

MARCH 2023

Restaurant Case Study

Business Cases for Data Science

■ ■ ■ Basket Analysis

Lukas
Gross
20221363

Beatriz
Carmo
20220685

Karim
Miladi
20220720

Tomás
Domingos
20221370

Tomás
Vicente
20221355



Table of Contents

1. Executive Summary	2
2. Business Needs and Requirements.....	2
3. Methodology	3
3.1. Data Understanding.....	3
3.2. Data Pre-processing.....	4
3.2.1. Data Cleaning.....	4
3.2.2. Feature Engineering.....	4
3.3. Modelling	5
3.3.1. Delivery Perspective	6
3.3.2. Dine-In Perspective	6
3.3.3. Part of the Day Perspective.....	7
3.3.4. Seasonal Perspective	8
3.3.5. Weather Perspective.....	8
3.4. Evaluation.....	8
4. Results Evaluation	9
5. Deployment and Maintenance Plans.....	9
6. Conclusions	11
7. References.....	11
8. Appendix	12

Table of Figures

Figure 1 - Average total amount by product family (Delivery)	12
Figure 2 - Average total amount by product family (Restaurant).....	12
Figure 3 - Numerical features distributions	13
Figure 4 - categorical features distributions	13
Figure 5 - Correlation matrix for quantitative and binary variables	14
Figure 6 - Boxplots for quantitative variables	14
Figure 7 - Visualizing relationship between product families (Delivery Perspective)	15
Figure 8 - Relationships between various product designations (Delivery Perspective)	16
Figure 9 - Relationships between product families (Dine-in).....	17
Figure 10 - Relationships between product designations (Dine-in)	17

1. Executive Summary

The restaurant business has become increasingly competitive, with many customers opting to order food to their desired location rather than dining in. This has put pressure on established restaurants to maintain their operations at normal levels amidst the continuous opening of new competing restaurants with new concepts and offerings. C, a company with several restaurants in Cyprus, is facing a similar challenge with one of its brands specializing in Asian cuisine struggling to maintain its profit margin and growth due to increased competition and changes in customer habits. To combat this, C plans to leverage its sales data to better understand customer consumption patterns and preferences.

With this in mind, the relations between the many menu products were analysed and the most common combinations were identified. This analysis was made according to different perspectives since it is normal for restaurants to have different attendance and order rates according to the time of the year. It is also usual for orders to differ depending on where the customers chose to eat: dine-in or at home via a delivery service.

With this analysis, it was possible to propose a new menu (with prices) to the restaurant; one with less items and products that (according to the analysis) the customers like to consume.

2. Business Needs and Requirements

The restaurant industry is a highly competitive one, particularly now that food delivery platforms are becoming more popular. A restaurant is a business that serves cuisine and, ideally, an enjoyable experience to those who dine there. As a result, it is anticipated that an institution like this adhere to the following standards:

- Quality food and drinks
- Friendly and polite servers
- Clean and cosy atmosphere
- Efficient kitchen service

In highly competitive business environments, such as the food service industry, it is crucial to remain agile and responsive to changing customer needs, preferences, and expectations. While the rise of food delivery apps has shifted the focus towards food quality and delivery efficiency, it is important to not overlook the importance of maintaining high standards across all aspects of the dining experience.

To succeed in this dynamic and rapidly evolving market, businesses must remain vigilant and attentive to the needs and habits of their customers. This can be achieved by gathering and analysing data on customer orders, preferences, and behaviour, as well as seeking out feedback and suggestions for improvement. By prioritizing customer satisfaction and continuously adapting to their changing needs and expectations, businesses can maintain a competitive edge and ensure long-term success.

There are a few more critical issues to consider, one of which is the menu's length, which is also one of the most essential. Large menus do not perform well most of the time because they require a wide range of products to be purchased (which can be a barrier to keeping ingredients fresh and also demand many different suppliers), a busier kitchen team (they must know how to cook every single item on the menu), and ultimately it can cause customers to take longer ordering and be indecisive about what they want to eat. A basic but delicious menu with items and dishes that the customers enjoy is therefore essential.

Customer segmentation and market basket analysis are critical components in attracting and retaining consumers across various platforms, including digital or dine-in services. These techniques allow restaurants to identify customer groups with specific needs and preferences, enabling them to create tailored campaigns that successfully meet these needs.

Furthermore, customer segmentation and market basket analysis can help identify menu items that are not popular with customers. The elimination of these menu products can improve menu efficiency, reduce food waste, and increase restaurant profitability. Restaurants can guarantee a more streamlined and successful dining experience, resulting to higher customer satisfaction, by strategically utilizing customer segmentation and market basket analysis.

3. Methodology

3.1. Data Understanding

Before making any changes to the dataset, it is critical to understand how it is structured and to identify any potential inconsistencies. To accomplish this, the dataset was thoroughly examined and analysed.

Several instances of duplicated rows within the dataset were discovered during the examination process, as well as the existence of missing values in two variables, namely **Age** and **DocIDHash**, with 4172 and 1001 missing values, respectively. Furthermore, several anomalies in the dataset were discovered, which were then substituted with the value 'NaN' and will be treated as missing data in future analyses.

The features in the dataset were classified into qualitative, quantitative, and binary traits to aid in the modelling process.

