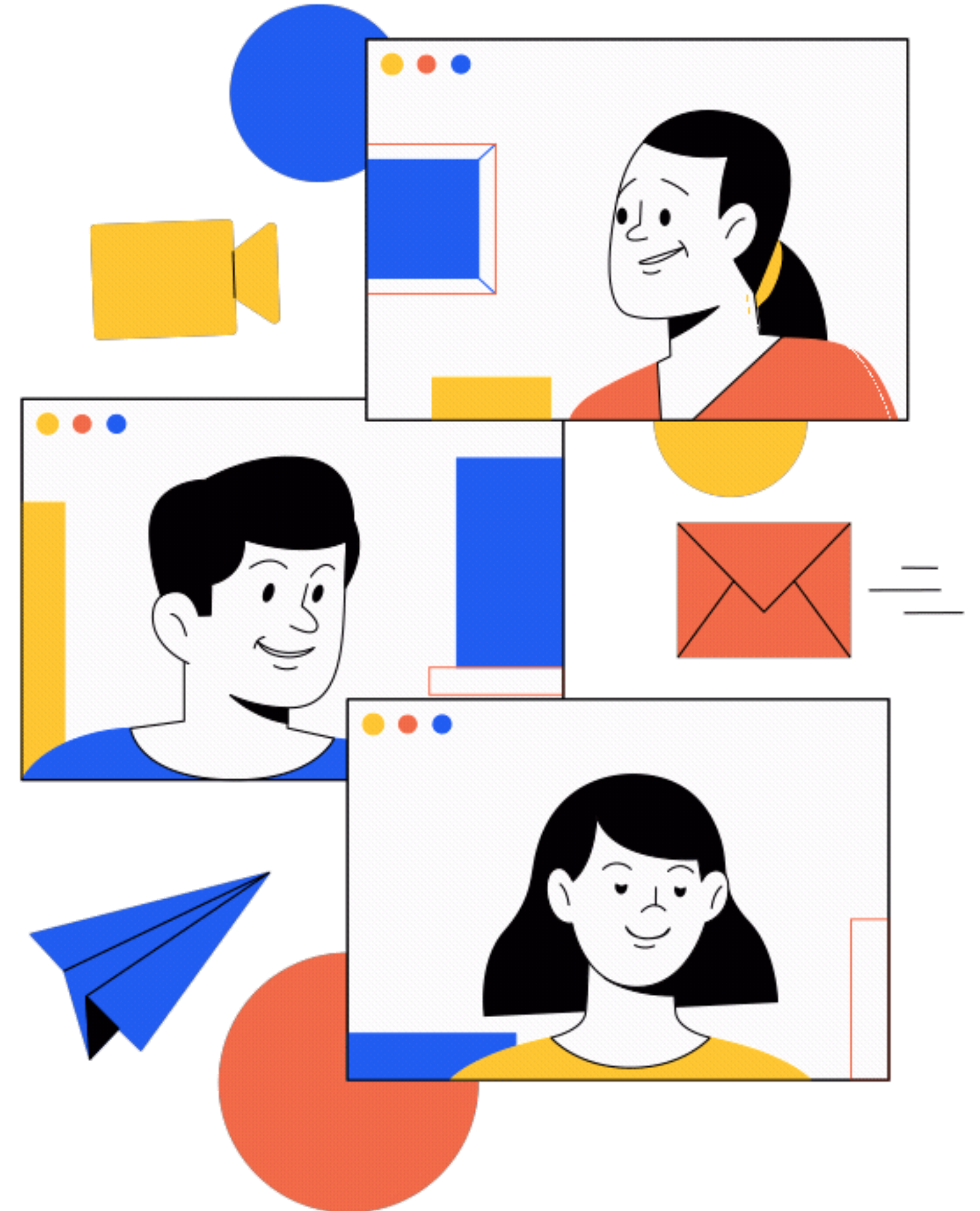


MARKETING ANALYTICS

Analyze business data, get insights and make actions!



MAIN CONTENTS

- 1 Business description
- 2 RFM analysis
- 3 Churn analysis
- 4 MBA
- 5 Sentiment Analysis





THE BUSINESS

We are a supermarket chain with physical markets and a e-commerce online. The chain's name is *Super, the great S*. We offer the client a large variety of products, from fresh fruit and vegetables to pet's food. All the products we offer in shop are also available on-line and, once bought, they can be shipped directly to the client or he can choose to pick up his groceries at designated pick up points.



MORE INFOS!

- **A lot of data**

We keep track of everything we sell (client's orders) and we cataloged all the items via an unique id. Once they receive their order, they can leave a review, both negative and positive.

