Reducing Food Wastage by Using Restaurants Takeaway Orders Dataset

Farjana Alam: 8558
Department of Computer Science
University of South-Eastern Norway
Kongsberg, Norway

Moin Shahriyar: 8563
Department of Computer Science
University of South-Eastern Norway
Kongsberg, Norway

***Abstract* — The main purpose of this report is to prevent food waste by limiting restaurant owners' grocery purchases using the past food order dataset. Analysis food order dataset is a comprehensive approach which is necessary to address the issue of restaurant food waste. In this report, we will analyze huge amounts of restaurant orders to predict the most popular food items for a given month using Big Data technologies. Our analysis will illustrate all the products used for each month and graphically represent the total number of ingredients needed for the restaurants. Due to this data-driven management process, restaurant owners can make significant adjustments to how they manage their restaurant's food waste.**

Keywords—Food waste, waste reduction, Overbuying, Big data, Big Data Technologies, Data Analysis.

I. Introduction

Global food waste is a multifaceted problem that begins with agricultural production and goes all the way to garbage. Each year, more than 30% of food is lost or squandered. Given the vast number of hungry people around the globe, this number is even more startling. Food waste is a problem for social justice as well as being inefficient. Global food waste has a significant negative influence on the environment. Food waste is a significant contributor to squandered natural resources and greenhouse gas (GHG) emissions [1].

When we think about food waste, we may envision food that was not consumed and discarded from our tables at a nice dining place but is wasted for a variety of causes like overproduction in the restaurant business. However, now is the time to focus on finding solutions [1].

Analysis of restaurant food orders will provide a fix for the issue. Restaurant owners need to keep all the ingredients for every dish on hand because of the large number of food orders that are placed daily from various food options. However, not every item will be ordered every day. If we can determine which dish is ordered the most frequently in a given month, the restaurant owners do not need to buy more ingredients than necessary, especially those ingredients which is easily degraded.

With the help of big data technologies, we can analyze, visualize, and estimate the number of ingredients needed for a particular month from a huge number of orders from an online collected restaurant orders dataset. Big data is seen as one of the most significant areas of future technology and is quickly catching the attention of various industries due to the tremendous value it can give to businesses. Companies presently face significant changes in how they manage their operations, customers, and business models because of the data-driven management revolution that has taken over [2]. In this project, we want to minimize food wastage in restaurants. The primary objective of our study is to reduce food waste at the restaurant by examining the restaurant menu card to determine which raw ingredient is most frequently utilized each month. As a result, a restaurant owner can regularly restock his or her food supply depending on the projection. In our restaurant dataset, for instance, there are typically five to six different forms of protein. However, if we can forecast the unit of protein used based on demand, the restaurant owner can purchase the raw material accordingly. In this way, Restaurant owners can limit their ingredient waste.

In Section II, we have researched some of the previous work related to our project from different papers and journals as well as we gave a brief introduction of the dig data technologies that we use in our project.

In Section III, the methodology of the project was discussed where we introduced the source of our data, a short description of the datasets, data analysis process, data storage, visualization, and system architecture of our project.

In Section IV, we demonstrate the results that we got after analyzing the data under the Result subsection. We also put our discussion part about the project in this section under the Discussion subsection where we represented how should we implement this project in real life and achieve our goal.

In Section V, we conclude our project by establishing our success in the analysis part and the usage of it to reduce food wastage for any restaurant owner or non-technical person.

In Section VI, we share the contribution of each member of this group.

In Section VII, we put all the references from where we used information and data to complete our project.

II. Background and Literature Review

A. Background

The restaurant industry is expanding quickly over the world. People appreciate trying out new foods and spending quality time with their loved ones more frequently. However, a significant global problem that many restaurants encounter is food waste. And the food waste generates by the restaurant business is massive. Food waste in the dine-in industry is quickly evolving into a major issue considering that it has recently contributed to roughly 12% of all food waste [2][3]. Additionally, with the rise in popularity of dining out, which is being fueled by rising wages and tourism, hospitality waste has emerged as a significant problem for both rich and developing nations [4]. The volume of food-related waste produced in this industry is routinely mentioned in the media, but it has not received enough academic attention yet [5]. Other experts have also highlighted that, despite being identified as a major topic, the issue of food-related waste in this area has received less attention [6][4]. Food can be wasted in the restaurant industry in various ways, whether cooked or raw. But since it is difficult to forecast the number of raw ingredients needed for each day in the restaurant industry, a lot of waste is produced every day.

B. Big Data Technologies

Big Data is becoming more and more prevalent, and recent developments in this area have been significant. To analyze big data, frameworks like Apache Hadoop and Apache Spark have become incredibly popular over the past few decades [7]. For our project, we decide to use Python programming language, PySpark, Pandas Library, MongoDB for Database, NumPy for scientific computing, and Matplotlib for data visualization.

- Apache Spark: Apache Spark is a framework for data processing that can swiftly conduct operations on very large data sets and distribute operations across several computers, either alone or in conjunction with other tools for distributed computing. These two characteristics are essential to the fields of big data and machine learning, which both call for immense computing resources to be mobilized to process enormous data warehouses. Spark also relieves developers of some of the programming responsibilities associated with these activities by providing an easy-to-use API that abstracts away most of the grunt work of distributed computing and big data processing [8].

- Python: There are lots of programming languages to choose from for data analysis. But among all of them, we choose python. Python has built-in analytical features so this is the ideal choice for our project. There are lots of benefits to using Python. Python has a huge open-source library collection like Pandas, NumPy, Matplotlib

- Pandas: For working with structured data sets typical of statistics, finance, the social sciences, and many other subjects, use pandas, a Python library of rich data structures and functions. For performing standard data manipulations and analysis on such data sets, the library offers integrated, simple procedures [9].

- Matplotlib: Most of the scientific journal illustrate figures to explain their findings. Matplotlib is one of the better options to use when it comes to producing visualizations. In order to create publication-quality images across user interfaces and operating systems, interactive scripting, and application development, Python includes a 2D graphics library called Matplotlib [10].

- MongoDB: MongoDB database is a document-oriented database. It is schema-free and includes a database, collection, and document. Many Collections can exist in the same database. Every Collection consists of several Documents. Anytime, without a prior definition, a collection can be built. It may also contain records from documents with various schemas. The type of the property can be any basic data type, such as numbers, texts, dates, and so on, or an array or hash, or even a sub-document [11].

C. Literature Review

Numerous papers have been written about managing restaurant food waste with various technologies such as the Internet of Things (IoT) network system for restaurant food waste management [12]. Wen et al. (2018) developed an IoT sensor system that provides how the help of the Integrated RFW management platform restaurant waste can be managed. There are, undoubtedly, several important literature studies about restaurant food waste reduction. The existing literature on food waste in the hotel industry has concentrated on a diversity of topics. including food waste quantification, waste composition, waste handling, doggy bags, customer attitudes, demographic considerations, composting, governmental regulations, and landfills [3]. However, most of the paper focuses on handling leftovers and the causes of the composition of food waste. For instance, Food waste in hospitality and food services [2].

Dhir et al. (2020) use a systematic literature review (SLR) approach that is executed through the search, assessment, and synthesis of peer-reviewed papers to investigate how food waste occurs in hotels and food services. Many scholars analyze different datasets to handle food waste more efficiently.

Panda et al. use predictive analytics to minimize food wastage using machine learning techniques. In this paper, scholars make the prediction using datasets from different organizations like schools, colleges, universities, and catering businesses and use public information like age, gender, payment, and education attributes [13].

Admittedly, there are some noteworthy literature reviews on this issue but most of them are focused on how to manage restaurant food waste [14] or how to reduce food wastage from customer plates [15] such as Sustainable consumption by reducing food waste written by De Los Mozos et al. and Food waste reduction from customer plate written by Kim et al.

But from the aforementioned that there is not much research that focuses exclusively on reducing raw food waste in the restaurant industry using our dataset.

