Tableau

Mohan Kumar Pulibanti

Mohan Kumar Pulibanti

What is Tableau?

Tableau is a powerful and widely used data visualization and business intelligence software. It allows users to connect, visualize, and share data in a way that enables insights and decision-making. Tableau is known for its user-friendly interface and the ability to create interactive, shareable dashboards.

Tableau installation:

Download Links:

Tableau Public: https://www.tableau.com/products/public/download
Tableau: https://www.tableau.com/products/trial

Download the executable from the above links and run that to install Tableau.

- Tableau is a licensed tool and comes with a 14-day free trial.
- Tableau Public is free, so it doesn't have any limitations in terms of the amount of data that you can use it for, it's completely free software. It has slight restrictions; for example, you cannot create extracts and save files directly onto your computer. When data security is not a big deal, we can go with this.

Tableau's public Home page looks as below:

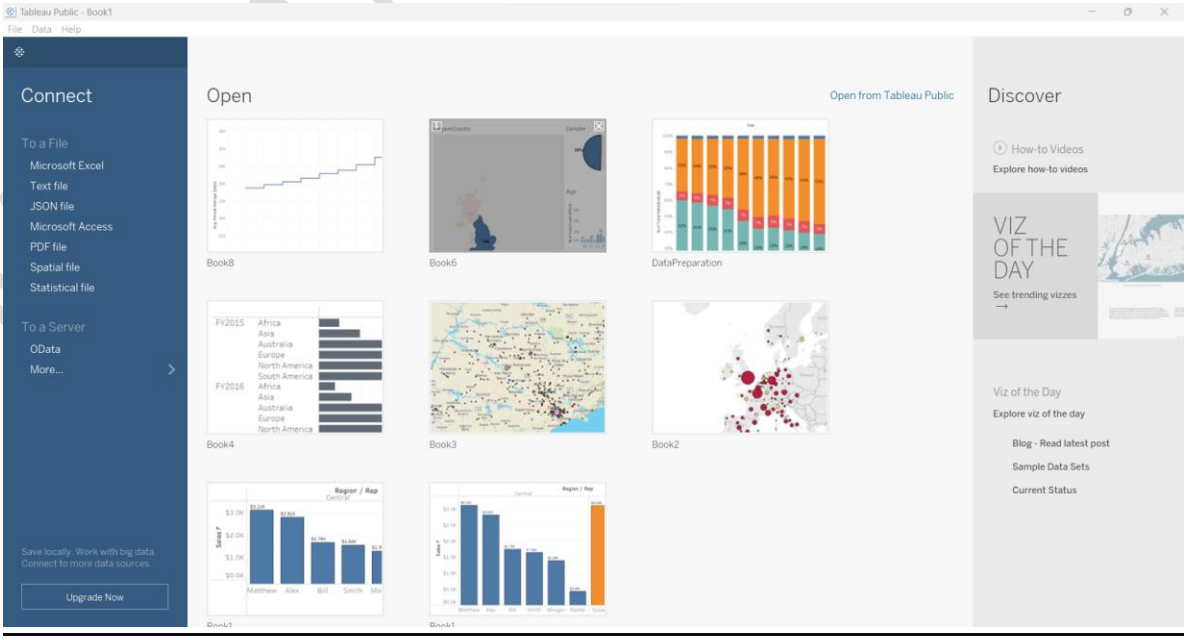


Tableau Basics: Your First Bar Chart

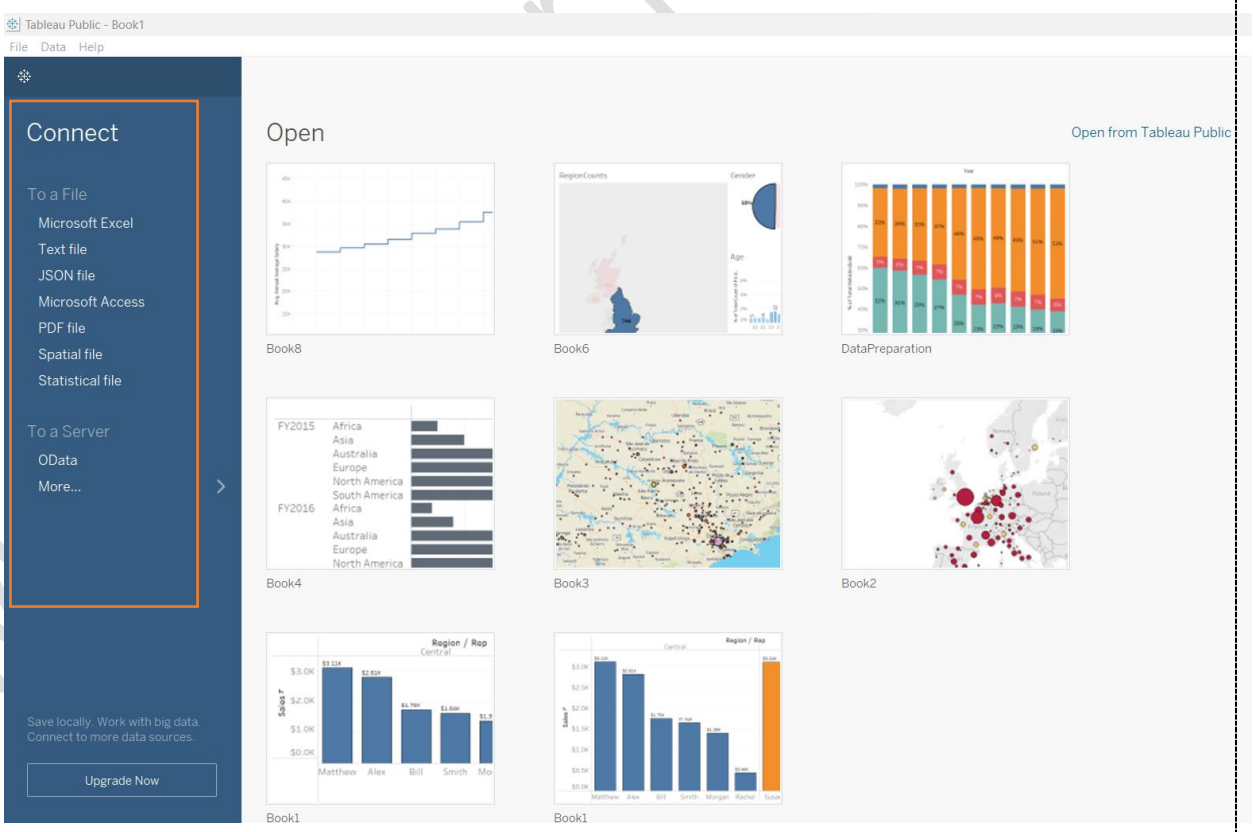
Business Challenge – Who gets the Annual Bonus?

The store operates in three regions and only the top-performing employee in each region qualifies for a bonus. Find out which three employees are eligible to get a bonus for this year. The criteria to choose a Top-performing Employee is the total value of sales.

Download Dataset here:
[Tableau_Data_Sets/OfficeSupplies_US_Data.csv at main · MohanPulibanti/Tableau_Data_Sets (github.com)]

Connecting Tableau to a Data File:

1) **When we open Tableau, we will get to a screen where we can connect our data file**



Like, click Text file to connect to a CSV file and then select the file you want to connect from the browse window.

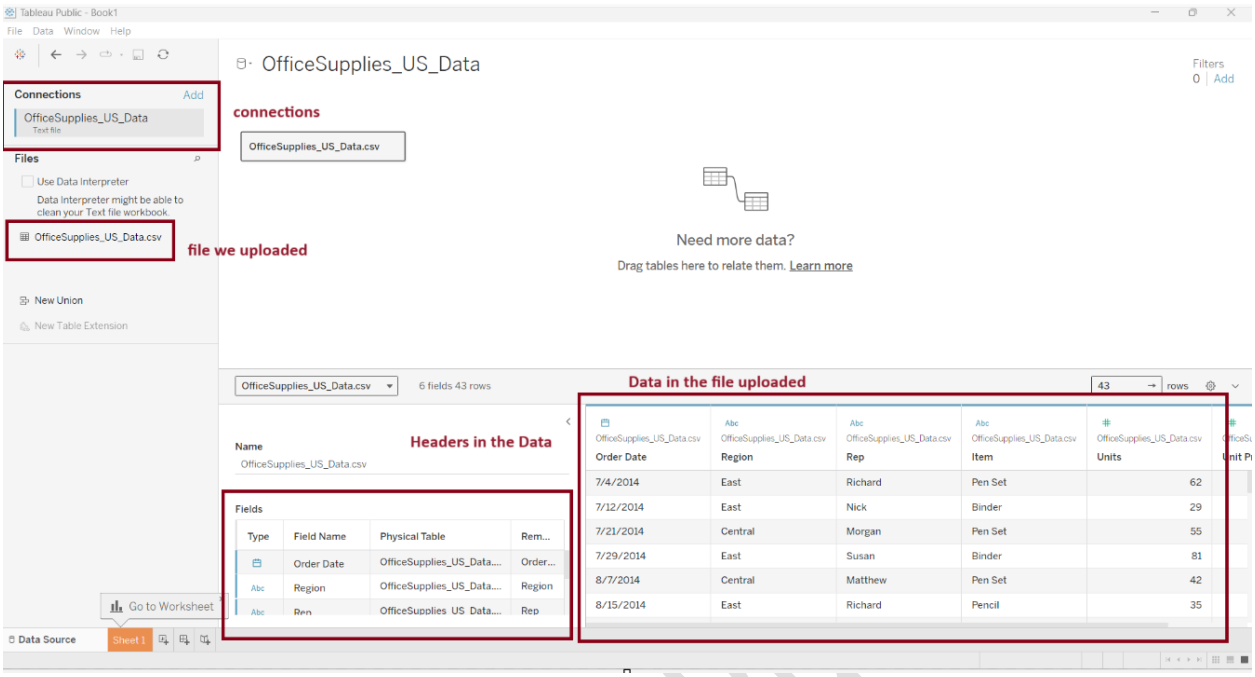


Tableau Public - Book1
File Data Window Help
OfficeSupplies_US_Data
Connections
Add
OfficeSupplies_US_Data
Text file
connections
OfficeSupplies_US_Data.csv
Files
Use Data Interpreter
Data Interpreter might be able to clean your Text file workbook.
OfficeSupplies_US_Data.csv
file we uploaded
New Union
New Table Extension
Filters
0 | Add
Need more data?
Drag tables here to relate them. Learn more
OfficeSupplies_US_Data.csv
6 fields 43 rows
Data in the file uploaded
43 rows
Headers in the Data
Name
OfficeSupplies_US_Data.csv
Fields
Type
Field Name
Physical Table
Order Date
Region
Rep
Item
Units
7/4/2014 East Richard Pen Set 62
7/12/2014 East Nick Binder 29
7/21/2014 Central Morgan Pen Set 55
7/29/2014 East Susan Binder 81
8/7/2014 Central Matthew Pen Set 42
8/15/2014 East Richard Pencil 35
Go to Worksheet
Data Source
Sheet 1

