

Master Degree Program in
Data Science and Advanced Analytics

Business Cases with Data Science

Case <2>: <MARKET BASKET ANALYSIS>

<Engjulla>, <Hasani>, number: <20221404>

<Danylo>, <Ivanchyshyn>, number: <20221692>

<Group B>

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa

<March>, <2023>

INDEX

1. EXECUTIVE SUMMARY	2
2. BUSINESS NEEDS AND REQUIRED OUTCOME	3
2.1. Industry Background	3
2.1.1.1. Restaurant Background	4
2.2. Business Objectives	4
2.3. Business Success criteria	4
2.4. Situation assessment	5
2.5. Determine Data Mining goals	6
3. METHODOLOGY	7
3.1. Data understanding	7
3.2. Data Preparation	9
New Menu Recommendations:	12
Delivery vs Eat in:	13
New Product Recommendations:	13
4. Recommendations based on different visualization insights and Deployment	14
5. CONCLUSIONS	15
Considerations for model improvement	15
5. APPENDIX	16
References:	19

1. EXECUTIVE SUMMARY

Customers are increasingly ordering food and beverages for delivery instead of dining in restaurants, which is increasing competition in the restaurant business. This trend has raised the strain on established restaurants to maintain operations in the face of the constant introduction of new eateries with diverse concepts and menus. C, a firm that operates many restaurants in Cyprus, faces a dilemma with one of its brands that specializes in Asian cuisine, specifically Chinese cuisine. Due to increased competition and a shift in consumer preferences, the brand is battling to retain its profit margin and steady growth. To address this obstacle, C intends to evaluate its sales data using market basket analysis to uncover clients' purchasing habits and preferences. This strategy will enable the organization to make data-driven decisions to enhance the performance of the brand and preserve its competitive advantage in the market.

By conducting a market basket analysis, our team of data scientists intends to assist Restaurant C in comprehending its consumers' interests and behavior. This technique involves studying client transactions to see which things are frequently purchased together, allowing us to discover product affinities and design sales and customer happiness plans. By doing this study, we can offer Restaurant C significant information into the performance of its Asian cuisine brand, creating new menus, and providing new marketing and sales cross selling strategies

Market basket analysis necessitates sophisticated technologies and voluminous amounts of data. Yet, the insights gleaned from this study can be invaluable for organizations seeking to remain competitive in a congested marketplace. By exploiting their sales data and employing market basket analysis, C is able to make data-driven decisions to sustain their market advantage.

2. BUSINESS NEEDS AND REQUIRED OUTCOME

During the Business Understanding phase of the CRISP-DM methodology, our team of consultants endeavored to get a thorough grasp of Restaurants C's business issue. We studied the restaurant's customer data to uncover patterns and trends that may aid in the development of effective client segmentation strategies. After comprehensive analysis of the data, we discovered the need for a more effective customer segmentation in order to better focus marketing activities and increase revenue. Our team advocated employing cluster analysis to categorize clients based on shared features, hence enabling more targeted and personalized marketing activities. By following this strategy, restaurant C will be in a better position to comprehend and adapt to the individual interests and behaviors of its broad client base, resulting in greater customer loyalty and profitability.

2.1. INDUSTRY BACKGROUND

According to Cyprus Statistical Service data, the restaurant and catering industry in Cyprus grew steadily from 2016 to 2019, with the number of restaurants increasing from 3,995 to 4,209 during this time period (1). The COVID-19 epidemic, on the other hand, had a severe impact on the sector, forcing numerous restaurants to close temporarily or permanently owing to lockdown measures.

The number of restaurant and catering establishments in Cyprus fell to 3,715 in 2020, an 11.3% fall from the previous year (2). The pandemic also had an effect on revenue, with turnover declining by 33.3% in 2020 compared to the previous year (3).

According to the European Bank for Reconstruction and Development, restaurants in Cyprus face increased competitiveness in this hard environment, with new venues emerging on a regular basis (4). As a result, in order to remain competitive, restaurants must constantly adapt to changing client tastes and develop new tactics.

Given this context, it is obvious that Restaurant C and other Cyprus establishments must work hard to maintain their market position and profitability.

2.1.1.1. Restaurant Background

Restaurant C is a restaurant chain in Cyprus that has various restaurants that are categorized under distinct brands according to their location, concept, and cuisine type. The company has been in business for more than two decades and has established itself as a major participant in the nation's highly competitive restaurant market. The restaurants provide a range of dining experiences, from casual to fine dining, and serve a mix of Asian, European, and Mediterranean cuisines. One of the company's brands specializing in Chinese cuisine has recently struggled to maintain its profit margin and growth due to increased competition and shifting customer preferences.

2.2. BUSINESS OBJECTIVES

Restaurant C intends to use sales data to acquire insights into a number of major issues. Creating set menus, offering new items, understanding replacement products, recommending/promoting cross-selling, consumer segmentation, and maybe exposing other areas for development depending on the findings are among the obstacles. Restaurant C seeks to obtain a better knowledge of its clients' consumption patterns and preferences by using market basket analysis on its sales data.

In addition to the challenges outlined above, Restaurant C aims to use the market basket research information to uncover popular combinations of items purchased together, which might be used to build set menus or promotions that appeal to customers. The data will also assist the organization in determining which items people frequently substitute for one another, providing insight into future menu modifications that could boost customer happiness and drive sales. Restaurant C may adapt its marketing and promotional activities to certain groups by segmenting its consumers based on their tastes and purchasing behaviors, resulting in more successful cross-selling and upselling opportunities.

Overall, the market basket analysis of Restaurant C's sales data has the potential to offer significant insights that can help the company handle its core difficulties and achieve its business objectives. The organization can make educated decisions that generate growth, profitability, and customer pleasure in the highly competitive restaurant industry by employing data-driven insights.

2.3. BUSINESS SUCCESS CRITERIA

Many competitors vie for market share in the restaurant industry, which is naturally competitive. Restaurant C, like many other businesses in its field, has a number of obstacles that could result in a business crisis. These problems include rising competition from new entrants, shifting customer preferences and habits, economic downturns that impact consumer purchasing, and unforeseen occurrences such as pandemics and natural disasters. One of Restaurant C's brands has experienced diminishing profits and slowed growth due to competition and shifting client preferences, posing a substantial threat to the company's financial stability and long-term viability.

Restaurant C must take proactive actions to solve its issues in order to limit these risks and avert a possible business collapse. These measures may include enhancing its marketing and promotional efforts to attract and retain customers, exploring new menu items and promotions that appeal to shifting consumer preferences, streamlining its operations to reduce costs, and leveraging technology to enhance efficiency and customer satisfaction. Restaurant C may reduce the danger of a business disaster and ensure its long-term survival in a highly competitive industry by adopting a data-driven approach to these initiatives.

Due to rising competition and shifting client preferences, one of Restaurant C's brands that specializes in Asian cuisine struggles to sustain its profit margins and growth. This threatens the entire company's financial viability, as this brand is a significant income provider. To avert a commercial crisis, Restaurant C must be proactive in addressing these obstacles. By utilizing its data to obtain insights into client behavior and preferences, the business may make decisions that are more likely to result in favorable outcomes.

