

**BUSINESS CASES WITH DATA SCIENCE**

**MASTER DEGREE PROGRAM IN DATA SCIENCE AND ADVANCED ANALYTICS – MAJOR IN BUSINESS ANALYTICS**

**Business Case 1 – Wonderful Wines of the World**

Group H

Hiromi Nakashima, m20201025

Manuel A. F. Carreiras, m20200500

Luis F. R. Agottani, m202000621

<Fourth student name>, number: <student number>

February, 2021

THIS IS JUST AN EXAMPLE OF THE INDEX. FEEL FREE TO CHANGE THE STRUCTURE

Table of Contents

[1. INTRODUCTION 1](#_Toc65267510)

[2. BUSINESS UNDERSTANDING 2](#_Toc65267511)

[2.1. Business Objectives 2](#_Toc65267512)

[2.2. Business Success criteria 2](#_Toc65267513)

[2.3. Determine Data Mining goals 2](#_Toc65267514)

[3. PREDICTIVE ANALYTICS PROCESS 3](#_Toc65267515)

[3.1. Data understanding 3](#_Toc65267516)

[3.2. Data preparation 3](#_Toc65267517)

[3.3. Modeling 4](#_Toc65267518)

[3.4. Evaluation 4](#_Toc65267519)

[4. RESULTS EVALUATION 6](#_Toc65267520)

[5. CONCLUSIONS 7](#_Toc65267521)

[5.1. Considerations for model improvement 7](#_Toc65267522)

[6. REFERENCES 8](#_Toc65267523)

[7. APPENDIX (OPTIONAL) 10](#_Toc65267524)

[Figure 1 - Variable Information 3](#_Toc65271688)

[Figure 2 - Data preparation process 4](#_Toc65271689)

[Figure 3 - Clusters Distribution over RFM classes 4](#_Toc65271690)

[Table 1 - Data Mining Goals 2](#_Toc65272043)

# 

# INTRODUCTION

This project we were designated to analyze a wine company dataset named Wonderful Wines of the World (WWW). WWW is a 7-year-old enterprise, which sells wine through three channels: catalogs, web site and physical stores (10 branches). The purchase can be done in the physical stores, telephone or online.

Into this moment, WWW keep your clients engaged by sending them a newsletter with the updates of wine world. Even that the database of WWW has only 4-year-old, the company recently organized a marketing activity which aggressively increased the database. One of the currently pain points is a lack of cross selling strategies which support the recurrent profit.

This project was developed with a 10.000 sample of the currently WWW’s database of customers that purchased in the last 18 months. And the report was adjusted in four main parts based in CRISP-DM methodology [CRISP-DM ARTICLE].

# BUSINESS UNDERSTANDING

At this step were defined the essential business guide lines to garante a good result of the project. In order to develop the best solution to WWW the business understanding were based in the currently reality of the company presented on the introduction.

## Business Objectives

The goals of WWW are:

* Improve the familiarity of the database by create a classification for each client to develop marketing strategies by profile;
* Be able to classify the new customers;
* Improve the Return on Investment by understanding the client value (ROI).

## Business Success criteria

Based on the business objectives describe, were defined two main results to guarantee the success of this project: identify the profile of the new customers since the first purchase, develop marketing strategies to reach all classification profiles and recurrent profit to the company.

## Determine Data Mining goals

Based on the business goals we translated to Data Mining language as shown in the table below (Table X).