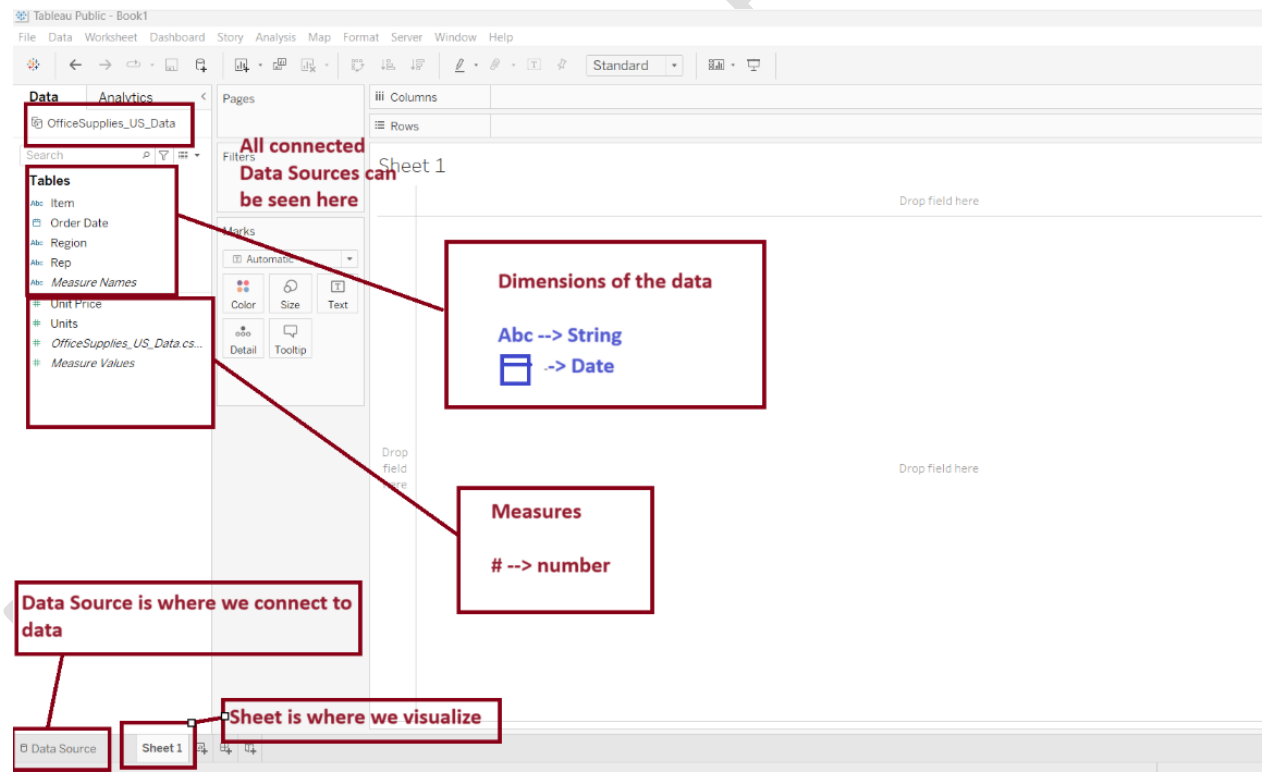
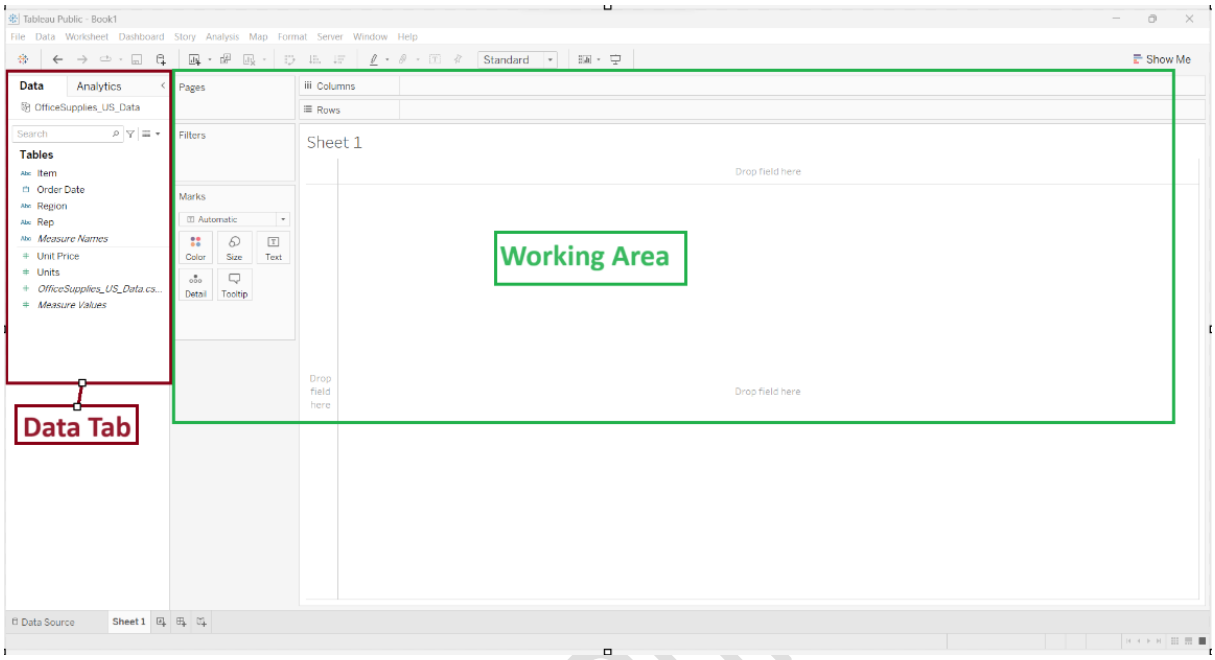
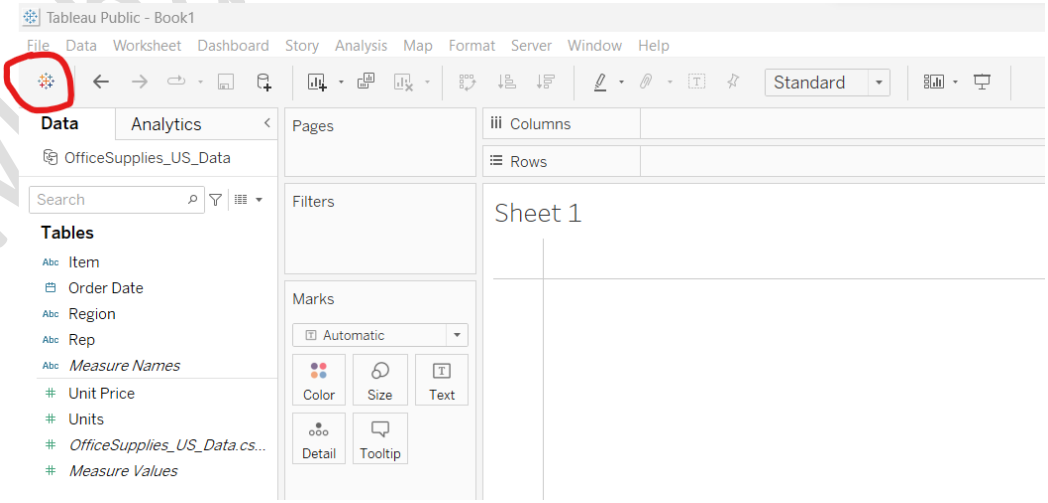


Tableau Public - Book1
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Data
Analytics
OfficeSupplies_US_Data
Search
Tables
Item
Order Date
Region
Rep
Measure Names
Unit Price
Units
OfficeSupplies_US_Data.cs...
Measure Values
Pages
Filters
Marks
Automatic
Color
Size
Text
Detail
Tooltip
Columns
Rows
Sheet 1
Drop field here
All connected Data Sources can be seen here
Dimensions of the data
Abc --> String
-> Date
Measures
--> number
Data Source is where we connect to data
Sheet is where we visualize
Data Source
Sheet 1

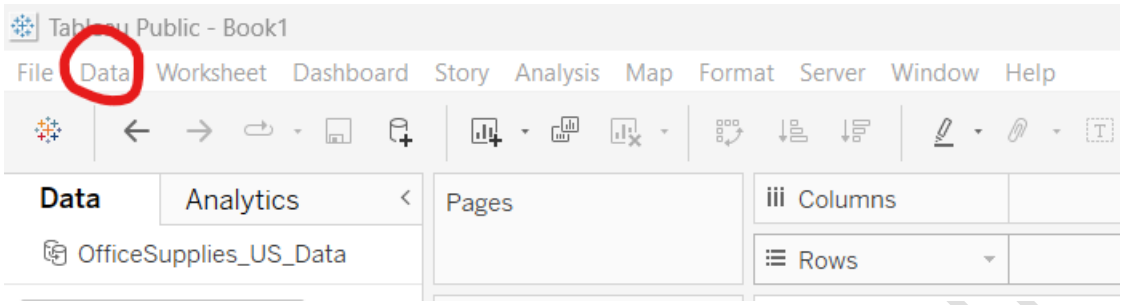


- Data tab is always broken down into two sections, dimensions and measures, and there's a very distinct difference between them. Dimensions and measures are two different roles that any data element can take. Tableau keeps any kind of numerical into measures and it puts all the categorical and kind of qualitative data into Dimensions.
- We can always drag dimensions to measure and measure to dimensions.
- Dimensions are independent variables and measures are dependent variables.
- In the above example, region is the independent variable and Units are the dependent variable as its grouped based on the region.

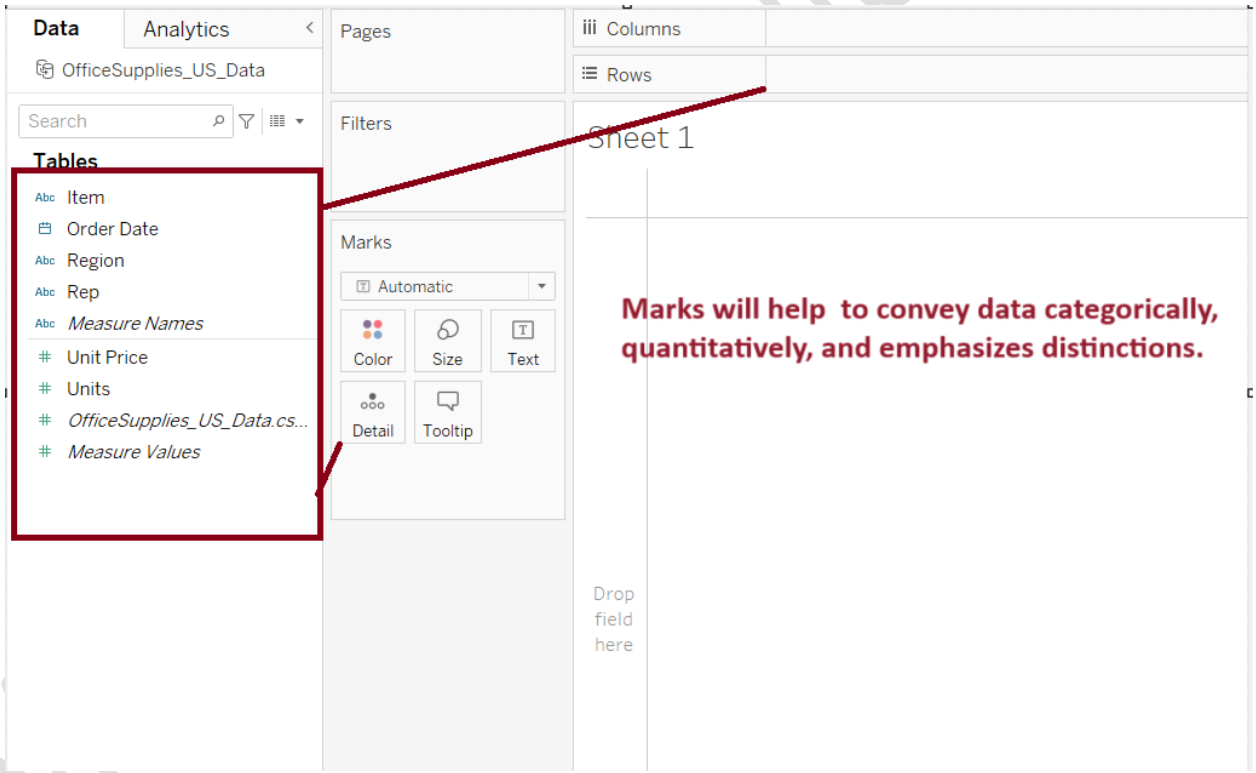
2) The second way to connect to a new data source is to click on the highlighted icon in the below diagram.



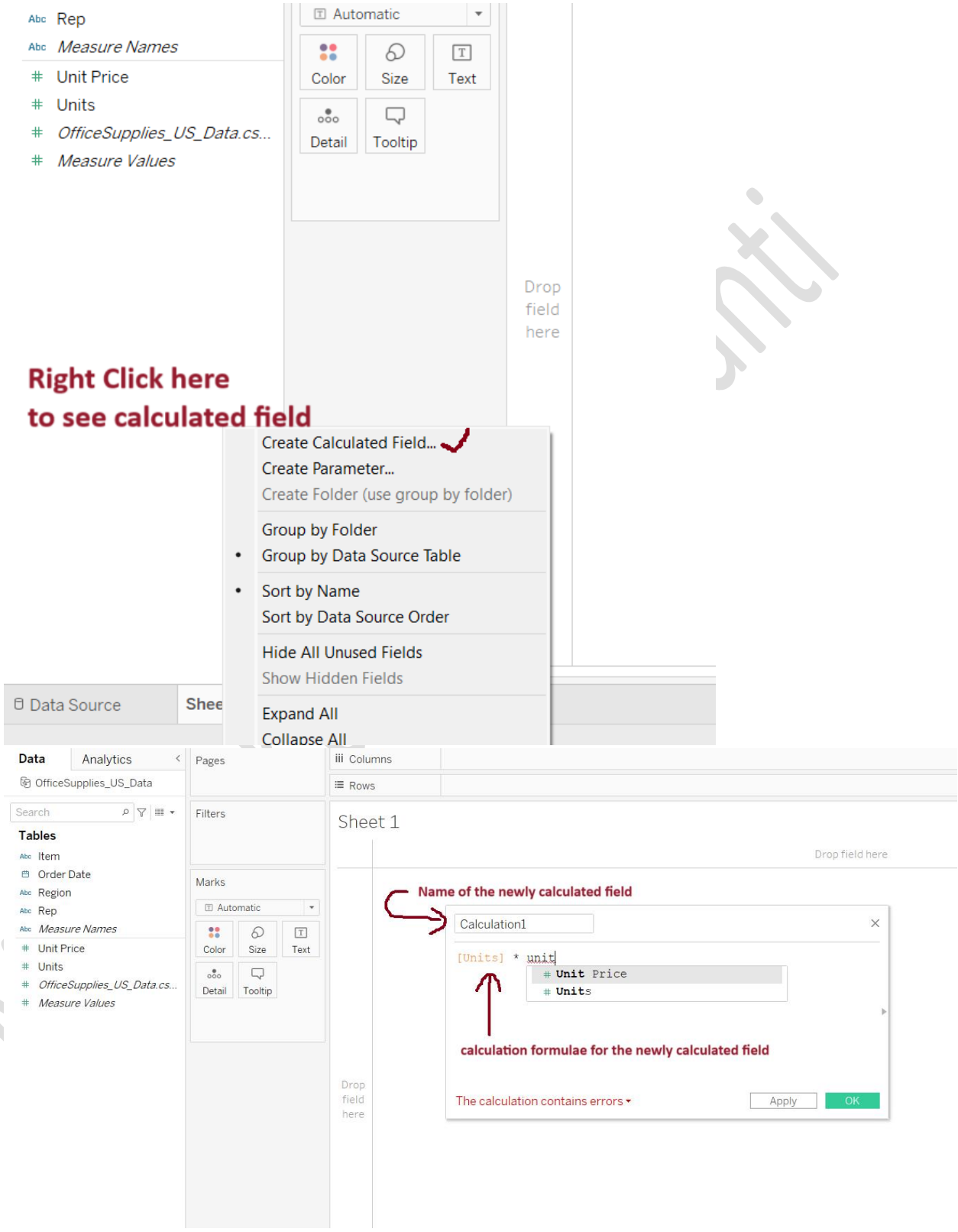
3) **Third option is from the menu Data → New data Source.**



- Drag the Dimensions and measures into rows and columns to see the visualization of the data. Marks will help to convey data categorically, quantitatively and emphasizes distinctions.



