BC Jr Business Analyst

Health QC

a journey towards being the #1 grocery convenience store

by Alejandro Domínguez Arquero

Outline

01 Objective

02 Methodology

03 KPIs

04 Recommendations

05 Assumptions & Next Steps

1. Objective

Conversion rate

Average order size

Product margin

2. Methodology

01

EDA - Explanatory Data Analysis

Using **Python**, null values detection was done, as well as data type identification. Also, different columns were created: revenue, product margin, time of the day, etc. In summary, data transformation to make it workable.

02

KPIs determination

Once data was clean, **KPIs** were identified: average order size, average product margin, among others. See point 3 in this report.

03

Tables relationship establishment

Using **SQL**, relationship between the two tables was established in order to link them through a query.

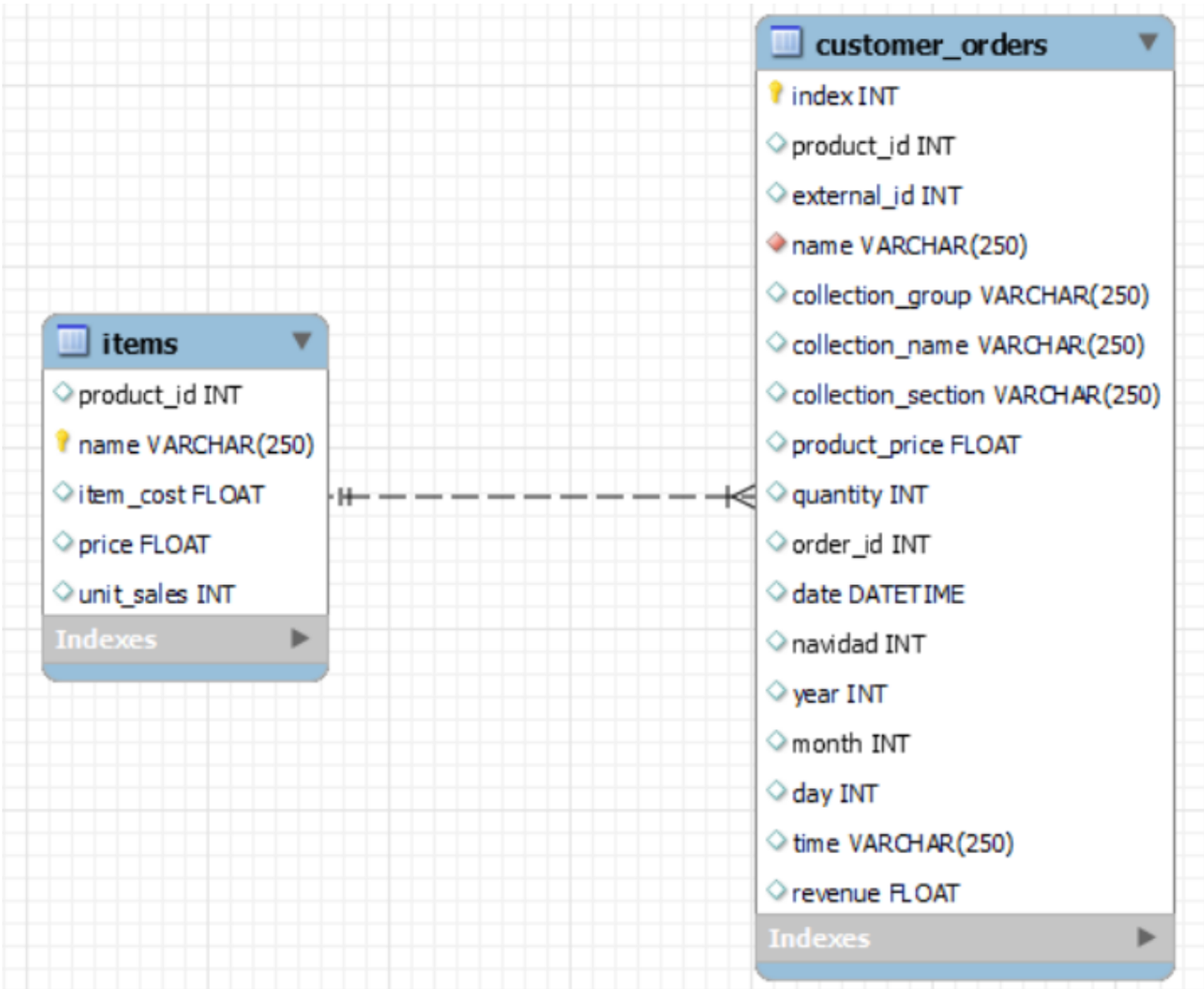
04

Insights identification

Several calculations using **Python** and **Excel** were done in order to identify different insights to address the objectives.