- **Clients**

To keep everything running, also each client has an unique id and we collect different information about them, such as their address, age and favorite physical Super store.

- **Services**

Furthermore, clients have the opportunity to sign up to one of our loyalty programs, that guarantee some benefits (free shipment, special offers, ...) that vary on the specific program they chose.

Given this context, the goal of the project is to **analyze the data available** to perform some well-known marketing tasks and provide some **taylor made marketing strategies** based on the results of these analyses.

RFM ANALYSIS

RFM stands for "**recency, frequency, and monetary value**". The idea is to segment customers based on when their last purchase was, how often they've purchased in the past, and how much they've spent overall.

It allows us to **target specific clusters of customers** with communications that are much more relevant for their particular behavior - and thus generate much higher rates of response, plus increased loyalty and customer lifetime value.





RFM STEPS 1/3

- **Recency**

It measures the time elapsed since a customer's last transaction. Customers who have made a recent purchase are generally more engaged and likely to respond to marketing efforts.

- **Frequency**

It provides valuable insights into customer engagement and purchasing habits. Customers with a high frequency value are those who make frequent transactions, indicating their active involvement with our supermarket.

- **Monetary**

It provides a measure of a customer's overall value and contribution to revenue. Customers with a high monetary value are those who have made significant purchases, indicating their higher spending potential and greater profitability.



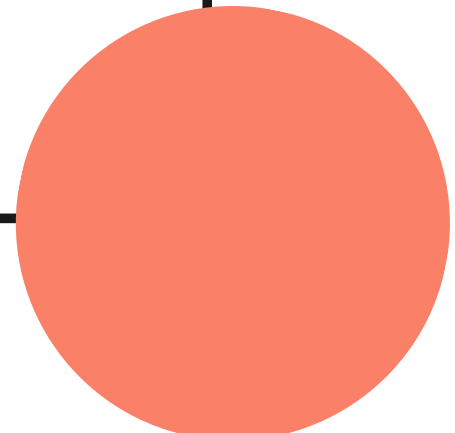
RFM STEPS 2/3

- **Active customers**

We have to define our active customers. In this context, **active customers** are clients that have purchased at least once during 2023.

- **RFM Variables**

For each customer we have defined **3 variables**:

- *Recency*: the time elapsed since a customer's last purchase;
 - *Frequency*: the number of times a customer has made purchases;
 - *Monetary*: the total amount of money spent by the customer;
- 



RFM STEPS 3/3

• RF Matrix

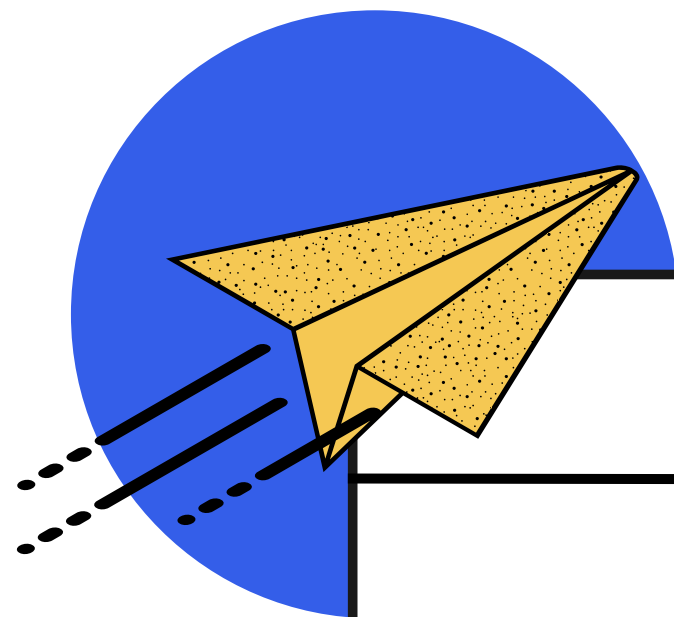
We have segmented the active customers in **5 categories** according their Recency and frequency values:

- 0: One-timer
- 1: Leaving
- 2: Engaged
- 3: Leaving Top
- 4: Top

• RFM Matrix

We have segmented the active customers in **7 categories** according their RF Matrix and Monetary values:

- 0: Cheap
 - 1: Tin
 - 2: Copper
 - 3: Bronze
 - 4: Silver
 - 5: Gold
 - 6: Diamond
- 

