

CANARA ENGINEERING COLLEGE

“Sudhindra Nagara”, BENJANAPADAVU

BANTWAL TALUK D.K – 574 219



LABORATORY MANUAL DATA VISUALIZATION LABORATORY (BAIL504)

PREPARED BY

Mr. SIJU V SOMAN

ASSISTANT PROFESSOR

DEPARTMENT OF AIML

1 Getting Started - Tableau Workspace, Tableau terminologies, basic functionalities.

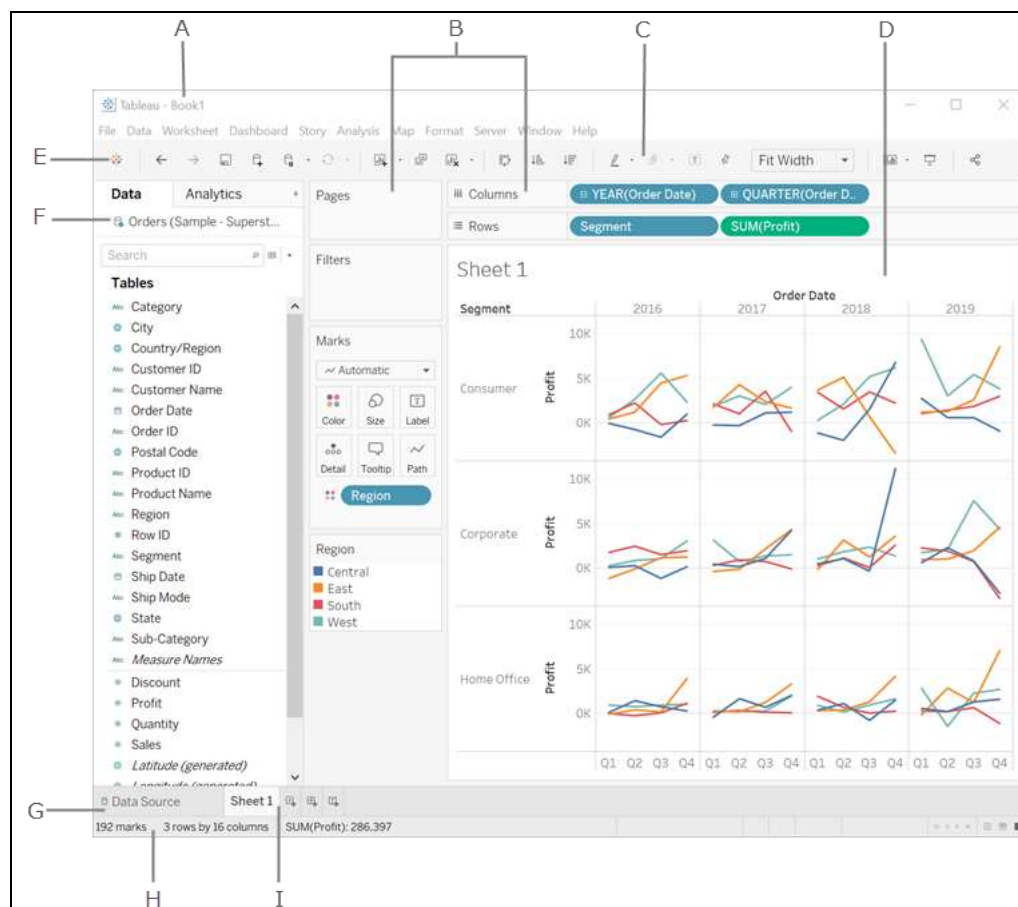
Data visualization is the practice of translating information into a visual context, such as a map or graph, to make data easier for the human brain to understand and pull insights from. It is the representation of information and data through use of common graphics, such as charts, plots, infographics, and animations. Data visualization is a powerful way for people, especially data professionals, to display data so that it can be interpreted easily. Data Visualization enables decision-makers of any enterprise or industry to look into analytical reports and understand concepts that might otherwise be difficult to grasp.

Tableau is a data visualization tool that provides pictorial and graphical representations of data. It is used for data analytics and business intelligence. Tableau provides limitless data exploration without interrupting flow of analysis. With an intuitive drag and drop interface, user can uncover hidden insights in data and make smarter decisions faster.

Tableau is a Business Intelligence tool for visually analyzing the data. Users can create and distribute an interactive and shareable dashboard, which depict the trends, variations, and density of the data in the form of graphs and charts. Tableau can connect to files, relational and Big Data sources to acquire and process data. The software allows data blending and real-time collaboration, which makes it very unique. It is used by businesses, academic researchers, and many government organizations for visual data analysis.

Tableau Workspace

The Tableau workspace consists of menus, a toolbar, the Data pane, cards and shelves, and one or more sheets. Sheets can be worksheets, dashboards, or stories.



A. Workbook name. A workbook contains sheets. A sheet can be a worksheet, a dashboard, or a story.

- B. Cards and shelves - Drag fields to the cards and shelves in the workspace to add data to your view.
- C. Toolbar - Use the toolbar to access commands and analysis and navigation tools.
- D. View - This is the canvas in the workspace where you create a visualization.
- E. Click this icon to go to the Start page, where you can connect to data.
- F. Side Bar - In a worksheet, the side bar area contains the Data pane and the Analytics pane.
- G. Click this tab to go to the Data Source page and view your data.
- H. Status bar - Displays information about the current view.
- I. Sheet tabs - Tabs represent each sheet in your workbook. This can include worksheets, dashboards, and stories.

Tableau terminologies

Terminologies	Description
Cross-tab	Another name for a text table or a table of numbers.
Dashboard	A collection of views shown in a single location where you can compare and monitor a variety of data simultaneously.
Data source	The underlying data that Tableau Reader is connected to. You can't change the data source in Tableau Reader.
Filter	A control on a view that limits the data shown in a view. For example, a filter on Region that only includes the West.
Marks	A visual representation of one or more rows in a data source. Mark types can be bar, line, square, and so on.
Packaged workbook	A type of workbook created in either Tableau Desktop or Tableau Server. These files contain both the workbook as well as copies of the referenced local file data sources and background images.
Pane	The row and columns areas in a view.
Repository	A folder located in your My Documents folder that stores workbooks.
View	The visual representation of your data in a worksheet or dashboard.
Workbook	A collection of one or more worksheets and dashboards.
Worksheet	A single view of data. Each worksheet can be connected to a single data source.

Basic Functionalities

Functionality	Description
Data visualization	Tableau can transform raw data into charts, graphs, and maps that are easy to understand. It offers a variety of charts, including bar charts, pie charts, line charts, scatterplots, histograms, and more.
Data analysis	Tableau can help you explore and analyze large datasets to find patterns, trends, and anomalies.
Dashboard creation	Tableau can help you create interactive dashboards that provide an at-a-glance view of key performance indicators (KPIs).
Operational reporting	Tableau can help you generate visually appealing reports that communicate data-driven insights.
Connecting to data sources	Tableau can connect to multiple data sources, including Excel, SQL Server, Web Connectors, and Google Analytics.
Data grouping	Tableau can help you group data points to simplify analysis.
Data functions	Tableau has a variety of functions, including number functions, string functions, date functions, type conversion aggregate functions, and logical functions.

2 Connecting to Data Source – Connecting to Database, Different types of Tableau Joins.

Dataset used: sample_superstore.xls (Contains 3 sheets: **Orders, Returns, People**)

- Connecting to Excel Files in Tableau:
 - Open Tableau and click on Connect in the left pane.
 - Under To a File, choose Microsoft Excel.
 - Browse and select your Excel file (sample_superstore.xls).
 - Tableau will display the sheets from the Excel file in the Data Source tab.
 - Drag the relevant sheets to the workspace.
- **sample_superstore.xls** Dataset: has three Excel sheets
 - **Orders:**
Row ID
Order ID
Order Date
Ship Date
Ship Mode.....
 - **Returns:**
Returned
Order ID
 - **People:**
Person
Region

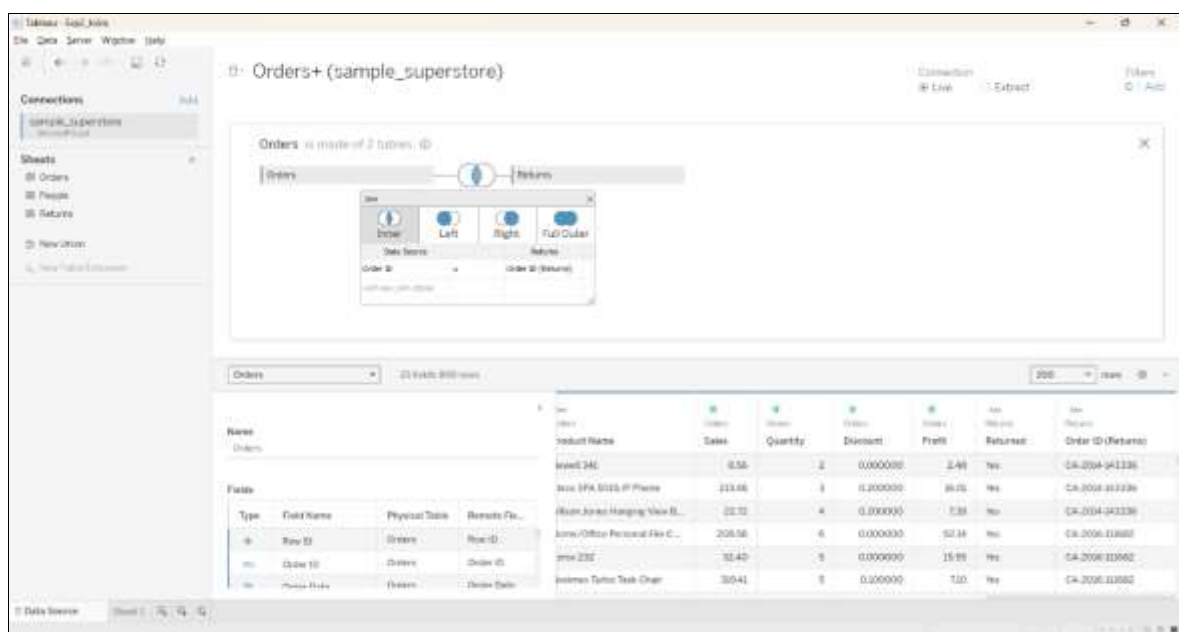
From these, the sheets named as “Order” and “Returns” have a relationship based on the field “Order ID”, and you can join them using this field.

- **Types of Joins in Tableau:**
 - a. **Inner Join:**
 - **Description:** Returns only records where there is a match in both tables.
 - **How to Create in Tableau:**
 - Double click on the Sheet “Orders” to bring it to the Relationship Window & to view the contents of the sheet.
 - Select “Orders” in the Relationship window, go to the drop-down menu in “Orders”, now click on “Open” in order to unlock the contents and to build a relation with other sheets.

- Now, check the other two sheets named as “People” & “Returns” for any common field that can be considered as the Primary Key to join both the sheets.
- As per this dataset, we have “Order ID” as the common field appearing in the sheets “Orders” & “Returns”. Hence, we can join these two sheets as follows;

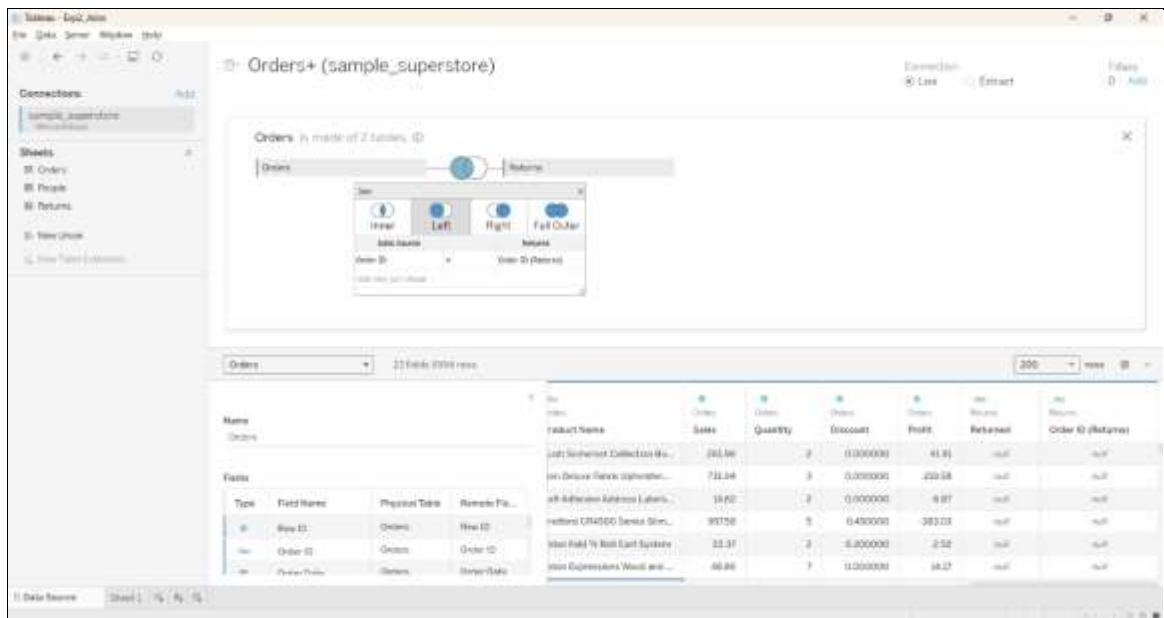
Drag the sheet “Returns” to the relationship window near to “Orders”

- Now the sheets will be joined automatically by default as “Inner Join”. To check that, we can look into the picture shown in between the line that connects both the sheets
- **Result:** You will see only orders where “Order ID” matches in both “Orders” and “Returns” table.



