

Coffe Vending Machine Project

We came across a dataset of coffee sales from a machine, containing fewer than 1000 rows and 6 columns.

1. CLEANING PROCESS IN EXCEL

The first thing we noticed was the .CSV file. We opened Excel to use Power Query to convert the file. Once we did that, we began the cleaning process.

Since it's a small project, everything was converted correctly. However, we noticed a couple of issues:

- We have two date columns. The first one gives us day/month/year while the other adds the hour. In this case, we should clean the last column to keep the hour and date separate.
- On the other hand, we found that the amount spent had lost the comma. Additionally, it's in Ukrainian Hryvnia. So, in this column, we need to return the decimal comma to its place and convert the currency back to Euros (even though the amount is going to be relatively cheaper than an actual Spanish coffee would be). However, for this purpose, we are going to keep the prices as they are.

We have some questions we want to answer through this analysis, such as:

- Which is the most frequently sold coffee?
- What's the maximum amount spent and the average amount spent?
- Are there any frequent clients based on the anonymous ID numbers?
- Are people more likely to pay with cash or card?
- What are the peak hours?
- How much is made from each type of coffee?

Once this was done, we created a table to search through filters for any rare values. Then, we started with pivot tables to answer the questions.

2. Data Analysis

We found some interesting results and began creating graphs.

- Which is the most frequently sold coffee? How much is made for each type?

Type	Profit	Operations
Latte	167,84 €	203
Americano with Milk	161,83 €	227
Cappuccino	146,19 €	177
Americano	88,91 €	145
Hot Chocolate	58,53 €	71
Cortado	53,18 €	86
Cocoa	23,28 €	28
Espresso	20,31 €	39
Grand	720,07 €	976

We found some interesting results and began creating graphs. But first, we already had some information to answer questions.

The most sold coffee is Americano with Milk, with a total of 227 out of 976 sales. However, it is not the highest revenue generator, as that title goes to the Latte.

- What's the maximum amount spent and the average amount spent?

Sells	MAX op	MIN op	AVG op
720,07 €	0,88 €	0,51 €	0,74 €

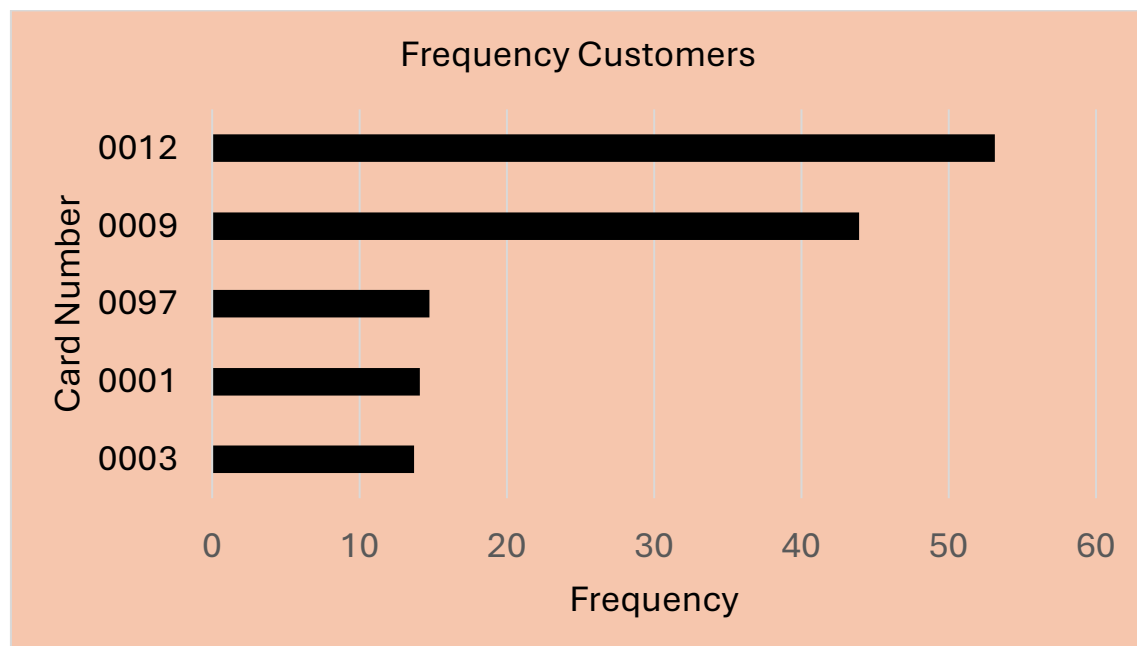
We found that the total revenue generated by the machine was 720 euros, with a maximum transaction amount of 0.88 euros, a minimum of 0.51 euros, and an average of 0.74 euros per transaction.

