

Web Marketing and Communication Management

A project by:

- Emanuela Elli (892901)
- Federica Madon (825628)
- Tommaso Strada (829351)



Business questions



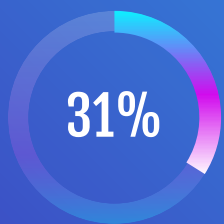
1) *Customer* relationship *valorization*:

- How do we increase the relationship profits?
- Which are the customer base features? How can it be divided?
- Which are the most valuable customers?
- Who are the likely churners? What is the reason for this?

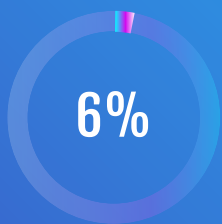
2) *Products valorization*.

- Which are the most purchased items?
- Which are the valuable items combinations?

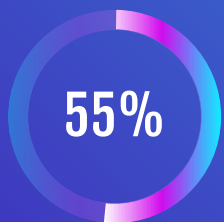
CUSTOMER EXPLORATION



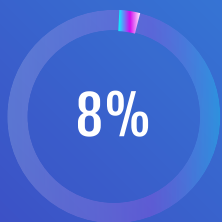
PREMIUM



PREMIUM_BIZ

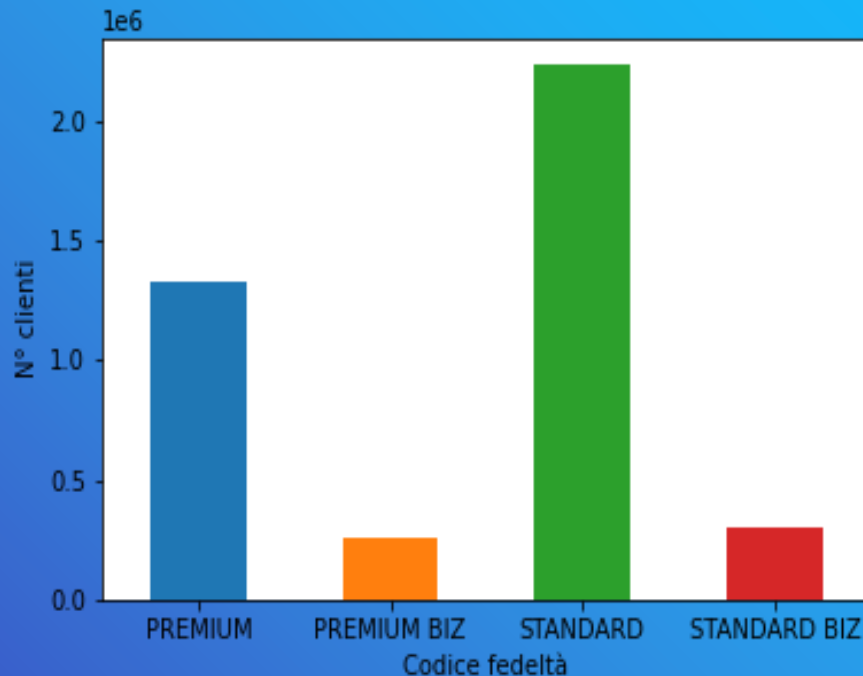


STANDARD



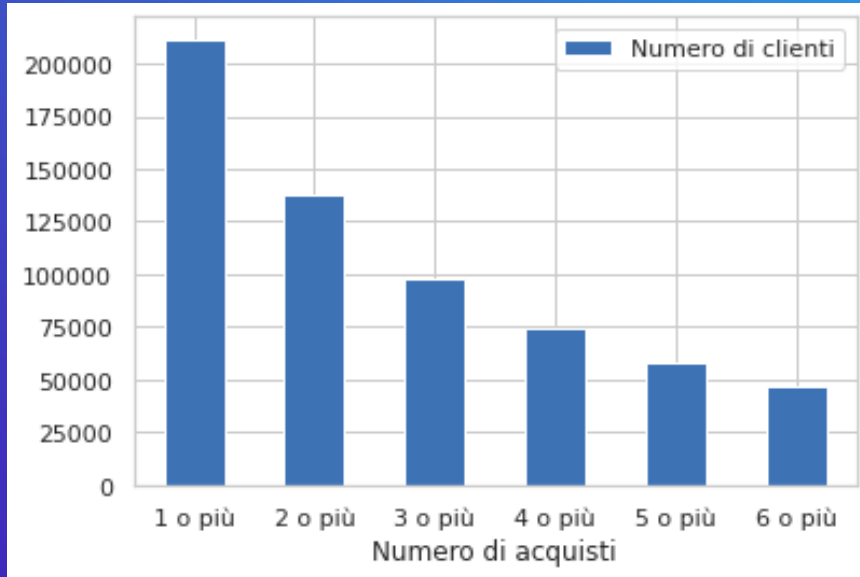
STANDARD_BIZ

Customers fidelity class count

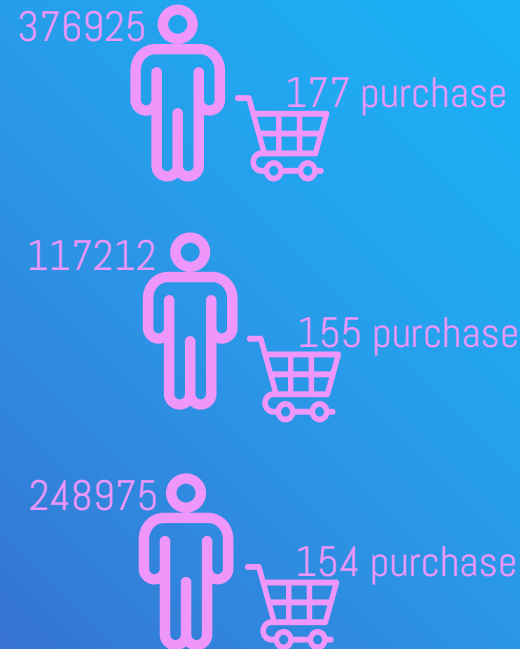


CUSTOMER EXPLORATION

Customer distribution for number of purchases



The top 3 customer for purchasing items are:



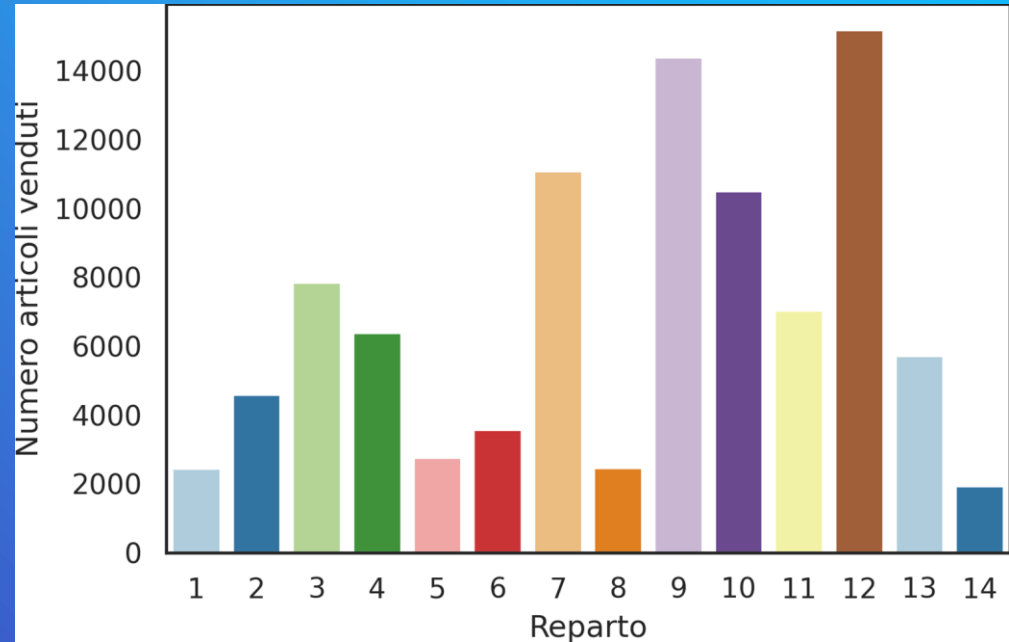
PRODUCTS EXPLORATION

After the data preparation step, there are:

- 902994 different bills
- 95826 different items sold among 14 departments

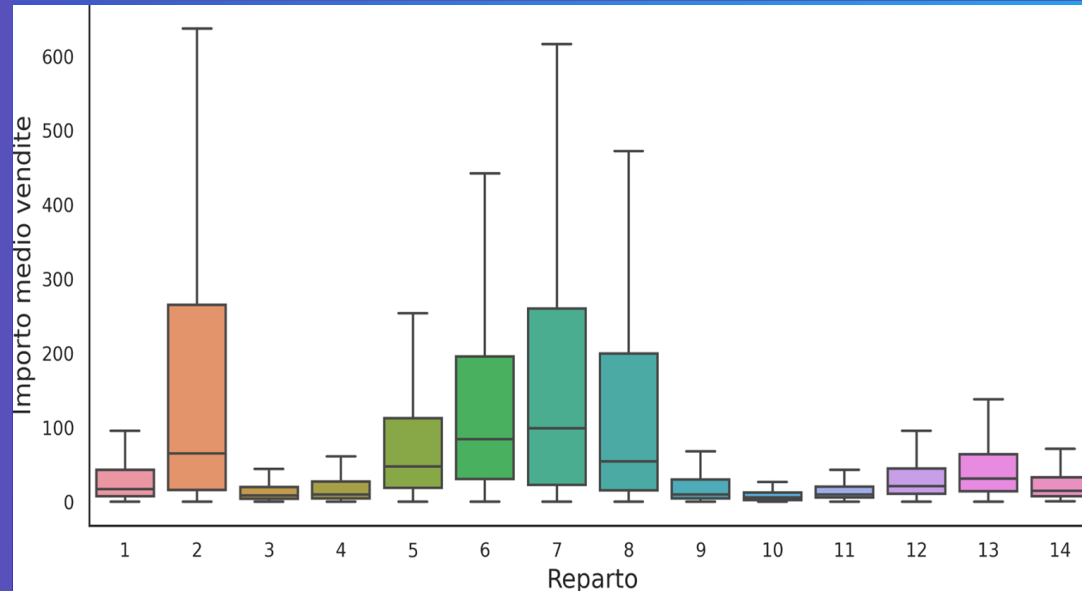
The most sold items belong to the 12 department

Product distribution for departments



PRODUCTS EXPLORATION

Average sales amount for departments



The department 7 items have the highest median of items sold.

