**Analysis of Pizza Sales Dataset**

**Case Study**

*by*

Rishika Kesharwani

Roll Number: (56)

**Department of Computer Application**

**United University**

**Prayagraj**

April 2023

BONAFIDE CERTIFICATE

This is to certify that this project report entitled “**Analysis of Pizza Sales Dataset” submitted** to **United University Allahabad**,is a bonafide record of work done by “**Rishika Kesharwani”** under my supervision from “26/04/2023” to “06/04/2023”.

Mr. Kuldeep Gupta

Department of Computer Application

Date: 25/04/2023

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. We aver that if any part of the report is

found to be plagiarized, we are shall take full responsibility for it.

Date:- 06/04/2023 Rishika Kesharwani

Roll no:- (56)

**TABLE OF CONTENTS**

ABSTRACT iii

NOMENCLATURE xvii

**CONTENTS**

1. Introduction ....................................................................... 1

2.Tools………………………………………………………………………………….2

2. Preparation before problem solving ................................. 7

3. Problem Addressed........................................................ 10

4. Solution of problem 1..................................................... 11

5. Solution of problem 2......................................................... 13

6. Solution of problem 3....................................................... 14

7. Solution of problem 4...................................................... 15

8. Solution of problem 5.......................................................16

9. Conclusion ………………………………………………………………….…...18

**INTRODUCTION**

This data set contains the details about the pizza sells in year 2015.

This data set contain 4 tables these are :-

1.**Order\_Details:-** This table contains the details of orders like order\_details\_id, order\_id, and pizza\_id.

2. **Orders**:- This table contains the information about order id, date, time, round of time, month and weekday basis on date.

3. **Pizza type**:- This table contains the information about the pizza types it contained the details like pizza type id, name, category and ingredients.

4. **Pizzas**:- This table contained the details of pizzas like pizza type id, pizza id, size and price.

These are all the details about the data set which are given. This data is not directly used in IBM Cognos. First of all we have to create data module.

Before we solve the problem we have a proper knowledge about the **IBM COGNS TOOLS** so these are some tools as follows :-

**Tools-**

We are using IBM Cognos for data analytics. So, first question arises is that-

**What is Data Analytics?**

Data Analytics is the process of collecting, processing, and analyzing data to uncover insights, trends, and patterns that can be used to make informed decisions. It involves using various tools and techniques to extract valuable insights from large and complex datasets.

So now we know about data analytics, basically we use analytics to understand the market and take the good move so that business can achieve organization goal.

So, now the question arrives,

**What is IBM Cognos tool and why we are using it?**

IBM Cognos is a business intelligence and performance management software suite that helps organizations analyses 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. And the reason behind using IBM Cognos is that it is a BI tool that is designed for enterprise-level organizations. It provides a comprehensive suite of capabilities for reporting, analysis, and planning. Cognos allows users to create dashboards and reports that can be accessed by stakeholders across the organization. It also offers advanced analytics capabilities, such as predictive analytics and data mining.

There are many things, we are going to use **Data Visualization** in our data. Here, data is taken from there save data module. IBM Cognos is very user interactive software it helps us to create very attractive Dashboard easily.

**What is Data Module?**

In IBM Cognos, a 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 isdesigned 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*.*

**What is Data Visualization?**

Data visualization is the representation of data in a graphical or visual form. It involves creating visual representations of data to communicate complex information clearly and effectively. Data visualization is an essential aspect of data analytics, as it helps to convey insights and trends in a more intuitive and understandable way.

So, we are using Data Dashboard for representing these data, using graphs, charts, etc.

**What is IBM Cognos Dashboard?**

Dashboard is a place where you can explain your data visually and make the communication easily to understand. It is Data Visualizing board for Cognos.

There are many tools which help you to make data more attractive. Here are some tools which we are going to use or some which are present in Dashboard:

**Column**

Use a column visualization to compare values by one or more columns, such as sales

for products or sales for products each month.

**Stacked column**

Use a stacked column visualization to compare the proportional contributions for each

item to the total, such as sales for products and sales for products each month.

**Bar**

Use a bar visualization to compare values by one or more columns, such as sales for

products or sales for products each month.

**Stacked bar**

Use a stacked bar visualization to compare the proportional contributions for each item

to the total, such as sales for products and sales for products each month.

