CASE STUDY 2

PIZZA Data Set

COLLEGE PROJECT REPORT

by

SHIVA PANDEY

ROLL NO. 70

Department of

COMPUTER APPLICATION

UNITED UNIVERSITY

PRAYAGRAJ

May 2023

BONAFIDE CERTIFICATE

This is to certify that this project report entitled **"CASE STUDY 2 on PIZZA DATASET"** submitted to **UNITED UNIVERSITY PRAYAGRAJ,** is a bonafide record of work done by "**SHIVA PANDEY**" under my supervision from "**26th APRIL**" to "**03th MAY**".

<div align="right">

Mr. Kuldeep Gupta

IBM Trainer

</div>

Mr. Prashant Shukla

Head of department

Computer science

Place: United University

Date: 03 / 05 / 2023

Declaration by Author(s)

This is to declare that this report has been written by me/us. No part of the report is plagiarized from other sources. All information included from other sources have been duly acknowledged. I/We agree that if any part of the report is found to be plagiarized, I/we shall take full responsibility for it.

Shiva Pandey

ROLL NO. 70

Place: United University

Date: 03 / 05 / 2023

TABLE OF CONTENTS

CHAPTER NO. TITLE   **PAGE NO.**

INTRODUCTION TO DATASET

Pizza Sales Dataset is a fictional dataset commonly used as an example in data analytics and business intelligence courses, including IBM Cognos. The dataset contains information about a pizza chain's sales over a period of time, including details about the pizzas sold, the customers who purchased them, and the stores where they were sold.

The dataset typically includes variables such as:

Date: The Date in which the Pizza was sold.

Time: The exact time on which the pizzas were sold.

Round of time: The time rounded off to hours.

Pizza ID: The IDs of the pizza that were sold.

Pizza Name: The name of the pizzas that were sold.

Pizza Size: The size of the pizzas that were sold.

Pizza Price: The price in which the pizzas were sold.

Quantity: The total quantity of pizzas sold.

Using this dataset, analysts can gain insights into various aspects of the pizza chain's business, such as which pizzas are most popular, which stores are performing best, and which customers are making the most purchases. These insights can then be used to make data-driven decisions that help improve the company's bottom line.

Explain the Tool which you are going to use.

COGNOS: IBM Cognos is a business intelligence and performance management software suite that helps organizations analyze and make informed decisions based on their data. The software provides tools for reporting, analysis, score carding, and monitoring, as well as planning, budgeting, and forecasting.

DATA MODULE: is a metadata layer that provides a simplified and unified view of data sources that are used to create reports and dashboards. The data module provides a user-friendly interface for creating relationships between data sources, defining calculations and aggregations, and specifying data governance rules. This makes it easier for business users to access and work with data, without requiring specialized technical skills.

A data module is designed to integrate with a wide range of data sources, including relational databases, flat files, and big data platforms. It provides a flexible and scalable approach to data integration, allowing users to connect to multiple data sources and combine them into a single view. This eliminates the need for users to have deep technical knowledge of the underlying data sources and makes it easier to work with data from different systems.

To create a data module in IBM Cognos, follow these steps:

• Open the IBM Cognos Analytics interface and navigate to the Data module option on the home screen.

• Select the data source that you want to use as the basis for your data module. This can be a relational database, flat file, or big data platform.

• Define the relationships between your data sources by dragging and dropping the relevant fields onto the canvas.

• Create any necessary calculations and aggregations by selecting the relevant fields and applying the appropriate mathematical or statistical functions.

• Define any data governance rules and policies that you want to apply to your data module.

Save your data module and use it as the basis for creating reports and dashboards.

Overall, a data module provides a user-friendly and flexible approach to data integration, allowing users to work with data from multiple sources without requiring specialized technical skills. By simplifying the process of data integration and governance, data modules can help organizations to make more informed decisions and gain insights into their data.

DASHBOARD: In Cognos, a dashboard is a visual representation of data that provides a quick and easy way to monitor and analyze key performance indicators (KPIs) and metrics. A dashboard typically consists of a set of graphical charts, tables, and other visualizations that help users to quickly understand the data and make informed decisions.

Cognos dashboards can be customized to fit the needs of specific users or groups. They can also be interactive, allowing users to drill down into the data to get more detailed information or to filter the data to focus on specific areas of interest.

Dashboards in Cognos can be created using a drag-and-drop interface, making them easy to build and customize. They can also be shared with others in the organization, making it easy to collaborate and make informed decisions based on the data.

Tools In Dashboard:

Summary: A summary is a type of visualization that displays a condensed version of data in a dashboard. It typically shows aggregated data in a simple format, such as a total or an average.

Dropdown List: A dropdown list is a user interface element that allows users to select an item from a list of options. It is often used in Cognos Dashboard to filter data based on user selection.

Heat Map: A heat map is a type of visualization that displays data in a matrix format where each cell is shaded according to its value. It is often used to represent the density of data points in a particular area.

Pie Chart: A pie chart is a circular chart that is divided into slices to represent numerical proportions. It is often used in Cognos Dashboard to show the relative proportions of different data points.

Crosstab: A crosstab is a table that displays data in a grid format where each row represents a category and each column represents a different variable. It is often used in Cognos Dashboard to compare data across different categories and variables.

Map: A map is a type of visualization that displays data geographically. It is often used in Cognos Dashboard to show the distribution of data across different regions.

Network: A network is a type of visualization that displays data as nodes connected by edges. It is often used in Cognos Dashboard to show relationships between different data points.

Column: A column is a type of visualization that displays data as vertical bars. It is often used in Cognos Dashboard to compare data across different categories.

Creating The DATA MODULE

NOTE: We have been provided four files that contain the data that we need for this case study, but the data is not yet in module form, so we are going to create a new Data module to use on the dashboard.