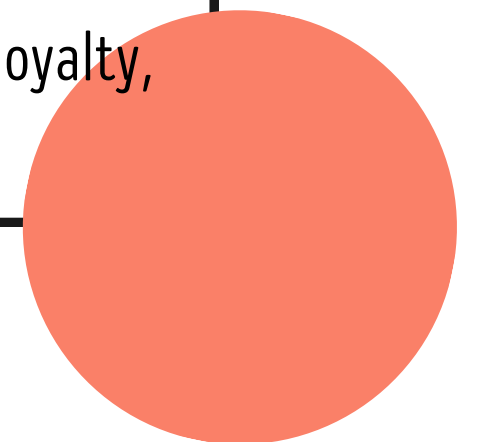


RESULTS

- **Customers segmentation**

The output of the RFM analysis is the **RFM matrix**, which provides a comprehensive view of customer segmentation based on recency, frequency, and monetary value. The RFM matrix categorizes customers into distinct segments, allowing us to tailor our strategies to meet the specific needs and behaviors of each group.

Each customer is assigned a combination of **RFM scores** that indicate their level of engagement and value to our supermarket. This segmentation enables us to develop targeted marketing campaigns, personalized offers, and customer retention initiatives that are designed to maximize customer satisfaction, loyalty, and overall business growth.



RFM MARKETING STRATEGIES

Action 1

Personalized Loyalty

Programs: Develop a tiered loyalty program that offers exclusive benefits to high-value customers. This could include special discounts, early access to sales, personalized offers tailored to their preferences, and invitations to exclusive events.

Action 2

Personalized Communication:

Establish direct and personalized communication channels with our high-value customers. Utilize targeted emails or SMS messages to showcase special promotions, discounts, or new product launches that align with their preferences. It is essential to utilize their preferred communication channels to ensure maximum message reception and engagement.

Action 3

Special Events and

Workshops: Organize exclusive events, workshops, or cooking classes specifically tailored for our high-value customers. These events could feature tastings, demonstrations by renowned chefs, or educational sessions on nutrition and healthy cooking. By creating an exclusive community for these customers, they will feel a sense of belonging and importance, making it highly unlikely for them to desire to disengage from it.

Best Action

Among the three ideas, we acknowledge that **the third one** may require a higher financial investment. However, it holds the potential for the highest level of engagement as it fosters the development of a dedicated community. This strategy aims to make our customers feel special and connected, leading them to be strongly motivated to remain a part of the community.



CHURN ANALYSIS

Churn rate (sometimes called attrition rate), in its broadest sense, is a measure of the number of individuals or items moving out of a collective group over a specific period.

The purpose of the churn is to **predict which customers are most likely to abandon** and allow us to intervene with appropriate retention initiatives.



CHURN STEPS 1/2

- **Active clients definition**

We have to define our **target variable (churn)**, so we want to analyze people who have purchase at least two times (not occasionally) and consider the last time that they purchased. In this context, **active customers** are clients that made more than one purchase.

- **Churners definition**

We set the **reference date** (2023-01-01), the **lookback period** and the **holdout period**. If a customer *did not complete at least two purchases* in the interval [lookback period, holdout period], **we consider him as churner**. So the variable churn contains a **boolean value**, where True means churner and False means not churner.



CHURN STEPS 2/2

- **Preprocessing**

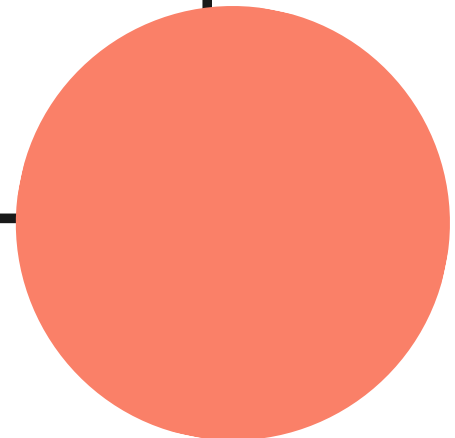
Many techniques were applied:

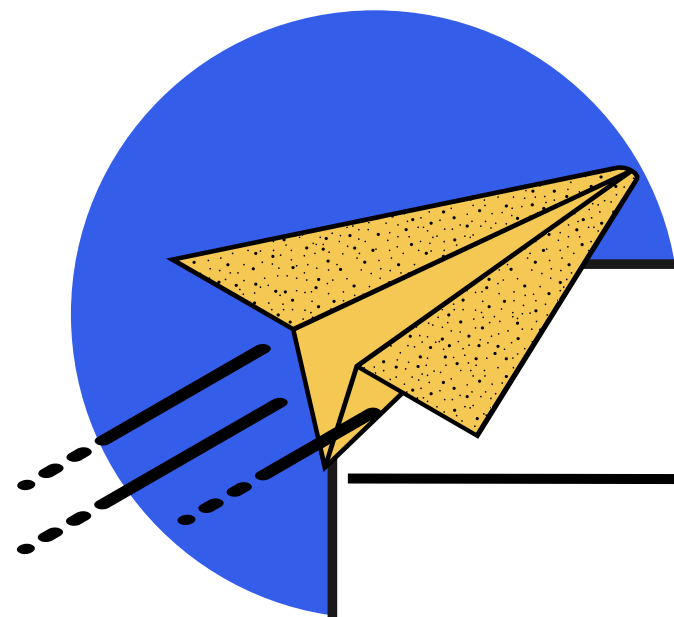
- Merging more tables
- Removing negligible columns
- Values replacing: from categorical to integer
- Oversampling: SMOTE
- Train, test and validation splitting

