Pizza Sales Analysis

A Comprehensive Report on Pizza Sales Trends for 2015

Submitted by:

Dhrumil Patel

Guided by:

Prof. Chintan Patel

Fly The Nest – Deets Digital

Acknowledgment

I would like to express my heartfelt gratitude to **Prof. Chintan**, my guide, for his unwavering support and expert guidance throughout this project. His invaluable insights and constructive feedback were crucial in the successful completion of this analysis.

I am also deeply thankful to **Fly the Nest** for providing the resources and platform to undertake this project. Their commitment to fostering innovation and learning has been instrumental in shaping this work.

I extend my sincere appreciation to my friends and family for their constant encouragement and motivation during the course of this project. Their belief in my abilities kept me focused and determined.

Additionally, I would like to acknowledge the data sources and tools that made this analysis possible, including the 2015 pizza sales records and data visualization platforms, which allowed for a comprehensive and detailed analysis.

Finally, I am grateful to everyone who contributed directly or indirectly to the success of this project.

Thank you all.

Abstract

The **Pizza Sales Analysis of 2015** provides a comprehensive examination of sales trends, customer preferences, and revenue patterns for the year, leveraging historical sales data sourced from Kaggle. This project identifies the best-performing pizza varieties, peak sales periods, and regional variations in demand. By employing statistical techniques and data visualization tools, the analysis uncovers actionable insights to enhance inventory management, marketing strategies, and operational efficiency.

The findings reveal that specific pizza types and promotional campaigns significantly influenced customer preferences, with peak sales observed during weekends and festive seasons. Additionally, the report highlights opportunities to address underperforming segments and improve customer engagement for sustained growth.

This project was conducted under the expert guidance of **Prof. Chintan** and supported by **Fly the Nest**, whose mentorship and resources were integral to its successful execution. The insights generated from this analysis provide valuable recommendations for optimizing future sales strategies in the pizza industry.

Contents

1. Introduction

1.1 Purpose of the Project

The primary purpose of this project is to analyse pizza sales data to uncover trends and insights that can inform decision-making in the pizza business. By analysing customer orders, pizza types, and revenue generation, the business aims to optimize pricing strategies, inventory management, and marketing efforts to boost profitability and customer satisfaction.

1.2 Background Information

The pizza industry is highly competitive, with customer preferences varying based on factors such as pizza type, size, and price. Understanding customer behavior and sales patterns allows businesses to tailor their offerings and marketing efforts to drive greater engagement and revenue.

1.3 Objectives

- Analyse historical pizza sales data.
- Identify trends in customer preferences, revenue, and demand patterns.
- Provide actionable recommendations to improve sales, customer engagement, and operational efficiency.

2. Data Collection and Source

2.1 Data Sources

The data used in this project was collected from a relational database containing information on orders, pizza details, pizza types, and pricing. The data includes information such as the number of pizzas ordered, their types, sizes, and total revenue.

2.2 Tables and Columns

The following tables were used for data collection:

2.2.1 Order Details Table

Column Name	Data Type	Description
order_details_id	int	Primary Key, unique identifier for order details
order_id	int	Foreign Key linking to orders table
pizza_id	text	Identifier for the pizza
quantity	int	Quantity of pizza ordered

2.2.2 Orders Table

Column Name	Data Type	Description
order_id	int	Primary Key, unique identifier for the order
date	date	Date the order was placed
time	time	Time the order was placed

2.2.3 Pizza Types Table

Column Name	Data Type	Description
pizza_type_id	text	Unique identifier for the pizza type
name	text	Name of the pizza
category	text	Category of pizza (e.g., Classic, Veggie, Chicken)
ingredients	text	List of ingredients in the pizza

2.2.4 Pizzas Table

Column Name	Data Type	Description
pizza_id	text	Unique identifier for the pizza
pizza_type_id	text	Foreign Key linking to pizza types
size	text	Size of the pizza (e.g., Small, Medium, Large, XXL)
price	double	Price of the pizza

3. Methodology

This section outlines the systematic approach followed for the analysis of pizza sales data. The project involved extracting data from a MySQL database and performing data analysis and visualization using **MS Excel**. Below is the detailed methodology:

3.1 Data Cleaning in MS Excel

Data cleaning is a crucial step to ensure the accuracy and reliability of the dataset. In this project, the following data cleaning steps were performed in **MS Excel**:

- **Removing Duplicates**: Any duplicate rows in the dataset were identified and removed to ensure that each record was unique.
- **Handling Missing Data**: Missing values in key columns such as order_id, quantity, and price were addressed. Rows with missing critical information were either corrected or removed.
- **Standardizing Formats**: Date and time columns were standardized to a consistent format, ensuring no discrepancies while performing further analysis.
- **Outlier Detection and Removal**: Outliers in the dataset (such as unusually high prices or quantities) were identified and treated to prevent skewed results.