STEP 1: Click on the NEW button on the screen and we will see the option to create a new Data module.
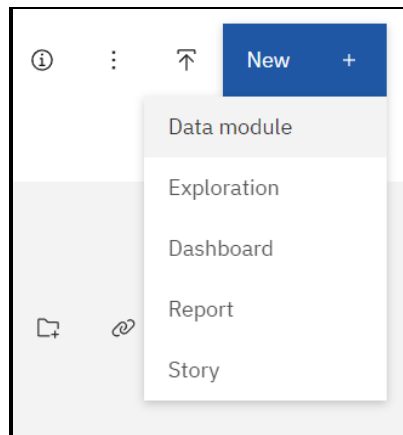
**STEP 2:** To create a new data module we will need to upload our own data that was provided to us. Click on upload symbol.
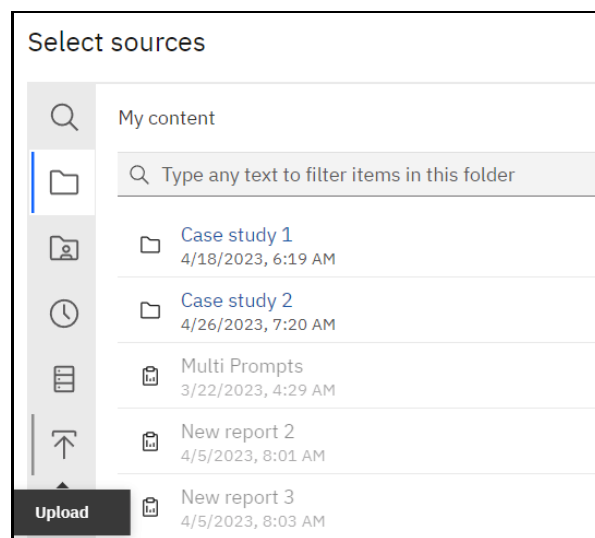
STEP 3: Browse to the location where we have the files and select all the files.
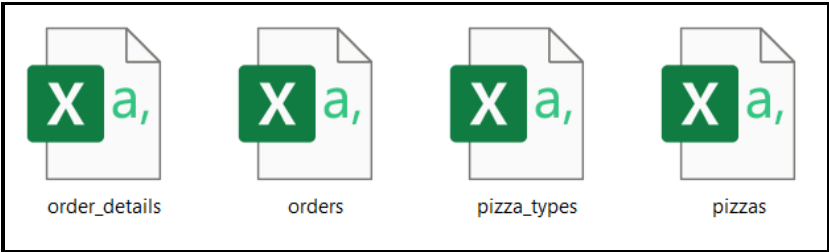
STEP 4: Now we have to create relations between these data sets. Click on the relationship tab on the top.
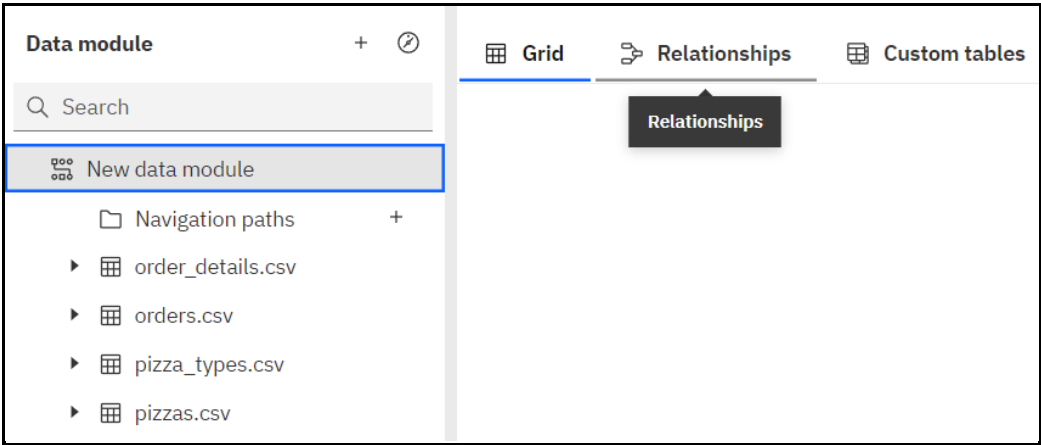
STEP 5: After clicking on relationships we will see all our files scattered on the screen.
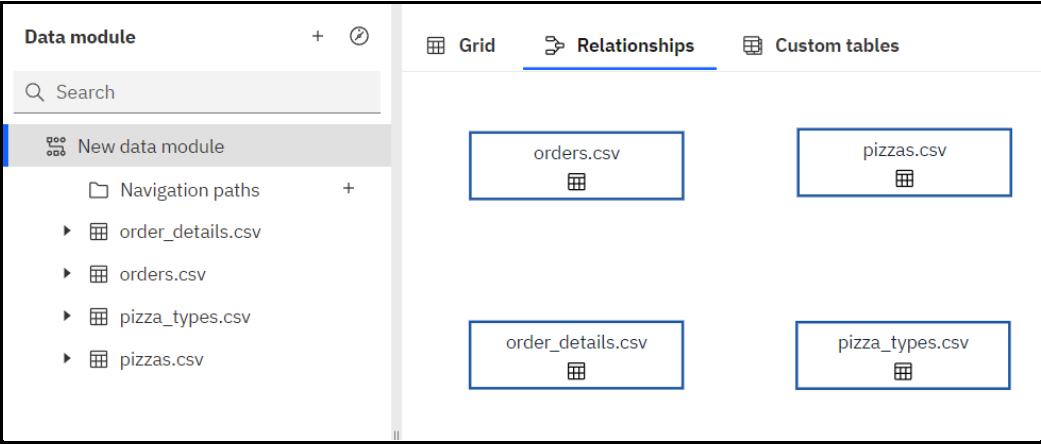
STEP 6: Right click on one of the files and click on relationship.
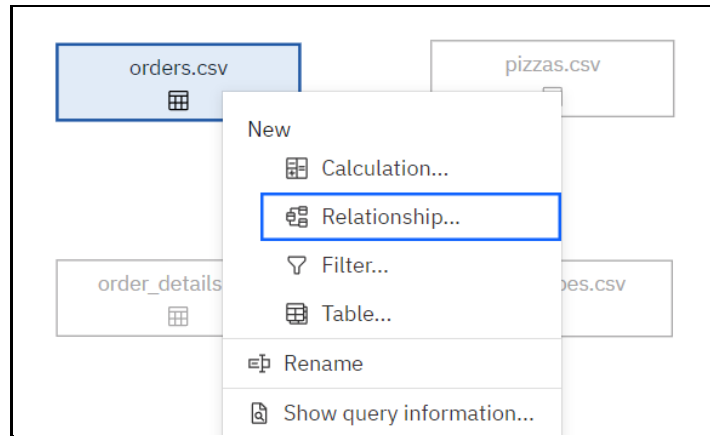
NOTE: So, to create relationship between two or more files we need to select the commons in both the files.

