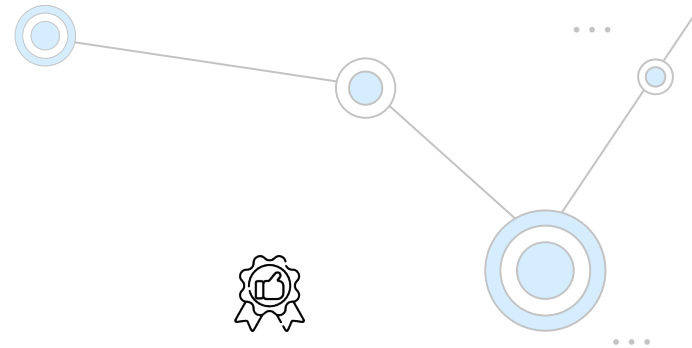
Cluster Analysis



Utilize RFM analysis and K-means clustering from RFM to segment customer based on their behavior



Recommendations


Insights and recommendation based on the analysis





02

Initial Analysis



UK Dominates Total Sales

This is obvious due to E-Commerce based in UK

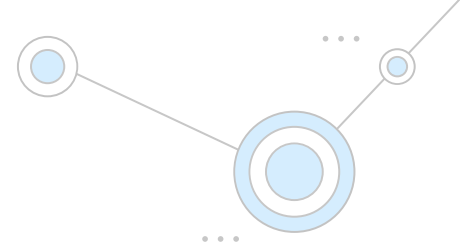
Top 10 Countries with Highest Sales



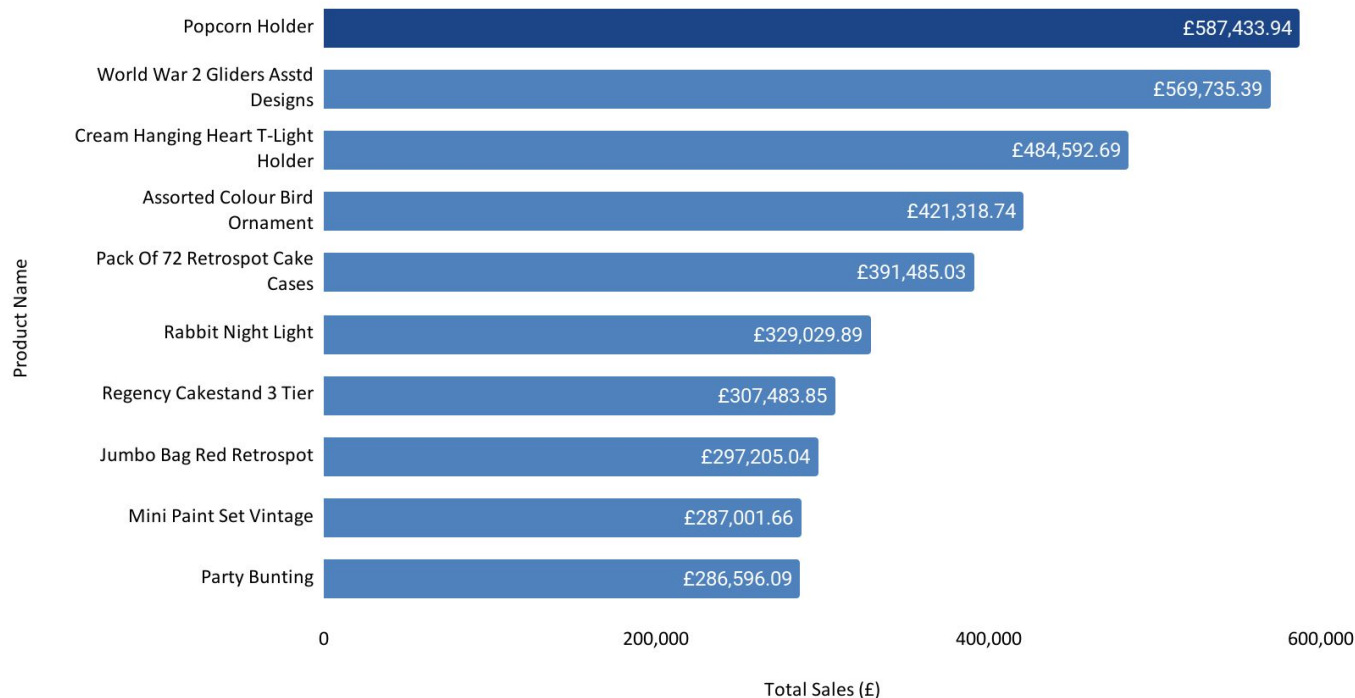
Besides the UK, notable sales contributions came from **Netherlands, Ireland (EIRE), Germany, France, and Australia**

Popcorn Holder is the best selling product

Top-selling categories: party supplies, kids toys, and home décor.