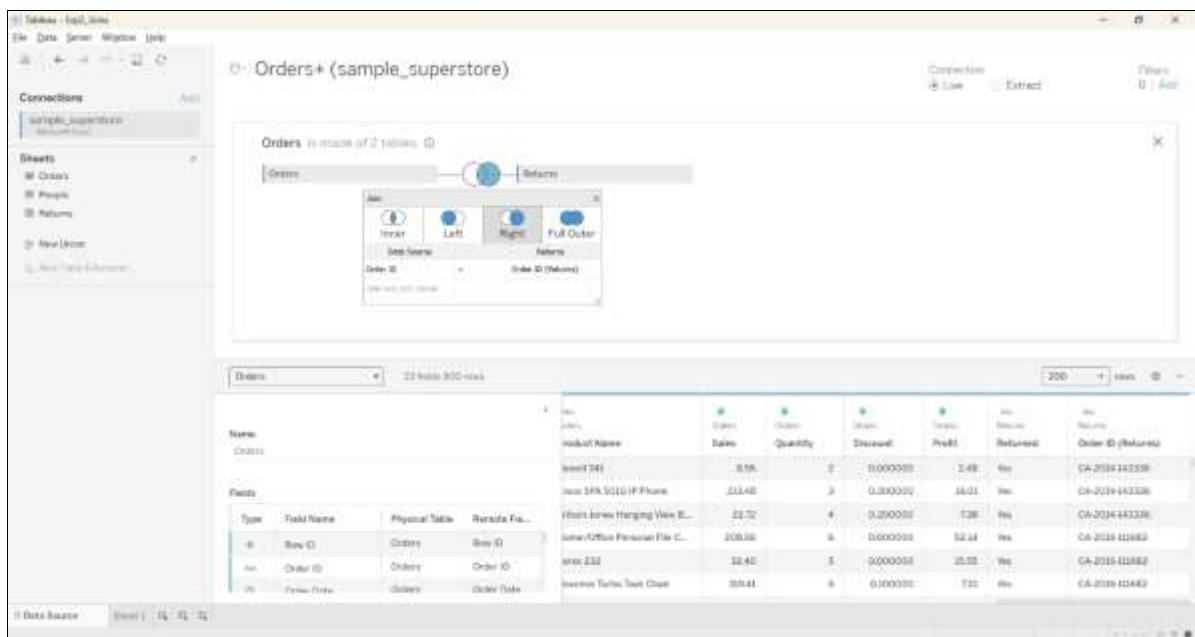
b. Left Join:

- **Description:** Returns all records from the left table (Orders), and matched records from the right table (Returns). If there's no match, NULL values are returned for fields from the right table.
- **How to Create in Tableau:**
 - ✓ In the join settings, select **Left Join**.
- **Result:** All orders will be returned, even if data missing in Orders. Returns information will be NULL for those without a match.



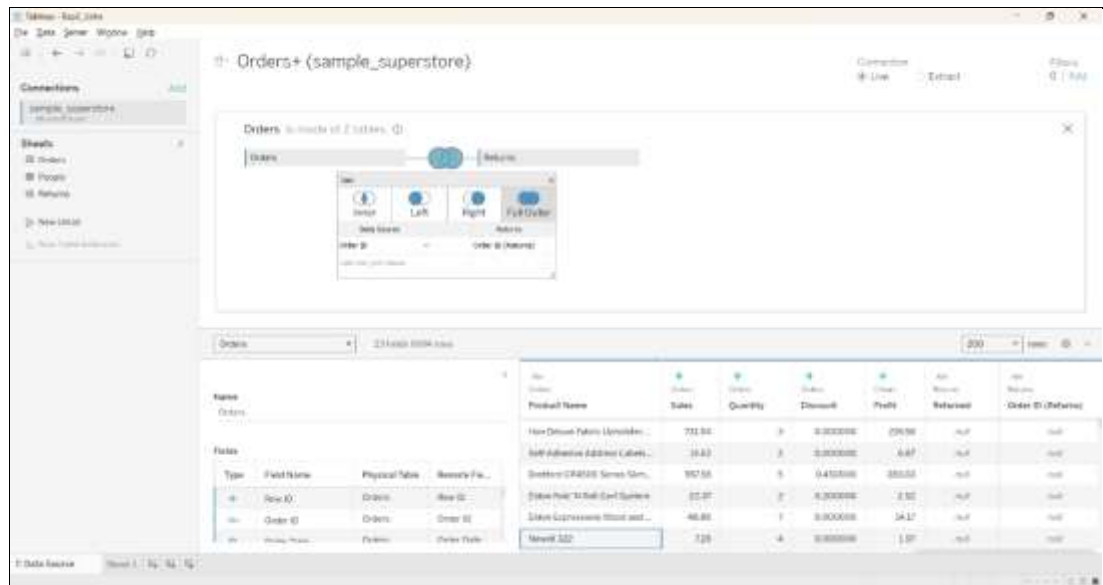
c. Right Join:

- **Description:** Returns all records from the right table (Returns), and matched records from the left table (Orders). If there's no match, NULL values are returned for fields from the left table.
- **How to Create in Tableau:**
 - ✓ Select **Right Join**.
- **Result:** You will see all orders, even if they don't have :Order ID". Order information will be NULL for those orders with no matching Order ID.



d. Full Outer Join:

- **Description:** Returns all records when there is a match in either the left (Orders) or right (Returns) table. If there's no match, NULL values are returned for the missing side.
- **How to Create in Tableau:**
 - ✓ Select **Full Outer Join**.
- **Result:** You will see all orders and all returns, even if they don't have a match in the other table. NULL values will appear where there's no corresponding record.



3

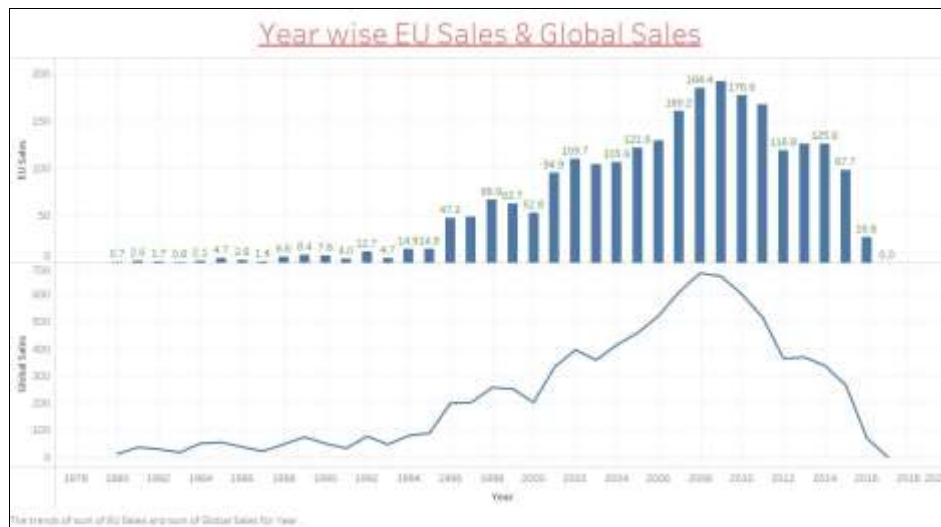
Creating a View - formatting charts, adding filters, creating calculated fields and defining parameters.

1. Click on “Connect → To a File → Text file”
2. Select the dataset file “vgsales.csv”

FORMATTING CHARTS

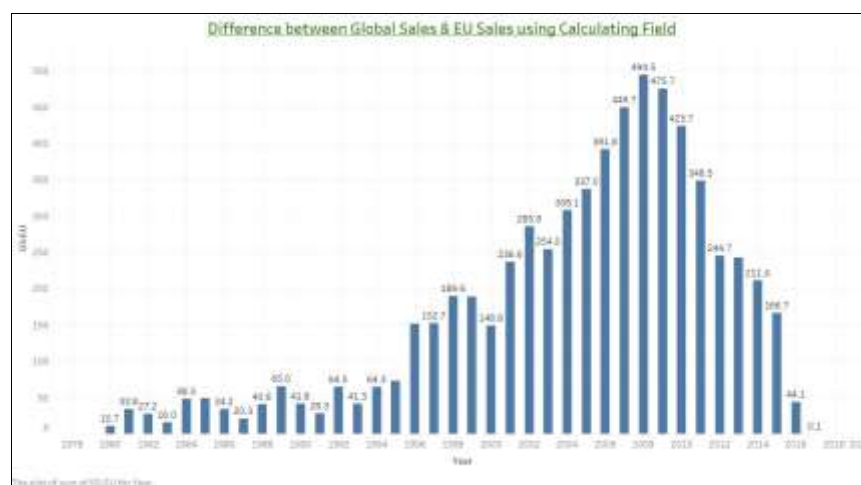
1. Drag and drop the field “Year” to the Columns.
2. Drag and drop the fields “EU Sales” and “Global Sales” to the Rows.
3. Now we can see two different Line Charts for both “EU Sales” and “Global Sales” year wise.
4. If we wish to change the graph of “EU Sales” from Line Chart to Bar Graph, then select “Sum(EU Sales)” from the “Marks” pane. By default it will be chosen as “Automatic” under the “Marks” pane. Click on that and change to “Bar”. Now you will see the graph of “EU Sales” as a Bar Chart.

- To indicate the values of each Bar, click on “Label” from the “Marks” pane and give a tick mark on “Show mark labels”, and if required you can change the text color of the label as per your choice there itself by clicking on “Font” and choosing a colour.



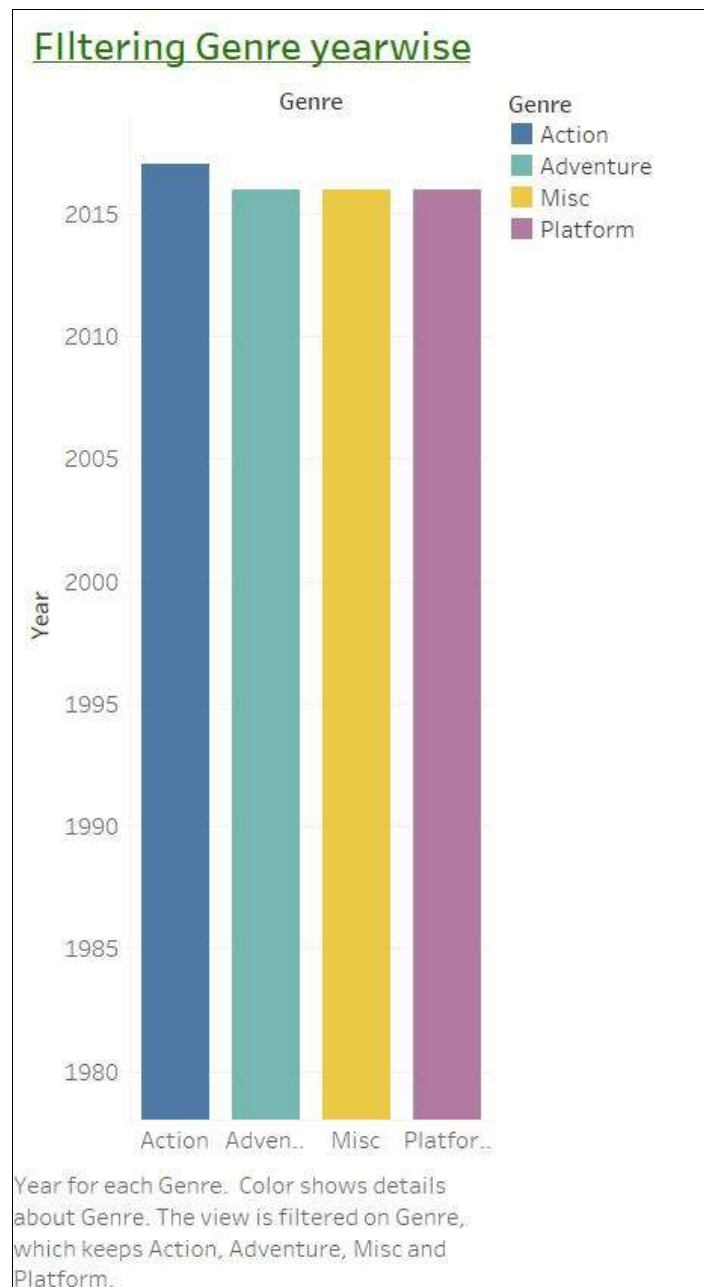
CREATING CALCULATING FIELDS

- Create a new Sheet by clicking on the symbol ‘+’ near Sheet2.
- Drag and drop the field “Year” to the Columns.
- Select “Global Sales” in the Data pane and right-click on “Global Sales”. From the drop down menu, choose “Create → Calculate Field”.
- Rename the title “Calulcation1” with a meaningful name to remember as “GS-EU”.
- Now, come to formulae window below and give the formulae as “[Global_Sales] – [EU Sales] “. Click on “Apply”, then “OK”.
- From the Data pane, Drag and drop “GS-EU” to the Rows, and visualize the Line Chart that got created.



