**Pizza Place End to End Data Analysis with PowerBI.**

**Background Information.**

A year's worth of sales from a fictitious pizza place, including the date and time of each order and the pizzas served, with additional details on the type, size, quantity, price, and ingredients.

**Questions and Recommended Analysis**

1. How many customers do we have each day? Are there any peak hours?
2. How many pizzas are typically in an order? Do we have any bestsellers?
3. How much money did we make this year? Can we identify any seasonality in the sales?
4. Are there any pizzas we should take of the menu, or any promotions we could leverage?
5. **Data Collection.**

* This data was gotten from the Maven Analytics Data Play Ground in CSV format containing 5 different files.

1. **Basic Data Transformation.**
2. **Renaming Tables and Columns as per the reporting format.**

I hade to rename the tables as they contained underscores especially the table column names.

They were all named with underscores in between them which I had to change to make it easier to use and to read

1. **Promoting Row to Header.**

I did this for the Pizza Type as during the importation of the dataset the first row was auto generated

1. **Data Type Check.**

All the data types for each column form each table had the correct data types.

1. **Replacing null with 0.**

There were no null values in this dataset.

1. **Data Modelling**

In this section I had t created new tables, measures and add new columns using the existing data to create new relationships to aid in answering some of the end goal questions.

Initial Relationships in the datasets were: -

* Pizza Type Table and Pizzas Table were related via Pizza Type Id
* Pizzas Table and Order Details Table were related via the Pizza Id
* Orders Table and Order Details Table were related via the Order Id

1. **Creating the Date Table**

First, I created a new Table to extract the timelines from the dates i.e.: - The Year, Month, Month No. Quarter, Day and Day Number.

This will aid in determining recurrent trends over a certain period in time.

Date Table =

ADDCOLUMNS(

    CALENDAR(MIN(Orders[Date]), MAX(Orders[Date])),

    "Year", YEAR([Date]),

    "Month", FORMAT([Date], "mmm" ),

    "Month No", MONTH([Date]),

    "Quarter", FORMAT([Date], "\QQ"),

    "Day", FORMAT([Date], "ddd"),

    "Day no", WEEKDAY([Date])

)

I created a new relationship between the Orders Table and the newly created Date Table using the date columns.

1. **Creating a Measures Table**

Created a table to hold all the measures I will be calculating from the dataset.

The first measure was calculating the **Total Sales** using the prices from the Pizzas Table and Quantity from the Order Details Table.

Total Sales =

SUMX(

    'Order Details',

    'Order Details'[Quantity] \* RELATED(Pizzas[Price])

)

**Challenge:** Since the price column was not in the Order Details table it forced us to pull it from the Pizzas table and that’s why we have used the related function.

To answer the question of peak hours of operation we can use the Time column from the Orders table **creating a new column with time** from around 9am to 12 pm – this is the time frame we get from the data when you arrange the column in ascending and descending order.

Therefore, I added a **new column Time slot** to the Orders table using the following DAX command.

Time Slot =

SWITCH(

    TRUE(),

    Orders[Time] >= TIME(9, 0, 0) && Orders[Time] <= TIME(12, 0, 0), "9 - 12AM",

    Orders[Time] >= TIME(12, 0, 0) && Orders[Time] <= TIME(15, 0, 0), "12 - 3PM",

    Orders[Time] >= TIME(15, 0, 0) && Orders[Time] <= TIME(18, 0, 0), "3 - 6PM",

    Orders[Time] >= TIME(18, 0, 0) && Orders[Time] <= TIME(21, 0, 0), "6 - 9PM",

    Orders[Time] >= TIME(21, 0, 0) && Orders[Time] <= TIME(23, 59, 59), "9 - 12PM"

)

Because we don’t have the names of the customers to know the number of customers per day, we have to calculate the **Avg. Orders in a day** using measures.

Avg orders in a day =

DIVIDE(

    DISTINCTCOUNT('Order Details'[Order Id]),

    DISTINCTCOUNT(Orders[Date])

)

To get the **Average Number of Pizzas** we can divide the Order Details Id with the Order Id

Avg Pizzas in an order =

DIVIDE(

    DISTINCTCOUNT('Order Details'[Order Details Id]),

    DISTINCTCOUNT(Orders[Order Id])

)

**Challenge:**

When it came time to visualize the number of orders for the amount of time, the 9am – 12pm and 9pm – 11:59 pm weren’t arranging ascendingly order so I had to created another column “**Time Slot Sorting**” using the DAX function below and use that to arrange the times in ascending order I then used the time slot column simply as labels.

Time Slot Sorting =

SWITCH(

    TRUE(),

    Orders[Time] >= TIME(9, 0, 0) && Orders[Time] <= TIME(12, 0, 0), 1,

    Orders[Time] >= TIME(12, 0, 0) && Orders[Time] <= TIME(15, 0, 0), 2,

    Orders[Time] >= TIME(15, 0, 0) && Orders[Time] <= TIME(18, 0, 0), 3,

    Orders[Time] >= TIME(18, 0, 0) && Orders[Time] <= TIME(21, 0, 0), 4,

    Orders[Time] >= TIME(21, 0, 0) && Orders[Time] <= TIME(23, 59, 59), 5

)

To get the **Number of Orders Per Day** in the visualizations I created a orders against days graph.

To order it I modified the time slot function to give me the day number then ordered it by day number in ascending number giving is a clear visual from Sunday to Monday.

**Answering the Questions**

1. **How many Customer do we have each day? Are there any peak hours?**

Since the data does not provide any unique identification for each Customer I assumed the number of unique orders to be the number of customers per day of which we have on average 60 orders in a day

Our peak hours of orders are from 12pm – 3pm, closely followed by 6pm to 9pm.

1. **How many Pizzas are Typically in an Order? Do we have any best sellers?**

There are on average 2 pizzas in an order

Our best sellers according to sales was the Chicken Category specifically The Barbecue Chicken Pizza making a total sale of $42, 768.00, The California Chicken Pizza making $41, 409.50, The Thai Chicken Pizza - $43, 434.25, The Classic Deluxe Pizza - $38, 180.50 and The Spicy Italian Pizza - $34, 831.25

Our best sellers according to Number of Orders was the classic garnering 29.02% of all the orders

1. **How much money did we make this year? Can we Identify any Seasonality in the Sales?**

We made $817, 860.05 in sales this year (2015).

Our sales seem to be seasonal with our lowest months being October followed by September and third lowest being December – we can also include February with each of these months having orders below 1, 700 in number.

Our Highest months are July, January, May, March and August in that order each having orders above 1, 800

1. **Are there any Pizzas we could take off the menu? Or any promotions we could leverage?**

We could take the Bottom 5 Worst performing pizzas off the menu, they are mostly from the Supreme and Veggie Category, they include The Brie Carrie Pizza - $11, 588.50, The Green Garden Pizza - $13, 995.75, The Spinach Supreme Pizza - $15, 277.75, The Mediterranean Pizza - $15, 360.50 and The Spinach Pesto Pizza - $15, 596.00

To promote this bottom 5 Pizzas, we can give discounts for people who buy more than 3 while also offering a free sample for those who buy more than 10 of our better performing options, its estimated that some people have a negative opinion about something they haven’t really tried themselves.

**Recommendations.**

1. We should transfer most of our morning shift employees to the afternoon and afternoon shift specifically from 12pm – 9pm.
2. We should consider reducing our operations for the months of September and October including reducing the number of suppliers stored for the months this will reduce costs in space and costs involved in storage.