DATA DRIVEN-STRATEGY

Business questions



Models



Actions data driven

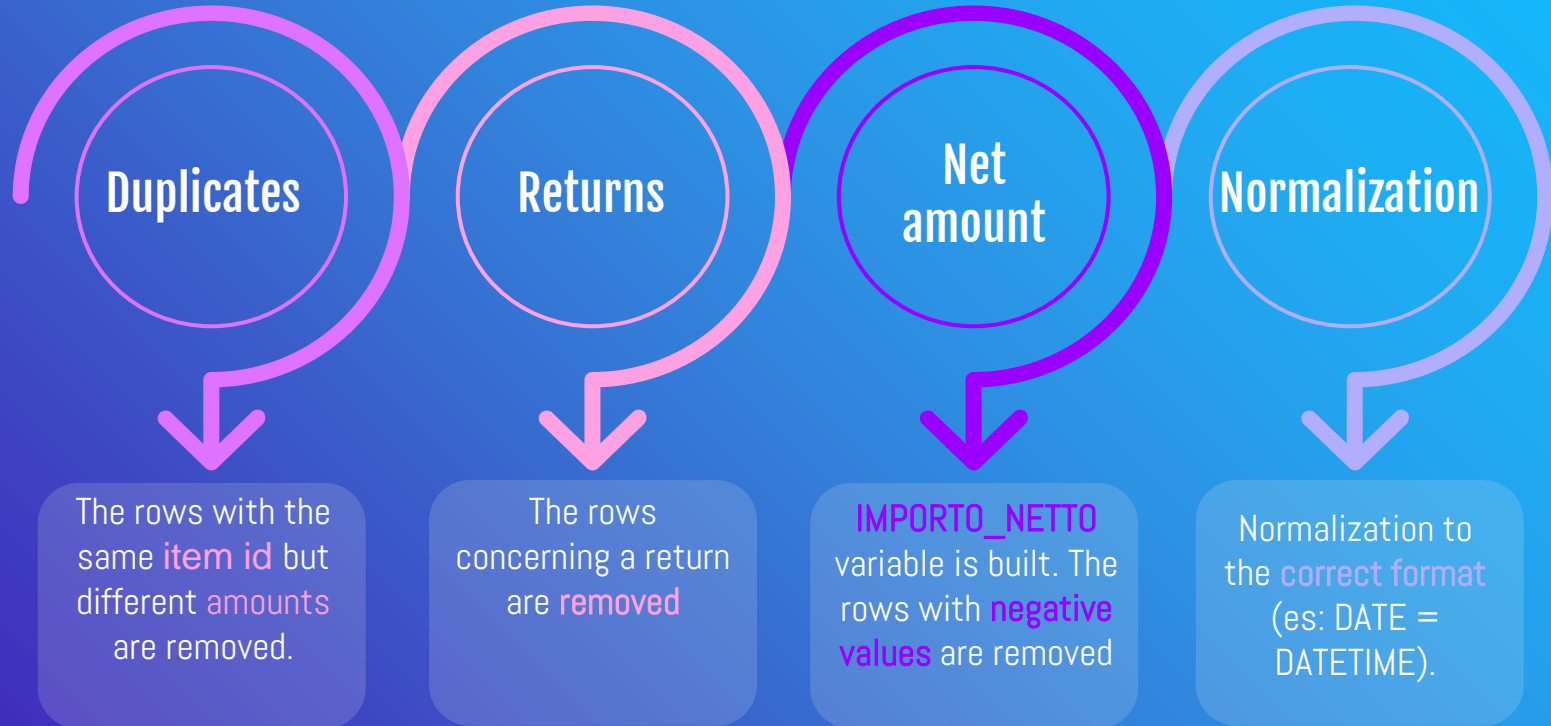


Data cleaning

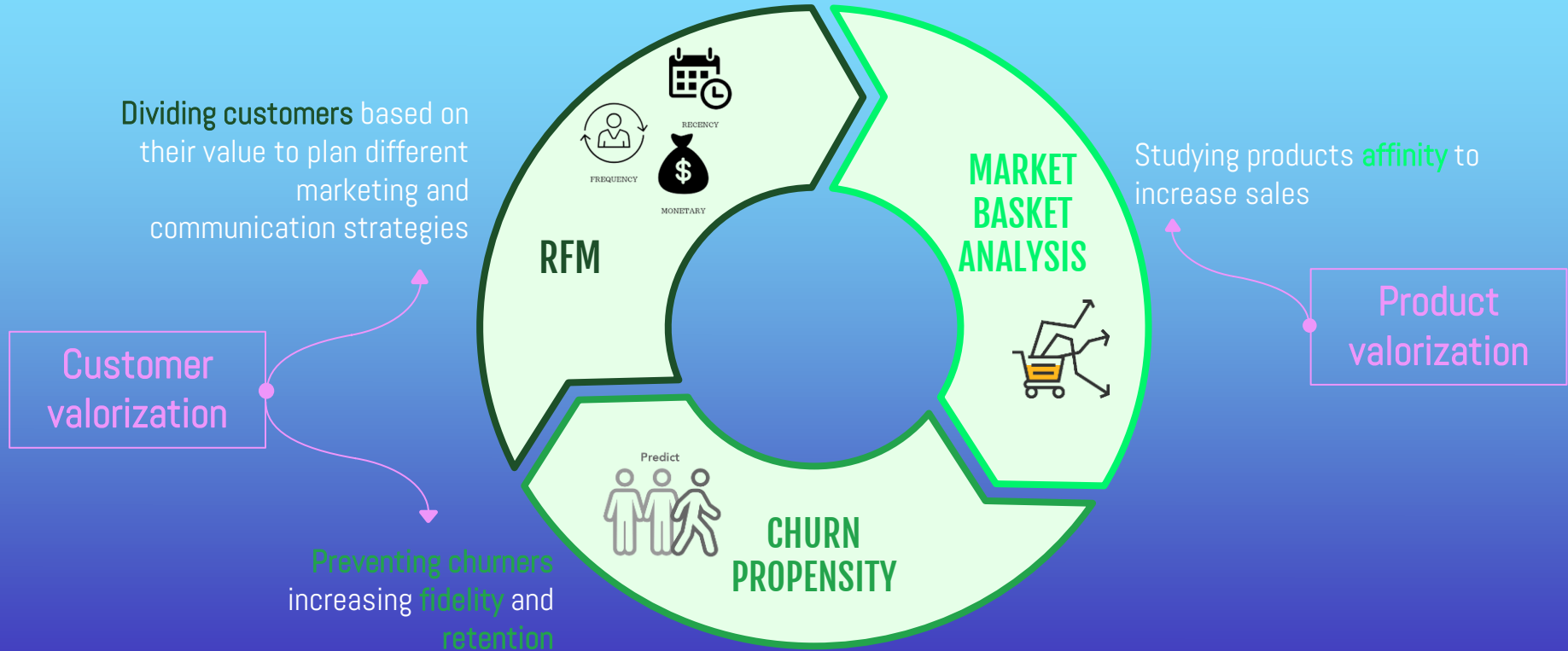


Evaluation

DATA PREPARATION



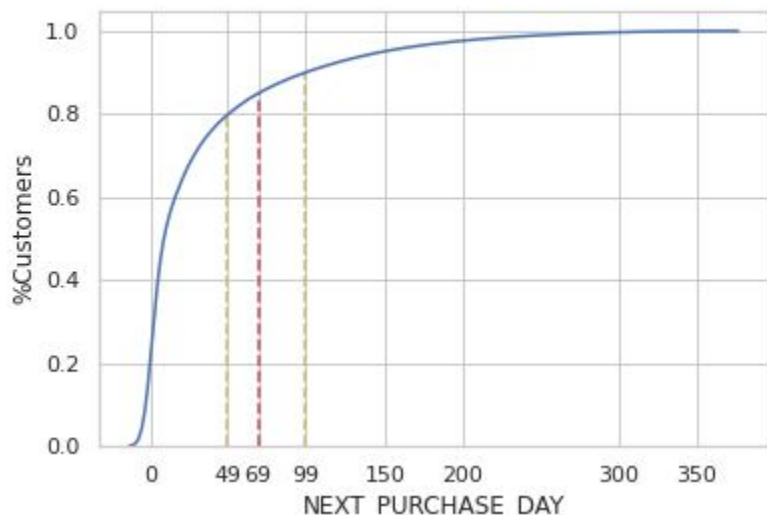
MODELS



RFM PREPARATION

Next Purchase Day Curve

85% of users buy within 69 days



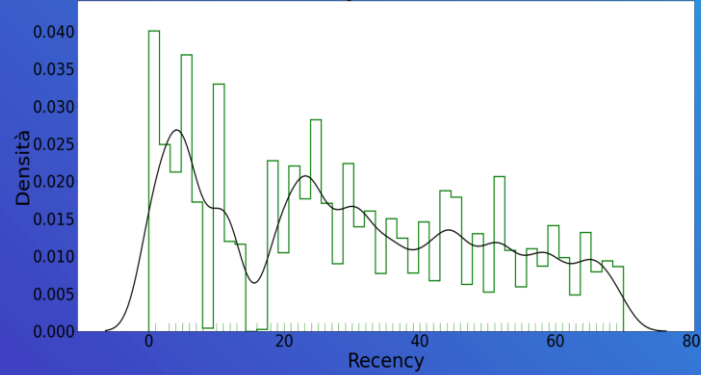
Preprocessing

- ❑ Customers splitting into active and inactive. The cut point is fixed on 2019-02-19. It is equal to 69 days before the last available date.
- ❑ Measure building: Recency, Frequency, Monetary Value
- ❑ Building RF Matrix e RFM Matrix.

