PIZZA SALES REPORT

Project Objective:

To analyze the pizza sales revenue, identify trends & patterns and provide insights to improve the pizza sales & business.

Tools used:

Tableau Desktop- Version 2024.3

MySQL Workbench-8.0 C.E

MS Excel- Version 2019

Steps involved:

- Data Collection
- Data Cleaning
- Data import to Tableau Desktop
- KPI CARD
- Data Visualization-charts required
- Report/Dashboard
- Data import on MySQL
- Data analysis using MySQL

Data Collection:

We have a dataset named Data Model-pizza_sales.xlsx with columns like order_details_id, order_id, pizza_id, quantity, unit_price, total_price, pizza_name, pizza_size, order_date, order_time, etc.

Data Cleaning:

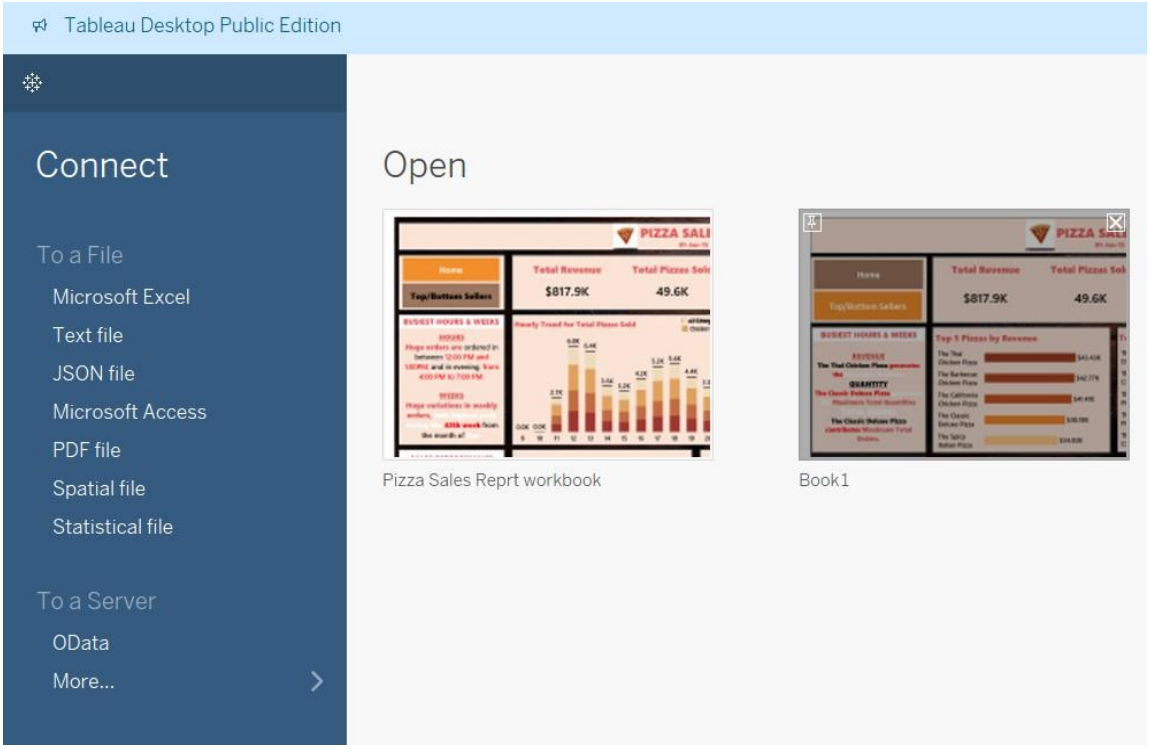
Use Excel or SQL tool to clean the data and clean the data to remove any errors.