- **Models training**

We applied different models to verify which one could be the best option.

In particular, we implemented a *Logistic regression*, a *SVM* and a *Decision Tree classifier*.





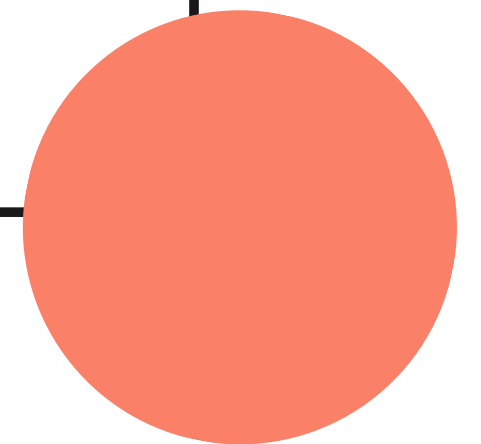
RESULTS

- **And the best model is...**

The best model is the **Decision Tree Classifier** (accuracy **0.79**) so we used it for the predictions.

Logistic regression accuracy: 0.70

SVM accuracy: 0.70



CHURN MARKETING STRATEGIES

Action 1

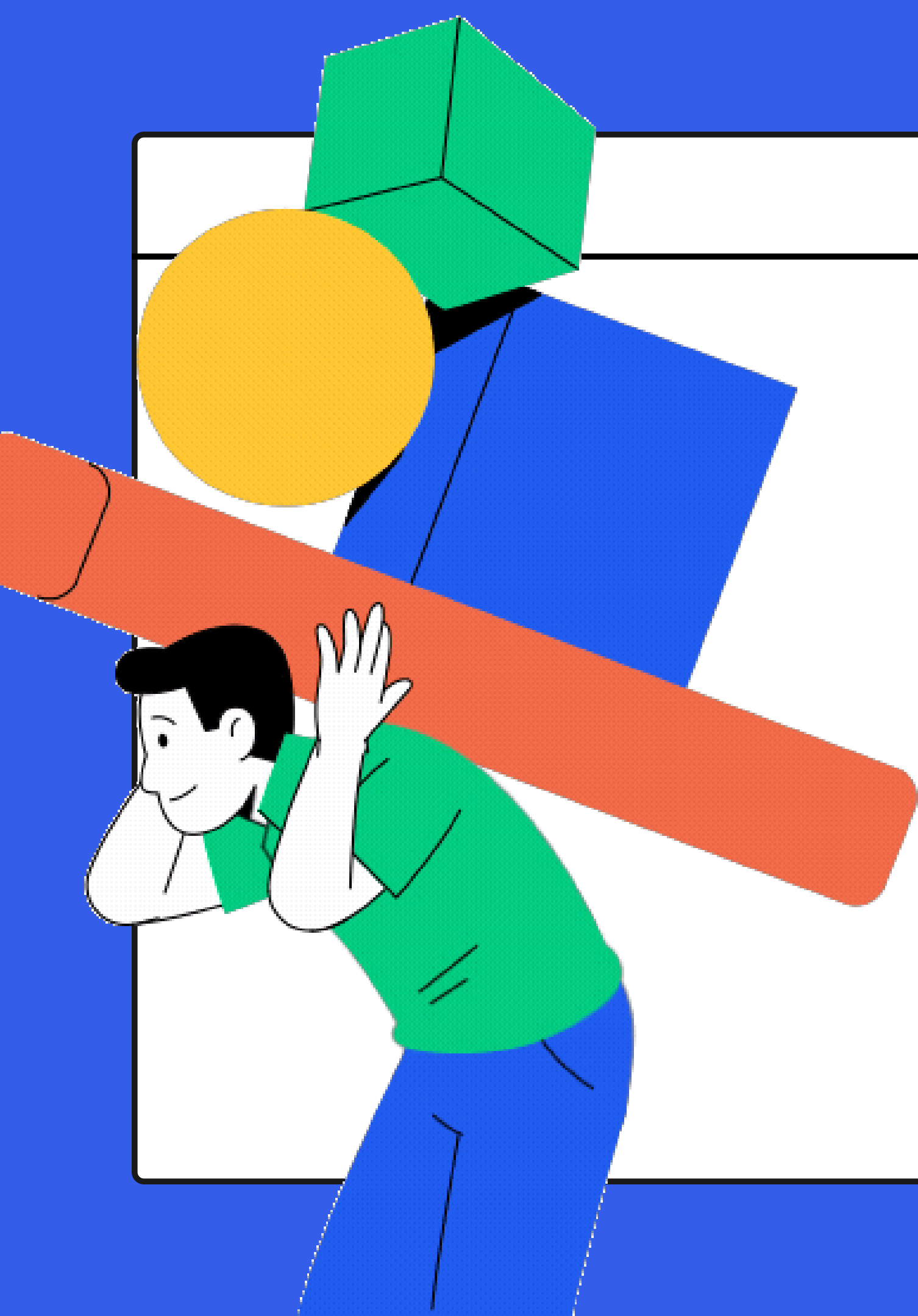
In general, we are interested to retain customers that are already clients, so we could implement some strategies that has to goal to keep the customer satisfied with our products/services. In this scenario, we could send clients regular discounts and coupons.