STEP 7: Select the other table that we want to make relationship with. Select order_details.
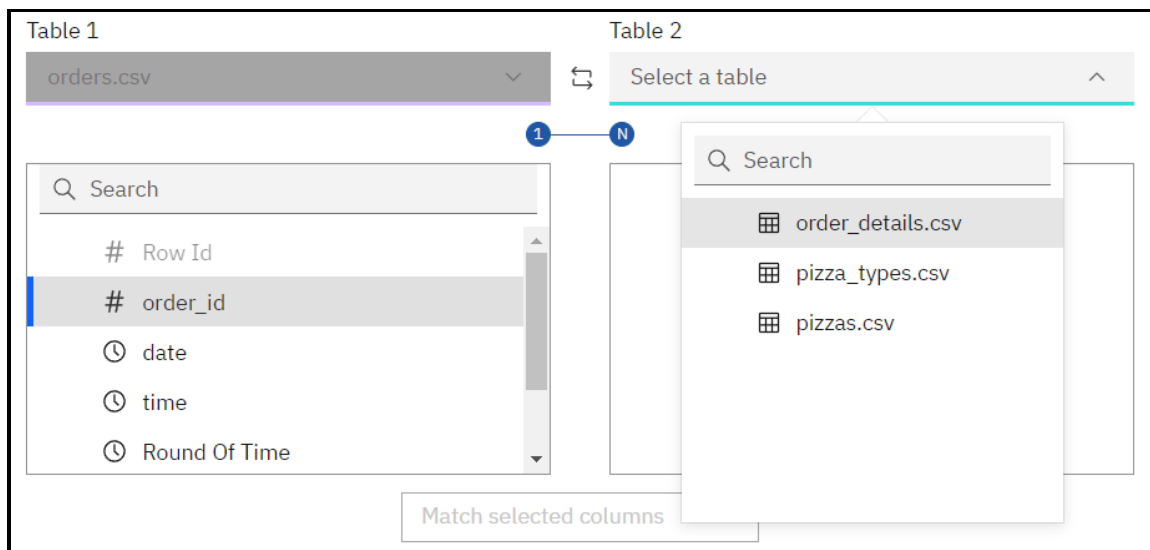
STEP 8: Select the common in both tables and click Match selected columns and then OK.
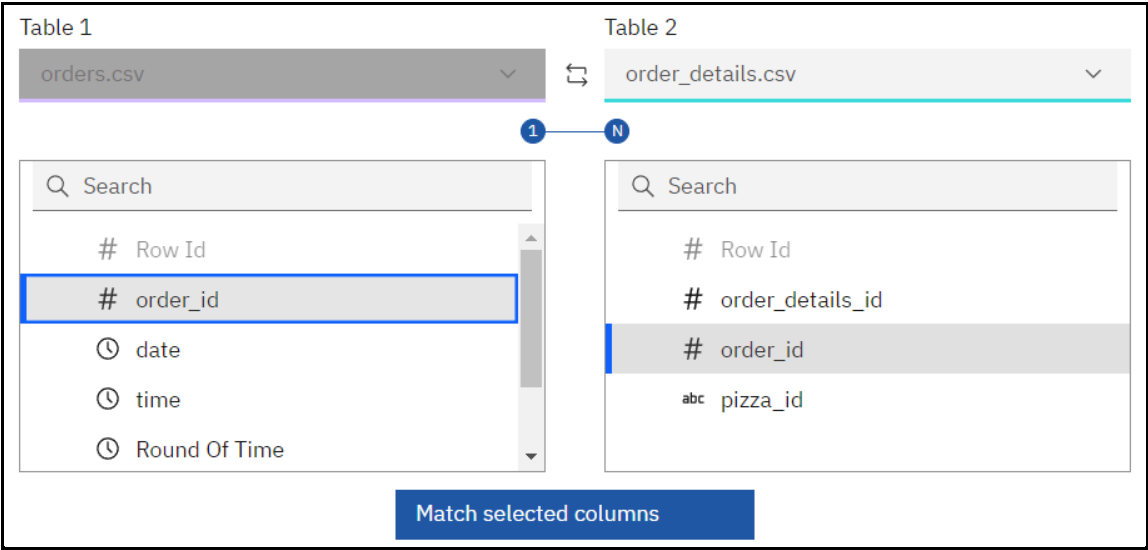
STEP 9: You can see that we created a 1 to many relationship between 'orders' and 'order_details'.
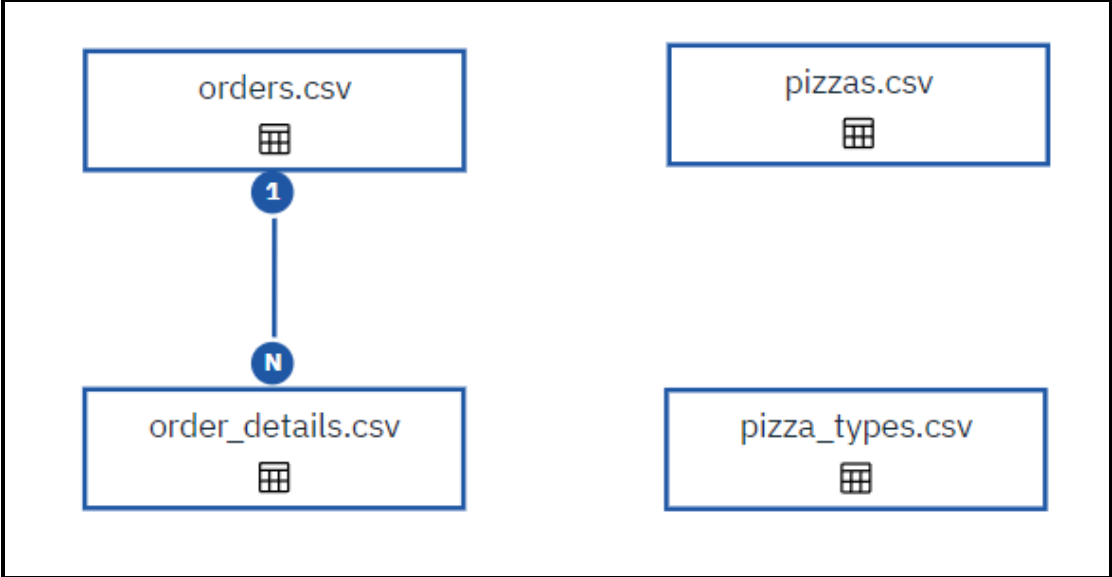
STEP 10: Similarly create relationships between "pizzas" "pizza_types" and "pizza" " order_details".
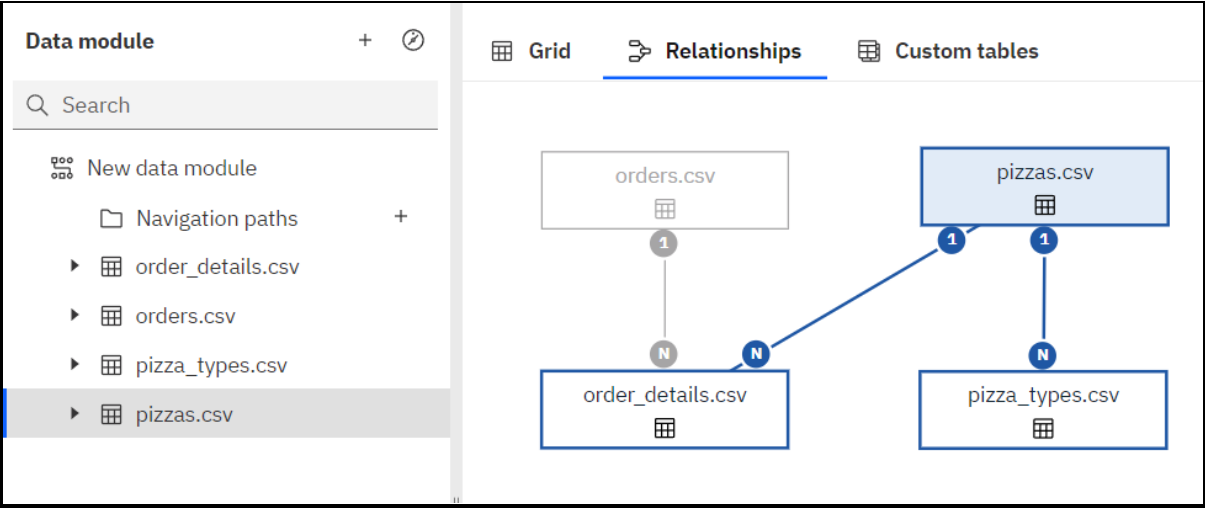
Now we are going to create a custom table that we will use to show Total sales on our dashboard. To create this table, we are going to select two tables "order_details" and "pizzas" because these tables contain the data 'quantity' and 'price' which we need to find out total sales.