Restaurant C should refine its marketing and promotional efforts to attract and retain customers, investigate new menu items and specials that correspond with evolving consumer preferences, streamline its operations to minimize costs, and utilize technology to improve efficiency and customer satisfaction. These initiatives could not only assist Restaurant C in overcoming its current obstacles, but also position it for future success in a highly competitive business.

Restaurant C can limit its risks and preserve its long-term viability in the restaurant sector by being proactive and data-driven. Taking a strategic approach to its difficulties will ultimately benefit not only the company, but also its shareholders, employees, and customers.

2.4. SITUATION ASSESSMENT

Using the CRISP-DM method, our team of highly skilled and experienced data scientists will conduct a comprehensive situation assessment. Our first step will be to outline Restaurant C's business problem, which consists of diminishing profits and growth in one of its established brands due to rising competition and shifting consumer preferences.

Once the problem has been identified, we will collect and evaluate the pertinent data, which will include sales data from the failing brand and data on client preferences and behavior. We will employ several data analysis approaches, such as market basket analysis, customer segmentation, to obtain insight into customer preferences and behavior.

The market basket research will enable us to determine which goods clients typically purchase together, allowing us to offer possible set menus and cross-selling opportunities. Customer segmentation will help us to identify distinct customer groups based on their interests and behavior, allowing us to customize our marketing and promotional activities for each group. Sentiment analysis will assist us in comprehending the opinions and levels of contentment of our consumers, allowing us to solve any potential issues and enhance overall customer satisfaction.

On the basis of the insights gained from our analysis, we will make recommendations regarding the Restaurant C's menu offerings, pricing strategies, promotional campaigns, and customer service, among other business-related aspects. Our ideas will be tailored to Restaurant C's unique circumstances and objectives in order to assist the company in remaining competitive and profitable in the extremely competitive restaurant sector.

In conclusion, our team of skilled and experienced data scientists will use the CRISP-DM method to conduct a comprehensive situation analysis and provide Restaurant C with actionable insights and recommendations. Our objective is to assist the company in resolving its business issue and retaining its leading position in the highly competitive restaurant market.

We recognize the necessity of producing high-quality solutions within the allotted budget in terms of expenses and benefits. We have accounted for the costs related with data collecting, software licensing, and labor. By using our experience and resources, we aim to provide clients with solutions that are both cost-effective and deliver substantial benefits.

2.5. DETERMINE DATA MINING GOALS

Based on the business goals and the data mining objectives, we will need to perform the following analyses to answer the questions posed by Restaurant C:

Market Basket Analysis: We will conduct an analysis of the transaction data to identify which products are frequently purchased together. This will help to understand customer preferences and identify any possible product combinations that may be suitable for new set menus.

Customer Segmentation: We will segment customers based on their purchasing behavior and characteristics. This will help to identify different groups of customers with specific preferences and purchasing behaviors, allowing for tailored marketing strategies and promotions.

Product Recommendation: Based on the market basket analysis and customer segmentation, we will develop a product recommendation system to suggest additional products that customers may be interested in. This will help to increase cross-selling and upselling opportunities.

Customer Behavior Analysis: We will compare the purchasing behavior of dine-in and delivery customers to identify any significant differences in their preferences and habits. This will help to identify areas for improvement in the delivery service and provide insights into ways to increase dine-in sales.

Strange Product Combinations: We will conduct an analysis to identify any unusual combinations of products that customers may be ordering together. This will help to identify potential product pairings or combinations that can be used in marketing campaigns or set menus.

Consumption Patterns: We will analyze the data to identify any patterns in consumption that may indicate tendencies or preferences among customers. This will help to identify opportunities for new products or menu items that are aligned with customer preferences.

Overall, the data mining goals of this project are to provide actionable insights and recommendations to Restaurant C to improve their operations, increase customer satisfaction, and ultimately increase profitability.

3. METHODOLOGY

This data set consists of transactional information from Restaurant C a popular restaurant in Nicosia, Cyprus. Each row in the dataset represents a document line, with the document number appearing in the same number of rows as the number of document lines in the original document. The dataset contains information about the items, including amounts and pricing, as well as certain client details, like the number of diners at the table, customer ID, city, and creation date. In addition, the dataset includes details about the employee who issued the document, whether the sale was for delivery or dine-in, as well as the day and time the document was issued.

3.1. DATA UNDERSTANDING

There are both numerical and category variables in the data. Quantity and TotalAmount are numerical variables, while DocNumber, ProductDesignation, ProductFamily, EmployeeID, IsDelivery, and CustomerCity are categorical variables. Depending on the context of analysis, the variables Pax, CustomerID, and CustomerSince can be considered as numerical and categorical variables.

The team of data scientists will conduct the analysis using a data mining strategy based on the CRISP-DM methodology. The first stage is to do data comprehension, which entails gathering and evaluating the dataset in order to discover any anomalies, such as missing numbers, outliers, and discrepancies. In addition to studying the distribution of the variables, identifying potential relationships, and comprehending the context of the data, this stage will also involve examining the context of the data and discovering potential relationships.

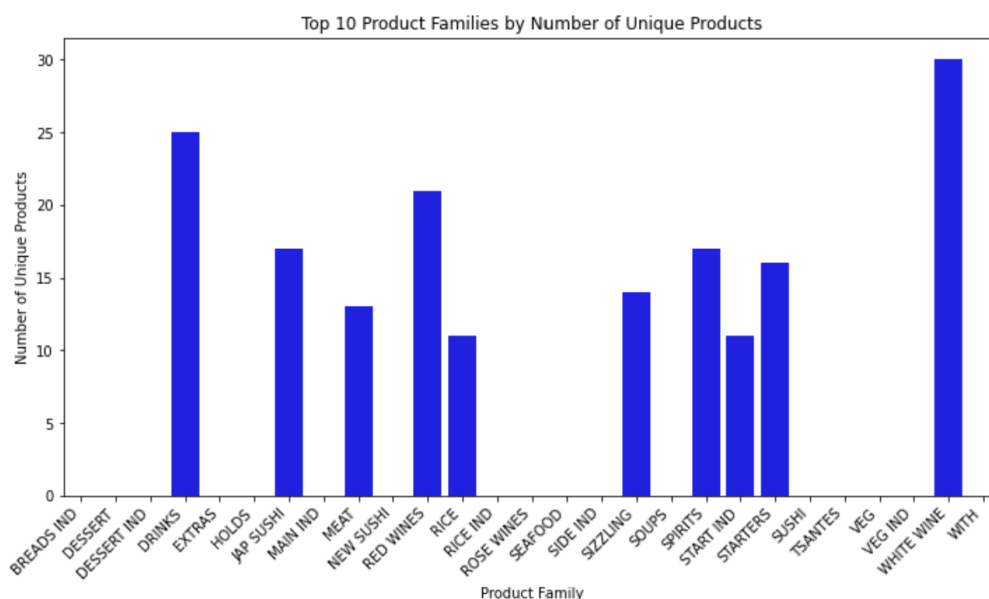
3.2. DATA PREPARATION

During the data preparation phase, we performed various operations to ensure that the data is suitable for analysis. Firstly, the dataset was loaded into a pandas dataframe using the `pd.read_csv` function with specified parameters such as the delimiter, column data types, and the `parse_dates` attribute for the InvoiceDateHour and CustomerSince columns. The TotalAmount column had commas that were replaced with dots and converted to float data type.