Action 2

We could engage more with customers and ask for a feedback more often.

Action 3

If churn happens, we could analyze why it happened and eventually offer the client some benefits in order to retain him, such as discounts, a better subscription plan (that is the loyalty plan) at a more convenient price.



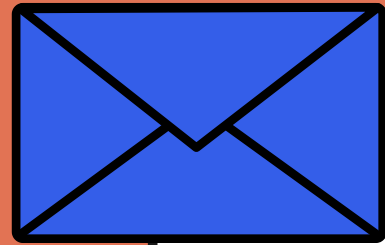
MBA

MARKET BASKET ANALYSIS

It is a **data mining technique** that is used to **uncover purchase patterns** in any **retail setting**.

MBA looks for **what combinations of products most frequently occur together in transactions**. These relationships can be used to:

- increase profitability through cross-selling,
- recommendations and promotions
- the placement of items on a menu or in a store.



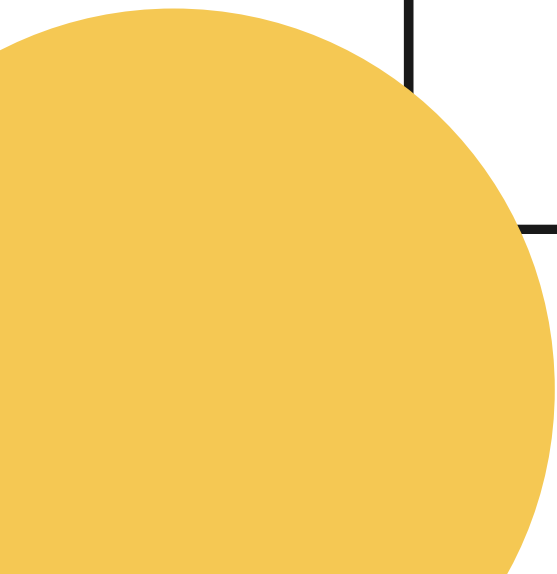
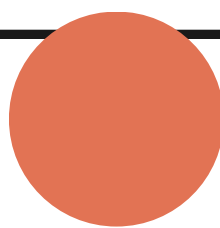
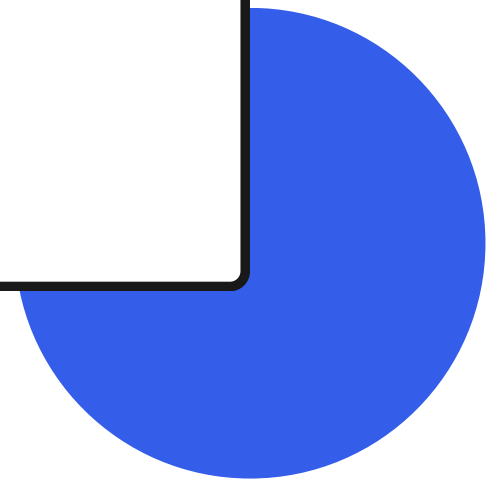
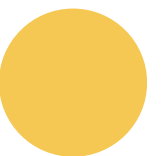
MBA STEPS 1/2

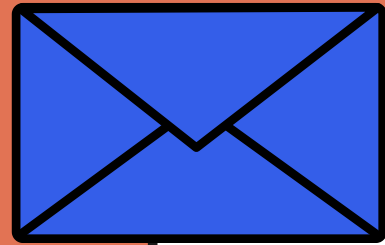
- **Data**

For this task we only need the **orders** table, that contains information about **what** was purchased, **when** and **by whom**.