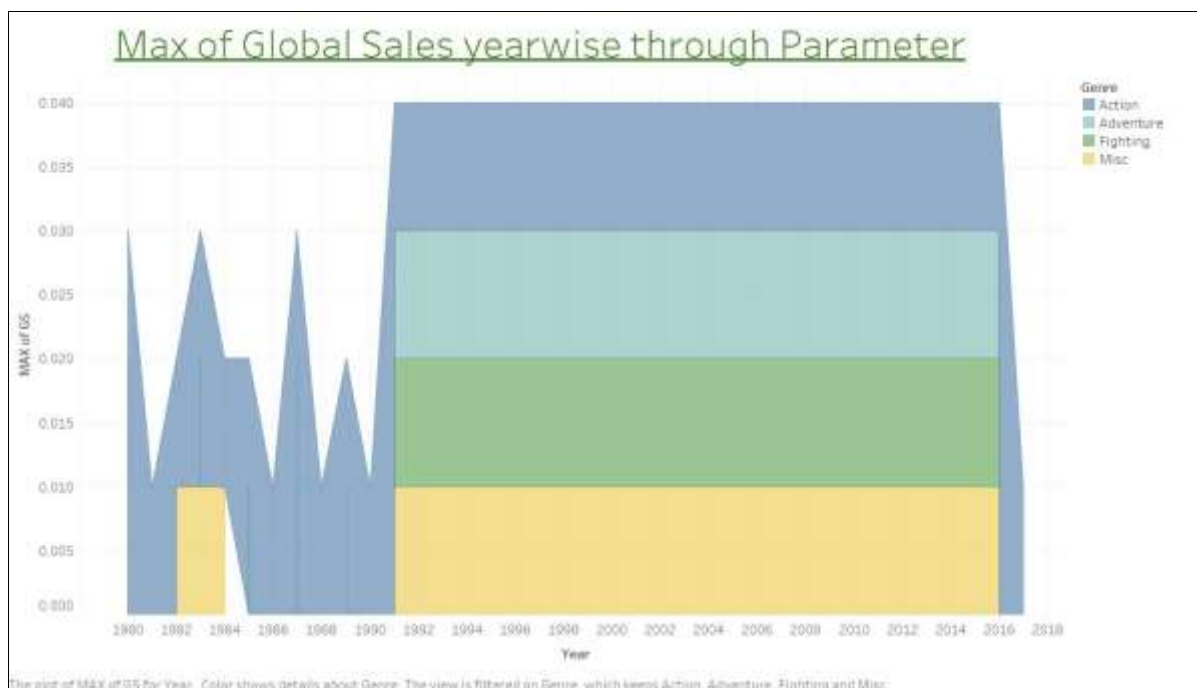
ADDING FILTERS

1. Choose a new Sheet by clicking on symbol '+' near the Sheet1.
2. Drag and drop the field "Year" to the Rows.
3. Now, if we want to view a year wise visual on some of the specific Genre, Drag and drop the field "Genre" to "Filters". Now all values are selected automatically. To select only specific Genre, first uncheck all by clicking on "None" and then select the specific "genre" from the list.



DEFINING PARAMETERS

1. From the Data pane, right-click on “Global_Sales” and choose “Create→Parameter”.
2. Rename the parameter name as any suitable name as “GSMAX” and go through the specifications in the parameter window and make changes if required. It can be to specify a range also if required.
3. Now, drag and drop the field “Year” to the Columns.
4. Now, to add a calculating field to the parameter “GSMAX”, right-click on “GSMAX” then choose “Create→Calculate field”. Now it opens a window for the formula to be typed in order to find the Maximum of Global Sales.
5. Rename the Calculation1 as “MAX of GS”. Then in the formulae window, use the pre-defined function as “max([GSMAX])”. Make sure to give the round braces after the function name “max”.
6. Now, from the Data pane, drag and drop “MAX of GS” into the rows. Modify the chart by choosing “Area Chart” from the Marks Pane.



4

Dashboard Design and Storytelling – Components of Dashboard, Understanding how to place worksheets in Containers, Action filters and its types.

DASHBOARD

A dashboard is a collection of several views, letting you compare a variety of data simultaneously. For example, if you have a set of views that you review every day, you can create a dashboard that displays all the views at once, rather than navigate to separate worksheets.

Like worksheets, you access dashboards from tabs at the bottom of a workbook. Data in sheets and dashboards is connected; when you modify a sheet, any dashboards containing it change, and vice versa. Both sheets and dashboards update with the latest available data from the data source.

We have two types of Dashboards;

1. Tiled Dashboard
2. Floating Dashboard

The “Tiled” dashboard has to be placed in an exact position as per the requirement. Whereas in the “Floating” dashboard, the objects can be placed anywhere in the window.

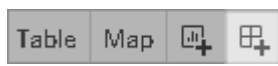
Components of a Tableau dashboard:

- **Layout containers:** Horizontal and vertical containers that can hold additional objects
- **Text:** A mini word processor for adding and formatting text
- **Images:** An image from your computer can be added to the dashboard
- **Web pages:** A web page can be embedded in the dashboard, but an internet connection is required to display it
- **Blank space:** Blank space can be added to the dashboard to separate elements that are too close together
- **Headers:** Headers are created when a dimension or discrete field is placed on the Rows or Columns shelves
- **Tags and notifications:** A tagging and notification system can be used to send messages, updates, and tasks to users

Creating a Dashboard

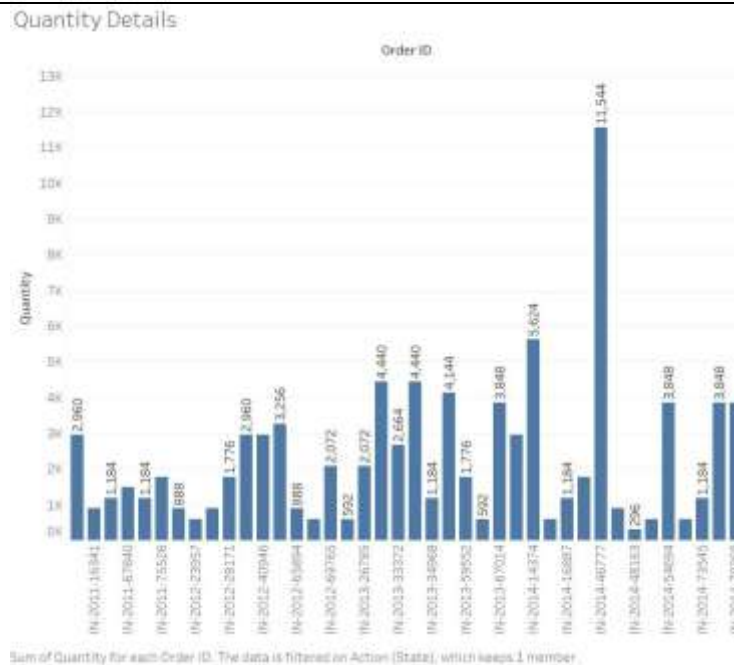
You create a dashboard in much the same way you create a new worksheet.

1. At the bottom of the workbook, click the New Dashboard icon:



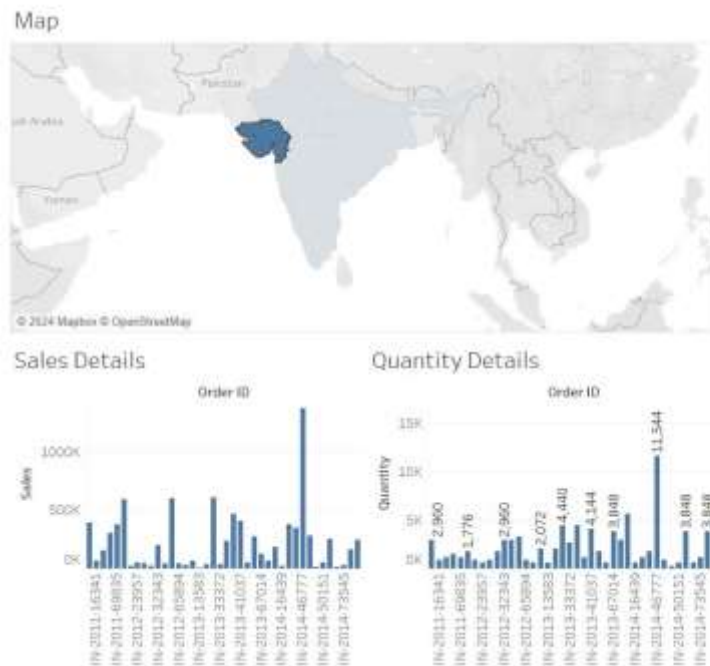
2. From the Sheets list at left, drag views to your dashboard at right
3. Add dashboard objects and set their options

❖ Steps to create multiple sheets for the Experiment 4



❖ Steps to create multiple Dashboard

- ✓ Create Dashboard-1 by clicking on '+' symbol at the bottom near to Sheet.
- ✓ Choose the type of the Dashboard as "Tiled".
- ✓ Insert a "Blank" object by double clicking on the object "Blank".
- ✓ Insert a "Horizontal" object onto the "Blank" object by dragging and releasing on the Top→Left to Right area.
- ✓ Insert a "Vertical" object onto the "Blank" object by dragging and releasing to the Bottom→Left, and one more on the Bottom→Right.
- ✓ Now, drag the sheet "Map" onto the "Horizontal container".
- ✓ Drag the sheet "Sales" onto the Left side "Vertical Container".
- ✓ Drag the sheet "Quantity" onto the Right side "Vertical Container".
- ✓ Rename the Dashboard-1 as "All 3 DB".



- ✓ Create Dashboard-2 just like mentioned above.

-
- | Order ID | Sales |
|---------------|-------|
| 10-2011-16841 | 4000 |
| 10-2011-16842 | 500 |
| 10-2011-16843 | 1500 |
| 10-2011-16844 | 3000 |
| 10-2011-16845 | 3800 |
| 10-2011-16846 | 6000 |
| 10-2011-16847 | 100 |
| 10-2011-16848 | 400 |
| 10-2011-16849 | 400 |
| 10-2011-16850 | 2000 |
| 10-2011-16851 | 200 |
| 10-2011-16852 | 6000 |
| 10-2011-16853 | 400 |
| 10-2011-16854 | 500 |
| 10-2011-16855 | 6000 |
| 10-2011-16856 | 400 |
| 10-2011-16857 | 100 |
| 10-2011-16858 | 100 |
| 10-2011-16859 | 100 |
| 10-2011-16860 | 100 |
| 10-2011-16861 | 100 |
| 10-2011-16862 | 100 |
| 10-2011-16863 | 100 |
| 10-2011-16864 | 100 |
| 10-2011-16865 | 100 |
| 10-2011-16866 | 100 |
| 10-2011-16867 | 100 |
| 10-2011-16868 | 100 |
| 10-2011-16869 | 100 |
| 10-2011-16870 | 100 |
| 10-2011-16871 | 100 |
| 10-2011-16872 | 100 |
| 10-2011-16873 | 100 |
| 10-2011-16874 | 100 |
| 10-2011-16875 | 100 |
| 10-2011-16876 | 100 |
| 10-2011-16877 | 100 |
| 10-2011-16878 | 100 |
| 10-2011-16879 | 100 |
| 10-2011-16880 | 100 |
| 10-2011-16881 | 100 |
| 10-2011-16882 | 100 |
| 10-2011-16883 | 100 |
| 10-2011-16884 | 100 |
| 10-2011-16885 | 100 |
| 10-2011-16886 | 100 |
| 10-2011-16887 | 100 |
| 10-2011-16888 | 100 |
| 10-2011-16889 | 100 |
| 10-2011-16890 | 100 |
| 10-2011-16891 | 100 |
| 10-2011-16892 | 100 |
| 10-2011-16893 | 100 |
| 10-2011-16894 | 100 |
| 10-2011-16895 | 100 |
| 10-2011-16896 | 100 |
| 10-2011-16897 | 100 |
| 10-2011-16898 | 100 |
| 10-2011-16899 | 100 |
| 10-2011-16900 | 100 |
| 10-2011-16901 | 100 |
| 10-2011-16902 | 100 |
| 10-2011-16903 | 100 |
| 10-2011-16904 | 100 |
| 10-2011-16905 | 100 |
| 10-2011-16906 | 100 |
| 10-2011-16907 | 100 |
| 10-2011-16908 | 100 |
| 10-2011-16909 | 100 |
| 10-2011-16910 | 100 |
| 10-2011-16911 | 100 |
| 10-2011-16912 | 100 |
| 10-2011-16913 | 100 |
| 10-2011-16914 | 100 |
| 10-2011-16915 | 100 |
| 10-2011-16916 | 100 |
| 10-2011-16917 | 100 |
| 10-2011-16918 | 100 |
| 10-2011-16919 | 100 |
| 10-2011-16920 | 100 |
| 10-2011-16921 | 100 |
| 10-2011-16922 | 100 |
| 10-2011-16923 | 100 |
| 10-2011-16924 | 100 |
| 10-2011-16925 | 100 |
| 10-2011-16926 | 100 |
| 10-2011-16927 | 100 |
| 10-2011-16928 | 100 |
| 10-2011-16929 | 100 |
| 10-2011-16930 | 100 |
| 10-2011-16931 | 100 |
| 10-2011-16932 | 100 |
| 10-2011-16933 | 100 |
| 10-2011-16934 | 100 |
| 10-2011-16935 | 100 |
| 10-2011-16936 | 100 |
| 10-2011-16937 | 100 |
| 10-2011-16938 | 100 |
| 10-2011-16939 | 100 |
| 10-2011-16940 | 100 |
| 10-2011-16941 | 100 |
| 10-2011-16942 | 100 |
| 10-2011-16943 | 100 |
| 10-2011-16944 | 100 |
| 10-2011-16945 | 100 |
| 10-2011-16946 | 100 |
| 10-2011-16947 | 100 |
| 10-2011-16948 | 100 |
| 10-2011-16949 | 100 |
| 10-2011-16950 | 100 |
| 10-2011-16951 | 100 |
| 10-2011-16952 | 100 |
| 10-2011-16953 | 100 |
| 10-2011-16954 | 100 |
| 10-2011-16955 | 100 |
| 10-2011-16956 | 100 |
| 10-2011-16957 | 100 |
| 10-2011-16958 | 100 |
| 10-2011-16959 | 100 |
| 10-2011-16960 | 100 |
| 10-2011-16961 | 100 |
| 10-2011-16962 | 100 |
| 10-2011-16963 | 100 |
| 10-2011-16964 | 100 |
| 10-2011-16965 | 100 |
| 10-2011-16966 | 100 |
| 10-2011-16967 | 100 |
| 10-2011-16968 | 100</ |