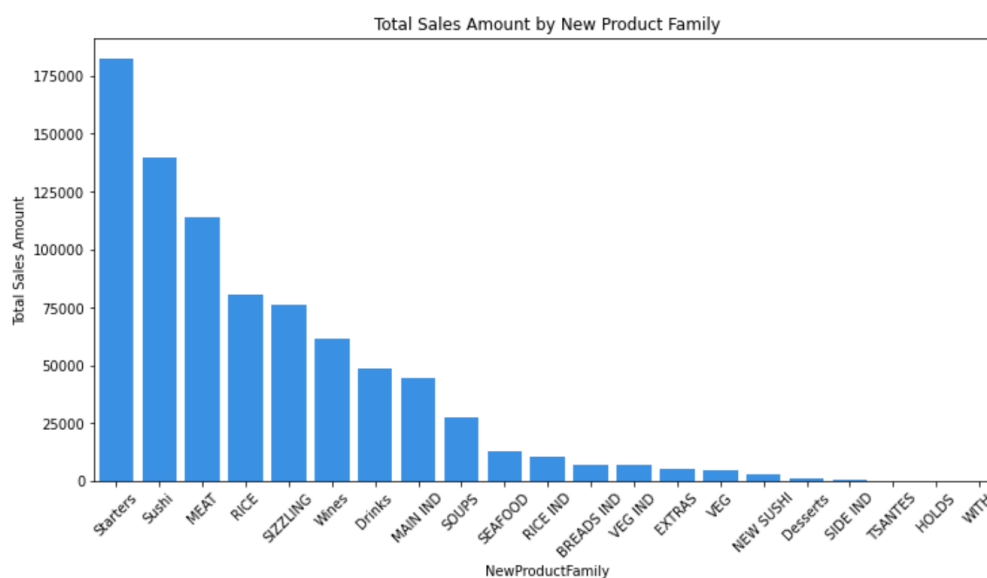
Before performing data exploration, the dataset was checked for missing values using the `isnull`

function. The grouped statistics of the dataset were computed using the groupby function on the ProductFamily column. The family_summary dataframe was created to obtain unique products in each family, the average quantity of products sold in each family, and the total quantity of products sold in each family. The family_summary dataframe was then sorted by the number of unique products in each family and the Num_Unique_Products, Avg_Quantity, and Total_Quantity columns were renamed for better readability.

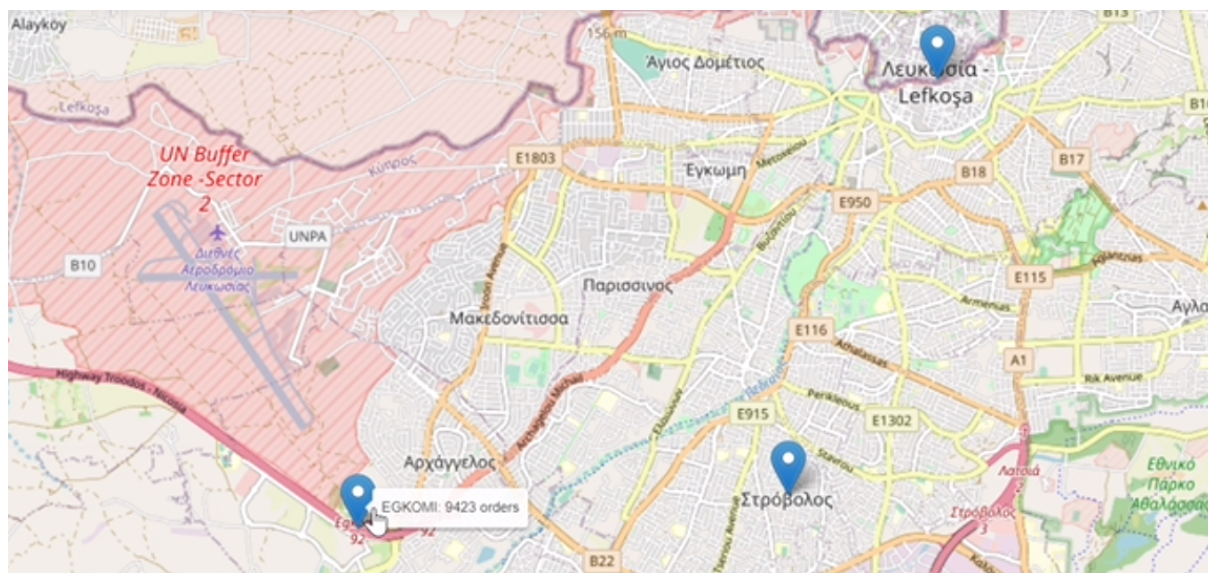
Top 10 product families by the number of unique products and top 10 product families by total quantity sold were then plotted as bar charts.



A mapping function was applied to the ProductFamily column to create the NewProductFamily column, grouping products by their type. A bar chart was created to show the total sales amount by NewProductFamily.



The ds_1 dataframe was created as a copy of the original dataset with the removal of outliers based on the Qty, TotalAmount, and Pax columns. The percentage of data kept after removing outliers in ds_1 was printed, and the TotalAmount distribution, sales by ProductFamily, relationship between TotalAmount and Pax, and the time series of sales were plotted. A countplot was created to show customer sales by the City..



A pivot table was created by pivoting the data with the ProductDesignation column as the columns, and the DocNumber column as the index, which was then filled with zeros if no sales of the product occurred in that document. The weekday and weekday_name columns were added to ds_1 to represent the day of the week of each sale. A new column was created to classify each sale into lunch or dinner time periods, and then the sales data was split into lunch and dinner dataframes. Descriptive statistics were computed for the entire ds_1 dataset, lunch_data, and dinner_data. Finally, the number of documents in which SPRING ROLL was sold was checked using the pivot table. We removed any duplicates and inconsistencies in the data. Next, we transformed the data by creating new features such as the total amount per document and the time of day of the transaction. We also applied filtering to exclude any irrelevant columns that did not contribute to the analysis. Moreover, we converted the categorical variables such as the product designation and customer city to numerical values using one-hot encoding to make them suitable for machine learning algorithms. Additionally, we addressed missing data by imputing them with relevant values such as the mean or mode of the respective features. Finally, we normalized the data to ensure that all features have a comparable scale. These steps were crucial to ensure that the data is in a suitable format for further analysis and modeling

