Marketing econometrics

TSE – M2 SE and D3S – 2021-2022

By Joanna MORAIS

Project n°1:
Behavior analysis & operating levers using a RFM segmentation

Context:

The brand "Feminiz" is subject to a strongest and strongest competition on the lingerie market. One of the first action was to develop a wider range of complementary products with bathing suits and homewear.
In parallel, a loyalty program was developed to connect with clients and reward them.
To go further in the perspective of differentiation from its competitor, "Feminiz" wishes to establish a RFM segmentation to communicate differently with its clients and to allocate optimally its marketing investments.

The lingerie brand "Feminiz" is thus commissioning you to make **a RFM analysis of its customers on the last year.**

Data:

The data available are receipts, the customers characteristics as well as reference tables on stores and products.

Expected work:

You should prepare a presentation of 15 minutes maximum (PowerPoint shape, in ppt of pdf format) for the steering committee of the marketing department on the 15/11/2021.

The outline of the presentation is imposed as follows:

1- Context & objectives of the meeting
2- Presentation of the RFM analysis results
3- Analysis of a particular segment of customers
4- Conclusion and your recommendations

You will be graded on content and form. Be professional!

Practical work guidelines

Organization:

During the 2 practical sessions (and then at home), you will work in groups of 4-5 students on Project n°1 which you will present at the oral on the 15/11/2021.

Each group organizes itself as it sees fit for the project and can parallelize the different components of the project (python code, data analysis, PPT presentation, overall organization) but each member must be comfortable with all the steps.

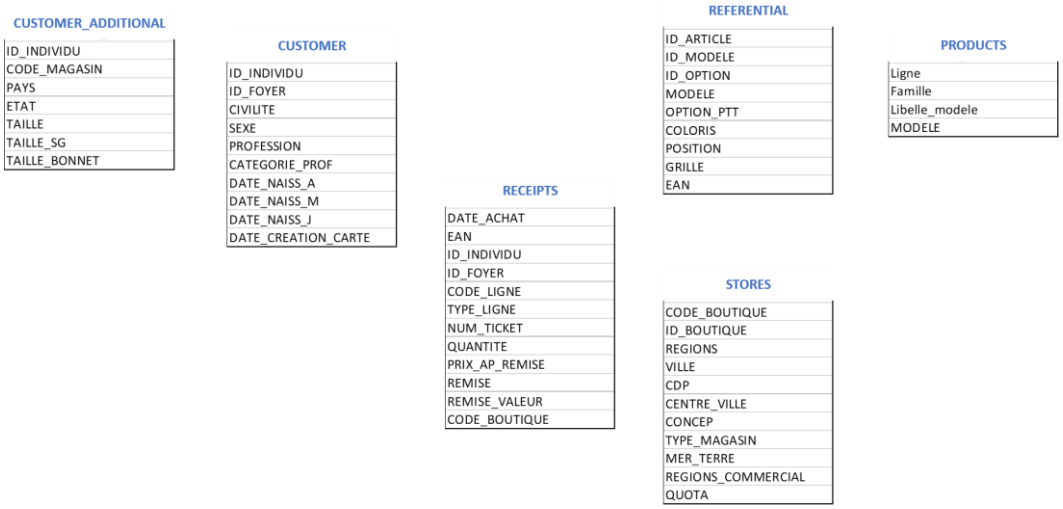
The overall objective of this project is to build an RFM segmentation for a lingerie brand (all groups have to build the RFM segmentation according to the following guidelines and according to what we will discuss during the practical sessions). Then each group will choose a particular segment, analyze it using the statistical tools of their choice and give marketing recommendations to communicate effectively with them.

Part 1: Data understanding

1. Load the data

- Import the 6 csv tables in a Jupyter notebook with Python 3 and look at the first rows of each table.
  - CUSTOMER
  - CUSTOMER_ADDITIONAL
  - STORE
  - REFERENTIAL
  - RECEIPTS
  - PRODUCTS

- Complete the following data relational schema with the link between tables.



