**A Training Report**

**On**

**“Pizza Sales Analysis”**

Submitted to the

**Rajasthan Technical University, Kota**

In Partial fulfillment of the requirement for the degree of

**MASTER OF COMPUTER APPICATIONS**

** **

**Name of Internal Guide Submitted by:**

Mr. Sunil Chauhan Rahul M Ramchandani

22CMSXX643

## Maharishi Arvind

### Institute Of Science And Management

**Affiliated to**

**Rajasthan Technical University, Kota**

**MCA (2022-24)**

Project Report

on

Pizza Sales Analysis

**~ A Data Analytics Project**

## MAHARISHI ARVIND

### INSTITUTE OF SCIENCE AND MANAGEMENT

#### Ambabari Circle, Ambabari, Jaipur – 302 039

Department Of Computer Science and Application

*Certificate*

This is to certify that the Industrial Project (MCA-451) work entitled “Pizza Sales Analysis” submitted by “***Rahul M Ramchandani”*** (22CMSXX643) to the Department Of Computer Science and Application of Maharishi Arvind Institute of Science and Management has been examined and evaluated.

Rahul M Ramchandani  **Internal Examiner**

Mr. Sunil Chauhan

**External Examiner Dr. Mahaveer Sain**

(HOD)

**Acknowledgement**

Undertaking a project is a huge task that can't be done alone. I want to express my gratitude to several people whose guidance and support made this project enjoyable. I am thankful to our Chairman, Dr. Bharat Parashar, and our Principal, Dr. DP Sharma, for providing excellent facilities and well-equipped labs.

A special thanks to our HOD, Dr. Mahaveer Sain, for consistently encouraging and advising me during my project work.

Meeting deadlines and receiving proper guidance are crucial for project success. I'd like to acknowledge the valuable contribution of the director of the company, and express my thanks to the team at DigiInk Solutions, including the trainer, for their support in a similar manner.

Additionally, I'd like to extend my thanks to the Department of Computer Science Engineering at MAISM, including both teaching and non-teaching staff, for their cooperation. A big shout-out to my parents for their continuous support. Lastly, I want to express my gratitude to my peers and friends.

#### **INDEX**

1. Introduction 11

2. System Study 14

2.1 System with limitations

2.2 Significance of the Project

2.3 Beneficiaries of the System

2.4 Feasibility study

3. System Analysis 17

Requirement Specification

i. Functional Requirement.

ii. Non Functional Requirement.

iii. User Requirement

iv. System Requirement

4. System Design 25

a) Data Flow Diagram

b) E-R Diagrams

c) Use Case Diagrams

d) Database Tables

5. Development 30

a) Environment

b) Coding Style

c) Coding Techniques

d) Coding

6. System Security 37

b. Checks and Control

c. Encryption, secure

7. Conclusion/Future Enhancement 40

8. Bibliography 49

#### **Figures INDEX**

Figure 1. DFD Diagram Level 0 26

Figure 1.1 DFD Diagram Level 1 27

Figure 2. ER Diagram 28

Figure 3 Use case diagram 29

Figure 4. Total Revenue 41

Figure 5 Daily Trend for orders 41

Figure 6 Monthly Trend for orders 41

Figure 7. Total sales by categories 42

Figure 8 Total sales by size 42

Figure 9. Top 5 Best Selling Pizzas 42

Figure 10 Bottom 5 pizza sales by revenue 43

Figure 11. Best Time for sales 43

Figure 12 Rush hour 43

Figure 13. Analysis Report 44

Figure 14 Order Analysis 44

Figure 15 pizza sales highest revenue 45

Figure 16 Pizza sales report Home 46

Figure 16.1 Pizza sales report Dashboard

#### **Tables INDEX**

Table 1 Orders 30

Table 2 Pizzas 30

Table 3 Order Details 30

**Brief Summary**

The "Pizza Sales Analysis" project aims to provide comprehensive insights into pizza sales to help businesses optimize operations and enhance customer satisfaction.

**Project Objectives**

1. **Understanding Customer Preferences**: The project focuses on identifying the most and least popular pizzas, as well as preferences for different pizza sizes and categories. This helps businesses tailor their offerings to meet customer demands more effectively.
2. **Optimizing Operations**: By analysing sales trends, the project seeks to improve operational efficiency, particularly in inventory management and staffing.
3. **Driving Business Growth**: Insights from the data analysis inform marketing strategies, product development, and overall business strategy, contributing to business growth.

**Methodology** The project utilizes a combination of SQL, Excel, and Power BI:

* **SQL**: Used for querying and managing data, enabling efficient data extraction and manipulation.
* **Excel**: Employed for data cleaning, basic analysis, and initial visualizations.
* **Power BI**: Used for advanced data visualization and deeper analysis, allowing the creation of interactive dashboards.

**Data Analytics Framework** Data analytics involves several key processes:

1. **Data Collection**: Gathering raw data from various sources, including point-of-sale (POS) systems.
2. **Data Cleaning**: Ensuring data accuracy and consistency by addressing errors and handling missing values.
3. **Data Analysis**: Applying statistical and computational techniques to identify patterns and trends in the data.
4. **Data Interpretation**: Drawing meaningful conclusions from the analysis.
5. **Data Visualization**: Presenting data in graphical formats for easy understanding.

**Key Areas of Analysis**

1. **Sales Performance**: The project analyzes total revenue, the number of pizzas sold, and average order amounts to assess overall sales performance.
2. **Product Popularity**: Identifying best-selling and least-selling pizzas helps in understanding customer preferences and managing inventory.
3. **Customer Preferences**: Analysis of pizza sizes and categories reveals which options are most popular, aiding in product development and promotional activities.
4. **Sales Trends**: Understanding daily and monthly sales trends helps businesses plan for peak periods and manage resources effectively.
5. **Operational Efficiency**: By analysing rush hour orders and other operational data, businesses can optimize staffing and resource allocation.

**System Analysis** The project involves a detailed system analysis, covering the requirements and design of the data analytics system:

* **Functional Requirements**: These include automated data capture, data storage, processing, and cleaning, as well as comprehensive sales analysis and reporting.
* **Non-Functional Requirements**: The system must be scalable, reliable, secure, and user-friendly, with features like real-time data access and customizable reports.
* **User Requirements**: The system should cater to the needs of various stakeholders, including business owners, marketing teams, inventory managers, and data analysts.

**Significance and Beneficiaries** The "Pizza Sales Analysis" project offers several benefits:

* **Informed Decision-Making**: Provides actionable insights into sales trends and customer preferences.
* **Operational Optimization**: Helps identify inefficiencies and optimize operations.
* **Enhanced Customer Satisfaction**: Allows businesses to better meet customer demands.
* **Competitive Advantage**: Enables businesses to leverage data-driven strategies.

