CURSO 19-20

Data Analytics with R MIBA, 2019

Boosting sales and achieving operational excellence at Vendex

Gonzalo de la Torre
Rubén Coca

CONTEXT

Do you feel thirsty or hungry, craving for your favourite chocolate bar or soft drink and you are on the go? Want a healthier option instead? At Vendex we know your day is frenetic and there aren't many opportunities to sit back and enjoy a break with a delicious snack or drink. We want to be there right when you most need it.

Vendex is a leading vending and coffee services company in Europe, with a turnover of 984 million euros on 2017 and employing approximately 5,500 people. Founded in 1967 and headquartered in London, Vendex. has grown its geographic market presence to 15 countries across Europe. Vendex serves more than 8 million consumers every day at its 165,000 point of sales addressing the growing need for out of home food and beverage services at the workplace and on the go.

Vendex roots stand for service excellence, a high-ˇquality product offering and innovative concepts for out of home food and beverage services.

Currently, Vendex is putting a strong focus on boosting sales and profitability by improving pricing and assortment strategy and achieving operational excellence through Big Data and Advanced Analytics.

DESCRIPTION OF THE DATA

Vendex has provided data from ~2500 machines including ~3 months of transactional data starting at 2017-01-02 consisting of records of every item being sold. Additionally, it has provided information about each product as well as the location characteristics of the machine.

There are three main sources of data:

1. **Transactional_data.csv** contains information of each transaction i.e. what product was sold, timestamp of the sale, etc.
2. **Product_data.csv** contains data regarding each product i.e. what category the product belongs to, what's the price, cost and tax associated etc.
3. **Machine_data.csv** contains information about the machine i.e. whether is a big or small model, and its location characteristics such as number of hotels, income of the area, number of passengers if train/metro station, etc.

Data fields

Transactional_data		
Column Name	**Data type**	**Description**
machine_id	character	Machine ID
timestamp	Date and time	Timestamp of the transaction including date and time
date	Date	Date of the transaction (included also in the column date)
column	Integer	Slot ID
product_name	character	Name of the product (disguished)

Machine_data		
Column Name	**Data type**	**Description**
machine_id	character	Machine ID
location_type	character	Type of location: transport such as an station, petrol station or others
num_vendex_nearby_300	character	Number of nearby machines in a 300m radius
train_AvgDailyPassengers	numeric	Daily number of passengers in the train station
train_AvgWorkingDayPassengers	numeric	Daily number of passengers in the train station in a working day
n_density_5km	numeric	Density within a radius of 5km
income_average	numeric	Average income of the area
total_number_of_routes_600	integer	Total number of routes (bus, metro, tram, etc.) in a radius of 600m
num_hotels	integer	Number of hotels in the area
num_hotels_45	integer	Number of hotels of 4 and 5 stars in the area
small_machine	Boolean	Indicator of small machine (1 – Small machine, 0 – Big machine)

Product_data		
Column Name	**Data type**	**Description**
product_name	character	Name of the product (disguished)
category	character	Category of the product
type_drink_snack	character	High level category of the product: snack or drink
price	numeric	Retail price of the product
cost	numeric	Cost for Vendex of the product
tax_rate	numeric	Tax rate applied on the product

OBJECTIVES OF THE CASE

The aim of this case is to **help Vendex grow the business** boosting sales and cutting costs by achieving operational excellence, leveraging Big Data and Advanced Analytics.

For that, Business development team has identified **5000 new potential locations** and require **where to place 1000 new machines** (700 of big size and 300 of small size) and **what assortment to place in them**. **Rationale for your decisions as well as an impact estimation will be expected**. Additionally, you will be asked to **identify other opportunities for Vendex** in all areas, keeping them mostly at a conceptual level with ballpark level impact estimation.

To help you on this goal, you will receive two challenges, containing some tasks to do with the data that will help you understand better the industry as well as build the final models.

- Challenge 1: General Overview and Data Treatment (**due 6th of March 2020**)
- Challenge 2: Introduction to Probability Models (**due 29th of March 2020**)

Finally, you will have to make a presentation (maximum 15 slides) on April 3rd, summarizing the key insights you will get from the challenges, the results of your model and some recommendations to improve Vendex sales with an impact estimation.

EVALUATION

Your final grade on the final business case presentation will be based on the following criteria:

- 20% - Presentation and communication skills
- 60% - Quality of identified opportunities (quality of the models developed and estimated business impact)
- 20% - Document presented (including strategies and implementation plan)