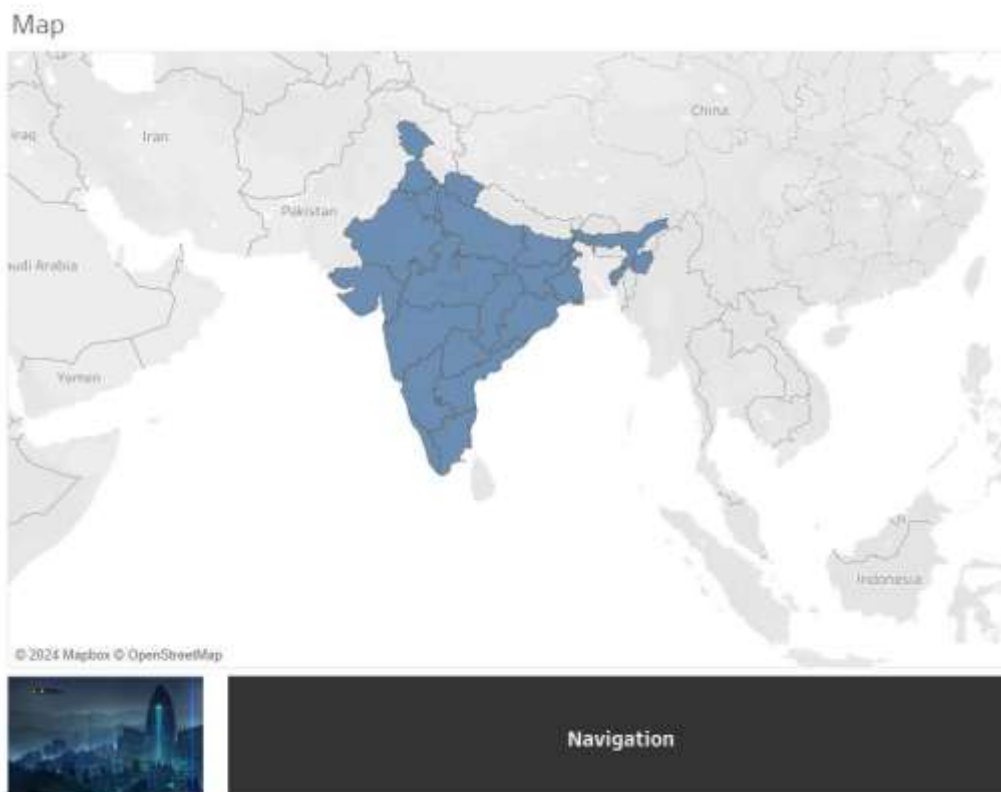
- [illegible]

- Page - 13 - of 43

- ✓ To add an Activity Filter in the dashboard, go to “All 3 DB”. Under the menu “Dashboard”, choose the sub-menu “Actions”, click “Add Action” → “Filter”. Now, in the window that displayed, select the following;

From “Source Sheet”, uncheck “Quantity”, “Sales”. Check only “Map”. Then under “Run action on” choose “Select”.

From “Target Sheet”, uncheck “Map” and check both “Quantity” and “Sales”. Then, from the option “Clearing the selection will”, choose “Show all values”. Then click OK.



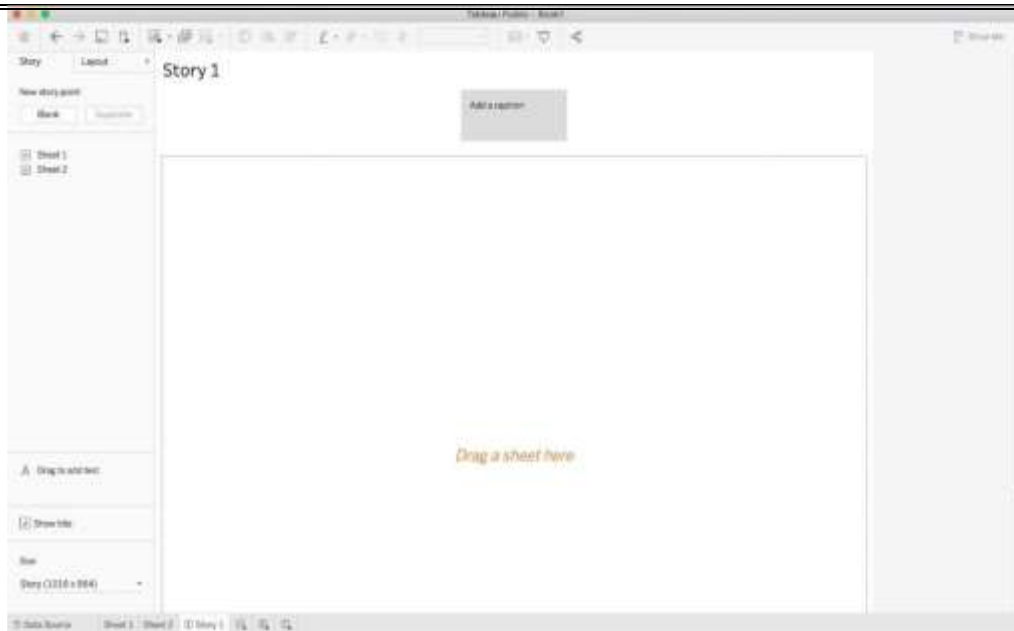
Now, to check the working of the action, go to “Presentation Mode” by clicking on the suitable symbol on top (or) press F7. Then click on a particular state from the map and you will see the changes occurring on the data in “Quantity” & “Sales” based on the state selected.

CREATING A STORY WITH TABLEAU

With Tableau, you are able to organize your data in order to tell a meaningful story. This is beneficial when you are doing a presentation, creating an article, or uploading to a website, as it helps your audience understand your data.

Stories are created through assembling the different worksheets and dashboards. We can highlight important data points, add text box and pictures to help convey our story. However, there are many different ways to tell a story.

- ❖ To begin, select “New Story” at the bottom right of your screen.



- ❖ Drag “Dashboard1” and “Dashboard2” on to “Drag a sheet here”. We can rename each slide in the story by clicking on “Add a caption” and then give a suitable name to identify that particular slide. Rename Story-1 if we want to have multiple stories.
- ❖ To create a new slide in this same story, click on “New Story Point→Blank”, and bring the necessary dashboard/sheet into the design area as per the requirement. Keep renaming each slide as and when we create in the same story.
- ❖ Steps to create a Story
 - ✓ Here in story, we can create different slides where each slide can be a Dashboard created before or can be a sheet created.
 - ✓ Create a story by clicking on the ‘+’ symbol combined along with a book symbol.
 - ✓ Rename “Add a caption” that appear on top with a suitable name to identify the specific slide/content.
 - ✓ Drag and drop the required Dashboard/Sheet into this story under a slide.
 - ✓ Create a second slide by clicking on “New Story Point”→”Blank”.
 - ✓ Repeat the above steps to add contents into each slide by simple drag and drop on to the story slides.
 - ✓ The default size set for the story can be changed by clicking on “Size” situated on the bottom left corner, followed by choosing an appropriate size.
 - ✓ The story created with different contents from a Dashboard/Sheet can be viewed using the Presentation Mode or by pressing “F7” in the keyboard.



5

Introducing Power BI –Components and the flow of work. Power BI Desktop Interface- The Report has five main areas.

Power BI includes the following components

- **Power BI Desktop** – This is used to create reports and data visualizations on the dataset.
- **Power BI Gateway** – You can use Power BI on-premises gateway to keep your data fresh by connecting to your on-premises data sources without the need to move the data. It allows you to query large datasets and benefit from the existing investments.
- **Power BI Mobile Apps** – Using Power BI mobile apps, you can stay connected to their data from anywhere. Power BI apps are available for Windows, iOS, and Android platform.
- **Power BI Service** – This is a cloud service and is used to publish Power BI reports and data visualizations

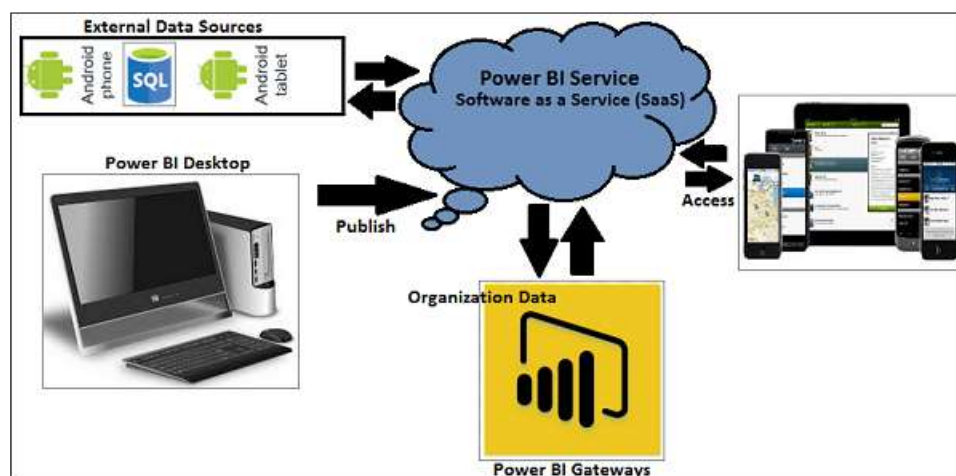





Fig. The Components of Power BI

Flow of work

- A typical Power BI workflow involves more than one type of content.
- A Power BI designer (yellow in the diagram) collects data from semantic models, brings it into Power BI Desktop for analysis, and creates reports full of visualizations that highlight interesting facts and insights.
- The designer pins visualizations from reports to dashboards, and shares the reports and dashboards with business users like you (black in the diagram).



Fig. Flow of work

-  A visualization (or visual), is a type of chart built by Power BI designers. The visuals display the data from reports and semantic models. Because they're highly interactive, you can slice, filter, highlight, change, and even drill into visualizations.
-  A semantic model is a container of data. For example, it might be an Excel file from the World Health Organization. It might also be a company-owned database of customers, or it might be a Salesforce file. And it might be all three if the designer combines them into a single model. Designers manage semantic models. The data contained in semantic models is used to build reports, dashboards, and apps that designers share with you.
-  A dashboard is a single screen with tiles of interactive visuals, text, and graphics. A dashboard collects your most important metrics, or a focused set of metrics, on one screen, to tell a story or answer a question. The dashboard content comes from one or more reports and one or more semantic models.



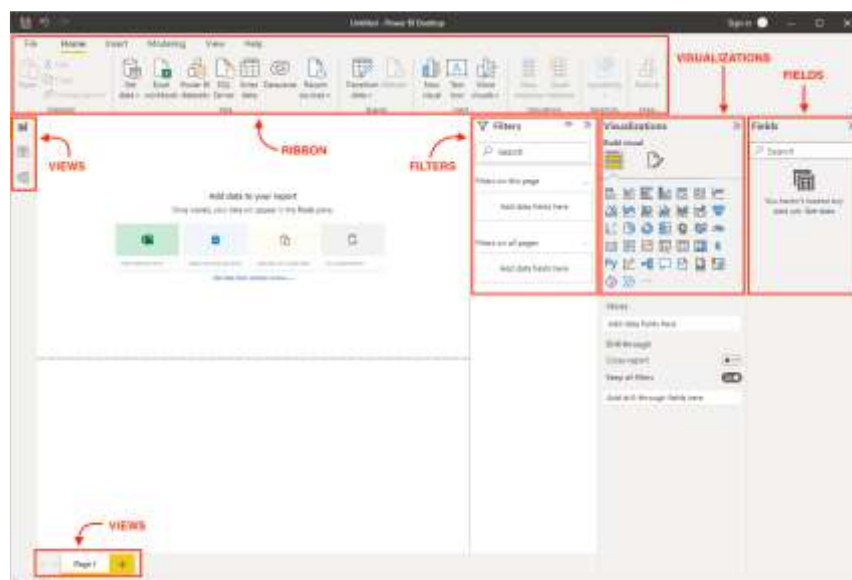
- A report is one or more pages of interactive visuals, text, and graphics that together make up a single report. Power BI bases a report on a single semantic model. Often, the designer organizes report pages to each address a central area of interest or answer a single question.



- An app is a way for designers to bundle and share related dashboards, reports, and semantic models together. Business users receive some apps automatically but can go search for other apps created by colleagues or by the community. For example, out-of-the-box apps are available for external services you may already use, like Google Analytics and Microsoft Dynamics CRM.

Power BI Desktop Interface-The Report has five main areas.

- **Ribbon** - the top ribbon contains most of the controls and options needed for building the report.
- **Views** - this is made up of the report view, the data view, and the model view.
- **Canvas** - this is the main design area where visualizations and other elements are added.
- **Page selector** - for navigation to other pages in the report.
- **Filters** - fields can be added here to filter the data.
- **Visualizations** - this contains the list of available visualizations.
- **Fields** - this section contains the tables and fields that are available in the data model.



The Major Components of Power BI Desktop Interface are

Power Query Editor

It is the process of cleansing and transforming data and permits users to access datasets connecting from multiple sources. It is included on the Power BI desktop. Business users may view the data from distinct databases like MySQL, SQL servers, DB2, and many more.

Power View

It is a data visualization tool that assists users in developing stunning charts, and colourful maps, that turn data into a story.