**Future Enhancements** Future enhancements could include integrating more advanced analytics techniques like predictive modeling and machine learning, expanding the analysis to include more detailed customer demographics, and improving real-time data analysis capabilities.

**1. Introduction**

The "Pizza Sales Analysis" project is an industrial venture aimed at understanding and optimizing pizza sales through data analytics. Conducted as part of the Master of Computer Applications (MCA) program at the Maharishi Arvind Institute of Science & Management, Rajasthan Technical University Kota, this project seeks to provide valuable insights into pizza sales, helping businesses refine their offerings and optimize operations.

This project, titled "Pizza Sales - A Data Analytics Project," was undertaken to provide valuable insights to pizza businesses by analyzing their sales data. The primary objectives include:

**1.Understanding Customer Preferences**: Identifying the most and least popular pizzas, and understanding the preferences for different pizza sizes and categories.

**2.Optimizing Operations**: Analyzing sales trends to improve operational efficiency and better manage inventory and staffing.

**3.Driving Business Growth**: Using insights from data analysis to inform marketing strategies, product development, and overall business strategy.

The project employs a combination of SQL, Excel, and Power BI to handle, analyze, and visualize the pizza sales data:

• SQL: Used for querying and managing the data stored in databases, enabling efficient extraction and manipulation of data.

• Excel: Utilized for data cleaning, basic analysis, and initial visualizations, providing a flexible platform for data manipulation.

• Power BI: Leveraged for advanced data visualization and deeper analysis, allowing for the creation of interactive and insightful dashboards.

By integrating these tools, the project creates a powerful workflow that transforms raw sales data into actionable insights, ultimately helping pizza businesses to make informed decisions and enhance their competitive edge.

**Data Analytics-**

Data analytics involves examining raw data to uncover trends, patterns, and insights that can inform decision-making. It encompasses various techniques, including statistical analysis, data visualization, and predictive modeling, to transform data into actionable knowledge.

The realm of data analytics has revolutionized the way businesses operate, providing insights and driving decisions that optimize performance and boost profitability. This project explores the application of data analytics in the pizza sales industry, leveraging tools like SQL, Excel, and Power BI to gain a deep understanding of sales patterns, customer preferences, and operational efficiency.

The core components of data analytics include:

1. **Data Collection**: Gathering raw data from various sources.

2. **Data Cleaning**: Ensuring the data is accurate and free from errors.

3. **Data Analysis**: Applying statistical and computational techniques to identify patterns and trends.

4. **Data Interpretation**: Drawing meaningful conclusions from the data analysis.

5. **Data Visualization**: Presenting the data in a graphical or pictorial format to make the insights easily understandable.

Relation to Pizza Sales Analytics

In this project, data analytics is applied to pizza sales data to gain a comprehensive understanding of customer preferences and sales trends. By analysing metrics such as total revenue, number of pizzas sold, and average order amounts, we can identify best-selling pizzas, popular categories, and peak sales periods. These insights enable pizza businesses to make informed decisions about inventory, marketing strategies, and product offerings, ultimately enhancing customer satisfaction and driving business success.

The analysis of pizza sales data through data analytics provides invaluable insights that can drive the success of pizza businesses. From understanding customer preferences to optimizing operational efficiency, data-driven decision-making is key to thriving in the competitive food industry. This project demonstrates the potential of data analytics in transforming pizza sales data into a strategic asset, guiding businesses towards growth and success.

In the context of pizza sales, data analytics plays a crucial role in understanding and optimizing various aspects of the business. By analyzing sales data, pizza businesses can gain insights into customer preferences, operational efficiency, and market trends. This project delves into several key areas of pizza sales analysis:

1. **Sales Performance**: By examining total revenue, number of pizzas sold, and the average order value, businesses can assess their overall sales performance. This helps in identifying peak sales periods and understanding customer spending habits.

2. **Product Popularity**: Analyzing which pizzas are the best and worst performers in terms of orders and quantity sold helps businesses understand customer preferences. This can guide inventory management and marketing strategies.

3. **Customer Preferences**: Exploring sales by pizza size and category provides insights into what sizes and types of pizzas are most popular. This information is vital for product development and promotional activities.

4. **Sales Trends**: Identifying daily and monthly sales trends helps businesses understand customer ordering behavior. For instance, determining which days of the week or months of the year see the highest sales can inform staffing and supply chain decisions.

5. **Operational Efficiency**: Analyzing rush hour orders and understanding the reasons behind high or low sales during certain times can help optimize operations. This includes better scheduling of staff and more efficient use of resources.

**2. System Study**

### 2.1 System with Limitations

The current system for analysing pizza sales typically involves basic data recording and minimal analytical processing. Limitations include:

* **Manual Data Entry**: Susceptible to human errors and inconsistencies.
* **Limited Analytical Capability**: Basic tools like spreadsheets offer limited analytical depth.
* **Lack of Real-Time Insights**: Existing systems often do not provide real-time data analysis.
* **Inefficient Data Visualization**: Visual representations are often static and not interactive, limiting the ability to explore data dynamically.
* **Fragmented Data Sources**: Data might be scattered across various systems, making comprehensive analysis challenging.

### 2.2 Significance of the Project

The "Pizza Sales Analysis" project significantly enhances the ability to understand and leverage sales data through advanced data analytics. Key benefits include:

* **Informed Decision-Making**: Provides actionable insights into sales trends, customer preferences, and product performance.
* **Operational Optimization**: Helps identify inefficiencies and optimize inventory management, marketing strategies, and pricing models.
* **Enhanced Customer Satisfaction**: By understanding customer preferences, businesses can tailor their offerings to meet demand more effectively.
* **Competitive Advantage**: Enables businesses to stay ahead of competitors by leveraging data-driven strategies.

### 2.3 Beneficiaries of the System

The primary beneficiaries of the "Pizza Sales Analysis" system include:

* **Pizza Business Owners**: Gain insights into sales performance, customer preferences, and operational efficiencies.
* **Marketing Teams**: Obtain data to create targeted marketing campaigns and promotions.
* **Inventory Managers**: Use sales trends to optimize stock levels and reduce waste.
* **Customers**: Benefit from improved product offerings and service based on data-driven decisions.
* **Data Analysts**: Gain experience and insights from working with real-world data and advanced analytics tools.

### 2.4 Feasibility Study

The feasibility study evaluates the viability of the "Pizza Sales Analysis" project from various perspectives:

* **Technical Feasibility**:
  + **Data Availability**: Comprehensive pizza sales data is available and can be integrated into the analysis system.
  + **Tools and Technologies**: Power BI, SQL, and MS Excel are suitable and widely used tools for data analysis and visualization.
  + **Skills Required**: The project team has the necessary skills in data analytics, SQL querying, and using visualization tools.
* **Economic Feasibility**:
  + **Cost-Benefit Analysis**: The investment in data analytics tools and training is justified by the potential increase in revenue and operational efficiency.
  + **Budget Allocation**: Costs related to software licenses and potential infrastructure upgrades are manageable within the project budget.
* **Operational Feasibility**:
  + **User Training**: Adequate training will be provided to users to ensure they can effectively utilize the new system.
  + **Integration with Existing Processes**: The system can be integrated with existing sales processes without significant disruption.
* **Schedule Feasibility**:
  + **Project Timeline**: A detailed project plan ensures that the analysis can be completed within the academic timeline.
  + **Milestones and Deliverables**: Key milestones are defined to track progress and ensure timely completion of the project.