**Bubble**

Use a bubble visualization to show relationships among columns that contain numeric

values, such as revenue and profit.

**Packed bubble**

Use a packed bubble visualization when you want to show relationships among columns that contain numeric values, such as revenue. It is similar to the bubble visualization but the bubbles are tightly packed instead of spread over a grid. A packed bubble visualization shows a large amount of data in a small space.

**Line**

Use a line visualization to show trends over time.

**Line and column**

Use a line and column visualization to highlight relationships between multiple data

series by combining bars and lines with one visualization.

**List**

Use a list visualization to create an overview the data in a hierarchical way.

**Point**

Use a point visualization to show trends over time.

**Area**

Use an area visualization to emphasize the magnitude of change over time.

**Pie**

Use a pie visualization to highlight proportions. Each slice shows the relative

relationship of each part to the whole.

**Tree map**

Use a tree map visualization to identify patterns and exceptions in a large, complex data asset.

**Table**

Use a table to show detailed information from your database, such as product lists and customer lists. A table shows data in rows and columns. Each column shows all the

values for a data item in the database or a calculation based on data items in the

database.

**Hierarchy**

Use a hierarchy when you want to see the data in rows and columns.

**Summary**

Use a summary visualization when you want to see the total for a measure or the count

for a categorical column.

**Radial bar**

In a radial bar visualization, each bar appears in a circle with longer bars that represent

larger values. Hover over a bar to see the details about it, such as the exact value

represented by the bar. Each bar starts at 12 noon and goes in a clockwise direction for

positive values and counterclockwise for negative values.

**Scatter**

Scatter visualizations use data points to plot two measures anywhere along a scale, not only at regular tick marks.

**Word cloud**

Use a word cloud visualization when you want to see a text-based visualization of a

given column. The text height represents the scale. The name itself is the different

members of the column.

**Network**

Use a network visualization when you want to see the connections among columns in

your data asset. A network visualization is a good choice to show connections,

networks, and points of intersection.

**Heatmap**

Use a heat map visualization to visualize the relationship between columns and you

want it to be represented in a matrix type view.

**Data player**

Use a data player to see an animation of the impact of a column on the other

visualizations.

**Preparation before Problem Solving-**

**Steps to create data module are given below** :-

Step 1:- Open **Homepage** of IBM Cognos.

Step 2:- Click on the **Menu Option** at the top left side and then click on new option.

Step 3:- Now upload the data.

Step 4:- Now select the **Template** and then click on OK.

Step 5:- Select the Table.

Step 6:- Make **Relationships** between the tables.

Now, the data module is ready.

Now to solve the case study we need to form a new table named as Total Sales:-

**Steps to create a new table**:-

Step 1:- Click on **custom table**.

Step 2:- Now, select the table which you want to customize select **order detail and pizza**.

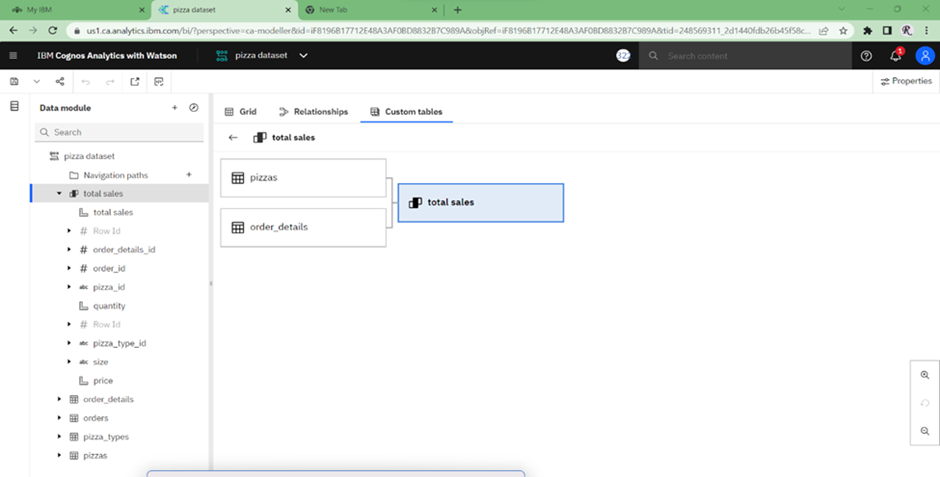
Step 3:- Select **joined view** and click on ok.