Power Map

It is a 3D map visualization tool to identify geospatial data on Map visuals. It helps organizations to examine the maximum sales production geographically, visualizing the demographic populations of specific regions.

Power Pivot

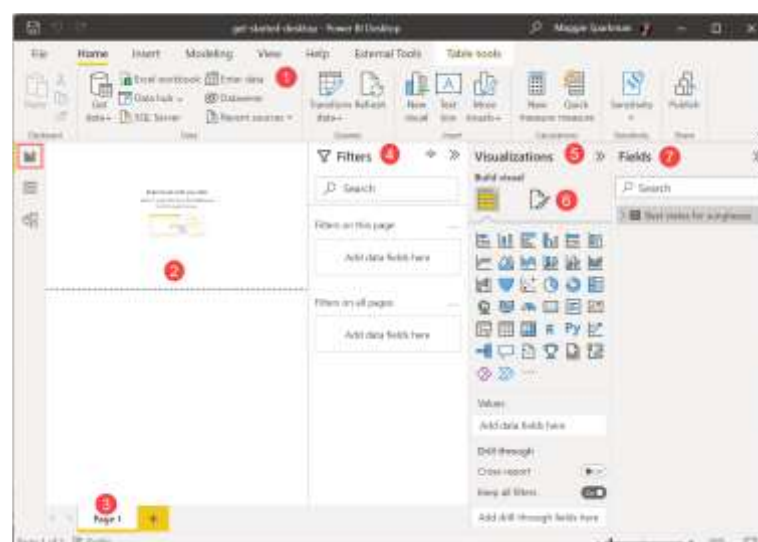
It is a Data Modelling technique that is used to create relationships between datasets. It performs complex computations by utilizing DAX functions.

Power Q & A

When dealing with giant datasets, it becomes crucial to get to know the in-depth details of the data. Luckily, it is done through natural language where users may ask questions and obtain the answer through Power Q & A.

Build reports :

In Power BI Desktop Report view, you can build visualizations and reports. The Report view has six main areas:

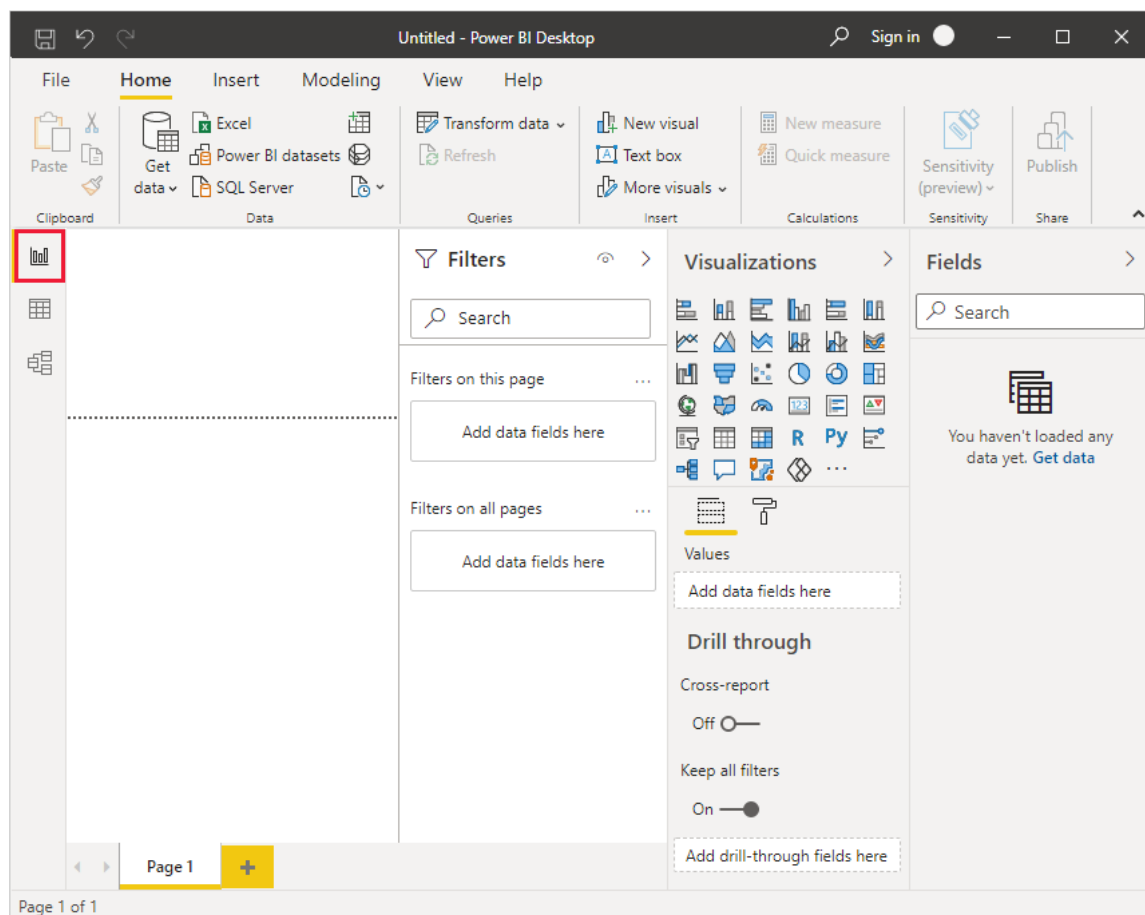


1. The ribbon at the top, which displays common tasks associated with reports and visualizations.
2. The canvas area in the middle, where you create and arrange visualizations.

3. The pages tab area at the bottom, which lets you select or add report pages.
4. The Filters pane, where you can filter data visualizations.
5. The Visualizations pane, where you can add, change, or customize visualizations, and apply drill through.
6. The Format pane, where you design the report and visualizations.
7. The Fields pane, which shows the available fields in your queries. You can drag these fields onto the canvas, the Filters pane, or the Visualizations pane to create or modify visualizations.

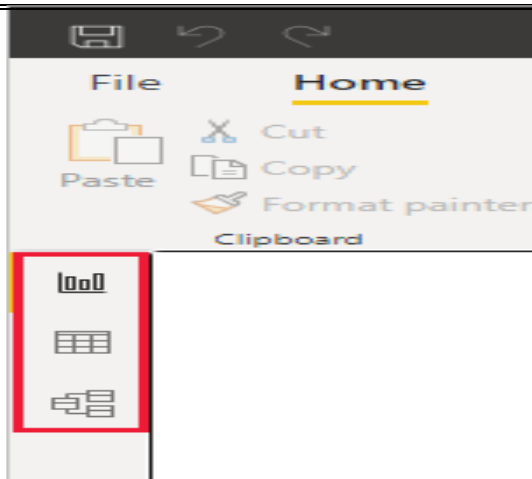
6 Querying Data from CSV - Query Editor, Connecting the data from the Excel Source, Clean, Transform the data.

Power BI Desktop also includes the Power Query Editor, which opens in a separate window. In Power Query Editor, you can build queries and transform data, then load the refined data model into Power BI Desktop to create reports.



Along the left side of Power BI Desktop are icons for the three Power BI Desktop views: Report, Data, and Model, from top to bottom. The current view is indicated by the yellow bar along the left, and you can change views by selecting any of the icons.

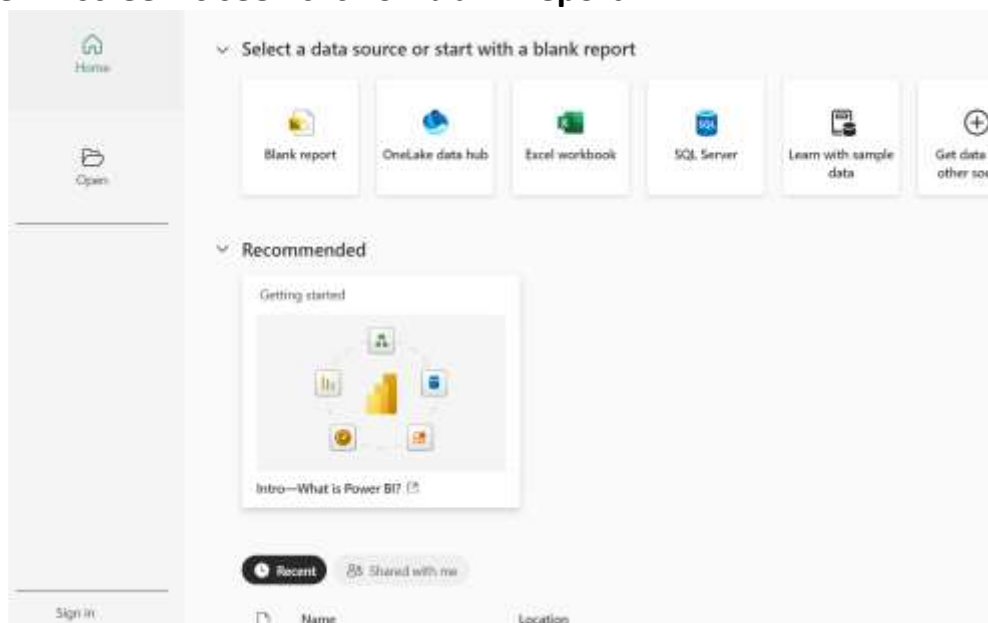
Report view is the default view.



Connect to data(Get Data from different Sources)

With Power BI Desktop installed, we can connect to the world of data. To see the many types of data sources available.

Once Power BI screen is seen click on blank Report

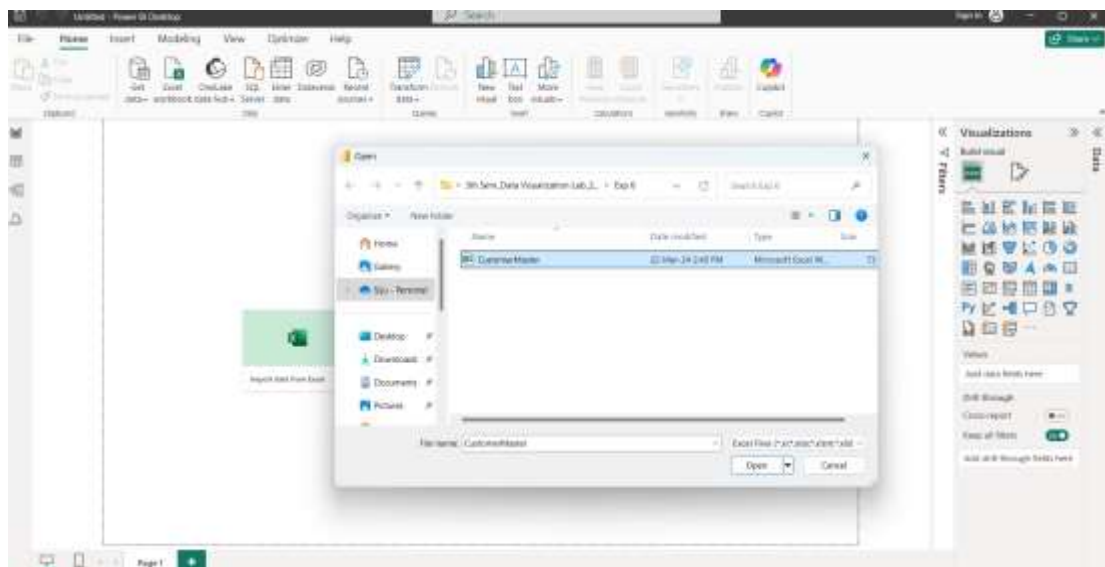
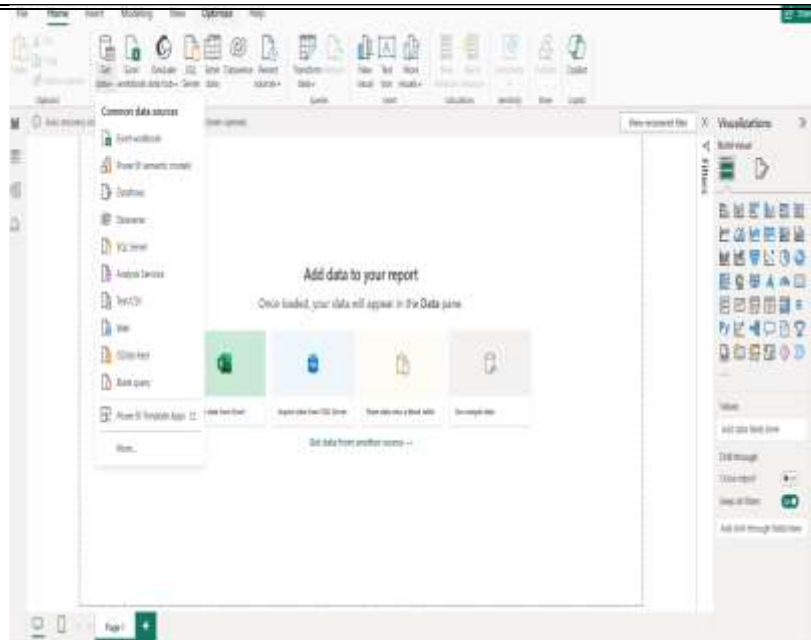


After clicking on blank Report the below screen appears.