Data Import to Tableau:

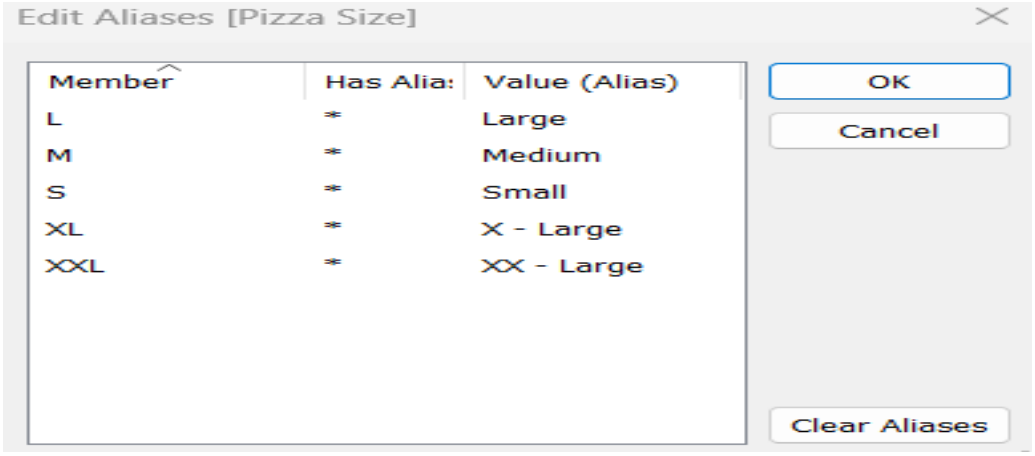
Open Tableau and connect your data source.

Load your prepared dataset into Tableau.

Tableau > Connect > Text File > Select the CSV File from the dataset saved folder.



- After the data load, again we must check the data types and data errors in the dataset for the better data visualization.
- Some Data Cleaning Happened in the dataset. We used alias in pizza size column, we have to convert L- Large, M-Medium, S-Small, XL-X-large, and XXL-XX-Large.



KPI CARD:

We want to generate the key performance indicators card to analyze the pizza sales data and gain insights . The main KPI'S for this pizza sales data are follows:

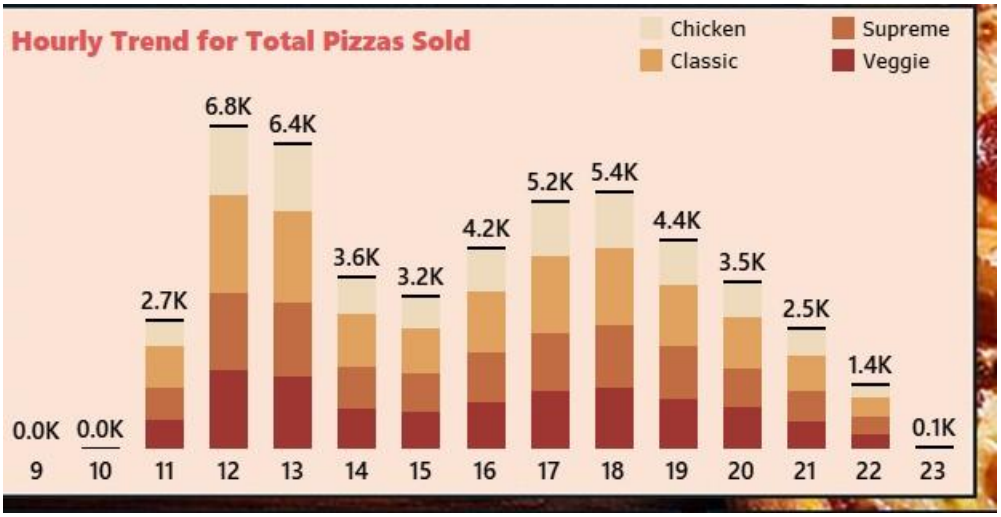
- **Total Revenue**: The sum of total price of all pizza orders.
- **Total Orders**: The total no of orders placed.
- **Total Pizza Sold**: The sum of the quantities of all pizzas sold.
- **Average Order Value**: The average amount spent per order, calculated by dividing the total revenue by the total number of orders.
- **Average Pizzas Per Order**: The average number of pizzas sold per order, calculated by dividing the total number of pizzas sold by the total number of orders

Total Revenue	Total Pizzas Sold	Total Orders	Avg Order Value	Avg Pizzas Per Order
$817.9K	49.6K	21.4K	$38.3	2.32

Data Visualization-Charts Required:

1. **Stacked Bar Chart – Hourly Trend for Total Pizzas Sold**

Created a stacked bar chart that shows the hourly trend of total orders over a particular period of time.