**3. System Analysis**

System analysis is a crucial phase in the development of any information system. It involves understanding the specific needs of the business, identifying problems, and proposing solutions to improve operations. The goal is to create a detailed blueprint that guides the development process, ensuring that the final system meets the intended objectives. For the pizza sales analytics project, this involves analyzing the current sales process, identifying key data points, and determining how the system will handle data collection, storage, analysis, and reporting.

**Requirement Specification-**Requirement specification is the process of documenting the system's requirements in a detailed and precise manner. This includes functional requirements that describe what the system should do, non-functional requirements that describe how the system should perform, user requirements that outline the needs of the end-users, and system requirements that specify the technical infrastructure needed to support the system.

**i. Functional Requirements**

#### Order Data Collection

**Order Data Collection** involves gathering all relevant data about each pizza order in a systematic and automated manner. This is a critical step to ensure the accuracy and completeness of the data that will be used for analysis.

1. **Automated Data Capture**:
   * The system should automatically capture data from the point-of-sale (POS) system whenever a pizza order is placed. This includes real-time data capture to ensure that no order is missed.
2. **Data Fields**:
   * The specific data fields to be captured include:
     + **Order ID**: A unique identifier for each order.
     + **Order Date and Time**: Timestamp of when the order was placed.
     + **Customer Details**: Information about the customer placing the order, including name and contact information.
     + **Pizza Details**: Information about each pizza in the order, including name, size, category (e.g., vegetarian, non-vegetarian), and quantity.
     + **Price Details**: Total price of the order, including any discounts or taxes applied.
3. **Integration with POS Systems**:
   * The system should seamlessly integrate with existing POS systems to facilitate data capture without requiring significant changes to current operations.
4. **Error Handling**:
   * Mechanisms should be in place to handle errors during data capture, such as network interruptions or system failures, ensuring that data integrity is maintained.

#### Data Storage

**Data Storage** refers to how the collected data is stored in a database, ensuring it is organized, secure, and easily retrievable for analysis.

1. **Database Design**:
   * The database schema should be designed to optimize data storage and retrieval. This includes creating normalized tables for orders, pizzas, customers, and sales to minimize redundancy and improve data integrity.
2. **Relational Database Management System (RDBMS)**:
   * The system should use a robust RDBMS (e.g., SQL Server, MySQL) capable of handling large volumes of data efficiently. The choice of RDBMS should also consider factors like scalability, security, and support for complex queries.
3. **Data Backup and Recovery**:
   * Implement regular data backup procedures to prevent data loss. The system should also have a data recovery plan in place to restore data in case of accidental deletion or corruption.
4. **Data Security**:
   * Ensure that the database is secured against unauthorized access. This includes using encryption, access controls, and regular security audits.

#### Data Processing and Cleaning

**Data Processing and Cleaning** involves preparing the raw data for analysis by correcting errors, handling missing values, and ensuring consistency.

1. **Data Validation**:
   * Implement validation checks to ensure that the data captured is accurate and consistent. This includes verifying that required fields are not empty, data types are correct, and values fall within expected ranges.
2. **Handling Missing Values**:
   * Develop strategies for dealing with missing data, such as imputation techniques (e.g., replacing missing values with the mean or median) or removing incomplete records if necessary.
3. **Data Transformation**:
   * Transform raw data into a format suitable for analysis. This includes aggregating data at different levels (e.g., daily, monthly sales) and creating derived metrics (e.g., average order value).
4. **Data Cleaning Tools**:
   * Use data cleaning tools and libraries (e.g., Python’s Pandas, SQL) to automate the cleaning process and ensure consistency.

#### Sales Analysis

**Sales Analysis** involves extracting meaningful insights from the data to understand sales performance and customer behavior.

1. **Total Revenue Calculation**:
   * Calculate total revenue over different periods (daily, weekly, monthly) to track overall sales performance.
2. **Sales Trends**:
   * Analyze sales trends over time to identify patterns and seasonal variations. This includes plotting sales data to visualize peaks and troughs.
3. **Top and Bottom Performers**:
   * Identify the top and bottom five pizzas in terms of the number of orders and quantity sold. This helps in understanding customer preferences and inventory management.
4. **Customer Segmentation**:
   * Segment customers based on their ordering behavior to tailor marketing strategies. This includes identifying high-value customers, frequent buyers, and those with specific preferences.
5. **Predictive Analysis**:
   * Use predictive analytics techniques to forecast future sales and identify potential growth opportunities.

#### Visualization and Reporting

**Visualization and Reporting** involves creating graphical representations of the data to make it easier to understand and communicate insights.

1. **Interactive Dashboards**:
   * Develop interactive dashboards using tools like Power BI or Tableau that allow users to explore data visually. Dashboards should include key metrics, trends, and comparisons.
2. **Custom Reports**:
   * Generate custom reports that can be tailored to the needs of different stakeholders. Reports should be easy to generate and export in various formats (e.g., PDF, Excel).
3. **Graphical Elements**:
   * Use various graphical elements such as bar charts, line graphs, pie charts, and heat maps to represent data visually. Each type of chart should be chosen based on the nature of the data and the insights it aims to convey.
4. **Real-Time Reporting**:
   * Implement real-time reporting capabilities to provide up-to-date information. This is particularly useful for monitoring current sales performance and making timely decisions.

#### User Interface

**User Interface (UI)** design focuses on creating an intuitive and user-friendly interface for interacting with the system.

1. **Ease of Navigation**:
   * The UI should be designed to ensure easy navigation, with clear menus and buttons that guide users through different functions.
2. **Responsive Design**:
   * Ensure that the UI is responsive and works well on various devices, including desktops, tablets, and smartphones.
3. **User Feedback**:
   * Implement features that provide immediate feedback to users, such as confirmation messages, error alerts, and progress indicators.
4. **Customization Options**:
   * Allow users to customize their interface based on their preferences. This includes saving custom views, setting up alerts, and personalizing dashboard layouts.
5. **Accessibility**:
   * Design the UI to be accessible to all users, including those with disabilities. This includes support for screen readers, keyboard navigation, and adjustable font sizes.

**ii. Non Functional Requirements**

Non-functional requirements define the system's operational attributes and constraints. These are critical for ensuring the system performs well under various conditions. For the pizza sales analytics project, non-functional requirements may include:

**Performance**:

* The system should handle large volumes of data efficiently. This includes fast data retrieval and processing times to ensure that reports and analyses are generated quickly.