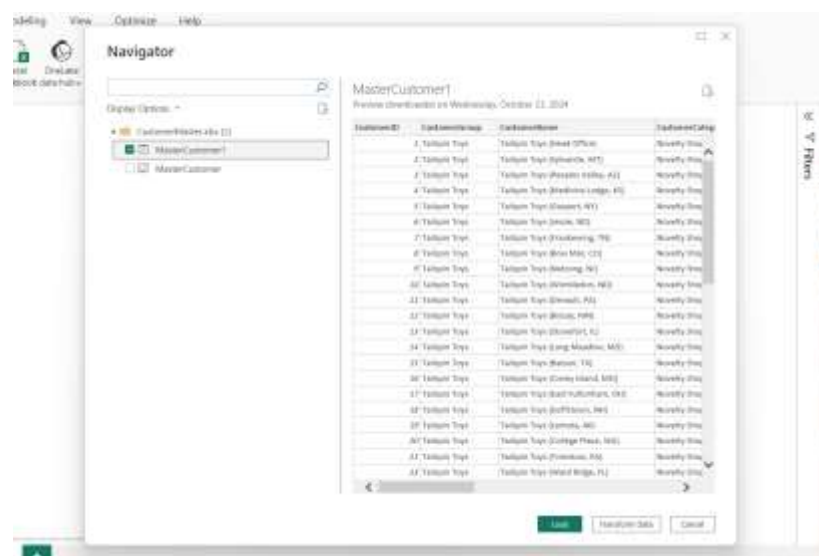
NOW TO GET DATA FROM DIFFERENT SOURCES -----The steps

Step 1: Select **Get Data** in the Power BI Desktop Home tab, and in the Get Data window, scroll through the list of All data sources.(like Excel,CSV,Oracle....)

On the Power BI Desktop Home tab, select Get Data > Excel workbook

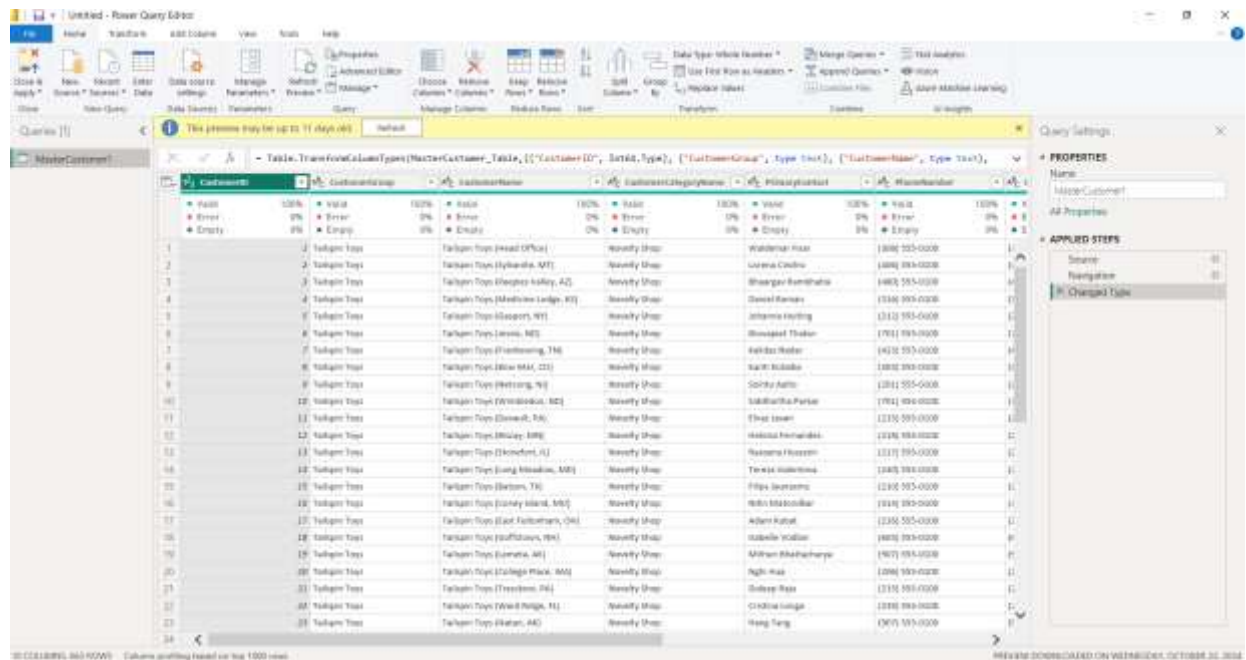


Step 2: Click on the file you need and open the file, once you open the file below window with navigator appears, select the file that appear as “Table” View format.



Step 3: At this point you can select Load to load the table, or Transform data to make changes in the table before you load it.

Step 4: When you select Transform data, Power Query Editor launches, with a representative view of the table. The Query Settings pane is on the right, or you can always show it by selecting Query Settings on the View tab of Power Query Editor.



Transforming the data

Once connected to a data source, you can adjust the data to meet your needs.

To transform the data, you provide Power Query Editor with step-by-step instructions for adjusting the data while loading and presenting it. Transforming doesn't affect the original data source, only this particular view of the data.

Transforming the data, includes **renaming columns or tables, removing rows or columns, or changing data types.**

Power Query Editor captures these steps sequentially under Applied Steps in the Query Settings pane.

Notice that the Applied Steps in Query Settings already contain a few steps. You can select each step to see its effect in the Power Query Editor

❖ Transforming the Data using excel file

- ✓ Check for the datatype automatically applied by PowerBI. If required, change it.
- ✓ Find the unnecessary columns and remove it which is not going to be useful for reporting. (Option are (1) Click on “Remove columns” from the ribbon (2) Right-

click on the column to be removed and choose “Remove” (3) or Go for “Choose Columns” from the ribbon and uncheck the columns that are not required for the reporting.

Remove Columns – “DeliveryMethod”, “WebsiteURL”

Choose Columns – Uncheck “PhoneNumber”, “FaxNumber”, “PrimaryContact”

- ✓ Split “ProvinceCity” separately. By right-click→Split Column→By Delimiter(mention the Custom delimiter as ‘(’).
- ✓ Now, we get a new column created with Province name followed by ‘)’. To remove this ‘)’ we have two options as mentioned below;
 1. Right-click on “CityProvince2”→Split Column→By Delimiter (and then choose custom delimiter as ‘)’) and give “OK”. Now we will get another new column with only ‘)’. Then delete that column which has only ‘)’.
 2. Right-click on “CityProvince2”→Replace Values. In this now, give “Values to Find” with the bracket ‘)’ in order to remove that bracket from this column. “Replace With” value to be nothing as we don’t want any replacement for ‘)’. It must be empty.
- ✓ Rename the column “CityProvince1” as “City”, “CityProvince2” as “Province”
- ✓ Since the data regarding “CustomerGroup” is repeating in two different columns, we can delete one column with the name “CustomerGroup” first and modify the other column as follows;
 - ➔ Right-click on “CustomerName”, Split Column→By Delimiter . It automatically choose ‘(’ as the delimiter, then give “OK”. Now the values are split into two columns.
 - ➔ Remove the ‘)’ from the “CustomerName 2” column. Rename the columns “CustomerName1” as “CustomerName” and “CustomerName2” as “Head Office”.
- ✓ Check for any “Null Values” in any of the columns by clicking on the drop down arrow near to column name. If there is any column with Null Value is present, then check whether to retain them as it is or to remove them by un-checking “Null” from the list.
- ✓ Instead of retaining “Null values” in the column “Head Office” we can create a new column where we can include a condition and generate the values as either “Head Office Exist” (or) “No Head Office”. Before that, sort the values in the column “Head Office” as Ascending, and now we can see all “Null” valued data appears on the top rows.
- ✓ The new column with a condition can be created as follows;
 1. Click on “Add Column”.
 2. In the window that opened, give a suitable name for this new column as “Head Office Status”.
 3. Then mention the condition in that window as “if Head Office equals null No Head Office” “else Head Office Exists”.
- ✓ Rename the table/query as “Customer” instead of “MasterCustomer1” near the Query Properties area.

❖ Transforming the data from text file

- ✓ Click on “Get Data→Text/CSV”, and choose the invoice.txt file from our local repository.
- ✓ Transform the data.
- ✓ Check for the data type. If changes required, then change it. PowerBI automatically identify the headers from the text file.
- ✓ Change the data type of the column “Sales” from “Whole number” to “Fixed decimal number” type.
- ✓ Remove the un-necessary column items as mentioned here. Remove the column “ TotalChillerItems” as it has only ‘0’ as the values.
- ✓ In this invoice.txt, the day, month & year are separated into 3 columns. But if we want to merge all these 3 columns into a single column as “Date”, then follow the steps as follows;
 1. Select the 3 columns by pressing Ctrl + “Day”, “Month”, “Year”, then right-click and choose the option as Merge.
 2. Now in the merge window, give Separator as “Custom”, then give the symbol ‘/’ followed by a new column name as “Date”.
 3. Change the datatype from “Text” to “Date”.
- ✓ As we have tables from two different files, we need to make a Relationship by considering the fields that are common in both. Here it “Customer” table, the field “Customer ID” is common with “Customer Code” in the “InvoiceData” table. This relationship is built automatically by PowerBI which can be verified by click onto Model View on the left side View Pane.

7 **Creating Reports & Visualizations** - Different types of charts, Formatting charts with Title, Colors.

- ✓ Open Power BI Desktop
 - ✓ Click on Get data in ribbon pane & Choose the Excel File(CustomerMaster.xls) & the Text File (InvoiceData.txt)
 - ✓ Clean & Transform the data as mentioned in Experiment 6 with the Date field as correctly changing to “Date” datatype. Only then the analysis can be done based on the Date.
1. Give a suitable color for the Canva Background as follows;
 - Go to “Visualizations→Format your report page”
 - Click on “Canvas Background”. Choose a suitable color and reduce the Transparency to see the chosen color in the Canva background.
 2. Add a title for the Canva with the name “**Customer & Invoice**” as we have the data from these two tables to generate the visuals. Title can be added as follows;

- Click on “Text Box” from the ribbon.
 - Type the title as “Customer & Invoice”, make it Bold, give suitable font colour and make it Center Justified. Drag and bring it to the middle of the Canva background.
3. Create a visual for the “City-wise Total-Sales” as follows;
- Go to “Visualizations→Build Visual”
 - Choose “Stacked Bar Chart”
 - Drag “City” to Y-axis & “Sales” to X-axis
 - Check on the data by choosing “Focus Mode” on the top right corner of the chart. When we hover the mouse, we can get the sales details of each city as a tooltip.
4. Create a visual for the “Province-wise Sum-of-Sales” as follows;
- Go to “Visualizations→Build Visual”
 - Choose “Stacked Column Chart”
 - Drag “Province” to Y-axis & “Sales” to X-axis
5. If we want to know the Sales details of the cities coming under a specific Province “Texas”, then just click on “Texas” from the chart. Now, we can see the sales details of a city coming under the province “Texas” by highlighting only that city in the previous chart. This feature in PowerBI is known as INTERACTION.
6. Bring a KPI Card on the Canva as a Key Performance Indicator using the Sales data as follows;
- First of all, make sure that none of the charts in the Canva is selected. If selected, then the new visual that we choose will over-write the existing selected chart.
 - Go to “Visualizations→Build Visual”
 - Choose “Card” as the KPI Card
 - Drag and drop “Sales” field on to this KPI Card to see the Sales details as the Key Performance Indicator in the Organization/company.
7. Suppose if we don’t want the Overall Sales data, but we want year-wise Sales data means, then we can go for an option called “Slicer”. Bring the “Slicer” on to the Canva as follows;
- Go to “Visualizations→Build Visual”
 - Choose “Slicer”
 - Go to “Data” pane and choose “Date→Date Hierarchy” and from this hierarchy, select only “Year”. Now it will display the details of how many year’s data is available with us from the dataset.

- Choose “Focus Mode”
- If we are not satisfied with the visual for the Slicer, then we can change them by clicking on “Format Visual→Slicer Settings” and then from the “Style” choose “Vertical List” as our preference.
- Now, to see the Sales data in the Year 2019, choose “2019” from the Slicer and suppose if we want to know the sales details of only “New York”, then click on “New York” under the “Province-wise Sales Chart”. Now, it shows the Sales of New York in the year 2019 in the “KPI Card” that we have placed on the Canva. This is another example for the INTERACTION among the data in PowerBI.



8 Dashboards - Filters in Power BI, Formatting dashboards.

Choose the Data Set

Open PowerBI→Blank Report→GetData→Excel Workbook

Then, select the dataset “HR Data.xlsx” from the local disk.

Choose “Table 1” from the Navigator Window & now choose the option “Transform”.

Transform the dataset

Add a new Column “Attrition Count” by adding new Conditional Column as shown below;

Add Conditional Column

Add a conditional column that is computed from the other columns or values.

New column name:

Attrition Count

Column Name	Operator	Value	Output
if Attrition	=	ABC 123 Yes	Then ABC 123 1

Add Clause

Else

ABC 123

0

OK

Cancel

Change the datatype of the newly created column “Attrition Count” as “Whole Number”

Click on “Close & Apply”

Canvas Design

1. Choose a background colour for the canvas as follows;

Visualization→Format your report page→Canvas Background(Choose a suitable color and reduce the Transparency from 100% to a level as per the colour requirement).

2. Bring a Title for the canvas as follows;

Click on “Text Box” from the Ribbon. Give the title as “Filters and Formatting Dashboard”. Change the Font size, make it to Bold, Underline and bring the text to middle. Give a background color for the Text Box from Effects→Background Color.

3. Bring a KPI Card on the canvas.

Select “Employee Count” for this KPI. Rename the default Title as follows;

Select the KPI Card→Format your visual (from Visualization Pane)→General→Title(make it ON)

Now, give the Text in the title as “Overall Employees”. Change the background colour/Text colour if required as per your choice.

Go to Visual Tab→ OFF the category label

4. Select “Pie Chart” from Visualization pane

Drag "Department" to LEGEND
 Drag "Attrition Count" to VALUES
 Rename the title as "Department-wise Attrition" from;
 Format Visual→General→Title(give name as Department wise attrition)
 Change the background as per your choice.

