

[

LAB

EXPERI

MENT

]

[LAB EXPERIMENTs]

Data Warehouse

November 6, 2023

GLA University

Name- Oshin Rathor

Class-Section V

University roll no.-2115200016

Roll no.-14

Course-B.Tech CS(DA)

Year-2023

**EXPERIMENT-01**

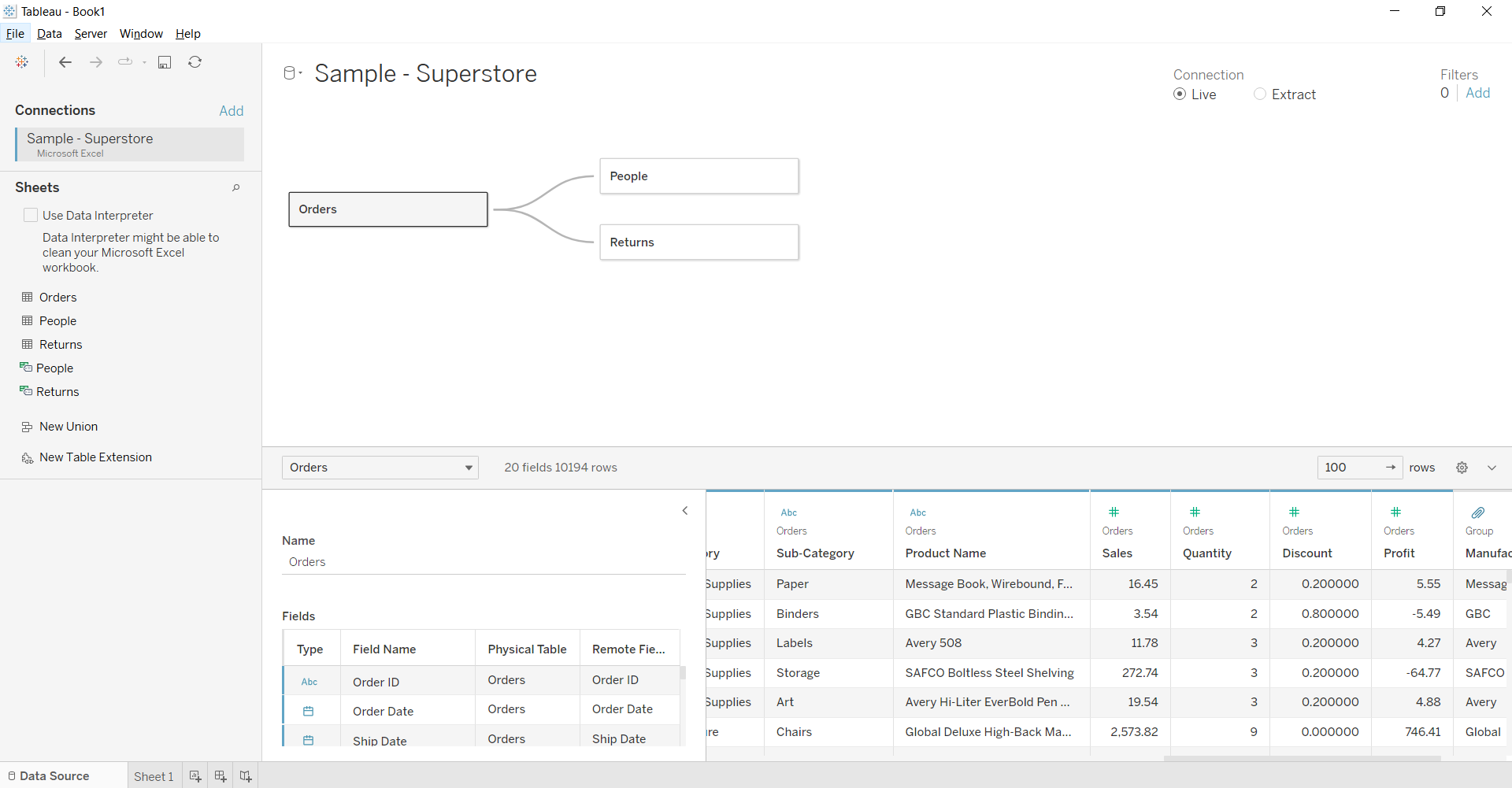
**1. Exploratory Data Analysis (EDA):**

o Import a dataset into Tableau and perform EDA to explore data distribution, correlations, and outliers. o Create various types of charts and visualizations to gain insights into the dataset.

A general outline of how to perform Exploratory Data Analysis (EDA) using Tableau. Please note that the specific steps may vary depending on your dataset and analysis goals.

Assuming you have Tableau Desktop installed and your dataset ready, here's how you can perform EDA:

1. **Connect to Your Dataset:**
   * Open Tableau Desktop. o Click on "Connect to Data" and select your dataset. Tableau supports various data formats such as Excel, CSV, SQL databases, etc.



1. **Data Source Tab:**
   * In the Data Source tab, you'll see your dataset's structure.
   * Examine the columns and their data types. Ensure that Tableau has correctly identified dimensions (categorical) and measures (numeric).
2. **Drag and Drop Dimensions and Measures:**
   * In the Data Source tab, drag and drop dimensions and measures onto the "Rows" and "Columns" shelves. o For example, you can drag a numeric column (measure) to the Columns shelf and a categorical column (dimension) to the Rows shelf.



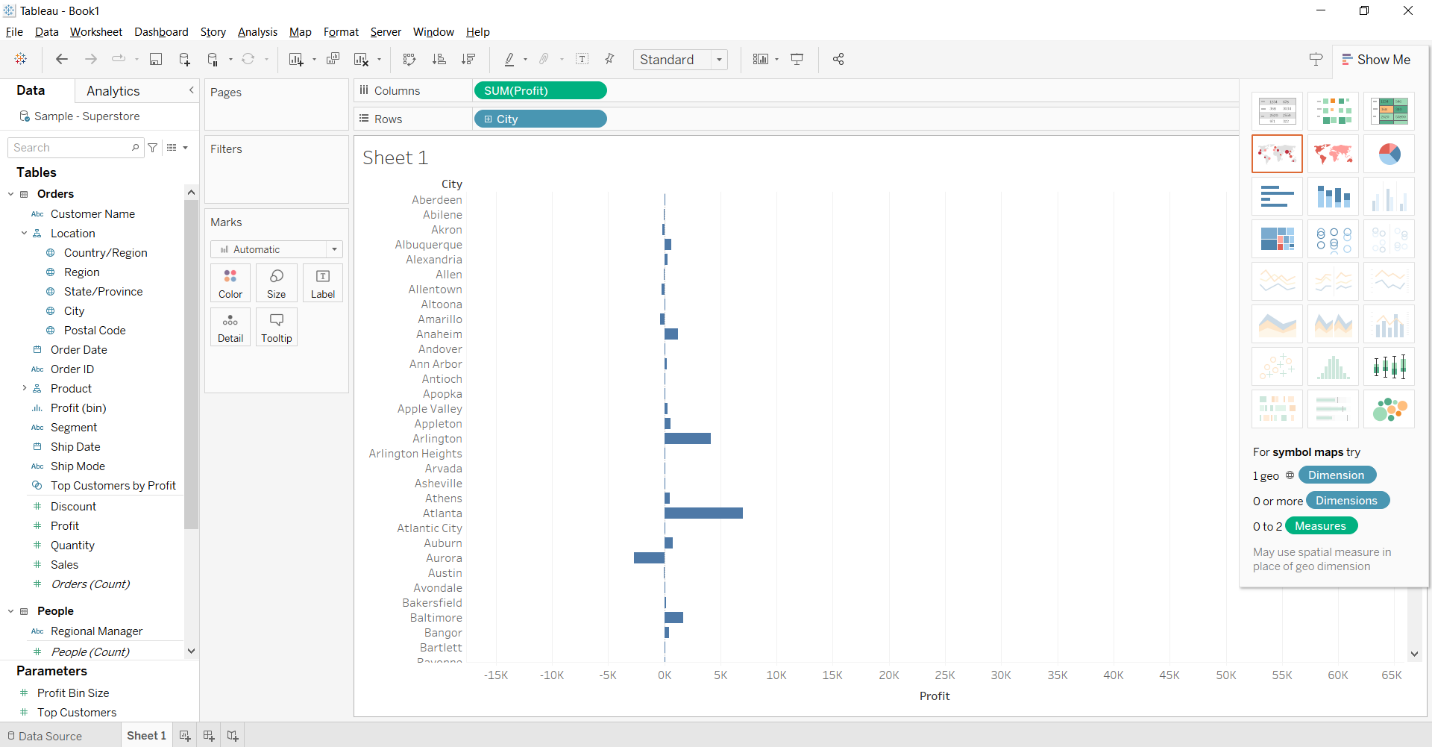
[

LAB

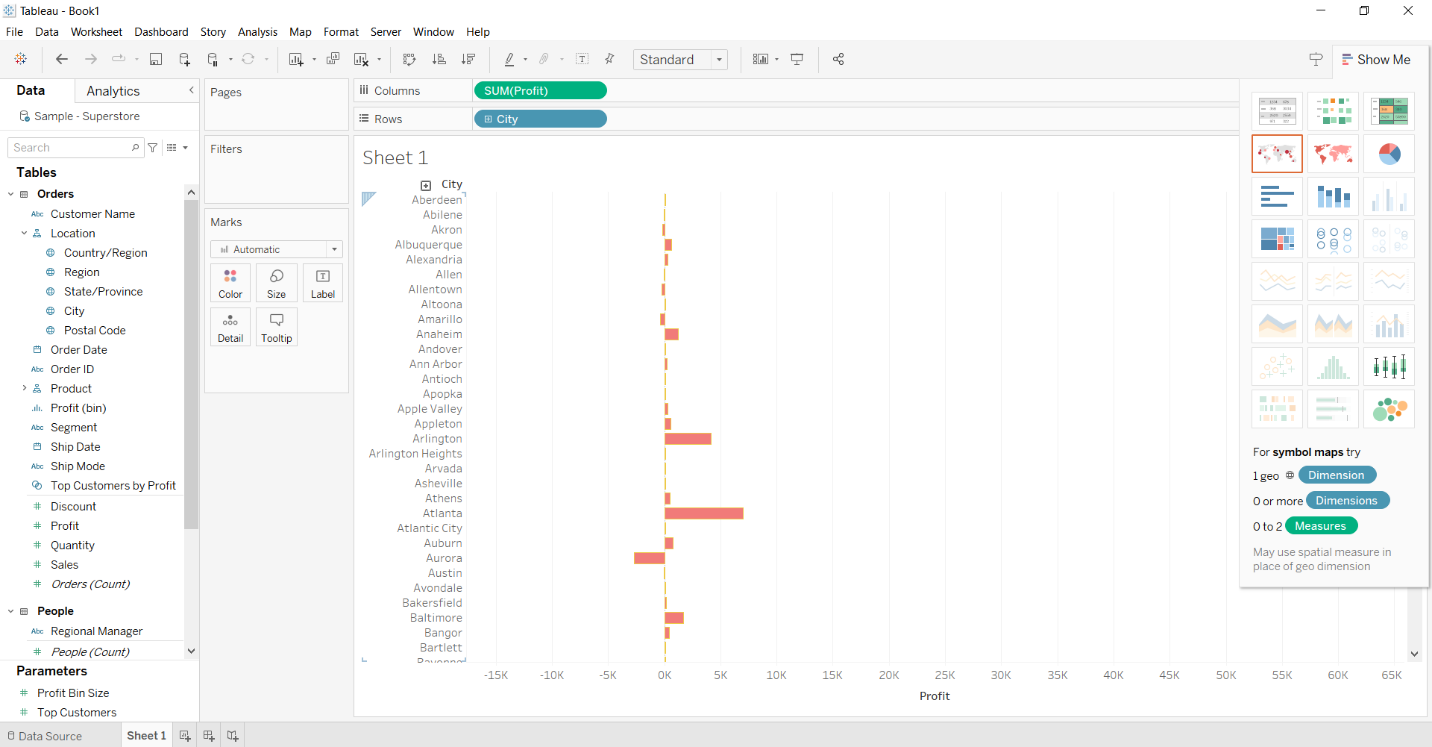
EXPERI

MENT

]