III. METHODOLOGY

A. Data Collection

The "Takeaway Food Orders" dataset from Kaggle [16], which consists of two Indian takeaway businesses in London, UK, is used to analyze food orders. In the dataset, each row is a single product within the order and contains ~200k rows of data

B. Dataset Description

a. Restaurant-1-orders: More than 13 thousand orders totaling over 75 thousand rows may be found in the dataset. 248 products are available for ordering. Order Number, Date of Order, Name of Product, Quantity of Product Ordered, Product Price, and Total Product are the six columns that make up this table. The dataset is in CSV format (Comma Separated Values).

b. Restaurant-2-orders: The dataset contains over 20 thousand orders which consist of around 120 thousand rows in total. There are 302 items for ordering. It contains six columns in total named Order Number, Date of Order, Name of Product, Quantity of Product in order, Product Price, and Total Product. The format of the dataset is also in CSV (Comma Separated Values).

c. Restaurant-1-ingredients: There are 3 columns and 248 rows altogether. The Object Name field in each row gives the name of the item. Each item from Restaurant-1 orders has its ingredients manually entered into the ingredient's column. The format of this dataset is CSV.

d. Restaurant-2-ingredients: There are 248 rows and 3 columns in total. Each rows contains the name of the item in the Item Name column. The ingredients are manually inserted into Ingredients column for each item from Restaurant-2 orders. This dataset is in CSV format as well

C. Data Analysis Process

After collecting the data, we analyzed the data by filtering it many ways such as we grouped our data by Date, Order Items, Total Orders and so on for. Because we have to determine how many items were ordered on which day. Then we figured out that we need Ingredients for each item as it was missing in the dataset and that was essential for our project. So, we inserted the main ingredients for each item in another dataset and planned to merge it with the main dataset which will be used for analyzing food wastage. Both Restaurant-1 and Restaurant-2 dataset.

D. Data Storage

After data collection and analysis were complete, we stored all datasets, including the ingredients dataset, on a local MongoDB server (localhost:27017). Using the Studio 3T platform and PyMongo, a Python distribution with utilities for working with MongoDB, we were able to accomplish that. Our datasets were given the names restaurant-1-ingredients, restaurant-1-orders, and restaurant_1_details. For restaurant 2, we used the same names where restaurant_1_details containing all the data for restaurant 1 and same goes for restaurant 2 which is merely a dataset that combines the ingredients and orders datasets. We cleaned and created two new datasets—one for ingredients and another for order—from this dataset that were then utilized for additional research. Studio3T was used to store these two datasets for each restaurant in MongoDB.

E. Data Visualization

For graphical data visualization we need to do some additional steps according to our need in our dataset file. For this part, the dataset first needs to be loaded into our text editor, Jupyter Notebook, before the data visualization process can begin. All of the data were in strings when we checked the dataset's schema. We need to calculate each day of month for counting the items. So, we divided Order date into Date and Time as it comprises both. The hour has now been separated from Time, while Day, Month, and Year have been separated from Date. Following that, we combine Day, Month, and Year into the timestamp standard format to transform Date from a string type into a timestamp type. To obtain weekdays for each date using the Timestamp function date_format(), we must accomplish that. Because which weekdays are the busiest, we can find out from that column, and we will analyze this further in our future work.

We import the ingredients dataset and add an empty Ingredients column to the primary dataframe. It is highly challenging to enter the ingredients for each food item because there are roughly 75 thousand rows for Restaurant-1 and 120 thousand rows for Restaurant-2. So, in order to accomplish that, we used a nested for loop. We have added the ingredients for each food item that has been ordered for both Restaurant-1 and Restaurant-2 after converting the spark dataframe into a pandas dataframe. We matched the string values from the ingredients dataset with the main dataset using the Numpy library attributes, and then we added the ingredients into the precise box where the Item Name matched. We used the Studio 3T platform to save our new Dataframe into MongoDB and gave it the names restaurant_1_details and restaurant_2_details.

The 15 columns in our combined dataframe provide all the information we require to conduct the project's analysis. However, we only need a few columns from all of them for the analysis of food waste. These columns were divided, and new dataframes called restaurant_1_ingredient, restaurant_1_order was generated. We saved these dataframes into MongoDB after eliminating the null values and purging the dataset of any unnecessary entries. After doing so, we tested the schema numerous times for all the data. The same was carried out for restaurant 2.

At the final stage of data visualization, our new dataframe restaurant_1_ingredient contained only 73093 rows, 6 columns and restaurant_2_ingredient 115293 rows and 6 columns where restaurant_1_order contains 74818 rows and 7 columns and restaurant_2_order contains 119183 rows and 7 columns.

F. System Architecture

Figure – 1 represents the system architecture of our project. This architecture represents all the software, databases, and platforms that we used to complete our project and how they interact with each other. The architecture starts from data collection process from Kaggle, an online open source platform and went through ETL (Extract, Transform and Loading) to Local MongoDB server using Studio 3T and finally finished at Data Visualisation Dashboard where we can show our analysis output.

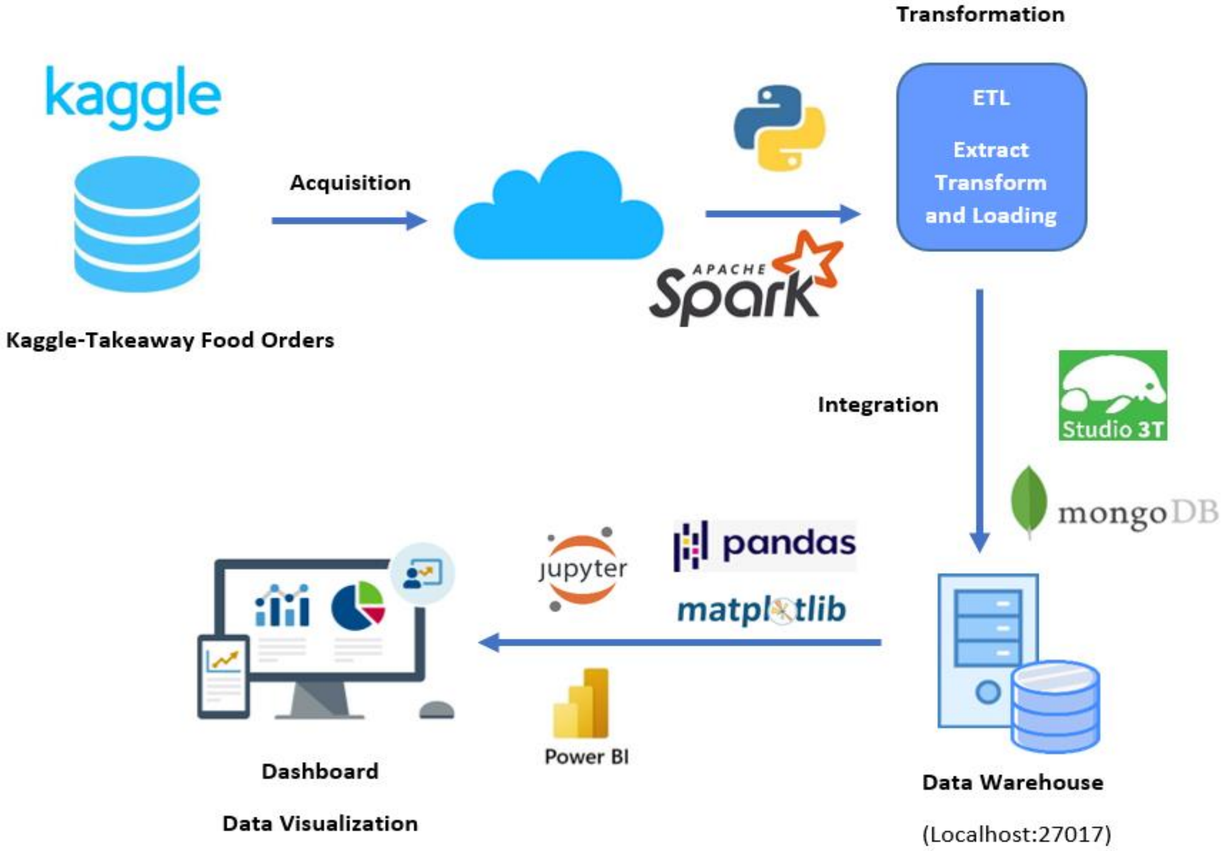

Figure 1 - System architecture

a) Data Flow Diagram: Figure-2 is the data flow diagram we created after segmenting our work into phases I, II, and III. This diagram is the workflow of our project it illustrates how we handled our dataset throughout the project. The terms "Data Collection," "Data Modeling and Cleaning," "Visualizing and Result Generating" can all be used to describe these three sections. For each phase, we went through several stages, but we can summarize them into three key phases that are described below.

   i. PHASE-I: In this phase we gather and then analysis the dataset which data columns required and need to be created for our project. There are a total of two datasets for the two restaurants food orders dataset. After gathering the data, we proceed to the main analytical stage. To better comprehend the data, we go through all the columns and rows here. After that, we quickly recognized a few columns that were crucial to achieving our objective. We also discovered that we require an additional dataset that includes ingredient lists for each dish from both restaurants. We completed phase I of the project by manually creating that dataset for both eateries.

ii. PHASE-II: In this phase we clean and modify our dataset to determine our output. All of the ingredient values from the ingredient dataset must now be inserted into the Main Dataset. The primary dataset has thousands of rows, though. As a result, accomplishing that via manual input is very impossible. In order to finish the work, we have built nested for loops for both eateries. After that, we conducted an analysis on our primary datasets to eliminate pointless rows and columns. The final step of this phase involved cleaning up our datasets and removing all the null values.

iii. PHASE-III: In this stage, we customized the dataset visualization to meet our needs. We accomplished that in two ways: first, we visualized the ingredients; second, we visualized the frequency of orders. We can have a sound hypothesis model to reduce food waste in restaurants by combining them. Therefore, we estimated how much of each ingredient would be required for a single restaurant in each month of the year. Finally, we went into great detail on how a restaurant might use this project to reduce food waste.