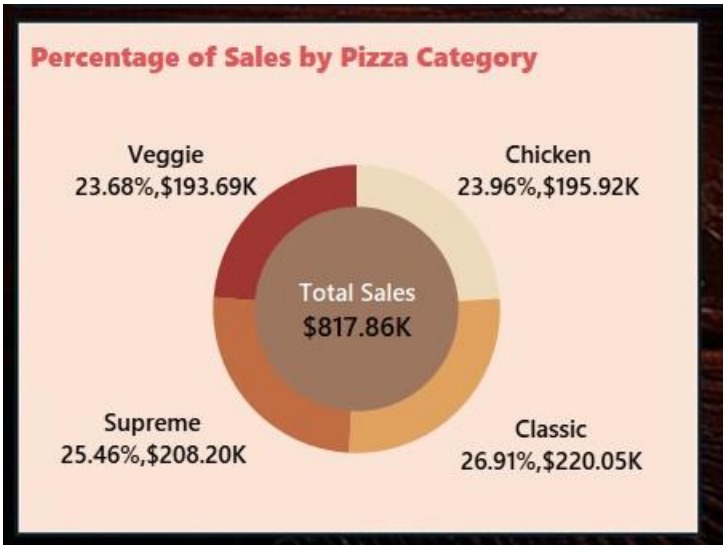


2. Line Chart – Weekly Trend for Total Orders

Created a line chart that displays the weekly trend for total orders throughout the year.

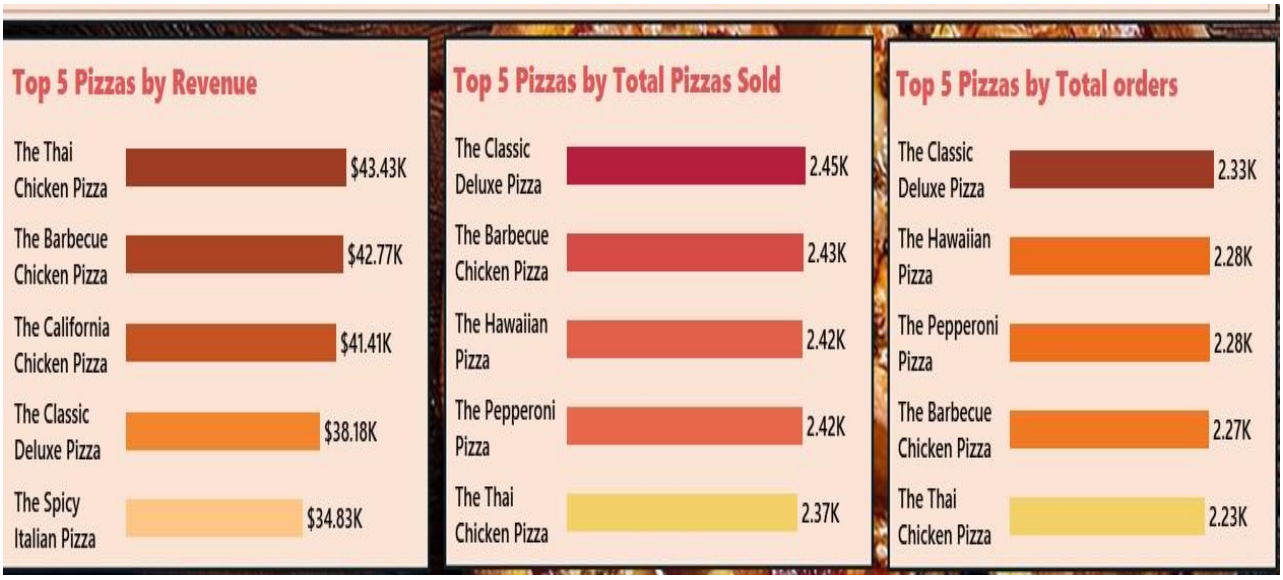
3. Pie Chart – Percentage of Sales by Pizza Category

Created a pie chart and this chart will give you insights into the popularity of various pizza categories and their contribution to overall sales.