**Scalability**:

* The system should be scalable to accommodate future growth in data volume and additional functionalities. This includes the ability to add new data sources or integrate with other systems as needed.

**Reliability**:

* The system should be reliable, with minimal downtime and high availability. This ensures that users can access the system whenever needed without interruptions.

**Usability**:

* The system should have a user-friendly interface that is easy to navigate. Users should be able to perform their tasks with minimal effort and training.

**Security**:

* The system should implement robust security measures to protect sensitive sales data. This includes access control, data encryption, and regular security audits.

**Maintainability**:

* The system should be easy to maintain and update. This includes clear documentation, modular design, and the ability to quickly address bugs or issues.

**iii. User Requirement**

User requirements are essential for designing a system that meets the needs and expectations of its end-users. For the pizza sales analytics project, understanding what users need from the system is critical for its success. Here, we'll delve deeper into the specific needs and expectations of the different user groups, their interactions with the system, and how these requirements translate into practical system functionalities.

1. **Sales Insights**
   * **Comprehensive Sales Reports**: Users need detailed reports that provide an overview of sales performance. This includes total revenue, number of orders, and quantities sold for different time periods (daily, weekly, monthly, yearly).
   * **Trend Analysis**: Users require tools to identify sales trends over time. This includes understanding seasonal fluctuations, peak sales periods, and long-term growth patterns.
   * **Product Performance**: Users need to know which products are performing well and which are not. This includes identifying the top-selling and least-selling pizzas by name, category, and size.
2. **Customer Preferences**
   * **Preference Analysis**: Users need insights into customer preferences, such as the most popular pizza sizes and categories. This helps in tailoring marketing strategies and menu offerings.
   * **Demographic Insights**: Users require information on customer demographics, such as age, gender, and location, to better understand their customer base and target marketing efforts effectively.
3. **Operational Efficiency**
   * **Peak Sales Times**: Users need to identify peak sales times to optimize staffing and inventory levels. This includes understanding hourly, daily, and weekly sales patterns.
   * **Order Processing**: Users need to streamline order processing to reduce wait times and improve customer satisfaction. This includes insights into average order processing times and bottlenecks.
   * **Inventory Management**: Users require data on inventory levels and turnover rates to ensure that popular items are always in stock and to minimize waste.
4. **Customizable Reports**
   * **Filter and Drill-Down**: Users need the ability to filter data and drill down into specific details. This includes filtering by date range, product category, size, and other relevant attributes.
   * **Custom Visualization**: Users should be able to create custom visualizations to highlight specific data points. This includes charts, graphs, and tables that can be tailored to individual needs.
   * **Export and Share**: Users need to export reports in various formats (PDF, Excel, CSV) and share them with stakeholders. This facilitates collaboration and decision-making.
5. **User-Friendly Interface**
   * **Intuitive Design**: The system should have an intuitive user interface that is easy to navigate. Users should be able to access the information they need quickly and without extensive training.
   * **Responsive Design**: The system should be responsive, allowing users to access it from different devices (desktop, tablet, mobile) and screen sizes.
   * **Interactive Dashboards**: Users need interactive dashboards that allow them to interact with data in real-time. This includes the ability to click on data points to view more details or to switch between different views.
6. **Real-Time Data Access**
   * **Live Data Updates**: Users require real-time access to sales data to make timely decisions. This includes automatic updates to reports and dashboards as new data is collected.
   * **Notifications and Alerts**: Users need notifications and alerts for critical events, such as stockouts, peak sales times, and unusual sales patterns.

**iv. System Requirement**

System requirements specify the technical infrastructure needed to support the system. These requirements ensure that the system operates efficiently, securely, and reliably. For the pizza sales analytics project, system requirements encompass hardware, software, networking, security, and development tools.

1. **Hardware Requirements**
   * **Server Specifications**: The server should have sufficient processing power (CPU), memory (RAM), and storage capacity (hard disk or SSD) to handle large volumes of sales data. This ensures fast data processing and retrieval.
   * **Backup and Redundancy**: Reliable backup solutions, such as external hard drives or cloud storage, to ensure data integrity and availability in case of hardware failure. Redundancy measures, such as RAID configurations, to protect against data loss.
   * **Networking Hardware**: Network switches, routers, and firewalls to ensure stable   
     and secure network connectivity. This includes high-speed internet connections for data transmission and remote access.
2. **Software Requirements**
   * **Database Management System (DBMS)**: A robust relational database management system (e.g., SQL Server, MySQL, PostgreSQL) to store and manage sales data. The DBMS should support efficient querying, indexing, and data integrity constraints.
   * **Data Analysis and Visualization Tools**: Software tools for data analysis and visualization, such as Power BI and Excel. These tools should integrate seamlessly with the database and provide advanced analytics capabilities.
   * **Operating System**: The server should run on a reliable operating system (e.g., Windows Server, Linux) that supports the required software and provides necessary security features.
   * **Middleware and APIs**: Middleware software and APIs to facilitate data integration and communication between different components of the system. This includes ETL (Extract, Transform, Load) tools for data processing.
3. **Networking Requirements**
   * **Internet Connectivity**: Reliable internet connectivity to enable data collection from various sources, remote access to the system, and cloud-based services. This includes high-speed broadband connections and backup internet links.
   * **Network Security**: Network security measures, such as firewalls, VPNs (Virtual Private Networks), and intrusion detection systems (IDS), to protect the system from unauthorized access and cyber threats. This includes secure access protocols (e.g., HTTPS) for data transmission.
4. **Development Tools**
   * **Programming Languages**: Knowledge of programming languages such as Python and SQL for data processing, analysis, and system development. Python libraries (e.g., pandas, NumPy) for data manipulation and SQL for database interactions.
   * **Version Control Systems**: Version control systems (e.g., Git) to manage code changes, collaborate with team members, and maintain a history of code modifications. This ensures efficient development workflows and code integrity.
   * **Integrated Development Environments (IDEs)**: IDEs (e.g., Visual Studio Code, PyCharm) to facilitate code development, debugging, and testing. These tools should support the required programming languages and libraries.
5. **Security Requirements**:
   * **Data Encryption**: Encryption of sensitive data to protect it from unauthorized access.
   * **Access Control**: Implementation of access control mechanisms to restrict access to authorized users only.
   * **Regular Audits**: Regular security audits to identify and address potential vulnerabilities.
6. **System Maintenance and Support:**

* **Documentation**: Comprehensive documentation for system architecture, data models, user guides, and maintenance procedures. This ensures that the system can be maintained and updated efficiently.
* **Technical Support**: Availability of technical support to address system issues, perform updates, and provide assistance to users. This includes helpdesk services, remote support, and regular system monitoring.
* **Disaster Recovery**: Disaster recovery plans to ensure business continuity in case of system failures, natural disasters, or other emergencies. This includes regular backups, data recovery procedures, and contingency plans.