.

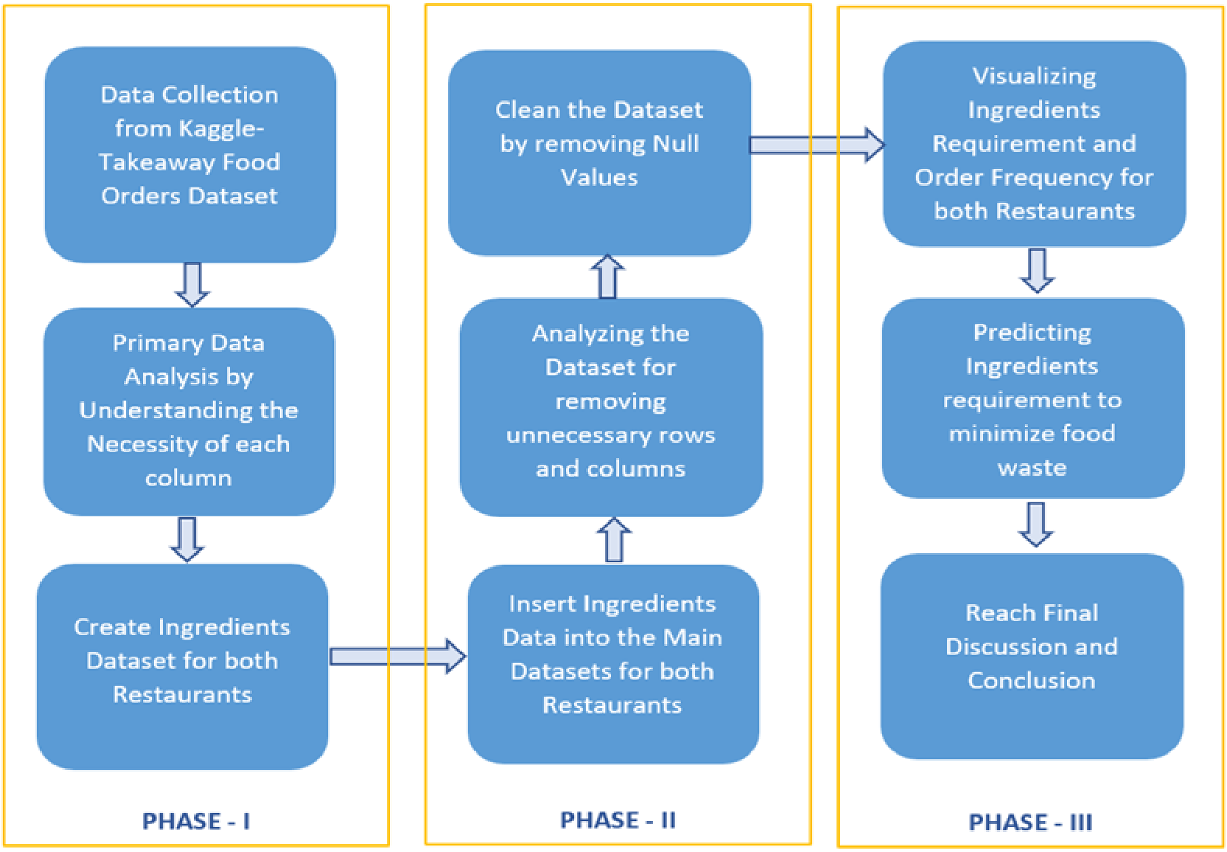

Figure 2 - Flow Diagram

b) Data Modelling: Following Figure- 3 is the data storage phase, we linked MongoDB and PowerBI to determine the relationships between our datasets. We built the relationships between the datasets using PowerBI, which was crucial for data visualization. We were able to determine what columns and data were required for our project using this modeling. As a result, we could filter our data more effectively.

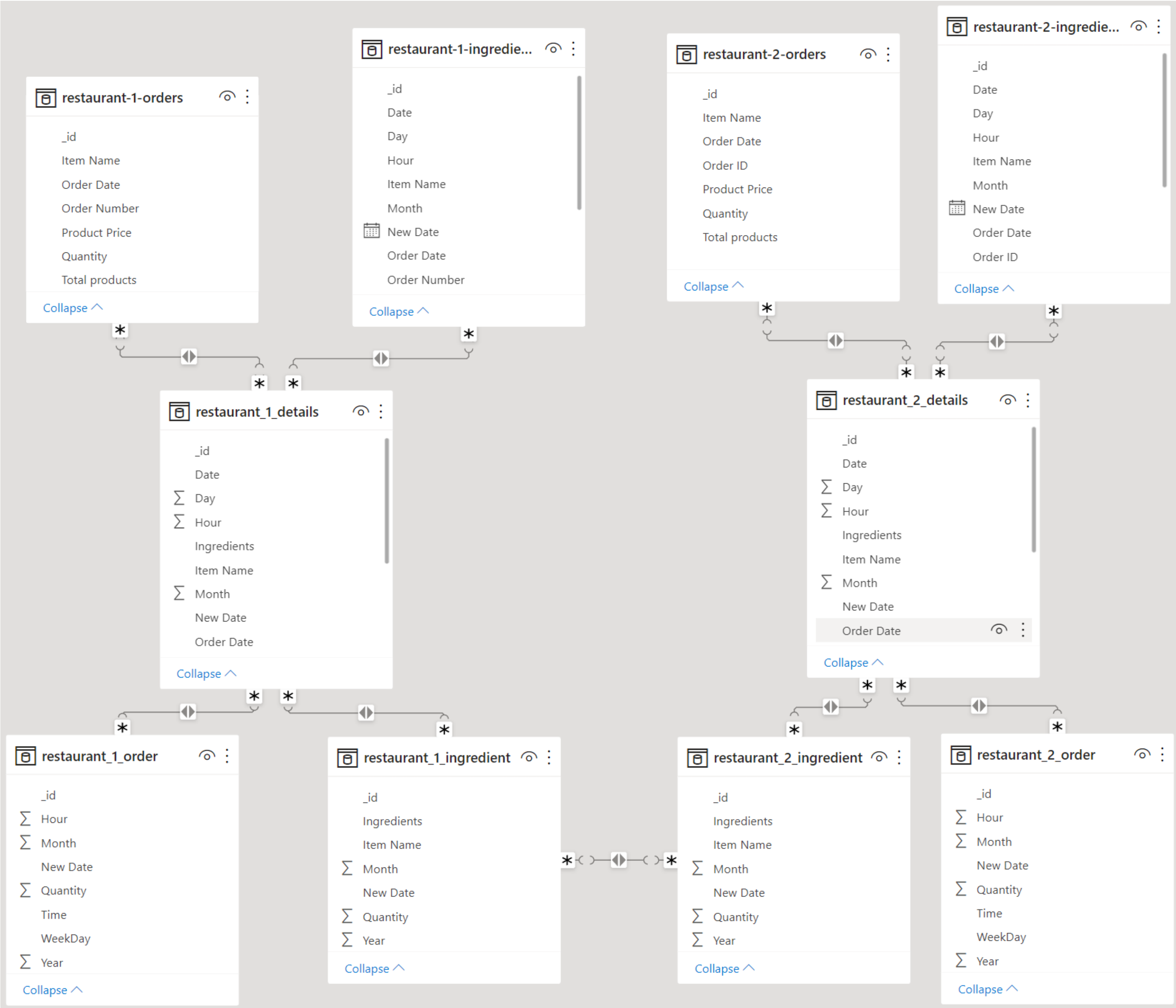

Figure 3 - Data Modelling

IV. RESULTS AND DISCUSSION

A. Result

We present our data visualization results in this section and briefly describe the outcomes.