STEP 11: Right click on the table 'order_details' and click on tables.
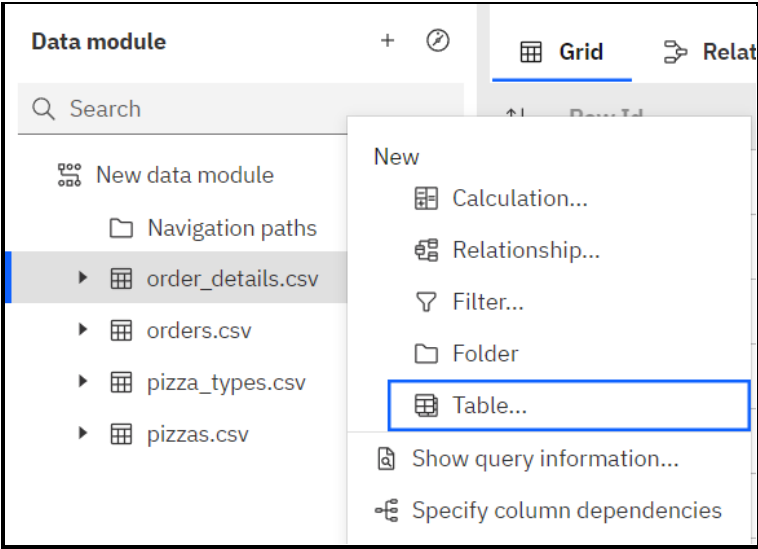
STEP 12: Click on add icon and add the table 'pizzas', then click on 'Joined view' and finish.
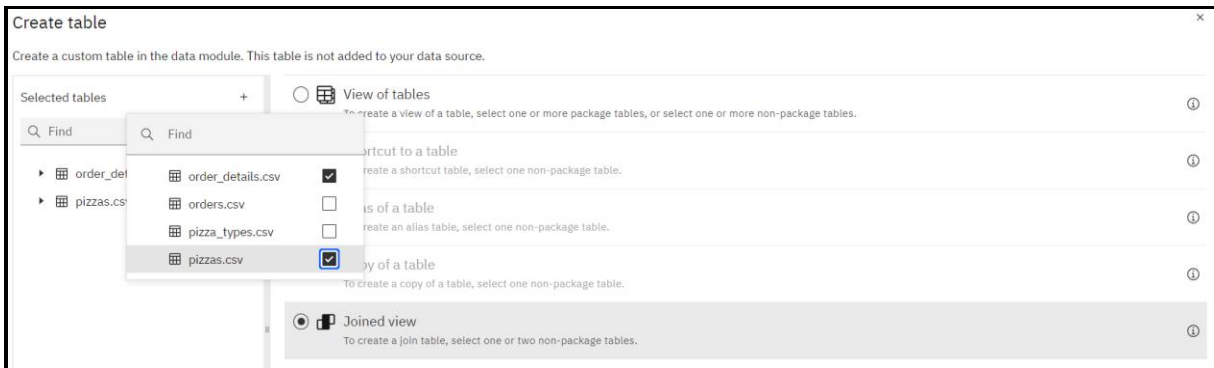
STEP 13: Now we can see our custom table on the Relationships tab.
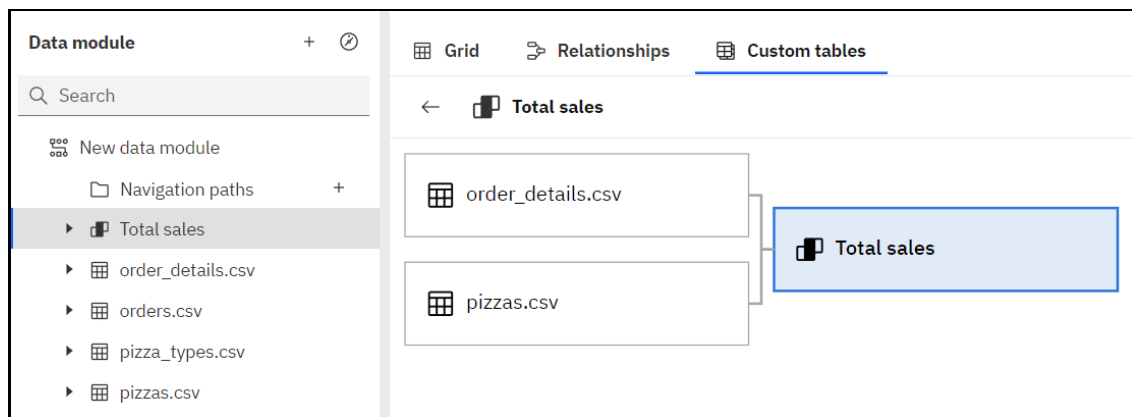
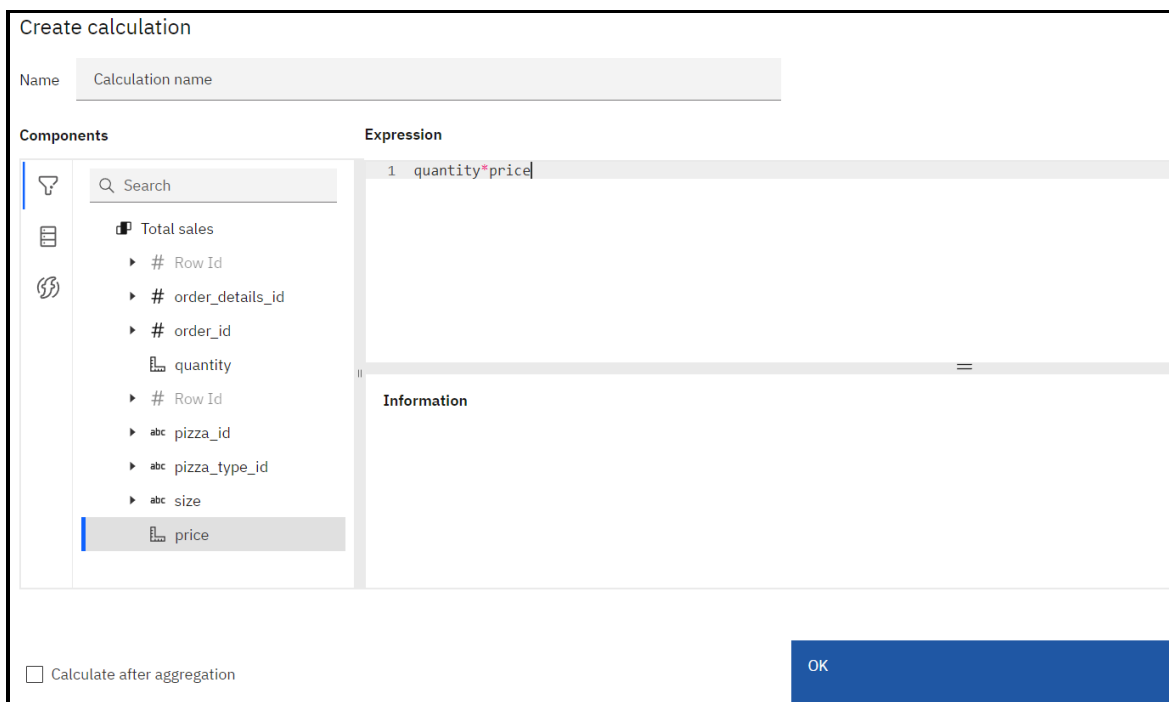**STEP 14:** Create relationships from new table to 'order_details' and 'pizzas'.
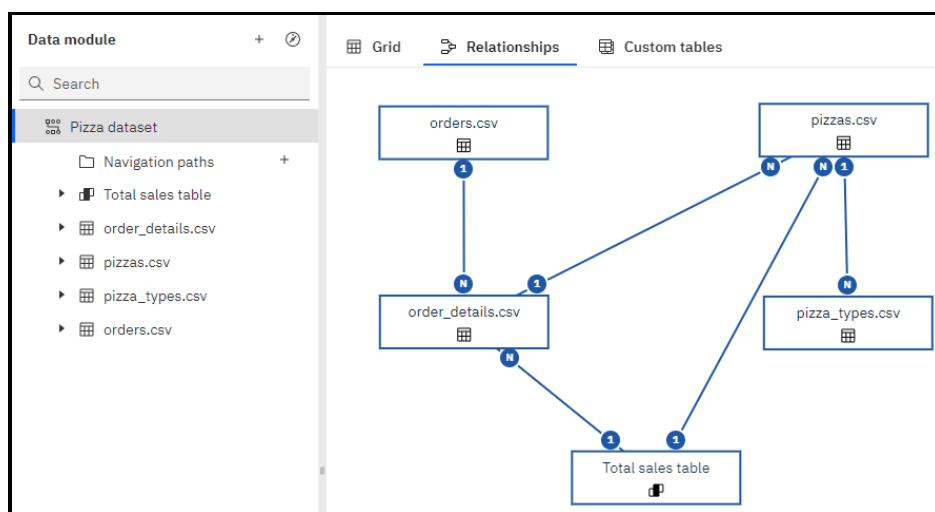
NOTE: So, to get total sales we are going to create a calculation on the new table 'Total sales'.

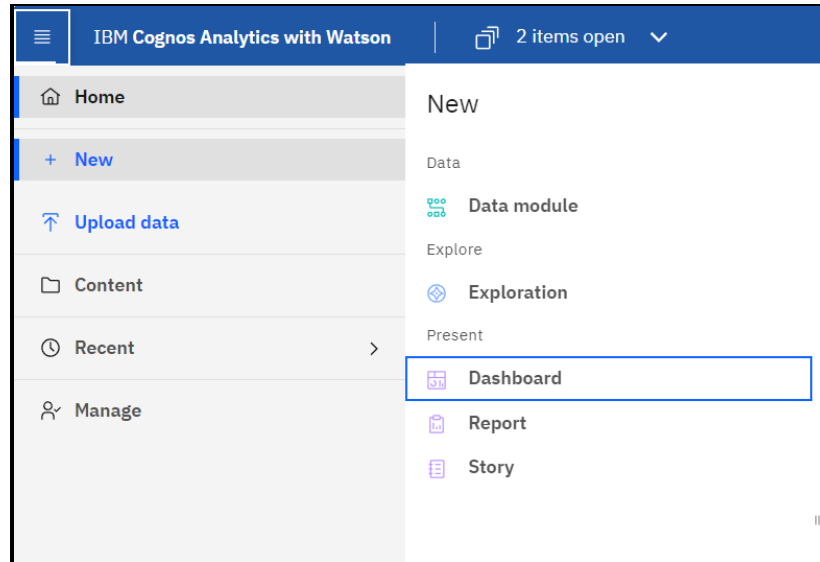**STEP 15:** Right click on 'Total sales' and click on calculation.
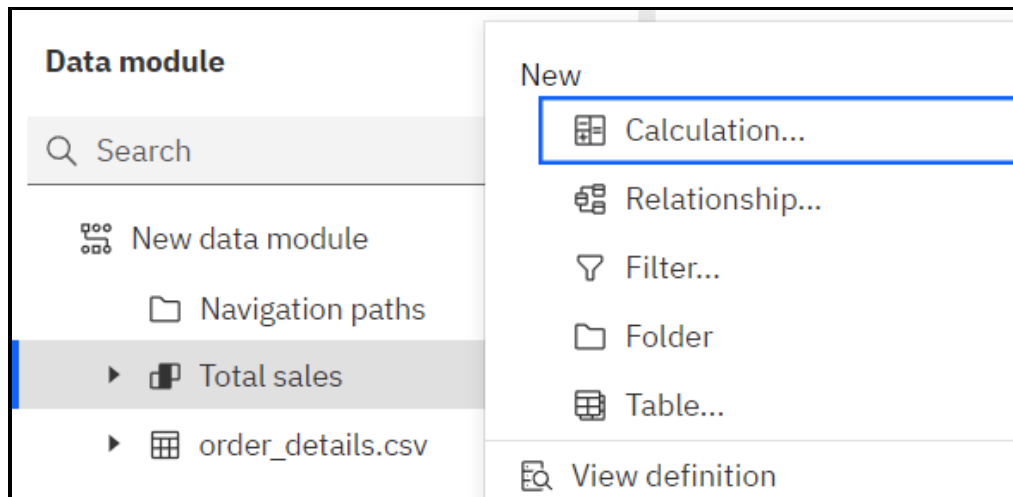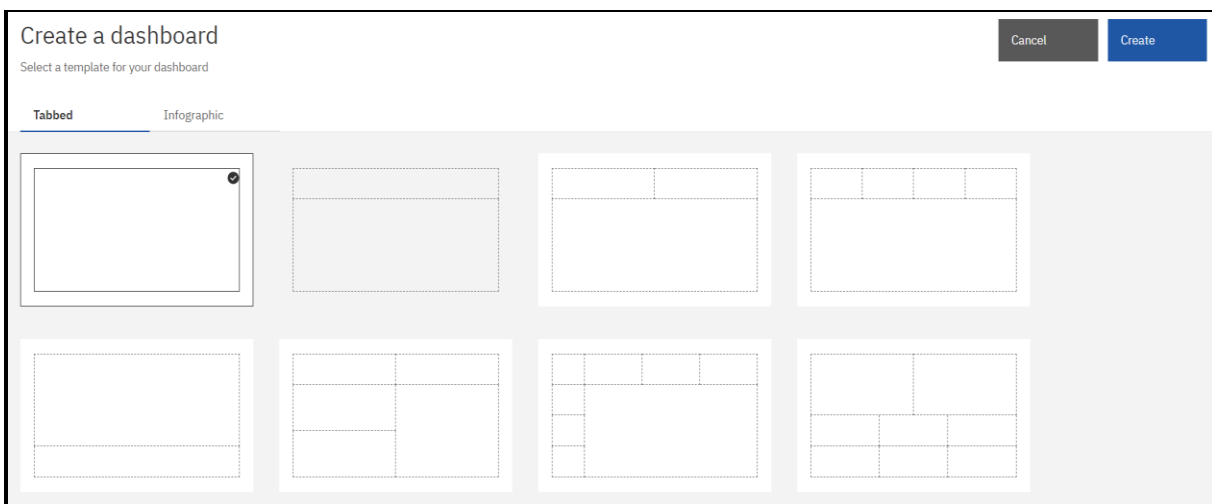