3.2 Data Extraction Using MySQL Queries

Once the dataset was cleaned, relevant data was extracted from the database using **MySQL**. The MySQL database contained multiple tables, including order_details, orders, pizza_types, and pizzas. SQL queries were written to retrieve the required data:

- **Total Number of Orders**: A query was executed to calculate the total number of orders placed by customers.
- **Total Revenue**: A SQL query was written to calculate the total revenue generated from pizza sales.
- **Top-Selling Pizzas**: Queries identified the most ordered pizzas by quantity and revenue.
- **Time-Based Analysis**: SQL queries were used to analyze the distribution of pizza orders by hour of the day, identifying peak sales periods.

The queries were optimized to ensure fast and efficient retrieval of large datasets, with proper use of **JOIN** clauses to combine relevant data from different tables.

3.3 Topics Used in Writing SQL Queries

This section outlines the key SQL concepts and techniques used in writing queries for this project:

3.3.1 Basic SQL Syntax:

- Simple **SELECT** statements were used to retrieve data from tables.
- The **FROM** clause was used to specify the tables from which data was retrieved.

3.3.2 Filtering Data:

- The **WHERE** clause was used to filter data based on specific conditions (e.g., filtering orders by date or price).
- Logical operators such as **AND**, **OR**, and **NOT** were used to combine multiple conditions.

3.3.3 Sorting Data:

- The **ORDER BY** clause was used to sort the results based on certain columns (e.g., sorting orders by revenue or quantity in descending order).

3.3.4 Aggregation Functions:

- Functions such as **COUNT()**, **SUM()**, and **AVG()** were used to calculate the total number of orders, the total revenue, and average order quantities.
- **GROUP BY** was used to group data by specific columns (e.g., by pizza type or category) and aggregate results accordingly.

3.3.5 Joining Tables:

- **INNER JOIN** and **LEFT JOIN** were used to combine related data from different tables. For example, joining the orders and order_details tables to retrieve detailed order information.
- The **ON** clause was used to specify the conditions for joining tables, ensuring that data was correctly matched between tables.

3.3.6 Subqueries:

- Subqueries were used in certain cases to fetch aggregated data and use it in the main query. This helped in scenarios like fetching the highest-priced pizza from a subquery and using it in the main query to calculate the total revenue.

3.3.7 Date and Time Functions:

- Date and time functions such as **DATE()** and **TIME()** were used to extract specific parts of the date and time, such as the day of the week or the hour of the day, for time-based analysis.

3.3.8 Logical Expressions:

- The **CASE** statement was used to categorize data or perform conditional logic (e.g., categorizing pizzas based on price range).

3.3.9 SQL Query Optimization:

- Indexing was applied to relevant columns to improve query performance.
- **LIMIT** was used to restrict the number of results returned when working with large datasets, which improved query efficiency.

By applying these SQL techniques, complex queries were developed to extract meaningful insights from the database, such as total revenue, top-selling pizzas, and peak sales periods.

3.4 Tools and Technologies

This project utilized the following tools and technologies:

- **MS Excel**: Used for data cleaning, organizing, and basic analysis.
- **MySQL**: Used to write and execute queries for extracting and analysing sales data from the database.
- **Kaggle**: Accessed for additional resources and dataset exploration.
- **MS Word**: Used for writing and formatting the final project report.

4.Interpretation & Recommendations

4.1 Retrieve the total number of orders placed.

Query Result: 21,350 orders were placed in total.

Interpretation: This indicates a high demand for the pizzas offered by the business. The total number of orders reflects customer engagement and the overall scale of operations.

Recommendation:

a. Focus on maintaining high service quality to retain this large customer base.
b. Implement a loyalty program to further increase repeat customers and enhance retention rates.

4.2 Calculate the total revenue generated from pizza sales.

Query Result: $817,860.05 revenue was generated from pizza sales.

Interpretation: The revenue figure demonstrates the success of the business in monetary terms. It suggests that the pricing and sales volume are aligned with business goals.

Recommendation:

a. Evaluate pricing strategies periodically to maximize profit while remaining competitive.
b. Consider identifying low-revenue days or times and launching promotions or discounts during those periods to boost sales further.

4.3 Identify the highest-priced pizza.

Query Result:

Name: The Greek Pizza

Category: Classic

Size: XXL

Price: $35.95

Interpretation: The Greek Pizza in XXL size is the premium offering in the product lineup, likely targeting customers willing to spend more for larger portions or premium ingredients.