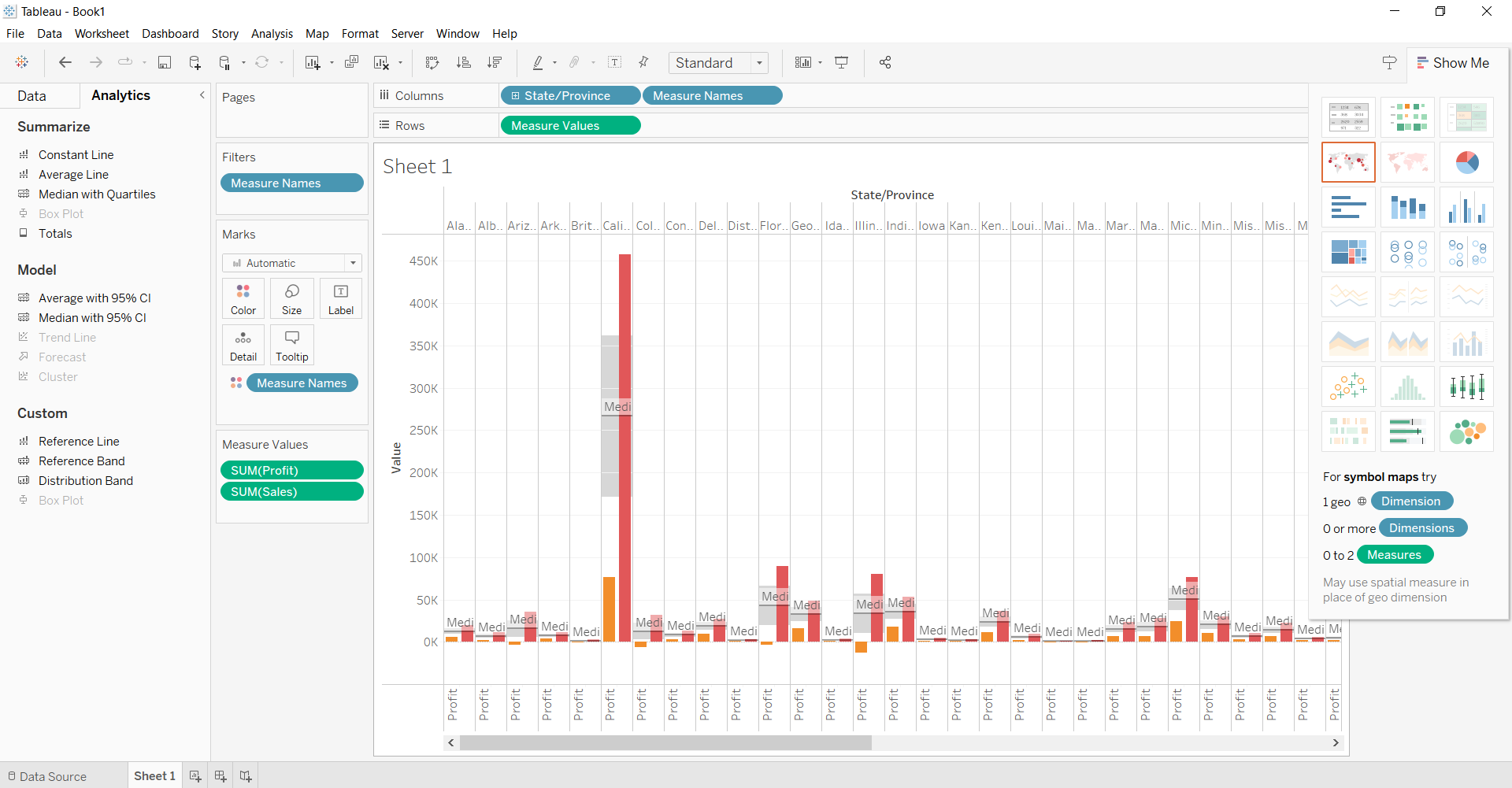


1. **Create Basic Visualizations:**
   * Tableau makes it easy to create various types of visualizations:
     + Drag and drop dimensions/measures onto the "Marks" card to choose the type of chart (e.g., bar chart, scatter plot, line chart).
     + Add additional dimensions to the "Color" or "Size" shelf to create more detailed visualizations.

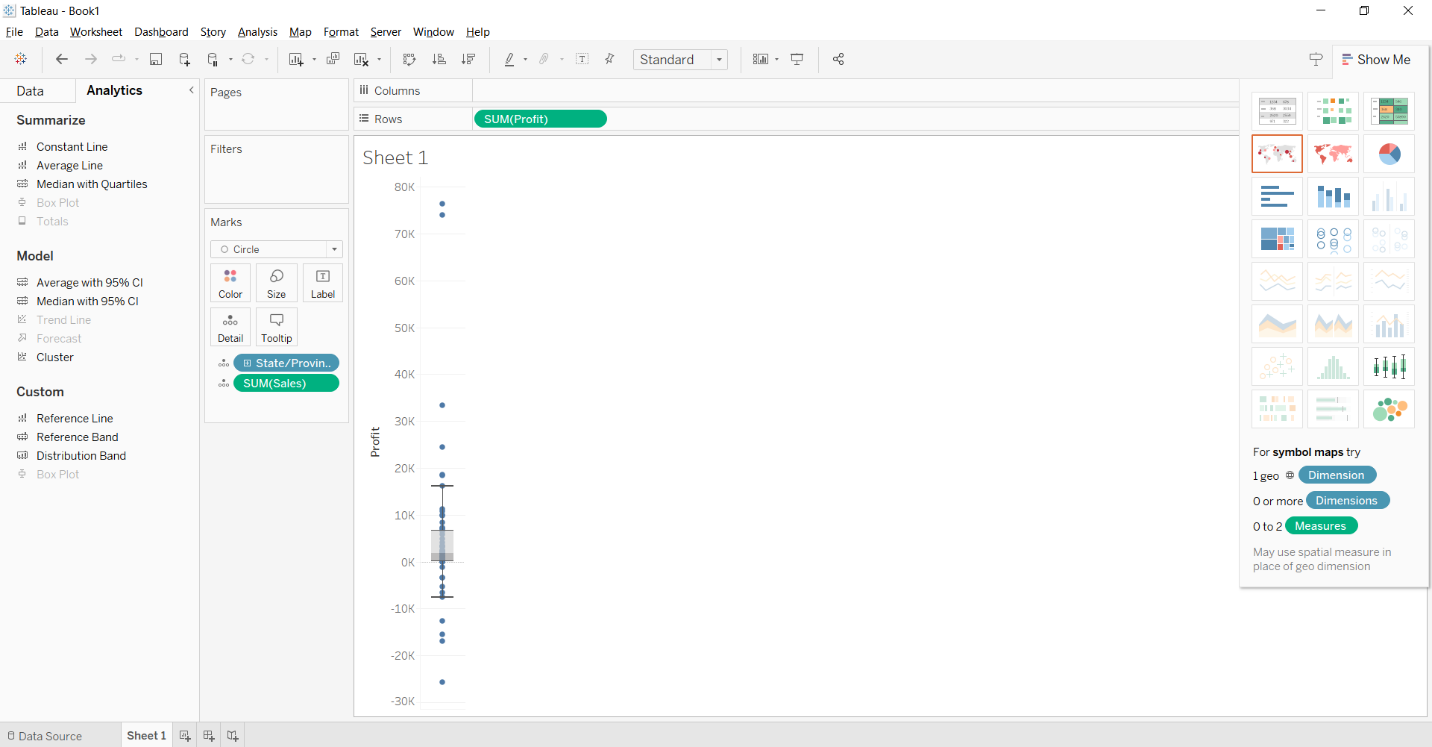


1. **Exploring Data Distribution:**
   * To explore the distribution of a numeric variable, you can:
     + Create a histogram by dragging a measure to the Columns shelf and selecting "Histogram" from the "Show Me" menu.
     + Calculate summary statistics like mean, median, and quartiles using the "Analytics" pane.