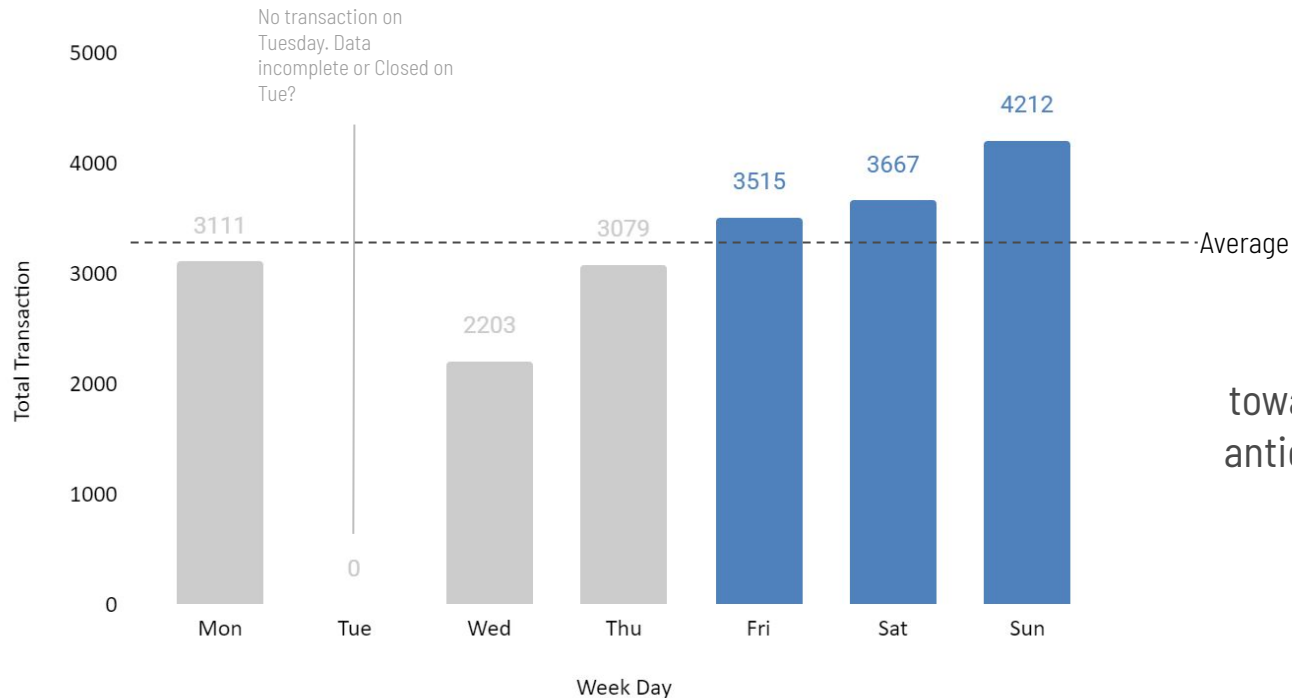


Top 10 Products with Most Sales



Prime Days: Thursday – Sunday

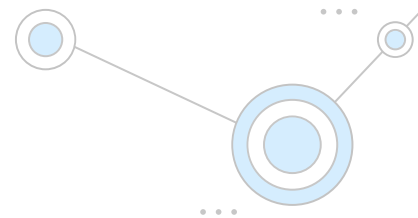
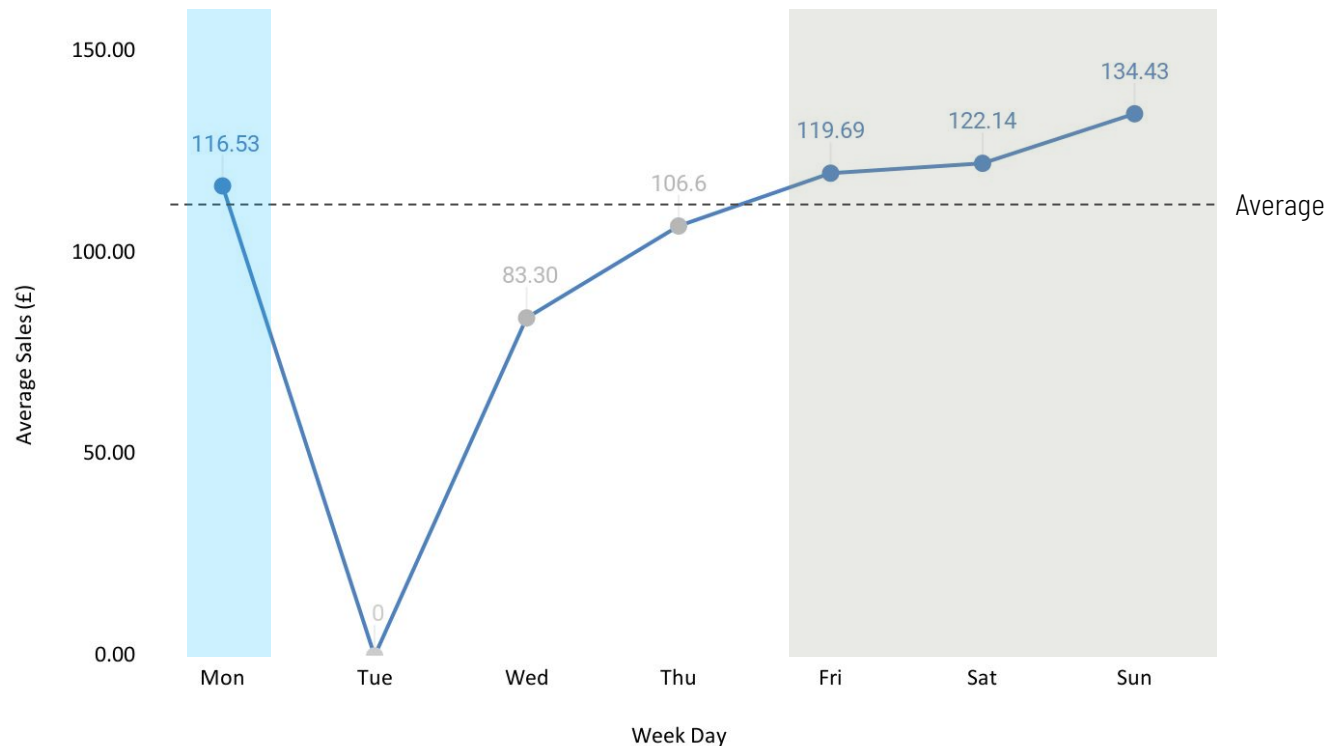
Transactions on the said days are higher than overall average
(average = 3,298)



Customers are more active towards the **weekend**, possibly anticipating weekend activities

Higher Average Sales on the Weekend

Average sales are higher compared to weekdays (average = £113.87)

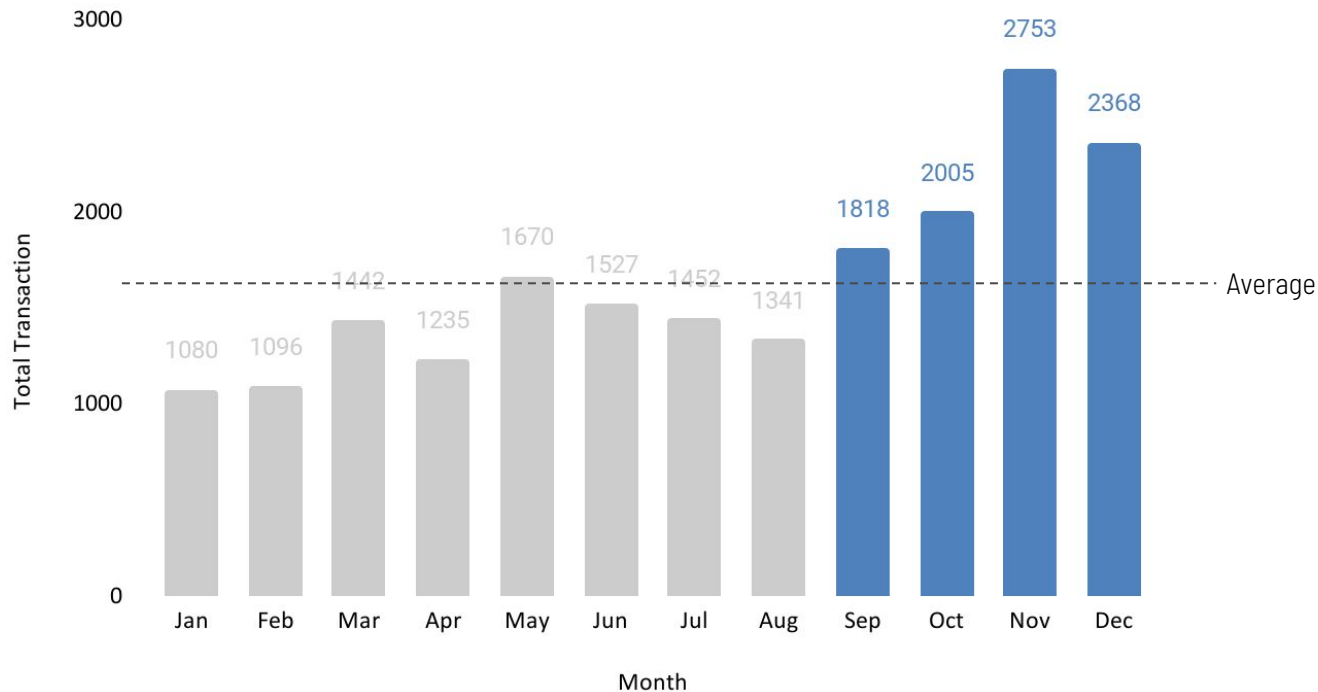


On **weekends**, customers not only shop more frequently, but they also spend more, either buying **higher-priced items** or **larger quantities**