- Are people more likely to pay with cash or card?

Type	money	%	OP
card	649,98 €	90,27%	887
cash	70,09 €	9,73%	89
Grand	720,07 €	100,00%	976

9 out of 10 customer pay with a card.

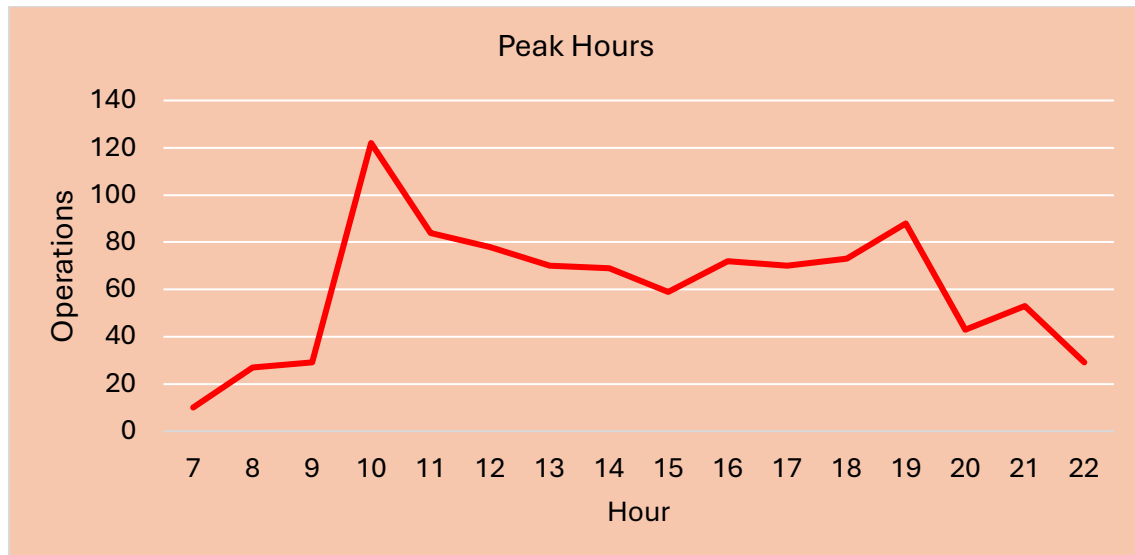
- Are there any frequent clients based on the anonymous ID numbers?



As we observed that 9 out of 10 customers pay with a card, we were able to track the cards to see if they were used more than once. We discovered that several customers did indeed

make repeat purchases with the same card multiple times over the months, with the best customer making over 50 transactions.

- What are the peak hours?