Step4:- Select the **matched column** and click on next.

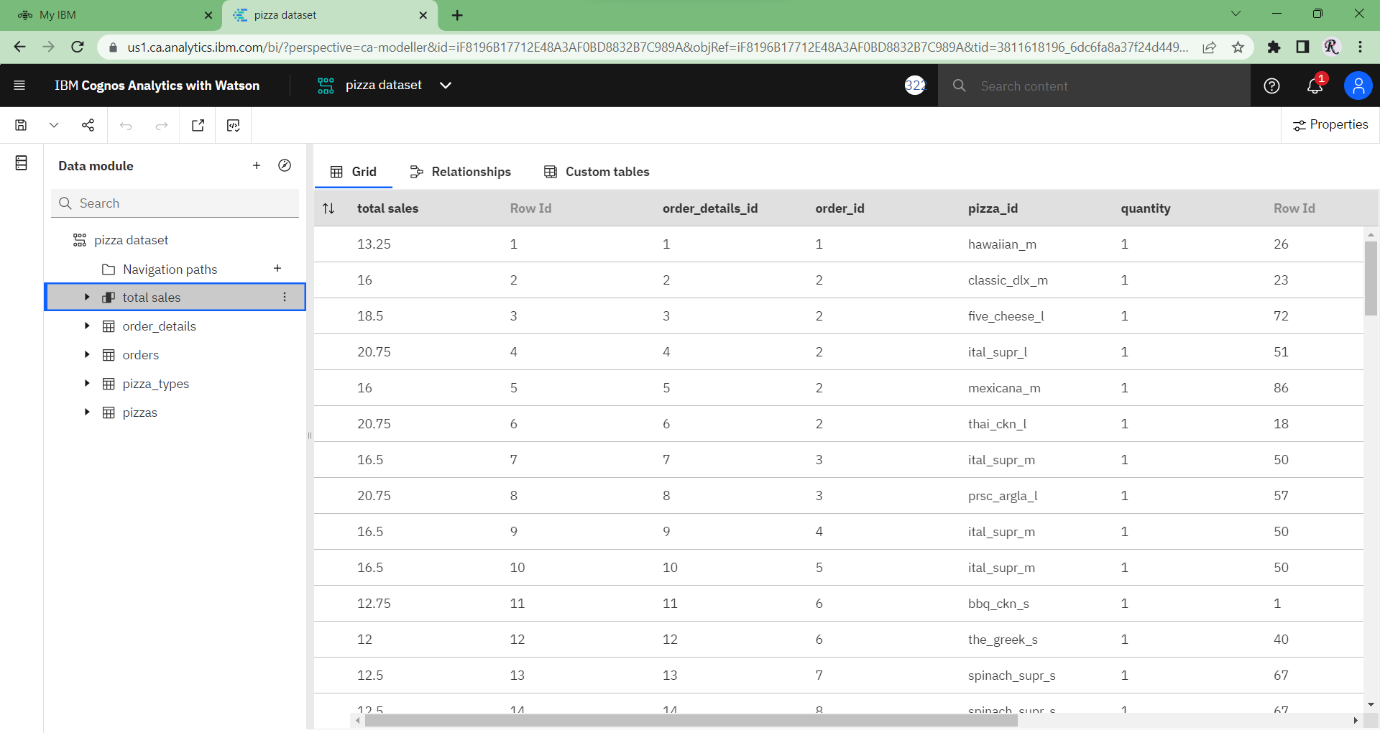
Step 5:- Name the table as **Total Sales** and save the table.

Step6:- Now, in source tab left click on total sales and select the calculation option.

Step7:- Now, type **Quantity** \* **Price** to get the **Total Sales**.



Our new table total sales is ready to use. Now, save this data module .

****

**Problem Statements -**

* What was the total revenue generated in 2015, organized by month and year?
* 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 years.
* Determine which pizza type was the most popular in the month of June.

**Solutions-**

1. To solve the first problem statement, we have chosen the column chart because we have to calculate the **Total Revenue** in each month of a particular year.

**Steps to solve the problem –**

Step1:- Go to new, then select the **Dashboard** option.