STEP 16: Multiply 'quantity' and 'price', click ok.

Steps to create Dashboard.

STEP 1: On the home page of IBM Cognos there is a Hamburger menu on the top left of the screen, by clicking on that and then new the user can see the option to create a new Dashboard

STEP 2: By clicking on the Dashboard icon a new window will show the user some predefined templates for the Dashboard
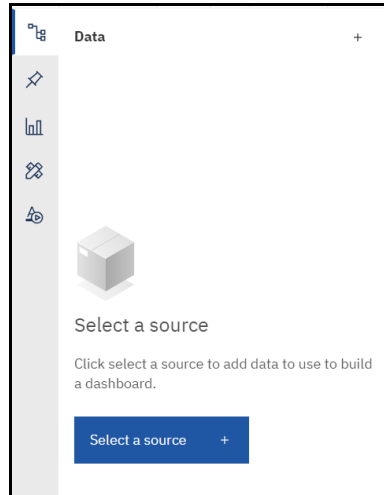
STEP 3: After choosing the template the user can click on the "CREATE" button on the top right of the screen, this will then create the dashboard and open the empty dashboard page
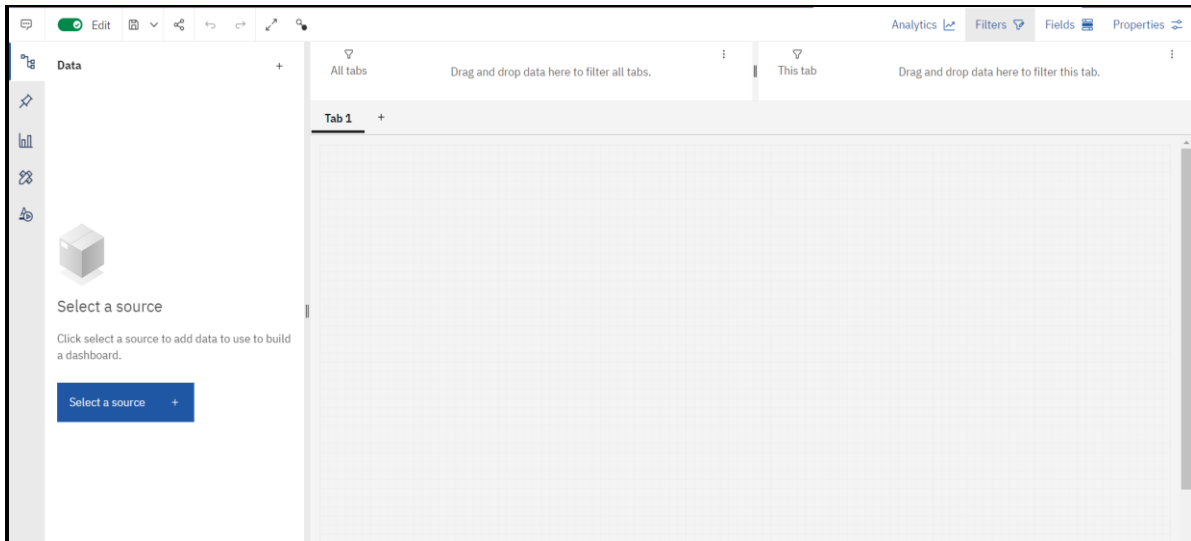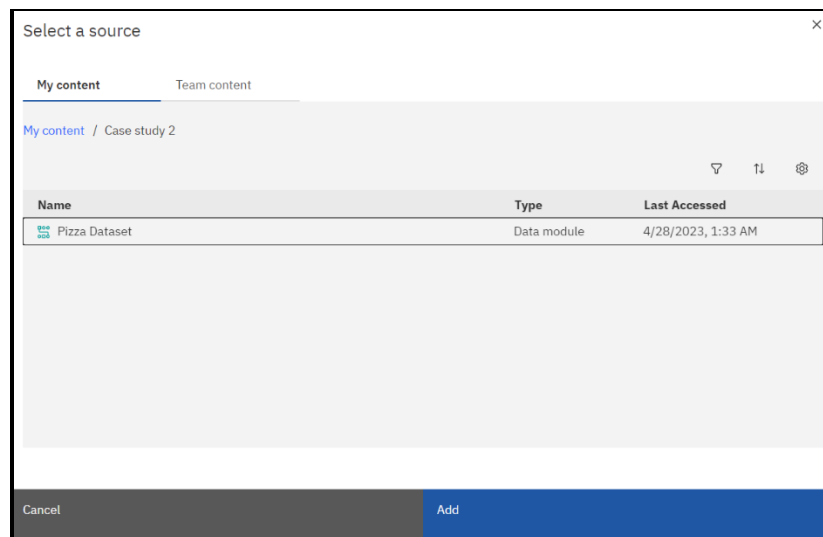
Steps to import the DATA.

STEP 1: To import the dataset that the user will use on the Dashboard, click on the "Select a source" button on the left side of the screen.

STEP 2: After clicking on the button a window will pop showing the location that user can import data from (Make sure to import data modules as Cognos tools only support data modules). Select Pizza Dataset.
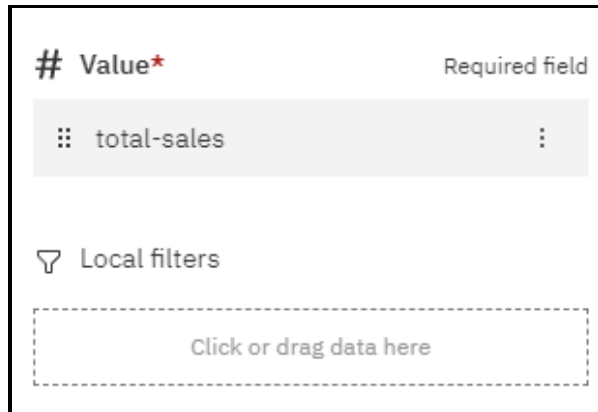
## Queries you have to resolve using Dashboard Components:

- Show Total Sales organized by months.
- Create a chart that displays the contribution of each category to total pizza sales.
- Identify the weekday with the highest pizza sales during a particular hour.
- Create a line chart that illustrates the fluctuation in total sales over the months.
- Determine which pizza type was the most popular in the month of June.

**QUERY 1:** Show Total Sales organized by months.

STEP 1: Drag and drop a 'Summary' tool from visualization tab onto the dashboard. On the field '#Value' we will add 'total-sales' data set. This will show us the total sales.