Monday is also worth noting given average sales is also slightly above average

Prime Months: September – December

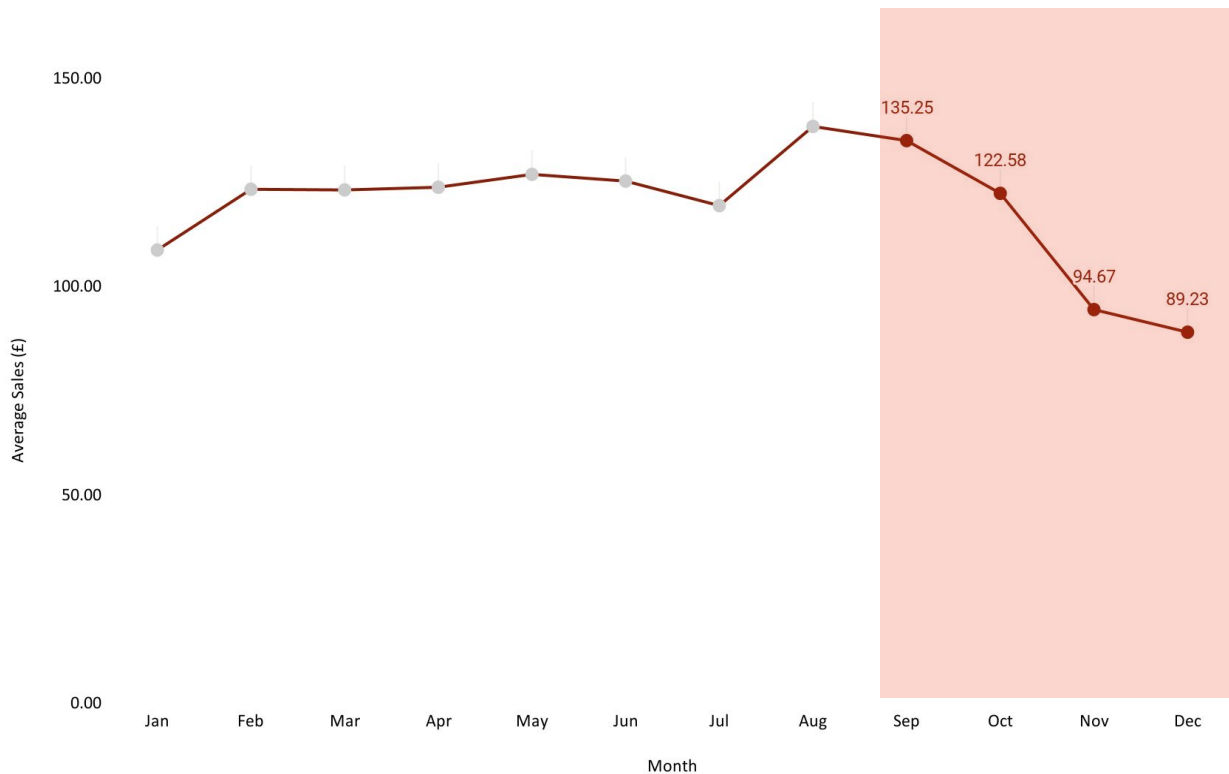
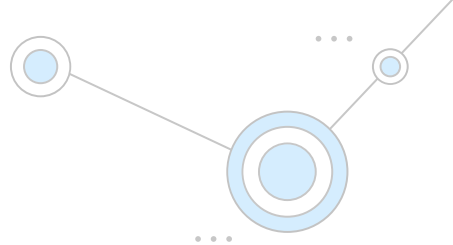
Transactions on the said months are higher than overall average (average = 1,649)



Customers are more engaging towards the **end of year**, possibly anticipating end of year holidays season and promotions

Average Sales Dropped for the Past 4 Months

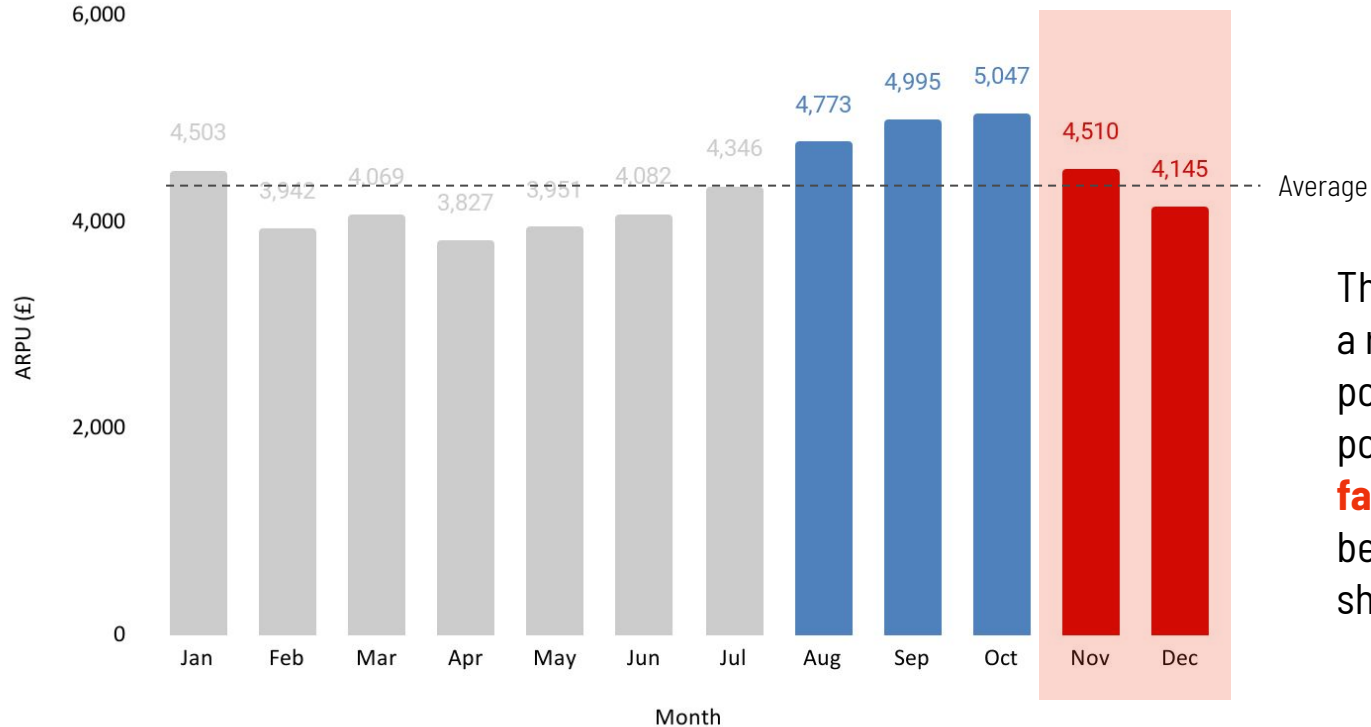
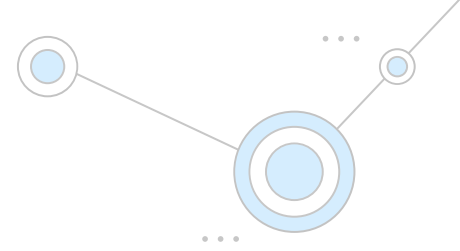
Although the total transactions for past 4 months are higher than other months



Customers may conduct transactions more during these times, but prefer products with **lower prices** or make purchases with **fewer quantities**

Average Revenue per User Decrease on End Year

Average Revenue per User (ARPU) increased from September - October before drop
(average = £4,349)

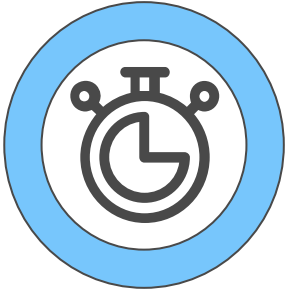


This sudden drop could be a result of post-promotional or post-holiday **spending fatigue**, or a shift in user behavior after the peak shopping season.