- **Preprocessing**

We proceed by making some **data manipulation** and creating a cross table: **one row for each order** and **one column for each product**. The values in the table ranges in $[0,1]$, where 1 indicates that a specific item was bought, 0 otherwise.



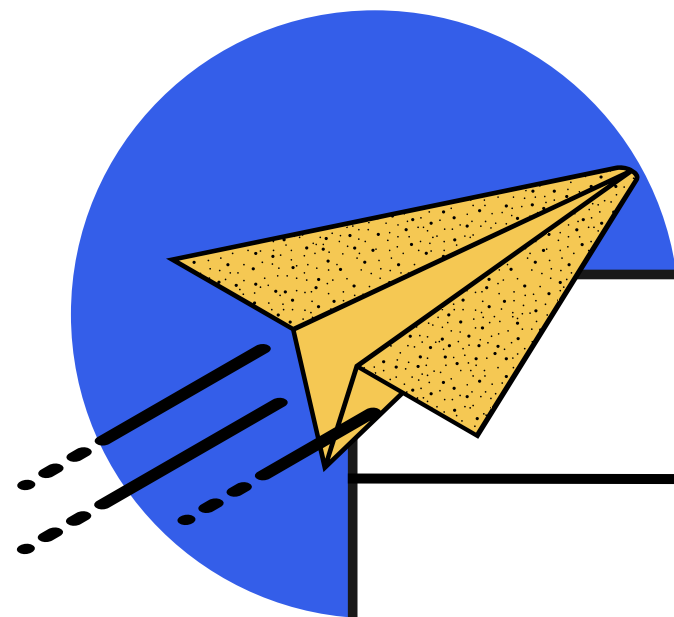


MBA STEPS 2/2

- **The Solution**

The Apriori algorithm creates an itemset is considered as "**frequent**" if it meets a user-specified support threshold. In this case we set the threshold at 0.003 therefore a frequent itemset is defined as a set of items that occur together in at least 0.2% of all transactions in the database.

Then we create the **association rules**. An association rule is a statement of the form "*if item A is present in a transaction, then item B is also likely to be present*". The strength of the association is measured using the confidence of the rule, which is the probability that item B is present given that item A is present.

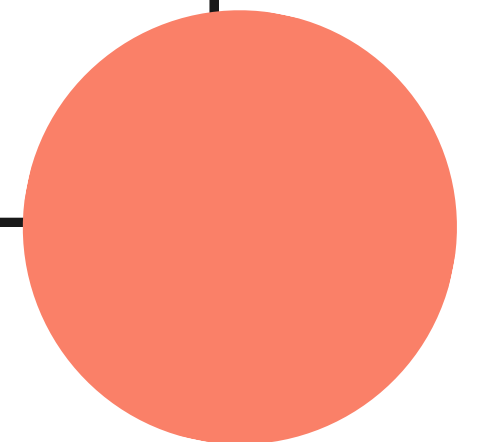


RESULTS

- **Lift value**

The lift value is the **metric we used for the evaluation of our rules**. The higher the lift value, the more a rule is considered to be effective in predicting the relationship between different product patterns.

All the rules we found have Lift value > 1 .



MBA MARKETING STRATEGIES

Action 1

We could choose to put items that relate to each other in a closer shelf, so that they can be found more easily by the customer.

Action 2

We could offer a discount on the consequent if the client buys the antecedent product.

Action 3

We could offer and recommend the client the antecedent at a lower price.



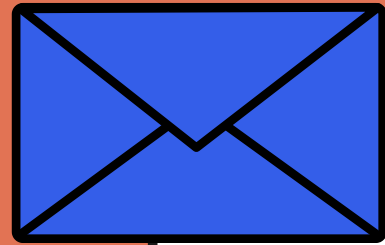
EXTRA STRATEGY

RFM + MBA

It could be useful to **provide only some clients more interesting offers**. For example, we can offer **high-value customer** specific offers based on their personal shopping cart.

For this reason, we can **integrate RFM and MBA analyses** to gain a comprehensive understanding of customer behavior and preferences, develop a customer segmentation strategy based on RFM scores and market basket analysis results to tailor marketing efforts.

We can utilize data-driven insights to optimize pricing strategies, promotional offers, and product assortment.