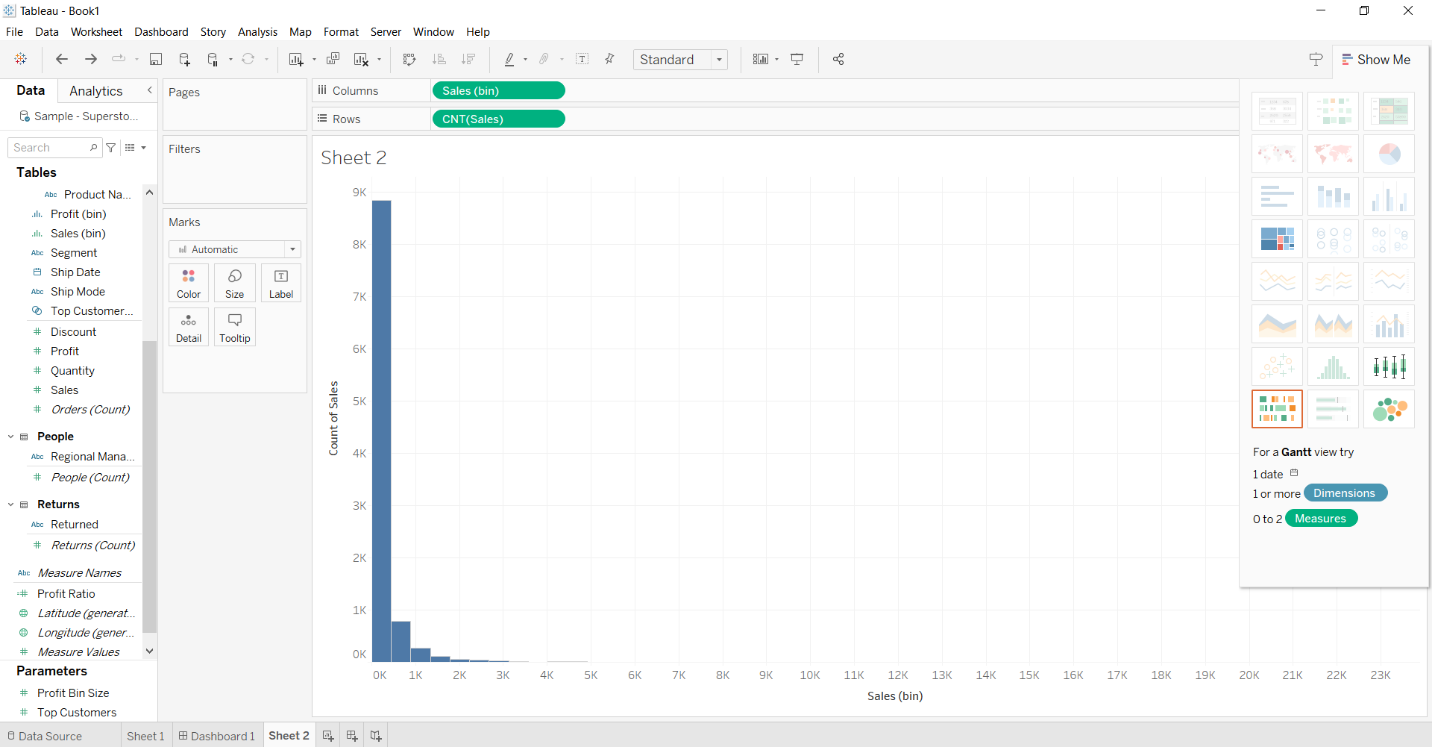
Median



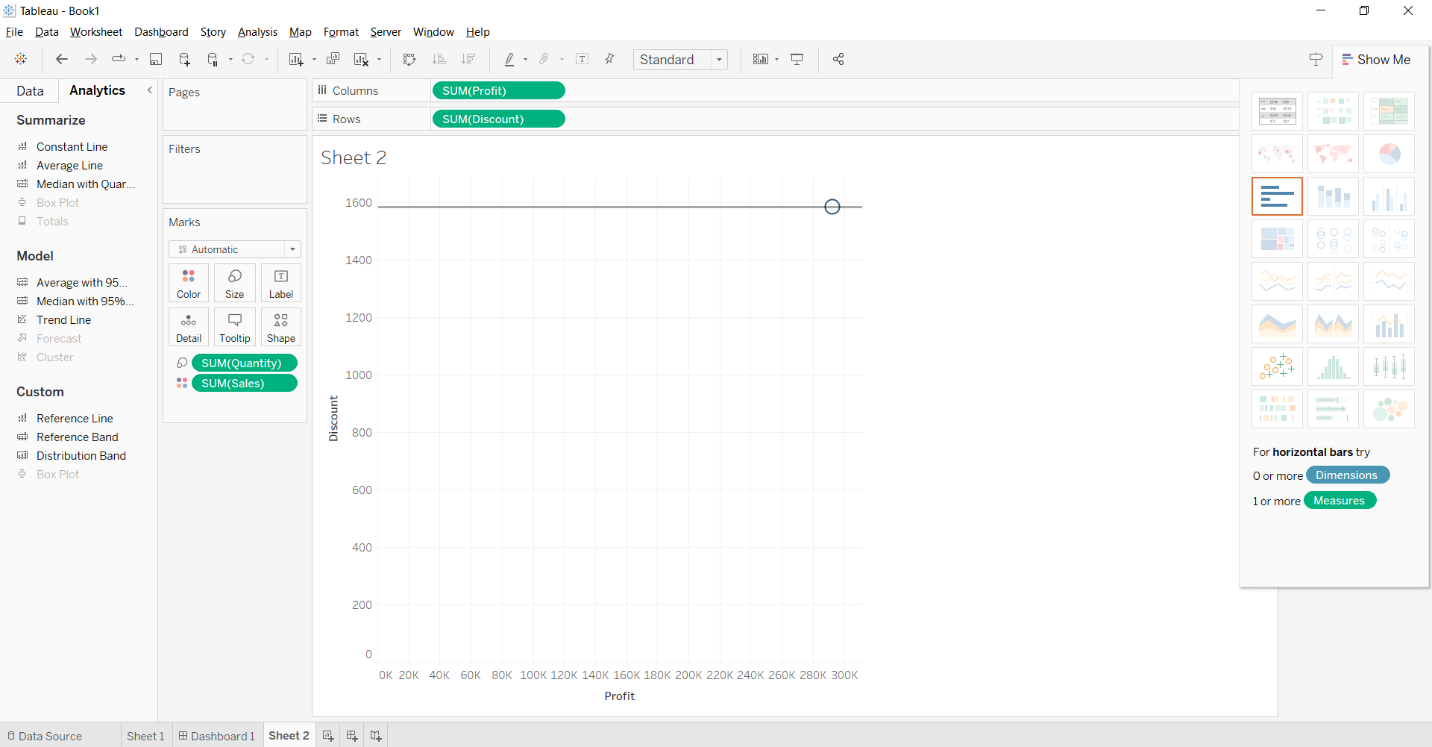
Quartiles



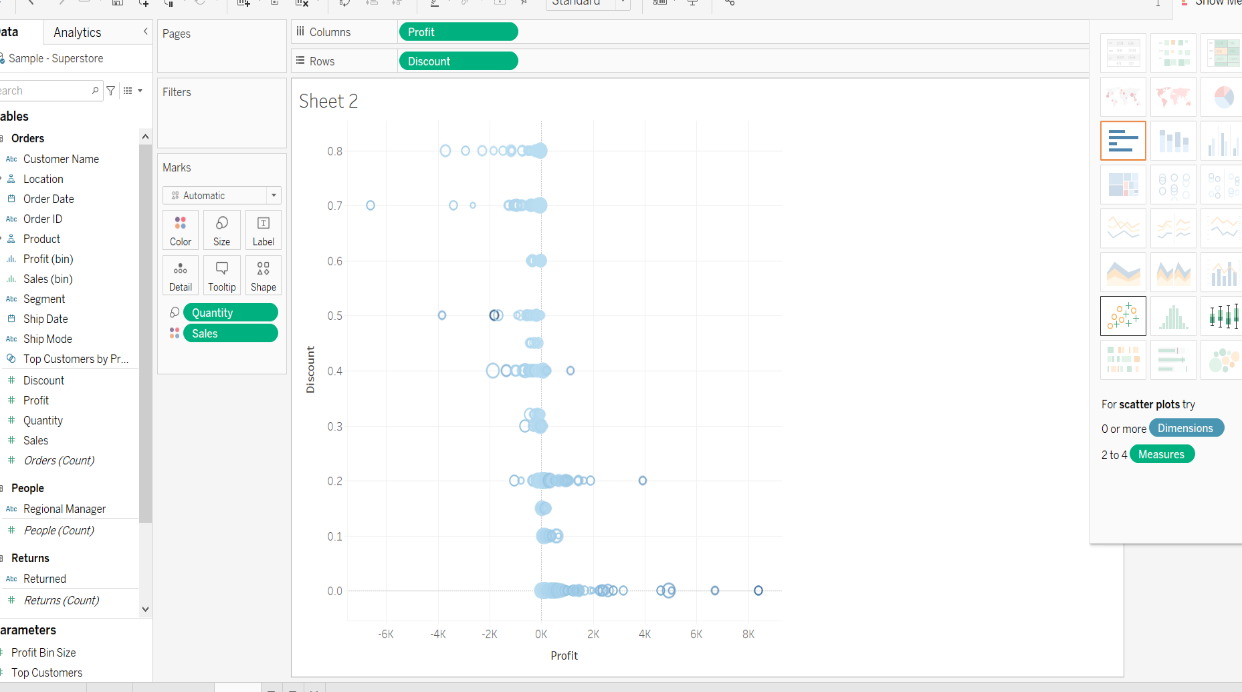
histogram



1. **Correlations:**
   * To analyze correlations between variables, you can:
     + Create a scatter plot to visualize relationships between two numeric variables.
     + Use the "Analytics" pane to add trend lines and correlation coefficients.

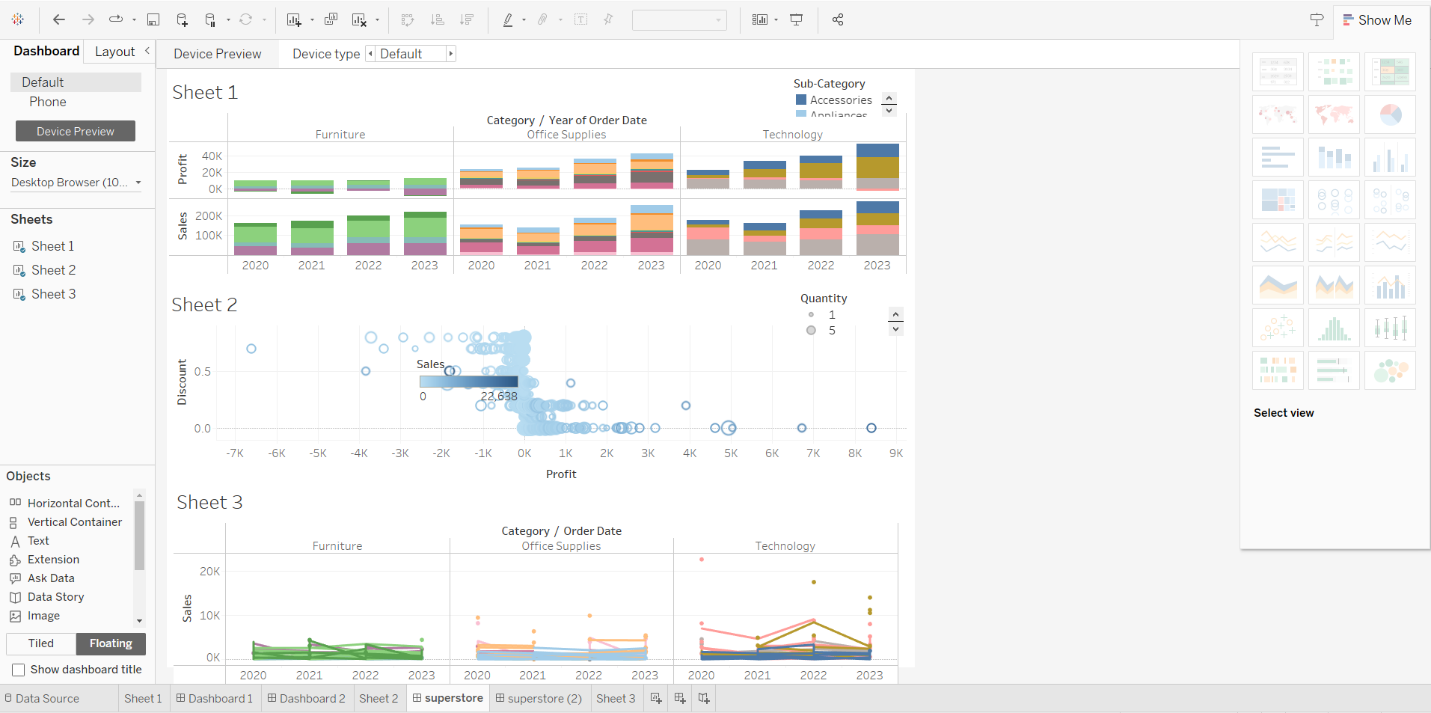


1. **Outliers:**
   * Identify outliers by visual inspection or using statistical techniques.
     + Create box plots, scatter plots, or histograms to detect extreme values.
     + You can also use Tableau's built-in clustering or distribution analysis tools.



[ Experiment -01 | ENTERPRISE BUSINESS INTELLIGENCE AND DATA WAREHOUSING LAB ] 1

1. **Dashboard and Story Creation:**
   * After creating individual visualizations, you can combine them into dashboards.
     + Use dashboards to provide an overview of your findings and allow for interactivity.
     + You can also create a story to guide viewers through your analysis step by step.



1. **Filtering and Drill-Down:**
   * Implement filters to allow users to interactively explore the data.
     + Add filters to your dashboard to enable users to drill down into specific subsets of data.
2. **Save and Share:**
   * Save your Tableau workbook. o Publish it to Tableau Server or Tableau Public if you want to share your findings with others.

[ Experiment -01 | ENTERPRISE BUSINESS INTELLIGENCE AND DATA WAREHOUSING LAB ] 2



[

LAB

EXPERI

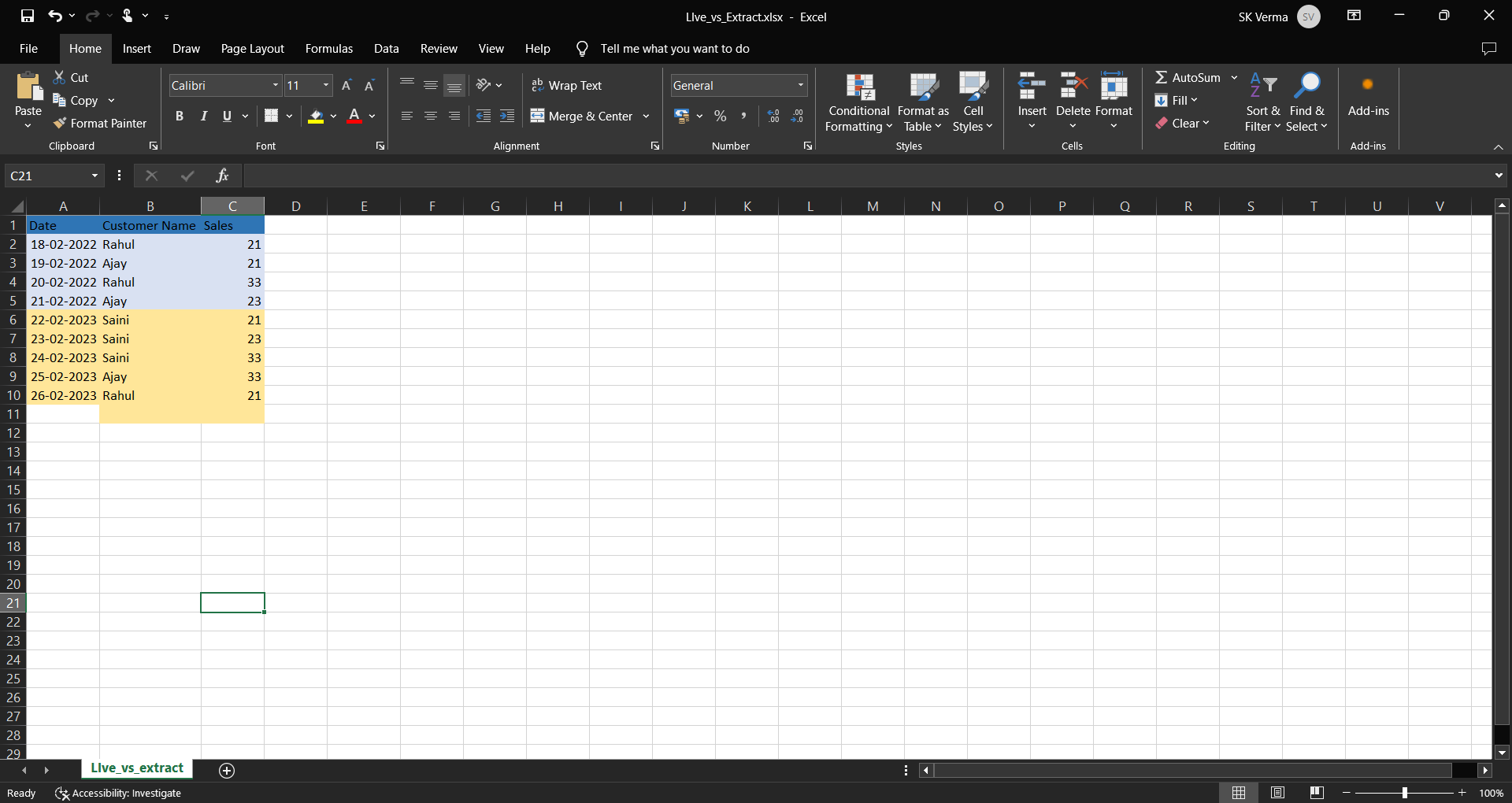
MENT

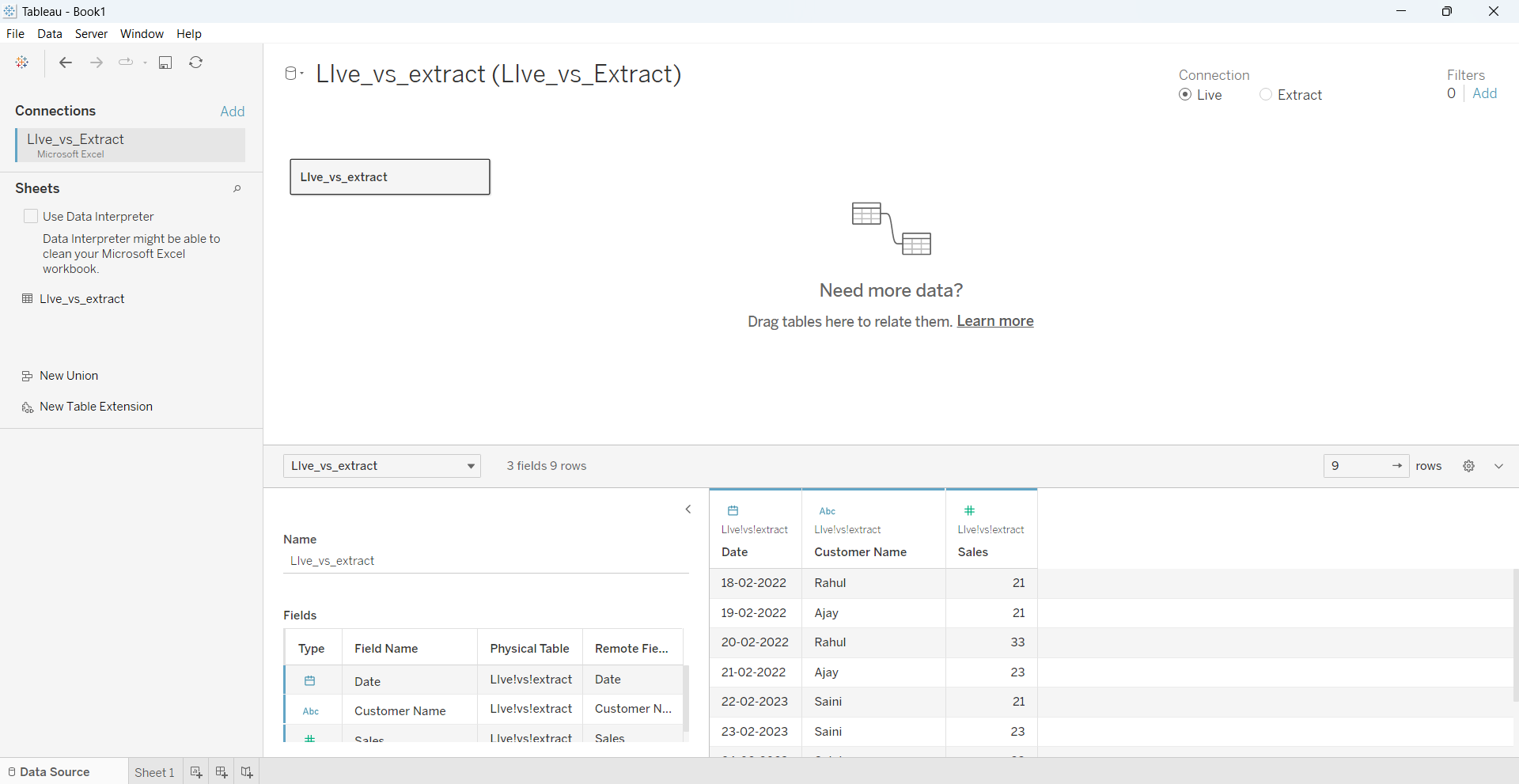
]