- Creating Calculated fields



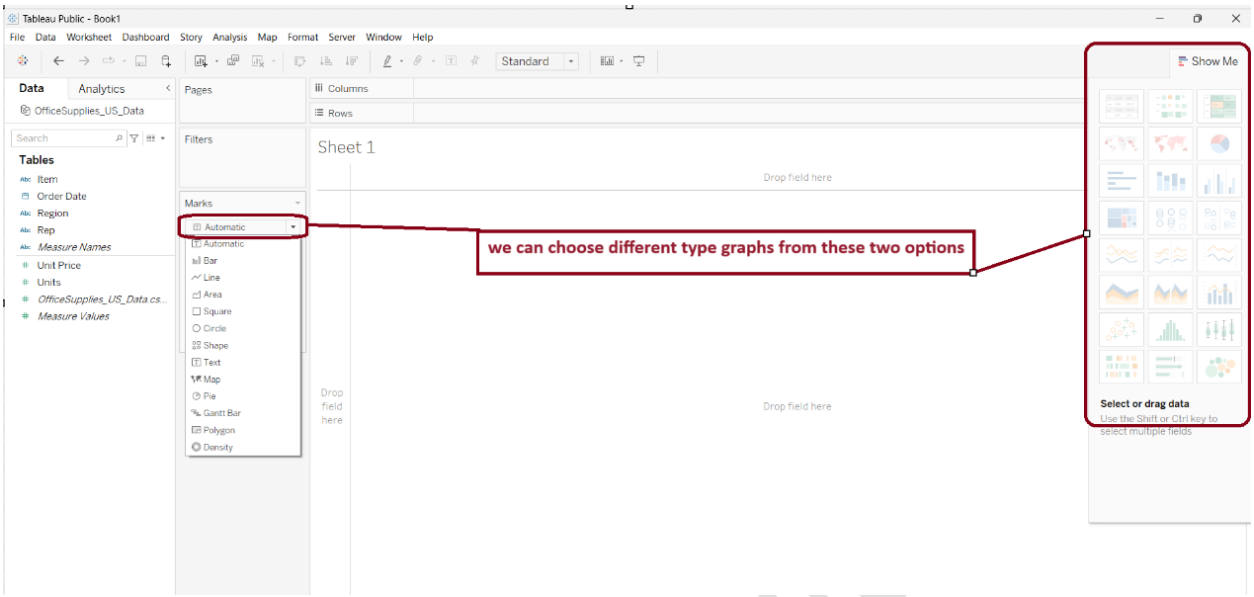
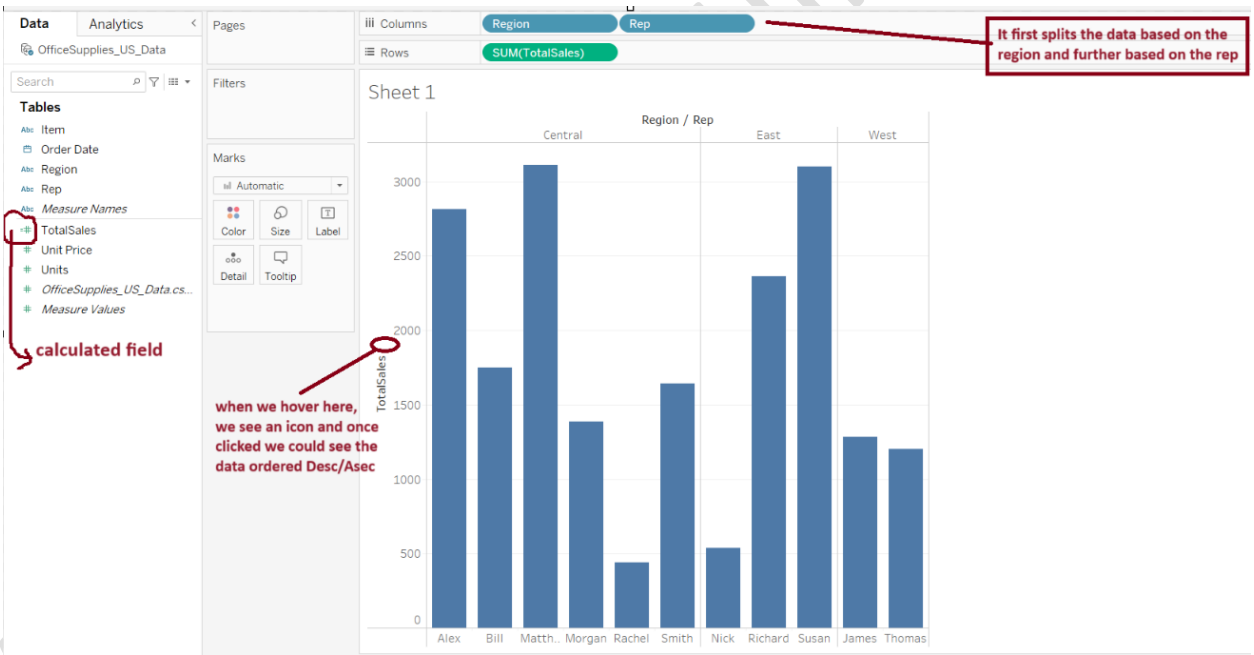


Tableau Public - Book1
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Data
Analytics
OfficeSupplies_US_Data
Tables
Item
Order Date
Region
Rep
Measure Names
Unit Price
Units
OfficeSupplies_US_Data.cs...
Measure Values
Pages
Filters
Marks
Automatic
Bar
Line
Area
Square
Circle
Shape
Text
Map
Pie
Gantt Bar
Polygon
Density
Columns
Rows
Sheet 1
Drop field here
Show Me
Select or drag data
Use the Shift or Ctrl key to select multiple fields
we can choose different type graphs from these two options


Data
Analytics
OfficeSupplies_US_Data
Tables
Item
Order Date
Region
Rep
Measure Names
TotalSales
Unit Price
Units
OfficeSupplies_US_Data.cs...
Measure Values
calculated field
Pages
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Columns
Region
Rep
Rows
SUM(TotalSales)
It first splits the data based on the region and further based on the rep
Sheet 1
Region / Rep
Central
East
West
TotalSales
3000
2500
2000
1500
1000
500
0
Alex
Bill
Matth..
Morgan
Rachel
Smith
Nick
Richard
Susan
James
Thomas
when we hover here, we see an icon and once clicked we could see the data ordered Desc/Asec

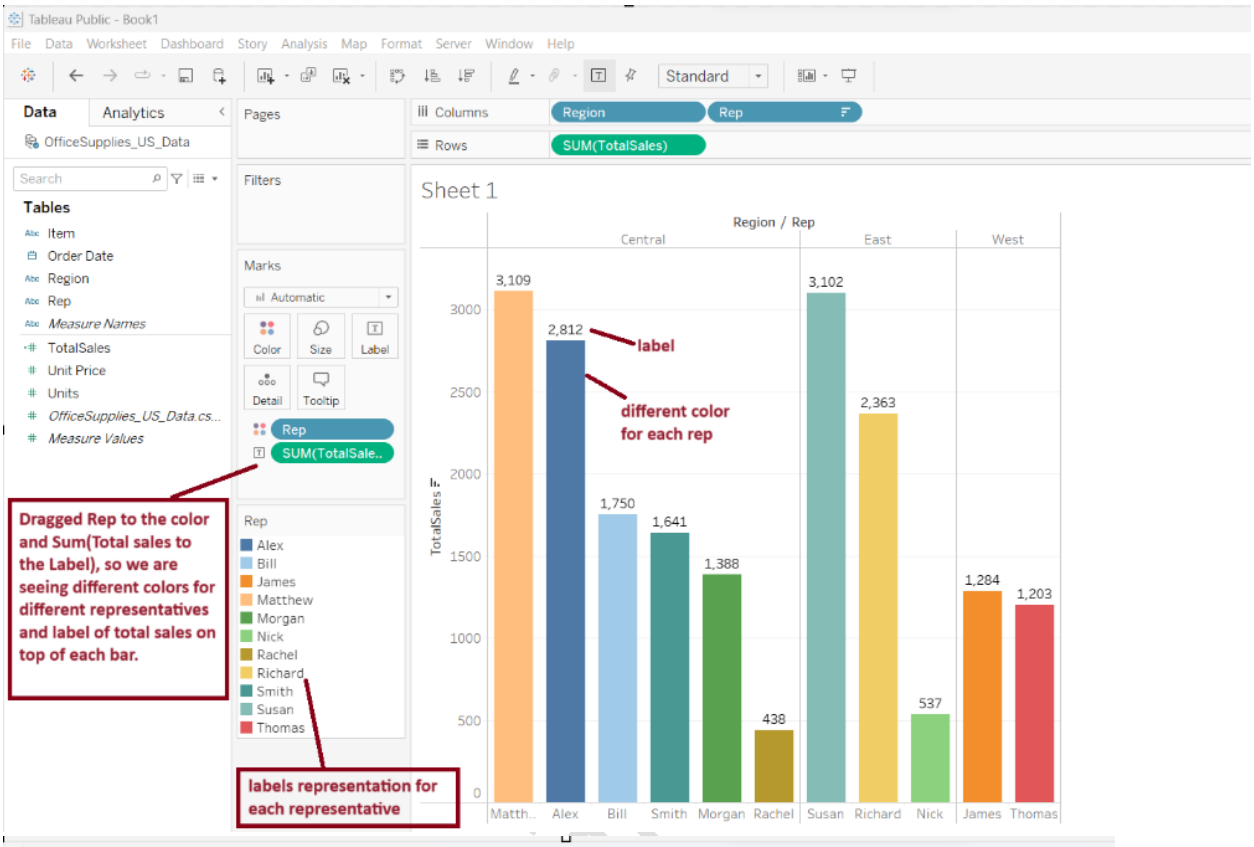


Tableau Public - Book1
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Data
Analytics
OfficeSupplies_US_Data
Tables
Item
Order Date
Region
Rep
Measure Names
TotalSales
Unit Price
Units
OfficeSupplies_US_Data.cs...
Measure Values
Pages
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Rep
SUM(TotalSale..
Columns
Region
Rep
Rows
SUM(TotalSales)
Sheet 1
Region / Rep
Central
East
West
3,109
2,812
1,750
1,641
1,388
438
3,102
2,363
537
1,284
1,203
label
different color for each rep
Dragged Rep to the color and Sum(Total sales to the Label), so we are seeing different colors for different representatives and label of total sales on top of each bar.
Rep
Alex
Bill
James
Matthew
Morgan
Nick
Rachel
Richard
Smith
Susan
Thomas
labels representation for each representative
Matth..
Alex
Bill
Smith
Morgan
Rachel
Susan
Richard
Nick
James
Thomas