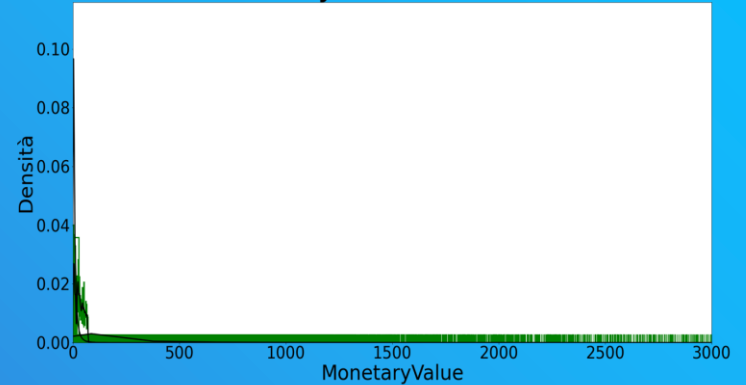
RFM

Distributions

Recency Distribution



Monetary Value Distribution

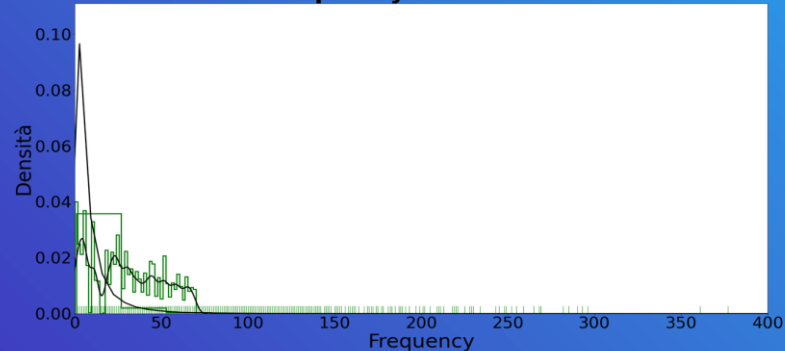


Time passed by since the last purchase or interaction with the item.