Step2:- Select the **Blank Template**

Step3:- Go to select source, select **Pizza Dataset**.

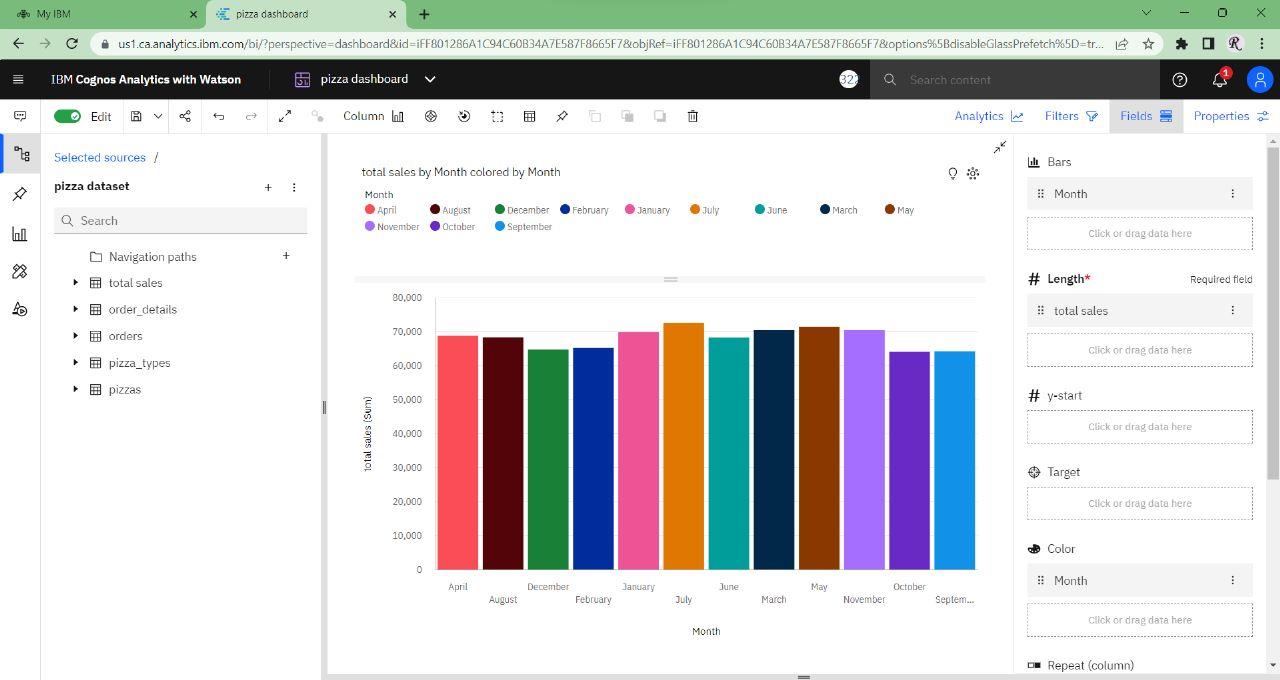
Step4:- Go to **Visualization Tab**.

Step5:- Select the **Column Chart,** drag and drop.

Step6:- As we need the total revenue generated organized by

month so we put, **Months** in **Bars** and **Total Sales** in **Length**.

Step7:- To show different Months we have choosed different colors.



July month have most sales in the year 2015.

**Total revenue generated in 2015, organized by month and year**. July month have most sale whereas October has least sale.