When the data frame was filtered using Boolean indexing, it was discovered that the value for **Customer City** was consistently recorded as 'NaN' whenever the associated **Customer ID** was 0. The same pattern was seen with the **Customer Since** variable, which was consistently recorded as 'NaN' in the instances mentioned above. In this instance, if the **Customer ID** column contains a 0 quantity, both the **Customer City** and **Customer Since** columns will be empty.

The variable **ProductFamily** was looked into to better understand how the restaurant organizes its menu items, and a dictionary was made to extract the **ProductDesignation** values associated with each **ProductFamily**. This enabled the identification of various menu categories or families, as well as the specific menu items that fit within each category. This data can help with menu planning and organization, as well as future data analysis. The creation of this dictionary proved to be an effective method for analysing data and getting insights into menu item organization.

To better understand the differences between the restaurant's delivery and dine-in services, an analysis of the average total number of items ordered by product family for each service was conducted. To visualize the results, two plots were created. (Figure 1 and Figure 2). The routines of customers who dine in and those who favour take-out were found to be considerably different, implying that separate analyses for each service should be conducted in the future.

A small inconsistency in the soup family was discovered during the data analysis process. It was specifically observed that the drink and food designations within this family had unusual frequencies that were not easily explained.

It is important to mention that the restaurant's financial difficulties were considered during the analysis. However, it is too soon to infer that the data inconsistencies are the result of "book

cooking," or accounting tricks used to make the company's financial results appear better than they are. To reach any firm findings, more research and analysis are required.

Following an examination of delivery habits, the locations from which customers made orders were investigated. Despite having delivery services accessible, some cities, including STROBOLO, TSERI, LAKATAME, LAKSTAMEIA, SYN., ANTHOYPOLIS, GERI, and PANO DEYTERA, had low frequency counts. Further research revealed that when *IsDelivery* was set to 0, these cities had no corresponding values in the *CustomerCity* column. As a result, it is reasonable to presume that delivery services are not available in these cities.

As a result, it is critical to investigate the distribution of variables. Histograms were used to evaluate the numerical features to accomplish this (Figure 3). Surprisingly, the months with the highest business activity were March, November, and December, rather than the expected summer months. Furthermore, the evening and afternoon time slots (lunch and dinner time) had the greatest workload.

To examine the distribution of categorical variables, several bar plots were created. (Figure 4). According to the plots, most customers prefer to dine-in at the restaurant rather than use the delivery service. Furthermore, the data shows that the restaurant has a higher volume of business outside of the holiday season.

In addition to this exploration, the correlation between variables was taken into consideration as well. However, there were no high correlated variables (Figure 5).

It is worth mentioning that this dataset included several inconsistencies that were not resolved or addressed during the analysis process. This was since the restaurant was facing bankruptcy, and it was expected that there would be some organizational issues that would be reflected in the data. Therefore, it was deemed necessary to include these inconsistencies in the analysis to provide a comprehensive overview of the restaurant's situation at the time of data collection.

3.2. Data Pre-processing

3.2.1. Data Cleaning

After exploring the data frame possible issues, it is important to solve them so the data will be ready to be process by the algorithms.

The first thing to be done was to remove duplicates and variables with too many missing values (*CustomerSince* and *CustomerCity*). The hour format was also changed to ease future processes.

As expected, the data had some outliers (Figure 6) that were removed manually after concluding it would remove less data from the original dataset.

The rules applied to removed outliers manually were as follows:

- **TotalAmount** < 760
- **Pax** < 100
- **Qty** < 25

After removing outliers manually, **99.991 %** of the original Customers were maintained.

3.2.2. Feature Engineering

Feature engineering was used on a dataset in this project to create new variables and transform existing ones for use in machine learning models.

The first transformation involved creating a new variable **Lunch/Dinner** based on a comparison of the variables **InvoiceHour** and **SunsetHour**. This variable was created to help differentiate between lunch and dinner meals.

The subsequent involved using the **strftime** method to convert the variables **SunsetHour** and **InvoiceHour** to a more readable format. This simplified data interpretation and improved visualization of trends over time.

The dataset was then divided into two separate data frames based on the **Lunch/Dinner** variable, with one containing lunch data and the other containing dinner data. This division assisted in separating the two meal types and allowing for more targeted analysis of each.

The **TimeStamp** variable was then split into two new variables, **Month** and **Time**. The variable **Month** was used to group data by month and identify seasonal trends, whereas the variable **Time** was used to group data by hour and identify patterns throughout the day.

Additionally, two new variables, **Season** and **PartOfDay**, were added to further categorize data by season and time of day. The **Season** variable was created by dividing the **Month** variable into four seasons, and the **PartOfDay** variable by dividing the **Time** variable into six parts of the day. These variables aided in providing additional context for data analysis and improving the accuracy of machine learning models.

The food items were ready for categorization after some pre-processing. The categories 'VALENTINES BUFET' and 'SPECIAL BUFFET' were efficiently assigned to the 'HOLIDAY' category after being discovered to be holiday-specific. The 'SOUPS' category was excluded from the analysis because it is a non-product family. Certain categories were given more intuitive names to make the data more understandable. The categories 'JAP SUSHI' and 'NEW SUSHI' have been merged and renamed 'SUSHI'. 'INDIAN SIDES' was created by combining the categories 'RICE IND', 'VEG IND', and 'BREADS IND'. To better reflect their contents, the categories 'RICE' and 'VEG' were renamed 'SIDES'. Finally, the category 'TSANTES' was renamed 'BAG' to improve descriptive accuracy.