**EXPERIMENT – TWO**

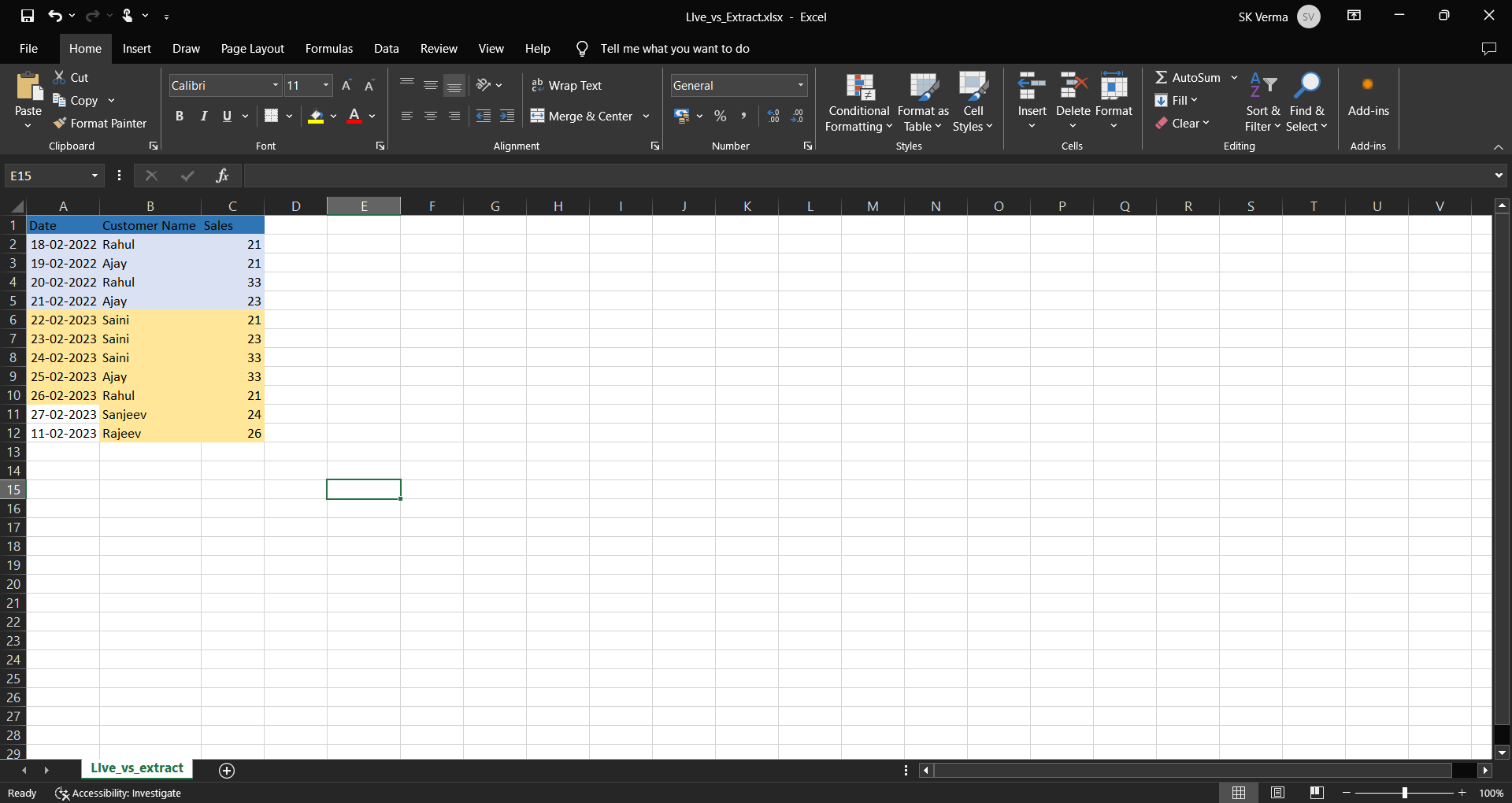
Objective: to perform data analysis on live vs extract dataset

1. CONNECT TO DATA SHEET



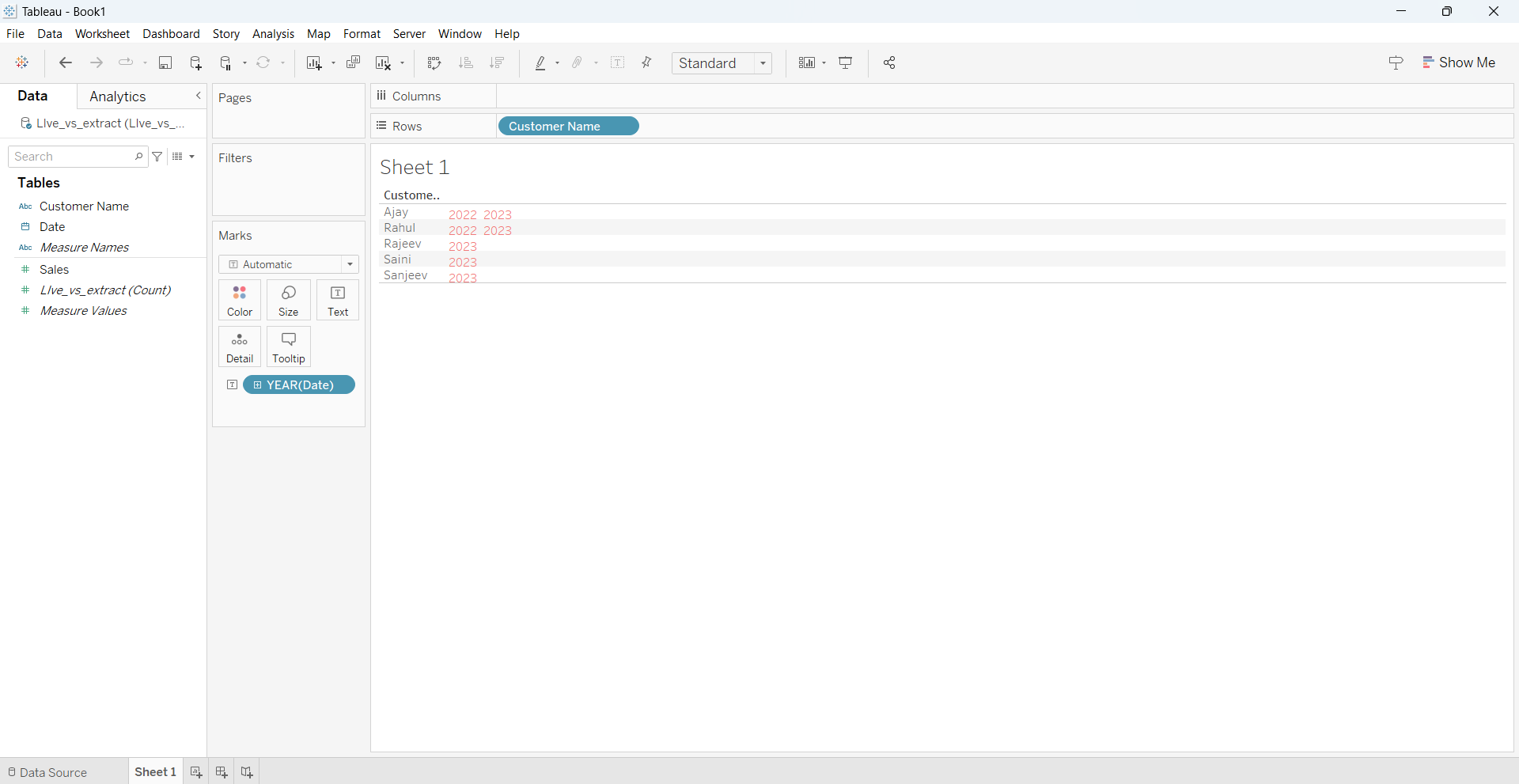


[ we have selected live connection so that we can do live changes direct from the excel to our data ]

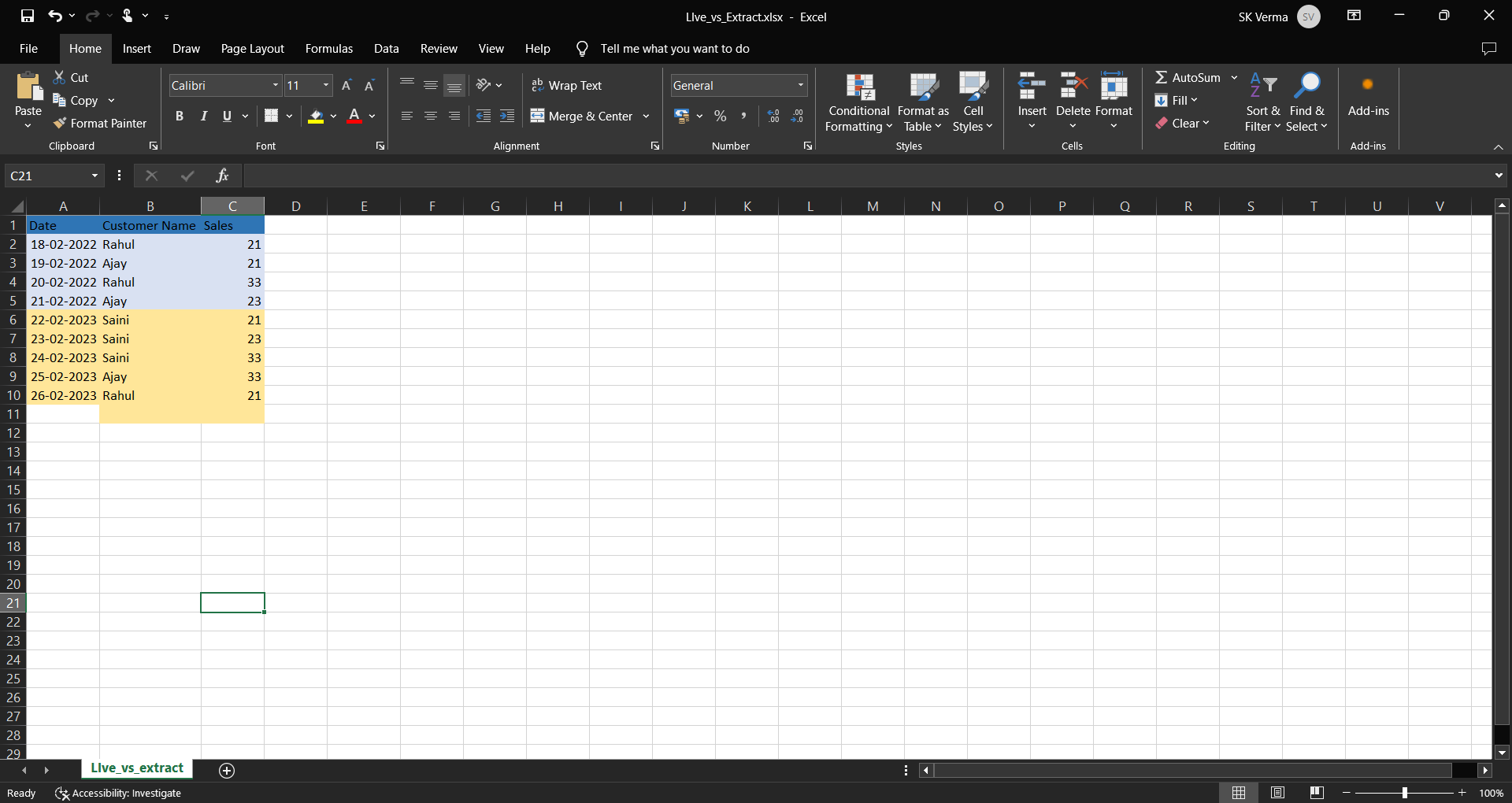
1. ADD NEW DATA

[ we have manually added new data in our excel sheet to see live changes ]