STEP 2: Now to show the data organized by months we will drag and drop a pie chart from the visualization tab. In filed Segments we will add Month and in #Size add total-sales.
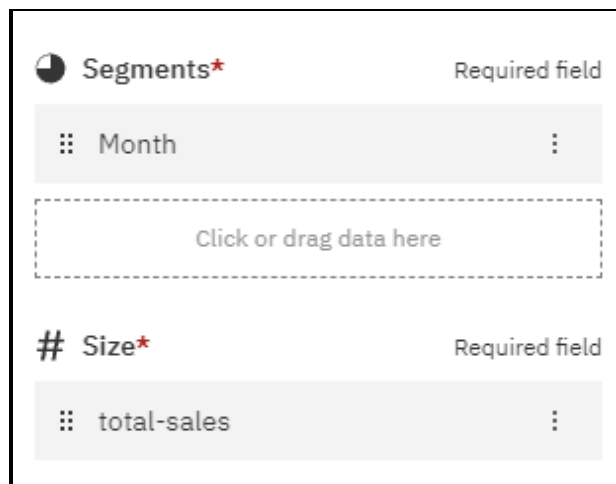
NOTE: So, the dashboard will show us the sales according to months. To get a specific month's sales we can select that month on the pie chart.
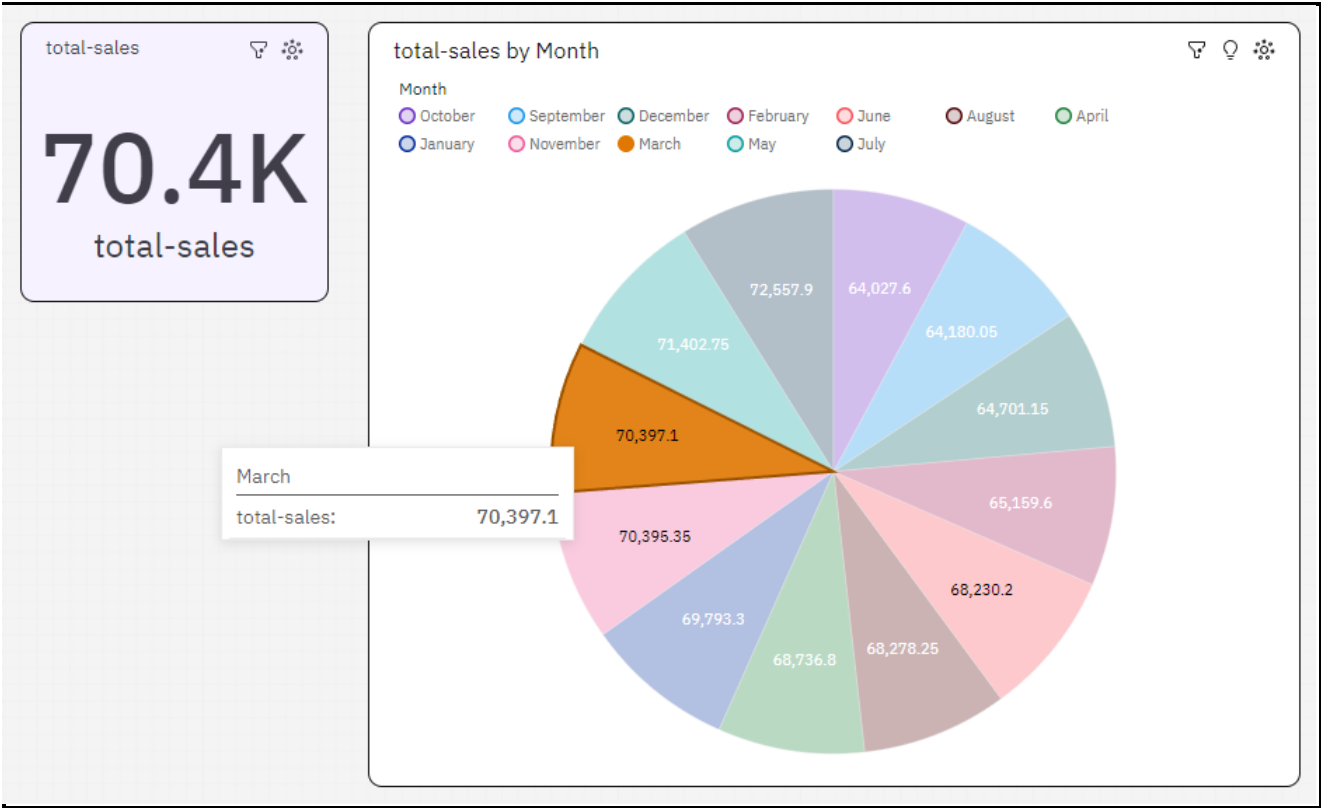
QUERY 2: Create a chart that displays the contribution of each category to total pizza sales.

STEP 1: Drag and drop a Column chart from visualization tab.

STEP 2: In the field 'Bars' add 'Category', in the field '#Length' add 'total-sales' and in field 'Color' add 'category'.
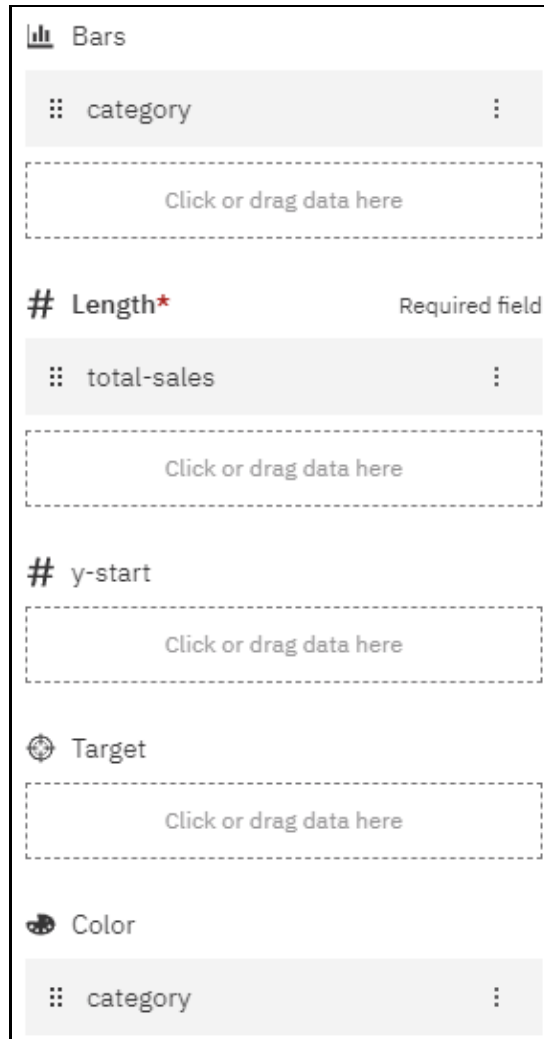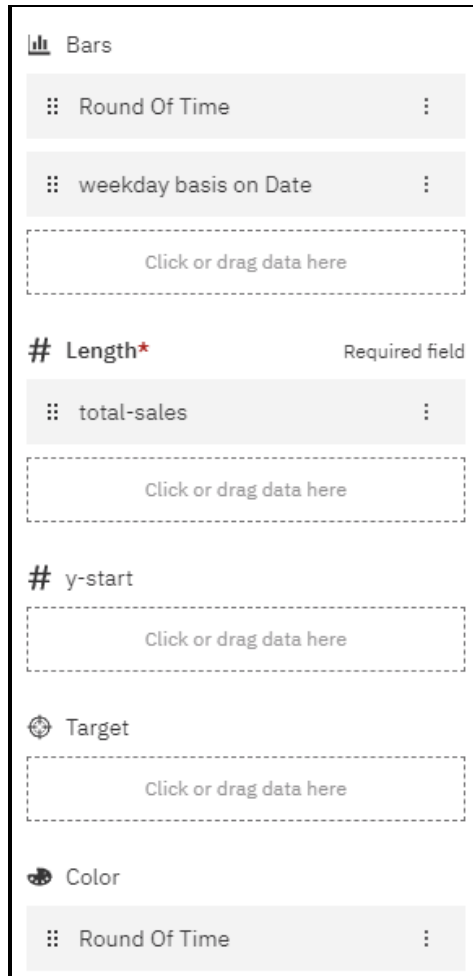
NOTE: The dashboard shows the total sales divided by different categories of pizzas sold. By looking at the dashboard we can understand the contribution of each category to the total sales of pizzas.
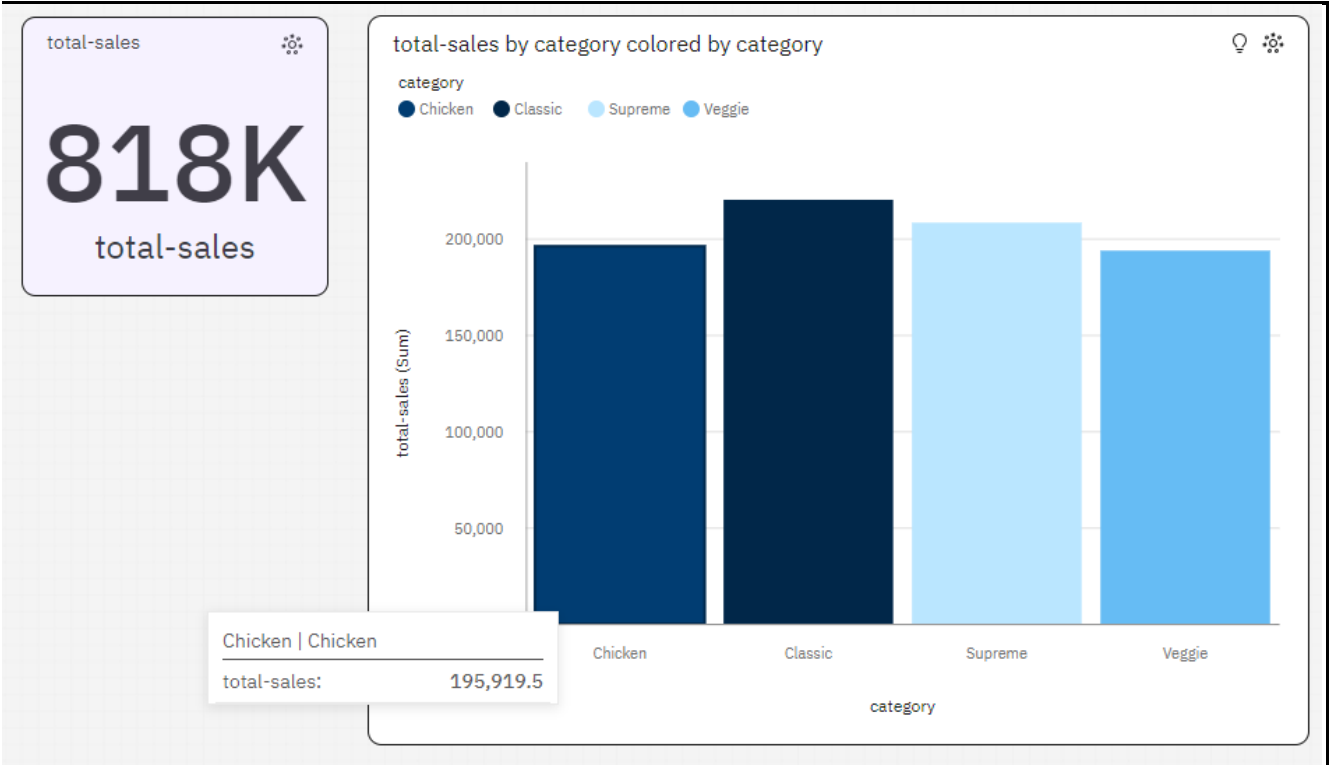
QUERY 3: Identify the weekday with the highest pizza sales during a particular hour.

