Web Marketing & Communication Management

Sara Nocco 892030
Federico Pirola 846725
Igor Schirripa 883614

Business questions:

1) Customer Relationship Valorization:

How to make the relationship with customers more profitable?

Exploiting customer base data to understand:

- What are the characteristics of the customer base?
- How can it be segmented?
- Which are the high-value customers to retain and which are the potential churners?
- Which way the customers behavior changed over the last 2 quarters?

2) Products Valorization:

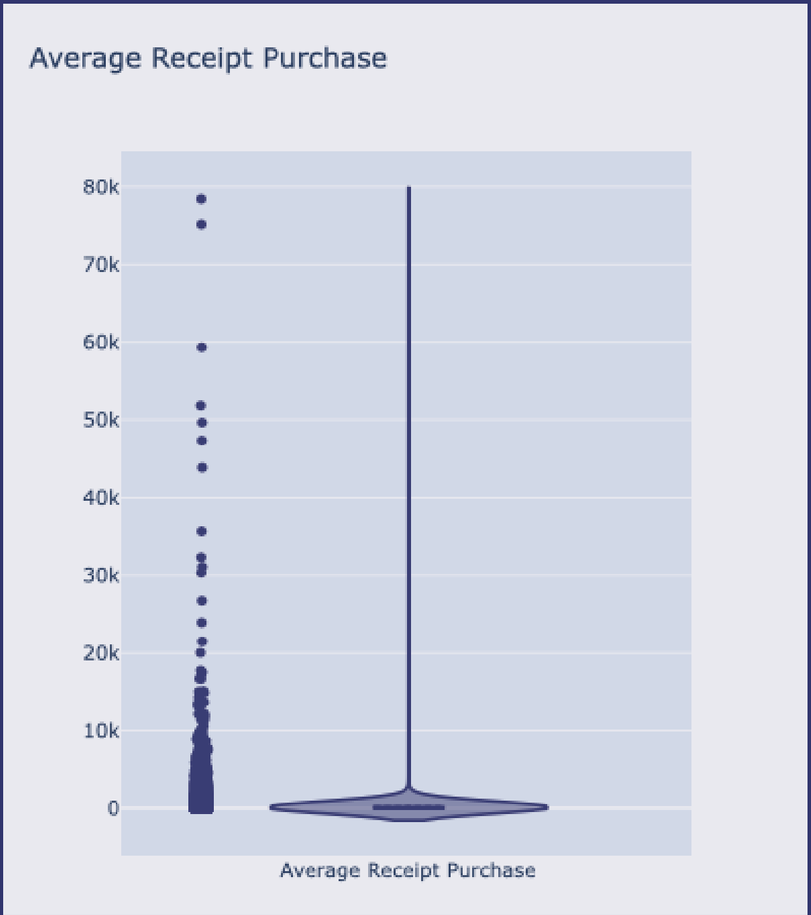
Which products are more likely to be sold together?

- Which department they come from and what is their price range?

Let's take a look at data...

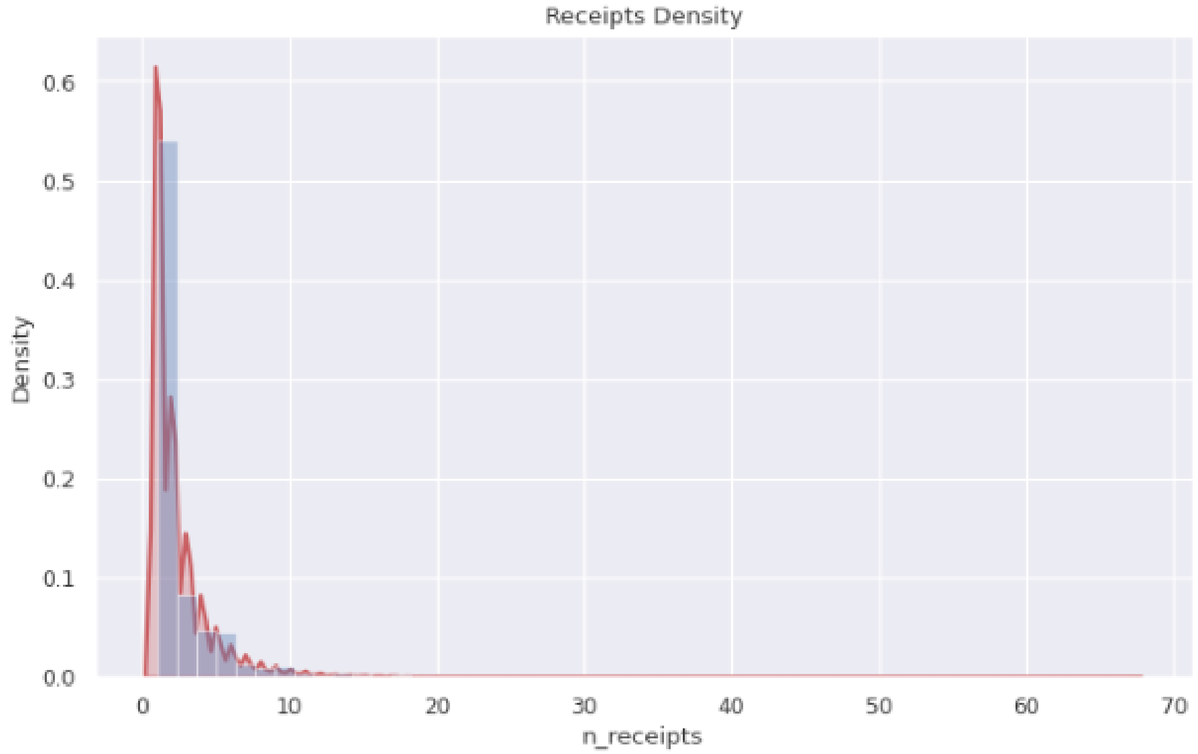
Data exploration:

Average expenditure is concentrated from 23 to 128 with median receipt close to 55. Outliers from 10k to almost 80k.



mean	142.70
std	679.99
min	0
25%	23.98
50%	54.10
75%	128.89
max	78442.96

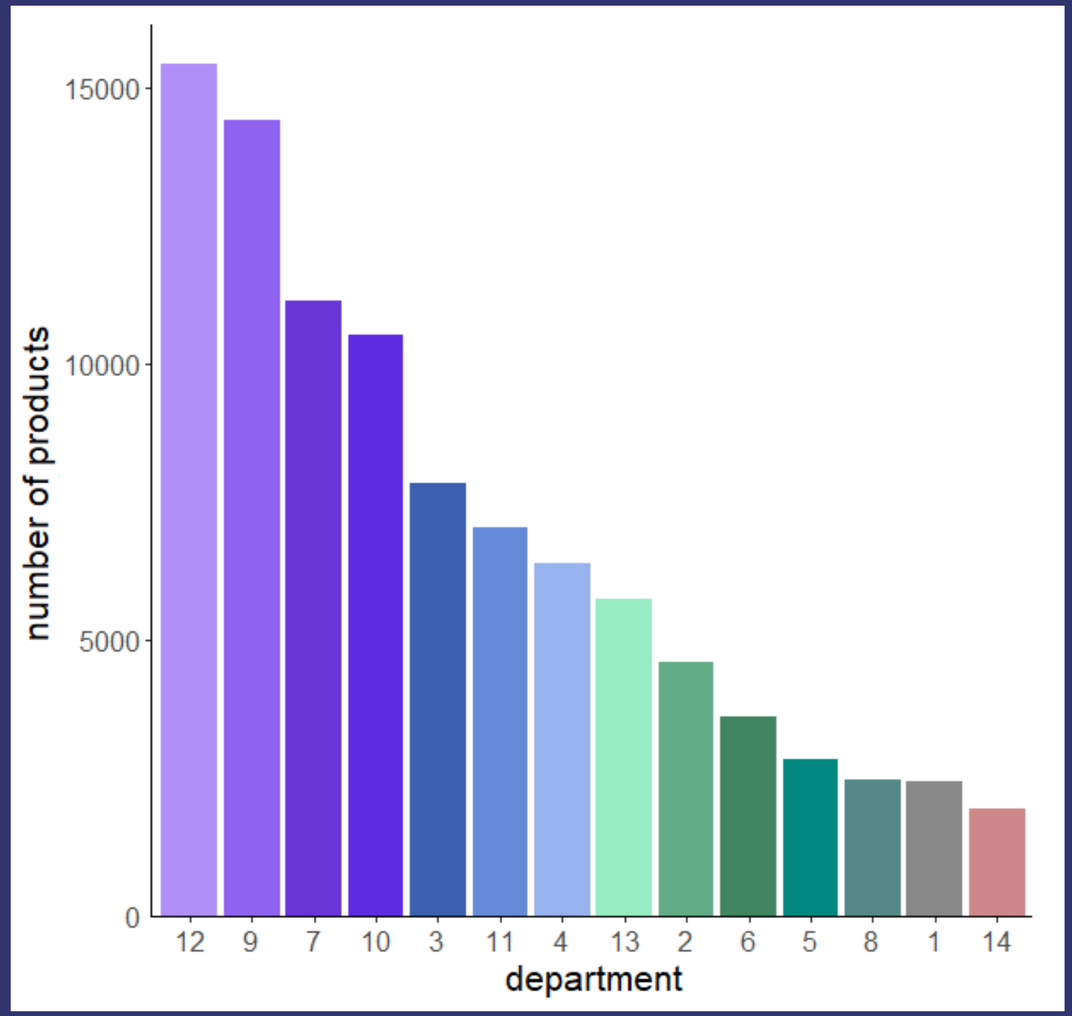
Customer purchasing behavior



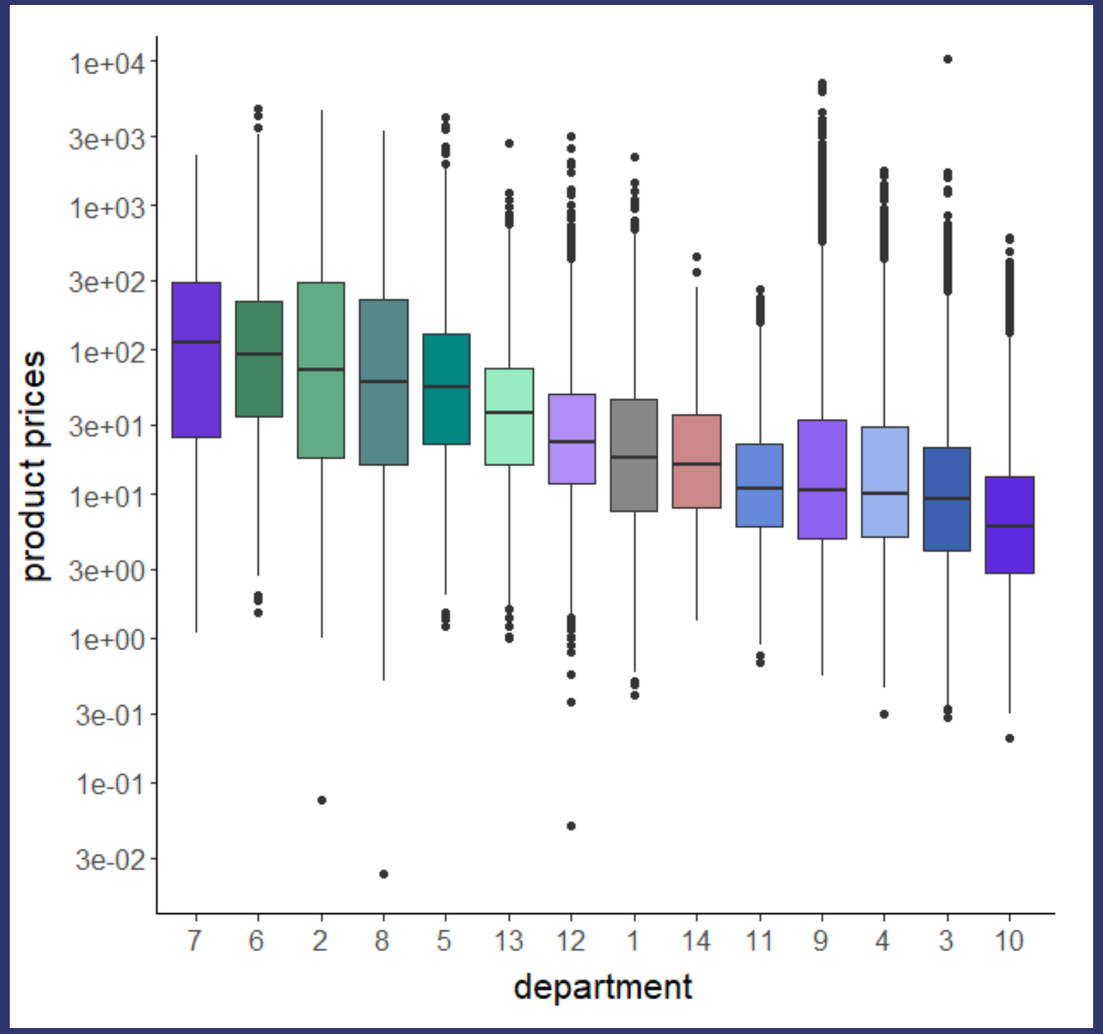
Most of the customers generate from 1 to 10 receipts.
Only a small portion of the customer base generates a number of receipts between 20 and 65.

Data Exploration

Products descriptive analysis

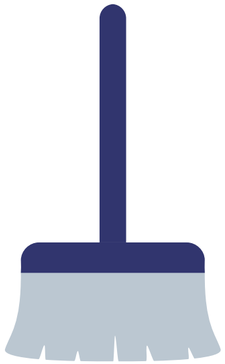


Department #12 is the one with the most product sold.



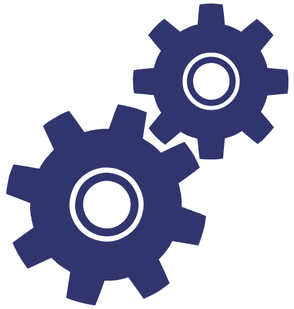
Whereas Department #7 is, on average, the department with most expensive products.

Data-driven strategy



Cleaning

Missing Values handling & inconsistency resolution



Preprocessing

Data type conversion, features transformation and creation

Exploration