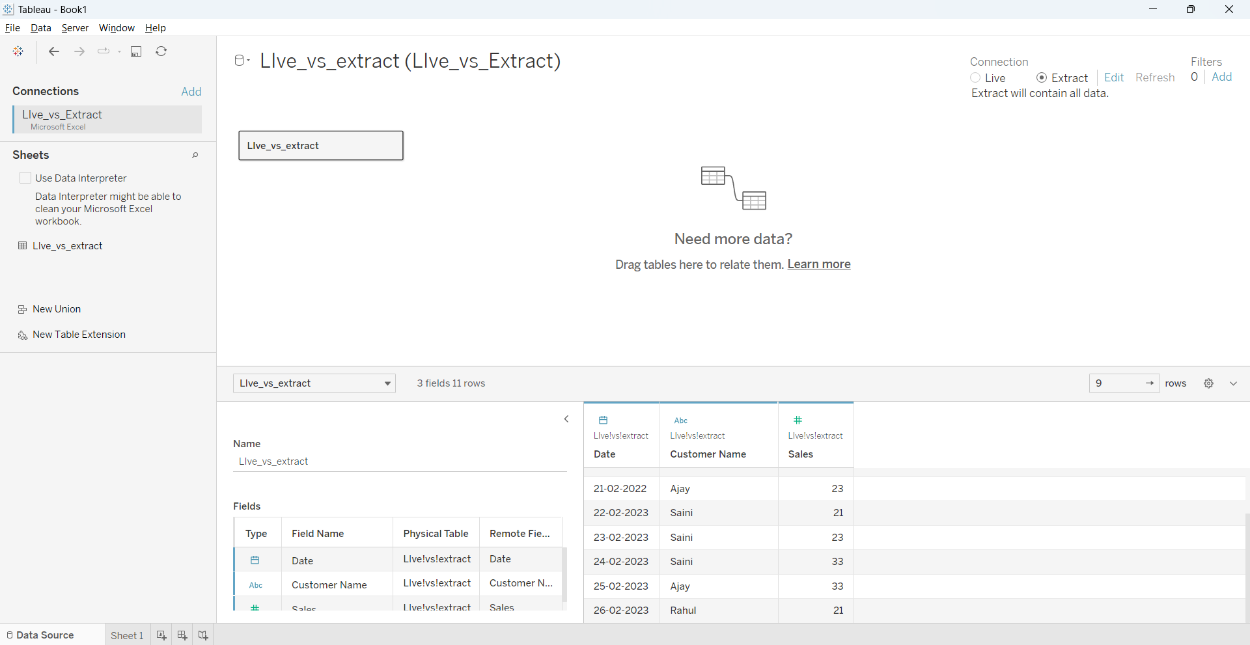
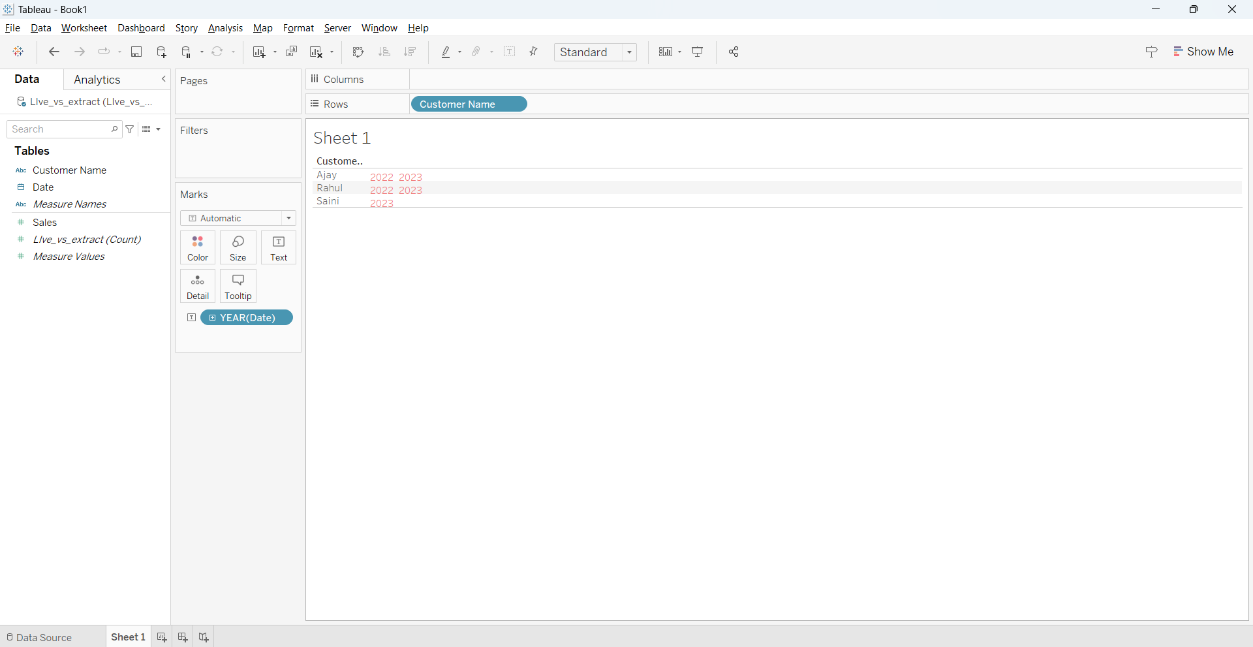
1. REFRESH TO LIVE CHANGE

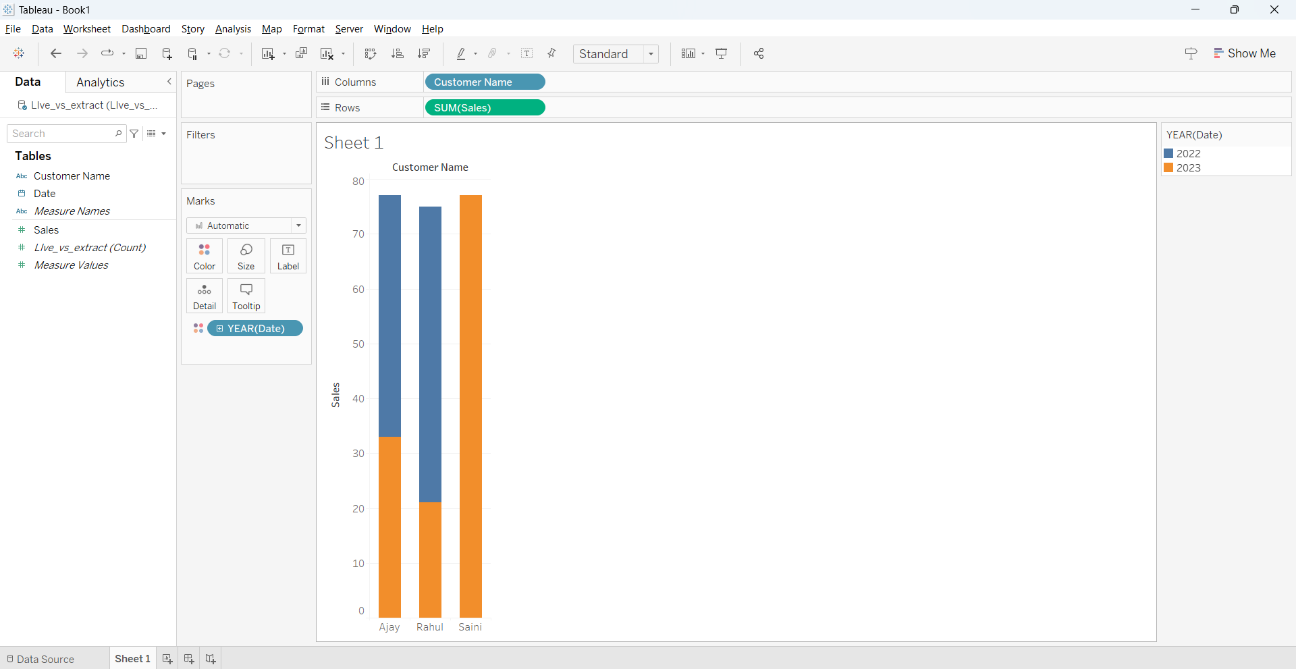


[ we have used REFRESH option to bring live changes made in our excel sheet to our tableau data ]

1. REMOVE THE DATA

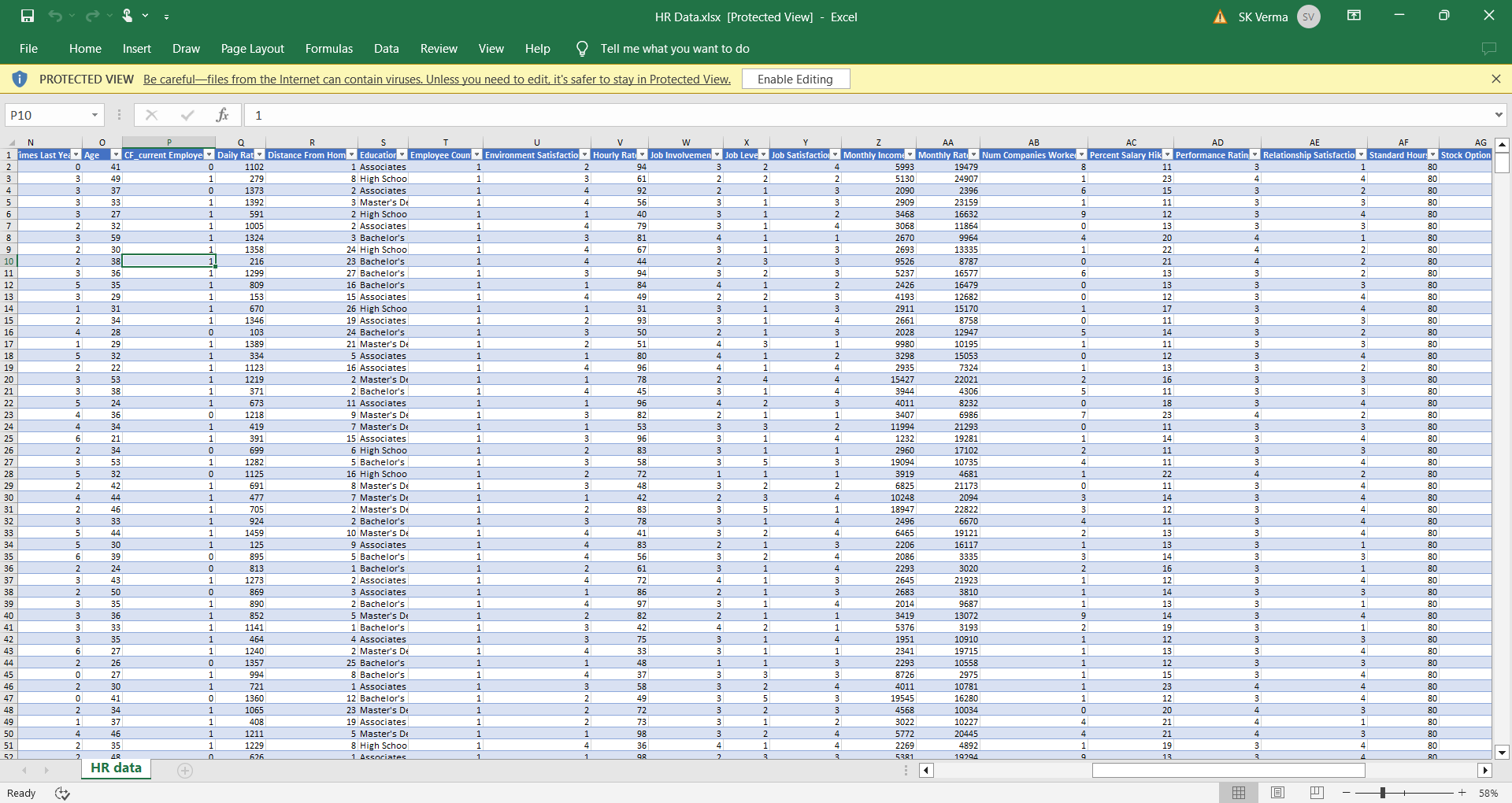
[ we have manually removed the data we added in our excel sheet ]

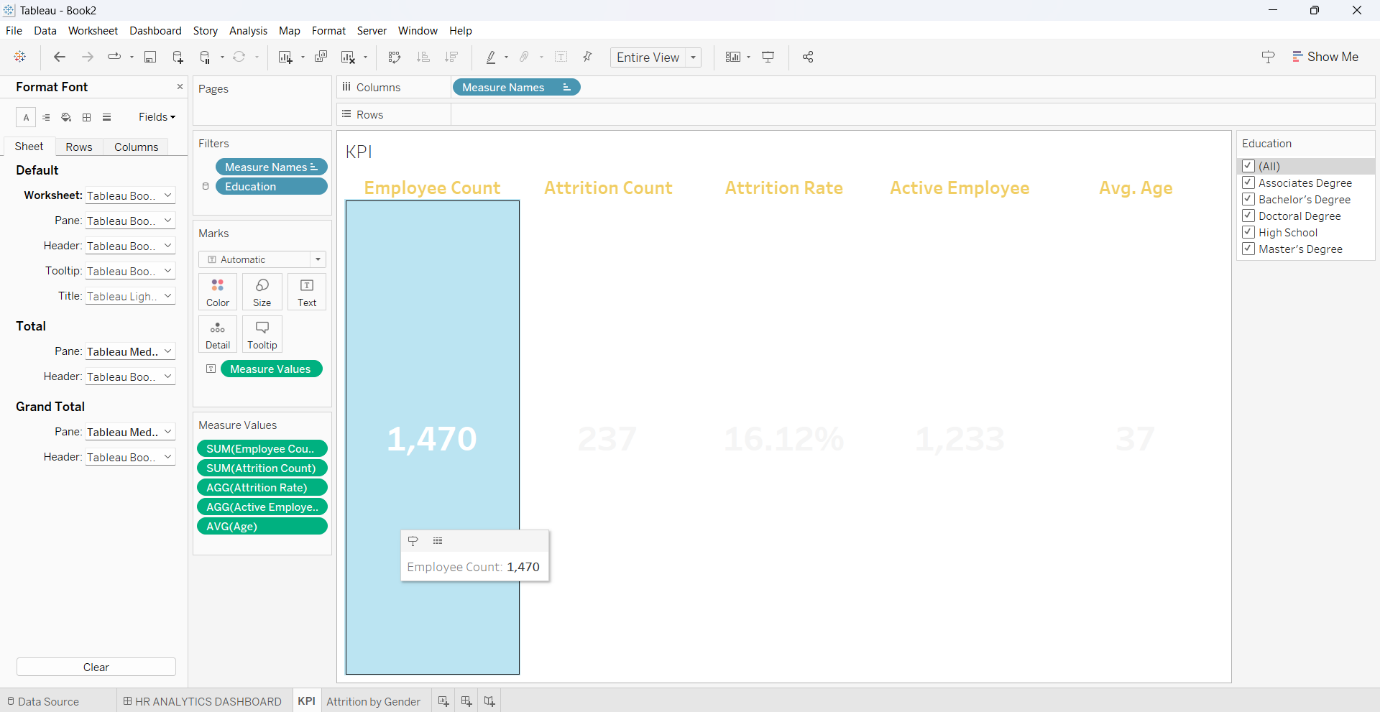
 [ we have selected extract connection to extract data we removed from our excel sheet]