3.3. MODELING - MARKET BASKET ANALYSIS

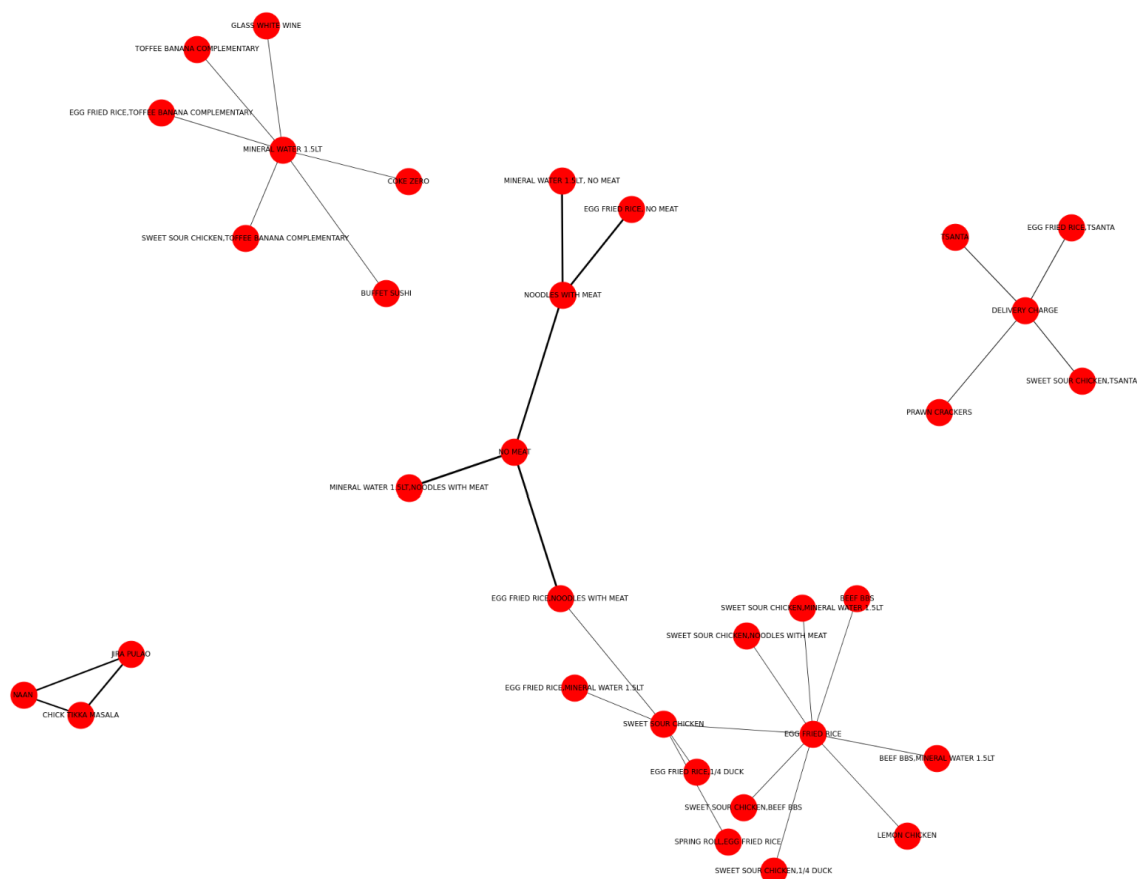
Market Basket Analysis utilizes the idea of association rules, which suggest the probability of a particular item being bought when another item is already present in the shopping basket. The

left-hand side of the rule, also known as the antecedent, indicates the prerequisite condition that must be met for the right-hand side or consequent to be valid. To conduct our analysis, we utilized the Apriori algorithm, a well-known technique for mining frequent itemsets that involves identifying collections of items that are frequently purchased together in a given dataset. We applied the Apriori algorithm to the restaurant's sales data and identified frequent itemsets that were present in at least 5% of the transactions.

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
39	(MINERAL WATER 1.5LT, NOODLES WITH MEAT)	(NO MEAT)	0.098140	0.097057	0.065637	0.668813	6.890954	0.056112	2.726387
40	(NO MEAT)	(MINERAL WATER 1.5LT, NOODLES WITH MEAT)	0.097057	0.098140	0.065637	0.676279	6.890954	0.056112	2.785918
35	(EGG FRIED RICE, NOODLES WITH MEAT)	(NO MEAT)	0.088389	0.097057	0.056338	0.637385	6.567142	0.047759	2.490089
36	(NO MEAT)	(EGG FRIED RICE, NOODLES WITH MEAT)	0.097057	0.088389	0.056338	0.580465	6.567142	0.047759	2.172908
34	(EGG FRIED RICE, NO MEAT)	(NOODLES WITH MEAT)	0.056338	0.158451	0.056338	1.000000	6.311111	0.047411	inf
38	(MINERAL WATER 1.5LT, NO MEAT)	(NOODLES WITH MEAT)	0.065637	0.158451	0.065637	1.000000	6.311111	0.055237	inf
41	(NOODLES WITH MEAT)	(MINERAL WATER 1.5LT, NO MEAT)	0.158451	0.065637	0.065637	0.414245	6.311111	0.055237	1.595142
0	(NO MEAT)	(NOODLES WITH MEAT)	0.097057	0.158451	0.097057	1.000000	6.311111	0.081678	inf
37	(NOODLES WITH MEAT)	(EGG FRIED RICE, NO MEAT)	0.158451	0.056338	0.056338	0.355556	6.311111	0.047411	1.464303
1	(NOODLES WITH MEAT)	(NO MEAT)	0.158451	0.097057	0.097057	0.612536	6.311111	0.081678	2.330390

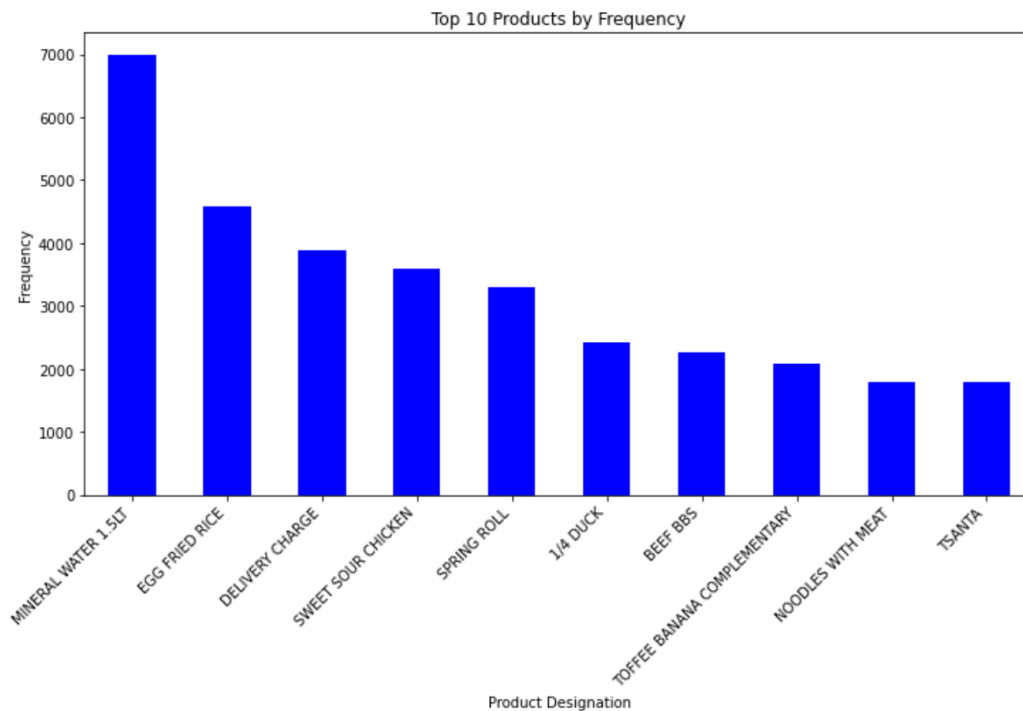
We then generated association rules using the 30 most frequent itemsets. Association rules are rules that indicate the likelihood of a set of items being purchased together. We generated rules based on both confidence and lift metrics. Confidence measures how often an association rule is found to be true, while lift measures the strength of the association between two items.