Recommendation:

a. Promote the highest-priced pizza as a luxury or gourmet option to attract high-spending customers.
b. Introduce combo offers including this pizza to increase its sales while encouraging larger order values.

4.4 Identify the most common pizza size ordered.

Query Result:

Name: The Classic Deluxe Pizza

Total Quantity Ordered: 2,416 pizzas

Interpretation: The Classic Deluxe Pizza is the most popular item on the menu, indicating strong customer preference for this particular option.

Recommendation:

a. Ensure consistent availability of ingredients for the Classic Deluxe Pizza to meet high demand.
b. Consider cross-selling or upselling options alongside this pizza, such as beverages, sides, or desserts.

4.5 List the top 5 most ordered pizza types along with their quantities.

Query Result:

The Classic Deluxe Pizza: **2,416 pizzas**

The Barbecue Chicken Pizza: **2,372 pizzas**

The Hawaiian Pizza: **2,370 pizzas**

The Pepperoni Pizza: **2,369 pizzas**

The Thai Chicken Pizza: **2,315 pizzas**

Interpretation: These top 5 pizzas dominate customer preferences, making them key drivers of business success. Their combined popularity shows the importance of offering a diverse menu catering to a variety of tastes.

Recommendation:

a. Highlight these top-selling pizzas in marketing campaigns and promotional materials.
b. Analyse the preferences of customers ordering these pizzas to identify potential patterns (e.g., time of purchase, additional items ordered) and optimize inventory and operations accordingly.
c. Use customer feedback to refine the recipes or introduce similar pizzas to expand the menu.

4.6 Join the necessary tables to find the total quantity of each pizza category ordered.

Interpretation: The Classic Deluxe Pizza is the most popular item, with 2,416 units sold. Chicken-based pizzas like The Thai Chicken Pizza and The Barbecue Chicken Pizza also rank among the top sellers, reflecting a strong preference for these flavours.

Recommendation:

a. Focus marketing efforts on promoting the Classic Deluxe Pizza as the pizzeria's signature dish.
b. Introduce combo offers or discounts featuring the top-selling pizzas to boost sales further.

4.7 Determine the distribution of orders by hour of the day.

Interpretation: Peak sales occur during lunch (12 PM - 1 PM) and dinner (6 PM - 8 PM), indicating that these are the primary revenue-generating time slots.

Recommendation:

a. Offer exclusive lunch and dinner deals to attract more customers during these hours.
b. Ensure staff and kitchen resources are adequately prepared for high demand during these periods.

4.8 Join relevant tables to find the category-wise distribution of pizzas.

Interpretation: Classic pizzas dominate orders (14,579), followed by veggie (11,449) and chicken (10,815) categories, suggesting a balanced preference among customers.

Recommendation:

a. Maintain a strong focus on classic pizzas while experimenting with new veggie and chicken-based options to cater to a broader audience.
b. Highlight the diverse menu in marketing campaigns to appeal to varied customer preferences.

4.9 Group the orders by date and calculate the average number of pizzas ordered per day.

Interpretation: Certain dates (e.g., January 5, January 29, February 22) and months show spikes in sales, often aligning with holidays or special occasions.

Recommendation:

a. Plan promotional campaigns, discounts, or themed events around these high-demand days and months.
b. Analyse holiday-specific patterns to introduce special menus or festive deals.

4.10 Determine the top 3 most ordered pizza types based on revenue.

Interpretation: The Thai Chicken Pizza (₹43,434.25) generates the most revenue, followed by The Barbecue Chicken Pizza and The Classic Deluxe Pizza.

Recommendation:

a. Feature these pizzas prominently in advertising, emphasizing their popularity and high customer satisfaction.
b. Experiment with premium pricing or gourmet variants for these pizzas to further maximize revenue.

4.11 Calculate the percentage contribution of each pizza type to total revenue.

Interpretation: The top contributors to revenue are *The Thai Chicken Pizza* (5.31%) and *The Barbecue Chicken Pizza* (5.23%). The bottom contributors include *The Brie Carre Pizza* (1.42%) and *The Green Garden Pizza* (1.71%).

Recommendation:

a. **Focus on High-Performing Pizzas**: Promote the top-performing pizzas through marketing campaigns and bundled offers.
b. **Revise Strategy for Low Performers**: Consider adjusting recipes, pricing, or marketing for low-performing pizzas to boost sales or remove them if profitability is not achievable.

4.12 Analyse the cumulative revenue generated over time.

Interpretation: The highest revenue day is *2015-11-27* with $4422.45, while revenue fluctuates significantly across other dates. High-revenue days often align with holidays, weekends, or specific promotional campaigns.

