ANALYSIS OF VENDING MACHINE SALES

Universidad: Universidad Francisco de Vitoria

Facultad: Derecho, Empresa, y Gobierno

Carrera: Business Analytics

Clase: 4ºA

Asignatura: Inteligencia Empresarial I

Nombres y apellidos: Manuel Rafael León Ruiz y Juan Fernández Cerezo

Fecha: 04/12/2023

Index:

Juan Fernández y Manuel León

Introduction

We have been hired by a **vending machine company**, in New Jersey, to carry out an estimation of product replenishment, in order to improve in all the value chain of the company. We have been in contact throughout the process with one of the managers, with knowledge in most of the fields of the company; technology, business, marketing,.. and he has asked us several questions, such as, How much should I buy from a product? What is the best selling product? Which is the most profitable machine?... All those questions are key points in the company, which will increase their profit margin and make them more efficient in the whole process of the activity. It will help them reduce costs in personnel and food, implement a better inventory management system for not losing possible clients, or know different kpis for knowing what to invest.

The successful implementation of these strategies can significantly enhance the efficiency of the vending machine company's operations. By determining the optimal quantity of each product to purchase, the company can strike a balance between meeting customer demand and avoiding unnecessary overstock, thereby minimizing storage costs and the risk of product expiration. This approach also ensures that popular products are consistently available, satisfying customer preferences and increasing overall sales.
Identifying the best-selling products is crucial for adapting the product mix to customer preferences. Through the analysis of sales data and market trends, the company can not only boost its current offerings but also stay ahead of emerging consumer preferences. Regular market research will provide valuable insights into changing customer behaviors and enable the introduction of new, innovative products that resonate with the target audience.
Profitability analysis for each vending machine allows the company to allocate resources strategically. By understanding the financial performance of individual machines, the company can make informed decisions about relocating underperforming machines to high-traffic areas, thereby maximizing revenue. Pricing strategies can be adjusted based on this analysis, ensuring that each machine strikes the right balance between profitability and affordability for the consumer.
Cost reduction strategies play an important role in improving overall profitability. Streamlining personnel costs through automation and optimizing replenishment routes can lead to substantial savings. Negotiating favorable terms with suppliers and adopting energy-efficient technologies further contribute to cost reduction, strengthening the company's financial position.
The implementation of a robust inventory management system is fundamental to preventing stockouts and minimizing the risk of losing potential clients.

User Persona:

<table>
<tr><th colspan="3">USER PERSONA</th></tr>
<tr><td rowspan="2">  </td><td>Personality</td><td>Hobbies</td></tr>
<tr><td>Organized and efficient

Ambitious

Innovative

Leadership</td><td>Sports: Tennis, Cycling

Read

Travel</td></tr>
<tr><td rowspan="2">Name: Joseph

Age: 56

Role: Operations Manager

Company: Vending_Machines

Ubication: New Jersey</td><td>Goals/Needs</td><td>Roles in the company</td></tr>
<tr><td>Increase vending machine sales

Improve customer satisfaction

Reduce operating costs

Accurate and up-to-date information on vending machine performance</td><td>Attend vending machine industry events

Resolve customer complaints

Optimize machine placement

See where to invest</td></tr>
</table>

Database:

Before getting into the process of the model, I will make a description of the database with each of its variables and the data engineering process we have carried out.

The initial database has 18 fields and 9617 records. The variables it has are as follows:

- Status: Represents whether the machine data is successfully processed.
- Device ID: Identification code of the machine.
- Location: Location of the machine. Indicates location of the vending machine
- Machine: Name of the machine. User-friendly machine name
- Product: Name of the product. Product vended from the machine.
- Category: Carbonated / Food / Non-carbonated / Water

- Transaction: Unique identifier for every transaction
- TransDate: Date of sale
- Type: Type of transaction (Cash / Credit)
- RCoil: Number of the machine coil
- RPrice: Price of the product
- RQty: Quantity of products sold in the transaction.
- MCoil: Spiral referenced to a product.
- MPrice: Minimum price at which a product can be sold.
- MQty: Minimum number of products that leave the machine per transaction (configured by the machine).
- LineTotal: Amount of money paid per transaction. LineTotal = MPrice*MQty
- TransTotal: Total amount paid by the user, adding all transactions.
- Prcd Date: Date where the payment is processed.