STEP 1: On the dashboard drag and drop a Column chart from visualization tab.

STEP 2: In the field 'Bars' add 'Round of time' as well as 'weekday basis on Date' , in the field '#Length' add 'total-sales' and in field 'Color' add 'Round of time'.

NOTE: The dashboard shows the total sales filtered by time and the weekdays. So, the column chart is showing both the rounded time and the day on which sales happened and we can see the most sales took place between **12 pm to 1 pm on Thursday** with 19k sales.
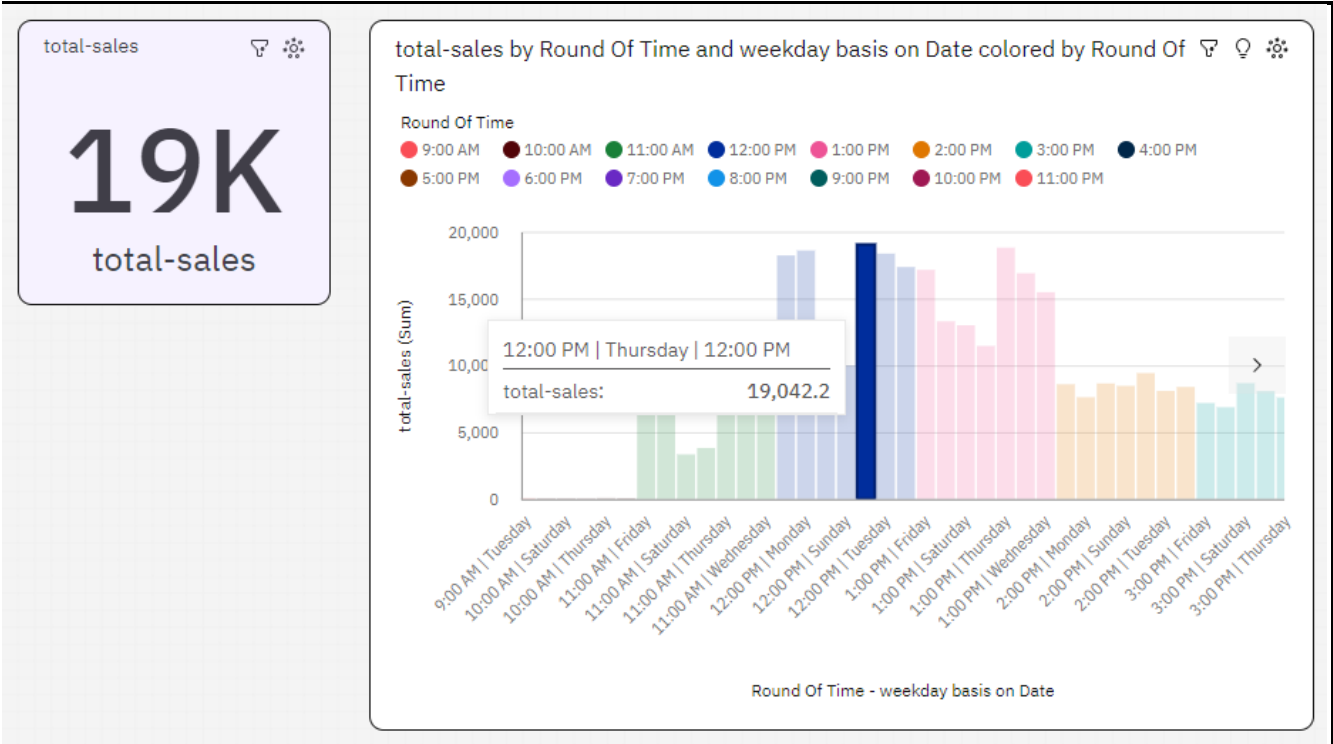
QUERY 4: Create a line chart that illustrates the fluctuation in total sales over the months.
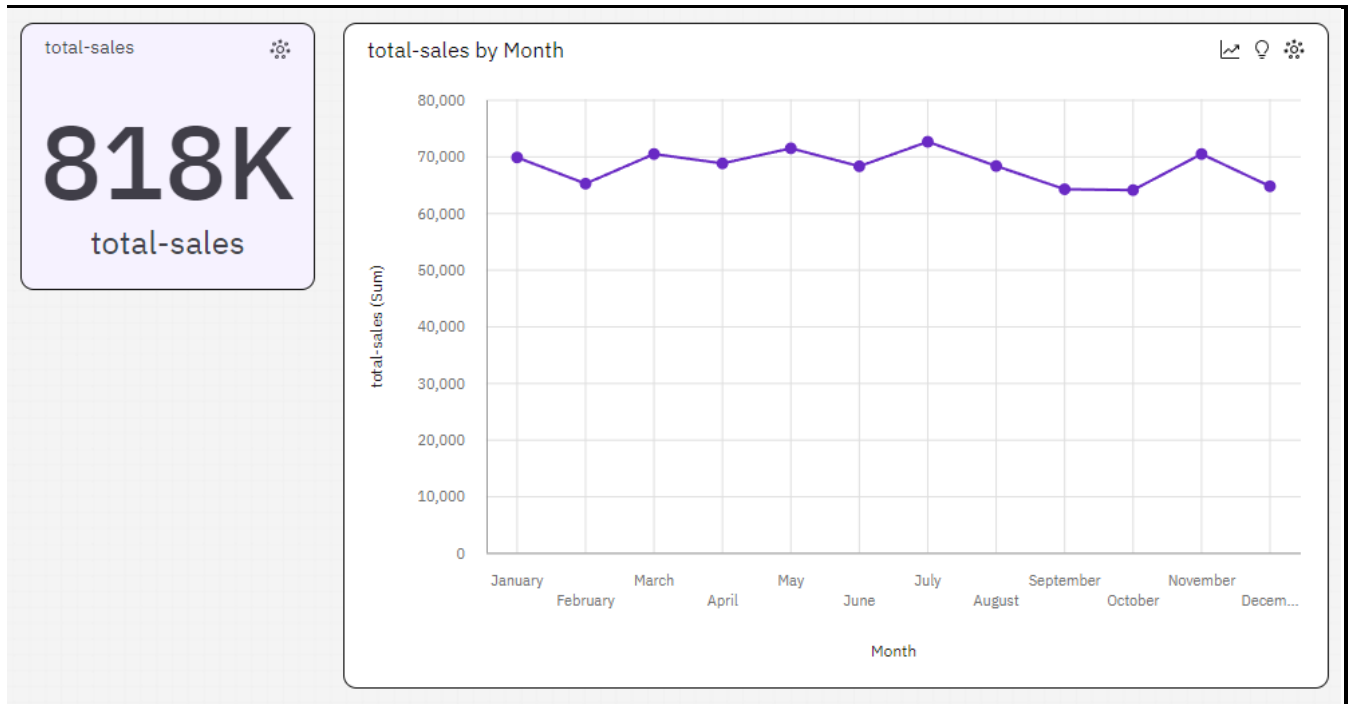
STEP 1: For this query we are already asked to use a Line chart so lets drag and drop a Line chart from the visualization tab.