1. SALES PER YEA

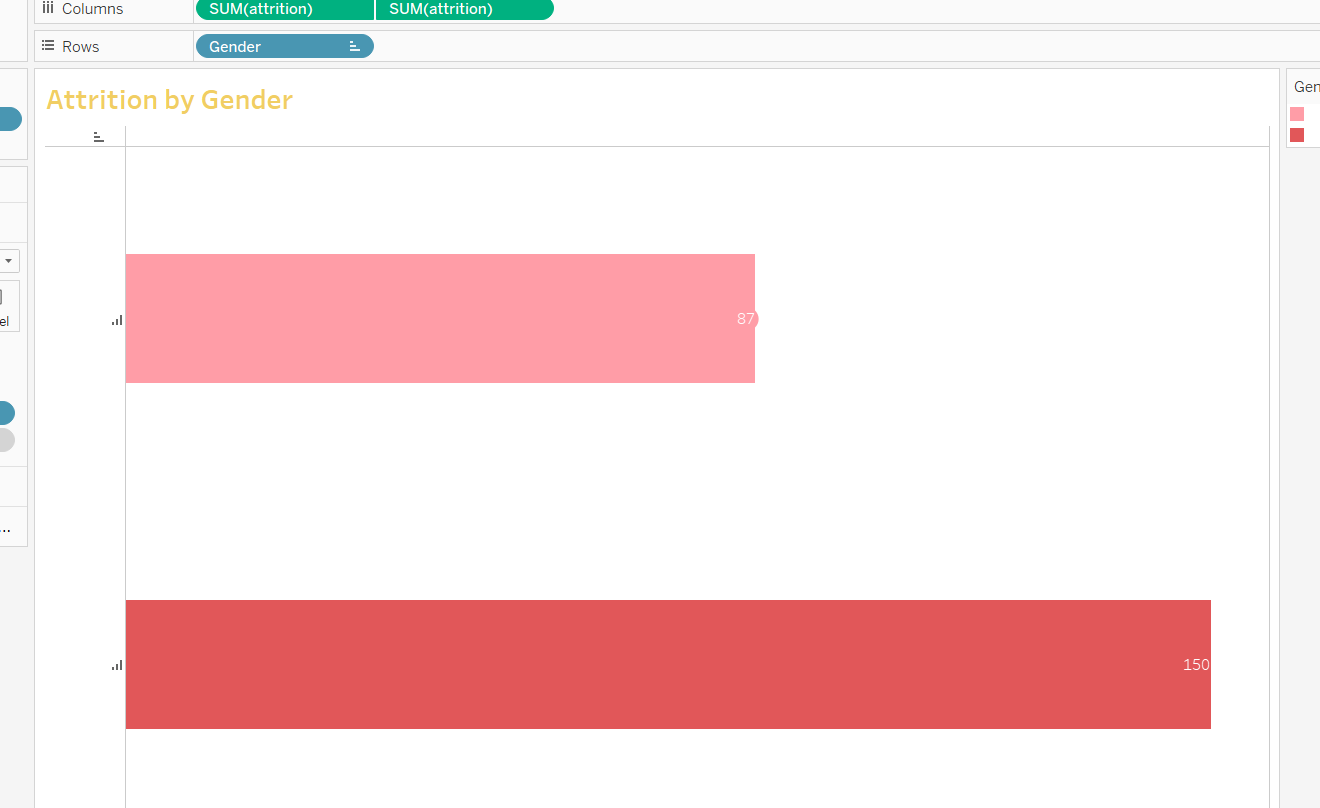
**EXPERIMENT – THREE**

**Create a Dashboard for HR Analytics**

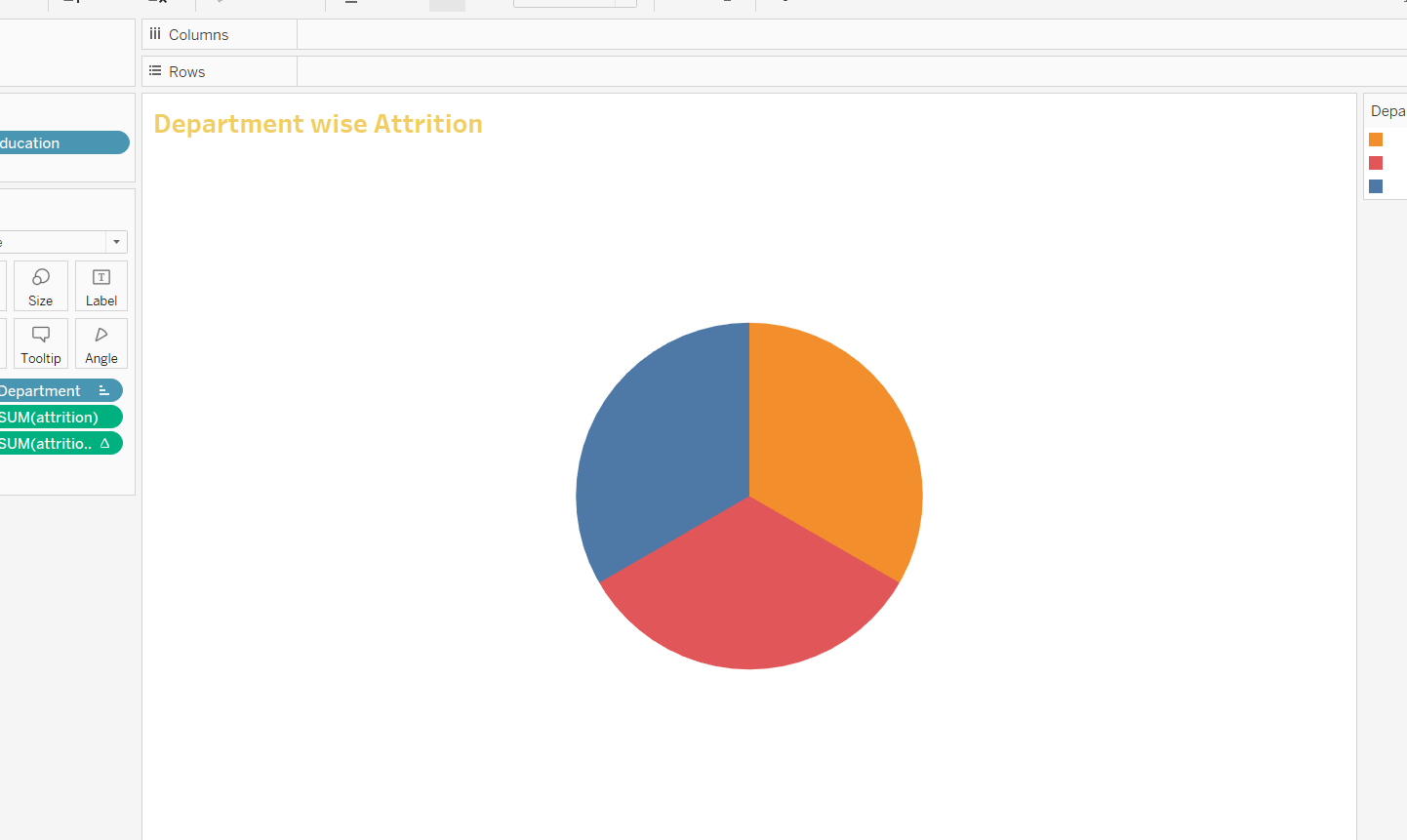
1. ****Data
2. KPI

****

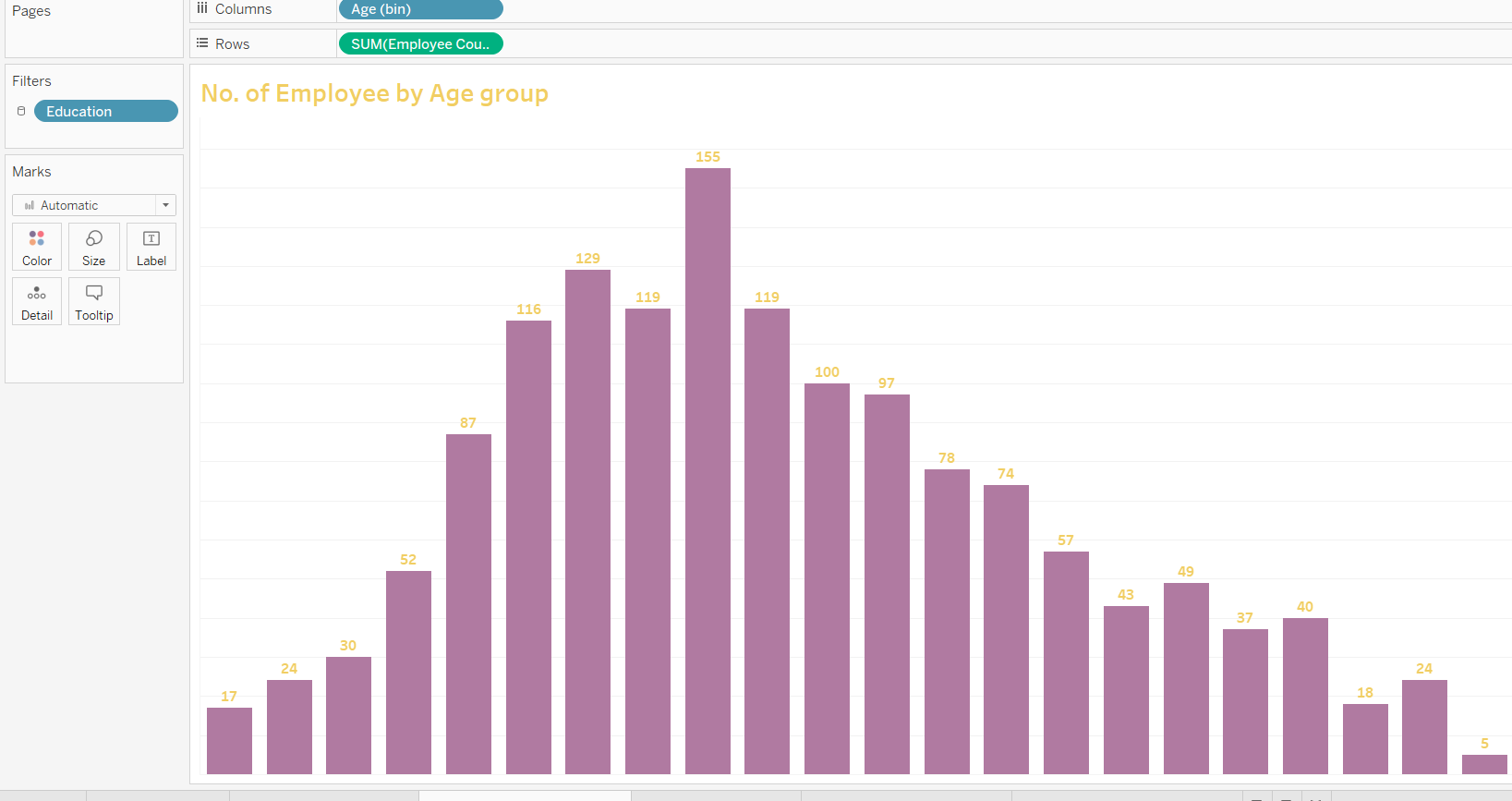
1. Attrition By Gender

****

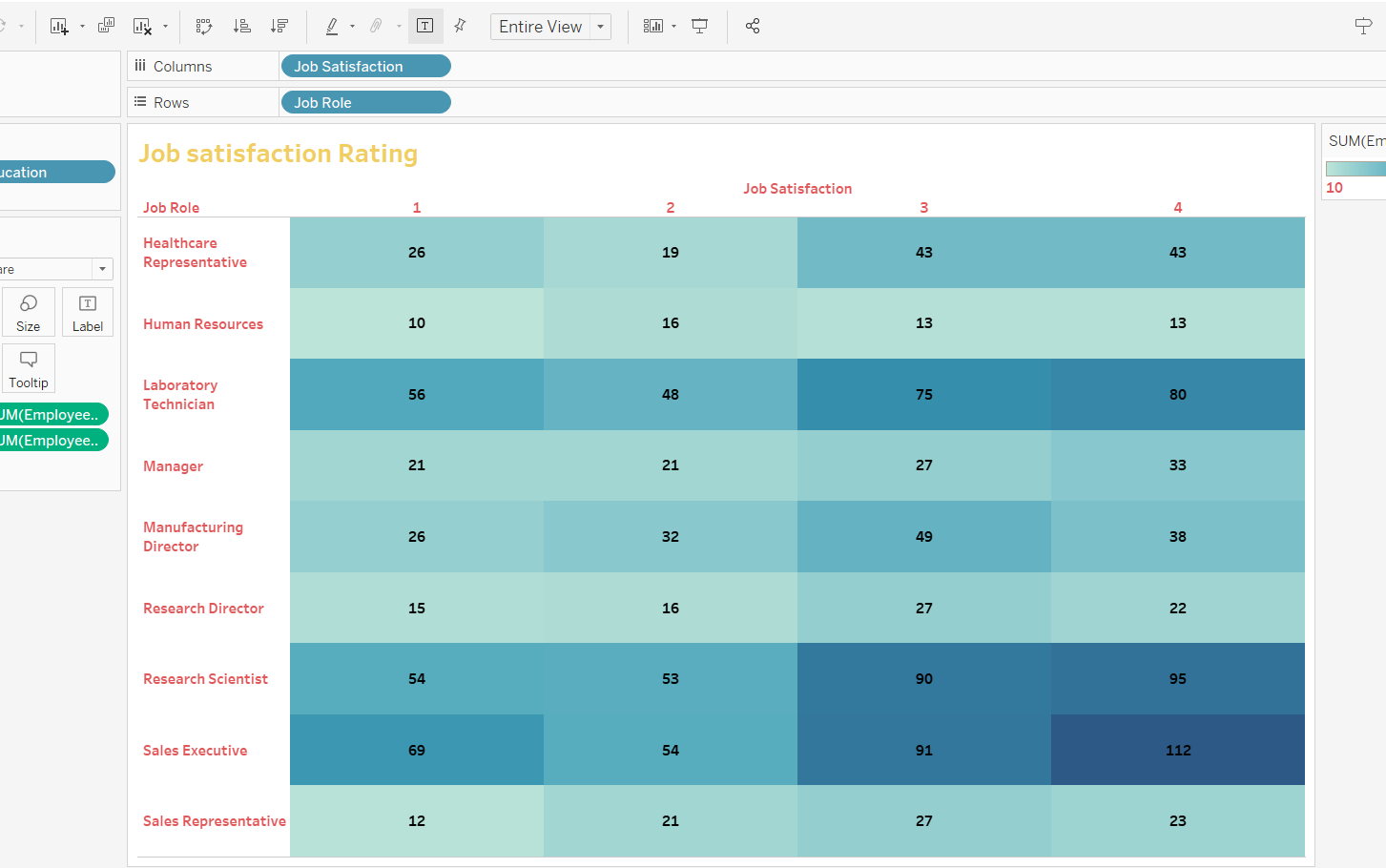
1. Department Wise Attrition

****

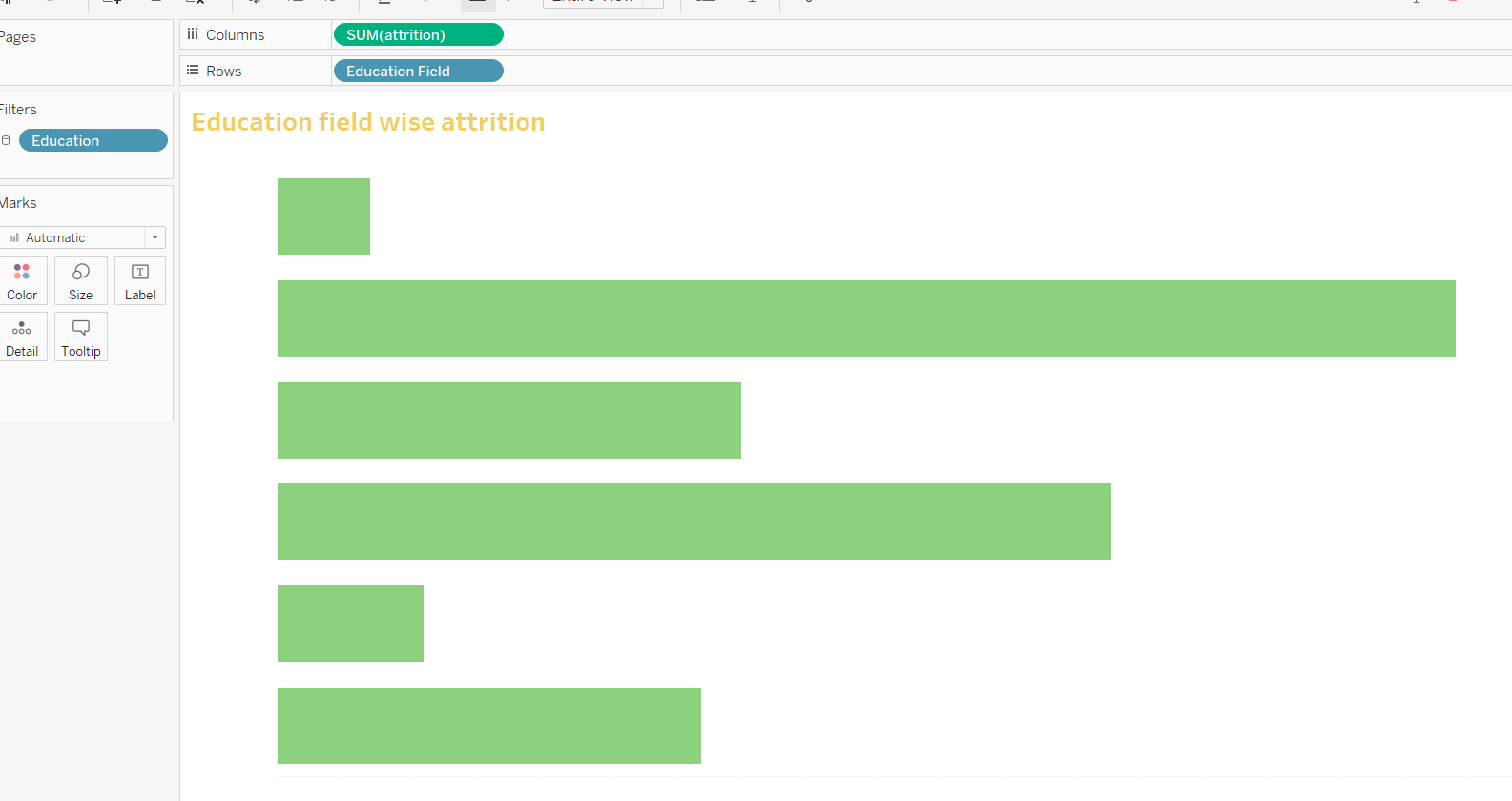
1. No. of Employee by Age Group

****

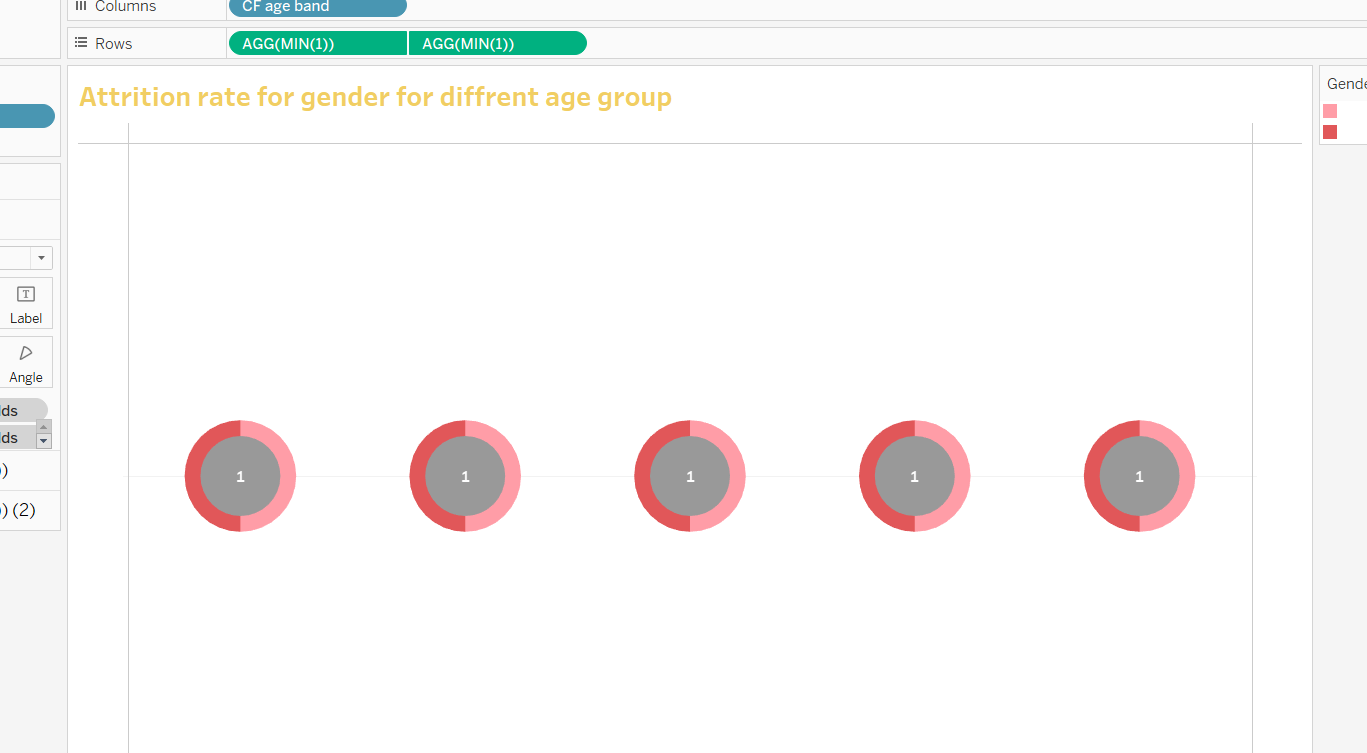
1. Job Satisfaction Rating

****

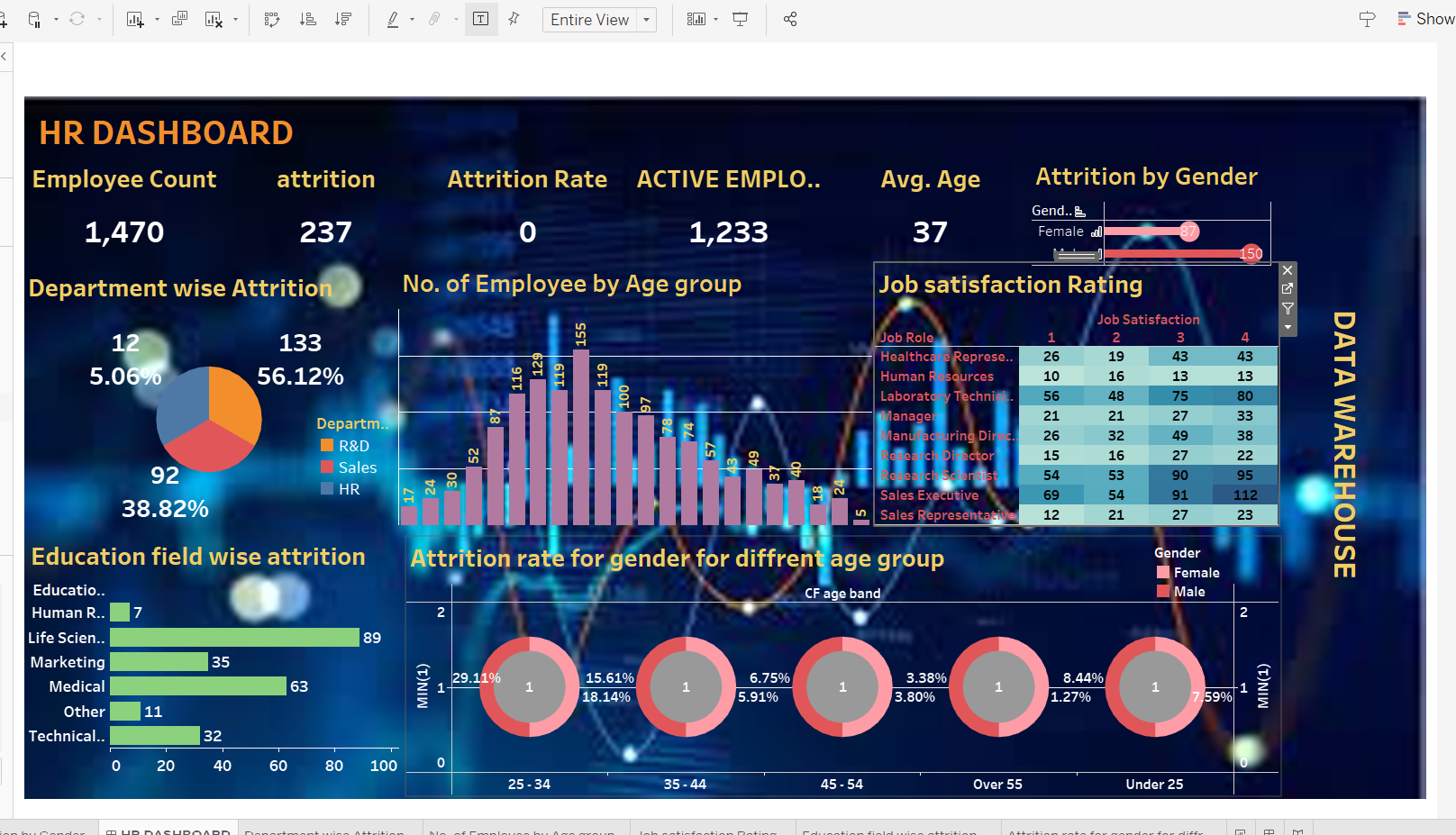