By thoroughly addressing these user and system requirements, the pizza sales analytics project aims to create a robust, efficient, and user-friendly system that provides valuable insights into sales performance, customer preferences, and operational efficiency. This will enable stakeholders to make informed decisions and optimize business operations effectively.

**4. System Design**

### A) Data Flow Diagram (DFD)

**Level 0 DFD (Context Diagram)**

****

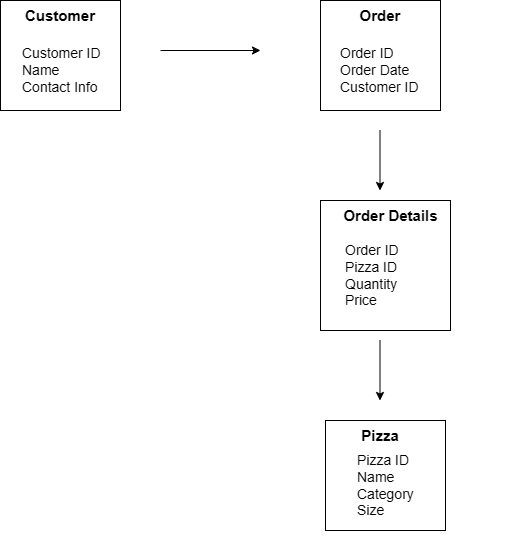
**Fig 1 Level 0 DFD**

**Level 1 DFD**

### 

**Fig 1.1 Level 1 DFD**

### B) E-R Diagram (Entity-Relationship Diagram)



**Fig 2 ER Diagram**

### C) Use Case Diagrams



**Fig 3 Use Case Diagram**

### D) Database Tables

Here are the primary database tables used in the system:

**Table 1- Orders**:

|  |  |  |
| --- | --- | --- |
| Column | Data Type | Description |
| OrderID | INT | Primary Key |
| OrderDate | DATE | Date of the order |
| CustomerID | INT | Foreign Key (Customers) |

**Table 2- Pizzas**:

|  |  |  |
| --- | --- | --- |
| Column | Data Type | Description |
| PizzaID | INT | Primary Key |
| Name | VARCHAR | Pizza name |
| Category | VARCHAR | Pizza category |
| Size | VARCHAR | Pizza size |

**Table 3- OrderDetails**:

|  |  |  |
| --- | --- | --- |
| Column | Data Type | Description |
| OrderID | INT | Foreign Key (Orders) |
| PizzaID | INT | Foreign Key (Pizzas) |
| Quantity | INT | Number of pizzas |
| Price | DECIMAL | Price of the order |

### 

**5. Development**

### a) Environment

**a) Software and Tools**

* **SQL**: SQL was used to manage and manipulate the pizza sales data. Created queries to extract relevant information, perform data cleaning, and prepare the data for analysis. SQL was also essential for creating complex joins, aggregating data by time periods (such as monthly or daily sales), and filtering data by categories like pizza type, sales region, or customer segments.
* **Excel**: Excel helped in exploring the data and performing initial analysis, including quick visualizations, pivot tables, and statistical calculations. It likely served as a flexible tool to validate SQL results and perform preliminary trend analysis before building a more robust visualization in Power BI. Excel might have been used for additional data processing steps, such as merging, formatting, and managing any data exports or imports.
* **Power BI**: Power BI enabled was used to create interactive dashboards and visualizations to display insights from the pizza sales data. Using Power BI, there was development of charts, graphs, and custom visuals to highlight key performance indicators (KPIs) such as total sales, top-selling pizzas, peak sales times, and customer trends. The tool provided a user-friendly interface for stakeholders to interact with and explore the data more deeply.

**b) Operating System**

* Windows 10 operating system was used in the project.

### b) Coding Style

**a) Consistency**

* **Naming Conventions**: Describe the conventions followed for naming tables, columns, variables, and files. For example, using snake\_case for SQL and camelCase for other scripts.
* **Indentation and Spacing**: Highlight the standard practices followed for readability, such as using 2 or 4 spaces for indentation.
* **Commenting**: Emphasize the importance of commenting on complex SQL queries or scripts, explaining the logic and purpose of different sections.

**b) Documentation**

* Discuss any internal documentation created, such as README files, documentation of data models, or comments within code that help explain the functionality and logic.

c) Coding Techniques

**a) Data Cleaning and Preprocessing**

* **Handling Missing Values**: Techniques like imputation, omission, or using default values for missing data.
* **Data Transformation**: used processes like normalization, aggregation, or filtering to prepare the data for analysis.
* **Data Integration**: data integration was done properly

**b) SQL Query Optimization**

* **Indexing**: some indexing strategies were used to speed up query performance.
* **Query Optimization**: Techniques such as avoiding SELECT \*, using WHERE clauses efficiently, and breaking down complex queries.

**c) Data Visualization**

* **Design Principles**: principles like clarity, simplicity, and accuracy were followed in creating visualizations and were well taken care of.
* **Chart Types**: multiple types of charts and visualizations used (e.g., bar charts, pie charts, scatter plots) were used.

d) Coding

SQL Queries-

use pizza;

Select \* from dPizza\_data;

-- Total revenue

SELECT SUM(total\_price) As Total\_revenue FROM pizza\_data;

--- Total Pizza sold

SELECT SUM(quantity) AS Total\_pizza\_sold FROM pizza\_data;

--Total orders

SELECT COUNT(DISTINCT order\_id) AS Total\_Orders FROM pizza\_data

-- Average Order

SELECT (SUM(total\_price) / COUNT(DISTINCT order\_id)) AS Avg\_order\_amt FROM pizza\_data;

--- Top 5 Pizza name by Order

SELECT Top 5 pizza\_name, COUNT(DISTINCT order\_id) AS Total\_Orders

FROM Pizza\_data

GROUP BY pizza\_name

ORDER BY Total\_Orders DESC;

-- Bottom 5 Pizza by Order

SELECT Top 5 pizza\_name, COUNT(DISTINCT order\_id) AS Total\_Orders

FROM Pizza\_data

GROUP BY pizza\_name

ORDER BY Total\_Orders ASC;

-- Top 5 Pizza by Quantity sold

SELECT Top 5 pizza\_name, sum(quantity) AS Total\_no\_order

FROM Pizza\_data

GROUP BY pizza\_name

ORDER BY Total\_no\_order DESC;

-- Bottom 5 pizza by Quantity sold

SELECT Top 5 pizza\_name, SUM(quantity) AS Total\_no\_order

FROM Pizza\_data

GROUP BY pizza\_name

ORDER BY Total\_no\_order Asc;

-- Total Amount Sold by Sizes

SELECT pizza\_name, pizza\_size, COUNT(pizza\_size) AS Amt\_sold

FROM Pizza\_data

GROUP BY pizza\_name, pizza\_size

ORDER BY Amt\_sold Desc;