Total transactions

Total or average value of transactions

Frequency Distribution



Recency



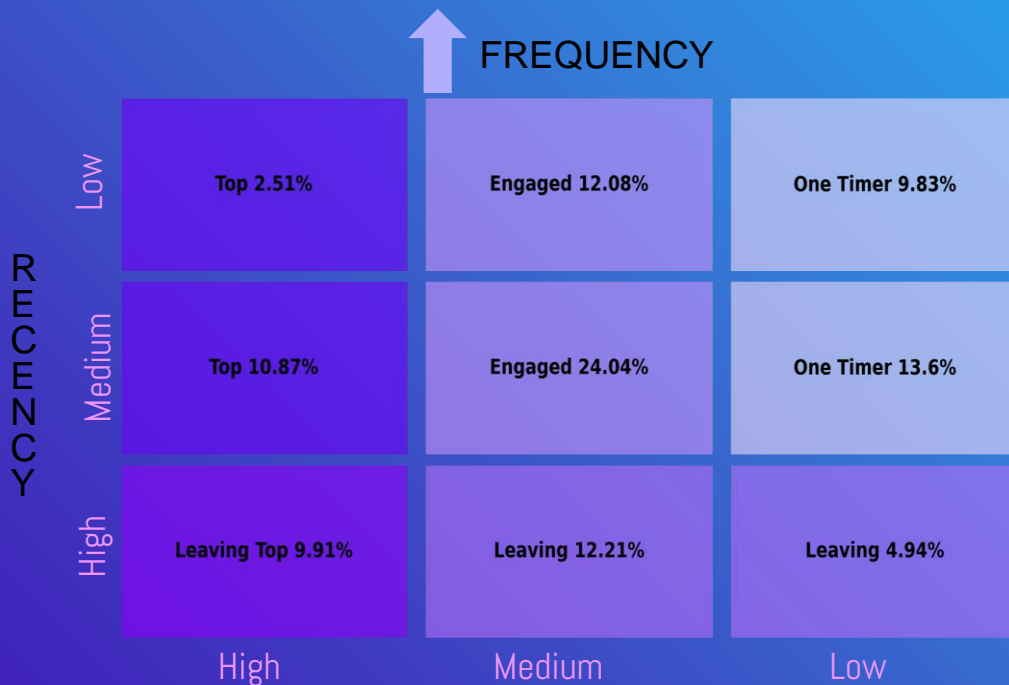
Monetary



RFM

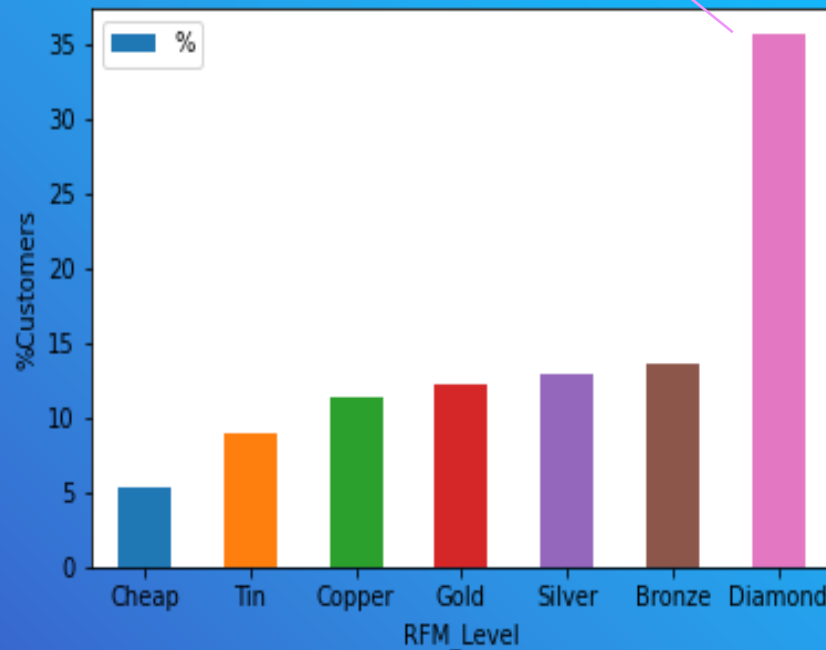
RF Matrix

Engaged(36.12%): Customers who recently buy with frequency on average.



RF Matrix

Diamond(35.56%): Customers with high Monetary value and Top RF



RFM INSIGHTS



HIGHER VALUE CUSTOMERS

Keeping this kind of customer, giving valuable items.



MEDIUM VALUE CUSTOMERS

Recommending other products and suggesting fidelity programs.