03

Cluster Analysis

RFM Framework



Recency

The recentness of customer's activity, like purchases or visits

Days since last purchase



Frequency

Frequency of customer's transactions or visits

Total transaction from each customer



Monetary

Value or purchasing power of customer

Total money spent from each customer

The Clusters



Loyal

Lowest Recency
Highest Frequency
Highest Monetary



Occasional

Moderate Recency
Moderate Frequency
Moderate Monetary



At Risk

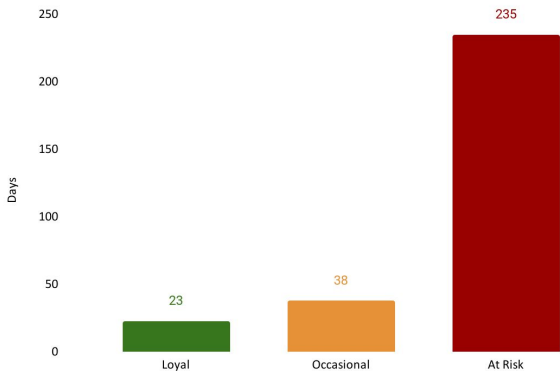
Highest Recency
Lowest Frequency
Lowest Monetary

Prime Cluster: **Loyal** Segment

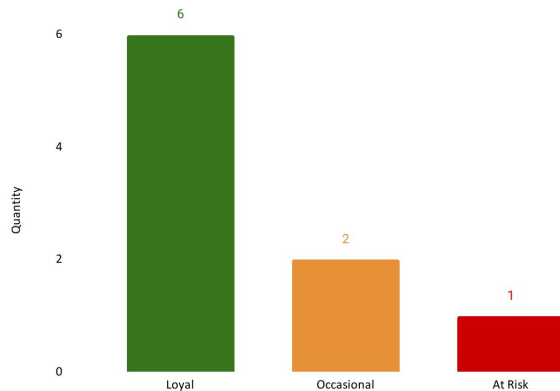
Having low recency, high frequency, and high monetary



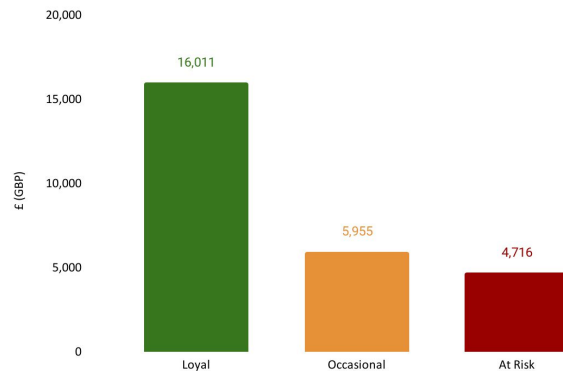
Recency



Frequency

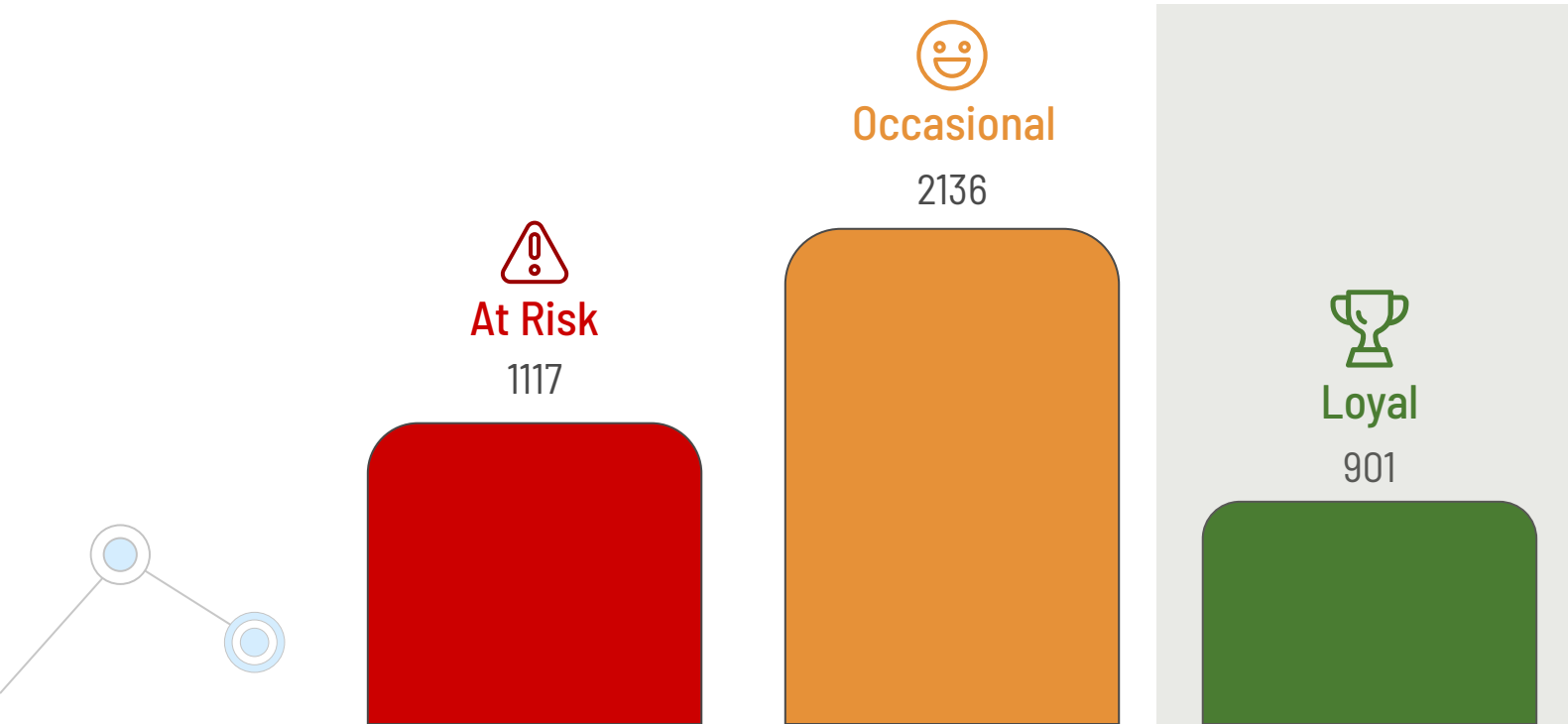
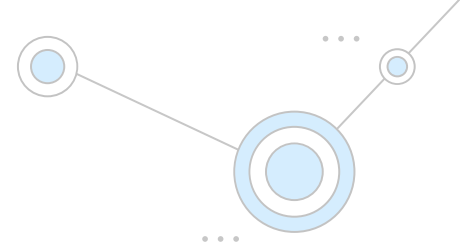


Monetary



Low Percentage of Loyal Cluster

Loyal Cluster has the lowest percentage among all segments, which is 22%



Exploring Loyal Cluster

Since we've found Loyal Cluster as the high-profile cluster, let's delve into their traits, such as:



**Preferred
Products**



**Product Price
Distribution**



**Purchase
Quantity
Distribution**

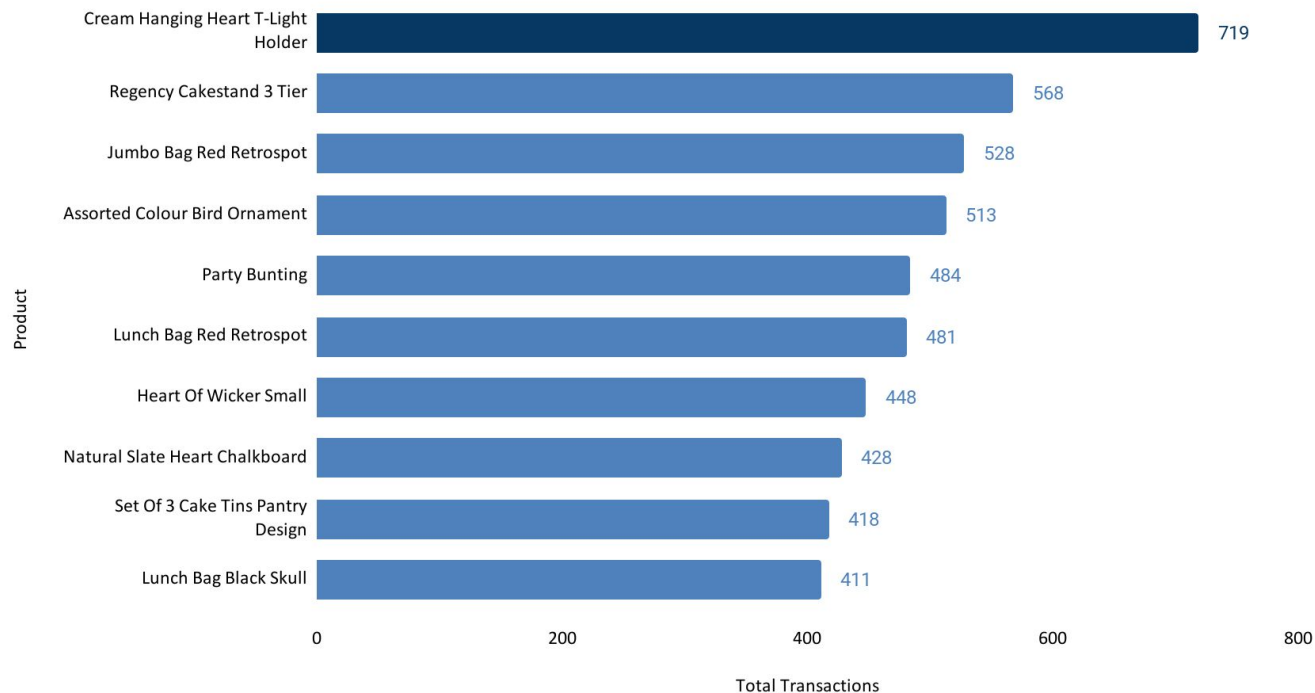


Purchase Timing

Cream Hanging Heart T-Light Holder: Best Seller

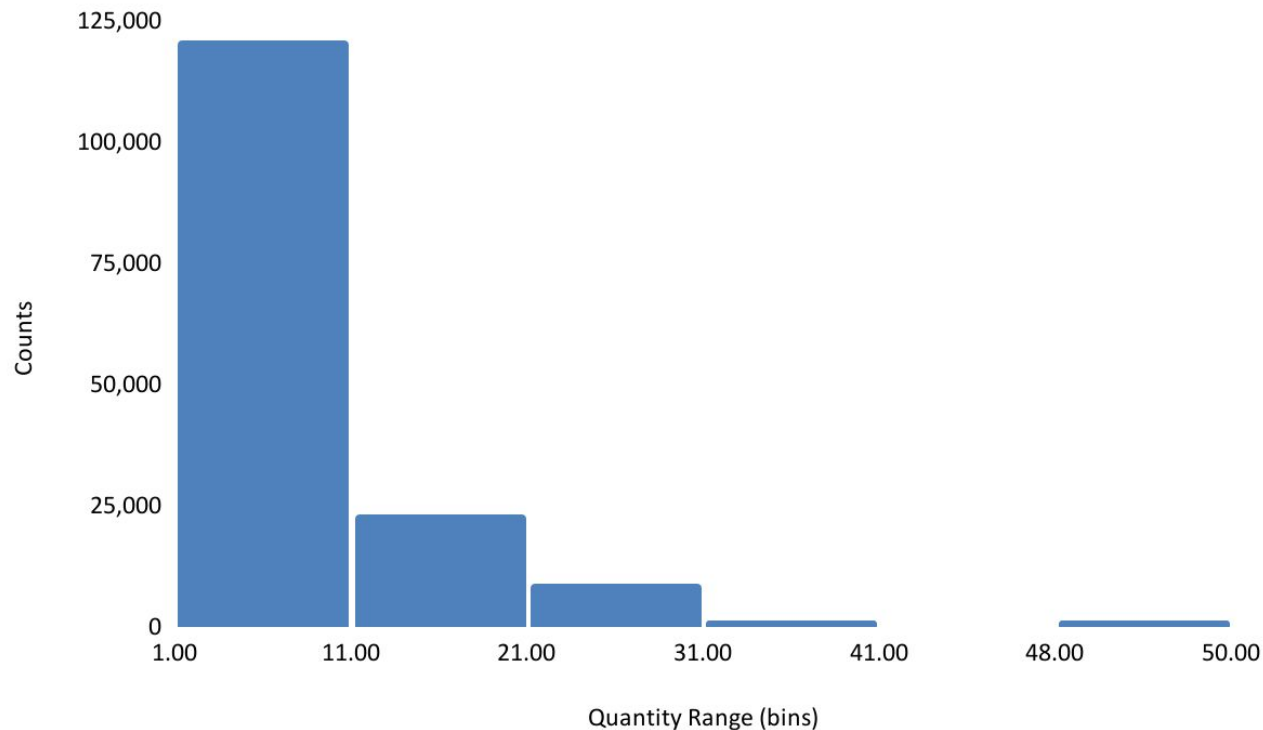
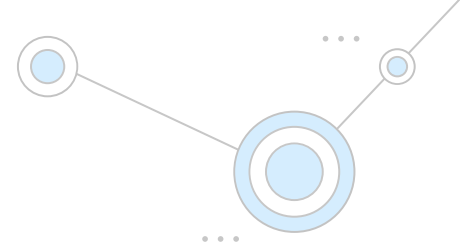
Most of these products are categorized as Home Décor, Party Supplies, and Kitchenware

Loyal Cluster: Top 10 Most Purchased Products



Loyal Cluster: Product Quantities

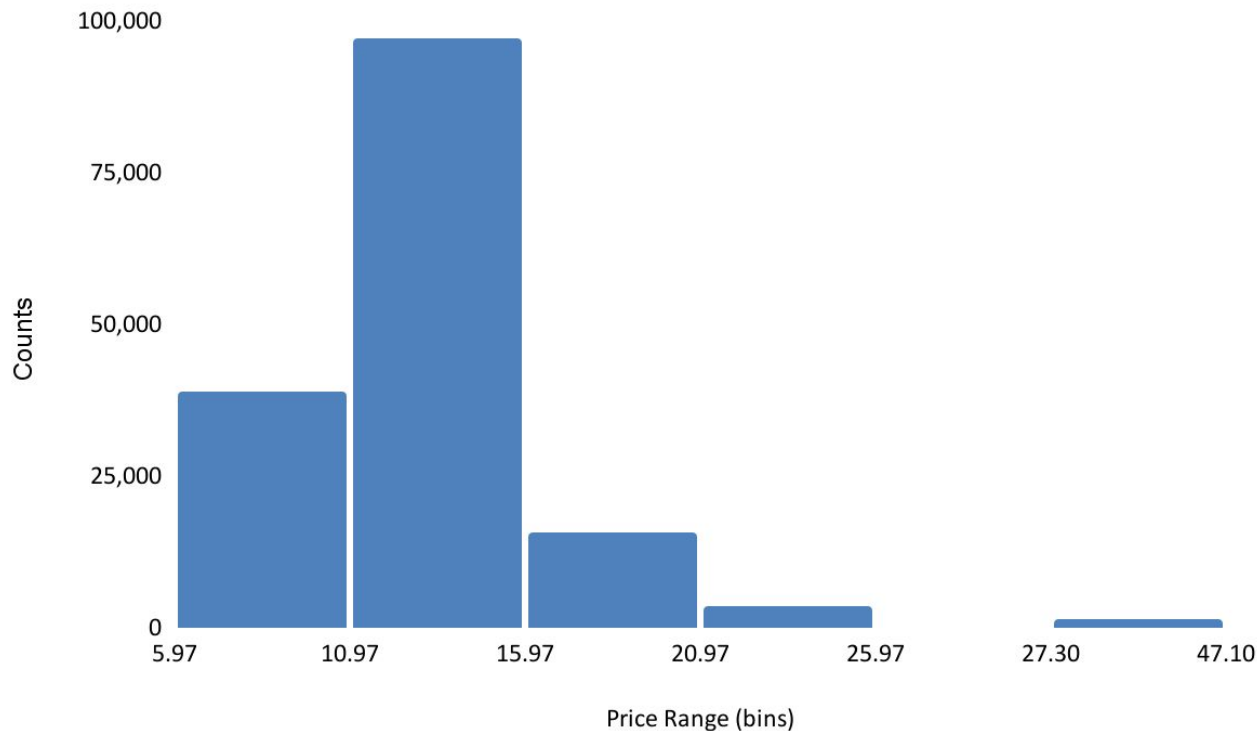
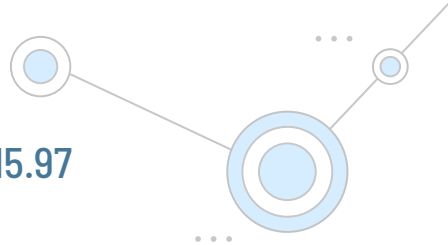
The typical order quantity for 77% of this segment falls within the 1-11 range