--Total Amount sold by Category

SELECT pizza\_name, pizza\_category, count(pizza\_category) AS Amt\_sold

FROM Pizza\_data

GROUP BY pizza\_name, pizza\_category

ORDER BY Amt\_sold Desc;

--Trend Pizza by the Day

SELECT DATENAME(DW, order\_date) AS order\_day, COUNT(DISTINCT order\_id) AS total\_orders

FROM pizza\_data

GROUP BY DATENAME(DW, order\_date)

ORDER BY total\_orders Desc;

--Trend Pizza by month

SELECT DATENAME(MONTH, order\_date) AS Month, COUNT(DISTINCT order\_id) AS total\_orders

FROM pizza\_data

GROUP BY DATENAME(MONTH, order\_date)

ORDER BY total\_orders Desc;

--% Sales by Pizza Category

SELECT pizza\_category, CAST(SUM(total\_price) AS DECIMAL(10,2)) as total\_revenue,

CAST(SUM(total\_price) \* 100 / (SELECT SUM(total\_price) from pizza\_data) AS DECIMAL(10,2)) AS Pct

FROM Pizza\_data

GROUP BY pizza\_category;

--% Sales by Pizza size

SELECT pizza\_size, CAST(SUM(total\_price) AS DECIMAL(10,2)) as total\_revenue,

CAST(SUM(total\_price) \* 100 / (SELECT SUM(total\_price) from Pizza\_data) AS DECIMAL(10,2)) AS Pct

FROM Pizza\_data

GROUP BY pizza\_size

ORDER BY pizza\_size;

**6. System Security**

### Checks and Control

Implementing checks and controls is crucial to ensure the security, integrity, and reliability of the system. Here are the key checks and controls for the "Pizza Sales Analysis" project:

### Data Integrity

**Data integrity** refers to the accuracy, consistency, and reliability of data throughout its lifecycle. In the context of the "Pizza Sales Analysis" project, data integrity ensures that all information related to pizza sales, such as order details, customer information, and sales metrics, remains accurate, consistent, and trustworthy from the moment it is entered into the system until it is analysed and reported.

#### Key Aspects of Data Integrity:

1. **Accuracy**:
   * Ensures that the data reflects the true values of the entities it represents. For example, the price of each pizza should match the actual selling price, and the quantities should reflect the actual number of pizzas sold.
2. **Consistency**:
   * Ensures that data remains consistent across different parts of the database and over time. For instance, if an order is recorded in the system, all related details (such as customer information, order date, and order amount) must be consistent across all records and tables.
3. **Completeness**:
   * Ensures that all required data is present and fully captured. In the context of pizza sales, every order entry should include all necessary fields like CustomerID, OrderDate, PizzaID, Quantity, and Price.
4. **Timeliness**:
   * Ensures that data is recorded and updated promptly. For example, sales data should be updated in real-time or at regular intervals to reflect the current status of orders and inventory accurately.
5. **Uniqueness**:
   * Ensures that each data entry is unique and not duplicated. For instance, each order should have a unique OrderID, and each pizza should have a unique PizzaID.
6. **Validity**:
   * Ensures that data conforms to defined formats and rules. For example, dates should be in the correct format, and quantities should be non-negative integers.

### Data Validation

**Data validation** is the process of ensuring that the data entered into the system meets predefined criteria and rules, thereby maintaining data integrity. In the "Pizza Sales Analysis" project, data validation helps prevent errors and inconsistencies during data entry and processing.

#### Key Components of Data Validation:

1. **Input Validation**:
   * **Format Checks**: Ensures that data is entered in the correct format. For instance, order dates should be in a valid date format (e.g., YYYY-MM-DD).
   * **Range Checks**: Ensures that numerical values fall within a specified range. For example, the quantity of pizzas ordered should be a positive integer.
   * **Type Checks**: Ensures that the data type of each input is correct. For instance, the price should be a decimal value, and the CustomerID should be an integer.
2. **Consistency Checks**:
   * **Cross-field Validation**: Ensures that related fields are consistent with each other. For example, the total price in an order should equal the sum of the prices of individual pizzas multiplied by their quantities.
   * **Referential Integrity**: Ensures that relationships between tables are maintained. For example, every PizzaID in the OrderDetails table should have a corresponding entry in the Pizzas table.
3. **Completeness Checks**:
   * Ensures that all mandatory fields are filled in. For example, an order entry should not be saved if the CustomerID, OrderDate, PizzaID, Quantity, or Price fields are empty.
4. **Uniqueness Checks**:
   * Ensures that there are no duplicate entries. For instance, each OrderID should be unique within the Orders table.
5. **Business Rule Validation**:
   * Ensures that data complies with business rules. For example, a special discount might only apply to orders above a certain amount, and this rule should be enforced during data entry.

### Relation to the "Pizza Sales Analysis" Project

In the "Pizza Sales Analysis" project, maintaining data integrity and performing data validation are crucial for accurate analysis and reporting. Here's how they are applied:

* **Order Entry**: When entering new orders, the system validates that all required fields are filled, values are within acceptable ranges, and data types are correct. This prevents invalid data from entering the database.
* **Data Processing**: As data is processed for analysis, consistency checks ensure that relationships between tables (such as Orders and OrderDetails) are maintained.
* **Reporting**: Accurate and consistent data allows for reliable reporting on sales trends, revenue, and customer preferences. Validation ensures that the reports generated reflect true business performance.

By implementing robust data integrity measures and thorough data validation processes, the "Pizza Sales Analysis" project ensures the accuracy and reliability of the insights derived from the sales data, ultimately aiding in better decision-making and business optimization.