Descriptive analysis of customers purchasing behavior & sold products

Modelling

Descriptive and predictive modelling answering the business questions in a quantitative and qualitative way

Models

1) Customer Relationship Valorization:

(Trimestral analyses using the following approaches)

Cluster Analysis:
customers behavioural segmentation

RFM Model:
customers deterministic segmentation &
movement between segments

Churn Model:
prediction of leaving customers

2) Products Valorization

(Annual Analyses)

Market Basket Analysis:
study of product affinities by purchases

Cluster Analysis

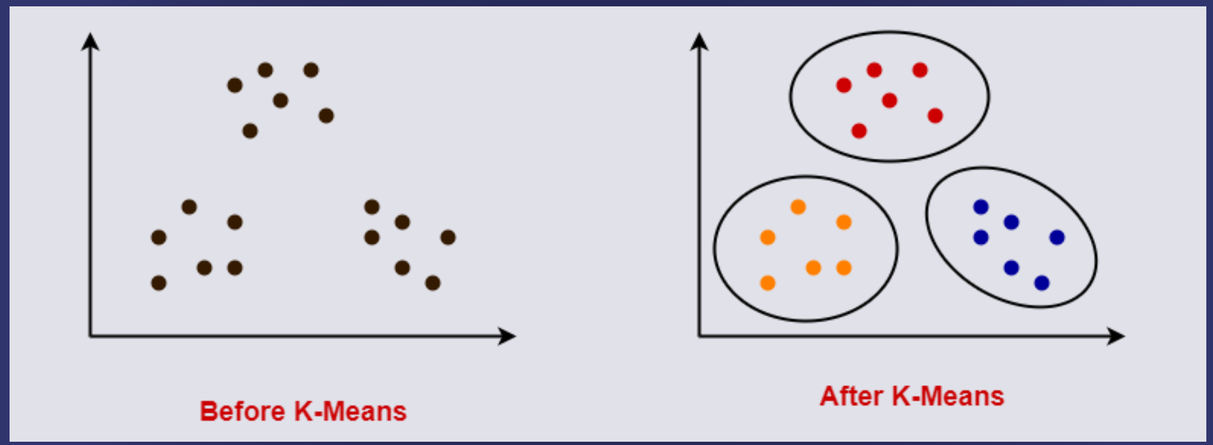
K - Means Algorithm

Reference period: 01/02/19 - 30/04/19

How it works:

K-means clustering uses “centroids”, K different randomly-initiated points in the data, and assigns every data point to the nearest centroid.

After every point has been assigned, the centroid is moved to the average of all of the points assigned to it.