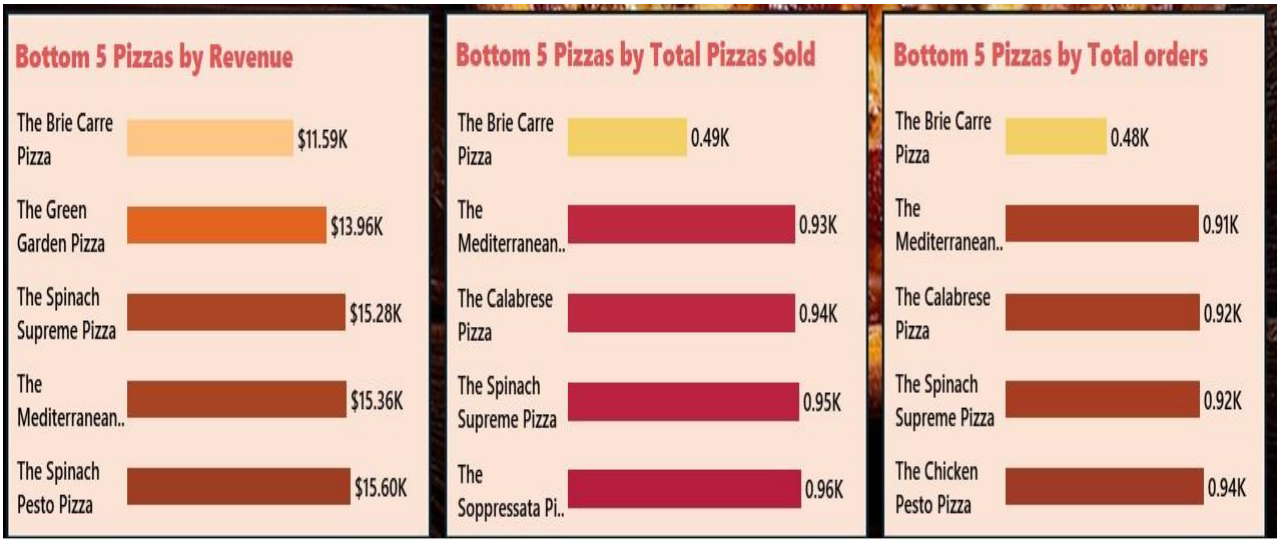
4. **Bar Chart – Top 5 Best Sellers by Revenue, Total Quantity, and Total Orders**

Created a bar chart highlighting the top 5 best-selling pizzas based on the revenue, total quantity and total orders. This chart will allow us to identify the most popular pizzas.



5. **Bar chart – Bottom 5 Best Sellers by Revenue, Total Quantity and Total Orders**

Created a bar chart highlighting the bottom 5 best-selling pizzas based on the revenue, total quantity and total orders. This chart will allow us to identify the less popular pizzas.



6. Some other charts used

Report/Dashboard:

Combine Visualizations into a Dashboard.

Arrange the charts in a sequence order.