We also explored the frequent itemsets to identify patterns in the data. We focused on itemsets with a length of 2 and a support of at least 0.2. We also looked at individual items and identified those that were frequently purchased. We created two network graphs, one interactive one, to closely see the patterns and customer orders, which were grouped together.



Findings:

Our analysis revealed several insights into the purchasing behavior of the restaurant's customers. We found that the most frequently purchased items were Noodles with beef, Spring Roll, and Mineral water. We also found that the top 10 products by total amount sold were dominated by delivery charges and plastic bags, indicating that these items contribute significantly to the restaurant's revenue.



Our analysis also identified several association rules that could be used to increase sales. We called them the best pairings where we gave suggestions for new menus and combos. Additionally, our analysis revealed that many customers tend to purchase more than one item at a time, and there are several products that are commonly purchased together. By promoting these products as a package deal or a meal combo, the restaurant could encourage customers to purchase more items and increase sales.

NEW MENU RECOMMENDATIONS:

According to figure 1, we found that the daily sales on Sunday to Thursday is about half of Friday and Saturday. And the sales in lunch time is significantly lower than dinner time, as shown on figure 2. We can raise our competitive power and attract customers by launching lunch set menu and weekday set menu, with the most associated products in a lower price. To further increase revenue, we can put more profitable products in the menu i.e., alcohol.

These are the possible menu list:

Chick Tikka Masala, Jira Pulao, Naan and wine or beverage.
Sweet Sour Chicken, Egg Fried Rice, Spring Roll and wine or beverage.

The analysis allowed us to identify which products tend to be purchased together, providing insight into the preferences of customers. By looking at lift and support values, we can determine which product pairs are strongly associated with each other and how much they are likely to be purchased together.

We can use this information to create new menu items that include these product pairs see figure 12 in appendix, such as a combo meal or a dish that includes two popular items. This would increase the

likelihood of these items being ordered and potentially lead to an increase in overall sales. Additionally, we can create promotions or discounts for these new menu items to attract customers and encourage them to try them out.

It is also important to consider the popularity of individual items and how they can be incorporated into the menu. For example, if a certain product has a high support value but a low lift value, it may be a good idea to feature it more prominently on the menu or include it as a side dish option with other popular items. Such as mineral water being one of the most frequent products in the restaurant, it can sometimes be given for free if we promote a meal combo of higher value and for 4 people, since that pax to revenue ratio created the more profit.

DELIVERY VS EAT IN:

Based on the graph showing the difference in number of dine-in customers versus delivery customers, it is clear that there is a significant opportunity for growth in delivery sales. However, it is important to consider how to increase delivery sales without negatively impacting the dine-in sales, which are currently the primary source of revenue for the restaurant.

One way to increase delivery sales while maintaining dine-in sales is to offer exclusive deals or promotions for delivery orders. This could include free delivery for orders over a certain amount, or a discount on the total order cost for first-time delivery customers. Another approach would be to promote the benefits of delivery, such as the convenience of having food delivered directly to one's doorstep, as well as the added value of the delivery fee and bags.

In addition, the restaurant could consider partnering with third-party delivery services, such as Uber Eats or Grubhub, to increase visibility and reach a wider audience. This would also allow the restaurant to take advantage of the marketing and advertising efforts of these platforms, without having to invest significant resources in their own marketing campaigns.

Overall, by focusing on promoting the benefits of delivery and offering incentives for customers to order delivery, while maintaining the quality of the dine-in experience, the restaurant can increase its revenue and market share in the delivery market without negatively impacting its existing customer base.

NEW PRODUCT RECOMMENDATIONS:

Based on the Apriori analysis we performed on the restaurant data, we found that one of the most frequent orders is "noodles with meat". However, it is interesting to note that many customers who ordered "noodles with meat" also removed the meat from their dish. This could suggest that there is a market for vegetable-based noodle dishes, as many customers seem to enjoy the noodles without the meat. Therefore, we recommend that the restaurant consider adding a vegetable noodle dish to their menu. This could appeal to a growing customer base that is interested in healthier and plant-based meal options.

4. RECOMMENDATIONS BASED ON DIFFERENT VISUALIZATION INSIGHTS AND DEPLOYMENT

Total Sales by Product Family: This chart shows the total sales amount for each product family. It can help the restaurant identify which product families are the most profitable, and possibly adjust their menu offerings or marketing efforts accordingly. Please see figure 3 on the appendix to see the best selling products.

Total Sales by Employee: This chart shows the total sales amount for each employee. It can help the restaurant identify their top-performing employees and potentially reward them for their efforts, as well as identify any employees who may need additional training or support. Please see figure 4 to see how 80% of the sales are done from the

Average Order Amount by Hour: The heatmap shown on figure 5, shows the average order amount for each hour of the day. It can help the restaurant identify peak ordering times and adjust staffing levels or promotions accordingly. Example: offering discounts on deliveries on less busy hours and

Sales by Customer City on a Map: This chart displays a map of Cyprus with markers representing the cities where the restaurant's customers are located. The size of each marker represents the total order amount for that city. It can help the restaurant identify which cities have the most sales and potentially target marketing efforts towards those areas. Please see figure 6 in the appendix.

Sales by Month and Product Family: This chart displays a stacked bar chart showing the total sales amount for each product family, broken down by month. It can help the restaurant identify any seasonal trends in sales and adjust their menu offerings or marketing efforts accordingly. Please see figure 7 in the appendix.

Sales by Day of the Week: This chart displays a stacked bar chart showing the total sales amount. It can help the restaurant identify which days of the week are the most popular for each delivery type and adjust staffing levels or promotions accordingly. Please see figure 1 below.

Sales by Customer City and Deliver Type: This chart displays a stacked bar chart showing the total sales amount for each customer city, broken down by delivery type. It can help the restaurant identify which delivery type is more popular in each city and adjust their delivery operations accordingly.

Revenue generated during the year and by day: This graph shows that there is a high correlation on generating higher revenue on the holidays. Holidays would be a great way to create more buzz for the restaurant and provide special offers for delivery.

Another suggestion based on the graphs and charts is to explore ways to increase sales during off-peak hours. The graph showing sales performance by time of day reveals that sales drop significantly during certain hours, particularly in the early afternoon and late evening. One possible strategy for increasing sales during these hours could be to offer discounts or special promotions to incentivize customers to visit during these times. Additionally, the restaurant could consider introducing new menu items or adjusting pricing during these hours to better appeal to customers.