FEATURES:

- Number of receipts
- Number of products
- Average expenditure
- Average discount rate
- Number of clicked mails

K-means image from https://www.gatevidyalay.com/

RFM

Reference period: 01/02/19 - 30/04/19

Recency	Frequency	Monetary Value
How recently was the customer's last purchase?	How often did this customer make a purchase in a given period?	How much money did the customer spend in a given period?

FEATURE -> Last purchase date

FEATURE -> Number of receipts

FEATURE -> Total net import

Segments:

Diamond
Premium Customers

Gold
High Value Customers

Silver
Medium-High Value Customers

Bronze
Average Customers

Copper
Medium-Low Value Customers

Tin
One-Timers

Cheap
Low Value Customers

Icons from https://www.flaticon.com/

Churn Model

Lookback Period -> 01/11/18 - 31/01/19

Holdout Period -> 01/02/19 - 30/04/19

FEATURES:

- Number of receipts
- Total amount spent
- Average discount rate
- Number of campaigns
- Opened e-mails rate (%)
- Clicked e-mails rate (%)
- Returned item (binary)
- Privacy subscription 1 (binary)
- Privacy subscription 2 (binary)
- Privacy subscription 3 (binary)

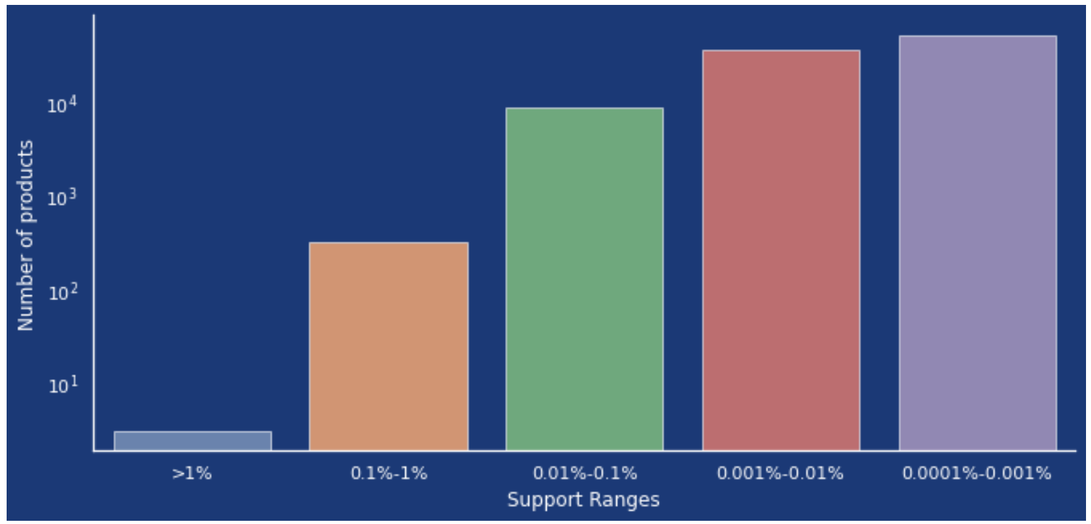
Not Churner	Churner	Inactive
Active fidelity card & at least 1 purchase	Active fidelity card but no purchases; Purchased but Inactive fidelity card	No purchases & Inactive fidelity card; Fidelity card activated after 01/11/18

Market Basket Analysis

Reference period:
01/05/18 - 30/04/19

Looking for joint purchasing patterns

Affinities among products with high support



Support >0.1% chosen based on the above distribution

Insights

Results:

1) Customer Relationship Valorization:

- Cluster Analysis

- RFM Model

- Churn Model

2) Products Valorization

- Market Basket Analysis

Cluster Analysis - Results

Buy till you die!

I just need to buy a couple of things

Loyal Customer (1.33%)

<u>Customers who buy very frequently</u>

num. receipts: 16.95
num. products: 104.8
avg. discount: 3.7 %
avg. bill: 177 $
num. tot clicks: 0.35

One-time buyer (80.41%)

<u>Customers who buy on average less than 2 times per quarter</u>

num. receipts: 1.63
num. products: 5.81
avg. discount: 2.5 %
avg. bill: 89 $
num. tot clicks: 0.09

Personas images from https://thispersondoesnotexist.com

I'm always looking for opportunities

Clicker (2.53%)

Customers who are very responsive to campaigns

num. receipts: 3.29
num. products: 13.94
avg. discount: 3.7 %
avg. expenditure: 107 $
num. tot clicks: 5.22

I'm interested in expensive products

Premium Customer (2.90%)

Customers who buy high-priced products

num. receipts: 1.75
num. products: 14.36
avg. discount: 6.5 %
avg. expenditure: 1221 $
num. tot clicks: 0.1

Balance is the key!