OUR IDEA

Our approach **combines MBA and RFM Analysis** to deliver personalized product recommendations and exclusive discounts based on customers' purchase behavior. By identifying the most frequently purchased products for each individual customer, we leverage market basket associations to suggest complementary items. Furthermore, by tailoring discounts based on their purchase frequency, we incentivize customers to continue buying and strengthen their loyalty to our brand. This powerful combination of MBA and RFM Analysis ensures a personalized shopping experience and drives long-term customer satisfaction.

1- Diamond Customers

We used RFM analysis to get high-values information (**Diamond**) customer because the strategy will be reserved to them.
We want to increase the fidelity of our best customers.

2- Product Occurrences

We used the orders table to get the number of occurrences of purchases for each product and customer to get the favorites products for each customer.

3- Consequent Products

We used MBA analysis to find the products that we can recommend to the customer based on his/her favorite products.

4- Personalized Discount

We analyzed the distribution of number of occurrences to define a custom discount that will be applied on the next purchase of the product that we recommend to the user.



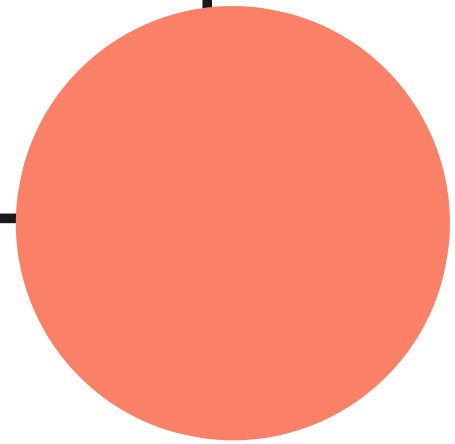
RFM
+
MBA



RESULTS

- **Customized recommendations and discounts**

For each diamond customer, we provide him a **personalized discount on the predicted consequent item based on his favorite product (antecedent)**. Furthermore, the value of the discount could be customized on the occurrences: this means that the more a client purchased a specific product, the more the consequent item recommended will be discounted:

- 20% if the occurrences are above the 0.75 quantile of the occurrences distribution;
 - 5% if the occurrences are between the 0.75 and 0.5 quantile;
 - 1% otherwise.
- 





SENTIMENT ANALYSIS

Sentiment analysis is the process of analyzing digital text to determine if the **emotional tone** of the message is positive, negative, or neutral. Today, we have large volumes of text data like emails, customer support chat transcripts, social media comments, and reviews.

***Extra:** Re-training of a large pre-trained language model (Electra-small)*



SA STEPS 1/2

- **Data**

We need the '*labelled_reviews*' to train the model and the '*customer_reviews*' to make predictions. Both will be cleaned to ensure the best possible condition (garbage in, garbage out principle). Lowercasing, tokenization, stopwords removal, stemming, and undersampling are the main actions performed.

- **Training**

Three models were trained: Multinomial Naive Bayes, Multinomial Logistic Regression, and Gradient Boosting Classifier. Each model underwent a search for the best parameters. **Logistic Regression** achieved the **highest accuracy** level (78.4%), so it was chosen for the marketing strategy.

- **Results**

The prediction resulted in

- 70.1% positive reviews
- 22.2% neutral reviews
- 7.7% negative reviews.

Therefore, the majority of customers provided positive comments.





SA STEPS 2/2

• Electra-small

It is a smaller variant of the Electra model.

Layer:

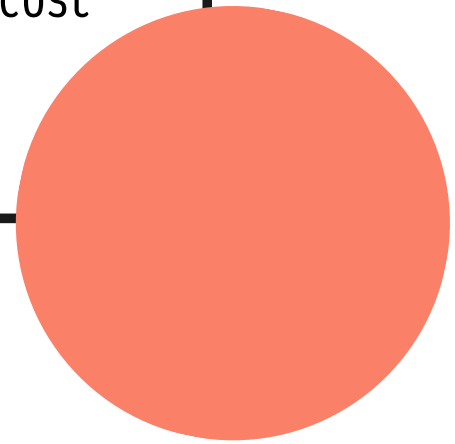
- Input
- Preprocessing
- Encoder (electra small)
- dropout
- dense
- softmax

• Results

The saved model, the one with the lowest loss on the validation set, achieved an accuracy of 80.7% on the training set and 79.5% on the validation set. This model was then used to make predictions on customer reviews.