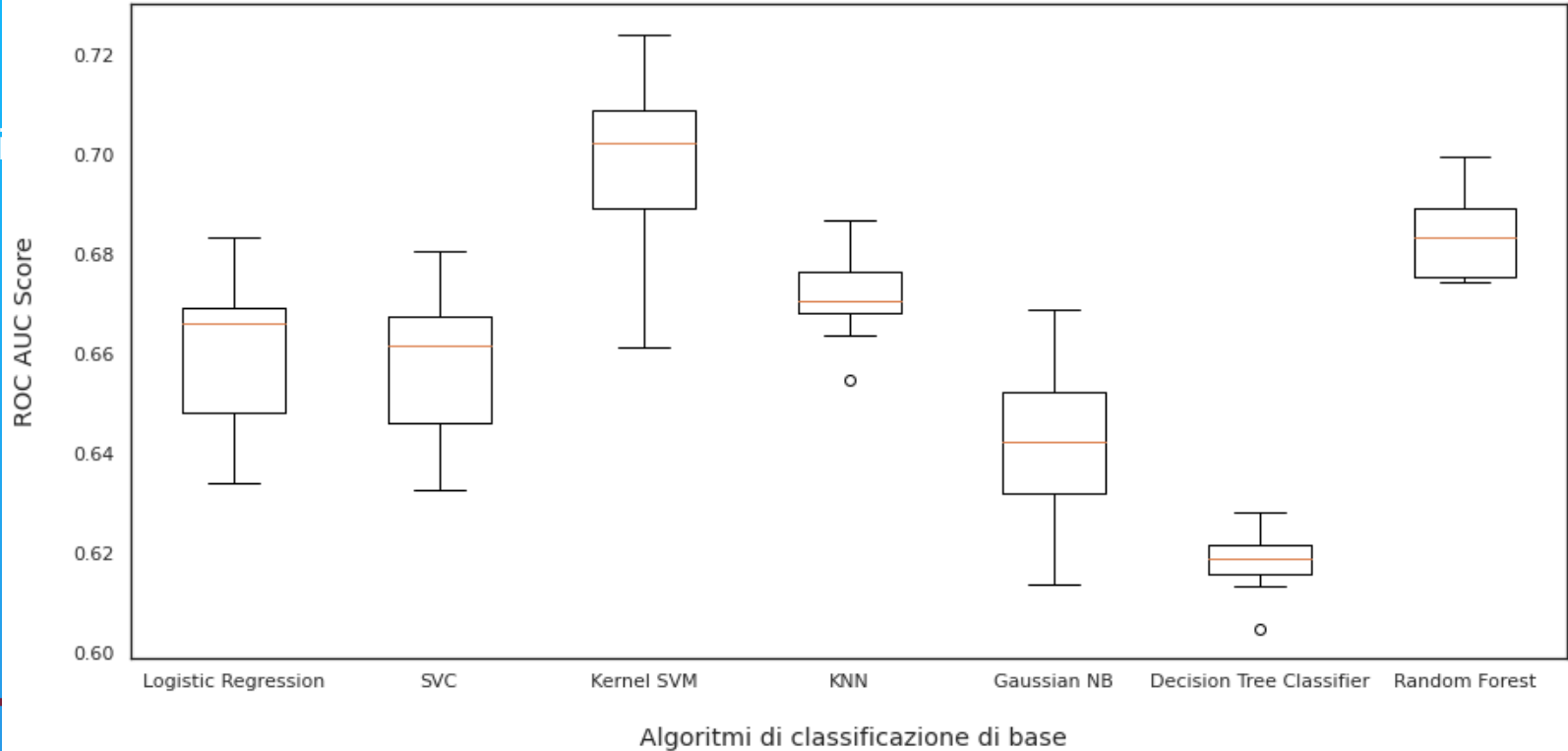


LOWER VALUE CUSTOMERS

Engaging this type of customer and increasing the retention

CHURN PROPENSITY

ROC AUC Comparison



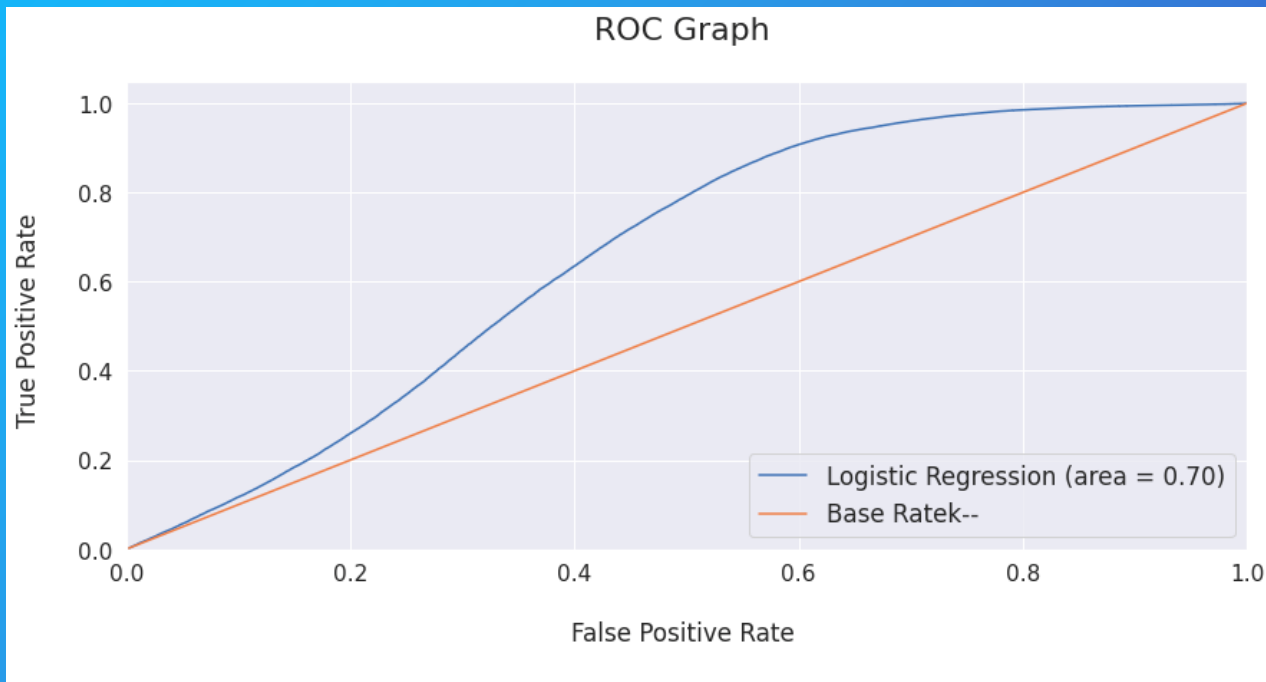
S
d
1



CHURN PROPENSITY

The **logistic regression** is an optimal model since it has the highest relative combination of *precision*, *recall* & *F1 score*.

The model is evaluated by ROC curve. It shows the model's capacity to distinguish the classes using the AUC average score.



The **orange line** shows the ROC curve of a random classifier (a good classifier stays far from this line).

As the figure shows, the optimized logistic regression gets a higher AUC score than the random model.