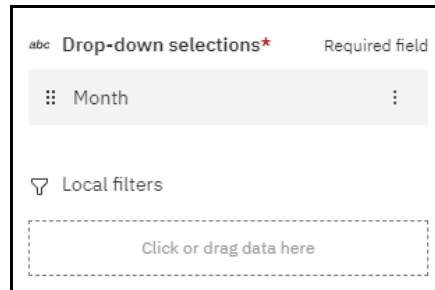**STEP 2:** In the field 'x-axis' add 'Month' and in the field 'y-axis' add 'total-sales'.

NOTE: The dashboard shows a clear Fluctuation in pizza sales on different months.

QUERY 5: Determine which pizza type was the most popular in the month of June.
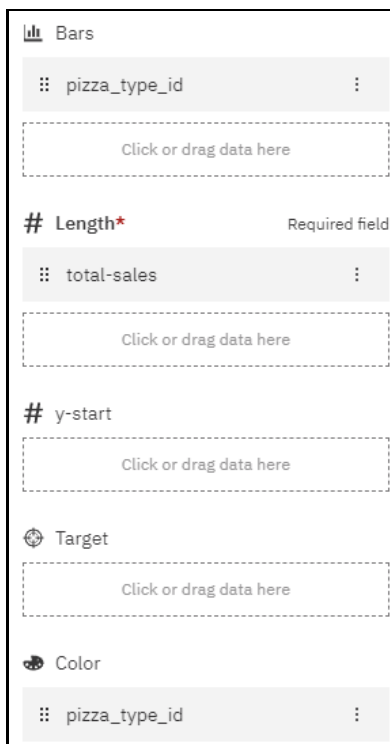
STEP 1: Drag and drop a 'Drop down list' from the visualization tab.

STEP 2: In the field 'Drop down selection' add 'Month'.

**STEP 3:** Now drag and drop a Column chart on the dashboard.

STEP 4: In the field 'Bars' add 'pizza-type-id', in the field '#Length' add 'total-sales' and in field 'Color' add 'pizza-type-id'.

STEP 5: On the Drop-down-list select the month June because we are supposed to show the best pizza type on June.

NOTE: The most popular pizza type on June was 'cali_ckn' with 3.9k sales, as we can see on the dashboard.
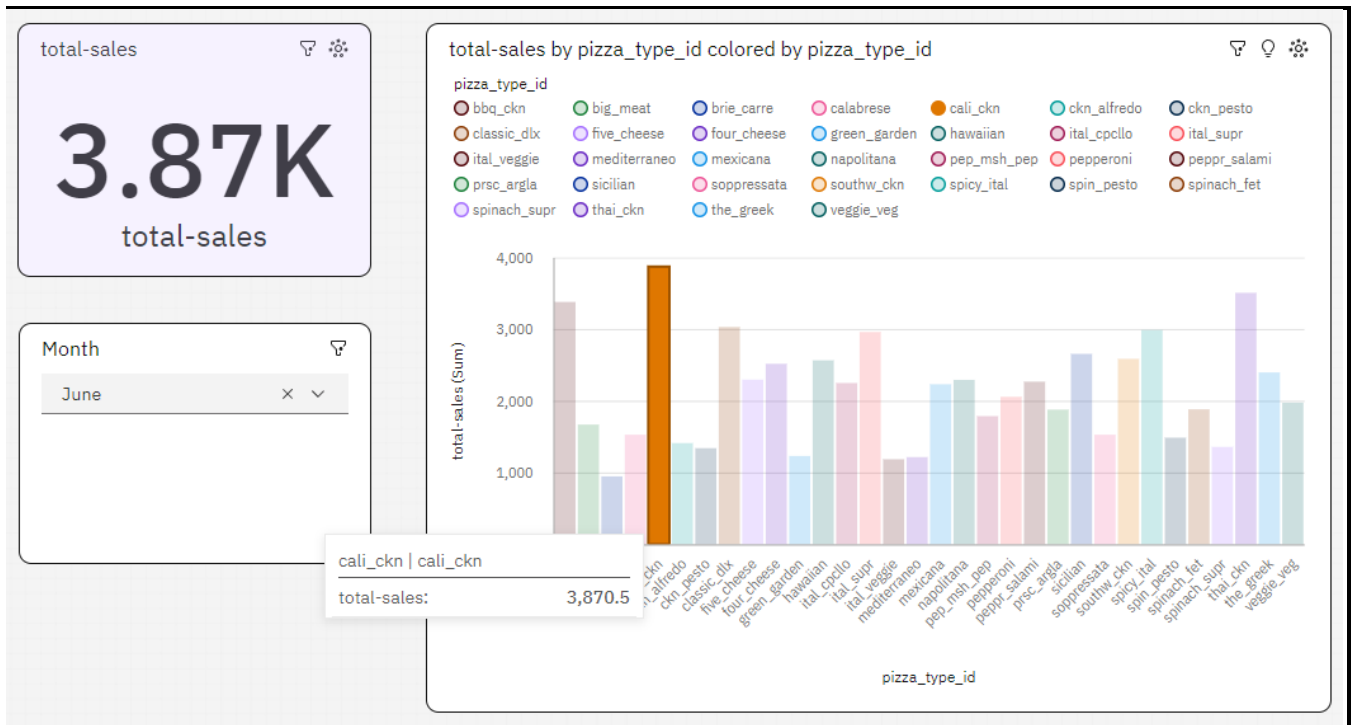
CONCLUSTION

In conclusion, the IBM Cognos data set for pizza sales provides valuable insights into various aspects of the business. The problem statements addressed in this analysis allowed us to gain a deeper understanding of the revenue generated, category contributions, peak sales hours, sales fluctuation over the years, and the most popular pizza type in a specific month.

Firstly, by examining the total revenue generated in 2015, organized by month and year, we were able to identify the financial performance patterns and assess the success of the business during that period.

Secondly, the chart illustrating the contribution of each category to total pizza sales helped us understand which categories played a significant role in driving sales and identify any areas that might require improvement or additional focus.

Thirdly, the identification of the weekday with the highest pizza sales during a particular hour enables the business to optimize its operations by allocating resources effectively and strategically planning promotional activities.

Fourthly, the line chart depicting the fluctuation in total sales over the years provided a visual representation of sales trends, allowing us to identify any long-term patterns, seasonality, or anomalies that may influence future decision-making.

Lastly, determining the most popular pizza type in the month of June allowed us to recognize customer preferences during that specific period and adjust marketing strategies or product offerings accordingly.

Overall, the analysis of the IBM Cognos data set for pizza sales presented valuable insights that can guide strategic decision-making, optimize operations, and enhance customer satisfaction. By leveraging the power of data analytics, the business can refine its strategies, allocate resources effectively, and stay ahead of the competition in the dynamic pizza industry.