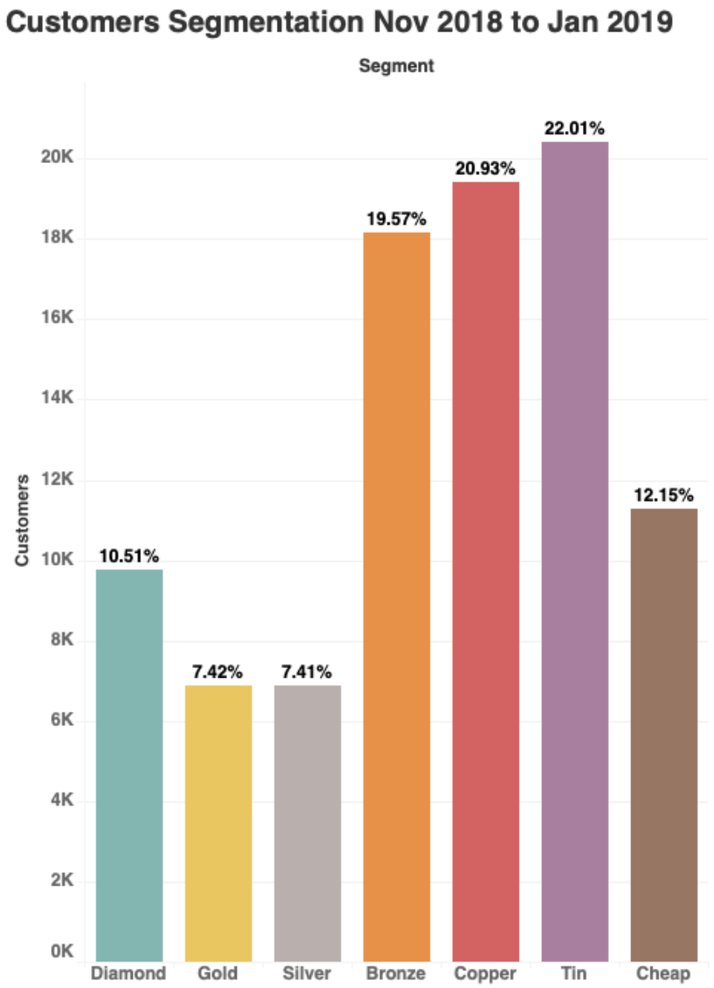
Average Customer (12.83%)

Customers with average purchasing behavior

num. receipts: 6.16
num. products: 29.4
avg. discount: 2.8 %
avg. expenditure: 128 $
num. tot clicks: 0.16

Personas images from https://thispersondoesnotexist.com

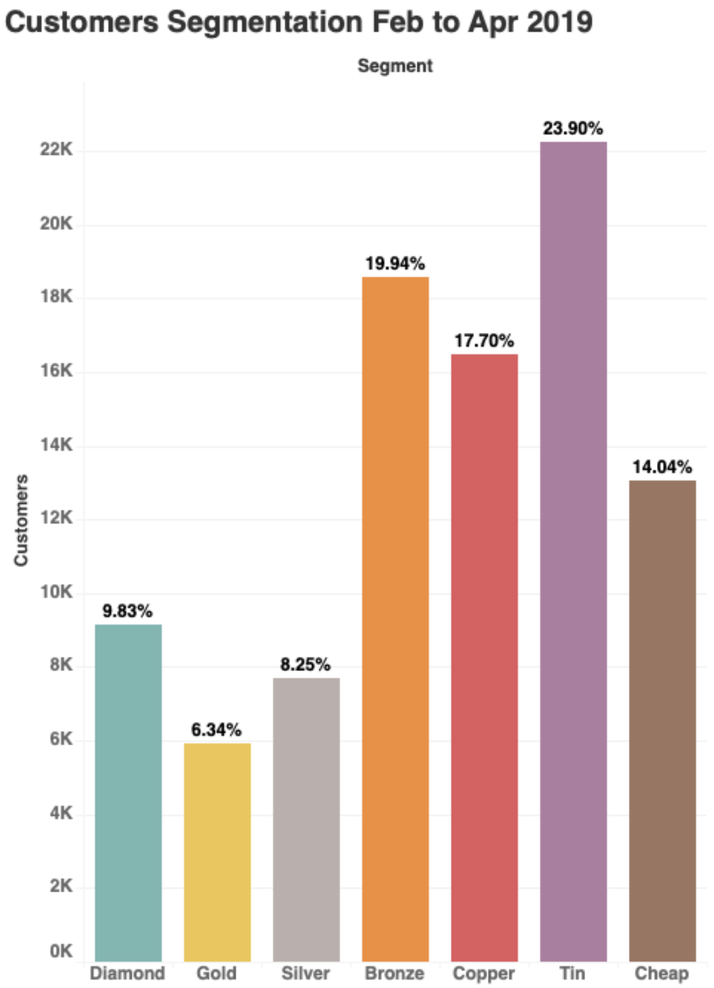
RFM - Results



Comparison of segments built over 2 consecutive trimesters

The percentage of customers in the **higher-value** segments **decreased** over the 2 periods