• Idea

Categorize the main issues identified in the reviews into macro-categories:

- shipping, problems related to the shipping, packaging and boxing
 - food, issues related to the product offered by the provider
 - money, related to the high cost of the product
 - customer_service
- 

SENTIMENT MARKETING STRATEGIES

Food

When customers reach a specific threshold, they become **eligible for a special discount** on a higher quality product. This allows them to enjoy a reduced price while upgrading their purchase to a superior option

Money

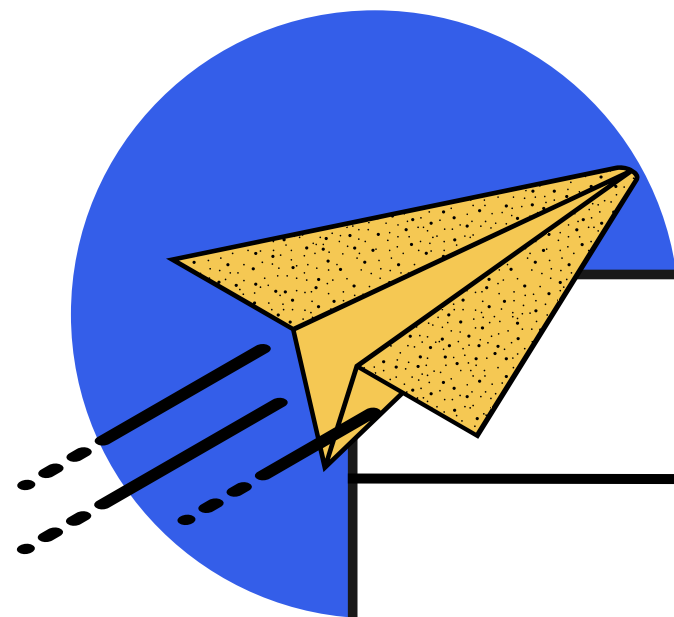
Customers are provided with a valuable discount coupon that they can use to avail a **15% off** their total purchase. This coupon serves as a monetary benefit, allowing them to save a significant amount on their transaction

Shipping

Customers will receive the added convenience of **free shipping** once they surpass a certain **spending threshold**. This perk encourages them to increase their purchase value, knowing that they will not have to pay for the shipping costs, ultimately enhancing their overall shopping experience.

Customer service

Since there are few customers who have complained about the customer service, we could send an email to **gather feedback on the negative aspects** in order to improve them.

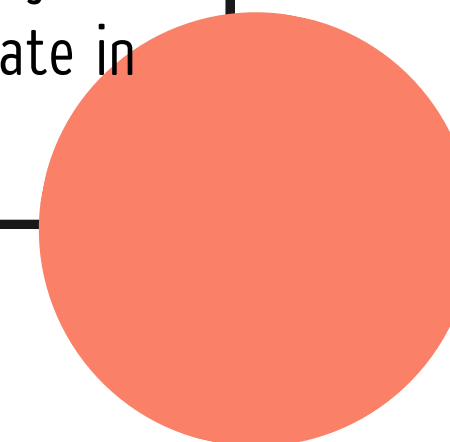


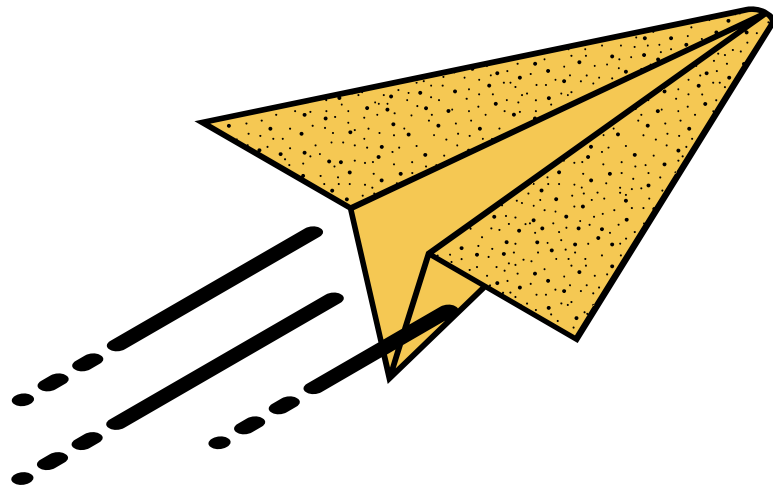
POSITIVE REVIEWS

Loyalty Program: offer customers with positive reviews the opportunity to participate in a loyalty program. So, for each positive review, customers will receive loyalty points that can be redeemed for winning prizes;

Testimonial Campaigns: ask customers with positive reviews to provide more detailed testimonials or reviews that can be used to promote the product or service in marketing campaigns;

Events or Workshops: organize exclusive events or workshops for the customers with the most number of positive reviews. This provides them with the opportunity to firsthand experience new products, receive personalized consultations, or participate in training sessions.





THANK YOU!

Cervini Stella 847060

Montalbano Daniel 897383

Sabino Giuseppe 852287