After observing the database and understanding it, we noticed that there were errors in the format of the variables, they were all set as alphanumeric so we changed the types; TransDate and Prcd Date to date type, MPrice, RPrice, TransTotal and LineTotal to monetary value in dollars, and Device ID, Transaction, RCoil, RQty, MCoil, MQty to numeric value. With the monetary values, we had to face another drawback, due to the fact that the numbers were all text and the commas disappeared, so we had the prices like $15 or 350$ when the correct price was $1.5 and $3.50 for this case, so we had to create another column in order to fixing it. Moreover we noticed some missing values that we rather input a value depending on similarities with other values or delete them when we had no clue about them. The other variables have been left as they were.

Model used:

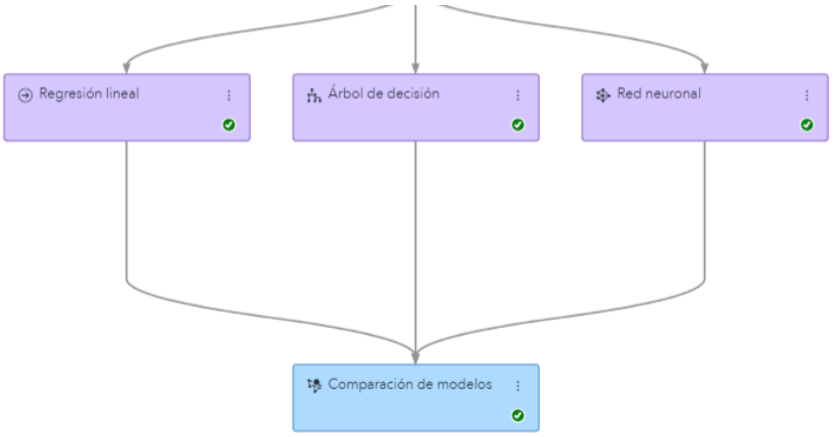
To do all what the manager demanded, we have studied different models according to the most important variables available. Subsequently, after obtaining our best model, we have carried out a forecast with this model and a study of how it can improve the whole activity and the relationship between different variables.

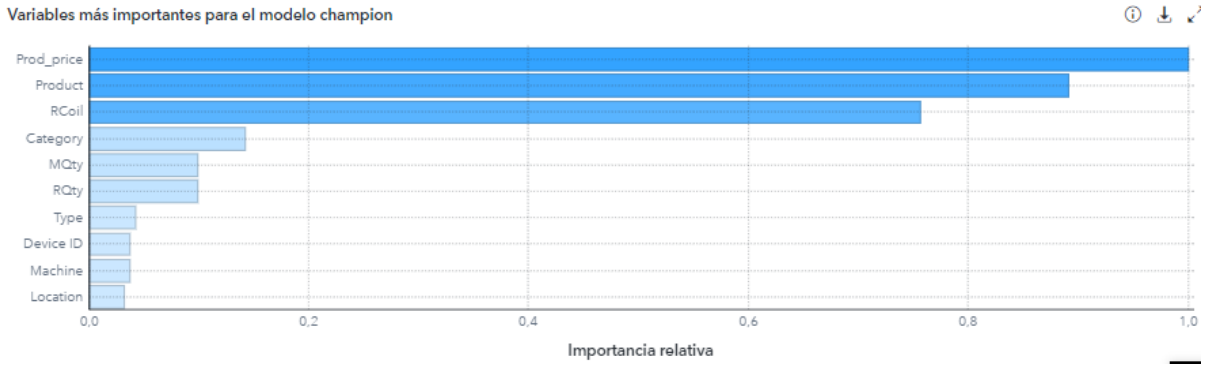
To determine the best model, we have carried out a pipeline in SAS with three models (Linear regression, Classification tree and Neural networks), establishing LineTotal as the objective variable. SAS obtained that the best model according to our variables is the linear regression model. In addition, we can see in the insights which are the most important variables in our model.