In Figure 4 and 5, Ingredient frequency of Restaurant 1 and 2 were visualized accordingly for a specific month of a specific year. The ingredients were put in the x-axis and the amount was in the y-axis.

In Figure 6 and 7, Ingredient frequencies for the whole timeline were demonstrated. We determined and graphically depicted the total ingredient amount required for the entire timeline (2015-2019) for both restaurants.

In Figure 8 and 9, we determined how frequently each restaurant accepts orders. We began by visualizing the frequency for each weekday starting with a certain week of a month.

In Figure 10 and 11, we have plotted the frequency of orders for each restaurant over the course of a full month. This clarifies the need for ingredients for each month of the year.

In Figure 12 and 13, mean value of both restaurants order frequencies was visualized from 2015 to 2019.

Figure 14 and 15 showed the frequency of ingredients for restaurant 1 and 2 respectively for each year from 2015 to 2019. These two diagrams enable us to forecast when each item will be needed throughout the year for both restaurants.

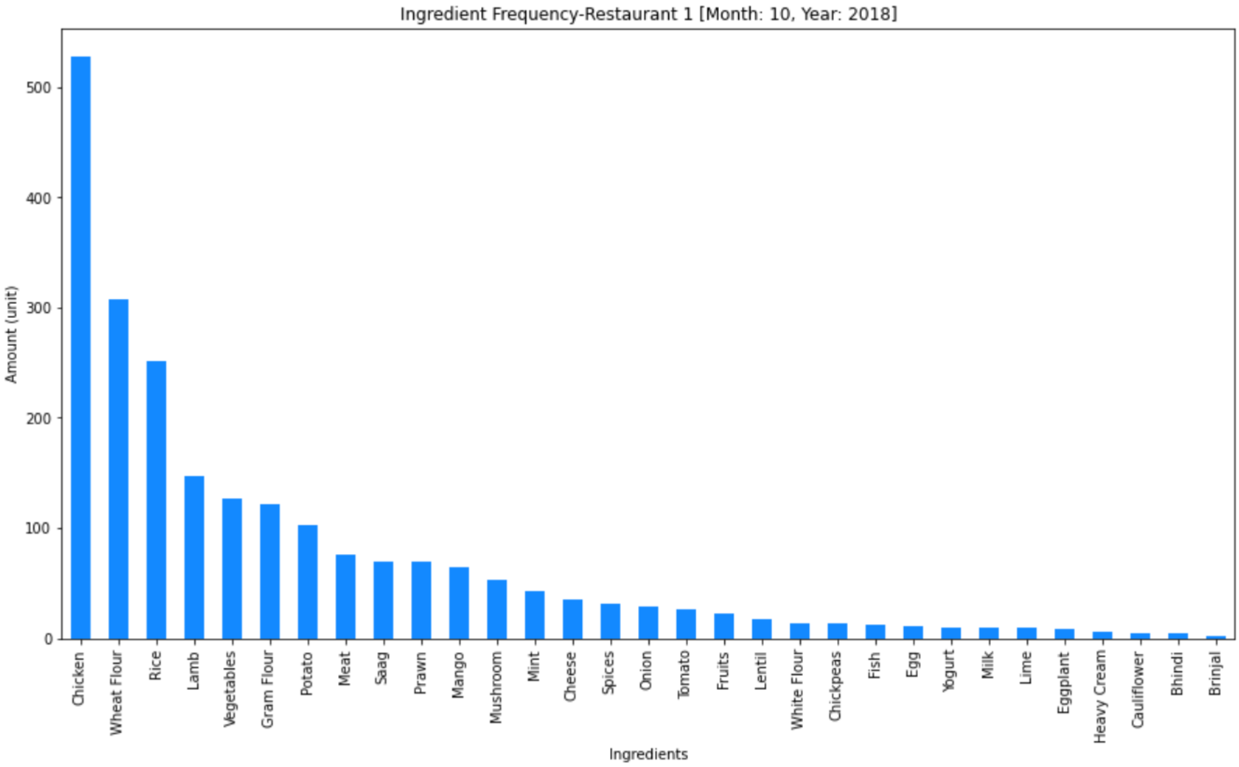

Figure 4 - Ingredient Frequency – Restaurant 1 (Month: 10, Year: 2018

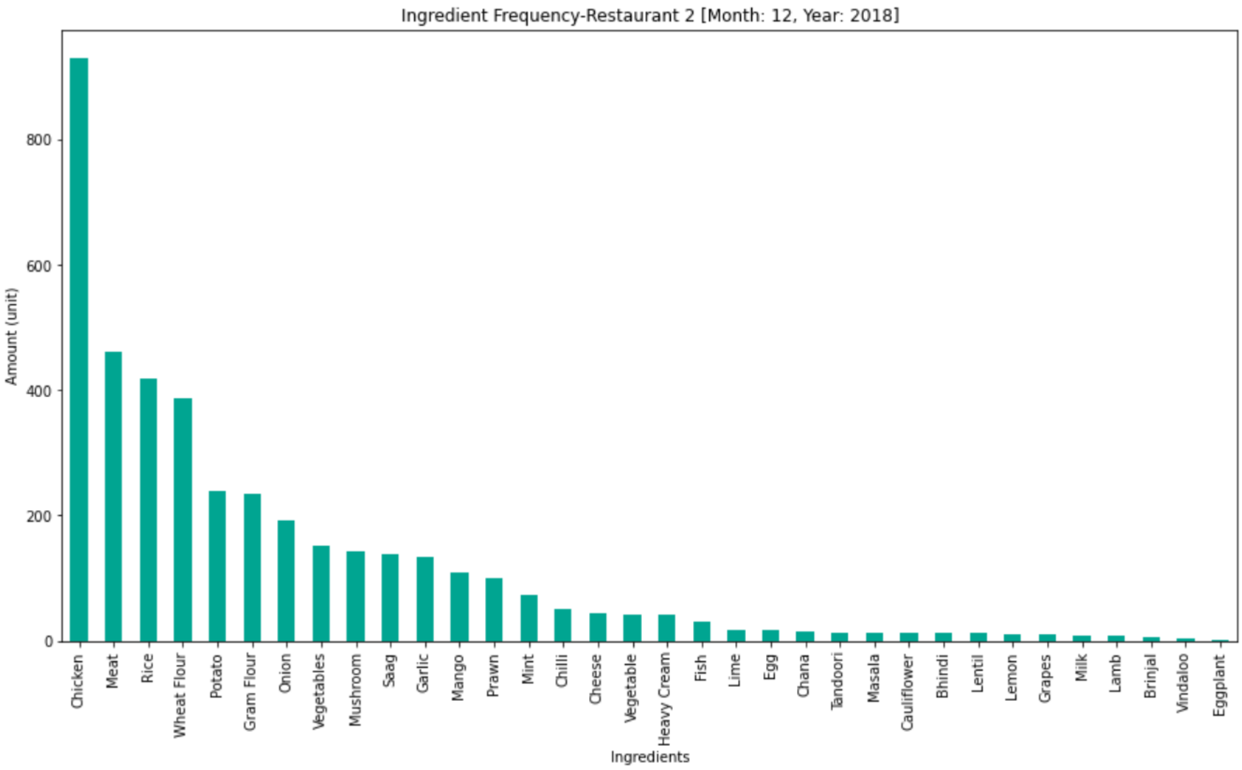

Figure 5 - Ingredient Frequency – Restaurant 2 (Month: 12, Year: 2018)

Here, we obtain the results of the frequency of ingredients for both restaurants for a specific month. It is obvious that chicken is the most popular ingredient for both eateries. Following that, more components such as meat, rice, gram flour, and wheat flour were added.