Following this, a list of product designations was created to be evaluated and an empty list to hold low-demand products. A dictionary of product families and corresponding items is then used for each product designation to create a data frame with dummy variables for each product. After that, the dummy variables are transformed into the corresponding total amounts for each product, and the data is grouped by order number to calculate the mean amount for each product. Products with mean quantities less than a certain threshold are identified as low-demand and added to the list. Finally, the percentage of low-demand products is calculated and removed. After removing low-demand products, the percentage of observations kept is **84.08%**.

3.3. Modelling

After removing the low-demand products, the data was ready for the modelling part of the project. In order to find the most convincing association rules with the **apriori algorithm** it was decided to split the data in order to be able to evaluate patterns from different perspectives. First, we split the dine-in and the delivery orders and analysed them separate from another. Within these two perspectives, association rules between the product families were looked into first. After the same was done for the product designations. This helped gaining deeper insights into the buying behaviour of the customers. It is important to add that two different perspectives for the dine-in orders were considered; one in which all products were included and one in which the product mineral water was removed from the data set.

This step was necessary because the proportion of mineral water in the orders was too high and this negatively influenced the associations, since a large part of these consisted of mineral water and the respective other product. Thus, it was assumed that mineral water is generally ordered in almost all orders and that there is not necessarily a correlation with other products.

After this step, it was decided to add three more perspectives, to gain even deeper insights into the data. The part of the day perspective, the seasonal perspective, the weather perspective and the holiday perspective.

3.3.1. Delivery Perspective

The analysis of the product families showed that in terms of support, SIDES, MEATS, and STARTERS were the most probable possible combos.

When it came to quantity, the most probable requested combos were **Sushi/Starters**, **Starters/Extra**, and **Starters/Sizzling**. Additionally, we found that the most valuable families were MEAT, STARTERS, SIZZLING, and SUSHI.

When examining unique customer **IDs**, it was discovered that MEAT, STARTERS, SIZZLING, and SIDES combos were the ones with the most distributed demand. Moreover, the analysis showed a strong dependence (**94%**) of EXTRAS on STARTERS, indicating that customers frequently order the first one with the latest.

Lastly, it was time to look at the lift of different food item pairs. The analysis revealed that an **Indian main dish** would multiply the probability of an **Indian side** to be ordered by **4.64**. This indicates that there is a strong association between Indian main dishes and Indian sides, and that customers are likely to order them together.

As a last step the association rules were plotted for a visual analysis of the links between the product families (Figure).

The analysis of the product designations revealed that the most probable possible combos included EGG FRIED RICE/SWEET SOUR CHICKEN/SPRING ROLL. In terms of quantity, the most probable requested combos were SPRING ROLL/LEMON CHICKEN or SPECIAL MONTPARNASSE RICE, SESAME PRAWN, STEAMED RICE and NOODLES WITH MEAT. Additionally, we found that the most valuable combos were 1/4 DUCK/BEEF BBS; SWEET SOUR CHICKEN/1/4 DUCK; SWEET SOUR CHICKEN/BEEF BBS.

When we looked at unique customer IDs, we found that Meat, Starters, Sizzling, and Sides combos were the ones with the most distributed demand. We also examined the confidence and lift of different food item pairs. Our analysis showed that EXTRA PANCAKES/EXTRA SAUCE were items that occurred together frequently, and JIRA PULAO/NAAN had a strong lift, indicating that they were often ordered together.

A plot was also created to better visualize these links (Figure).

3.3.2. Dine-In Perspective

The analysis for the product families of the dine-in perspective showed strong correlations with the delivery perspective. It revealed that SIDES, MEATS, and STARTERS were the most probable possible combos in terms of support, which aligns with the previous finding.

It was also found that SUSHI/STARTERS, STARTERS/EXTRA, and STARTERS/SIZZLING were the most probable requested combos, resulting in much higher revenue, which is consistent with the previous finding on quantity.

Furthermore, it was concluded that MEAT, STARTERS, SIZZLING, and SUSHI were the most valuable families in terms of total amount, which is in line with the previous finding on valuable combos. It was impossible to find any unique customer IDs related to the product families in this perspective.

In terms of confidence, it was observed a strong dependence (**94%**) of **Extras** on **Starters**, which aligns with the previous finding.

Additionally, the analysis revealed that MAIN IND, INDIAN SIDES, and START IND were all families that multiplied the chances of each one being selected in the same order, indicating a strong association between these families, which is consistent with the previous finding on lift.

As a last step we plotted the association rules for a visual analysis of the links between the product families (Figure).

After analyzing the existing relationships between the various families of products, the same was process was applied to the product designations.

It was found that EGG FRIED RICE/SWEET SOUR CHICKEN/SPRING ROLL were the most probable possible combos. When it came to quantity, SPRING ROLL/LEMON CHICKEN or 1/6 DUCK, SESAME PRAWN, STEAMED RICE and BABY PORK were the most probable requested combos.

Additionally, it was discovered that the most valuable combos were 1/4 DUCK/BEEF BBS; SWEET SOUR CHICKEN/1/4 DUCK; SWEET SOUR CHICKEN/BEEF BBS. The analysis did not yield any results for unique customer IDs.