Resumen del proyecto

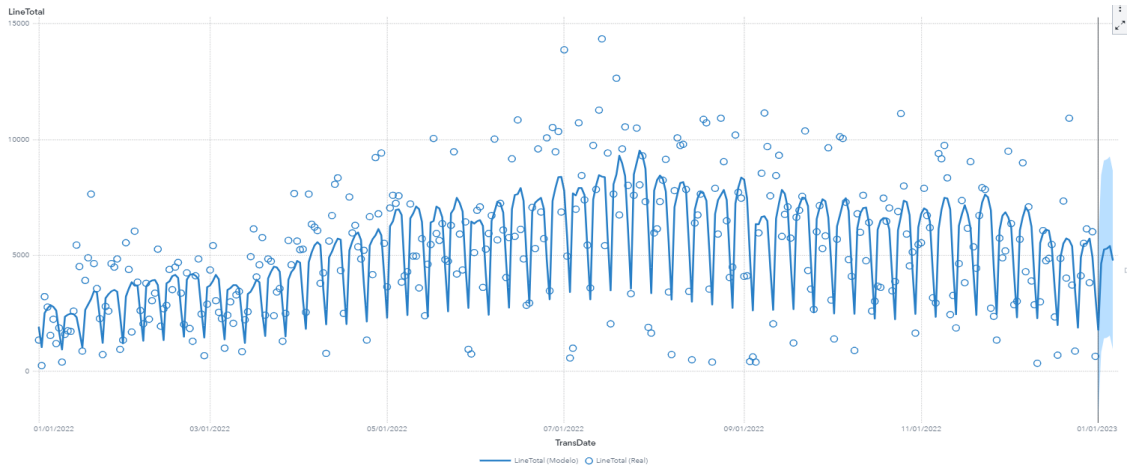
The champion model for this project is Regresión lineal. The model was chosen based on the Error cuadrático medio for the Entrenamiento partition (56.72). The five most important factors are Prod_price, Product, RCoil, Category, and MQty.

Objetivo del proyecto:	LineTotal	**Champion del proyecto:**	Regresión lineal
Promedio global:	198,5667	**Creado por:**	9004075@alumnos.ufv.es
Pipelines:	1	**Modificado:**	28-nov-2023 17:28:28





After doing the pipeline, we forecasted the evolution of the total revenue (LineTotal=MQty*MPrice) for the following days:



Juan Fernández y Manuel León

Visual Solution:

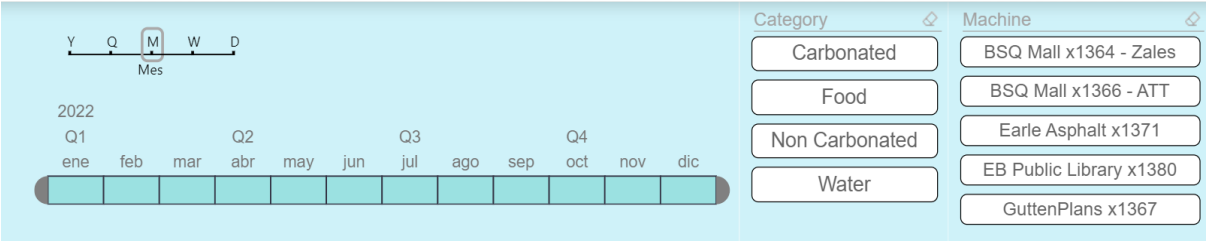
Additionally, we will also provide, in a more visual manner, how the company can track the sales of its vending machines and thus anticipate the supply needs, whether daily, monthly, or annually. This can serve as a solution to optimize inventory or purchasing. We aim to integrate this through a fully interactive Dashboard that we have created in Power BI, which would form the basis of our solutions to the possible questions or challenges that arise in this industry.

To achieve this, we started by loading all the data from the, previously mentioned, database into Power BI. Despite having done the necessary cleaning, as mentioned before, we also had to make some format changes so that Power BI could understand the data properly for the later Dashboard creation. Additionally, in order to visualize some KPIs, we had to create new measures such as "Average price per product" or "Distinct products." We also created a new column to display only the days of the week in text format to facilitate visualization.