Overall, the insights provided by the graphs and charts can help inform strategies to improve the restaurant's performance and better meet the needs of its customers.

These charts can provide valuable insights into the restaurant's sales and operations, helping them make data-driven decisions to improve their business

5. CONCLUSIONS

In conclusion, Market Basket Analysis is a powerful technique used by retailers and marketers to gain insights into consumer behavior and make informed decisions about product placement and promotions. By analyzing transactional data, Market Basket Analysis can identify patterns in purchasing behavior, such as which products are frequently purchased together or which products are often purchased in a particular sequence. This information can be used to optimize product placement, create targeted promotions, and develop personalized recommendations for customers.

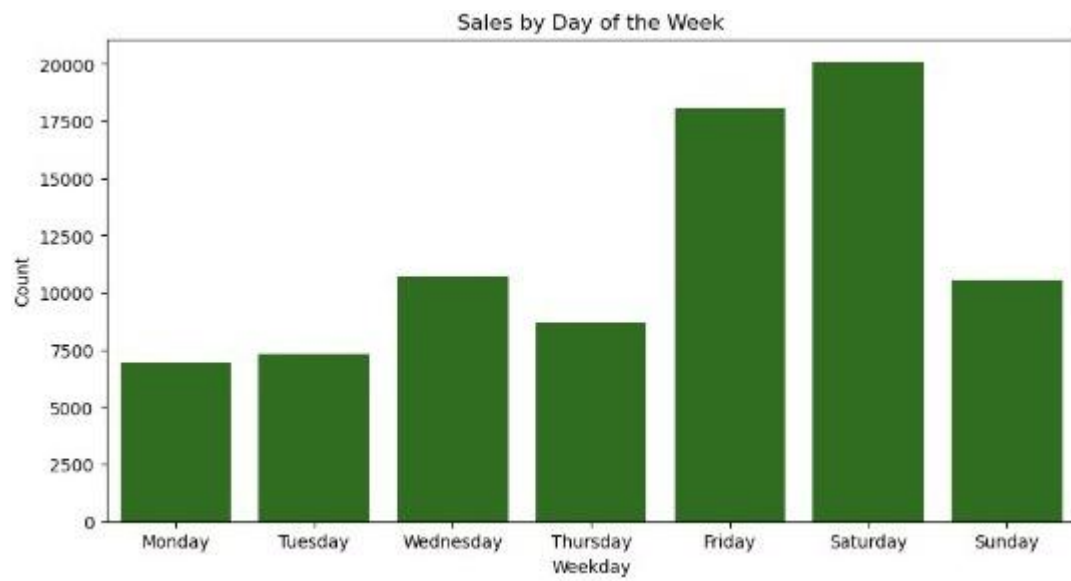
In this project, we applied Market Basket Analysis to a restaurant's sales data to identify patterns in customer ordering behavior. Using the Apriori algorithm, we were able to identify frequent itemsets and generate association rules with high levels of confidence and lift. By analyzing these rules, we were able to identify popular menu items and make recommendations for product bundling and substitution.

CONSIDERATIONS FOR MODEL IMPROVEMENT

There are several other algorithms that can be used for market basket analysis and frequent itemset mining, such as the FPGrowth algorithm, the Eclat algorithm, and the PrefixSpan algorithm. Each of these algorithms has its own strengths and weaknesses and may perform differently on different datasets. It may be worth experimenting with these algorithms to see if they can improve the accuracy and efficiency of the model. Additionally, feature engineering techniques, such as creating new variables based on domain knowledge, could be used to enrich the dataset and potentially improve the performance of the model. Finally, incorporating external data sources, such as weather data, could provide additional insights into customer behavior and further improve the accuracy of the model.