However, the confidence and lift of different food item pairs was examined. The findings showed that the items paired with EGG FRIED RICE tended to be order together frequently.

Moreover, it was found that CHICK TIKKA MASALA/JIRA PULAO/NAAN had a strong lift, suggesting that these items were often ordered together.

Once again and to aid the interpretation of the association rules, a plot was created to illustrate these links (Figure).

3.3.3. Part of the Day Perspective

In this perspective, the goal was to analyse the association rules of the afternoon evening and night. These parts were chosen because the exploration of the data indicated that there were the most orders in these parts of the day.

The **afternoon perspective** of the data revealed that the most probable possible combos during this time were EGG FRIED RICE with SPRING ROLL, SWEET SOUR CHICKEN, and BEEF BBS. When it came to quantity, STEAMED RICE with SPRING ROLL, NOODLES WITH MEAT with SPRING ROLL, and SPRING ROLL with SESAME PRAWNS were the most probable requested combos during the afternoon.

In terms of total amount, the most valuable combos during this time were 1/4 DUCK/BEEF BBS, 1/4 DUCK/SWEET SOUR CHICKEN, and SWEET SOUR CHICKEN/BEEF BBS.

There were no unique customer IDs in this perspective.

Finally, the consideration of the confidence showed that if BEEF BBS was ordered, then the probability of EGG FRIED RICE being ordered was **74%**.

The **evening perspective** of the data showed that there were no significant differences in combos compared to the afternoon perspective. However, in terms of quantity, there were no averages greater or equal to 3 compared to the afternoon perspective. This suggests that there is less group eating and more dining in pairs during the evening.

When it comes to total amount, although the frequencies are higher for dining in the evening, the average amount spent is smaller.

The most valuable combos during this time are 1/4 DUCK/SWEET SOUR CHICKEN and SWEET-SOUR CHICKEN/BEEF BBS.

The **night perspective** also cites great similarities with the two previous perspectives; however, this perspective is also the least profitable which indicated that the focus should be in the afternoon service since it is the most profitable. It is important not to neglect this later part of the day, however, and try to make it more profitable.

3.3.4. Seasonal Perspective

The objective of this perspective was to identify association rules within the dataset for the four distinct seasons of the year, namely Winter, Spring, Summer, and Fall.

This analysis wasn't as immediate and easy as the one done in the previous perspectives. But, despite the difficulties, it was clear that some products should be advertised more during specific seasons.

This is the case of the JIRA PULAO/NAAN combo, that is less present on Spring and Winter and therefore it would be interesting to coerce clients into buy this combo more times.

During Fall and Summer, LEMON CHICKEN/EGG FRIED RICE is one of the combos with less requests. Hence, just like before, it should be advertised more to make customers chose these items more frequently.

3.3.5. Weather Perspective

The data was also analyzed through a weather perspective lens. However, there weren't any major changes in comparison to the previous views.

3.4. Evaluation

After analysing the data trough different perspectives, it became clear that, as expected, the restaurant's menu is far too extensive and that are items that are rarely asked by clients while others have a high request rate.

The findings reveal some differences between dine-in and delivery customers. For instance, customers placing delivery orders tend to order specific item combinations, such as Egg Fried Rice with Sweet Sour Chicken and Spring Rolls, while certain products such as drinks and wines are less popular for delivery orders. On the other hand, dine-in customers often order item sets at the same price point to increase the total amount, and certain product combinations, such as Chicken Tikka Masala with Naan Bread and Jira Pulao, are more preferred.

When it comes to delivery orders, the most requested combos were as follows (according to different measures):

- **Support:** EGG FRIED RICE/SWEET SOUR CHICKEN/SPRING ROLL
- **Quantity:** SPRING ROLL/LEMON CHICKEN or SPECIAL MONTPARNASSE RICE, SESAME PRAWN, STEAMED RICE and NOODLES WITH MEAT

- **Total Amount:** 1/4 DUCK/BEEF BBS; SWEET SOUR CHICKEN/1/4 DUCK; SWEET SOUR CHICKEN/BEEF BBS
- **Unique Customer IDs:** Meat, Starters, Sizzling, and Sides
- **Confidence:** EXTRA PANCAKES/EXTRA SAUCE
- **Lift:** JIRA PULAO/NAAN

Considering this, it would be beneficial to include these items and combos in a future menu.

There aren't many differences between delivery and dine-in orders. But the ones that exist should be taken into account in the future. The most popular product combinations in dine-in orders were as follows:

- **Support:** EGG FRIED RICE/SWEET SOUR CHICKEN/SPRING ROLL
- **Quantity:** SPRING ROLL/LEMON CHICKEN or 1/6 DUCK, SESAME PRAWN, STEAMED RICE and BABY PORK
- **Total Amount:** 1/4 DUCK/BEEF BBS; SWEET SOUR CHICKEN/1/4 DUCK; SWEET SOUR CHICKEN/BEEF BBS
- **Confidence:** the combos mentioned on support, where the item paired with EGG FRIED RICE
- **Lift:** CHICK TIKKA MASALA/JIRA PULAO/NAAN

As before, these items and combinations must not be forgotten when elaborating a possible new menu.

4. Results Evaluation

While applying the **apriori algorithm** to be able to get the links and combos between the various menu items, it was clear that some products had a lower request rate while others appeared connect with specific sides or starters. The relationships between the many items were clear and easy to identify while using the algorithm and taking into considerations the different measures (support, lift, confidence, etc.).