2. Methodology

pandas matplotlib



3. KPIs

Total margin
49K

Total revenue
234K

Total Margin %
21 %

Number of orders
19K

Total quantities
120K

References
837

3. KPIs

Av. order size
12,32€

It **varies** depending on the **time of the day**: slows down during the night: **11,63€**, and reactivates during the afternoon: **13,27€**.
Also, on average, orders with drinks category have the highest av.order size, followed by the ones with frozen products.

Again, this metric is **higher** in the **daytime** than during the **night** (7,2 units versus 5,8 units).
Also, a fact is being repeated: drinks group present the highest average quantities per order, being the **beer cans** at the top.

Av. q/order
6,29u

Av. categories /order
1,89 cat

Each order, on average does **not even reach 2 or more collection groups**.
This means, people tend to buy, on average, several quantities of products on the same category, and **sometimes adding other categories**.

This is the metric that does not significantly varies, no matter the time of the day.
However, when we look at category level, we see huge differences, being **cleaning category** the clear **winner**, followed by **hygiene and beauty** (each represent 4% of total revenue). **Maybe is this the precedent of the Pharma& Beauty (Retail) vertical?**

Av. margin/ order
2,59€

4. Recommendations

01

Product portfolio reorganization:

The current portfolio consists of **847** different products, which are categorized in one of the 12 collection groups. **Drinks** category brings the **54% of total revenue** and is present in **34%** of the **orders**. However, from the total products, there is just a **17% of drink products**. If market is demanding drinks, an **increase in supply** should be implemented.
Also, **breakfast** and **general food** collection groups (second and third revenue stream) bring **higher gross margin %** (25% and 24% respectively, compared to 17% in drinks group). So, in terms of portfolio, **focus to keep in suply the most demanding products**.

Complementary products:

02

Average collection groups per order is **1,89** (which varies between 1.54 during the night and 2,13 during the noon). On the other hand, **TOP 25 products** (which represent **3%** of **total products**) bring **29,5% of total revenue**.
Taking these two facts togehter, for these **TOP products**, which mainly belong to **drinks group**, should be presented with **complementary product**s belonging to **other group categories** (breakfast and/or frozen products.)

4. Recommendations

03

Packs - product bundles:

In terms of revenue, **ice cubes** represent **1,7%** of **total revenue**. On average, orders with ice cubes products, have **2,5 extra products**, which are mainly drinks. Taking into account Recommendation #2, a **4-product pack** should be offered (ice cubes + alcoholic drink + blend + snack). Same situation for other kind of products.

04

Promotion multiple times:

TOP1 and TOP2 products in terms of quantity are **beer cans**, representing **10% on average of total quantities delivered**. Also, orders with those products, on average have **9 units** (**far from** the **6,29** average q/order seen in 3.KPIs). In such orders, in which the same product is added +7 times, a **promotion** of type "10x9 - take 10, pay 9", could be implemented without damaging the product margin.

5. Assumptions & Next Steps

Customer orders table - Assumptions

Product identification number: external_id.
Primary key identificator (for SQL): was not given. AN index number per row was used.
Product_price was taken as the right price data (divergences between that number and Items table number).

Items table - Assumptions

Product id does not match with external id in the customer orders tab.
Primary key used is the product name (as its a unique value and matches with the customer orders table).
Item cost is the only data used (price and unit sales do not match the other table).

Next Steps:

Revenue increases a 75% from November to December. Although a clear pattern is followed during the two months (orders increase in Thursday until Sunday), I have not identified the cause of the orders' increase from one month to the other. In the case of December, the orders categorized under "Navidad" represent just 2,5% of total orders. So, either the "Navidad" category is wrong, or the increase is due to seasonality.