5. APPENDIX

Graph 1



Graph 2

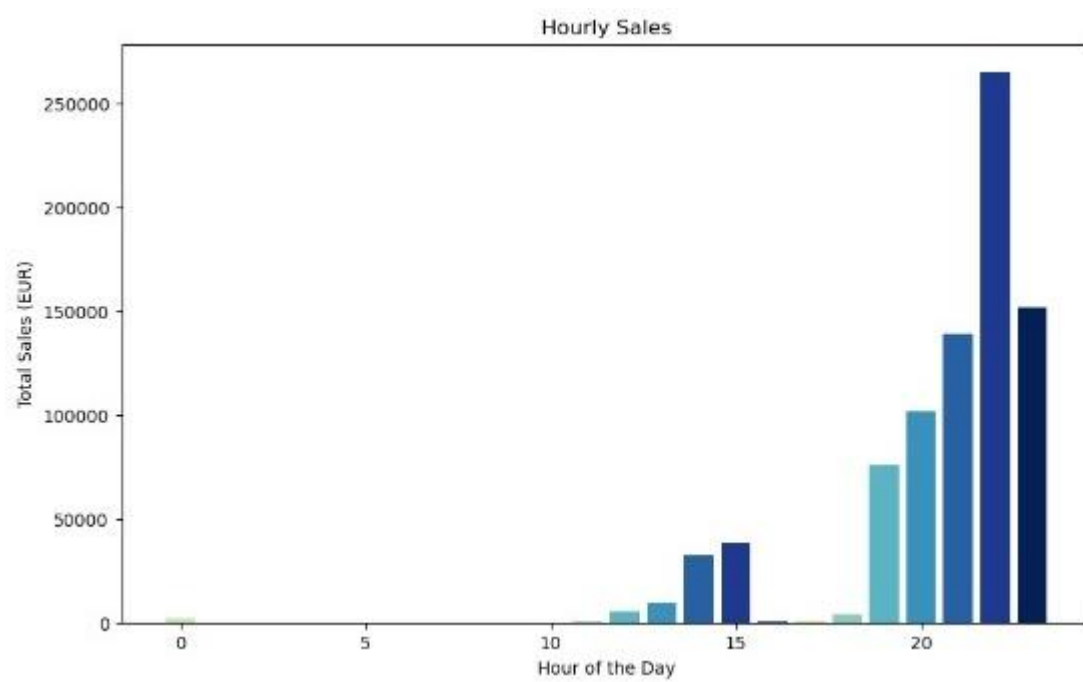


Figure 3

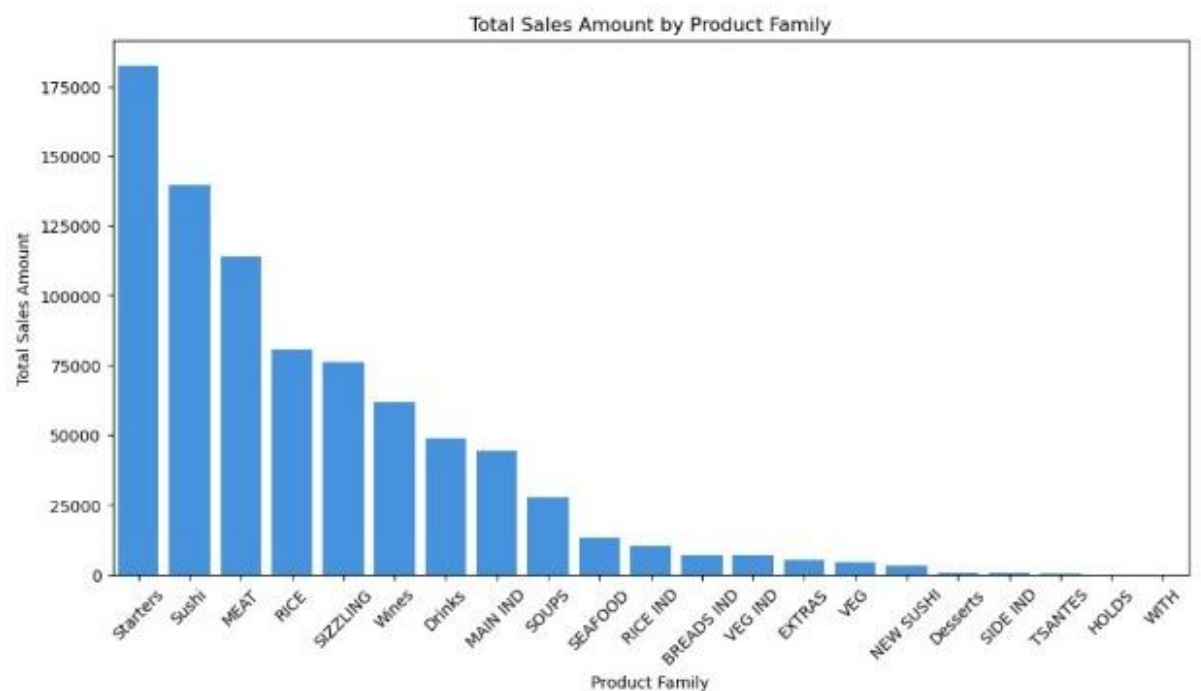


Figure 4

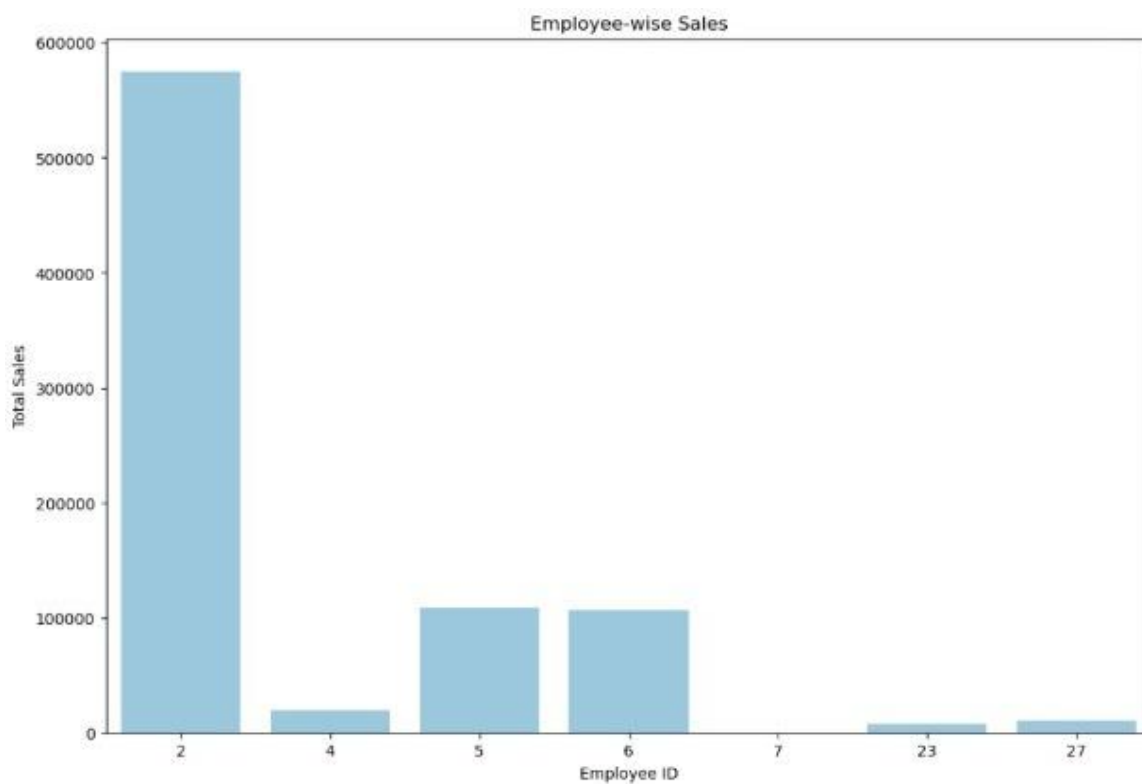


Figure 5

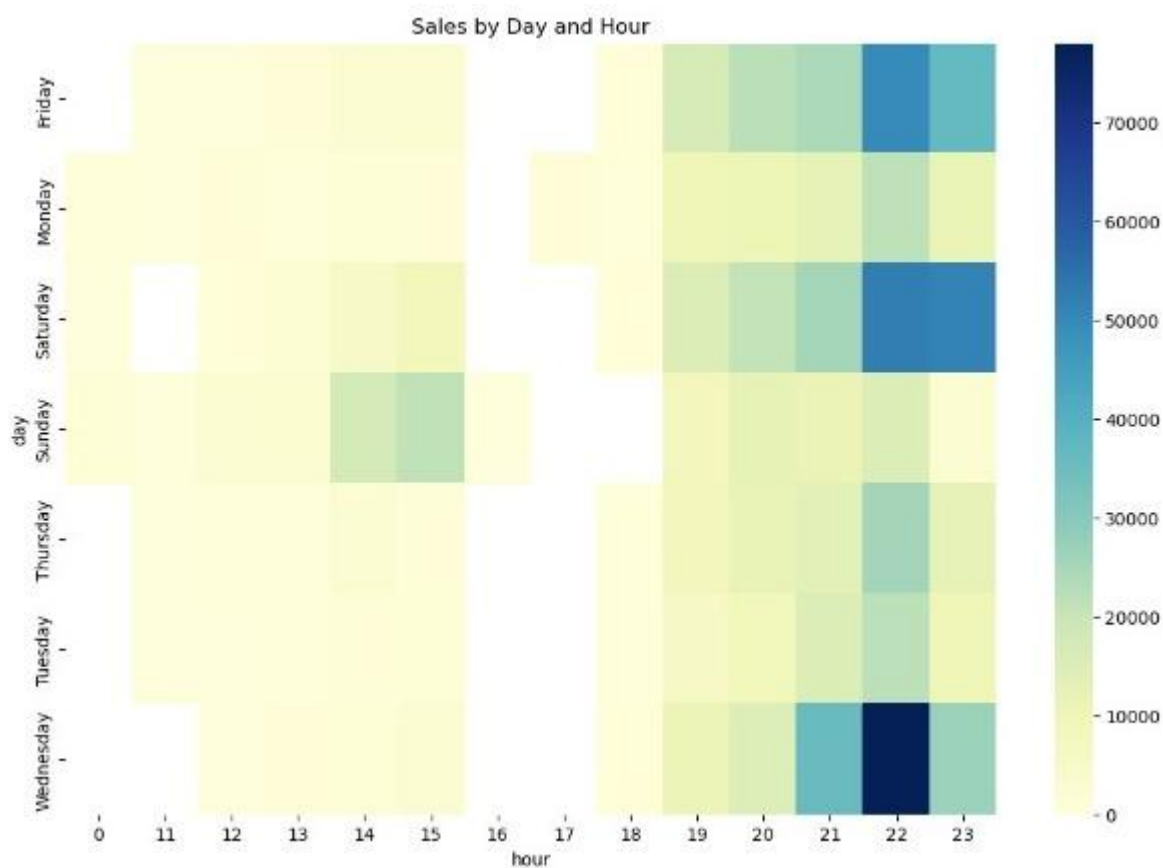


Figure 6

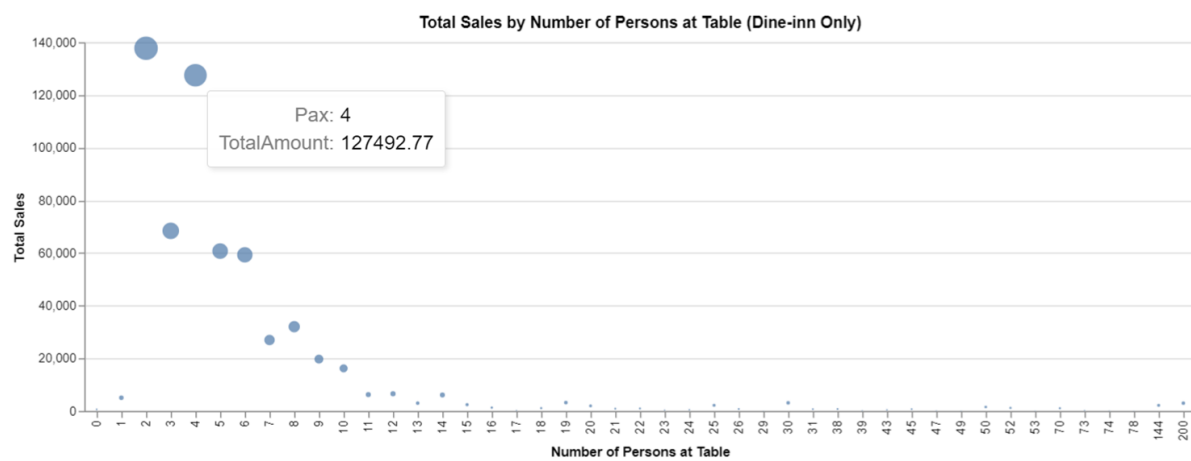
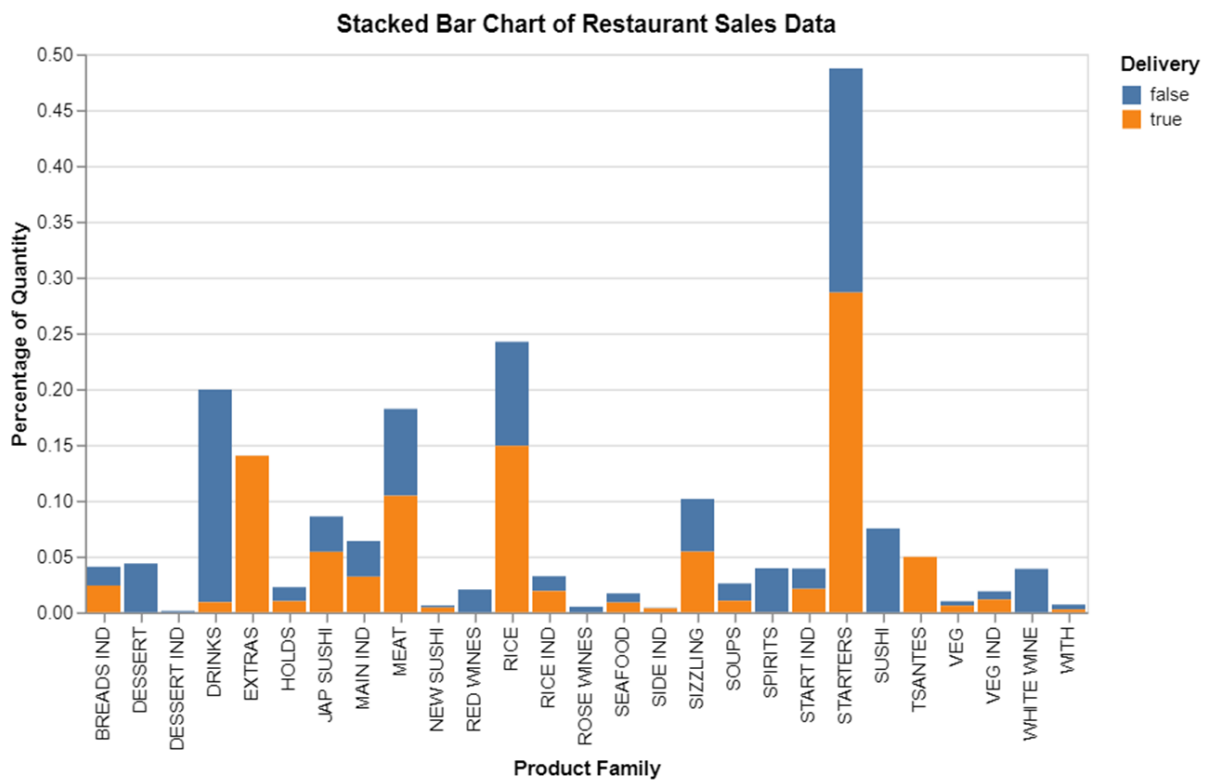


Figure 7



REFERENCES:

- 1) Cyprus Statistical Service (2019). Statistics on Accommodation and Food Service Activities, 2019. Retrieved from [https://www.mof.gov.cy/mof/cystat/statistics.nsf/All/7B03F810B98EB62EC2258539002328D7/\\$file/ACC03_2019.pdf](https://www.mof.gov.cy/mof/cystat/statistics.nsf/All/7B03F810B98EB62EC2258539002328D7/$file/ACC03_2019.pdf)
- 2) Cyprus Statistical Service (2021). Statistics on Accommodation and Food Service Activities, 2020. Retrieved from [https://www.mof.gov.cy/mof/cystat/statistics.nsf/All/7B03F810B98EB62EC2258539002328D7/\\$file/ACC03_2020.pdf](https://www.mof.gov.cy/mof/cystat/statistics.nsf/All/7B03F810B98EB62EC2258539002328D7/$file/ACC03_2020.pdf)
- 3) Cyprus News Agency (2021). Restaurant Turnover Drops by 33.3% in 2020, Hit by COVID-19 Pandemic. Retrieved from <https://www.stockwatch.com.cy/en/article/oikonomia-kypros/restaurant-turnover-drops-333-2020-hit-covid-19-pandemic>
- 4) European Bank for Reconstruction and Development (2019). Cyprus Country Assessment. Retrieved from <https://www.ebrd.com/documents/country-assessments/cyprus-country-assessment.pdf>