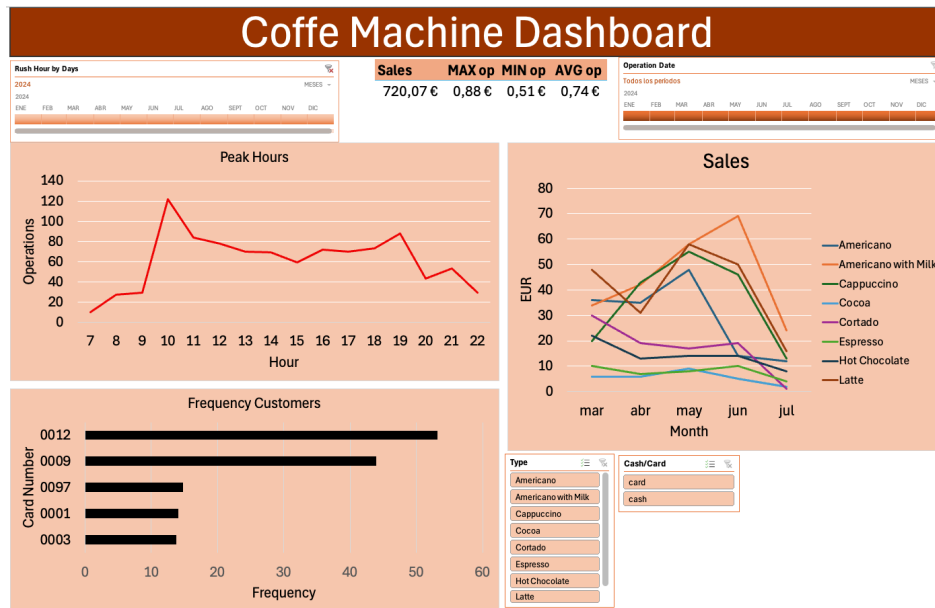


We observed that transactions occurred between 7 AM and 10 PM. We identified a peak hour around 10 AM and some activity around 7 PM.

3. Dashboard

We created a dashboard to display various data. It has an easy-to-understand interface with different dynamic filters. It allows us to see peak hours, the evolution of sales by product, and frequent customers. Additionally, it shows total sales as well as the average, maximum, and minimum sales.

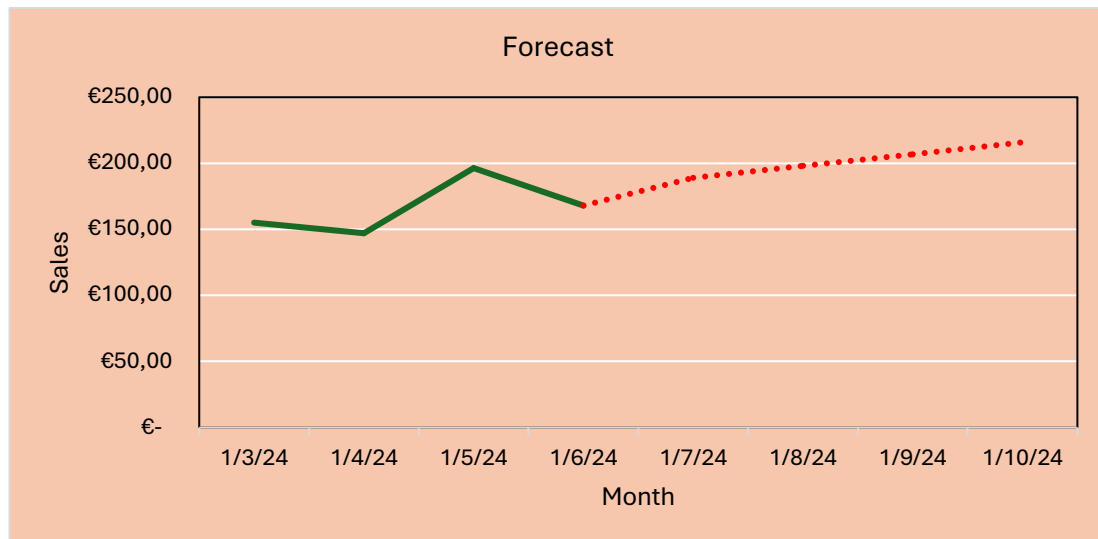
The dashboard is dynamic, enabling us to filter by coffee type, payment method, or timelines. These can be filtered by days to analyze sales or by peak hours to affect the hourly sales graph.



4. Conclusions

The analyzed period consists of 138 days (5 and a half months). The average daily coffee sales are 7 per day. Additionally, no trend is observed on specific days. Short-term forecasts are stable.

Unfortunately, we cannot determine the location of the machine. However, we can see that the machine is functioning correctly on an operational level.



Due to the objective of this analysis, the improvement proposals are:

- Ensure proper maintenance of the machine.
- Use a QR code to receive feedback.
- Remove products with decreasing demand (initially, the Cortado is at risk).
- Optimize inventory during the identified peak hours.