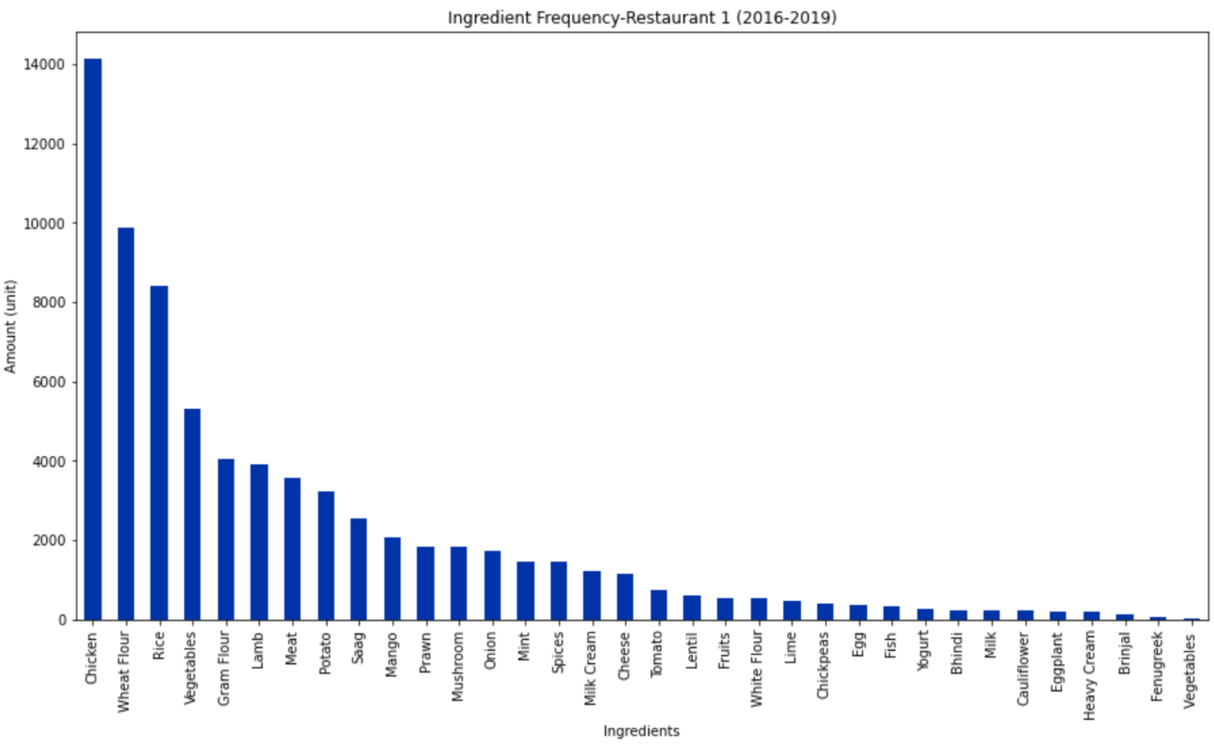

Figure 6 - Ingredient Frequency – Restaurant 1 (2016 - 2019)

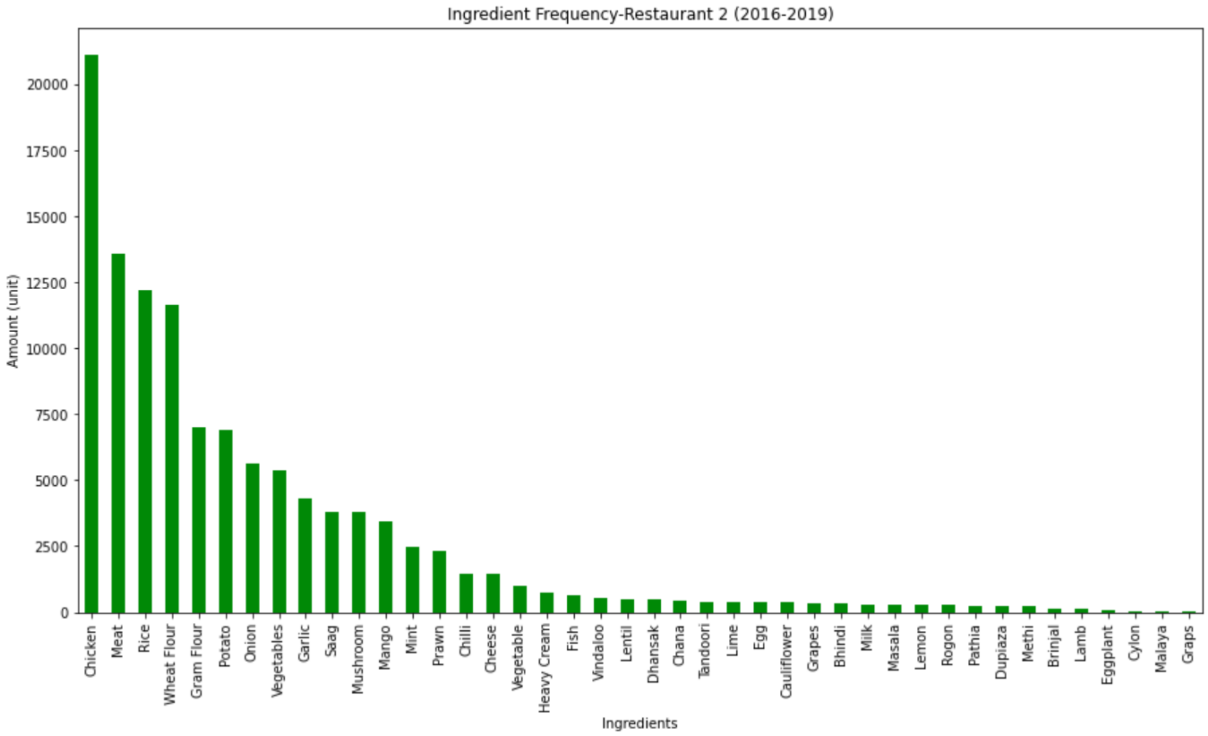

Figure 7 - Ingredient Frequency – Restaurant 2 (2016 - 2019)

In these two figures, ingredients were displayed along the x-axis, while frequency was plotted along the y-axis. By doing this, we can quickly understand the most important ingredients. We can see for both restaurant Chicken, Wheat Flour, Meat, Rice, and Lamb are the top needed ingredients. As these ingredients are used so frequently in almost every dish so these should be more in stock. Other ingredients can be ordered in a limited amount.