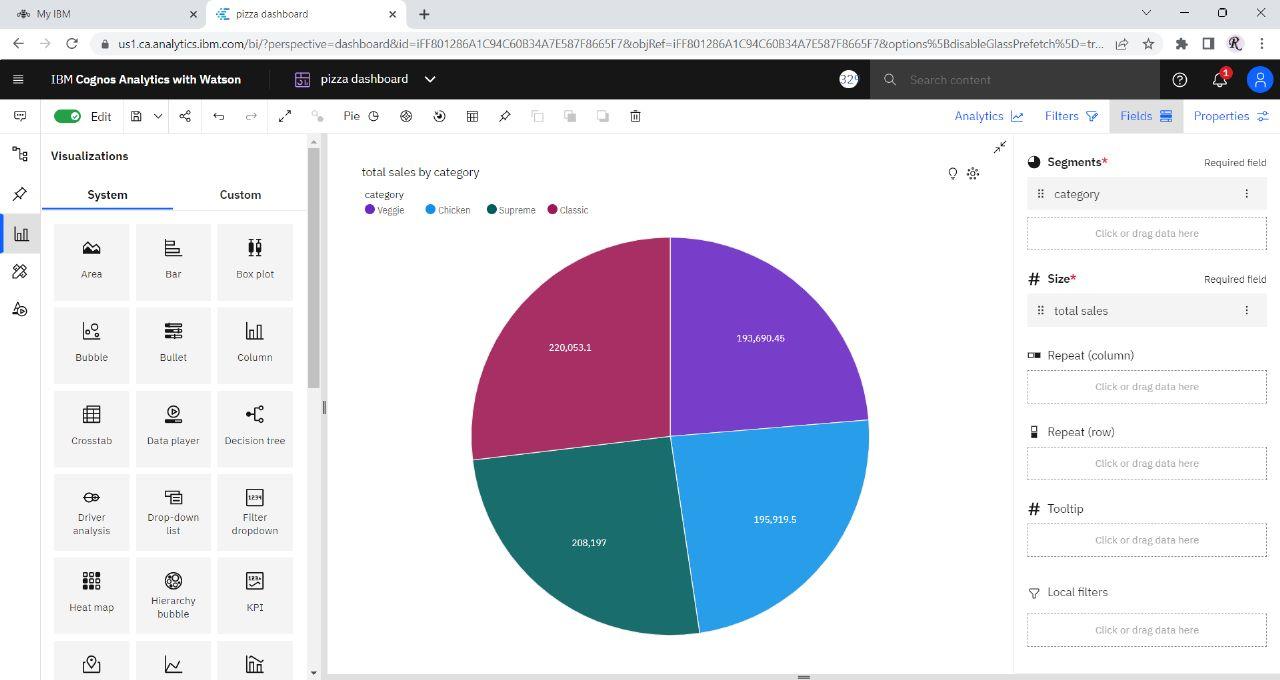
1. In this problem statement, we have to displays the contribution of each category of pizza sales therefore we use **Pie Chart**.

**Steps to solve the problem**-

Step1:- Go to **Visualization Tab**.

Step2:- Select the **Pie Chart,** drag and drop.

Step3:- Now to show the Total Sales by category we put, **Category** in **Segments** and **Total Sales** in **Size**.



**The contribution of each category of Total Pizza Sales**.

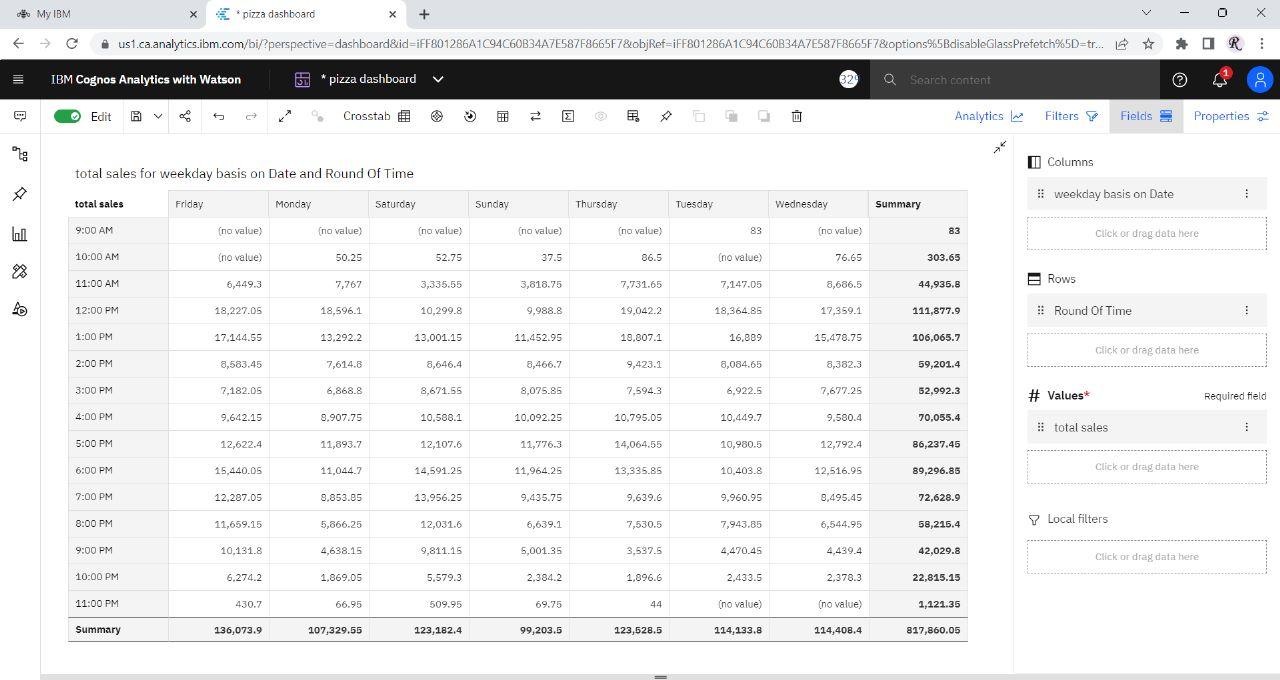
1. In this problem, we have to show the weekday with the highest pizza sales during an hour therefore we use **Cross Tab** to get the highest sales of pizza by comparing the weekdays and time**.**