5. Select "Stacked Column Chart" from Visualization pane

Drag "Age Band" to X-AXIS.
 Drag "Employee Count" to Y-AXIS.
 Drag "Gender" to LEGEND.
 Rename the title as "Age & Gender wise Employee count".

6. Select a "Donut" from Visualization pane

Drag "Gender" to LEGEND.
 Drag "Attrition Count" to VALUES
 Rename the title as "Gender wise Attrition count"

7. Select "Slicer" from Visualization Pane

Go to "Format your visual"→Visual→ Style(give it as TILE).

Drag "Department" to "Filters on this visual".
 Change the title as "Department".
 Drag "Department" to the "FIELD"
 Format your visual→Visual→Slicer Settings→choose style as "Tile"
 Format your visual→ Visual→Selection→choose "Show – Select All" as ON



9 Analysis of revenue in sales dataset:

- i) Create a choropleth map (fill the map) to spot the special trends to show the state which has the highest revenue.
- ii) Create a line chart to show the revenue based on the month of the year.
- iii) Create a bin of size 10 for the age measure to create a new dimension to show the revenue.
- iv) Create a donut chart view to show the percentage of revenue per region by creating zero access in the calculated field.
- v) Create a butterfly chart by reversing the bar chart to compare female & male revenue based on product category.
- vi) Create a calculated field to show the average revenue per state & display profitable & non-profitable state.
- vii) Build a dashboard.

Choose the Dataset "Revenue Sales Data.xlsx". As the data is clean, there is no need of a Transformation. So, we can load the data directly to canvas.

1. Bring the "Filled Map" to the Canvas Background that highlight the state which has the highest revenue.

Drag "State" to the "LOCATION".

Drag "State" to the "LEGEND" in order to see the states in different colour in the map.

Drag "Revenue" to the "TOOLTIP" in order to see the revenue details of each state as we move the Mouse over the state.

Choose "Maximum" from the "Revenue" in the TOOLTIP, in order to see the states which has the maximum revenue.

2. Create a Line Chart to show the revenue based on the month

Drag "Revenue" to the Y-AXIS

Drag "Month" to the X-AXIS

3. Create a Stacked Column Chart to create a Bin of size 10 on Customer Age to show the revenue related to that age group

Right-click on the field "Customer Age", choose "New Group".

Give the "Bin Size" as 10 and click "OK". Now a new field with the name "Customer Age(Bin)" will appear on the Data Pane.

Drag "Customer Age Bin" to "X-AXIS"

Drag "Revenue" to "Y-AXIS"

4. Create a Donut chart to show the percentage of revenue per region by creating zero access in the calculated field.

Drag "State" to "LEGEND"

Drag "Revenue" to "VALUES"

To show the percentage of revenue inside each revenue shown, go to
"Format Visual→Visual→Detail Labels→Options→Position→Inside"

5. Create a Butterfly chart by reversing the Bar chart to compare Female & Male revenue based on product category.

Bring a Stacked Column Chart to the canvas.

Drag "Product Category" to "X-AXIS"

Drag "Revenue" to "Y-AXIS"

Drag "Customer Gender" to "LEGEND"

6. Create a calculated field to show the average Revenue per state & display Profitable and Non-Profitable state.

Click on "New Measure" from the ribbon(HOME Tab)

Then, give the formula as given below;

```
AveragePerState=AVERAGEX(VALUES(SalesTable[State]),  
CALCULATE(sum(SalesTable[Revenue])))
```

Bring a "Table" to the canvas.

Drag "State" to the "COLUMNS"

Drag "AveragePerState" also to "COLUMNS"

Now, to display whether its Profitable or Not-Profitable, as we don't have any field with such data, we need to create a new column as follows;

Go to "Modelling→New Column". Here give the formula as follows;

```
ProfileStatus = IF(SalesTable[AveragePerState]>1000,"Profitable","Not-  
Profitable")
```

(Considering AveragePerState is greater than 1000 as Profitable)

Now, drag "ProfitStatus" to the "COLUMNS".

7. Give a Title to this Dashboard as "Revenue Sales Dashboard" and change the background as per your choice.



10

Analysis of GDP dataset:

- Visualize the countries data given in the dataset with respect to latitude and longitude along with country name using symbol maps.
- Create a bar graph to compare GDP of Belgium between 2006 – 2026.
- Using pie chart, visualize the GDP of India, Nepal, Romania, South Asia, Singapore by the year 2010.
- Visualize the countries Bhutan & Costa Rica competing in terms of GDP.
- Create a scatter plot or circle views of GDP of Mexico, Algeria, Fiji, Estonia from 2004 to 2006.
- Build an interactive dashboard.

Connect → To a File → Microsoft Excel → “GDP by Country per Year_data.xlsx”

Double-click on the sheet which has been verified & validated located on the left hand side of the Tableau.

GDP by Country...ar_data Table2

- Visualize the countries data given in the dataset with respect to latitude and longitude along with country name using symbol maps.

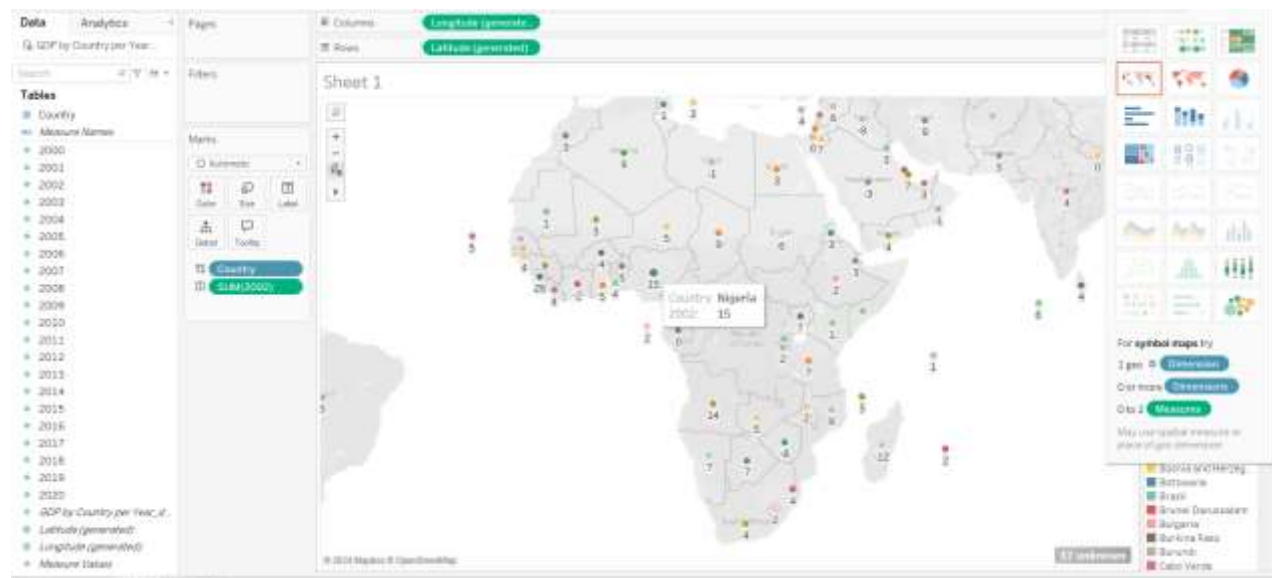
Drag “Latitude” to the “ROWS”

Drag “Longitude” to the “COLUMNS”

Drag “Country” into “Color Marks Pane”

Drag any Year Measured Value to the LABEL

Now, we can see the GDP of each country during that particular year chosen.



ii) Create a bar graph to compare GDP of Belgium between 2006 – 2026.

Drag “Measure Names” to “Filter” and select the years from 2006 to 2012.

Drag “Country” to “Filter” and choose only “Belgium”

Drag “Measure Names” & “Country” into the “Columns”.

Drag “Measure Values” into the “Rows”.



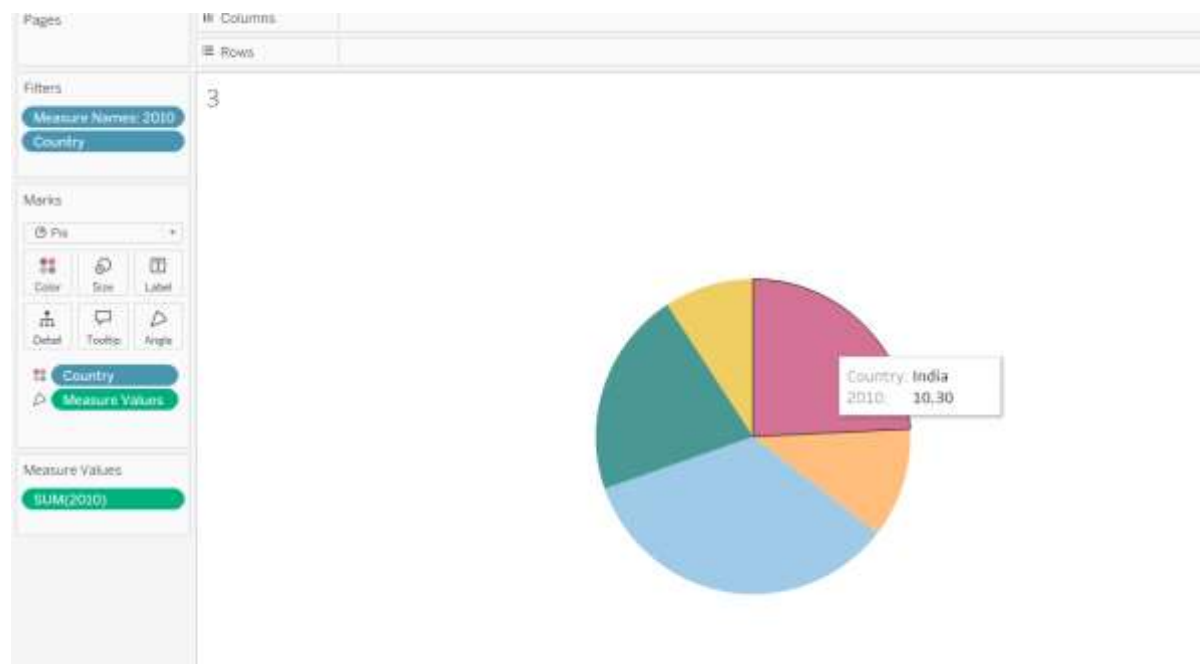
iii) Using pie chart, visualize the GDP of India, Nepal, Romania, South Asia, Singapore by the year 2010.

Drag “Country” to the “Filter Pane” and choose “India, Nepal, Romania, South Asia, Singapore”

Drag “Measure Name” to the “Filter Pane” and choose the year “2010”.

Select option of chart as Pie (instead of automatic in Marks Pane) and Drag “Country” into Color frame.

Drag “Measure Values” into the “Angle”.



iv) Visualize the countries Bhutan & Costa Rica competing in terms of GDP.

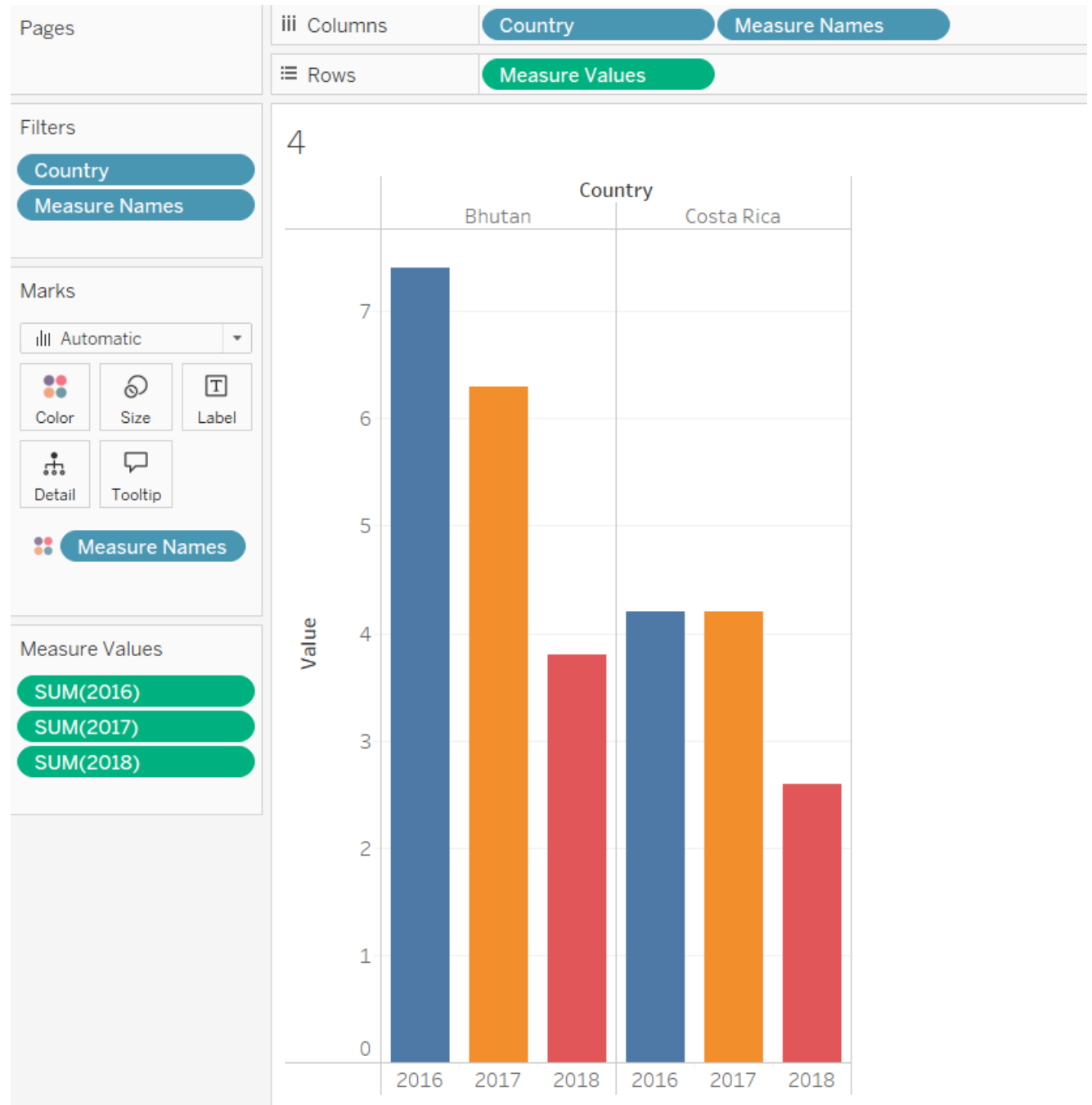
Drag “Country” into the “Filter Pane” and choose Bhutan, Costa Rica.