**7. Conclusion/Insights**

**Useful Insights –**

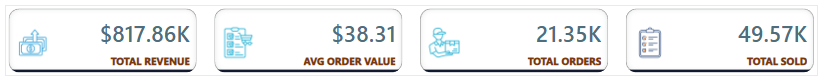


Figure 4 Total Revenue

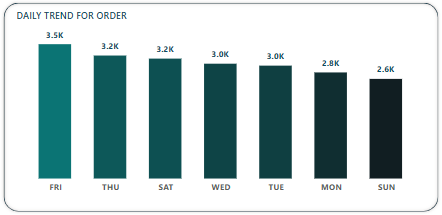
****

Figure 5 Daily Trends for Order

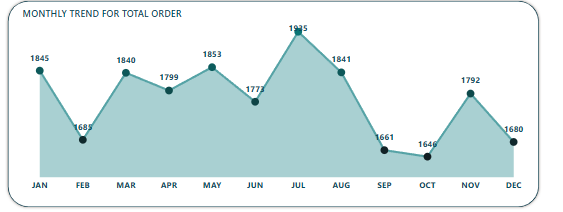
****

Figure 6 Monthly Trends for Order

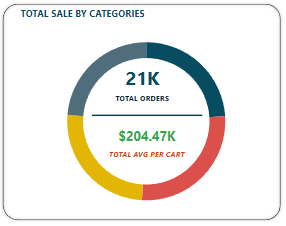


Figure 7 Total Sales by categories

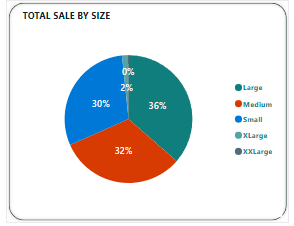


Figure 8 Total Sales by size

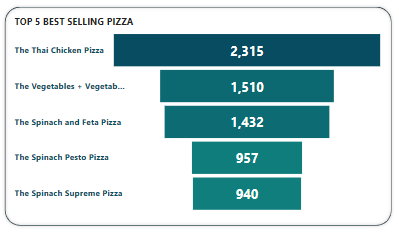


Figure 9 Top 5 Best selling

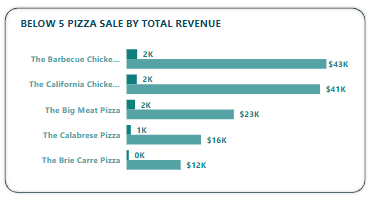


Figure 10 Bottom 5 Pizza Sales by Revenue

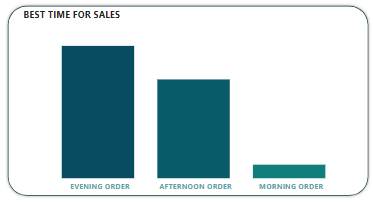


Figure 11 Best Time for Sales

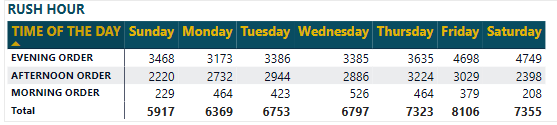


Figure 12 Rush Hour

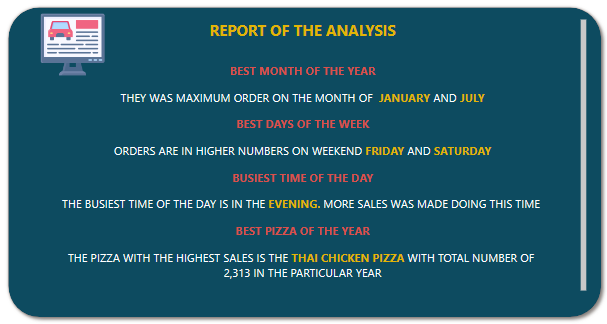


Figure 13 Analysis Report

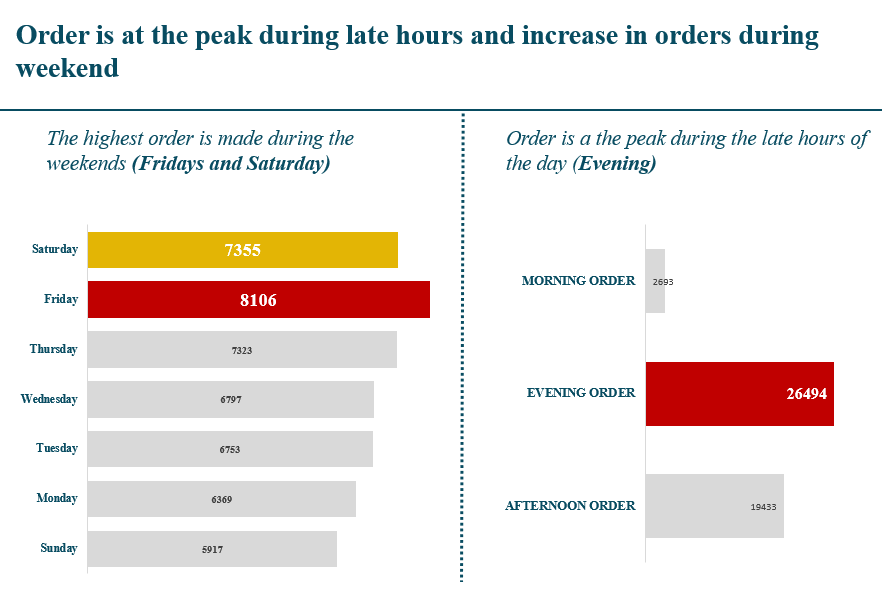


Figure 14 Order Analysis

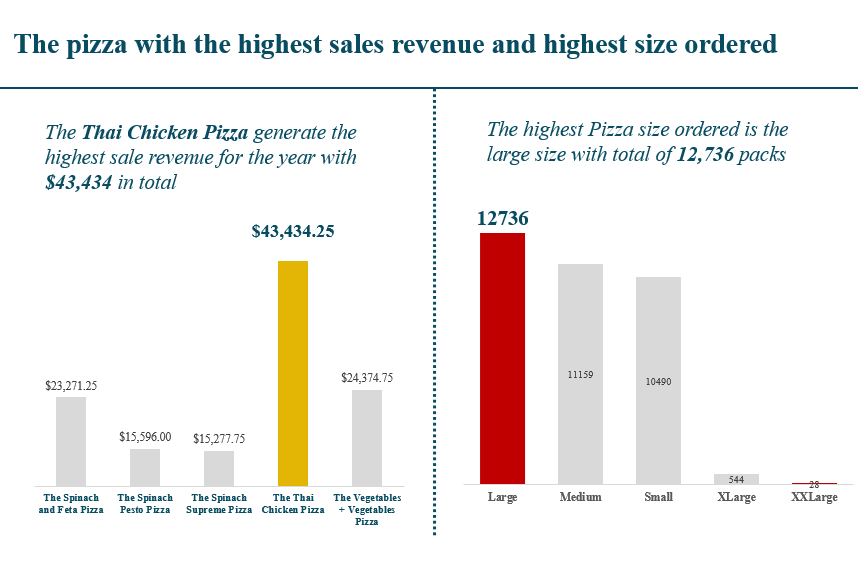


Figure 15 Pizza Sales Highest Revenue

**Power BI Dashboard –**

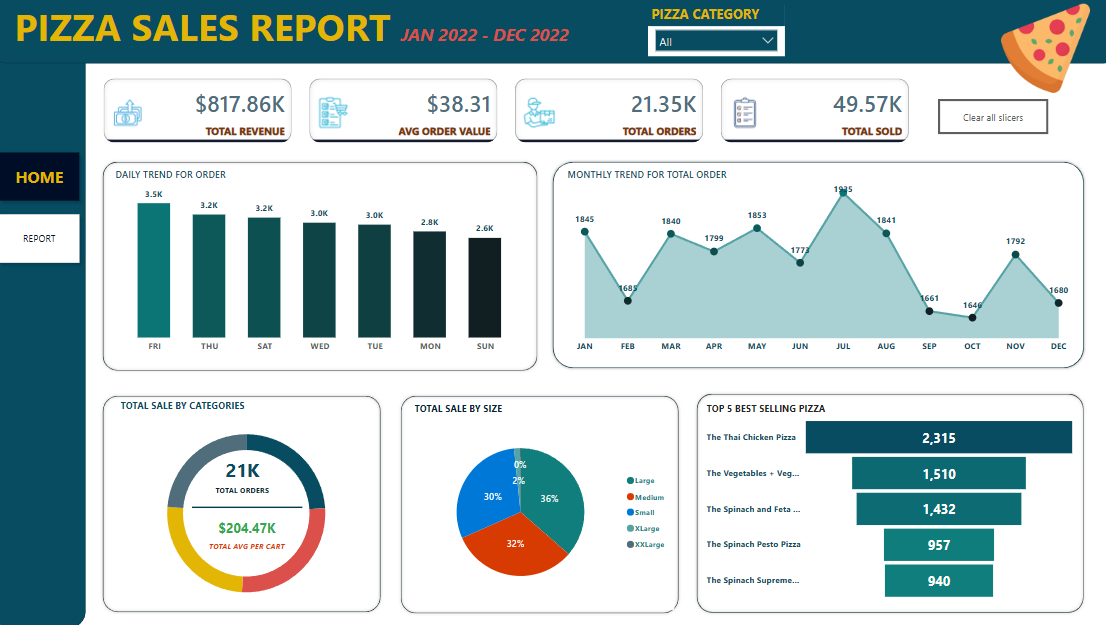


Figure 16 Pizza Sales Report Dashboard

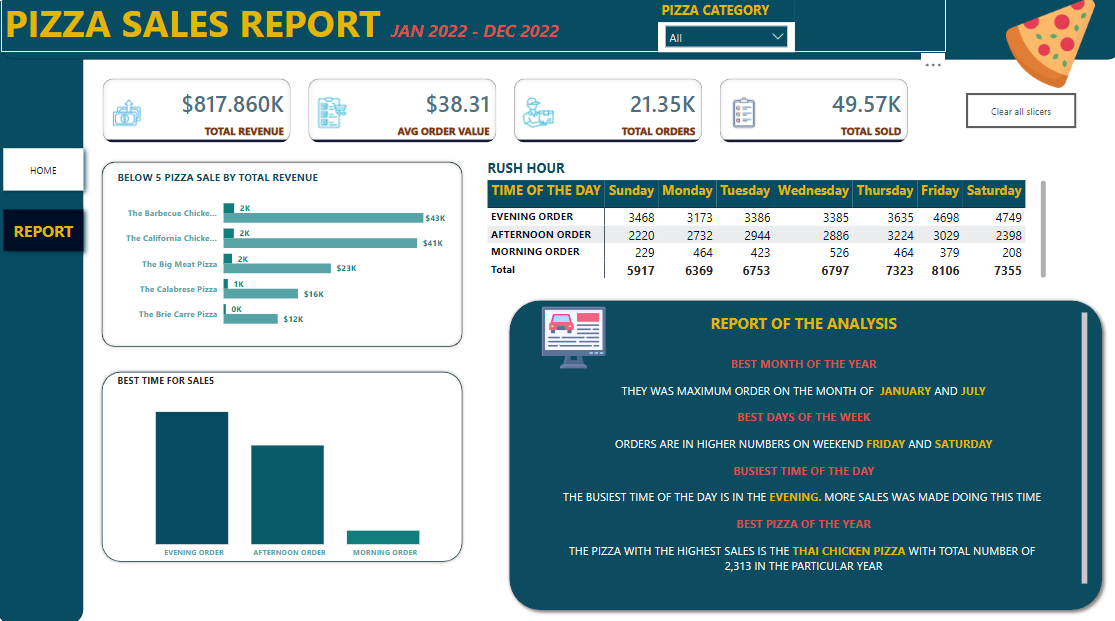


Figure 16.1 Pizza Sales Report Dashboard

**Additional Insights-**

1. They have sold a total of 50,000. pizzas. This data can guide inventory and production planning.
2. The average order amount is $38, indicating the customers’ spending habits. This could inform pricing and promotions.
3. The Classic Deluxe Pizza are customer favorites, and they should consider marketing them more prominently.
4. The Classic Deluxe Pizza are top performers in terms of quantity, showing strong demand.
5. Large Size and Classic dominate sales, suggesting they should prioritize these sizes and categories.

**Monthly Trend Analysis**

July has the highest number of orders.The data shows that July has the highest number of orders with a total of 1,935 orders. To understand why July has the maximum number of orders, these are the following reasons:

* Weather Influence: July is often characterized by warm weather, which can lead to more outdoor activities, gatherings, and parties where pizza is a popular food choice.
* School Holidays: Many schools have summer vacations in July, leading to increased family activities and higher demand for convenient food options like pizza.