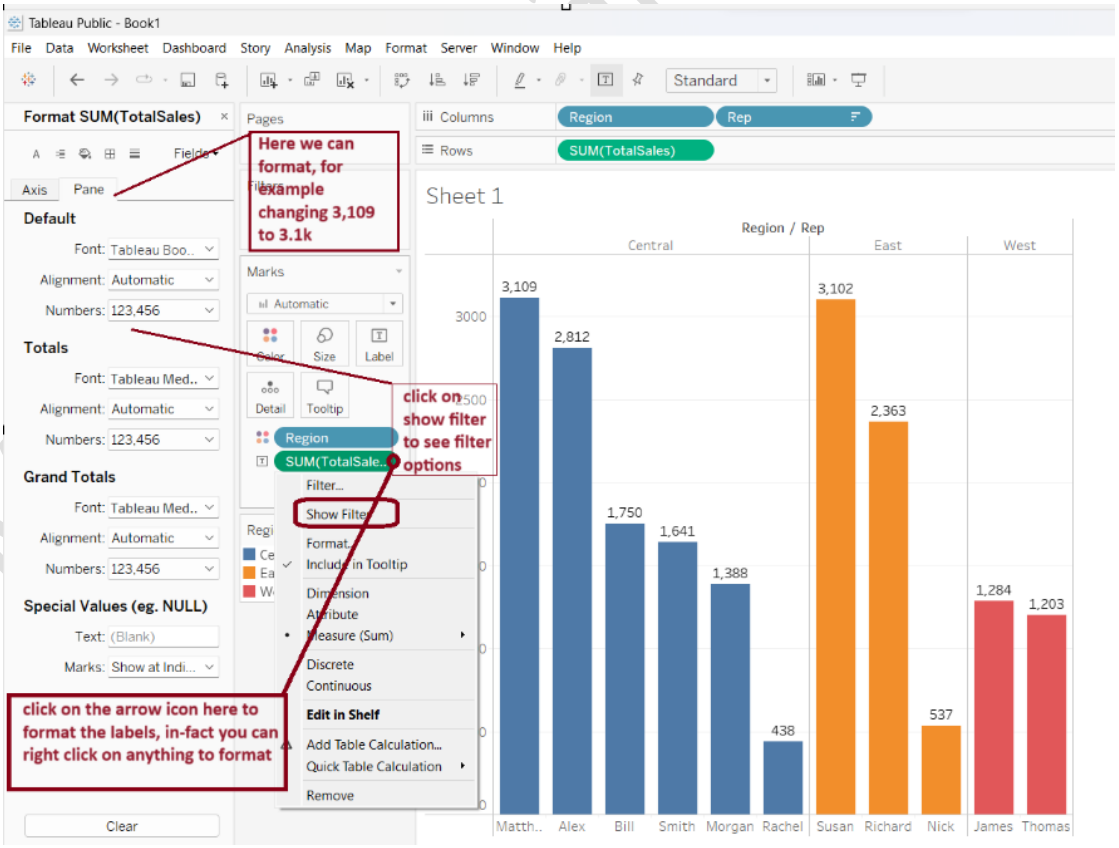
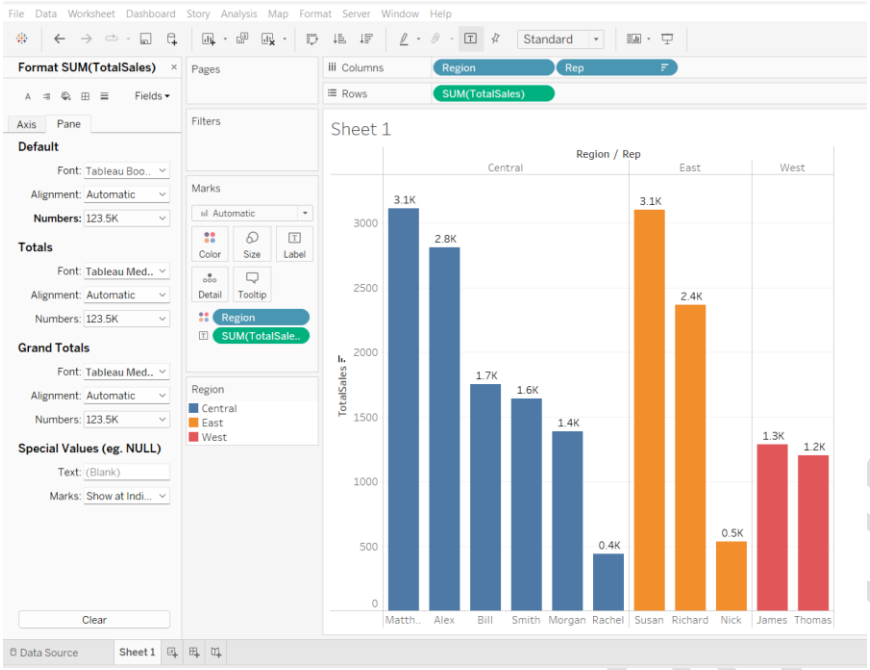
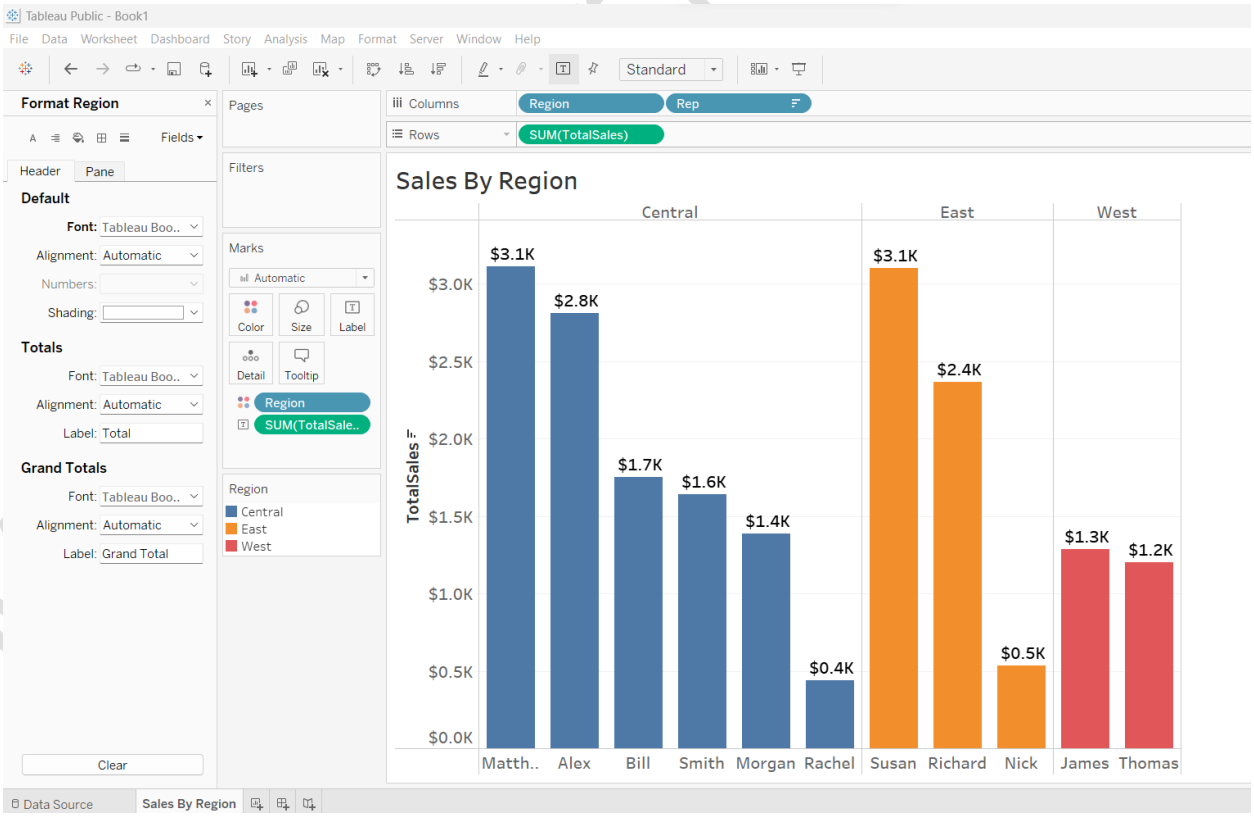


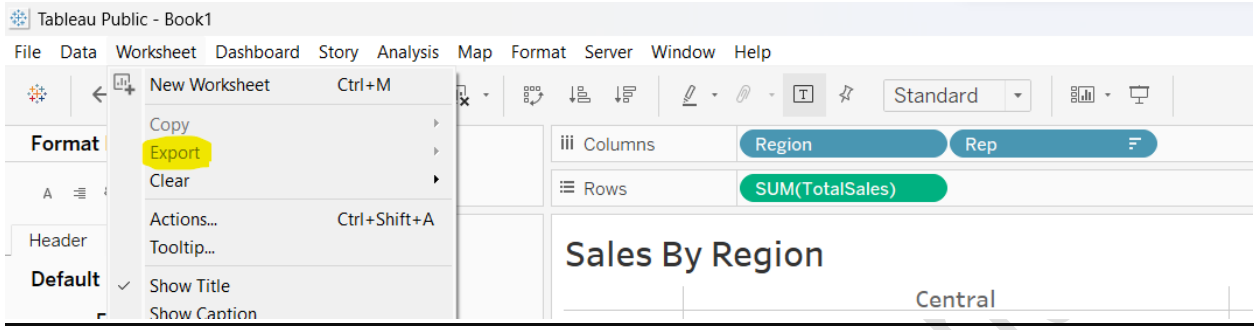
Tableau Public - Book1
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Format SUM(TotalSales)
Axis
Pane
Default
Font: Tableau Boo..
Alignment: Automatic
Numbers: 123,456
Totals
Font: Tableau Med..
Alignment: Automatic
Numbers: 123,456
Grand Totals
Font: Tableau Med..
Alignment: Automatic
Numbers: 123,456
Special Values (eg. NULL)
Text: (Blank)
Marks: Show at Indi...
Clear
Here we can format, for example changing 3,109 to 3.1k
click on show filter to see filter options
click on the arrow icon here to format the labels, in-fact you can right click on anything to format
Pages
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Region
SUM(TotalSale..
Filter...
Show Filter
Format...
Include in Tooltip
Dimension
Attribute
Measure (Sum)
Discrete
Continuous
Edit in Shelf
Add Table Calculation...
Quick Table Calculation
Remove
Columns
Region
Rep
Rows
SUM(TotalSales)
Sheet 1
Region / Rep
Central
East
West
3,109
2,812
1,750
1,641
1,388
438
3,102
2,363
537
1,284
1,203
Matth..
Alex
Bill
Smith
Morgan
Rachel
Susan
Richard
Nick
James
Thomas



After Applying different formats



Export your Worksheet [Not available in Tableau public]



Click on Export → Image / Data / Cross to Excel and then copy and paste in your document.

Time Series, Aggregation, and Filters

Download Dataset here:
[Tableau_Data_Sets/Long_Term_Unemployment_Statistics.xlsx at main · MohanPulibanti/Tableau_Data_Sets (github.com)]

Data Extracts in Tableau [not available in public]:

Data Connections are two types in Tableau.

- Live
  Allows users to directly connect to a data source, providing real-time access to the latest information without the need for manual data refreshes.
- Extract
  1) Tableau's Extract Data Connection enables users to create a static snapshot of data from a source, optimizing performance by storing a subset or the entire dataset locally, which is particularly useful for large datasets or situations where real-time connectivity is not essential.
  2) We can refresh manually to reflect changes in the source data.

- ➢ To extract data, In Tableau Sheets, right-click on the data source and click "extract data" and save to a location. The data will be extracted into a separate file and the data source icon gets changed to a new one in the worksheets.
- ➢ Right-click on the data source and you will see the "use Extract" option gets checked.
- ➢ If we change anything in the source file to which we connected, nothing will change in the tableau as we extract the Data. To get the changes in the extract file, right-click on the data source → Extract → refresh.

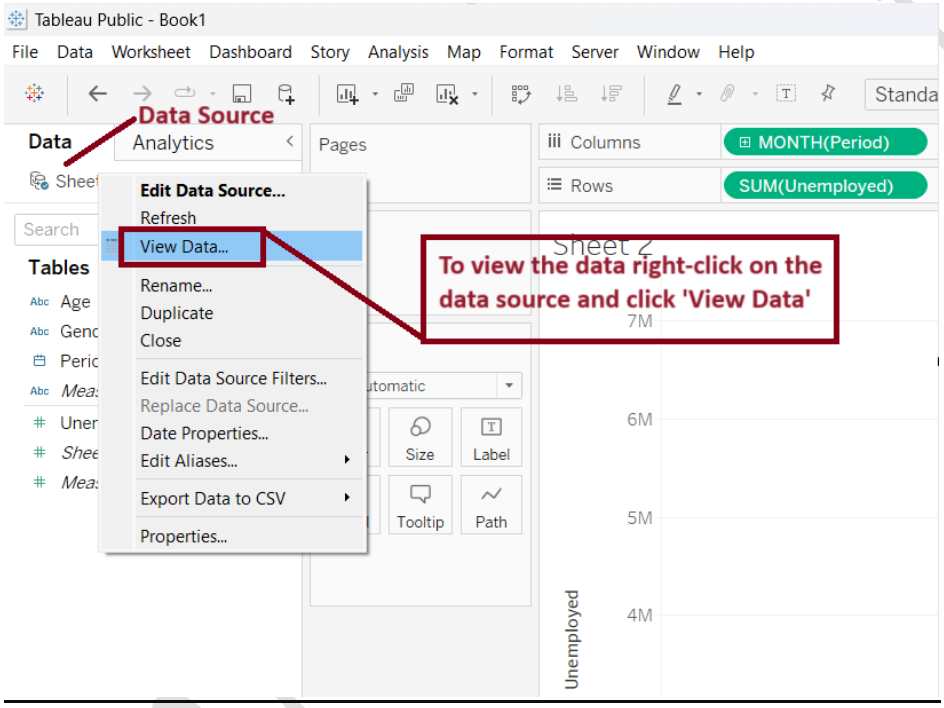
Mohan Kumar Pulibanti

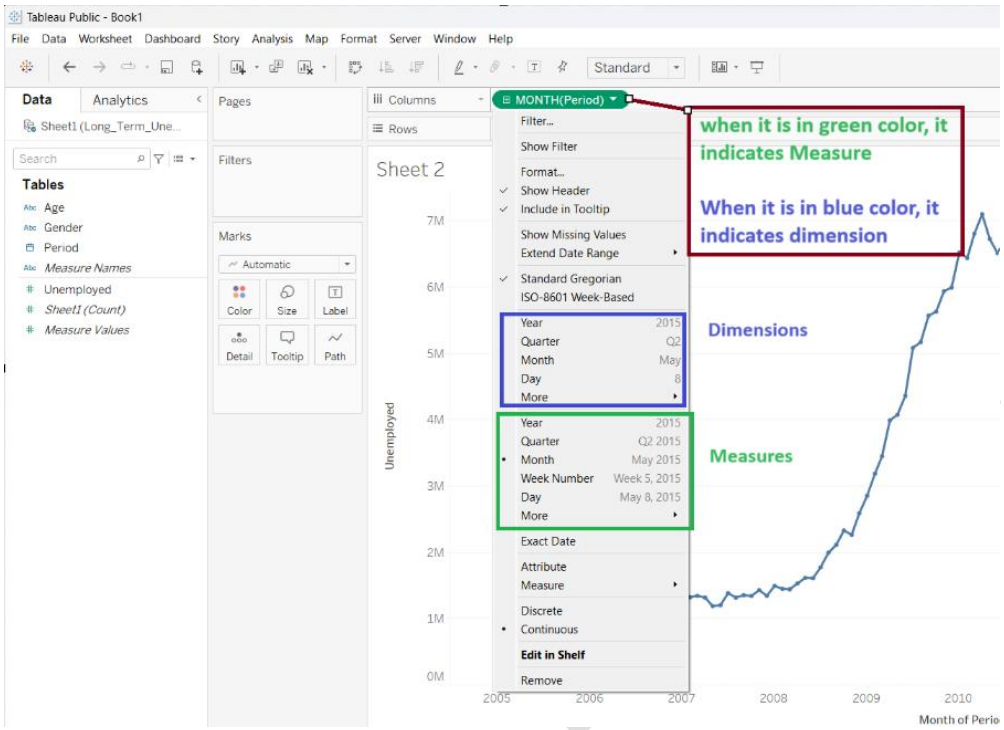
- To come back to the normal connection(live), Right-click on the data source and uncheck “Use Extract”.