Drag “Measure Names” into the “Filter Pane” and choose the years 2016,2017,2018

Now, drag “Country” and “Measure Names” in column

Drag “Measure Values” in Row

For better view, drag “Measure Names” into the Color frame in Marks pane



v) Create a scatter plot or circle views of GDP of Mexico, Algeria, Fiji, Estonia from 2004 to 2006.

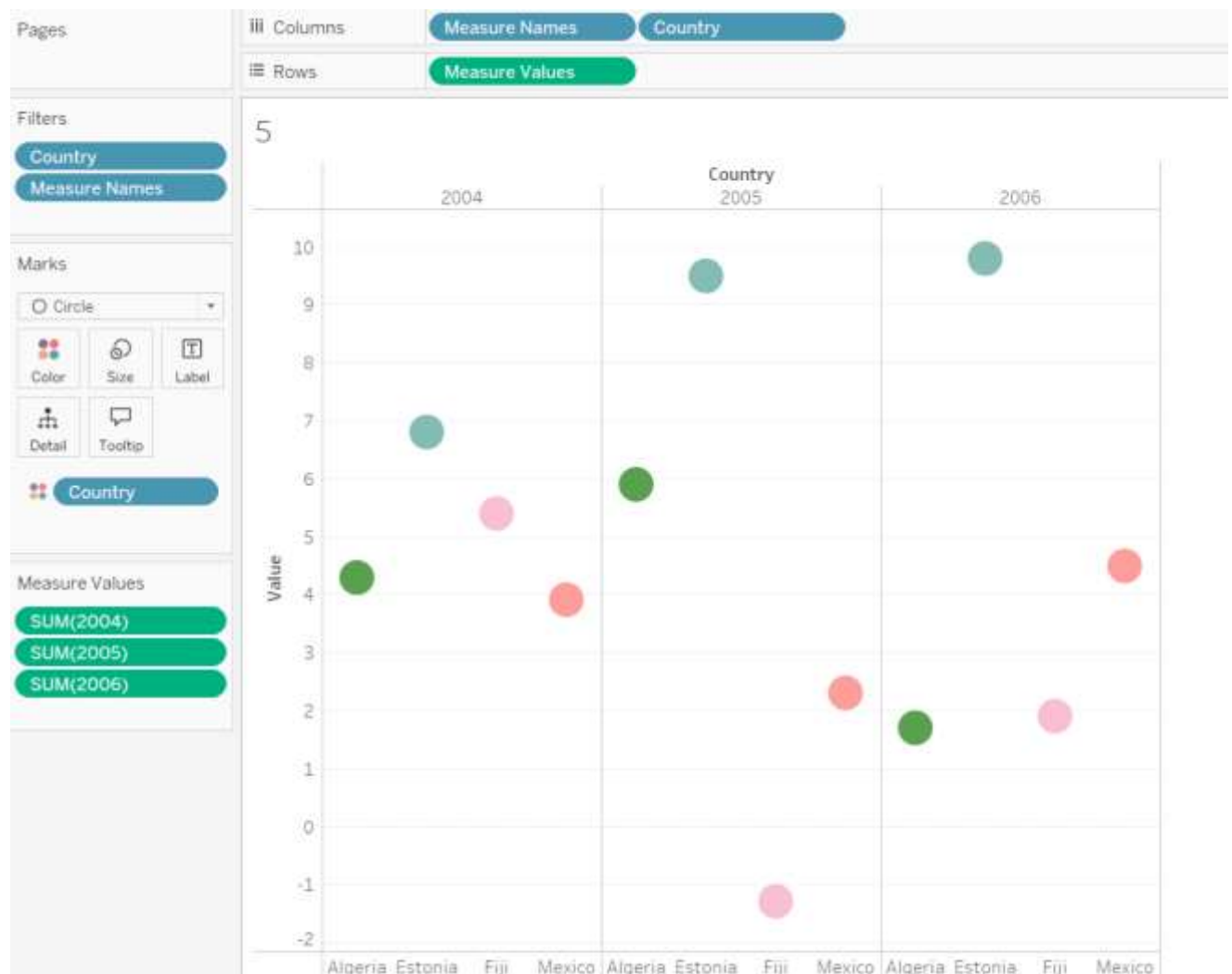
Drag “Country” into the “Filter” and choose “Mexico, Algeria, Fiji, Estonia”

Drag “Measure Names” into the “Filter” and choose the year 2004,2005,2006.

Drag “Measure Names” and “Country” to the “COLUMNS”

Drag “Measure Values” to the “ROWS”

For a better view, drag “Country” to the “Color” in the Marks Pane.



vi) Build an interactive dashboard.

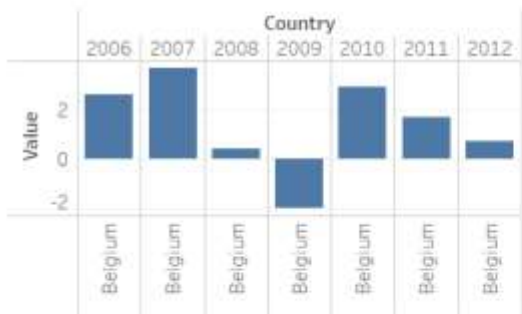
Create a new Dashboard.

Choose Floating type Dashboard

Drag the Sheets mentioned from Task 2 to Task 5 into the Dashboard

GDP Dashboard

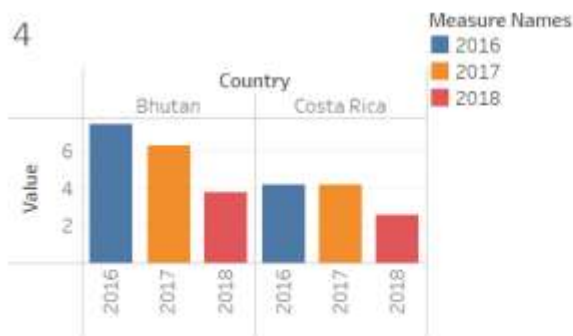
2



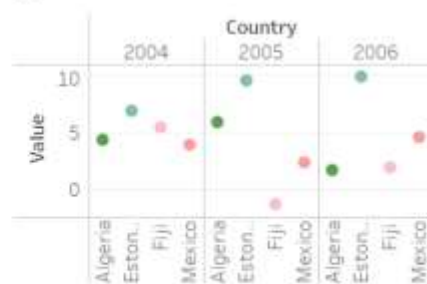
3



4



5



11

Analysis of HR Dataset:

- i) Create KPI to show employee count, attrition count, attrition rate, attrition count, active employees, and average age.
- ii) Create a Lollipop Chart to show the attrition rate based on gender category.
- iii) Create a pie chart to show the attrition percentage based on Department Category- Drag department into colours and change automatic to pie. Entire view, Drag attrition count to angle. Label attrition count, change to percent, add total also, edit label.
- iv) Create a bar chart to display the number of employees by Age group,
- v) Create a highlight table to show the Job Satisfaction Rating for each job role based on employee count.
- vi) Create a horizontal bar chart to show the attrition count for each Education field Education field wise attrition – drag education field to rows, sum attrition count to col,
- vii) Create multiple donut chart to show the Attrition Rate by Gender for different Age group.

Choose the Dataset “HR Analytics Data.csv”. As the data is clean, there is no need of a Transformation. So, we can load the data directly to canvas.

Change the Table name from “HR Analytics Data” to “HR” by just double-clicking on the table name, for our easy access purpose.

1. Create KPI to show employee count, attrition count, attrition rate, attrition count, active employees, and average age.

Create a background colour for the Canvas and give a title as “HR Dashboard”.

a) Bring a KPI Card to the dashboard for Employee Count.

Create a New Measure as follows;

Employee Count = COUNT(HR[EmployeeNumber])

Drag and drop this new measure “Employee Count” to “Fields”.

b) Bring a KPI Card to the dashboard for Attrition Count.

Create a New Measure as follows;

Attrition Count = COUNTROWS(FILTER('HR',HR[Attrition]="Yes"))

Drag & drop this new measure to the “Fields”.

c) Bring a KPI Card to the dashboard for Attrition Rate

Create a New Measure as follows;

Attrition Rate = DIVIDE([Attrition Count],[Employee Count],0) * 100

Drag & drop this new measure to the “Fields”.

d) Bring a KPI Card to the dashboard for Active Employee

Create a New Measure as follows;

Active Employees = [Employee Count]- [Attrition Count]

Drag & drop this new measure to the “Fields”.

e) Bring a KPI Card to the dashboard for Average Age

Create a New Measure as follows;

Average Age = AVERAGE(HR[Age])

Drag & drop this new measure to the “Fields”.

2. Create a Lollipop Chart to show the attrition rate based on gender category.

Bring a Line & Stacked Column chart as Lollipop chart is not available in PowerBi.

Drag “Gender” to “X-AXIS”

Drag “Attrition Count” to “COLUMN Y-AXIS”

Drag “Attrition Rate” to “LINE Y-AXIS”

3. Create a pie chart to show the attrition percentage based on Department Category- Drag department into colours and change automatic to pie. Entire view, Drag attrition count to angle. Label attrition count, change to percent, add total also, edit label.

Bring a Pie Chart to the Canvas.

Drag "Department" to "LEGEND"

Drag "Attrition Count" to "VALUES"

4. Create a bar chart to display the number of employees by Age group.

Create a BIN with size 10 to get the Age group as follows;

Right-click on "Age" → New Group

Rename as "Age Group", give Bin Size as 10 and click OK.

Bring Stacked Bar Chart to the canvas.

Drag "Age Group" to "Y-AXIS"

Drag "Employee Count" to "X-AXIS"

5. Create a highlight table to show the Job Satisfaction Rating for each job role based on employee count.

Bring Matrix to the canvas

Drag "Job Role" to "ROWS"

Drag "Job Satisfaction" to "COLUMNS"

Drag "Employee Count" to "VALUES"

6. Create a horizontal bar chart to show the attrition count for each Education field
Education field wise attrition – drag education field to rows, sum attrition count to col.

Bring Stacked Bar Chart to the canvas

Drag "Education Field" to "Y-AXIS"

Drag "Attrition Count" to "X-AXIS"

7. Create multiple donut chart to show the Attrition Rate by Gender for different Age group.

Bring Donut - 1 to the canvas

Drag "Gender" to "LEGEND"

Drag "Attrition Rate" to "VALUES"

Now drag "Age Group" to "Filters on this Visual" and select the age 30, to show the attrition under the Age 30.

Repeat this by creating Donut-2 by choosing another Age group as under 50.



12

Analysis of Amazon Prime Dataset:

- Create a Donut chart to show the percentage of movie and tv shows
- Create a area chart to shows by release year and type
- Create a horizontal bar chart to show Top 10 genre
- Create a map to display total shows by country
- Create a text sheet to show the description of any movie/movies.
- Build an interactive Dashboard.

Choose the Dataset "Amazon-Prime-Titles.csv". Transform the data as follows;
 Rename the field "listed in" as "Genre"
 Remove "Duration", "Rating", "date_added", "cast", "director"
 Click "Close & Apply"

Change the Table name from "Amazon-Prime-Titles" to "Amazon" by just double-clicking on the table name, for our easy access purpose.

- Create a Donut chart to show the percentage of movie and tv shows.

Bring a Donut into the canvas.

Drag "type" into "LEGEND". (If the type is showing other than MOVIE & TV Show, then transform and choose only these two values)

Drag "show id" into "VALUES"

2. Create an area chart to shows by release year and type.

Bring Area Chart into the canvas.

Drag "Release Year" into "X-AXIS"

Drag "type" into "Y-AXIS"

Drag "type" into "LEGEND" to analyse both TV Shows & Movie separately.

3. Create a horizontal bar chart to show Top 10 genre.

Bring Clustered Bar Chart into the canvas

Drag "Genre" into "Y-AXIS"

Drag "title" into "X-AXIS"

To display only Top 10 Genre, do the following steps;

In Filters, choose Filter Type as "Top N", mention the value "10" under show items , Drag "Genre" into "By Value".

4. Create a map to display total shows by country.

Bring Filled Map into the canvas.

Drag "Country" into "LOCATION"

Drag "Show ID" into "TOOLTIP" and change it to "Count" of "ShowID" under ToolTip

5. Create a text sheet to show the description of any movie/movies.

Bring Table into the canvas.

Drag "title" into "COLUMNS"

Drag "description" into "COLUMNS"

6. Build an interactive Dashboard.

Change the Background colour of the canvas & give a title as "Amazon Prime Dashboard" for the tasks performed above.