* Tourism Season: July is a peak travel month, with more tourists and visitors who might prefer ordering in rather than cooking, boosting pizza sales.
* Marketing Campaigns: Companies may strategically plan major marketing campaigns and promotions during the summer to attract more customers, including new product launches or limited-time offers.
* Local Festivals and Fairs: Local events, festivals, and fairs that occur during July can increase the number of orders as attendees look for convenient meal options.

**Daily Trend Analysis**

The data shows that Sundays have the lowest number of orders with a total of 2,624 orders. The reasons why Sundays have the lowest orders and other weekdays have more orders are:

* Workday Convenience: During weekdays, people may have less time to cook due to work commitments, leading to a higher reliance on food delivery services.
* Office Orders: Many offices and workplaces order pizza for lunch meetings, team gatherings, or Friday treats, contributing to higher weekday sales.
* Friday Socializing: Friday often marks the beginning of the weekend, and people may celebrate by ordering pizza, leading to a spike in orders.
* Religious Observances: For some people, Sunday is a day of religious observance and rest, which might result in fewer commercial activities, including ordering food.
* Family Dinners: Families might prefer cooking and eating together on Sundays, seeing it as a day for home-cooked meals and family bonding time.
* Meal Planning: By Sunday, individuals may be preparing for the week ahead, engaging in meal prepping or grocery shopping, which can lower the demand for takeout.

**8. Bibliography**

* Microsoft Power BI Documentation  
  <https://docs.microsoft.com/power-bi>  
  This is the official documentation site for Power BI, offering guides, tutorials, and best practices.
* W3Schools SQL Tutorial  
  https://www.w3schools.com/sql/  
  W3Schools provides a comprehensive SQL tutorial with examples and exercises for learning SQL.
* Excel Easy  
  <https://www.excel-easy.com/>  
  This site offers tutorials and tips for various Excel functions and data analysis techniques.
* Google
* Power BI Community  
  https://community.powerbi.com/  
  The Power BI Community site includes forums, blogs, and resources where users share tips and solutions.