The Loyal Customer cluster generally makes **small** to **moderate** purchase quantities

Loyal Cluster: Product Prices

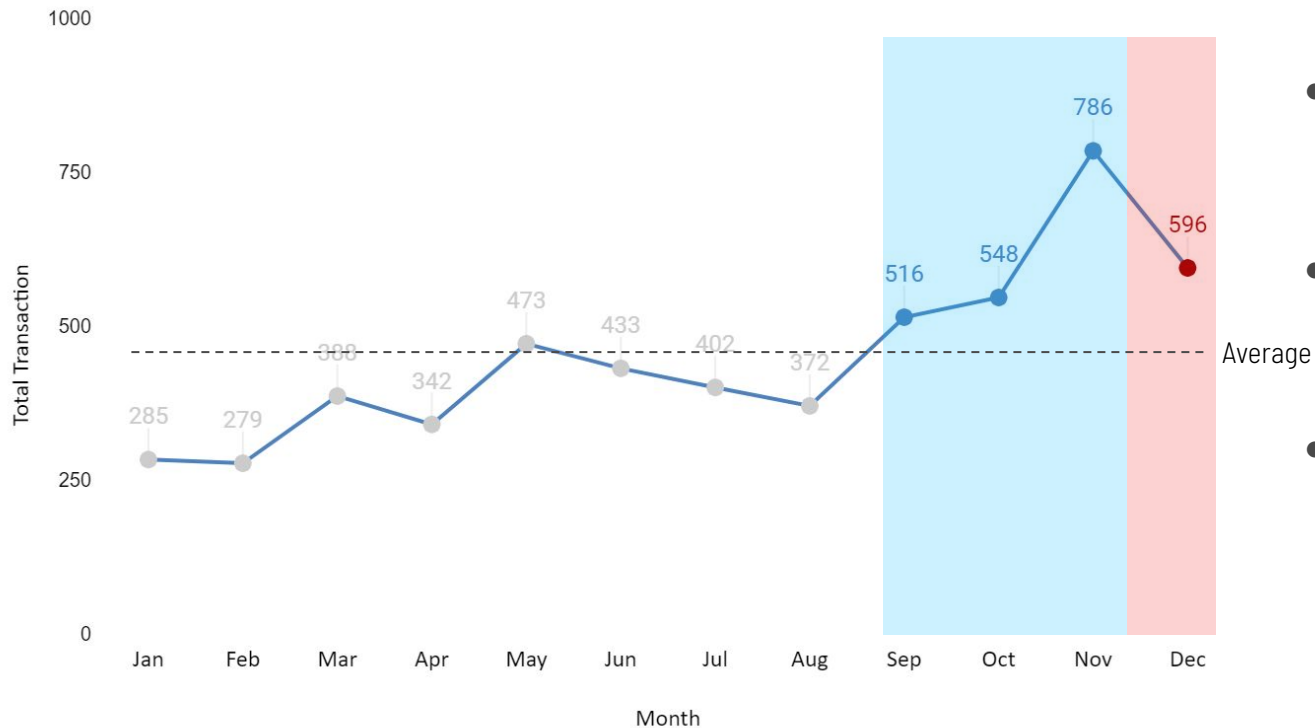
68% customers of this segment typically buys products priced between £10.97 and £15.97



The Loyal Customers cluster finds products priced between **£10.97** and **£15.97** is most appealing to buy

Prime Months: September – November

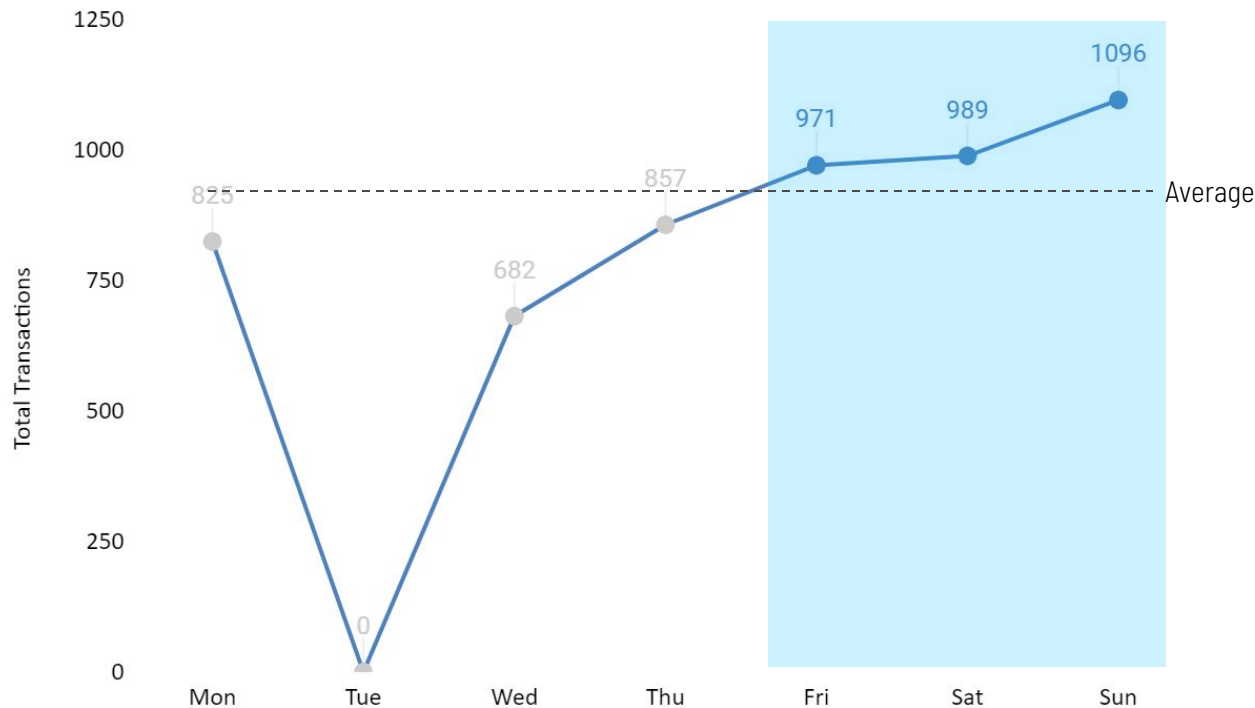
Transaction surged on September to November, but **dropped significantly on December** (average = 452)