L**ower-value** segments **increased** over the 2 periods

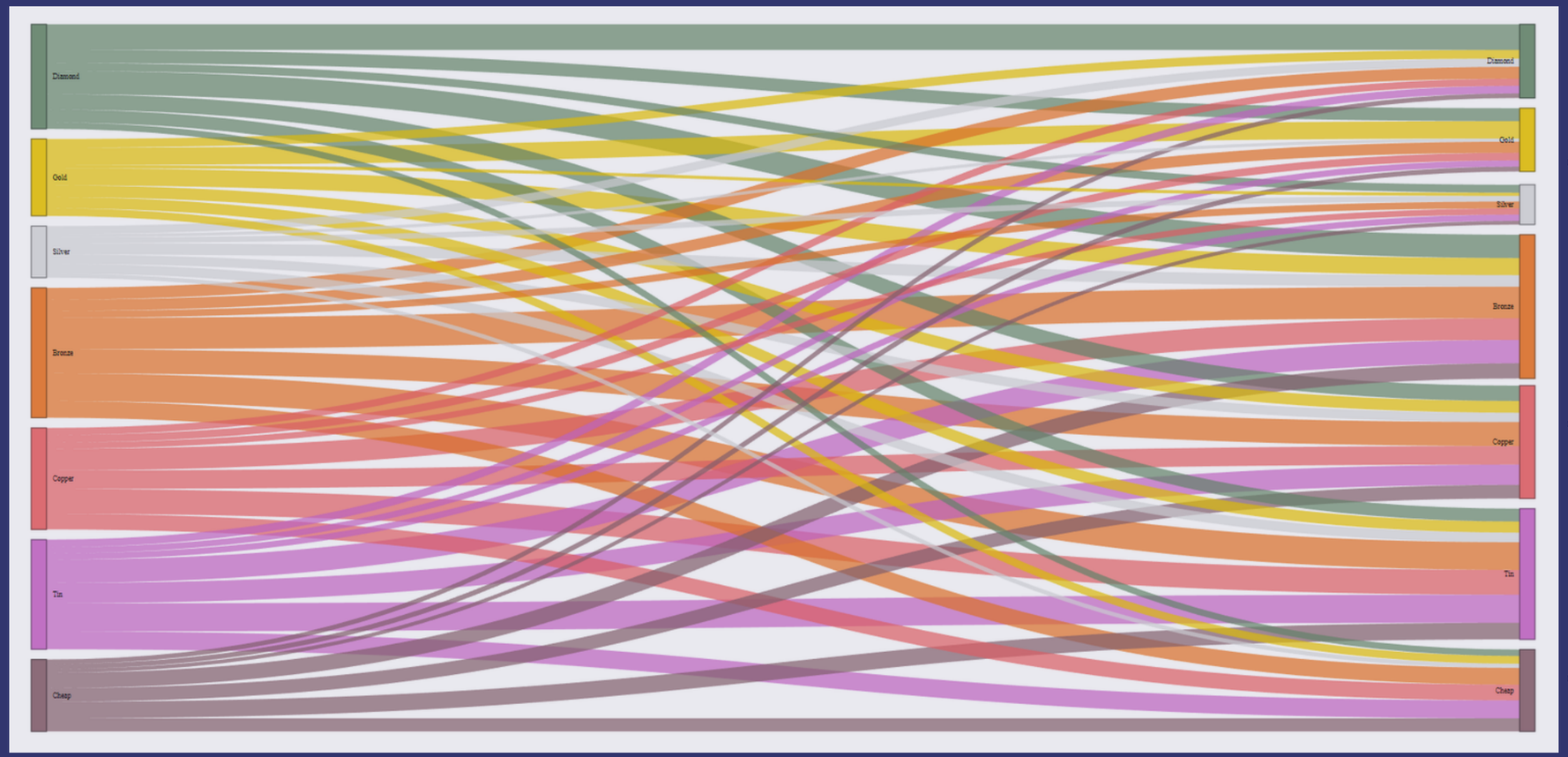


Let's take a look closer

Customer journey among segments

01/11/2018 to 31/01/2019

01/02/2019 to 30/04/2019



For interactive visualization click here

Actionable tips

- **Stayed in their high value segment**
  --> Upsell higher value products. Ask for reviews. Engage them
- **From high to higher value segment**
  --> Reward them. Can be early adopters for new products and will promote the shop
- **From high to lower value segment**
  --> Win them back via renewals or newer products, don't lose them to competition. Reconnect with them
- **From medium to higher value segment**
  --> Offer membership / loyalty program, recommend other products

Churn Model - Results

MODEL	ACCURACY	SPECIFICITY	RECALL
Logistic	0.7389	0.5908	0.8003
Tree	0.7871	0.9739	0.7096
Random Forest	0.8225	0.8369	0.8165
Neural Network	0.7570	0.9747	0.6667

CHURNER
29.32% (19.162 customers)

NOT CHURNER
70.68% (46.184 customers)

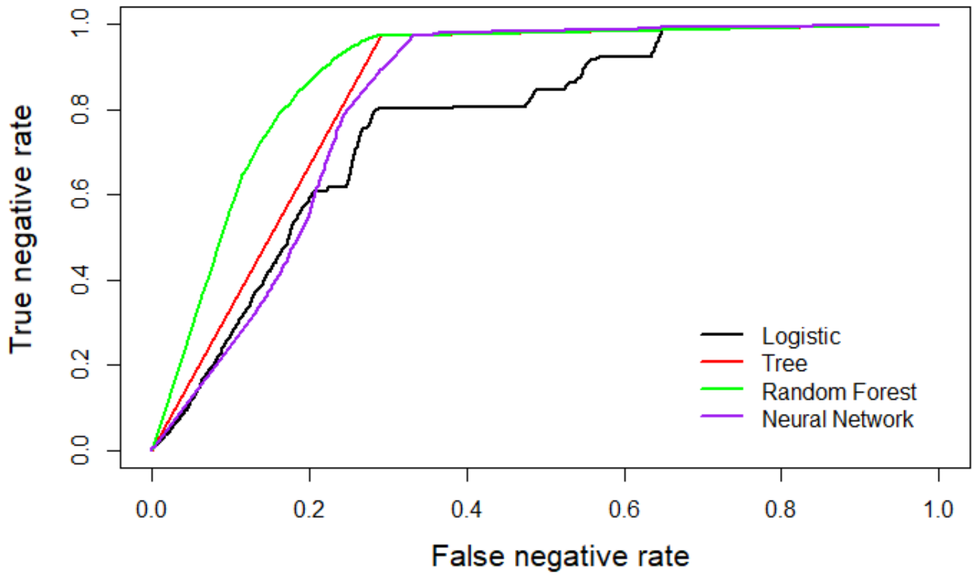
AREA UNDER THE CURVE (AUC)