Working with time-series

- A time series is a series of data points ordered over time.

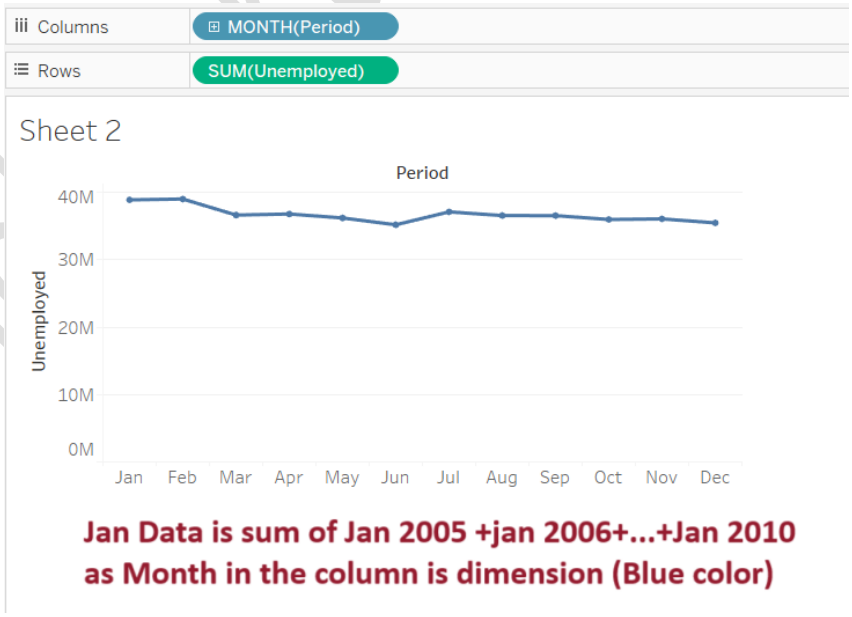
For the above dataset different Tableau plot looks as below.



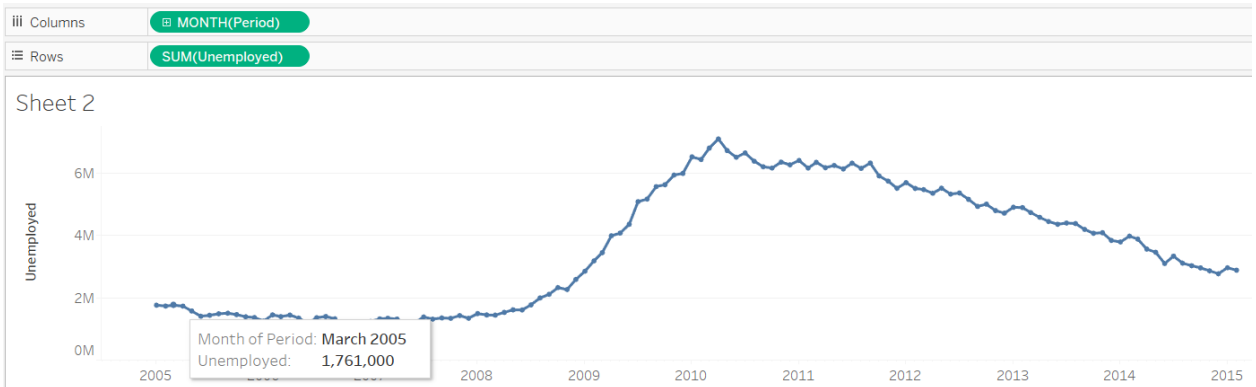


The graph above is showing Unemployed count at different times (2005-2010)

- When column is a dimension (Month) → it calculates the unemployed for each month (accumulating all the year), in the below graph the January column value is the sum of (January 2005 + jan 2006 +….+jan 2010)



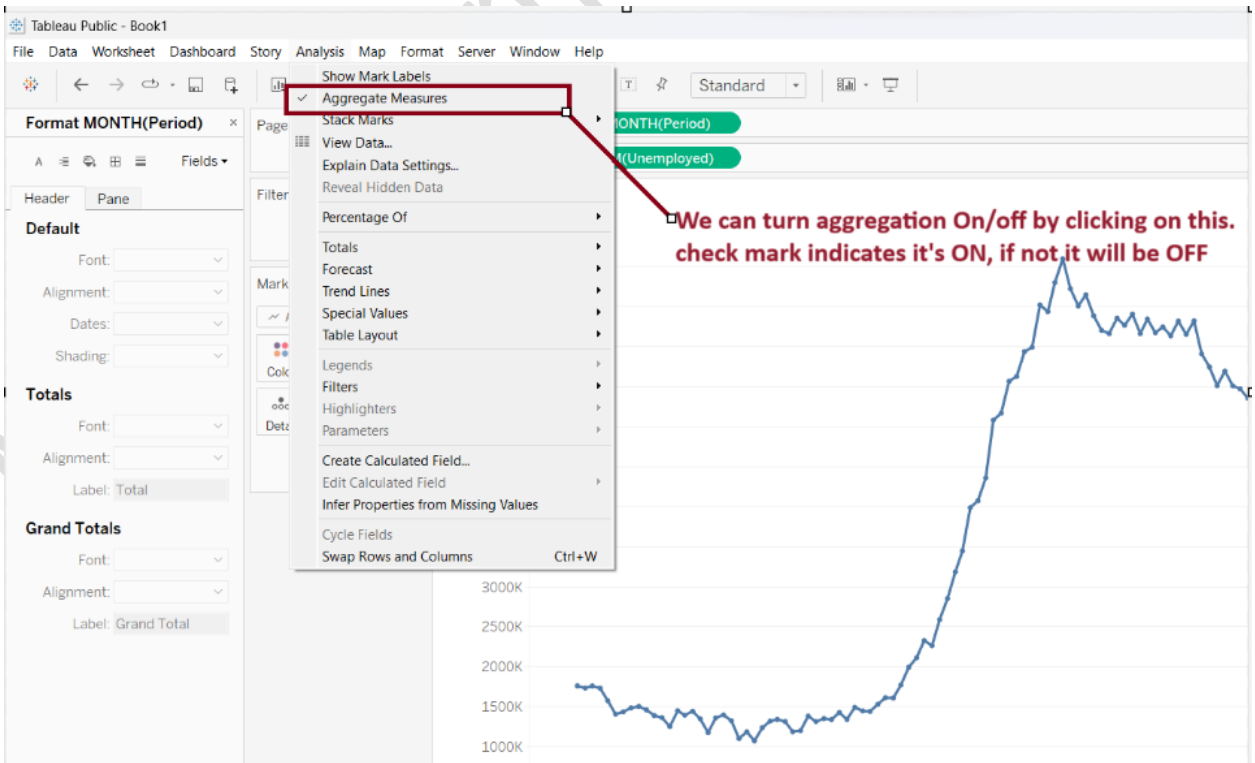
- When column is a measure → it shows the data for every month across all the years. Below is the graph for reference.

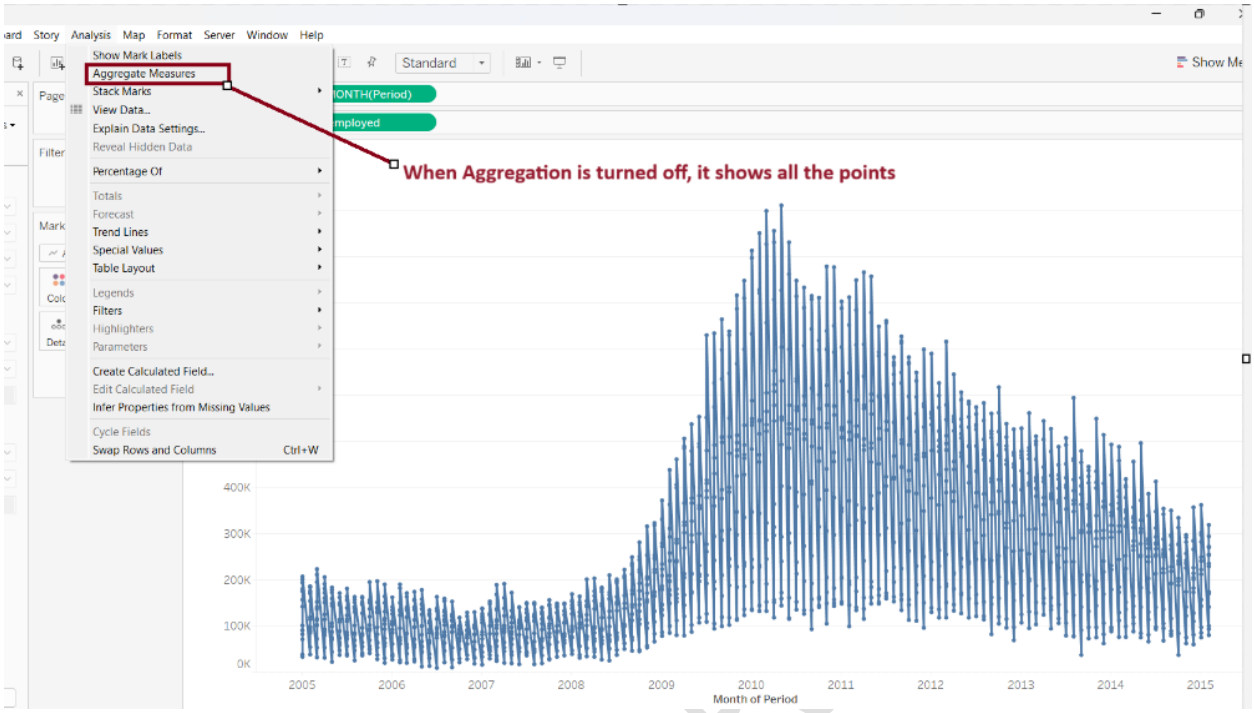


Here the data is continuous and is shown for every month across all the years, as Month in column is measure (Green color)

Aggregation

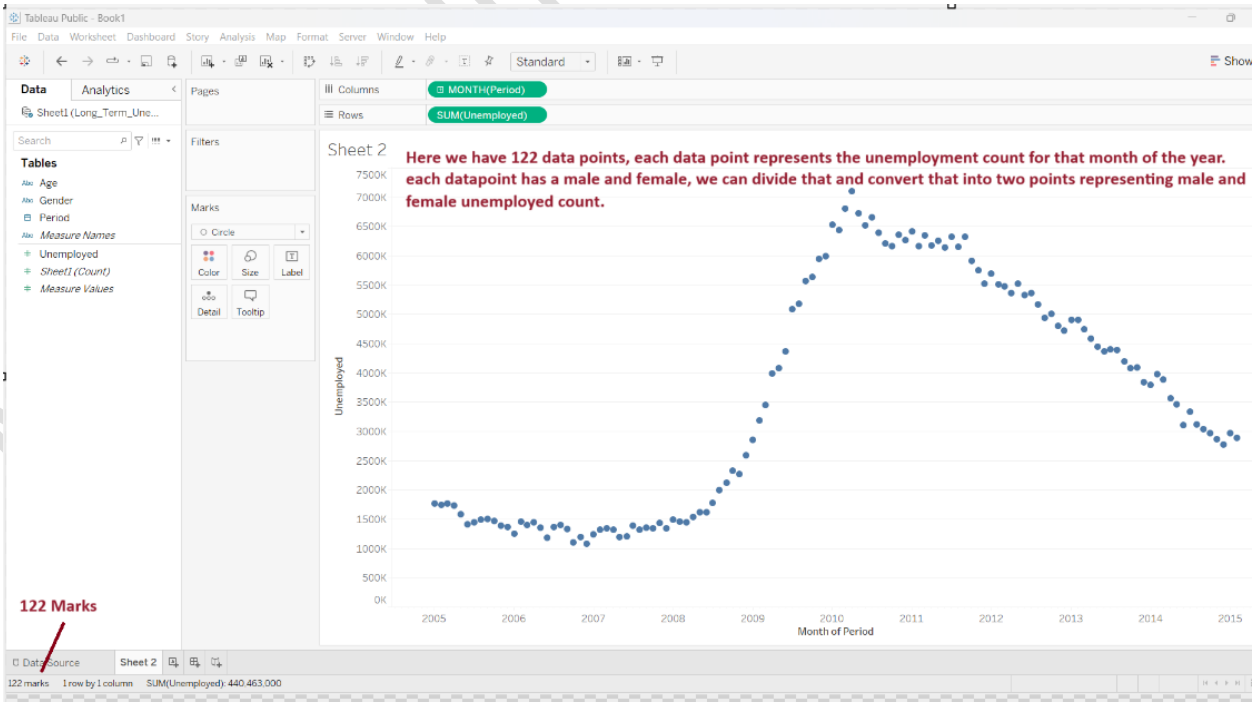
- Tableau automatically understands and aggregates data based on the rows and columns that we select.
  in the Above diagram, in the rows, SUM is applied to the unemployed.
- We can turn Aggregation ON/OFF by selecting Analysis → aggregate measures.

