ifood

Data Business Analyst –Case

Rodrigo Reis   - Notebook click here

Case

The objective of the team is to build a predictive model that will produce the highest profit for the next direct marketing campaign, scheduled for the next month. The new campaign, sixth, aims at selling a new gadget to the Customer Database. To build the model, a pilot campaign involving 2.240 customers was carried out. The customers were selected at random and contacted by phone regarding the acquisition of the gadget. During the following months, customers who bought the offer were properly labeled. The total cost of the sample campaign was 6.720MU and the revenue generated by the customers who accepted the offer was 3.674MU. Globally the campaign had a profit of -3.046MU. The success rate of the campaign was 15%. The objective is of the team is to develop a model that predicts customer behavior and to apply it to the rest of the customer base. Hopefully, the model will allow the company to cherry pick the customers that are most likely to purchase the offer while leaving out the non-respondents, making the next campaign highly profitable. Moreover, other than maximizing the profit of the campaign, the CMO is interested in understanding to study the characteristic features of those customers who are willing to buy the gadget.

Link for file

Objectives

01 **EXPLORE THE DATA**

Exploratory data analysis.

02 **A CUSTOMER SEGMENTATION**

Based on customers behaviors

03 **PREDICTIVE MODEL**

Maximize the profit of the next marketing campaign

04 **CONCLUSIONS**

Concludes from the logistic regression solution.

01

EXPLORE THE DATA

First analyses, made from the dataset.

DATA ANALYSIS OF PROCESSED

Variables

41 different variables

Input

Total of 2240

Null values

24 null values in income

Outliers

there were no outliers

"You can't control what you can't measure"

—TOM DEMARCO

02

A CUSTOMER SEGMENTATION

Based on customers behaviors

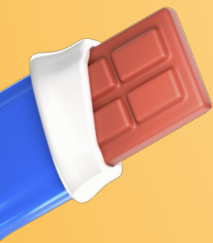# What is the relationship between marital status and spending?

Single

Singles are consuming 13% more meat than married people and today they are the ones who consume the most meat through the App.

Married

Married and the second consumer who consumes the least in the App, if it is above the lonely

Marital Status

Widow

Today it is the one who consumes the most premium products and fruits.

Alone

Alone have a very good income but their consumption is very low compared to other marital states.

Alone

Alone with Master have the best income, if we don't count the Absurd.