Logistic: 0.7695

Tree: 0.8418

BEST PERFORMING MODEL ← **Random Forest: 0.8874**

Neural Network: 0.8202

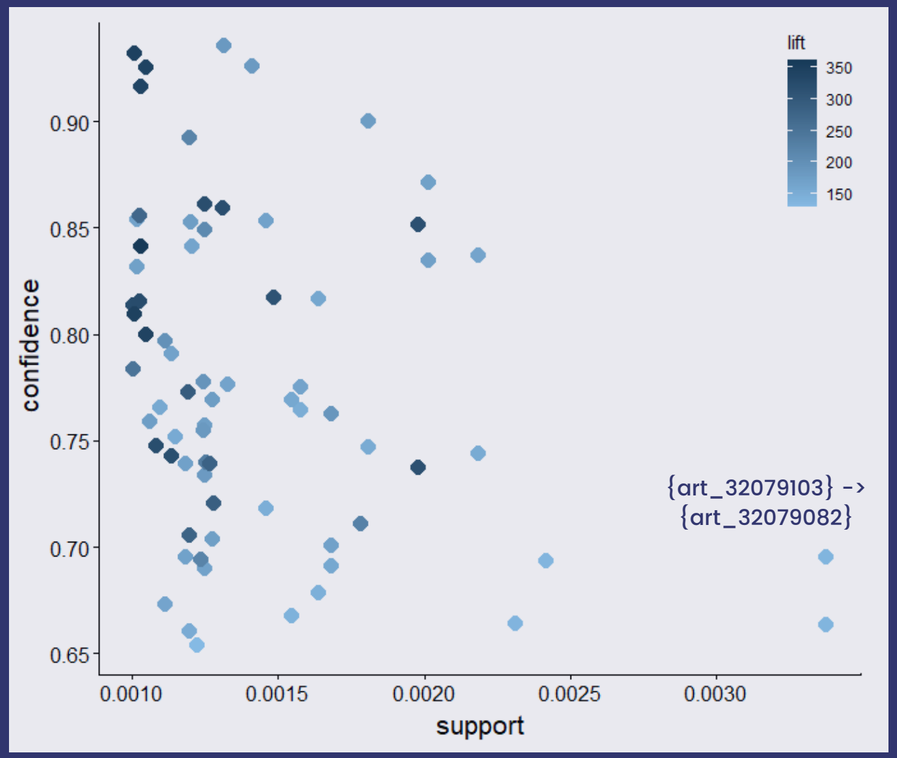


Market Basket Analysis - Results

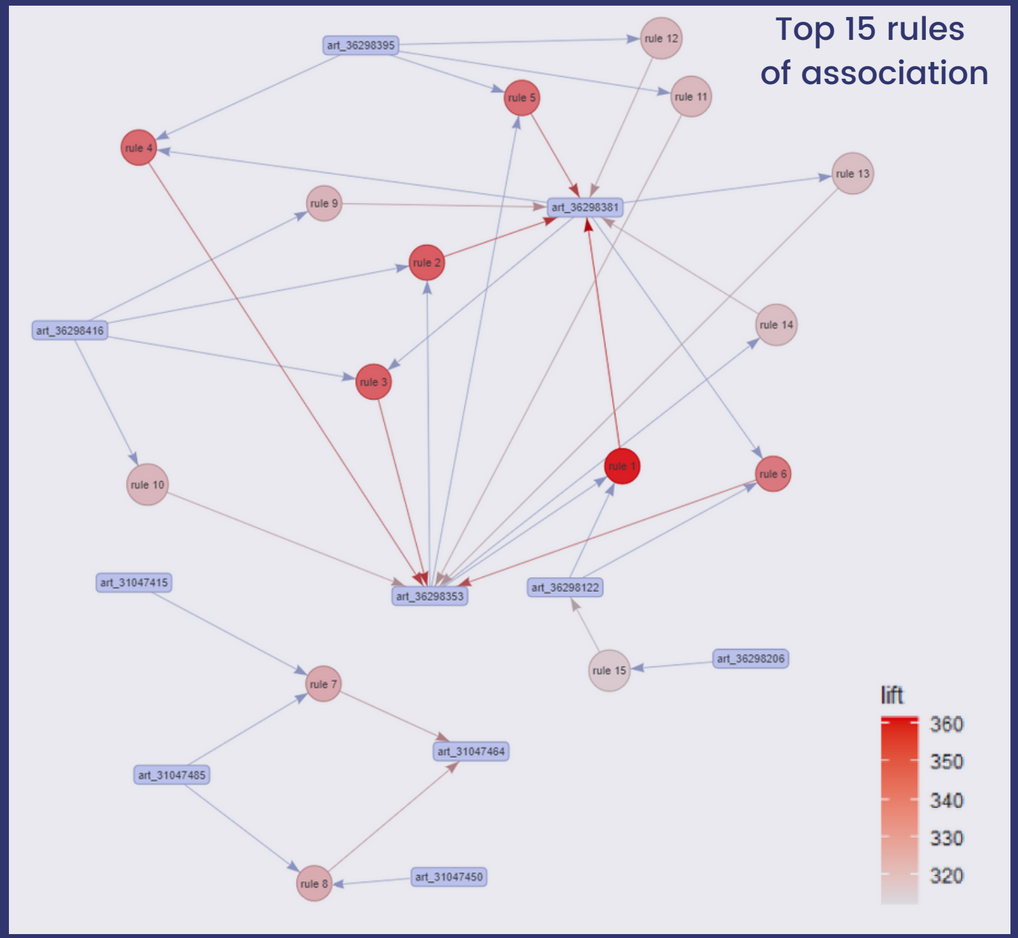
Affinities among products with high support (>0.1%)