Because of the nature of the restaurant business, it is crucial to be able to identify patterns in the orders and also to know which items are not ordered as frequently. This will help in the creation of a new menu, something that is also very important for this specific restaurant since it has a big menu.

5. Deployment and Maintenance Plans

Restaurants should all have great service, food, staff and also a good menu. A good menu, most of the times, small and with the items that customers love the most.

After analysing the data, it was found that there are certain food combinations that are more popular among delivery and dine-in customers. To capitalize on this, the restaurant should consider implementing a cross-selling strategy to offer customers the most profitable and demanded food combos. This can be done by training staff to suggest complementary items during the ordering process, or by incorporating a recommendation system into the restaurant's online ordering platform.

Furthermore, the restaurant's extensive product range may be hindering its ability to provide high-quality dishes consistently. To ensure the quality of the food is maintained, the restaurant

should consider streamlining its menu offerings to focus on the most profitable and popular dishes. This can help the restaurant to better allocate its resources and reduce food waste.

Keeping this in mind, the proposed menu consists of less items than the original one. However, the menu still has many items since it is dangerous, and it might not be smart to remove that many items – the restaurant still needs usual clients to consume there and also removing too many items might result in excluding some of the customers' favourite dishes.

Following this premise, the best way to deal with a still not so small menu is to try and control the customers' orders. This can be done by suggesting side dishes and appetizers from the moment one selects a main dish. These suggestions should be done according to the most popular combos found in the previous analysis, so their probability of success is higher.

Of course, customers must still have the option to choose other menu items. But presenting them interesting food combinations might incite them to choose them. This will make it easier for the team in the kitchen to get organized and prepare the dishes. The main idea is for client to choose their meal sequentially. Starting by selecting a starter, then a side and finally a drink and an appetizer (within a certain category: Chinese, Indian, Japanese) and be more reluctant to explore other combos that are not mentioned directly in the menu. On top of this, increasing the value of individual items would be an effective way to encourage customers to choose those combinations.

However, this strategy doesn't work as well in the delivery part of the business as it does in the dine-in. Customers who order via delivery already know what they want to eat and normally they order only for themselves (single order). To target these customers, it would be interesting to just present the single items as they are and not focus on cross-selling so much. Delivery clients are more certain of what they wish to eat and harder to convince to choose another dish. Therefore, it is better to stick with what is already working.

Another good strategy is to increase the prices of food during dinner since its when the restaurant has more movement. Contrary to what would be expected, the prices during lunch are higher than those during the dinner services. This problem should be solved as soon as possible to get more profit from the restaurant.

Of course, to maintain a successful business, it is not enough to make one single change. The business must be constantly looking for new ways to make profit. This includes collecting data regularly and performing new basket analysis to understand how the customers' preferences change and what dishes should be added or removed from the menu.

Changing the menu from time to time is also very useful because it will attract new customers who may like the new dishes better than the old ones, but also usual customers who are now curious to taste the new proposals. By regularly monitoring the costs and profitability of each dish, the restaurant can make informed decisions about which products to remove or adjust in order to maximize profits. It is also recommended that the restaurant implement a regular review of its menu offerings to ensure the most profitable and effective product range is being offered to its customers.

In a dynamic world like the one the restaurants find themselves in, it is essential not to stay the same for too long. Change is inevitable and also very important and beneficial when done correctly and in the right time.

One, however, must not forget about the restaurant's staff. It is of extreme importance that they are informed of every change in course and that they understand why they are crucial to the business. It is also useful to educate them on the possible future changes. A good team makes a good business, so having happy and encouraged employees is key.

6. Conclusions

After conducting a thorough analysis on customer segmentation and product offering, it has been concluded that there are notable differences between the preferences of dine-in and delivery customers. The study indicates that delivery customers tend to gravitate towards particular combos and items such as egg fried rice, sweet sour chicken, and spring rolls, while dine-in customers are more inclined towards purchasing sets of items at the same price rather than individual items.

Furthermore, the study recommends implementing cross-selling strategies to boost the total amount of orders. For example, suggesting side dishes, drinks, or desserts to customers during the ordering process can increase the value of each order.

In addition, the research indicates that the current product offering is too extensive, and a more limited yet effective range of items can bring better results. The analysis proposes identifying the most in-demand and profitable products within each family and removing products with low demand. However, it acknowledges that determining profitability can be challenging due to the limited data on costs.

To improve the product offering and satisfy customers' preferences, the study recommends conducting surveys or collecting feedback from customers to understand their likes and dislikes. This approach can help the restaurant to refine its menu and focus on offering the most popular and profitable items, ultimately enhancing customer satisfaction and loyalty.