1. Education Field Wise Attrition

****

8.Attrition Rate by Gender for Different age Group



1. Dashboard

****

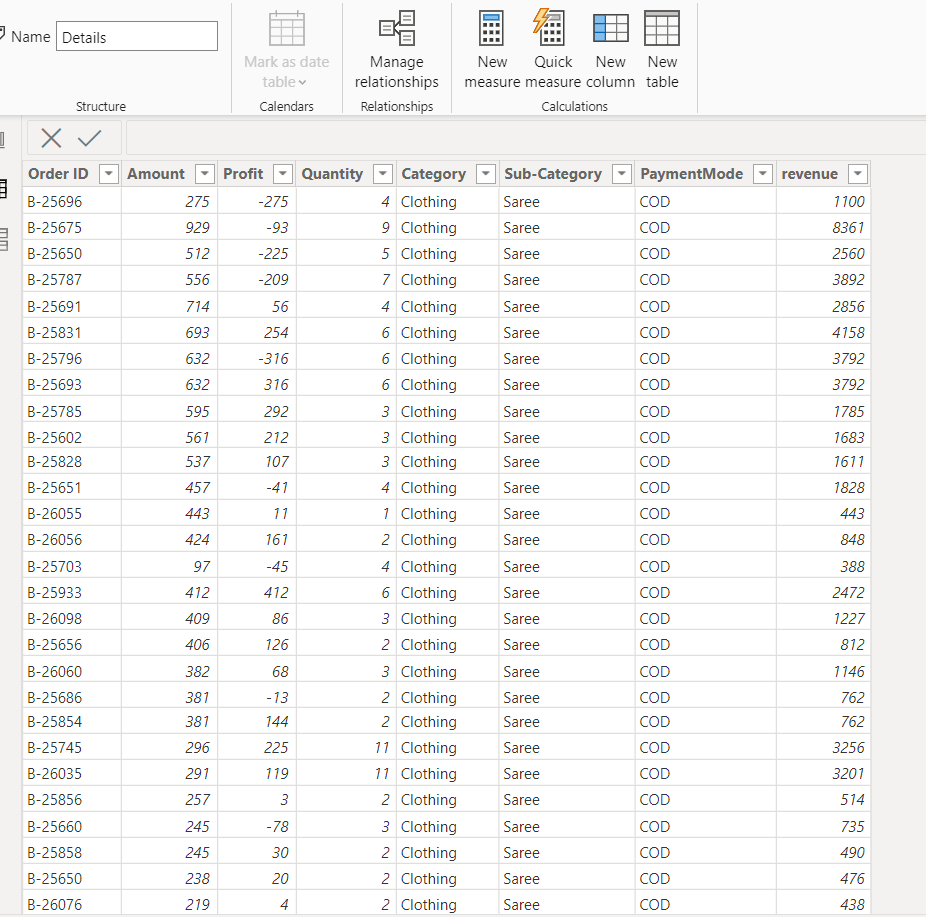
EXPERIMENT -FIVE

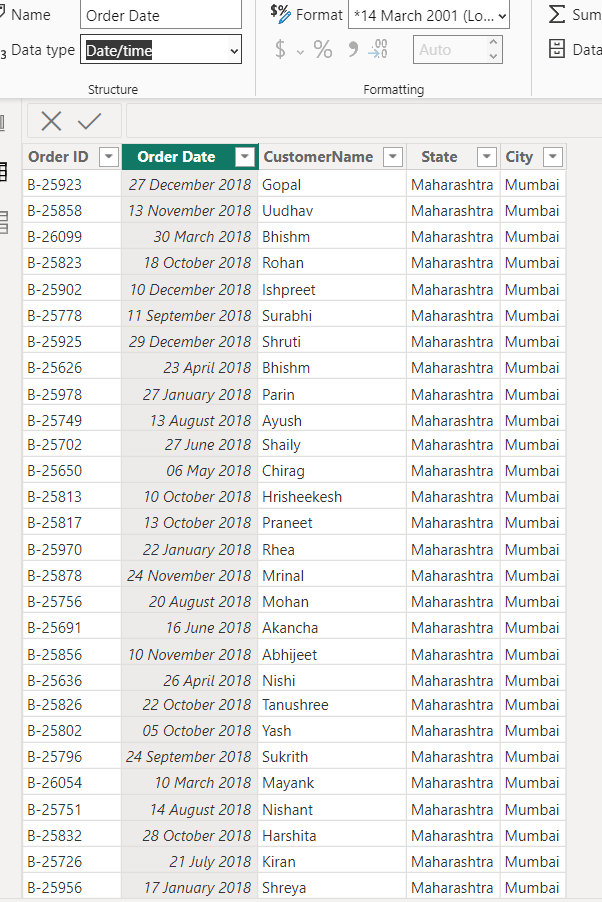
Problem statement:

You are tasked to make a sales report for a company with two tables given in the dataset with name order and details of customer. You have to analyse which category is performing better. In which city more revenue has been generated.

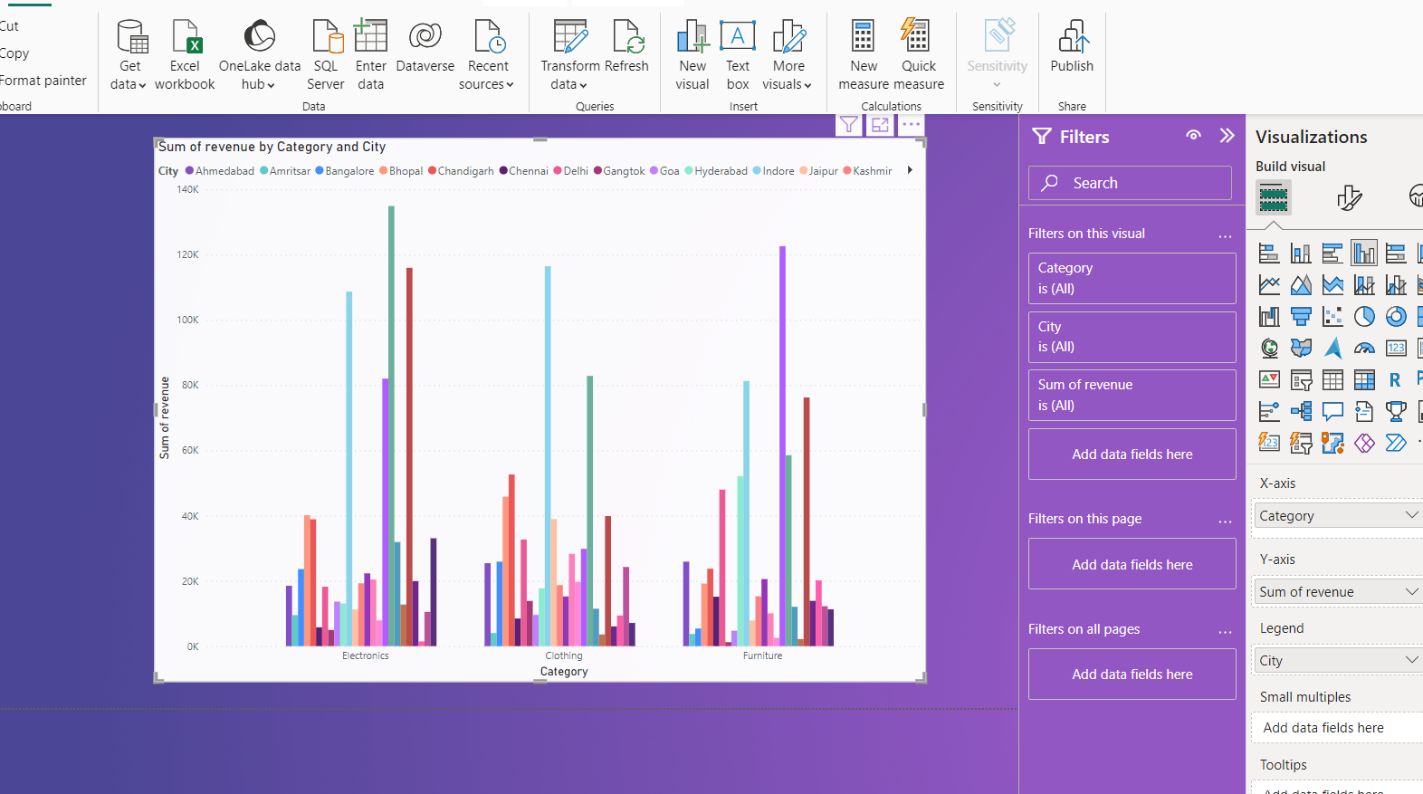
Solution:

Dataset ;