Associations with confidence >65% & support >0.1%

71 Non redundant rules between sets of two & three products



Article 36298381 & 36298353 are the most 'driven' products



Market Basket Analysis - Results

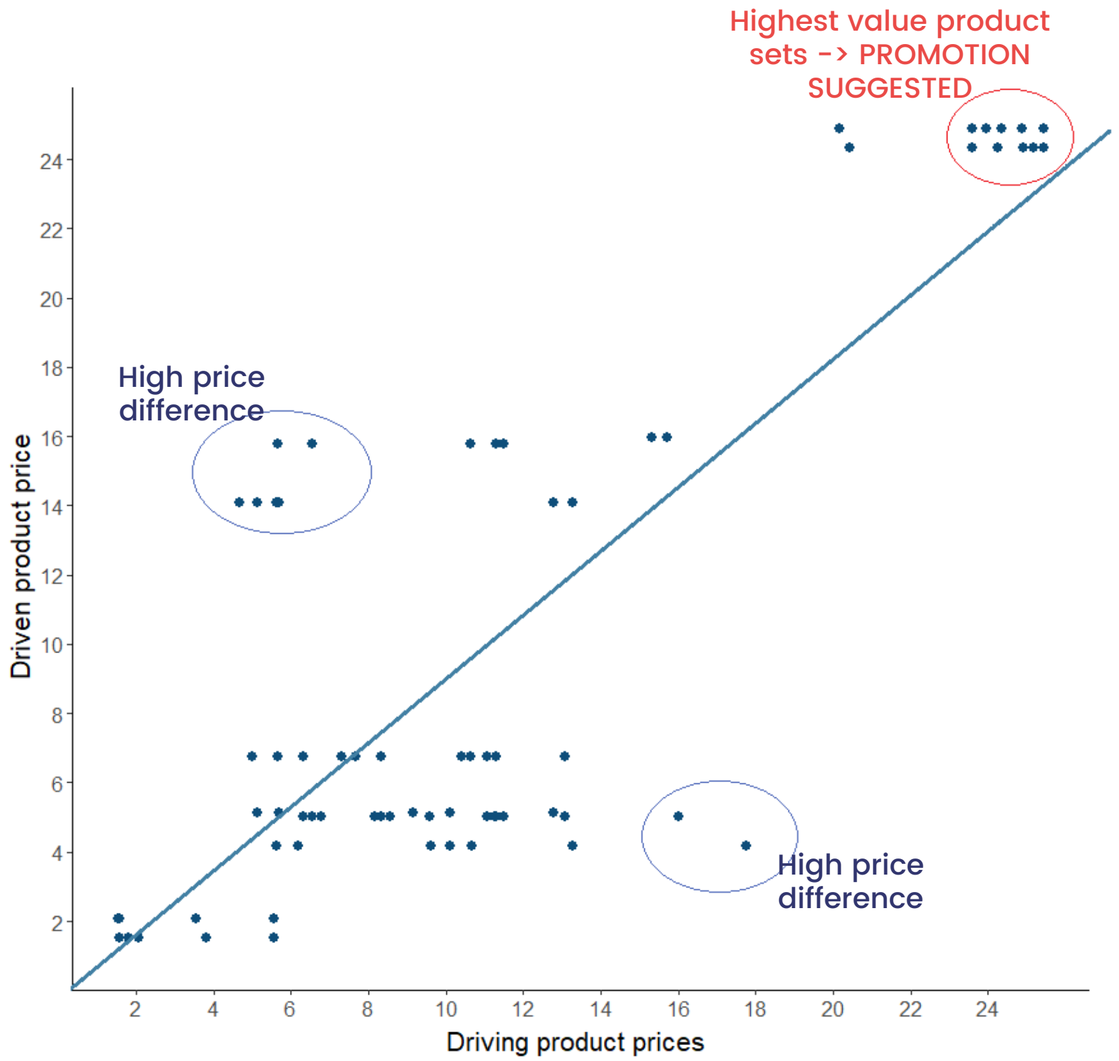
-> Meaningful affinities between **department #3** products **only**

-> The majority of the product sets have **driving** and **driven** products **around** the **same price**

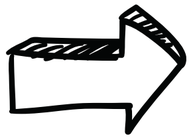
-> Only **few sets** present a **high price rate** between driving and driven products (ex. {art_31618300, art_31618321, art_31618405})

-> **Cross selling actions** focused on **high value product sets** (ex. {art_36298395, art_36298381}, having respectively a 25.43$ and 24.35$ **average** price)

-> **Layout optimization** for strategic collocation of products in **department #3**

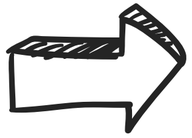


Summary Marketing Strategies



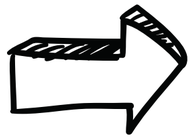
Up-selling actions personalized for each high value segment Customer

Optimization of the relationship with customers can be obtained via personalized up-selling actions before and during the purchasing moment. This strategy can be implemented to increase the value, the life time and the satisfaction of already valuable customers.
Loyal customers are indeed the most susceptible to product recommendations.



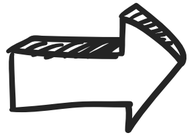
Level up: from Tin and Copper to Bronze

The majority of the customer base resulted to fall in the "Tin" and "Copper" segments.
It would be valuable to shift as many customers as possible to more profitable segments.
This can be achieved engaging the customers by making limited time offers and recommendations based on past purchases. Other effective strategies include: providing on-boarding support, offering free trials, sending personalized emails to reconnect, offering renewals.



Potential churner forecasting based on the developed model and consequent caring actions

Retaining potential churners is less expensive than acquiring new customers -> It is crucial to implement caring actions not to lose them. Examples of potential targeted actions: addressing special offers, rationalization and diversification of the marketing contact strategy (either via call-centers, e-mails, etc.) according to the risk of churning.



Cross selling actions focused on high value product sets

Recommendation of products based on the found association rules, to integrate "next-best-product" to buy actions.
Cross selling can be implemented both in the pysichal store and in the website, by means of specific e-commerce actions