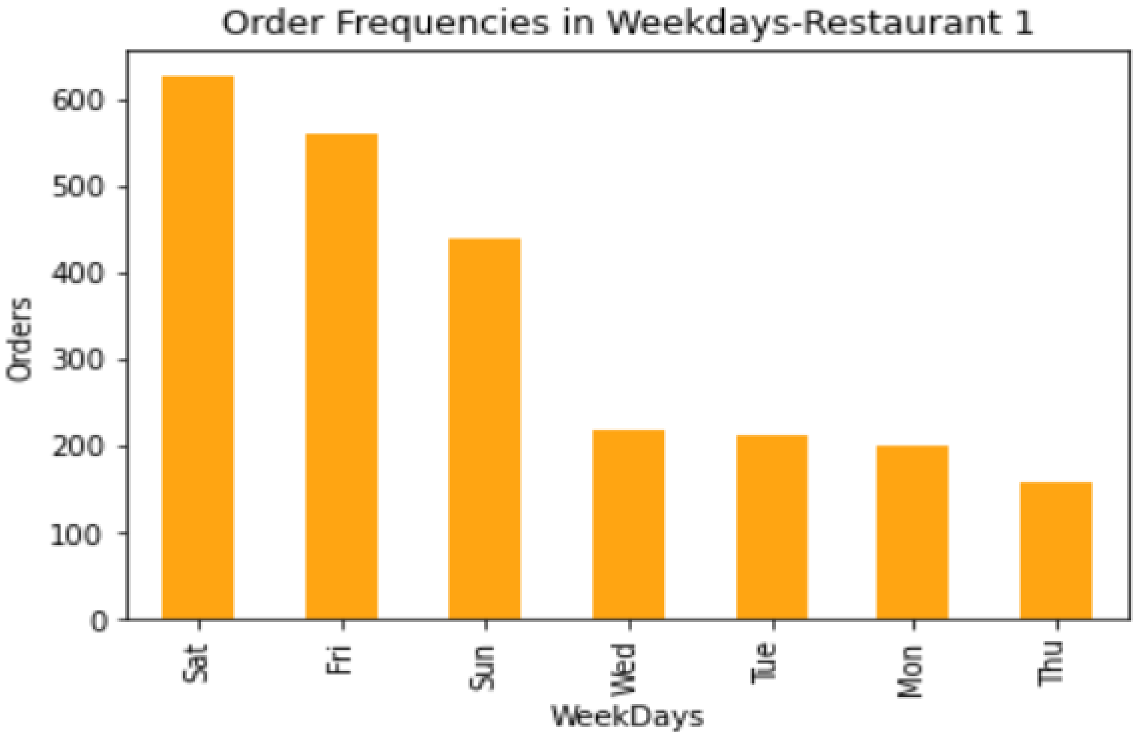

Figure 8 - Order Frequencies in Weekdays - Restaurant 1

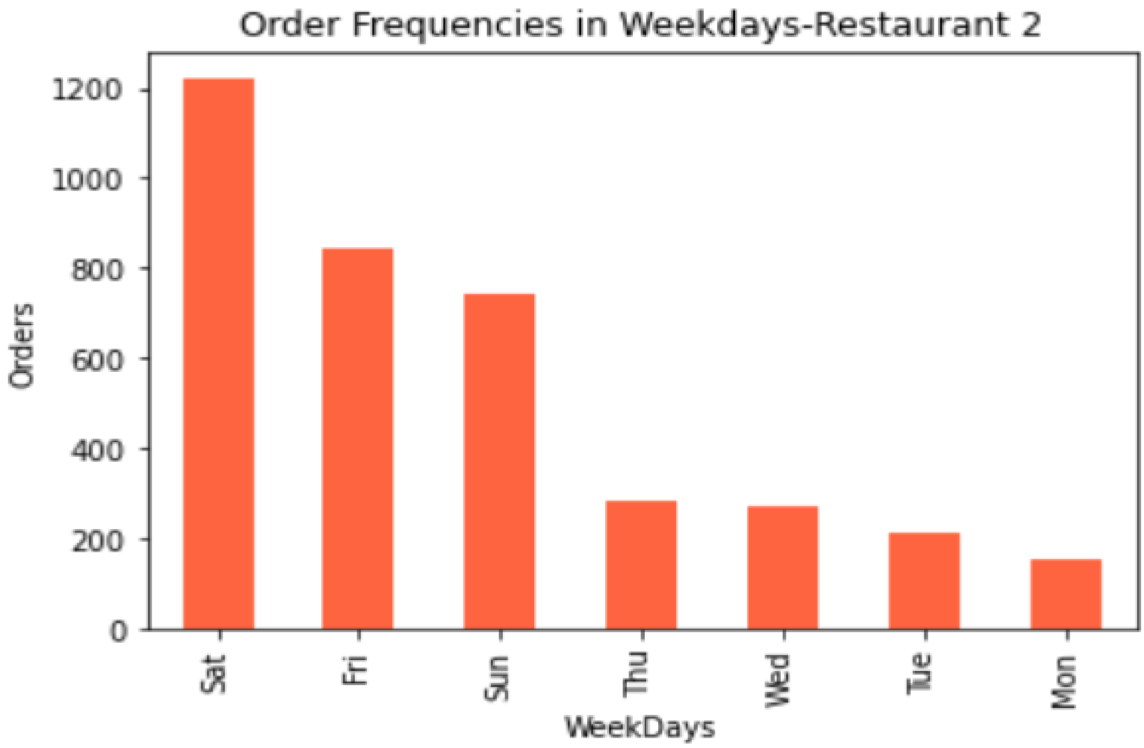

Figure 9 - Order Frequencies in Weekdays - Restaurant

How many orders per day were placed for each restaurant can be determined from these two diagrams. Weekdays were put in the x-axis and orders amount was in the y-axis. From this type of plot, we can estimate the frequency of order for each day. We can notice that during Saturday and Friday the order frequency is at its peak level. Other weekdays seem average. So, we can estimate on the weekend, Saturday and Friday, both restaurants need more ingredient than weekdays.

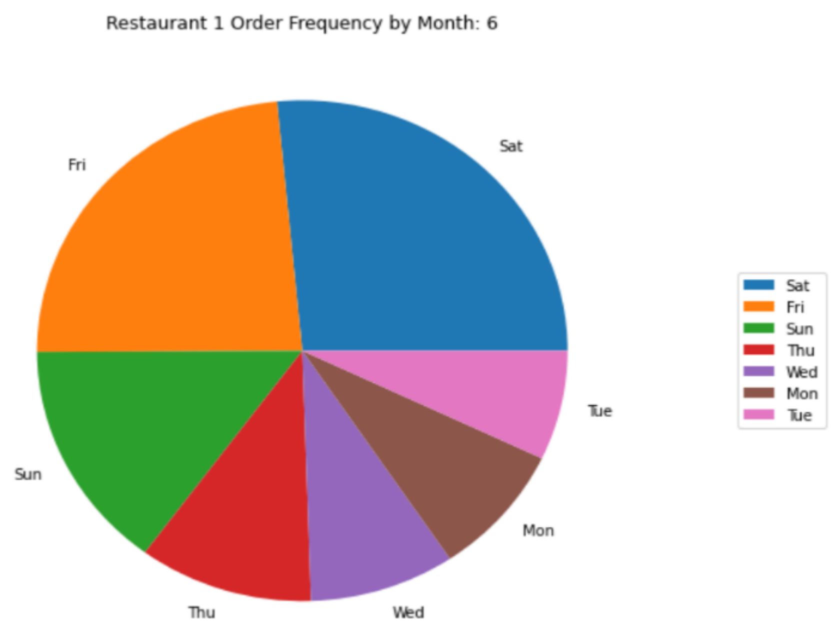

Figure 10 - Restaurant 1 Order Frequency by Month: 6

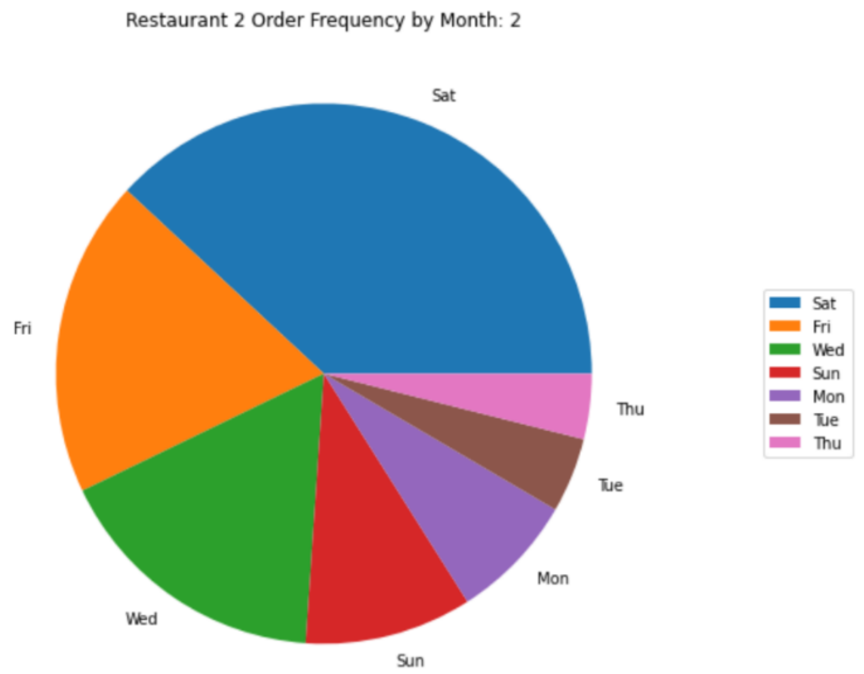

Figure 11 - Restaurant 2 Order Frequency by Month: 2

For both restaurants, we can see according on order frequencies, Friday and Saturday are the busiest days of the week during the whole month. So, we may anticipate that a restaurant will require the most ingredients on all the Fridays and Saturdays of a month.