7. References

- <https://sevenrooms.com/en/blog/why-to-have-limited-menu/>
- <https://github.com/toomingos/Restaurant-Case-Study/blob/e3eb591b9a36e29714854556f77889d2fd2225c8/Step-by-Step%20Product%20Removal%20Explanation.pdf>

8. Appendix

Average Total Amount by Product Family (Delivery)

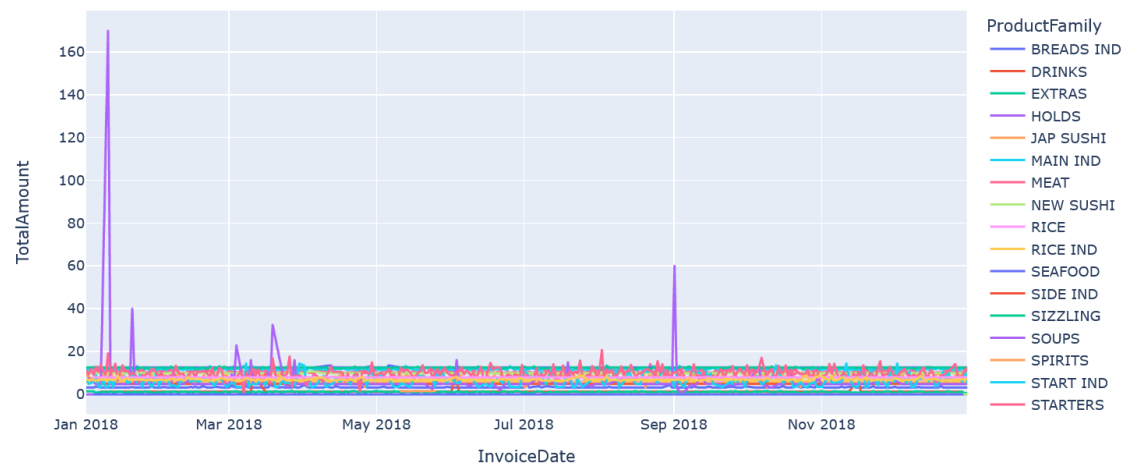


Figure 1 - Average total amount by product family (Delivery)

Average Total Amount by Product Family (Restaurant)

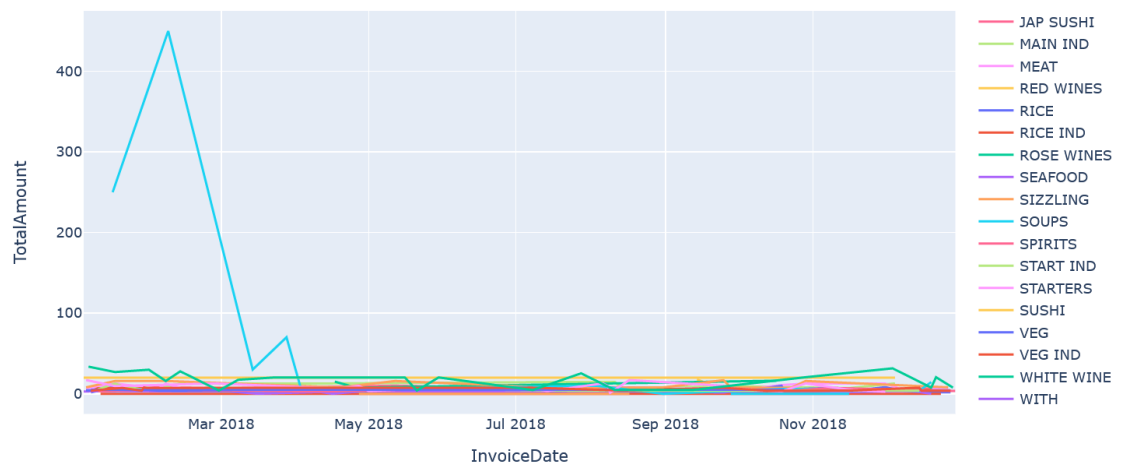


Figure 2 - Average total amount by product family (Restaurant)