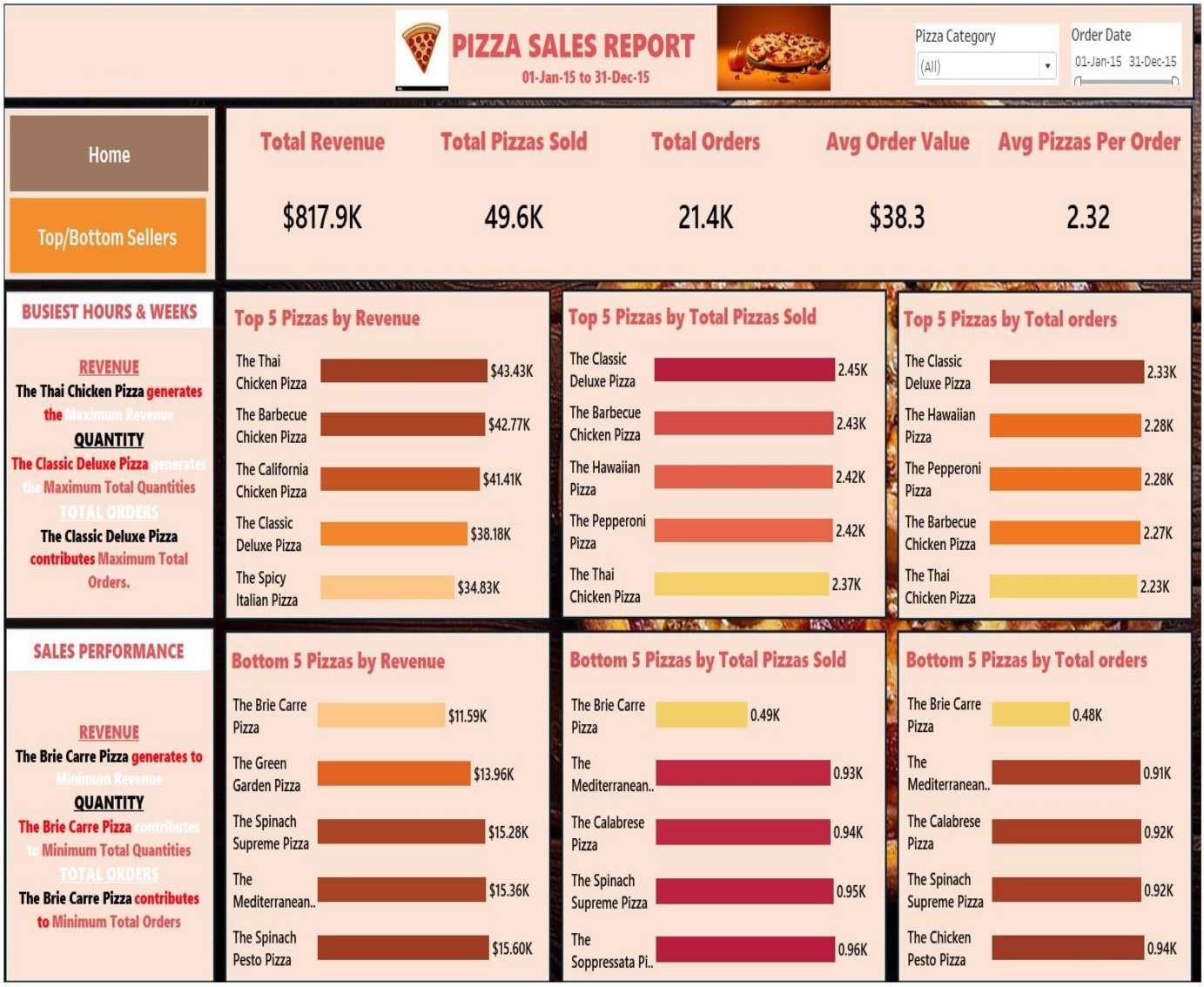
Add filters and interactivity to allow users to explore the data.

