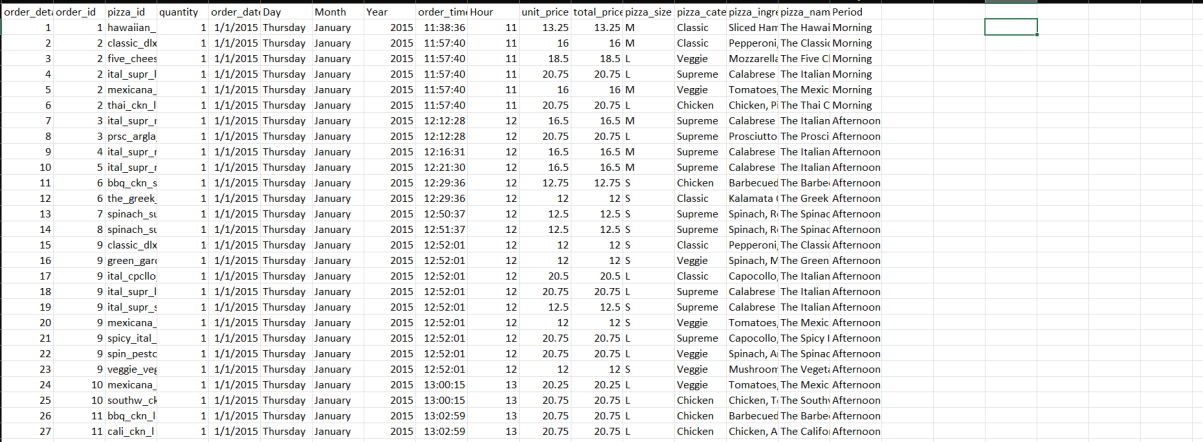
Pizza Restaurant Data Analysis and Visualization Project

The Pizza Restaurant Sales dataset was obtained via Kaggle, and it provides insights into the operations of the business throughout 2015. For this data, we are playing the role of a BI Consultant hired by the business to use their data in order to improve operations. They have transactional data collected throughout the year, but they have not been able to put it into very good use. They want us to analyze the data and put together a report to help them find opportunities to drive sales and increase work efficiency. There are 12 different columns and a total of 48620 rows. Columns include information such as Order\_Id, Order\_Details\_Id, pizza\_id, quantity, order\_date, and more. The data was downloaded in an XLS format then converted in a CSV to be uploaded into Big Query. My goal was to answer the questions Plato’s Pizza had and provide recommendations on how to improve operations. The questions the restaurant had are:

* What days and times do we tend to be busiest?
* How many pizzas are we making during peak periods?
* What are our best and worst-selling pizzas?
* What’s our average order value?
* How well are we utilizing our seating capacity? (we have 15 tables and 60 seats)

**Excel Sheet Editing**

Before we begin working with the data in Big Query, there are a few edits that I have made to the CSV file to make querying and visualization easier. First, I used the TEXT function such as =TEXT(E2,”dddd”), =TEST(E2,”mmmm”), and =TEST(E2,”yyyy”) to pull the Day Name, Month Name, and Year from the date column. Then, I used the HOUR function to pull only the hour from the time into its own column. Finally, I used an IFS function to group each row into a time period based on the hour column such as =IFS(AND(J2>=6,J2<12)=TRUE,AND(J2>=12,J2<18)=TRUE,”Afternoon”,AND(J2>=18,J2<24)=TRUE,”Evening”)



**Data Analysis**

1. **What days and times do we tend to be busiest?**

I first began by running these two queries.

Text

Description automatically generated

For this question, I wanted to work with the Day and Quantity with Quantity being summed and renamed into Orders. Then, they were grouped by the day column and ordered in descending order to show the top results first. The same applies to the second query, however Month is now used instead of the Day column.

Table

Description automatically generated

Table

Description automatically generated

The busiest day of the week would be Friday since they have 8,242 orders made throughout the year. The busiest month would have to be July with 4,392 orders placed. By looking at this, we can conclude that the busiest times of the year would be Fridays in July.

Chart, treemap chart

Description automatically generated

Chart, bar chart

Description automatically generated

1. **How many Pizzas are we making during peak periods?**

I used these queries to answer these questions.