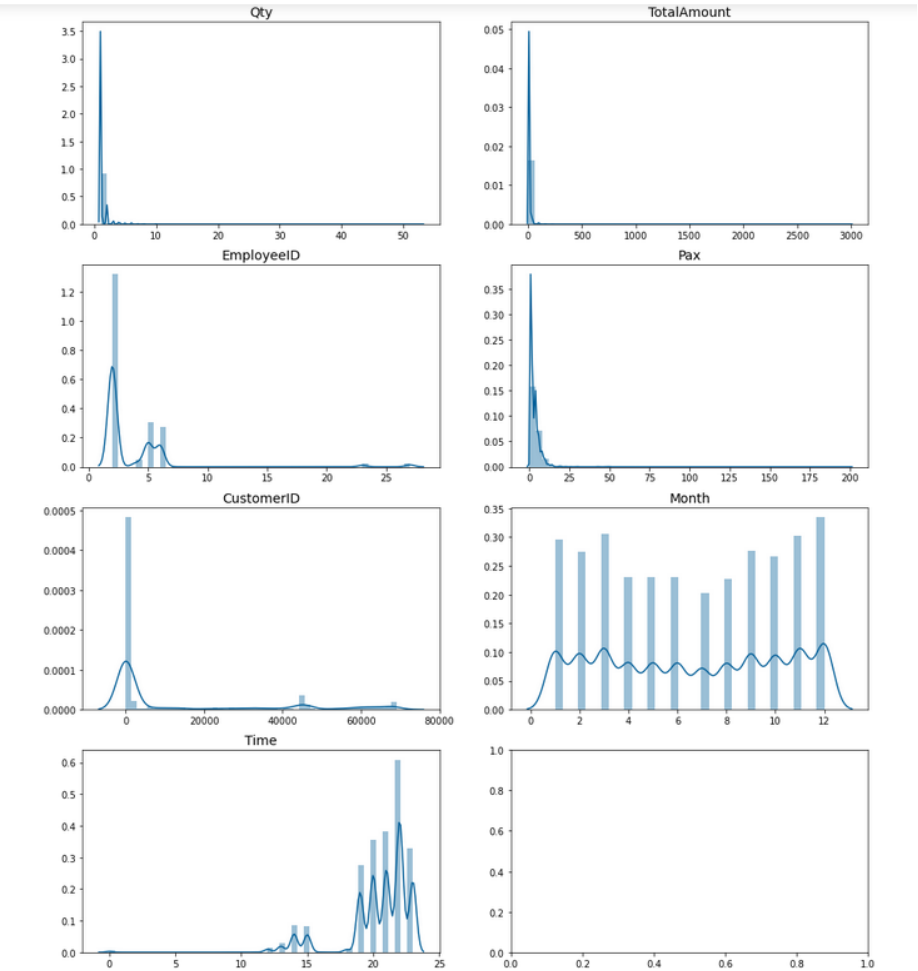


Figure 3 - Numerical features distributions

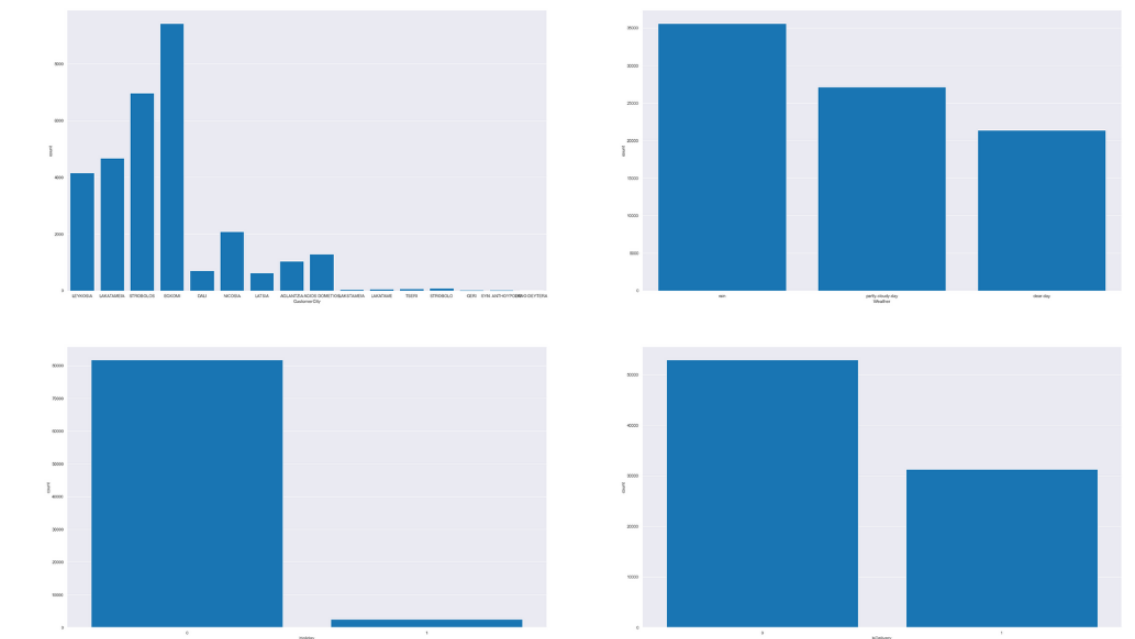


Figure 4 - categorical features distributions

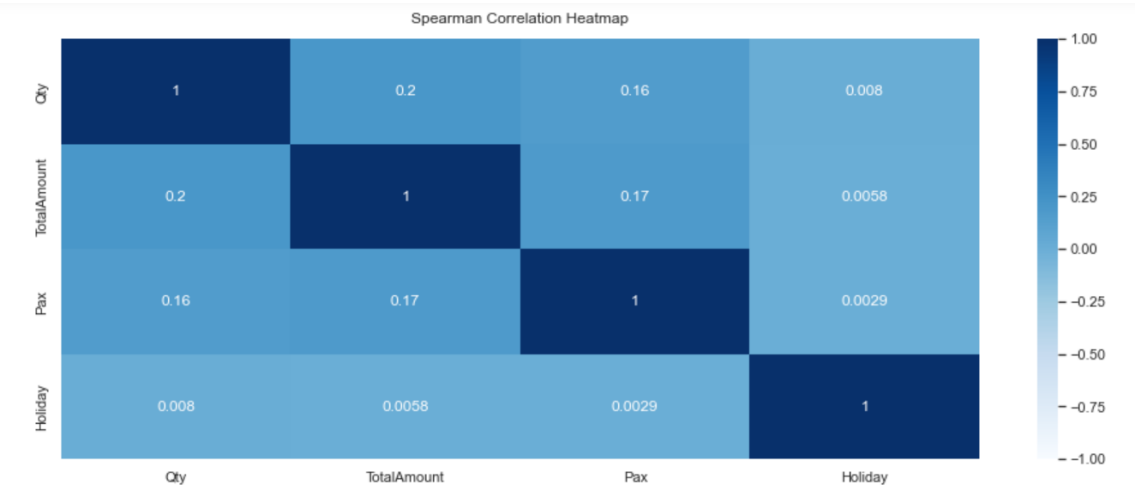


Figure 5 - Correlation matrix for quantitative and binary variables

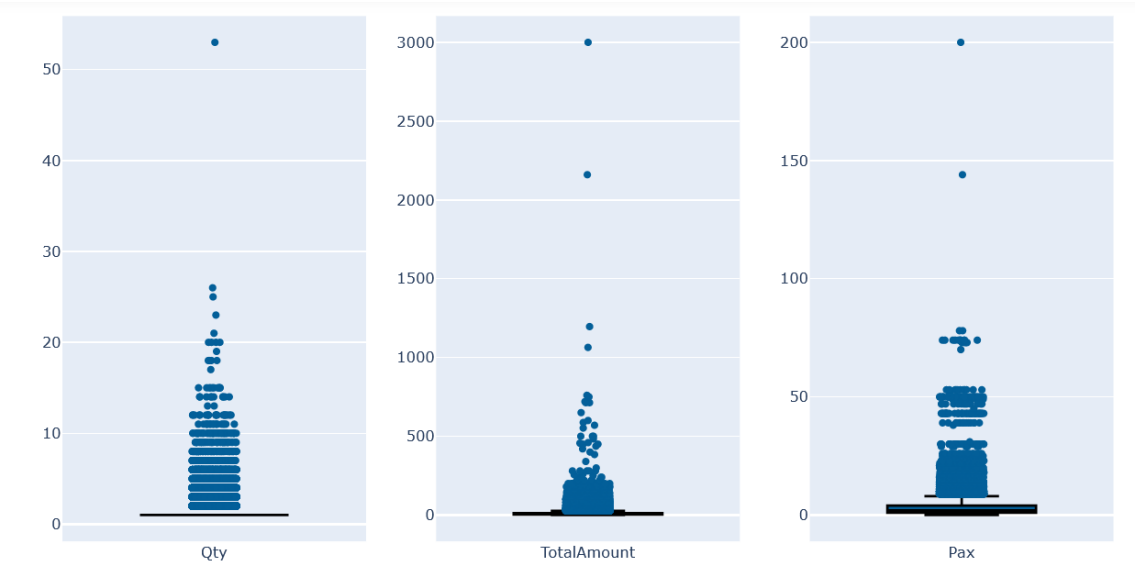