Add text boxes to provide context and insights on key findings

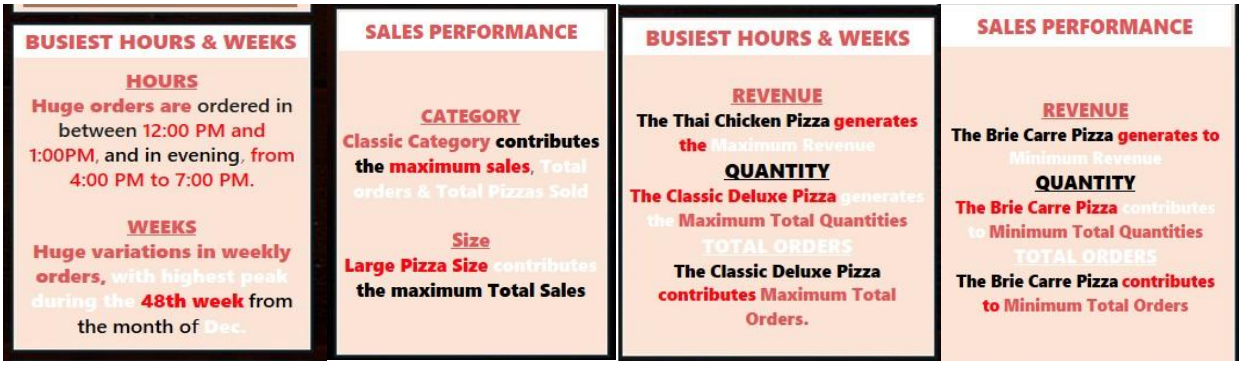
Home Page

PIZZA SALES REPORT
01-Jan-15 to 31-Dec-15
Pizza Category
(All)
Order Date
01-Jan-15 31-Dec-15
Home
Top/Bottom Sellers
Total Revenue
$817.9K
Total Pizzas Sold
49.6K
Total Orders
21.4K
Avg Order Value
$38.3
Avg Pizzas Per Order
2.32
BUSIEST HOURS & WEEKS
HOURS
Huge orders are ordered in between 12:00 PM and 1:00PM, and in evening, from 4:00 PM to 7:00 PM.
WEEKS
Huge variations in weekly orders, with highest peak during the 48th week from the month of Dec.
Hourly Trend for Total Pizzas Sold
Chicken
Classic
Supreme
Veggie
0.0K 0.0K 2.7K 6.8K 6.4K 3.6K 3.2K 4.2K 5.2K 5.4K 4.4K 3.5K 2.5K 1.4K 0.1K
9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
Weekly Trend for Total Orders
Avg-402.8
491
171
4 10 16 22 28 34 40 46 52
SALES PERFORMANCE
CATEGORY
Classic Category contributes the maximum sales, Total orders & Total Pizzas Sold
Size
Large Pizza Size contributes the maximum Total Sales
Percentage of Sales by Pizza Category
Veggie
23.68%,$193.69K
Chicken
23.96%,$195.92K
Total Sales
$817.86K
Supreme
25.46%,$208.20K
Classic
26.91%,$220.05K
Percentage of Sales by Pizza Size
Large 45.9%
Medium 30.5%
Small 21.8%
X-Large 1.7%
XX-Large 0.1%
Total Orders & Pizzas Sold by Pizza Category
Total Orders
Total Pizzas Sold
Chicken (11,050) 8,536
Classic (14,888) 10,859
Supreme (11,987) 9,085
Veggie (11,649) 8,941