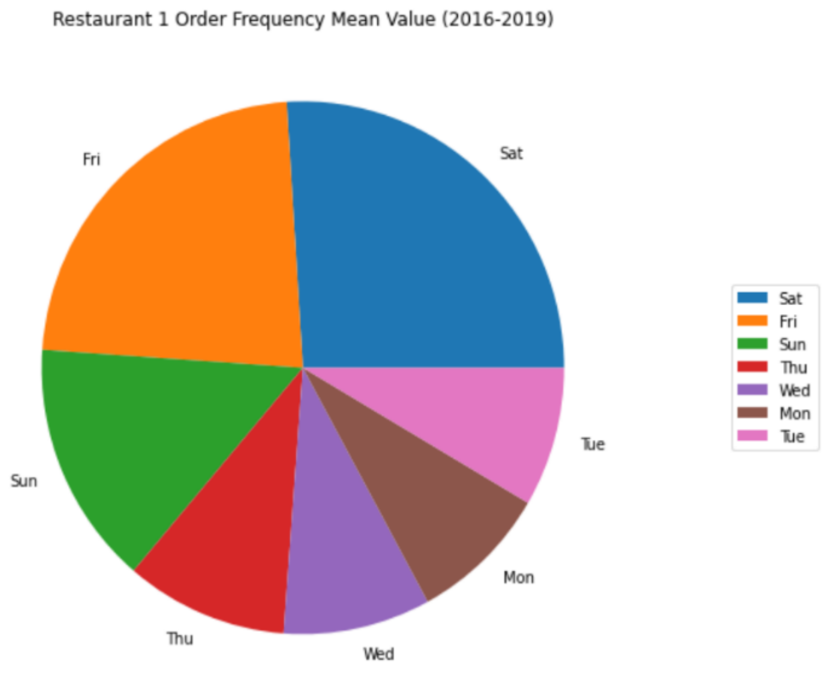

Figure 12 - Restaurant 1 Order Frequency Mean Value (2016 - 2019)

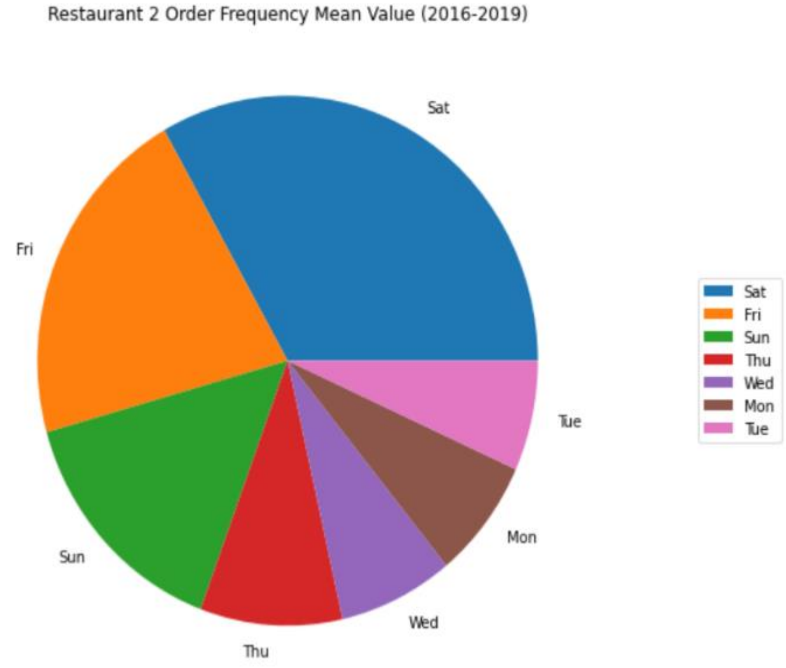

Figure 13 - Restaurant 2 Order Frequency Mean Value (2016 - 2019)

We can also see from the mean order frequency that for both eateries, Saturday and Friday saw the highest volume of orders. On these two days, we may anticipate that the ingredients will be used at their highest rate to fulfill the highest volume of orders.

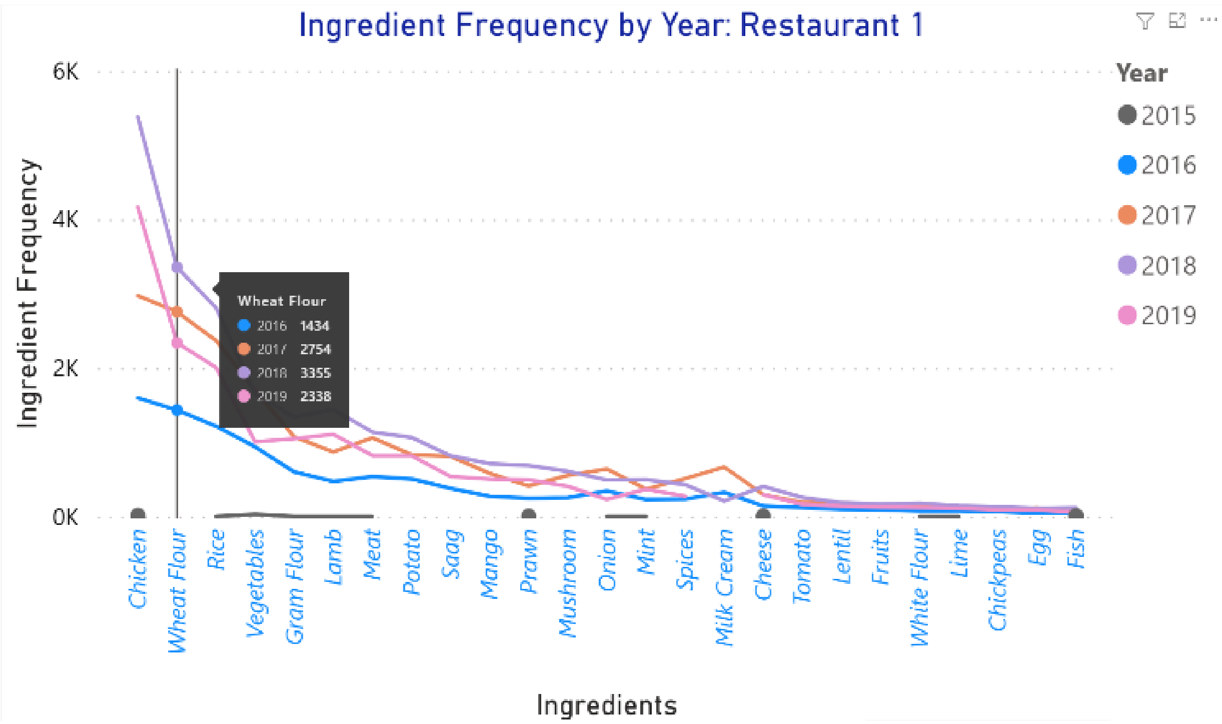

Figure 14 - Ingredient Frequency by Year: Restaurant 1

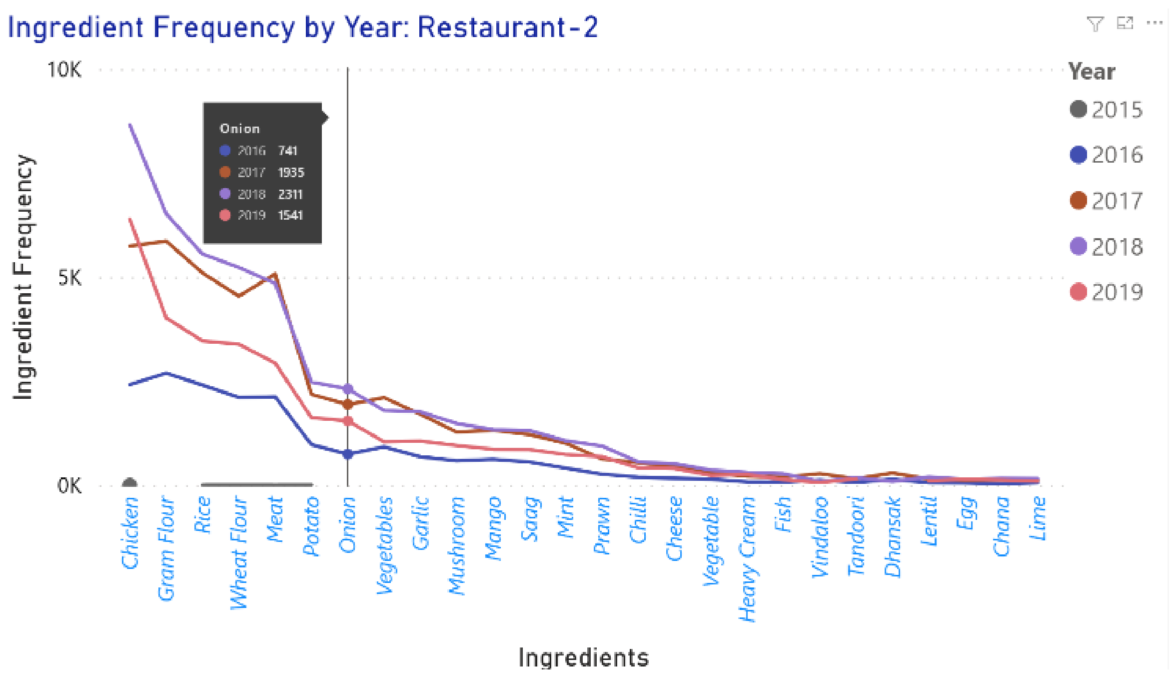

Figure 15 - Ingredient Frequency by Year: Restaurant 2

In these two diagrams, Ingredients are kept on the x-axis while ingredient frequency is kept in the y-axis in this graph. However, multiple lines for all years from 2015 to 2019 were drawn in the y-axis. So that we can detemine ingredient requirement for each year.