Text

Description automatically generated

The first query is similar to the ones used in the last question. We use the Hour and Quantity columns with Quantity being summed up. Then, it is grouped by the Hours and ordered by the Quantity descending to show the top results first. For the second query, we still use the same columns, however now we use a CASE statement to put the hours into different groups. Hours 6 – 11 would be grouped into Morning, Hours 12-17 would be grouped into the Afternoon, and Hours 18-24 would be grouped into Evening. The new column would be renamed into Time Period. The Quantity column would be summed up again and renamed into Orders. Then, they were grouped by the Time Period. Running the queries gives these results.

Table

Description automatically generated with low confidence

A picture containing application

Description automatically generated

We can see that 12 PM would be the busiest times at 6,776 order made during the hour. 9 AM is the least busy since there are only 4 orders made in total. Afternoon periods are the busiest since they made 29,468 orders made in the time frame during 2015. Pizzas are hardly ever ordered in the morning. If we combine this with what we found in the previous question, we can conclude that the restaurant is busiest on Fridays in July in the Afternoon around 12pm.

Chart, line chart

Description automatically generated

Chart, bar chart

Description automatically generated

1. **What are the best and worst selling pizzas?**

I used the following queries to answer this question.

Text

Description automatically generated

For the first one, I wanted to see which of the pizzas gets ordered the most. I am using the pizza\_name column renamed as Pizza Name, pizza\_ingredients renamed as Ingredients, and quantity summed renamed as Orders. Then, they were grouped by the pizza name and ingredients in a descending order to show the top results first.

Graphical user interface, text, application

Description automatically generated

Chart, treemap chart

Description automatically generated

The best selling pizzas would be the Classic Deluxe, Barbecue Chicken, Hawaiian, Pepperoni, and Thai Chicken Pizzas.

Table

Description automatically generated

The worst selling Pizzas are the Soppresseta, Spinach Supreme, Calabrese, Mediterranean, and Brie Carre.

Next, I wanted to see what the best and worst selling Pizza sizes were. I would be using the pizza\_size column renamed as Pizza Size, and quantity summed once again renamed as Orders. The columns were grouped by the Pizza Size and ordered descending.

Application

Description automatically generated with medium confidence

Chart, bar chart

Description automatically generated

Large Pizzas sold the best with 18,956 total orders, while Extra-Extra Large Pizzas sold the least with on 28 total orders.

Now, I wanted to look at all the pizzas again by their category such as Classic, Supreme, Veggie, and Chicken. I selected the pizza\_category column and renamed it into Category, and I summed the quantity column renaming it into Orders. They were grouped by the category and order descending as usual.

A picture containing table

Description automatically generated

Graphical user interface, application

Description automatically generated

We can see that Classic category pizzas sold the best at 14,888 orders and Chicken category pizzas sold the least at 11,050 orders. Based on what we have seen. We can infer that the best selling pizzas would be the large Classic Deluxe, and the worst selling pizza would be the extra-extra-large Brie Carre pizza.

1. **What is our average order value?**

The average order value is calculated by dividing the total revenue by the total amount of orders. Therefore, I will be summing the total\_price column and rounding it by 2 into order to get the revenue. Quantity would be summed, and the columns would be renamed Revenue and Orders respectively. I used a nested query here to have everything in one step. My formula would be the sum of the total\_price column divided by the sum of the quantity column rounded to 2 decimal places. This new column would be renamed as the Average\_Order\_Column.

Text

Description automatically generated

The average order value 16.5 dollars.

1. **How well are we utilizing our seating capacity?**

Finally, for this query I will be looking at the Hour, Quantity, and Order ID columns. Quantity will be summed and renamed as Pizzas, then I will be getting a distinct count of the order\_id column and rename it as Total Orders. Since July is our busiest month and Friday tends to be the busiest week day, filtering the results so that only a Friday in July will appear. As such, I have specifically chosen July 17th, 2015. Then, the columns were grouped by the hour.

Table

Description automatically generated

We can see that at 9pm, there were 13 total orders made. Since the restaurant has 15 tables and 60 seats, they are utilizing their seating capacity very well.

Chart, line chart

Description automatically generated

**Results**

* Orders increase on Thursday, Friday, and Saturdays.
* Busiest during the Afternoon and Evenings.
* Customers prefer to order large pizzas and Classic is the best-selling Category.
* Extra-Extra-Large is the least ordered size and Chicken is the least popular category.
* The most orders made are usually at 12pm and 9-10am have the least number of orders.
* Classic Deluxe is the best-selling pizza while Brie Carre is the worst selling.