- Pre-holiday events starts in ... **September**, leading shoppers to grab gifts early
- Customers might anticipate holiday chaos and finish shopping early (in November)
- December's **busy logistics** (Christmas & New Year's) has a probability of delays, one of reasons transactions to drop

Prime Days: Thursday – December

Transactions volumes are higher around these days for the Loyal Customers segment (average = 903)



Customers in the Loyal Customers segment buy items in anticipation of **weekend activities**, implying that the loyal customer cluster remains active and **very engaged** during the weekend period



03

Recommendations



Six Proposed Actions

With Loyal Cluster identified as high-profile cluster, we formulated business recommendation to leverage on this valuable segment



Bulk Pricing & Loyalty Point System



E-commerce Page Optimization



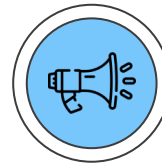
Exclusive Section for Products under £15



Email Campaign for Year End Sale



Curated Product Bundling



Weekend Campaign




Bulk Pricing & Loyalty Point System

Benefits: Increased **Revenue** and **Customer Satisfaction**

- Introduce **discounts** for purchasing specific quantities of **top-selling products**
- Threshold quantities should be strategically chosen to encourage bulk purchases **without sacrificing profit margins** too much
- Implement a **Bronze, Silver, and Gold** customer tier system based on **purchase frequency**
- Higher tiers unlock greater benefits, **incentivizing increased purchases**

Action Required:

- **Cost of Goods Sold (COGS) analysis** by product to determine profitable bulk pricing
 - **Discount sensitivity analysis** to ensure attractive discounts level without significantly impacting profitability
 - Define clear **purchase frequency requirements** for each tier to incentivize repeat orders and reward high-value customers
 - Design **attractive and differentiated benefits** for each tier
- 




Exclusive Section for Products under £15

Benefits: Increased **Sales** and **Customer Experience/Engagement**

- Create an independent "**Under £15!**" section on the e-commerce home page
- Feature **top-selling** and **similar items** under £15 from their preferred categories within this section
- Host **monthly events** within the "Under £15!" section

Action Required:

- Develop **algorithms** for **dynamic product selection** within the "Under £15!" section
 - Collaborate with **Marketing and Product Management** teams to **optimize** product selection & category presentation
 - Develop a **plan for promoting** within the "Under £15!" section, like themed sales
- 




Curated Product Bundling

Benefits: Increased **Sales**, **Average Order Value (AOV)**, **Simplified** Purchase

- **Create pre-packaged bundles** like "Party Poppers for New Year's Eve" featuring **a mix of relevant products** within different categories
 - (Potential: Home Décor, Party Supplies, and Kitchenware) accompanied by exclusive New Year cards
- Offer **customized bundles** based on customer data such as **interest and behavior**

Action Required:

- **Craft attractive visuals** and **compelling messages** to promote the bundles across marketing channels
 - **Conduct market research** to identify trending products and upcoming events (e.g., holidays, gift-giving occasions) that can inform bundle creation
- 




E-Commerce Page Optimization

Benefits: Increased Engagement and Sales

- Highlight **top-selling products** for the Loyal Customers cluster within the **first section** of their dedicated landing page
- Feature **similar products** related to each top seller alongside it, encouraging **further exploration and purchases**

Action Required:

- **Leverage personalization** techniques to dynamically populate the page with products tailored to Loyal cluster's behavior
 - **A/B test** different placement strategies for optimal performance
 - Work with **Product Management teams** to optimize product selection and cross-selling suggestions
- 




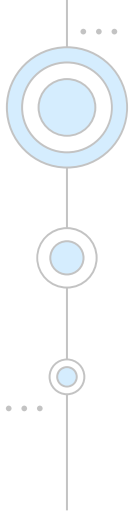
Email Campaign for Year End Sale

Benefits: Increased **Engagement, Sales and Conversion**

- Send targeted push emails from **September to November**, focusing on **Thursdays to Sundays** when engagement is typically higher
- Leverage these periods to deliver personalized messaging that "**nurtures**" customers towards **purchasing** their **top-selling products**

Action Required:

- Define the optimal **email frequency** and **scheduling** for sending push notifications during peak season (through surveys and next analysis)
 - **Design personalized push emails** featuring top-selling products and interesting offers
 - Design **visually appealing** email templates that are mobile-responsive and optimized for different email clients
- 



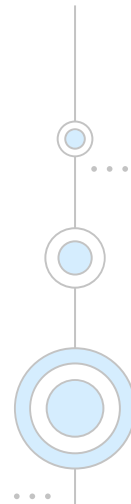
Weekend Campaigns

Benefits: Increased **Customer Engagement** and **Brand Awareness**

- Introduce a limited-time **"Happy Weekend"** campaign offering exclusive perks to our most valuable customers - the Loyal Customers segment
- Campaign Options: **Double Loyalty Points, Best-Selling Product Discounts, Reduced/Free Delivery Costs**

Action Required:

- Identify the campaign option with the **highest acceptance/use rate** through **A/B tests and surveys** after the event trial
- Collaborate with **Marketing and Finance** departments to **assess campaign costs** and potential **impact on profitability**
- Utilize **CRM system** to schedule and manage email and push notifications for the campaign





Thank You!



Thank You!

Do you have any questions?

satryo.sunup@gmail.com



[Satryo Sunu](#)
[Prabowo](#)

Resources

Dataset

[E-commerce Business Transaction
\(kaggle.com\)](#)

Data Analysis

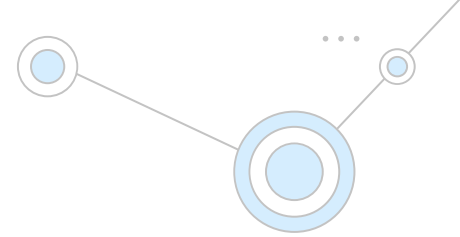
[Google Colab](#)
[Unclustered Sheet #1](#)
[Unclustered Sheet #2](#)
[Clustered Sheet #1](#)
[Clustered Sheet #2](#)

Appendix

A decorative graphic featuring a large, light blue, irregular blob in the center. The word "Appendix" is written in a bold, black, sans-serif font within this blob. Surrounding the blob are several network-like structures. On the left, a vertical line of three nodes (small blue circles with white outlines) is connected by a thin grey line. Below this, a larger node (a blue circle with a white outline and a blue center) is connected to two other nodes. On the right, a vertical line of three nodes is connected by a thin grey line. At the bottom, a larger node is connected to two other nodes. All nodes are blue circles with white outlines and blue centers. The lines are thin and grey.