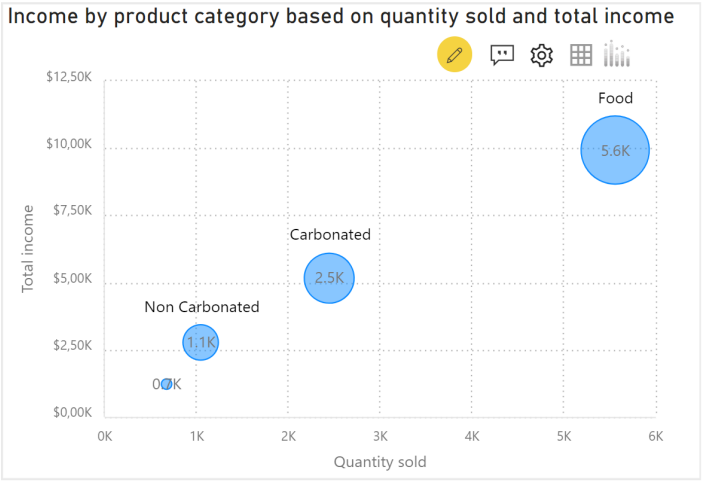
As mentioned, the Dashboard is the pillar of our work, and, apart from showing a snapshot of it in this document, we recommend viewing the file separately to understand the solutions provided. Nonetheless, we will explain its content:

The dashboard displays the sales of five vending machines located in New Jersey, situated in different locations such as a shopping mall, a library, or a gas station.

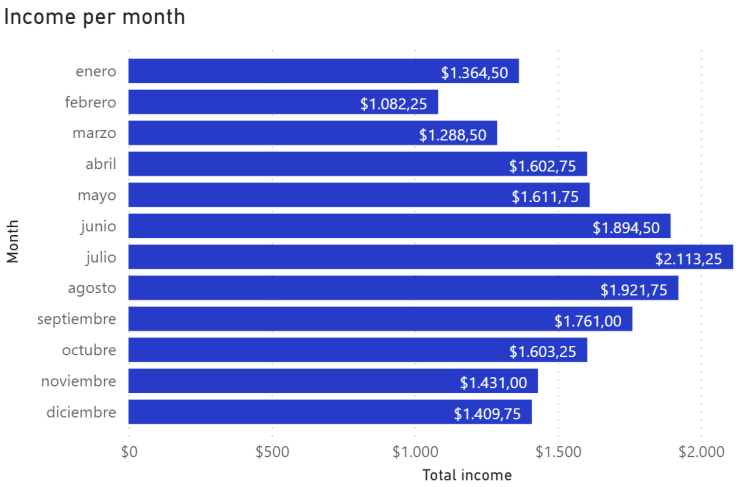
At the top, there are three filters, allowing the manager to filter by **temporality**, choosing quarters, months, weeks, or specific days (slicer); **type or types of product** (as two can be chosen simultaneously) (chiclets), differentiating between food, carbonated beverages, non-carbonated beverages, or water; and the specific **machine** or machines (chiclets). This provides a more specific view in all fields vital for decision-making for both machines and products over time:



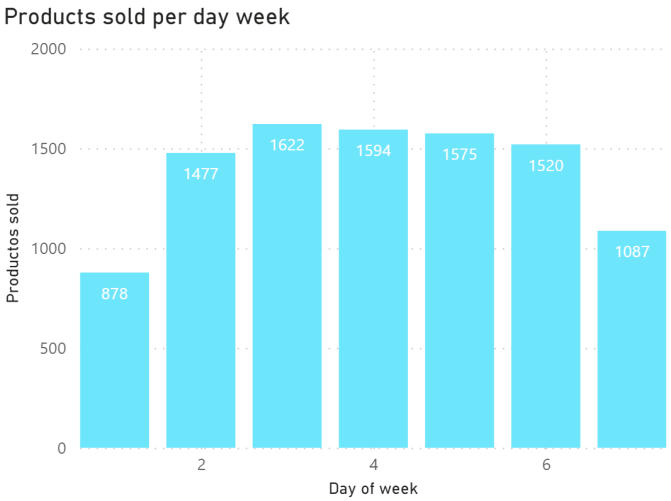
Further down, the first chart is a **scatter plot** that serves to observe the relation between the quantity sold of a type of product and the total revenue for that type and comparing it to the other types. This helps identify which product type can provide the highest margin of income per sale:



The next chart is a **horizontal bar chart** with the months of the year and their total revenues. Summer months usually generate the most revenue, indicating where more resources need to be invested:



The last chart in this row is a **bar chart** using the previously created column for the day of the week, showing which day of the week products are sold the most. Typically, Tuesday to Saturday are the days with the highest sales, although this data differs when filtering the two machines located in shopping malls:



In the second row, there are three different KPIs. The first is the **average price** of all products sold in the machines. The second KPI refers to the total number of **distinct products** among all machines, although it can be filtered for each one. And the last KPI is the sum of the **total products sold**. Undoubtedly, these three measures will assist in decision-making:

Average price per product	Disinct products	Total products sold
$1.99	184	9753

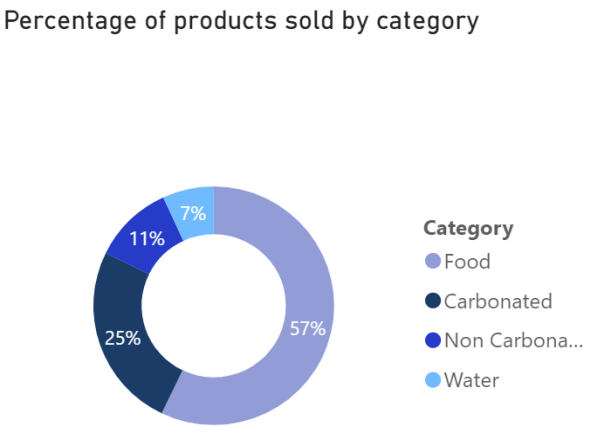
In the last row, there are four charts. On the ends, there are two **tables**. The table on the left shows the sum of sales for all items sold in different vending machines, serving as a guide for estimating the supply of products needed for a quarter, a month, etc. On the other end, there is a table showing the total revenue for each machine, providing an estimate of where to concentrate efforts:

Product	Suma de RQty
Zone Perfect - Chocolate Chip Cookie Dough	4
Zevia Zero Calorie Energy - Mango Ginger	6
Zevia Organic Tea - Peach	4
Wonderful Pistachios - Variety	196
Vitamin Water - XXX Acai BB Pomegranate	17
Vitamin Water - Squeezed Lemonade	25
Vitamin Water - Ice Cool Blue Lavender	11
V8 Vegetable Juice - Original	28
TruBar - Daydreaming about Donuts	7
TruBar - Cookie Dough	7
Teddy Grahams - Go Paks	21
Takis FUGEO mini	83
Takis - Hot Chilli Pepper & Lime	97
Sunkist Soda - Orange	290
SunChips Multigrain - Salsa	205
Total	**9753**

Suma de LineTotal	Machine
$7.001,25	GuttenPlans x1367
$6.827,75	EB Public Library x1380
$1.827,25	Earle Asphalt x1371
$1.670,75	BSQ Mall x1366 - ATT
$1.757,25	BSQ Mall x1364 - Zales
$19.084,25	

Juan Fernández y Manuel León

The next chart is an intuitive and visual **circular chart** showing the percentage of sales by product category among total sales:



Lastly, there is a **HeatMap** where we can observe the top 12 best-selling products. By hovering over a product, we can see its exact sales quantity. This chart will undoubtedly help us understand which products are crucial and which require more restocking:

Conclusion:

Having explained the Dashboard, we realize its importance in helping the company's manager make decisions regarding advertising or optimizing inventory, provisioning, and distribution management, thereby reducing costs. It provides a comprehensive and specific, when needed, view of the business.

Juan Fernández y Manuel León

So, overall, we have learnt a lot throughout the process of doing this assignment, from the transformation of the data extracted, or the understanding of it, to the final visualization, using useful tools such as Excel, SAS Viya, or Power BI. So we believe this work can easily be implemented for these kinds of businesses and would be so helpful at the time of taking decisions. We hope you liked it.

Having explained both the model and the visualization solutions, it becomes evident that its indispensability is key for the decision-making, particularly for advertising strategy formulation and the optimization of inventory, provisioning, and distribution management. This, in turn, causes a consequential reduction in operational expenditures. The Dashboard, with its capacity to give a comprehensive and, when needed, granular view of the business landscape, turns out to be an invaluable tool in the strategic decision-making process of the company's leadership.

Throughout the iterative processes for the execution of this assignment, from the meticulous transformation of extracted data to its comprehension and ultimate visualization, the role played by tools such as Excel, SAS Viya, and Power BI is clearly shown. So, taking this into consideration we are convinced that the integration of this framework could cause benefits for enterprises for this type of businesses, particularly for decision-making.

Overall, the knowledge acquired through this assignment goes beyond just manipulating data; it represents a shift in how businesses strategically positioned themselves. The practicality of the gained knowledge, as we commented, lays the groundwork for making well-informed decisions. So, we believe that you'll find the value in this work aligned with your evaluation.