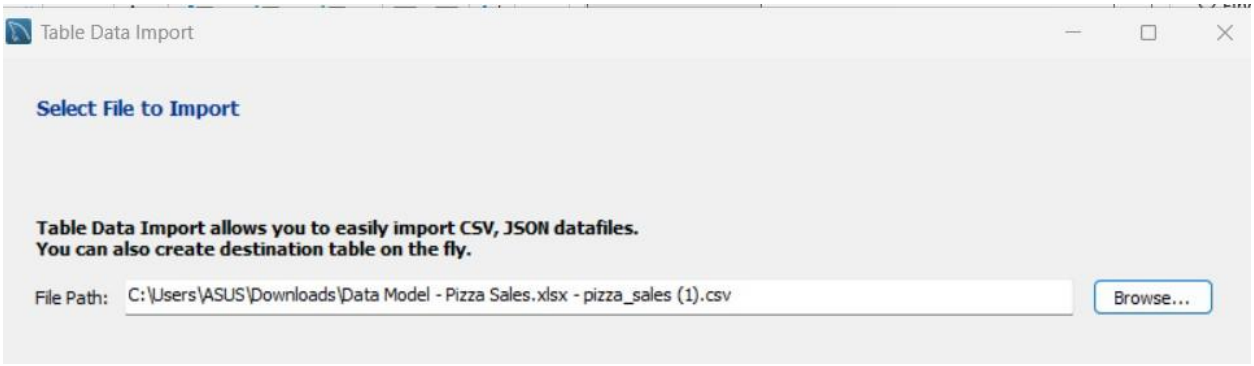
Top/ Bottom Sellers Page

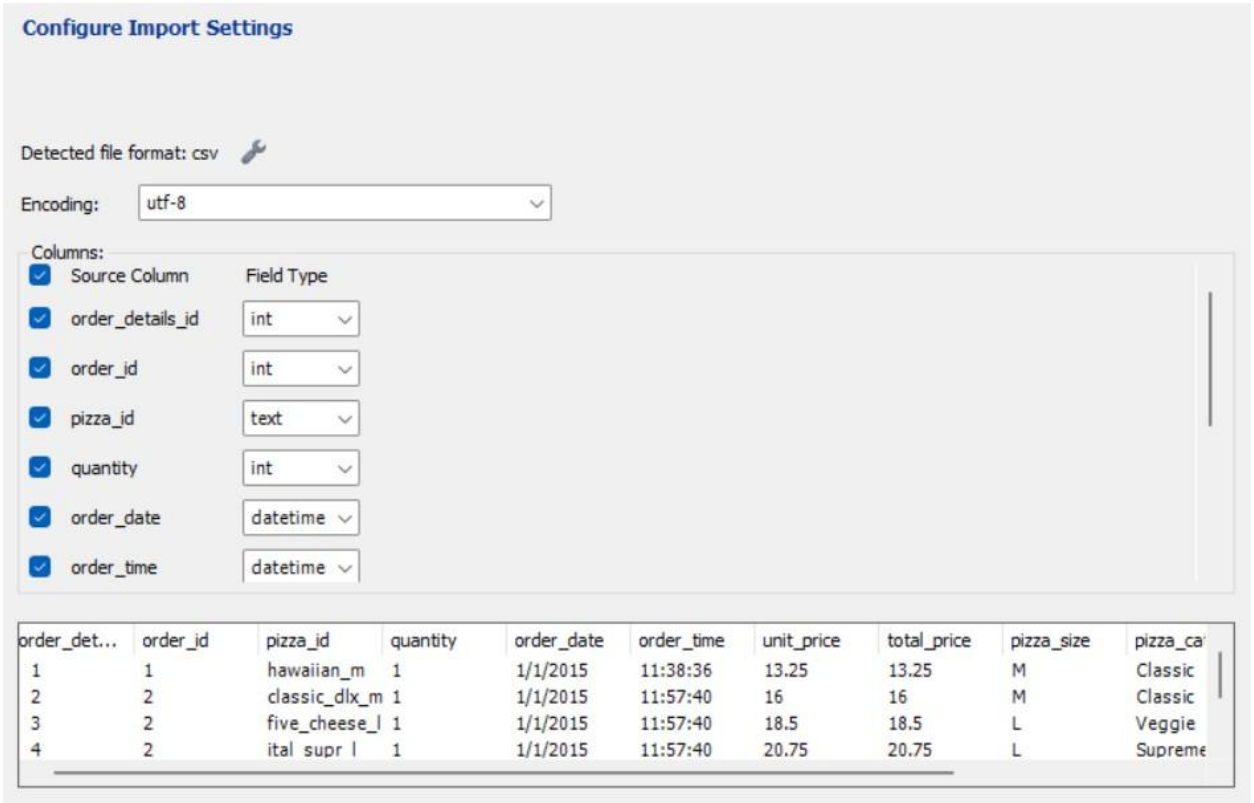


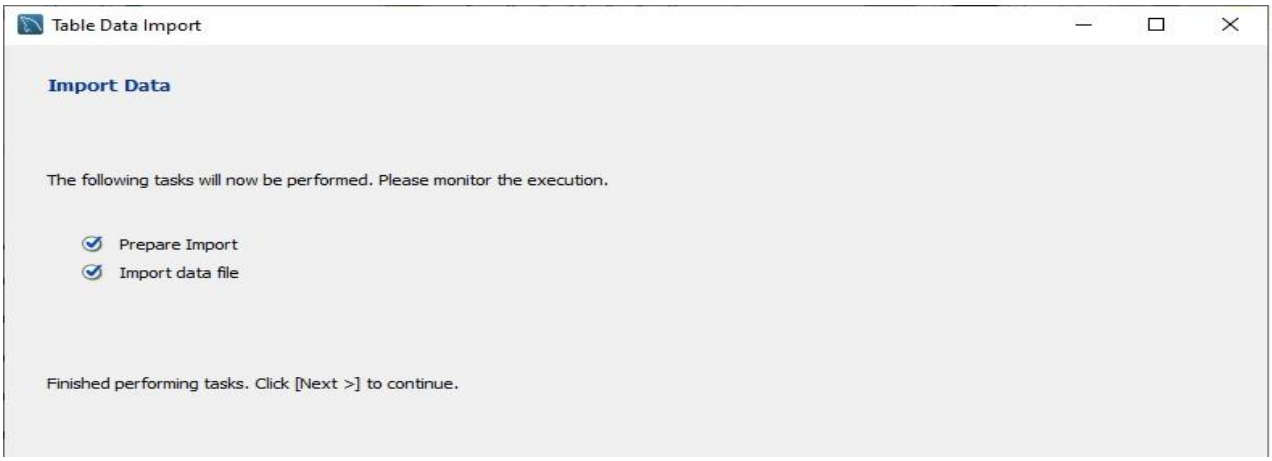
Text Boxes with insights



Data Import on MySQL:







Data Analysis using MySQL:

Analysis of different SQL queries are done by using MySQL workbench.