MARKET BASKET ANALYSIS

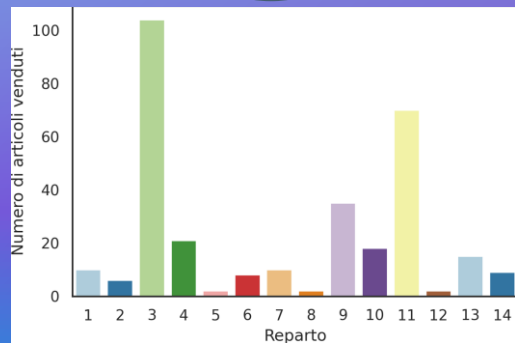


Overview of high support articles

188938 different bills- 62582 different products

Only items with support level higher than 0.1% are selected

There are 312 items with a high support level



DateTime starts from 19/02/2019 for being coherent with the other models

MARKET BASKET ANALYSIS

390 interesting itemsets with support $> 0.1\%$ belong to the third department

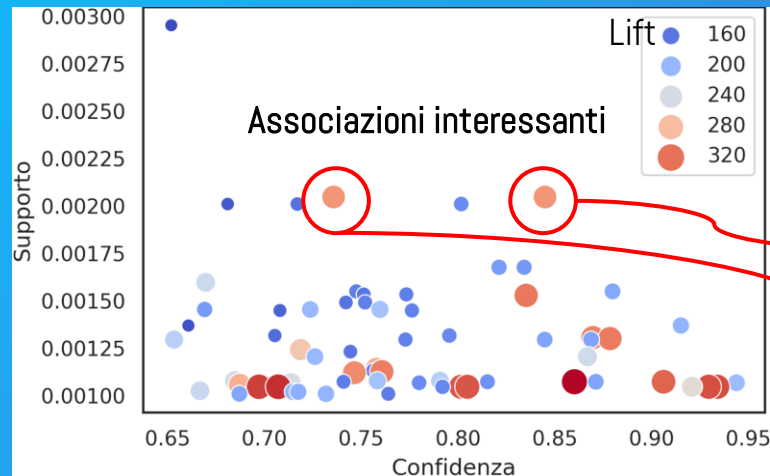
Apriori Algorithm

Association Rules

Pruning rules

Association rules

66 relevant associations with confidence > 0.65 and lift > 1



36298353→36298381

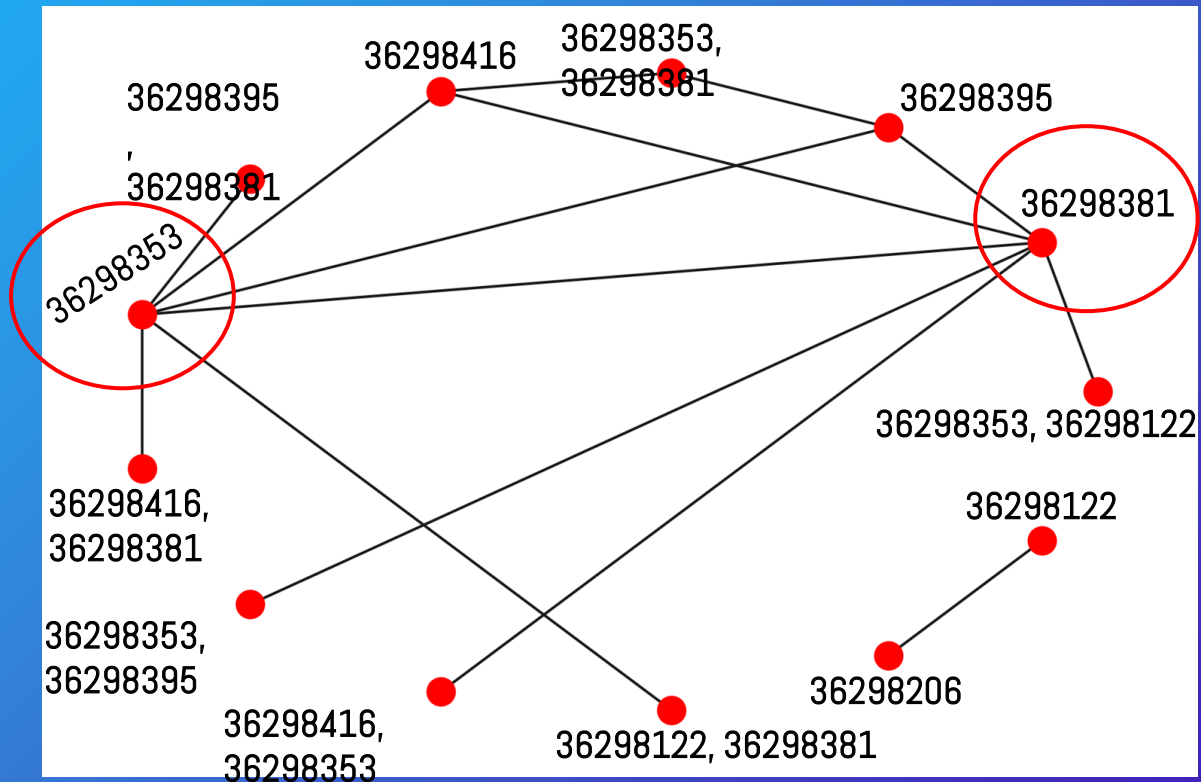
36298381→36298353

MARKET BASKET ANALYSIS



- Items **36298353** and **36298381** are the most "driven" products.
- Item **36298353** is the most popular among the items with the highest support.
- The association(**36298353**, **36298122**) → 36298381 is the one with the highest lift (> 354)
- The customers following the rule with the highest lift, buy the three products every **17 days** on average

Top 15 association rules



DATA DRIVEN ACTIONS



Move Silver and Bronze customers to Gold level

Following the Gold level, the majority of customers are divided into Silver and Bronze levels. It is convenient to engage this type of customer, increasing loyalty and developing up/cross-selling actions (creating personalized promotional item bundles in both physical and online stores). Some possible strategies include customizable email campaigns, sales offer based on customer habits, developing a strong relationship with customers beginning with the onboarding process, and offering rewards.



CRM

Customer relationship management to boost loyalty, advocacy, and sales opportunities



Layout optimization

Optimize product allocation by dividing items into leading and pull categories.



Churners prediction and caring action

Keeping potential churner costs less than acquiring a new one. Indeed, it is critical to plan and implements caring actions in order to avoid losing them. Some specific actions include: providing special offers, rationalizing and diversifying the contact marketing strategy based on churner risk, and so on..



Insight integration in the marketing automation

To supplement the contact strategy with the 'next best product to buy' based on the relationship between the leading and pulling categories.



Promotions

Promotions can be defined by calculating the average latency for certain product associations.



Thanks for your time!