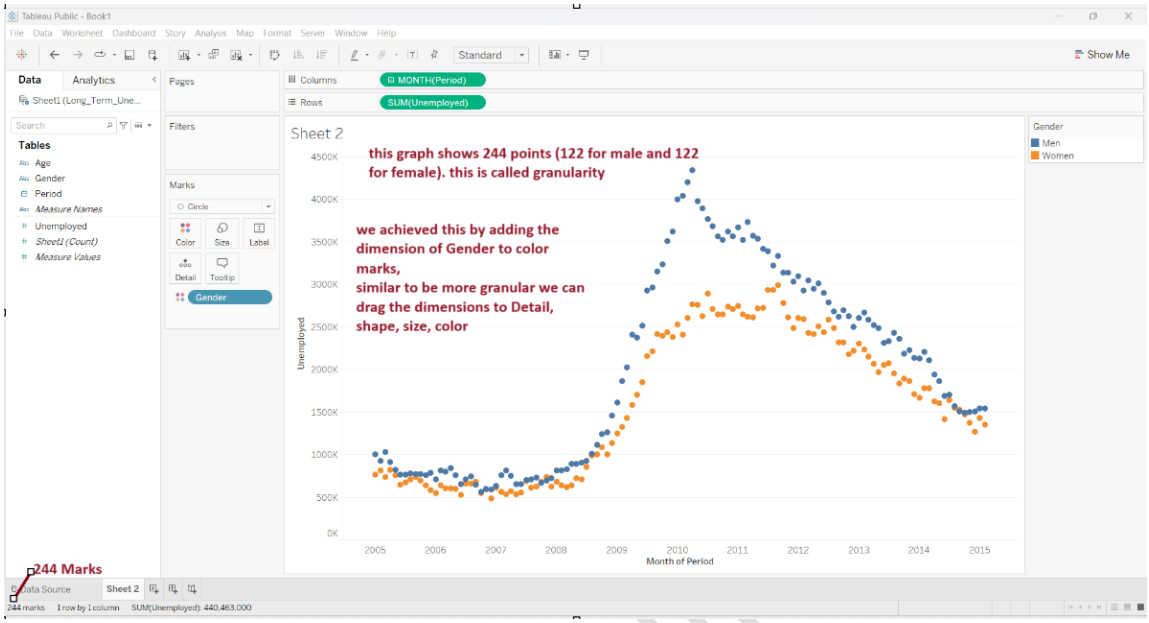




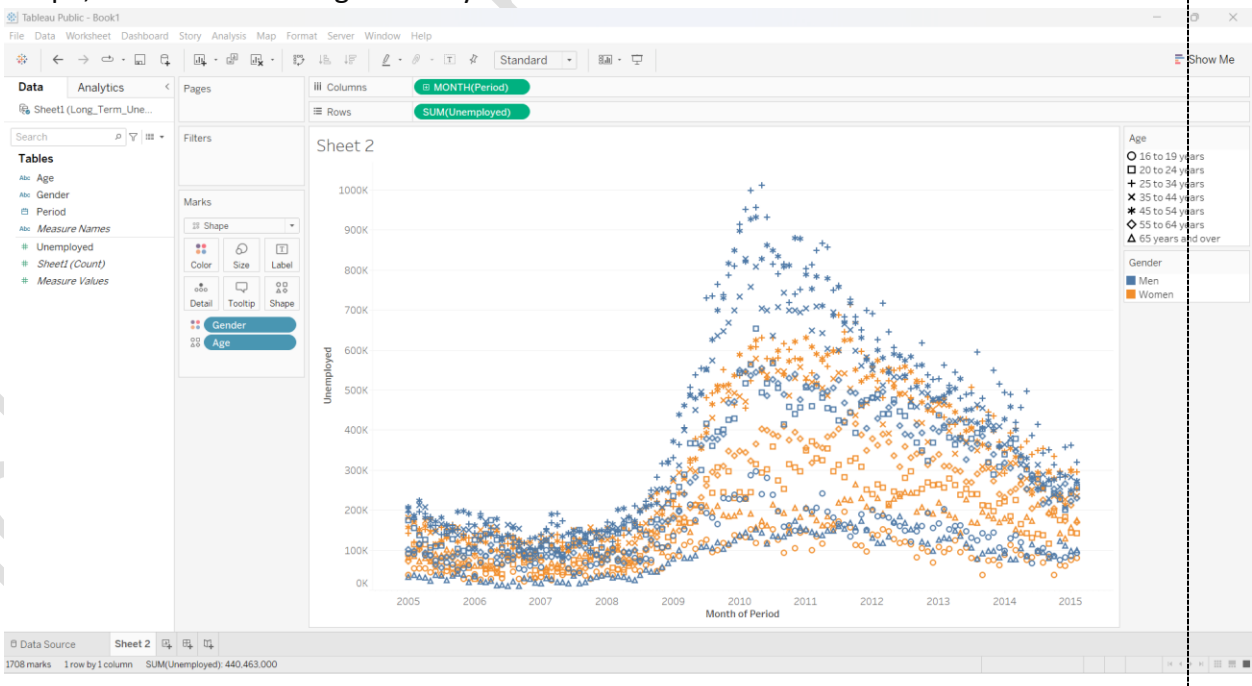
Granularity:

- "granularity" refers to the level of detail or the degree of aggregation in your data.

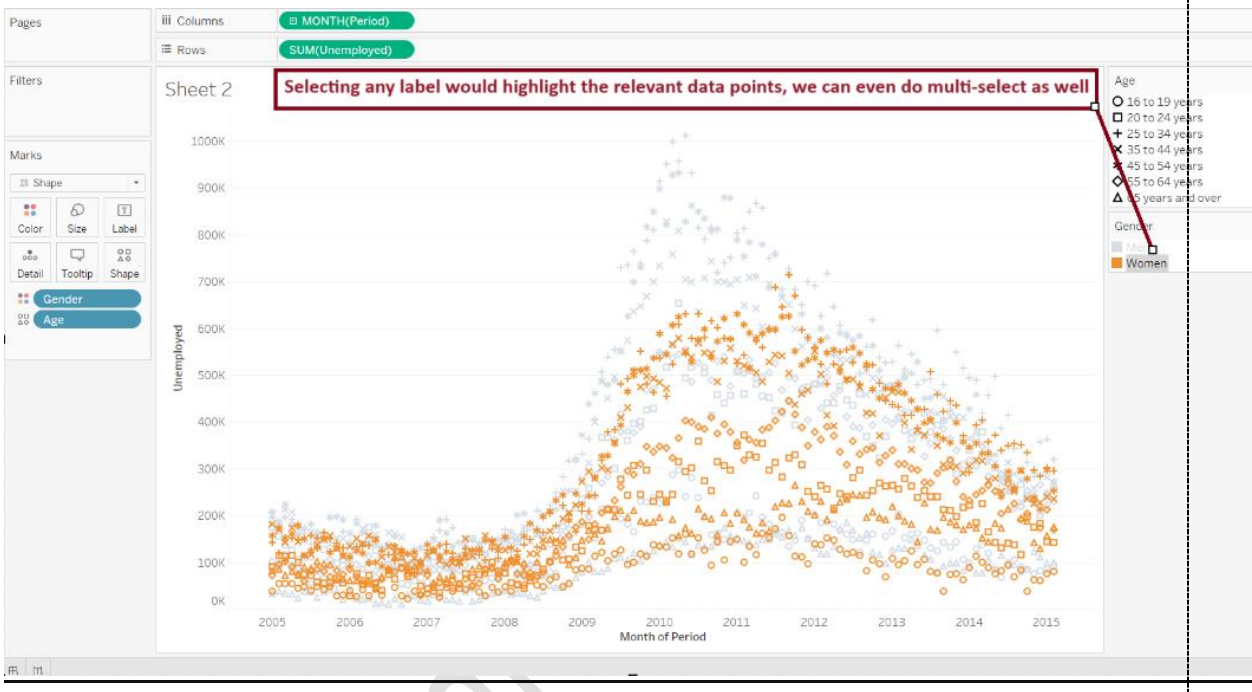




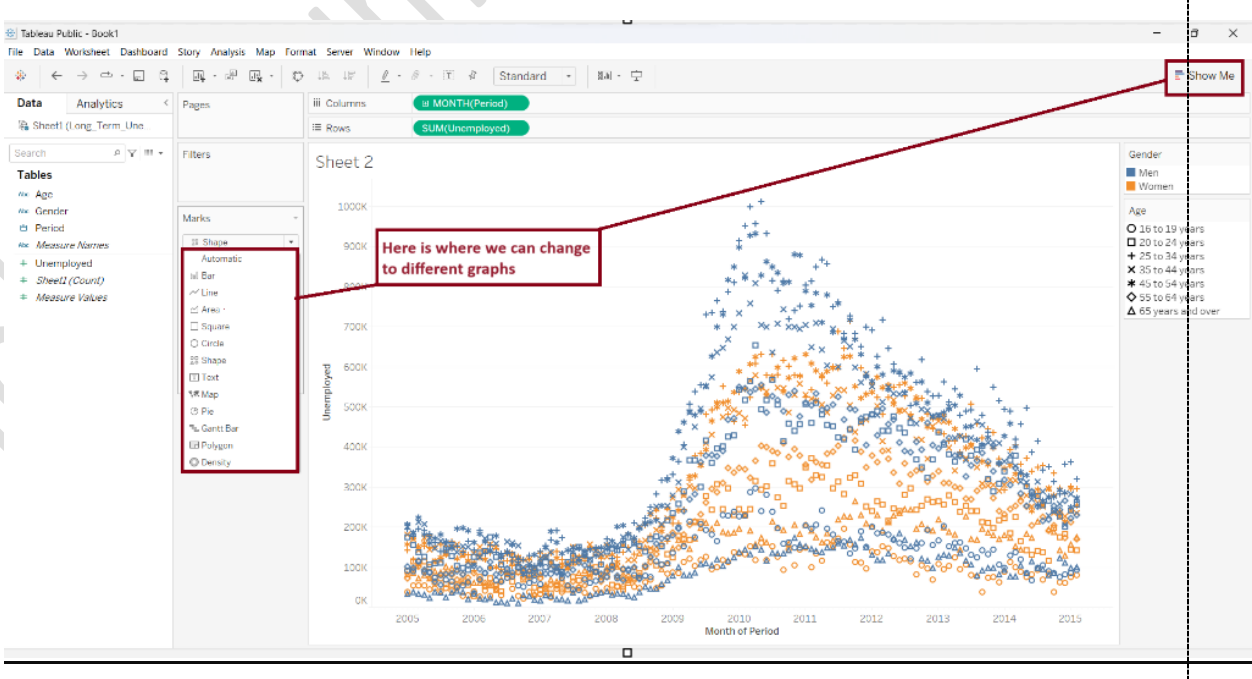
Shape, color are used for granularity.