The All-SQL queries are present in another document.

Conclusion:

Data Analysis of pizza sales data are analyzed by using Tableau and MySQL Workbench, and Excel.

We gained some insights after data analysis, the key insights are follows:

- Peak orders are between 12:00 PM and 1:00 PM, Evening- 4:00 PM To 7:00PM.
- 48th Week from the month of Dec has highest peak weekly orders.
- Large pizza Size generated huge total sales. In Category wise Classic category contributes maximum sales, total orders and total pizza sold.
- The Thai chicken pizza contributes the maximum revenue and the brie carre pizza contributes minimum revenue.
- The Classic Deluxe pizza contributes to maximum total orders and total quantities.

COFFEE SALES

Project Objective:

To analyze the coffee sales data, identify sales trends & purchasing patterns of customer and to build a machine learning model by using ml algorithm.

Tools used:

1. Python- Spyder-5.3.1
2. Machine Learning
3. Excel-2019

Steps involved:

1. Problem Statement
2. Data Collection
3. Data Cleaning
4. Exploratory Descriptive Analysis (EDA)
5. Machine Learning Modeling
6. Model Evaluation & Conclusion

Problem Statement:

To analyze the coffee sales data, find out the monthly sales, day sales, weekly sales & purchasing patterns of customer and also do the time series EDA.

Data Collection:

The coffee sales dataset is collected from the data source.

The coffee sales dataset consists of 6 columns. Columns are date, datetime, cash_type, card, money, and coffee_name.

Date – Date of the purchased coffee product.

Datetime- The datetime provides the information of date along with time of purchased coffee product.

Cash_Type- The cash_type column has two unique values- cash, card. That means purchased coffee product is either in cash mode or card mode.

Card – The card column gives information about if we purchased coffee product in card mode it will read the card number.

Money - The money column gives information about cost of purchased coffee products.

Coffee_Name- The name of the coffee product. 8 different types of coffee products are present.

Data Cleaning:

After Loading the coffee sales dataset, we must clean the dataset. We found zero duplicates in the dataset and 89 missing values in the column 'card'. All 89 missing values in the column 'card' are from cash users only. After that we followed different data issues like feature engineering, etc. We cleaned the dataset.

Exploratory Data Analysis:

Conducted the EDA to visualize the sales trends.

We plotted bar plots and line charts to analyze the time series dataset.

We used matplotlib and seaborn libraries to visualize the data.

Machine Learning Modeling:

After EDA we must select the ml algorithm and build a machine learning model.

We choose **the Multiple Linear Regression Model**. Because Linear Regression is a statistical method used to model the relationship between a dependent variable and one or more independent variables. It makes predictions for continuous or numeric variables such as sales, salary, stock price, money, etc. Multiple Linear Regression is an extension of linear regression.

And other supervised learning algorithms like SVM (Support Vector Machine), K-Nearest Neighbour, also used for predictive modeling, but linear regression model is simple and easy way to do prediction of dependent variable.

After Choosing Machine Learning Algorithm, we must split the data into train data and test data. Then define the features and target variable.

After that train the model and make the predictions.

Model Evaluation & Conclusion:

Evaluating a Multiple linear regression model is crucial to understand how the model fits the data and how accurate its predictions are. Here are some common evaluation metrics and techniques for linear regression.

1. Mean Square Error

   MSE measures the average squared difference between actual and predicted values. Low MSE indicates a better fit.

2. R-Squared

   R-Squared indicates the proportion of the variance in the dependent variable that is predictable from the independent variable.

Conclusion:

After the time series EDA and builded a machine learning model, we concluded and gained some insights.

- **The most customers are preferred for the Americano with Milk and Latte coffee products.**
- **Americano with Milk, Latte and Cappuccino are top selling coffee products.**
- **Latte coffee product contributed high revenue and Espresso coffee product generated low revenue.**
- **Tuesday and Wednesday has the highest sales of the week, and on other days sales are similar.**
- **In hourly sales analysis, we observed three peak hours within each day. The peak hours are 10:00am, 11:00 am and 7:00pm.**