For both restaurants, we can establish which ingredient was utilized at what range and in which year. The use of ingredients can be compared between years as well as eateries. For instance, some ingredients had a growth in utilization during the past several years, while others lost favor due to a decrease in usage. For instance, consumption of chicken rose from 2016 to 2017, but it fell from 2017 to 2018. These two graphs, such as this one, allow us to study numerous types of data.

B. Discussion

By examining food orders, we hoped to reduce restaurant food waste as part of our experiment. To compare the two restaurant datasets analytically, we chose two. We have so far created data visualizations for restaurant-1 ingredient and restaurant-2 ingredient. Additionally, we determined the overall ingredient frequency over the timeline.

We have also calculated and displayed the order frequency for weekdays. We took into account average order frequencies as well as monthly frequency for each eatery. If the dataset has the data, we may determine the ingredient frequency for every given month.

Finally, we predicted the frequency of ingredients for restaurants 1 and 2 for every year.

The number of ingredients needed by a restaurant at a particular time of the year is determined by all of these analyses. A restaurant manager or owner can quickly learn when and what ingredients are necessary for his or her restaurant during a certain season by utilizing these analyses. Using that information, it is simple to estimate what type of ingredient should be ordered and in what quantity at when period of the year or month. If a restaurant manager or owner can do it effectively, they can save a lot of raw materials with a short expiration date since they will order just enough. As a result, the ingredient won't be squandered because it was ordered in excess for a certain season. This project is also adaptable to many methodologies. A restaurant owner can also forecast which weekdays will see the highest volume of orders. By being aware of that, more staff members, such as a chef, kitchen assistant, dishwasher, waiter, and so on, can be employed to efficiently manage the busiest hour. Additionally, by combining the knowledge of reducing food waste with the expertise of hiring additional workers, a restaurant's profit can be increased. With the help of our initiative, it can be easily predicted, and a restaurant owner can use this information to adopt changes that will increase revenue and earn positive client feedback.

V. CONCLUSION

Data is one of the most valuable commodities in the world today, yet it has no worth until we use it to our advantage. If we can manage these massive volumes of data, we can find solutions to our issues in the future. Food waste is one of today's most concerning problems. There are various ways that food might be wasted, but we concentrate on the wastage that results from overbuying groceries in restaurant. Although people are trying to reduce restaurant food waste by various ways like using the food donation applications for instance Throw no More app is used in Norway to give way rest over food available in restaurants, but it would be more efficient if we could take some analytical measures before replenishing food storage. These measures can help to reduce food wastage by preventing overbuying of restaurant raw ingredients.

In section III (A), we mentioned the datasets we use to achieve our goal. From the datasets, we do the data visualization approach to simplify the outputs.

From our data management restaurant food order analysis project, we can easily estimate and visualize graphically the frequency of each dish orders for each month and the frequency of each dish in total dataset. The greatest benefit of the visual depiction is how simple it is for non-technical individuals to comprehend the analysis. As a result, a conventional businessperson, such as a restaurant owner, can simply comprehend which ingredient he or she needs to purchase during which month. As a result, we can limit the excessive food purchasing, which leads to minimize food waste.

To increase a restaurant's profit, we will attempt to use analysis data on the frequency of orders and culinary ingredients in the future. We will do it by forecasting the staffing requirements on weekdays during particular times of the month. For instance, more people visit restaurants over the course of holidays. We will take such elements into account and use them in our project. so that a restaurant owner can make the most profit possible. For any non-technical individual to use our software, we will strive to build a graphical user interface as well.

VI. CONTRIBUTION

The contribution of the project members is outlined in the below table.

Data Collection	8558, 8563
Create Spark Session, Load Dataset	8563
Create new datasets for Ingredients and manually inserted all the ingredient for each item.	8558
Separate time and date from order date column, Convert date schema string to timestamp format	8563
Finding the Weekdays from Date	8563
Merge Order datasets and Ingredient Datasets and create Details Datasets	8558
Create connection with MongoDB and Saving dataframes to Mongo DB	8558
Dataset Cleaning	8563
Generate bar chart and Pie chart for all the ingredients	8558
Create user defined functions to generate the specific month or year bar chart.	8558
Generating mean value of order frequency (2015-2019) for both restaurants	8558
Connect the MongoDB with Power BI	8563

Dataset Modeling	8563
Ingredient Frequency for all years visualization	8558
Report: Abstract, Introduction, Background Literature Review, Result and Discussion (50%) Acknowledgement, Conclusion	8558
Report: Methodology, Result and Discussion (50%)	8563

VII. References

[1] "Global Food waste in 2022." https://www.greenly.earth/blog-en/global-food-waste-in-2022 (accessed Sep. 09, 2022).

[2] A. Dhir, S. Talwar, P. Kaur, and A. Malibari, "Food waste in hospitality and food services: A systematic literature review and framework development approach," *J. Clean. Prod.*, vol. 270, p. 122861, Oct. 2020, doi: 10.1016/J.JCLEPRO.2020.122861.

[3] C. Tostivint *et al.*, "Food waste quantification manual to monitor food waste amounts and progression," 2016.

[4] L. en Wang *et al.*, "The weight of unfinished plate: A survey based characterization of restaurant food waste in Chinese cities," *Waste Manag.*, vol. 66, pp. 3–12, Aug. 2017, doi: 10.1016/J.WASMAN.2017.04.007.

[5] V. Filimonau, A. D.-T. management, and undefined 2019, "Food waste management in hospitality operations: A critical review," *Elsevier*, Accessed: Sep. 09, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0261517718302449

[6] L. Principato, C. Pratesi, L. S.-I. J. of Hospitality, and undefined 2018, "Towards zero waste: An exploratory study on restaurant managers," *Elsevier*, Accessed: Sep. 09, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0278431917306138

[7] A. Gupta, H. K. Thakur, R. Shrivastava, P. Kumar, and S. Nag, "A Big Data Analysis Framework Using Apache Spark and Deep Learning," in *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, 2017, pp. 9–16. doi: 10.1109/ICDMW.2017.9.

[8] "What is Apache Spark? The big data platform that crushed Hadoop | InfoWorld." https://www.infoworld.com/article/3236869/what-is-apache-spark-the-big-data-platform-that-crushed-hadoop.html (accessed Sep. 11, 2022).

[9] W. McKinney, "pandas: a Foundational Python Library for Data Analysis and Statistics," 2011.

[10] J. D. Hunter, "Matplotlib: A 2D Graphics Environment," *Comput. Sci. Eng.*, vol. 9, no. 03, pp. 90–95, May 2007, doi: 10.1109/MCSE.2007.55.

[11] Z. Wei-Ping, L. Ming-Xin, and C. Huan, "Using MongoDB to implement textbook management system instead of MySQL," *2011 IEEE 3rd Int. Conf. Commun. Softw. Networks, ICCSN 2011*, pp. 303–305, 2011, doi: 10.1109/ICCSN.2011.6013720.

[12] Z. Wen *et al.*, "Design, implementation, and evaluation of an Internet of Things (IoT) network system for restaurant food waste management," *Waste Manag.*, vol. 73, pp. 26–38, Mar. 2018, doi: 10.1016/J.WASMAN.2017.11.054.

[13] S. K. Panda and M. Dwivedi, "Minimizing Food Wastage Using Machine Learning: A Novel Approach," *Smart Innov. Syst. Technol.*, vol. 159, pp. 465–473, 2020, doi: 10.1007/978-981-13-9282-5_44.

[14] C. Martin-Rios, C. Demen Meier, and S. Pasamar, "Sustainable waste management solutions for the foodservice industry: A Delphi study," *Waste Manag. Res.*, vol. 40, no. 9, pp. 1412–1423, Sep. 2022, doi: 10.1177/0734242X221079306/ASSET/IMAGES/LARGE/10.1177_0734242X221079306-FIG2.JPEG.

[15] W. Kim, C. Che, and C. Jeong, "Food Waste Reduction from Customers’ Plates: Applying the Norm Activation Model in South Korean Context," *L. 2022, Vol. 11, Page 109*, vol. 11, no. 1, p. 109, Jan. 2022, doi: 10.3390/LAND11010109.

[16] "Takeaway Food Orders - Updated | Kaggle." https://www.kaggle.com/datasets/henslersoftware/19560-indian-takeaway-orders?select=restaurant-1-orders.csv (accessed Sep. 14, 2022).