**Steps to solve the problem –**

Step1:-Similarly, go to **Visualization Tab.**

Step2:- Select the **Cross Tab,** drag and drop.

Step3:- Now in fields, select **Weekdays basis on Date** in **Columns, Round of Time** in **Rows and Total Sales** in **Values.**

****

**Weekday with the Highest Pizza Sales during A particular hour.**

**Thursday has the most highest Pizza Sales around 12:00 PM**

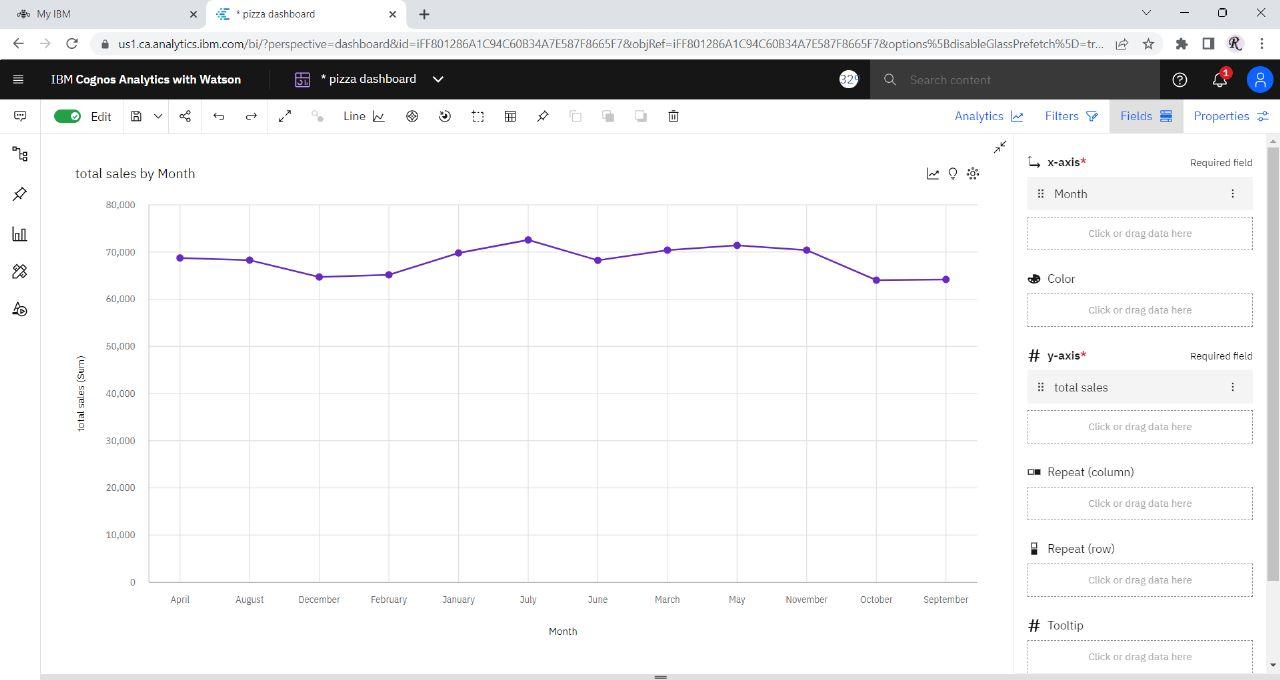
1. In this problem we have to show the fluctuation in total sales over the year.