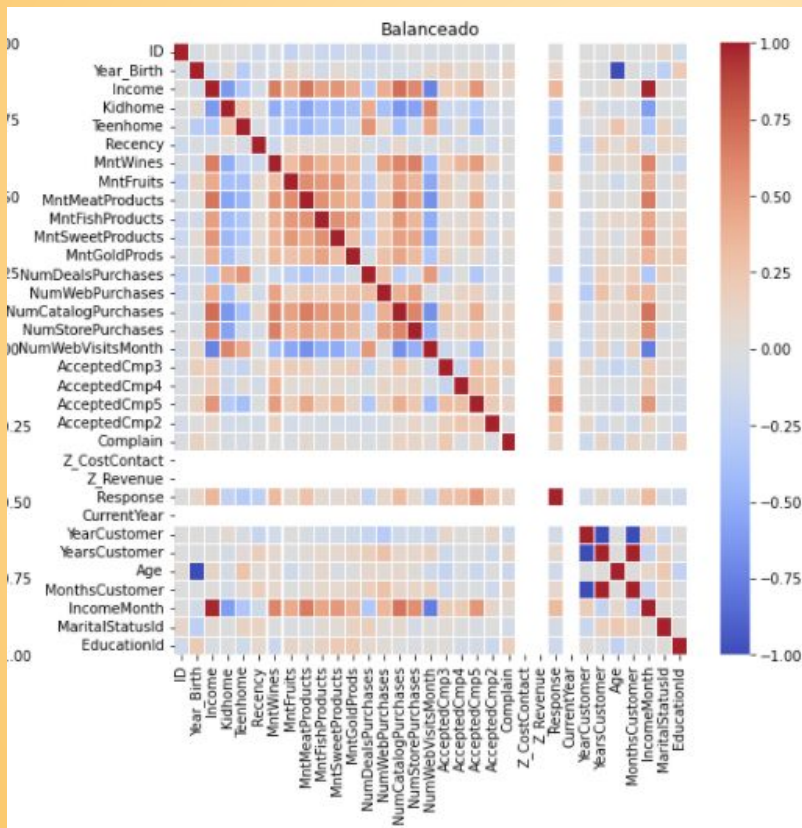# 6.46% Customers

Buy something with the first campaign
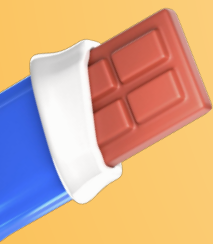
03

PREDICTIVE MODEL

Maximize the profit of the next marketing campaign

Data Balanced Correlation Matrix

Data Balanced Correlation Matrix

With the data balanced, we can look at the correlation matrix and identify which variables are most strongly related to each other.

Classifier using Logistic Regression

With all the data prepared and after a thorough exploratory analysis, I will build a classifier using Logistic Regression.

```
Relatório de Classificação:
              precision    recall  f1-score   support

           0     0.9909    0.7371    0.8454       445
           1     0.1931    0.9032    0.3182        31

    accuracy                         0.7479       476
   macro avg     0.5920    0.8202    0.5818       476
weighted avg     0.9390    0.7479    0.8110       476

Acurácia: 0.7479

AUC: 0.8202
```
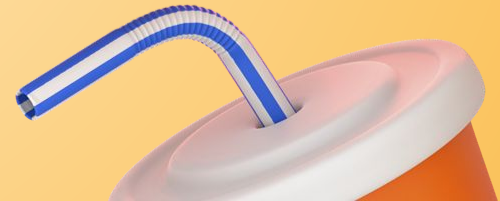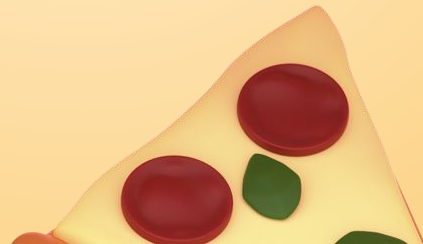

Normalized Confusion Matrix

Data to predict when the customer will buy with the first campaigns.

04
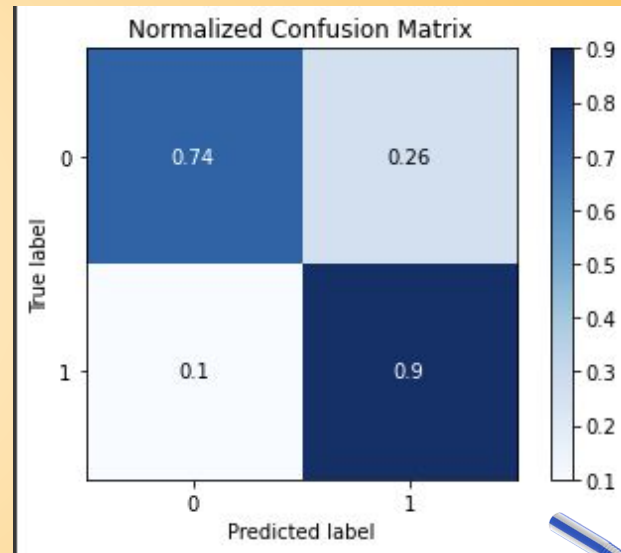
CONCLUSIONS

Concludes from the logistic regression solution.

CONCLUSIONS

As you can see, this is not a trivial problem. Despite the data being very absent and clean, with few values or variables, PCA unbalance and transformation demand conscious work.

A classified space was not identified, a good result, even more with other results besides optimizing its parameters.

THANKS!

Does anyone have any questions?

rodrigo.reis12@gmail.com
+1 604 209 1053 (WhatsApp)
https://github.com/rodrigoreisbr-tech
https://www.linkedin.com/in/rodrigoreisbr/