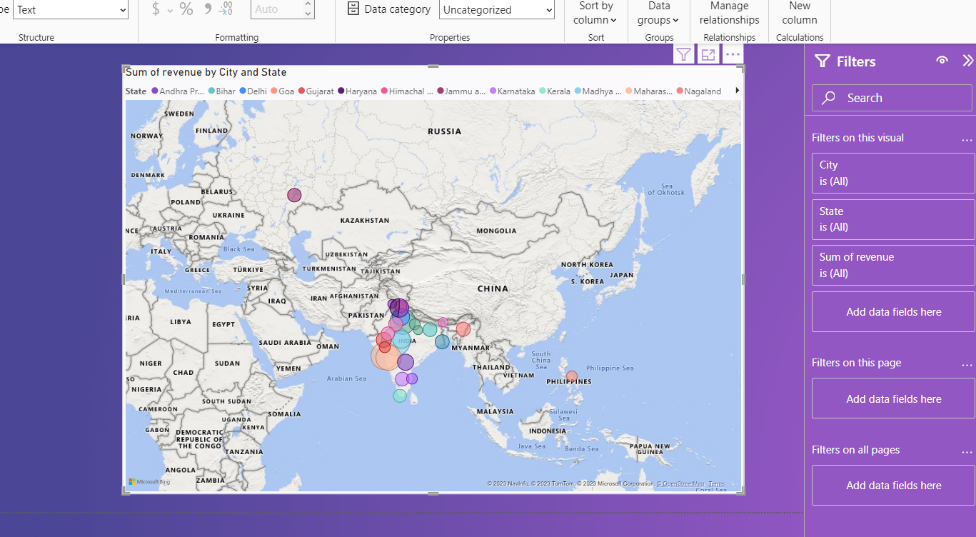
 detail .csv

order .csv

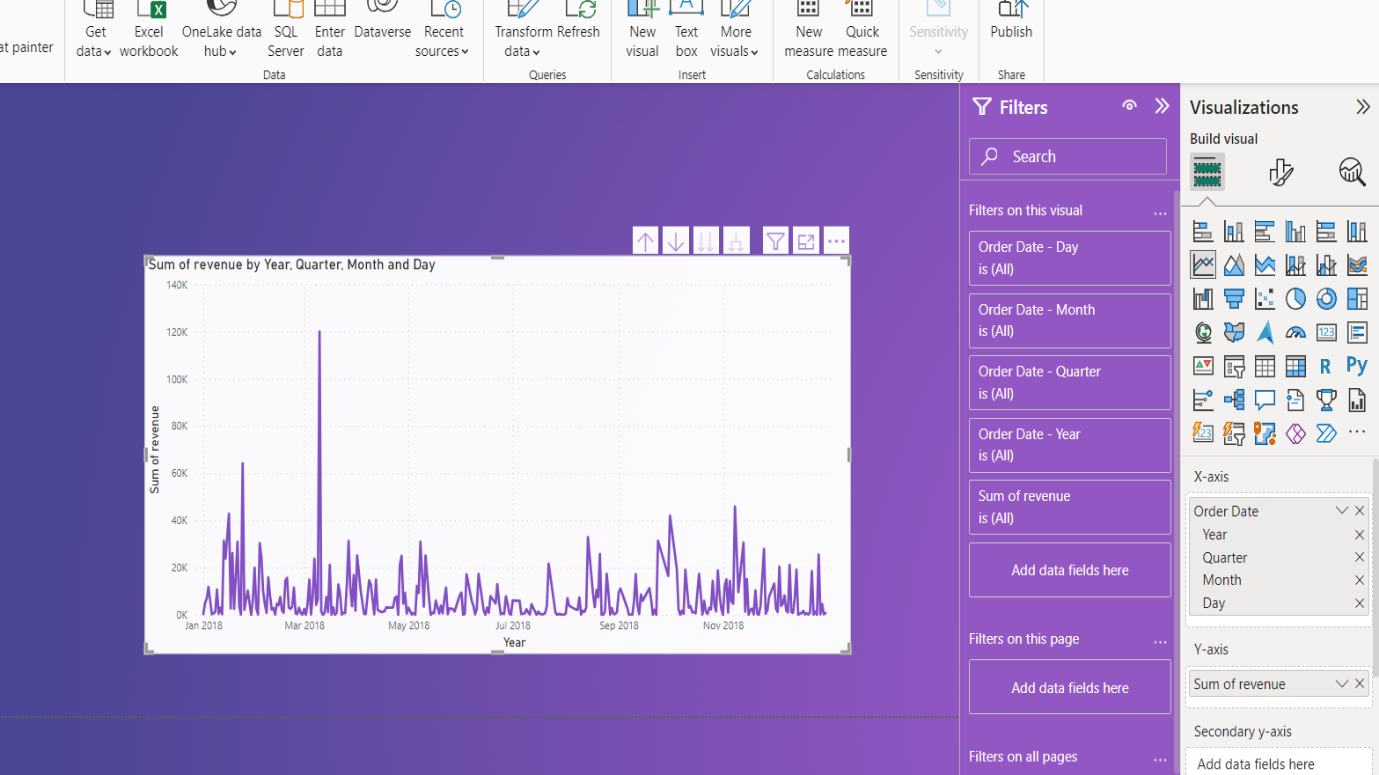
OBSERVATIONS –



Mumbai generated the most revenue by selling electronic items



Most of the revenue is generated in India rather then in other country’s



The maximum profit the company got is in April 2018

* Electronic generate most of the revenue throughout the year then other category .
* Most of the people used cod as there payment mode ,
* Printers and sarees are the most profitable sub-category

Final report