**Steps to solve the problem –**

Step1:- Go to **Visualization Tab.**

Step2:- Select the **Line Chart,** drag and drop.

Step3:- Now to show the fluctuation, put Month in X-axis, Total Sales in Y -axis.



**Fluctuation in Total Sales over the years**

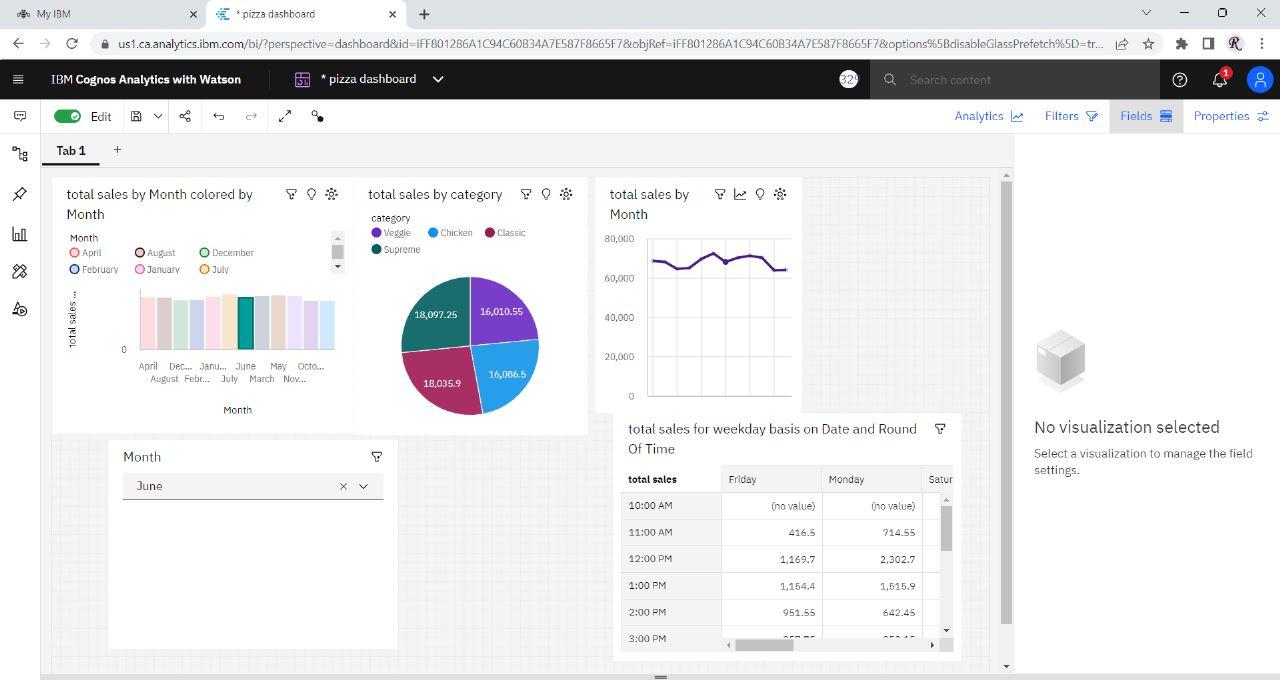
1. First of all, focus on the **Pie Chart** to determine which pizza type is most popular in the whole year. Now according to the problem statement, we only need which pizza type is most popular in month of June so we use **Dropdown.**