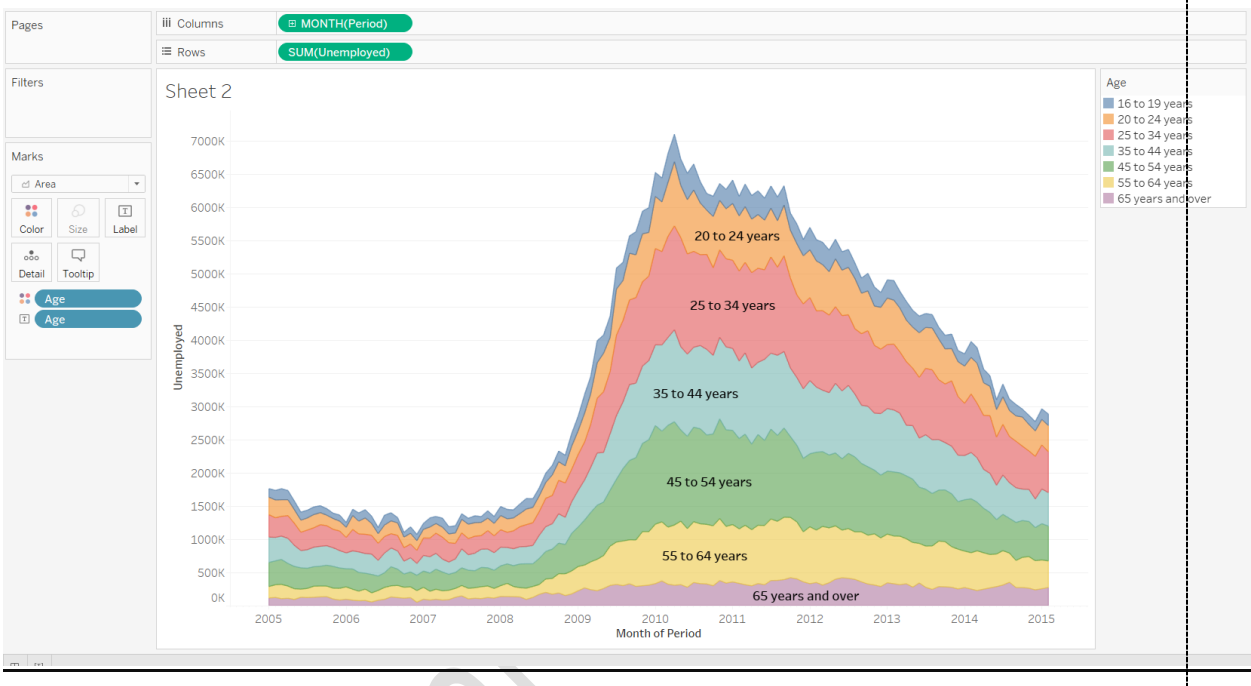
Highlighting the Datapoints:



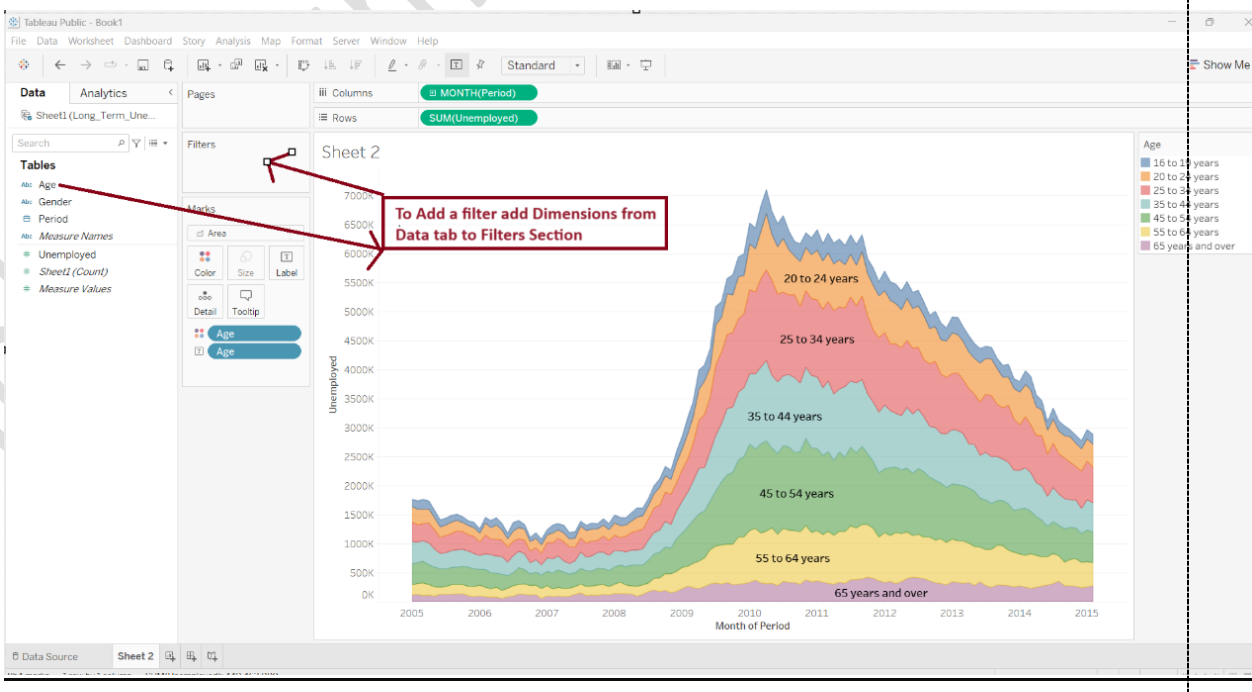
Area Chart:

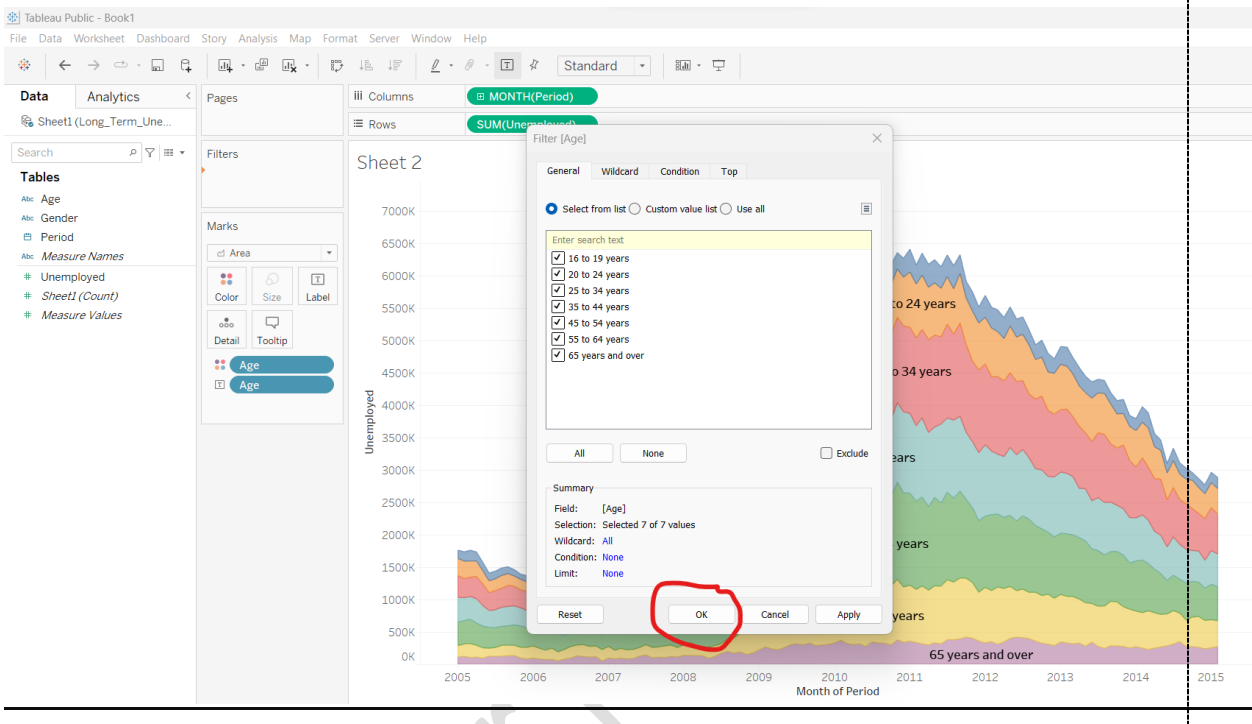


Below is how the area graph looks like.

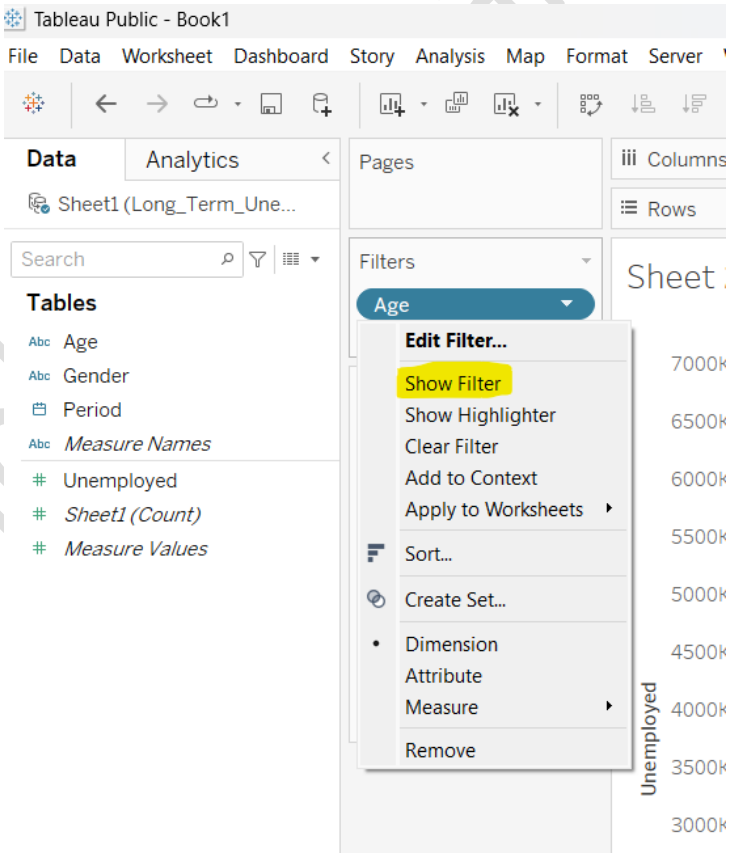


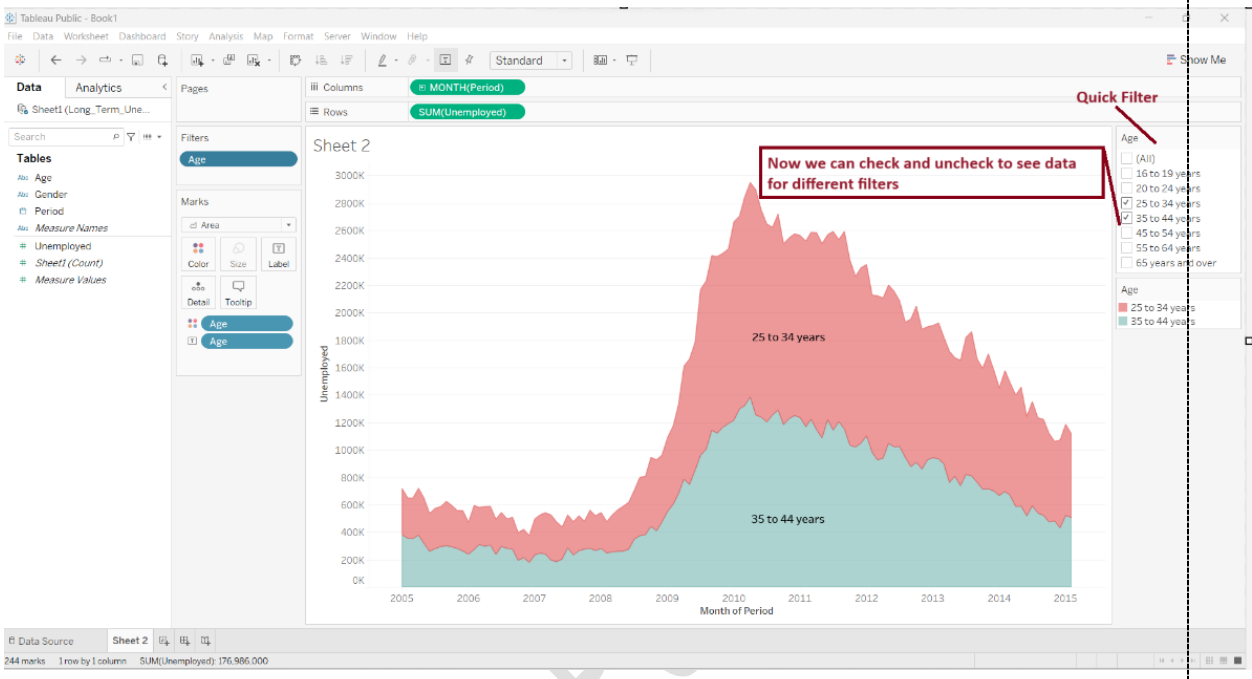
Adding a Filter and Quick Filter:



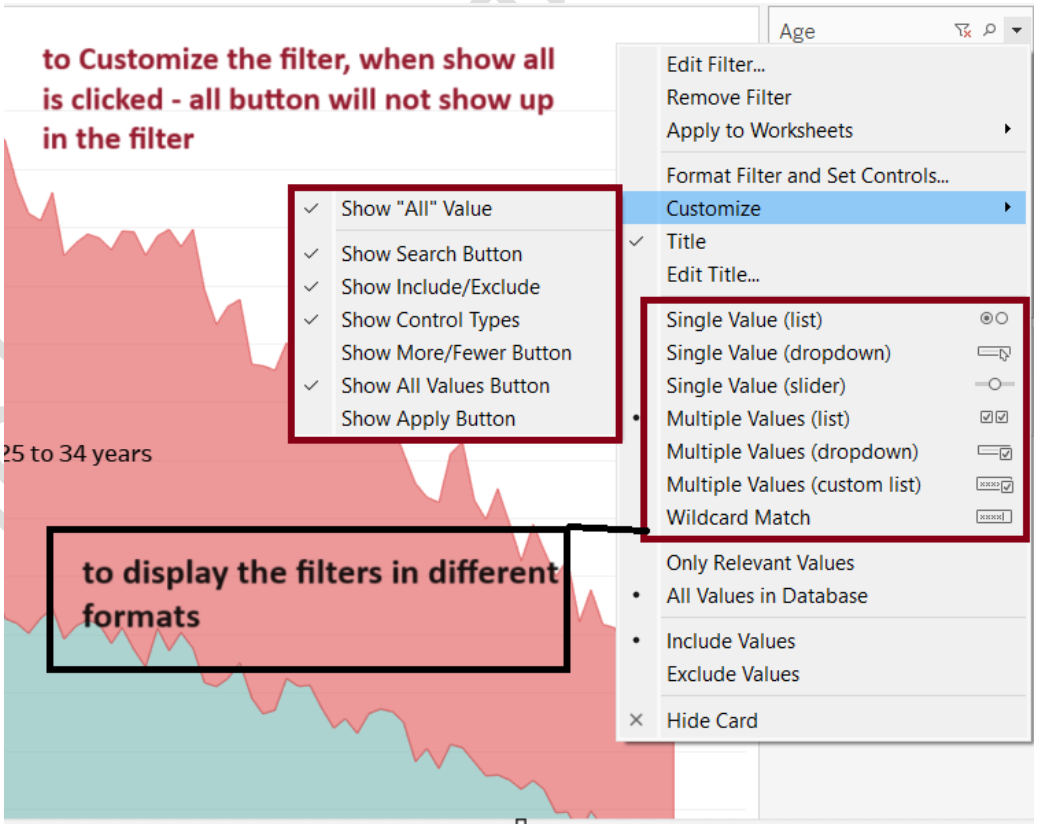


To add a Quick Filter, click on show filter.