2. Selection and merging of the data

Build the CUSTOMERS_INFO and the RECEIPTS_INFO tables containing all the information as follows:

- Join the *CUSTOMER's* dimensions (CUSTOMER and CUSTOMER_ADDITIONAL)
    - Drop ID_FOYER
    - Keep Code_magasin from CUSTOMER_ADDITIONAL
    - Code_magasin will be rename MANAGING_STORE not to confuse it with Code_magasin from RECEIPTS table

- Join the *RECEIPTS's* dimensions (STORE, REFERENTIAL, PRODUCTS and RECEIPTS)
    - Keep only Ligne and Famille from PRODUCTS
    - Keep only REGIONS, CENTRE_VILLE, TYPE_MAGASIN and REGIONS_COMMERCIAL from STORE

- What is the type of merging to use? What are the tables of reference?
- What are the merging keys of each table?

3. Understanding the data

Before starting the audit of a database, it is important to understand how the different types of data are linked.

- To help you, you can study the data of a particular customer: deduce from the analysis of the customer 174591 a rule for identifying a purchase receipt.
- From the 2 databases previously built, study the consistency of the following variables with relevant univariate distribution and cross repartitions.

    - Features to analyze in RECEIPTS_INFO:
        - REGIONS
        - CENTRE_VILLE
        - TYPE_MAGASIN
        - REGIONS_COMMERCIAL
        - MODELE
        - LIGNE
        - FAMILLE
        - EAN
        - PRIX_AP_REMISE
        - QUANTITE
        - REMISE_VALEUR
    - Features to analyze in CUSTOMERS_INFO:
        - Age
        - Seniority

- Analyze a particular customer to understand missing data
    - Focus on the *MODELE* variable and especially on the modalities ACCES, DIVE, DIVERS, FAVO, FAVORI, PACK, PLV
    - Identify potential inconsistencies and submit adequate management rules

Part 2: Data preparation

4. Data preparation: cleaning and creation of indicators

Scope: on the last 12 months available

- Define and apply the rules needed to clean the data (missing values, outliers…)

RECEIPTS_INFO table:

- Calculate per visit:
  - Turnover (VISIT_VALUE)
  - Number of products (NB_PRODUCTS)
  - Average amount per product (AVG_PRICE)

- Calculate per individual (from the precedent table per visit):
  - Number of visits (NB_VISITS)
  - Average turnover (AVG_VISIT_VALUE)
  - Cumulated turnover (CUMUL_VALUE)
  - Average number of products per basket (AVG_NB_PRODUCTS_VISIT)

- Calculate per individual:
  - Recency of the last purchase (RECENCY)
  - Number of distinct stores (NB_STORES)
  - Number of different lines (NB_LINES)
  - Number of different families (NB_FAMILIES)
  - Number of gifts (NB_GIFTS)
  - Share of visits in the managing store (= number of visits in the MANAGING_STORE / total number of visits)

CUSTOMERS_INFO table:

- Calculate per individual:
  - Age (AGE)
  - Seniority (SENIORITY)

- Gather all features in a MASTER table (at the individual level)

Part 3: Build the RFM segmentation

- ☐ Analyze the distribution of the following variables:
  - Amount,
  - Frequency
  - Recency

- ☐ Determine thresholds in order to split them in low/medium/high groups.
- ☐ Create the final RFM segments

Part 4: Analysis of segments

Analyze dimensions: Eclecticism / Reactivity / Diversification / Attachment

CATEGORY	INDICATOR
CLIENTS	AGE
CLIENTS	GENDER
CLIENTS	SENIORITY
PURCHASE	CUMULATED PURCHASE AMOUNT
PURCHASE	AVERAGE BASKET
PURCHASE	NUMBER OF VISITS
BEHAVIOR	RECENCY
BEHAVIOR	LOCATION OF THE MANAGING STORE
BEHAVIOR	PURCHASE PART OF THE MANAGING STORE
PRODUCTS TYPE	NB STORES
PRODUCTS TYPE	NB DIFFERENT FAMILIES
PRODUCTS TYPE	NB DIFFERENT LINES
PROGRAM	NB GIFTS

Bonus :

5. Migration

Apply the same treatment on the previous period

- ☐ Recalculate the segments dropping the last 6 months of data.
- ☐ Cross the two segmentations to quantify migrations

Part 5: Marketing recommendations

Focus on one segment and give marketing recommendations to the marketing direction using on statistics tools.