Dashboard

[Link](#)



E-Commerce Dashboard

Country All

Date 12/1/2018 to 12/9/2019

Revenue 970  50,681,745

527,762

Total Transactions

£61.1M
Revenue Earned

5,436K
Items Sold

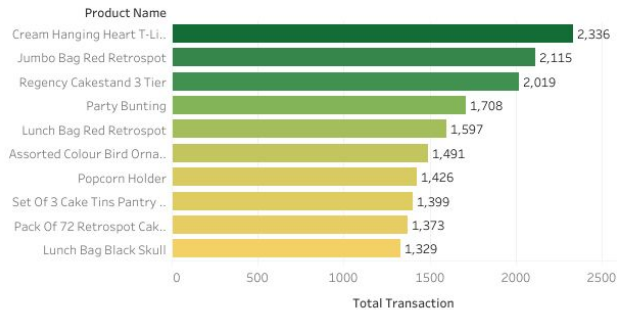
Sales Performance



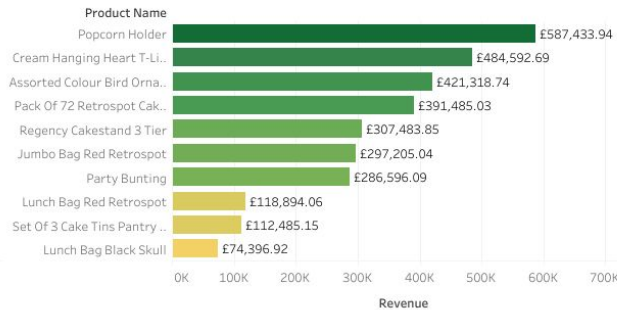
Profit by Region



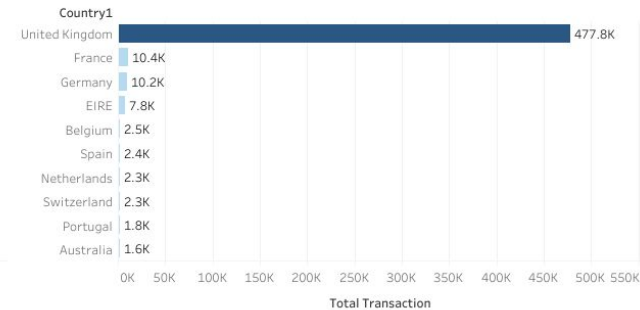
Top 10 Product Purchased



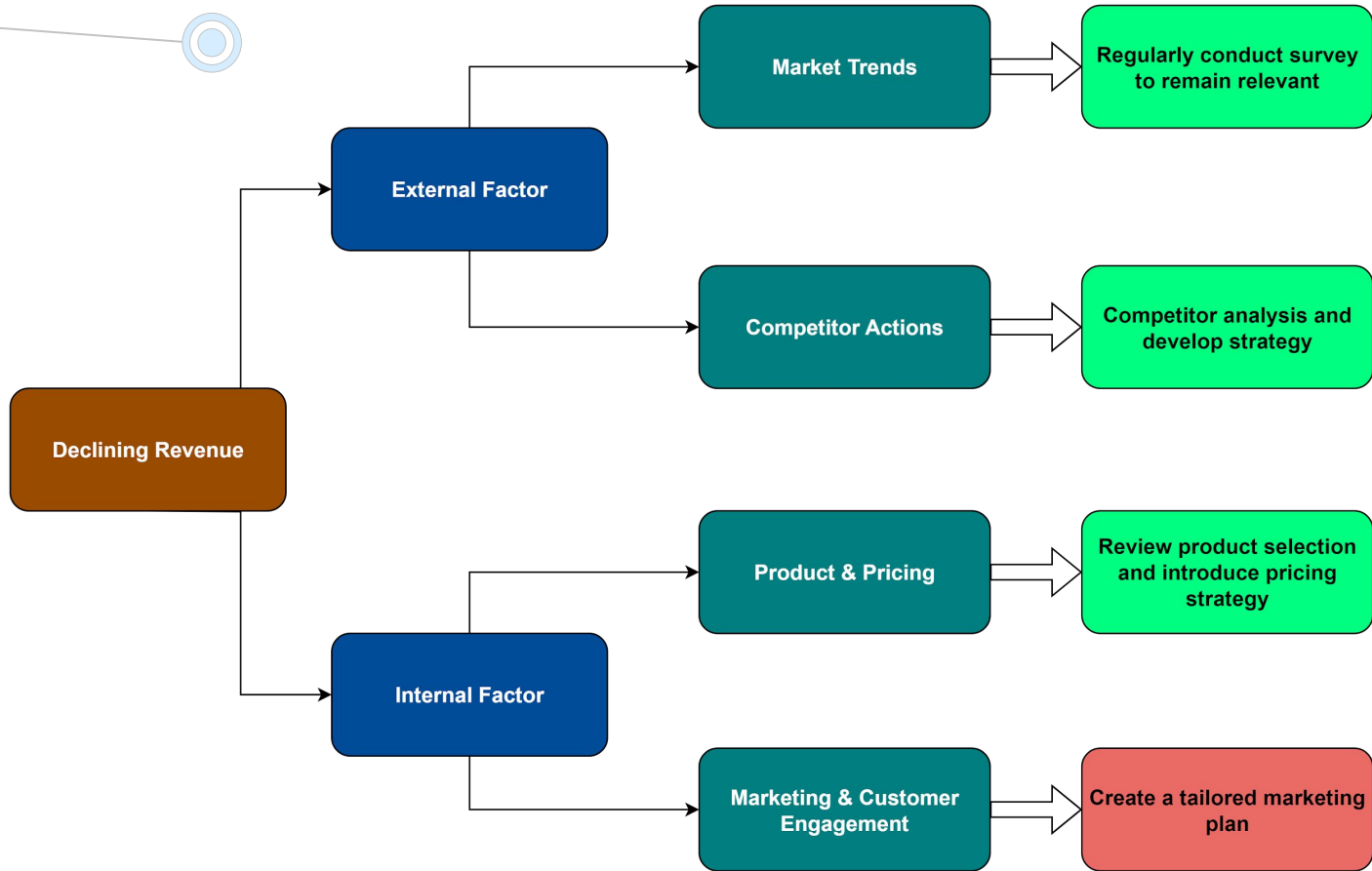
Top 10 Product Revenue



Top 10 Country Transactions

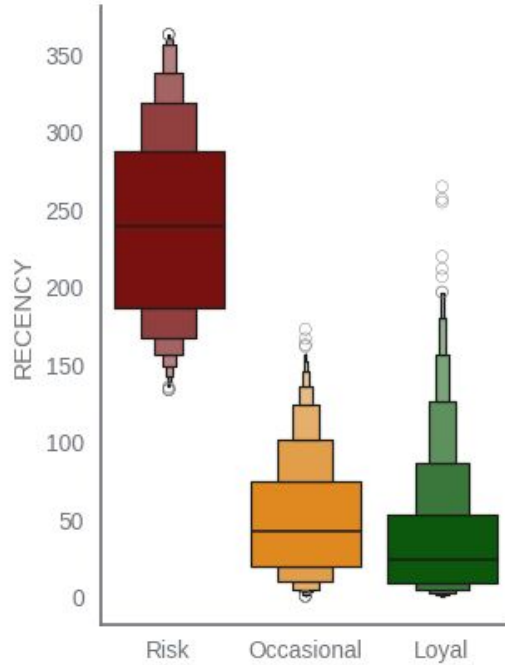


ROOT CAUSE ANALYSIS

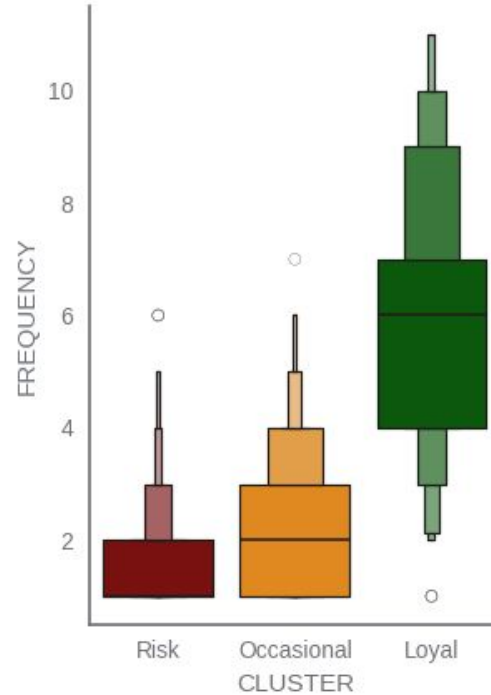


Data Distribution Each Cluster

Distribution of Recency across Clusters



Distribution of Frequency across Clusters



Distribution of Monetary across Clusters