Recommendation:

a. **Leverage Peak Sales Periods**: Analyse the top revenue days to identify underlying factors (e.g., promotions, events, or seasons) and replicate successful strategies.
b. **Improve Low Revenue Days**: Strategize to boost sales during low-performing days by introducing discounts, loyalty programs, or special offers targeting weekdays.

4.13 How many orders were placed on the day with the highest revenue?

Interpretation: On 2015-11-27, 115 orders were placed, generating $4422.45 in revenue. This indicates a peak in customer demand on that day.

Recommendation:

a. Analyse the reason for the surge, like promotions or events, and replicate it on other days.
b. Plan inventory and staff better for high-demand days to meet customer expectations.
c. Leverage successful strategies to boost sales on similar days in the future**.**

4.14 Determine the top 3 most ordered pizza types based on revenue for each pizza category.

Interpretation: The best-performing pizza in each category (e.g., Thai Chicken Pizza in Chicken, Classic Deluxe Pizza in Classic) demonstrates that different customer segments prefer specific options.

Recommendation:

a. Create combo offers or promotions featuring the top performer from each category to encourage cross-category sales.
b. Introduce "Chef's Special" days where top pizzas from each category are offered at discounted prices

4.15 How does revenue vary by month?

Interpretation: January's negative revenue suggests a data issue, possibly due to refunds or incorrect reporting. July shows the highest revenue (72,557.9), likely due to seasonal factors or promotions. Revenue in September, October, and December is relatively lower, indicating potential seasonal slowdowns or missed

opportunities for promotions. The remaining months show a fairly consistent revenue range between 64,000 to 71,000.

Recommendations:

a. Fix Data Issues: Investigate and correct January's negative revenue.
b. Promotions for Low Months: Increase marketing efforts or offer discounts during September, October, and December to boost sales.
c. Replicate Success in High Months: Analyse and replicate the successful strategies from July (such as special offers or promotions) to increase revenue in other months.
d. Seasonal Offers: Introduce holiday or seasonal-themed offers to drive sales during traditionally slower months.

4.16 How does revenue vary by seasonality?

Interpretation: Spring and Summer have the highest sales, with Spring slightly ahead. These seasons seem to attract more customers. Winter and Fall have lower sales, with Winter being the slowest.

Recommendations:

a. Increase Winter Sales: Offer special promotions or discounts during Winter and Fall to attract more customers.
b. Leverage Spring and Summer Success: Use what worked in Spring and Summer to boost sales during other seasons, such as special deals or popular menu items.

5. Limitations of the Study

- **Data Incompleteness**: The dataset may contain missing or incomplete entries, which can affect the accuracy of some analysis results. For instance, missing order details or incomplete customer data could skew some findings.
- **Scope of Data**: The analysis was based on data from a single pizza business, limiting the generalization of the findings to other businesses or broader market trends.
- **Data Granularity**: Some factors, such as specific customer preferences or detailed transaction-level insights, could not be fully explored due to the limitations in the available dataset.
- **Seasonality Consideration**: While seasonal trends were observed, a deeper exploration of specific holidays, events, or marketing campaigns was not possible with the available data.

6. Conclusion

This project successfully analysed pizza sales data, uncovering key insights related to customer preferences, popular pizzas, and sales trends. Through the use of **MySQL** queries and **MS Excel** for data manipulation, the project identified the top-selling pizzas, peak sales hours, and revenue-generating trends. The analysis revealed that pizzas like **The Classic Deluxe Pizza** and **The Thai Chicken Pizza** are customer favourites, while peak sales times are observed during lunch and dinner hours. Recommendations have been provided to further optimize sales, such as promoting high-performing pizzas, offering combo deals, and focusing on time-based discounts. Overall, this analysis can help businesses enhance their sales strategies, improve inventory management, and increase customer satisfaction.

7. References

7.1 Kaggle Dataset:

- Source: Pizza Sales Dataset
- Description: The dataset was obtained from Kaggle and contains details of orders, pizzas, and sales transactions for analysis.

7.2 Tools Used:

- **MySQL**: For querying the dataset and extracting key insights.
- **MS Excel**: For data cleaning, organization, and basic analysis.
- **MS Word**: For preparing and documenting the final project report.

7.3 Documentation:

- MySQL Documentation: https://dev.mysql.com/doc/
- Microsoft Excel Help: https://support.microsoft.com/excel

7.4 Fly the Nest Materials:

- Source: "Fly the Nest" Data Science Course- SQL Materials

7.5 Additional Resources:

- Database Query Tutorials: W3Schools SQL Tutorial
- YouTube Tutorials: SQL & Excel Tutorials on YouTube