Figure 6 - Boxplots for quantitative variables

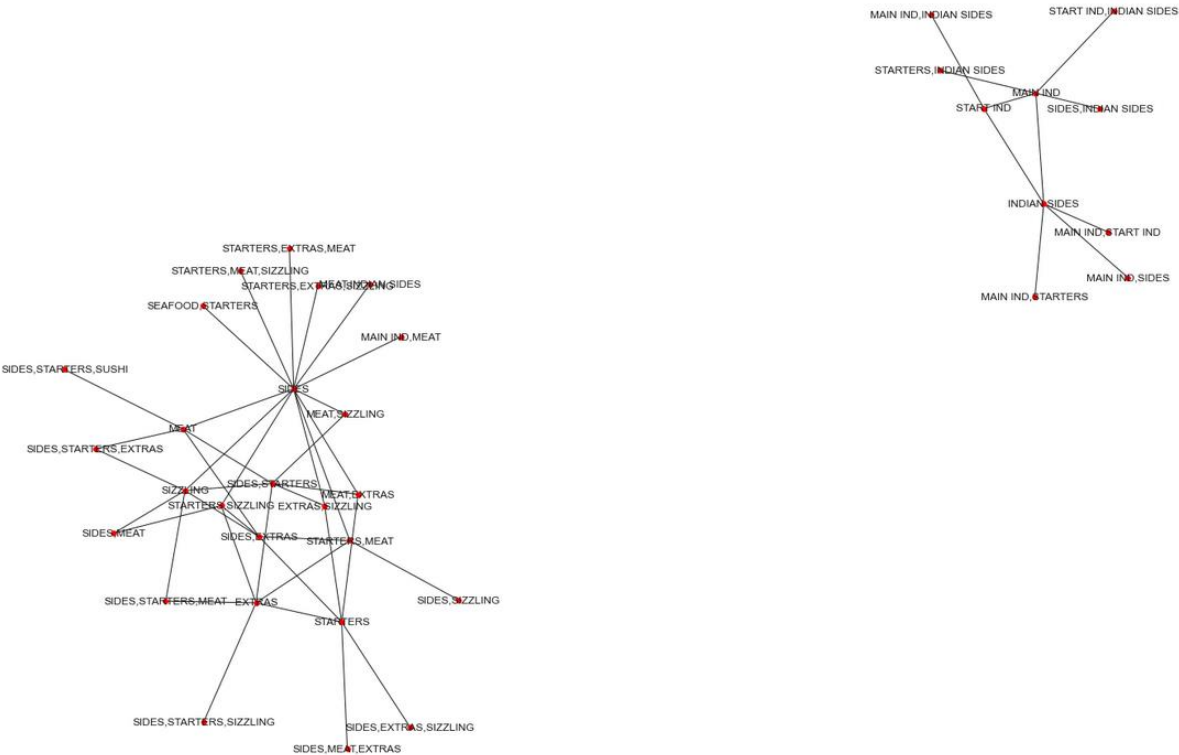


Figure 7 - Visualizing relationship between product families (Delivery Perspective)

Figure 9 - Relationships between product families (Dine-in)



Figure 10 - Relationships between product designations (Dine-in)