Table - Data Mining Goals

|  |  |
| --- | --- |
| Business Goal | Data Mining Goal |
| Classify the currently clients by profile | Clustering the clients |
| Ranking the clients to understand the ROI | Apply the recency, frequency and monetary value (RFM) |
| Identify the new customer profile | Apply a predictive model |

# PREDICTIVE ANALYTICS PROCESS

## Data understanding

The initial process of understanding the problem were explained in the introduction and the data structure (Figure X/METADATA PRINT). Were identified 30 columns and 10001 entries, where all of them presented to be numeric features. Although after a better understand of the metadata and the features, we recognized that 10 features were binary then we converted them to Boolean type. Also, on this step some columns were drop due to the insignificance for the project (Data Exploration, Notebook Reference)

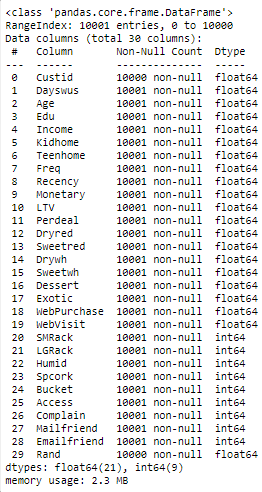


Figure – Variable Informatio

## Data preparation

The diagram below presents the data preparation steps followed to reach the final model in this project (Figure X).

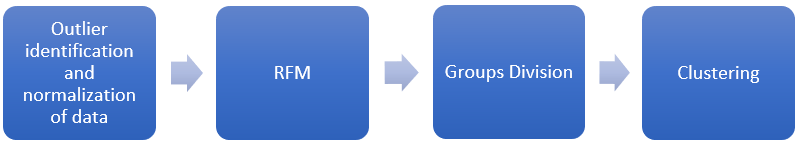


Figure - Data preparation process

Firstly, we identified the outliers and normalized the data. Subsequently were develop a RFM analysis to classify the value of each customer, by this classification we reached to 5 clusters which the client quality. After we clustered the clients by groups to reach a better profile clustering of the client. The result is presented below (Figure X).

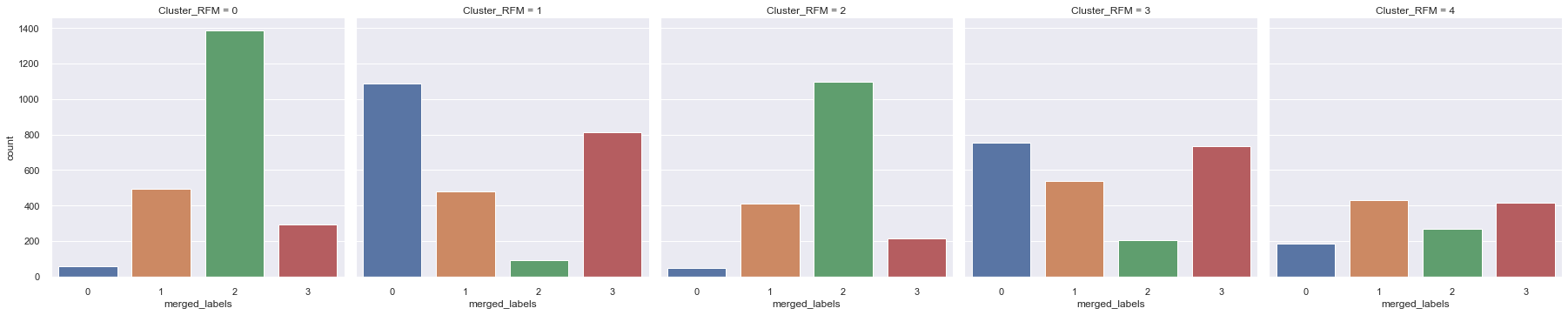


Figure - Clusters Distribution over RFM classes

## Modeling

Bla, bla.

## Evaluation

Results described in technical terms (e.g., reached an Accuracy of 95%).

# RESULTS EVALUATION

Describe the degree to which the model meets the business objectives. If that cannot be done without the application of the model in a real environment, describe how could that be done.

Assess the data mining results in respect to the business success criteria.

# CONCLUSIONS

Final remarks on the project.

## Considerations for model improvement

Bla, Bla

# REFERENCES

Author, A. A., Author, B. B., & Author, C. C. (Year). Title of article. *Title of Periodical, volume number* (issue number), pages.

# APPENDIX (OPTIONAL)