**Steps to solve the problem –**

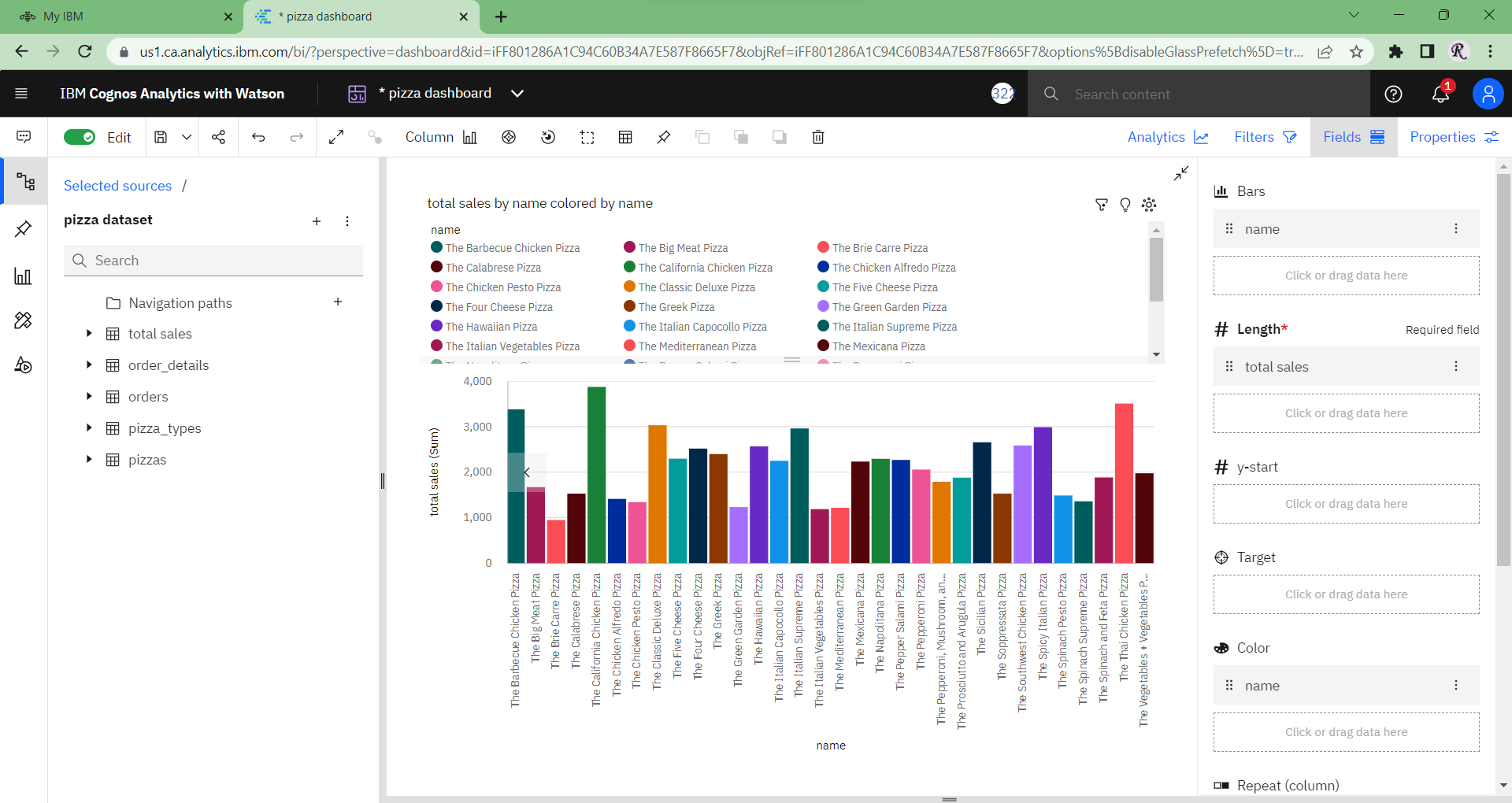
Step1:- Go to **Visualization Tab.**

Step2:- Select the **Filter Dropdown,** drag and drop.

Step3:- Now in fields, put **Month** in Column.



**Supreme Pizza is the most famous pizza on the basis of categories in the month of June.**



**The California Pizza is the most famous Pizza on the basis of name wise in the month of June.**

**Conclusion-**

This **Pizza Dataset** is all about the **Sales** of Pizza. It contains the sales information, customer details like which type of Pizza is the most famous one. By the help of this data, and the graphs they can understand the behaviour of the business whether they are performing well or not and if not, then company can get a huge amount of profit as now onwards, they know that in which field they have to work hard, they can take better decision with the help of this data and can aquire their organizational goals.