More options in Quick Filter.

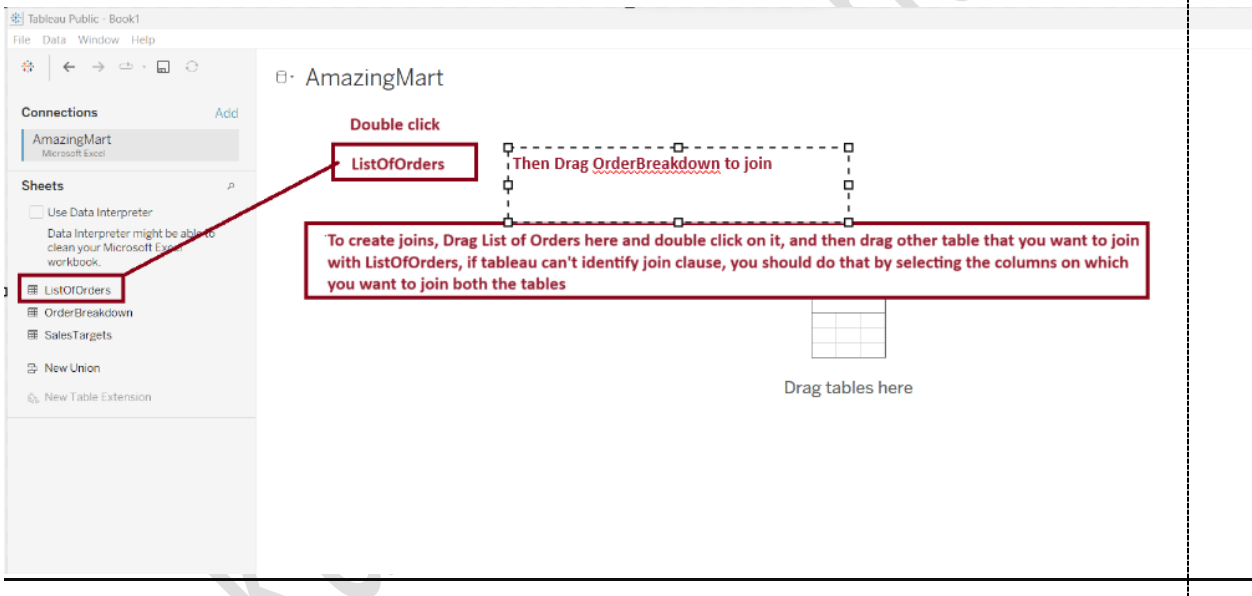


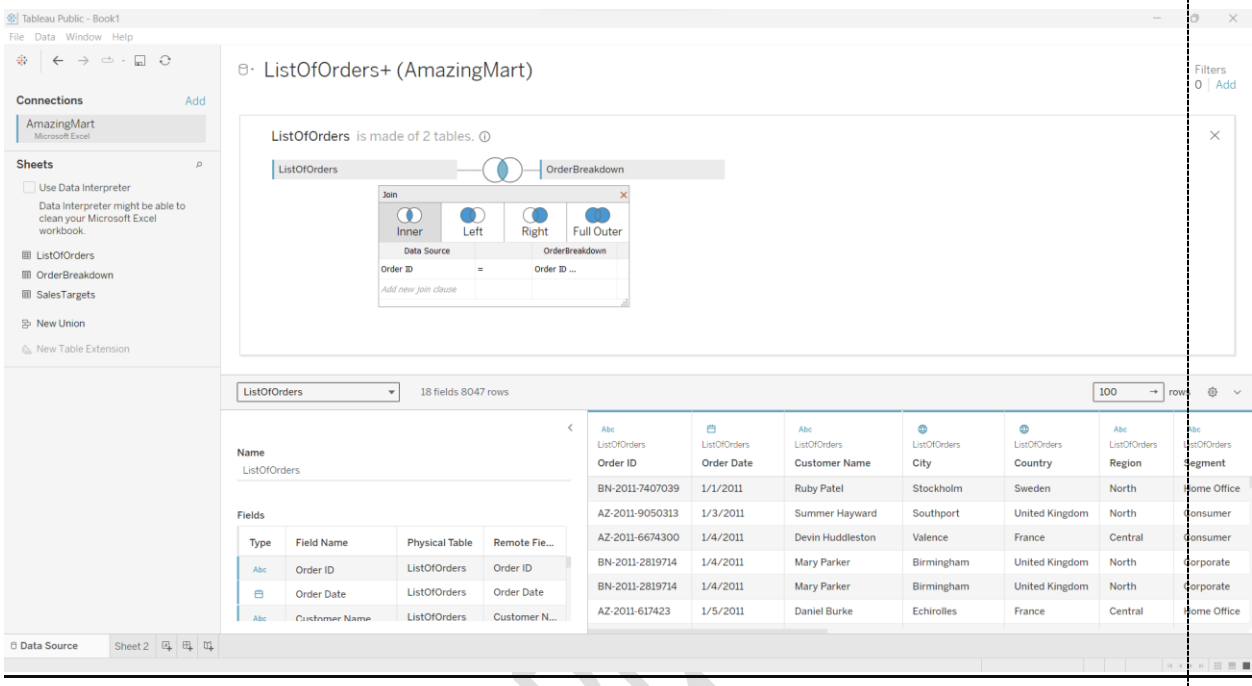
Maps, Scatterplots and Dashboard

Download Dataset here:

[Tableau_Data_Sets/AmazingMart.xlsx at main · MohanPulibanti/Tableau_Data_Sets (github.com)]

Creating Joins in Tableau:

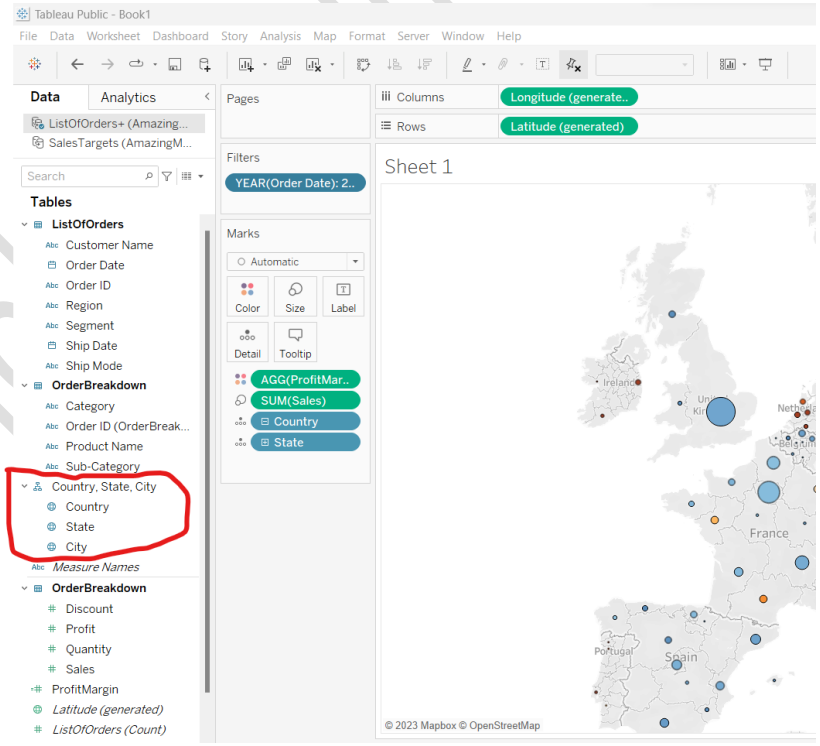




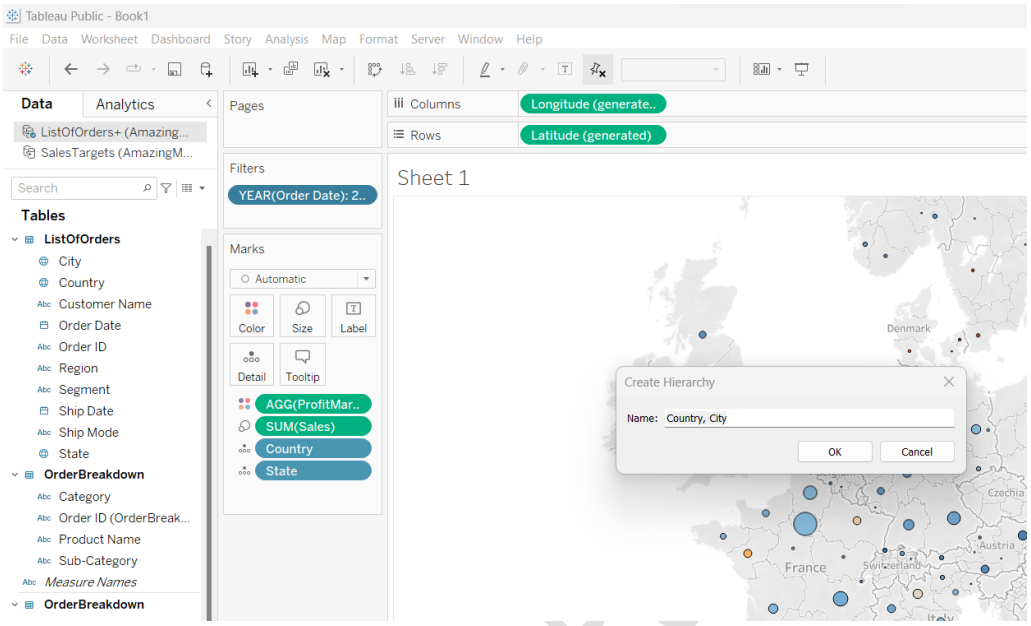
Creating a Map, Working with hierarchies

Creating Hierarchies:

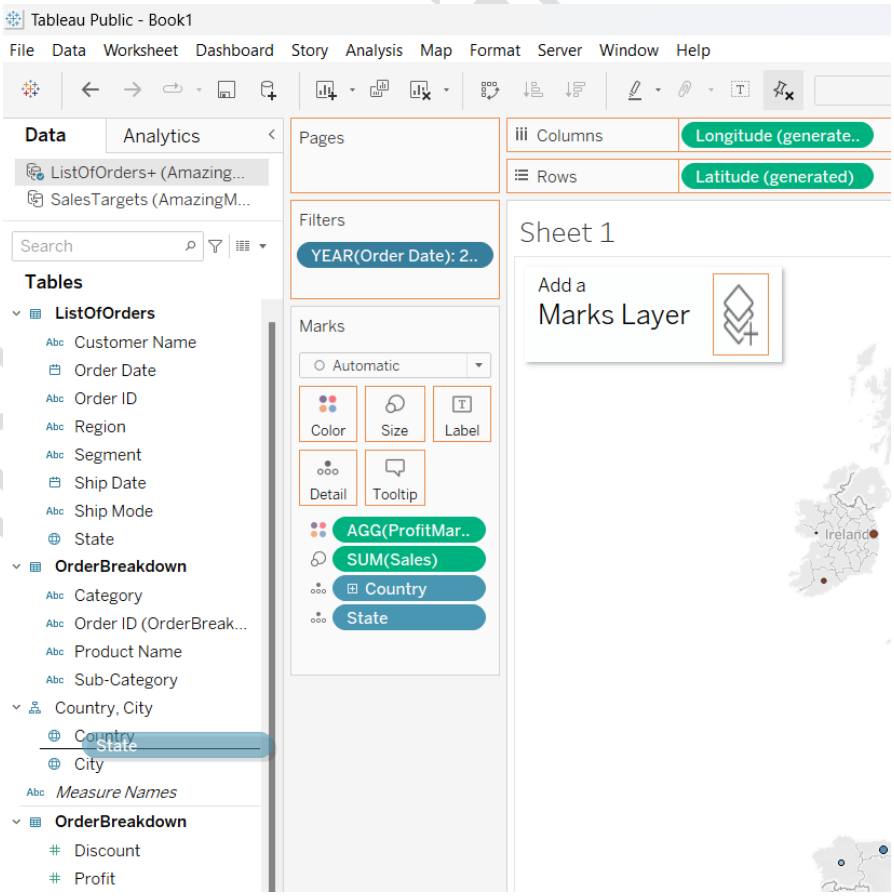
- By Default, Tableau shows hierarchies, we can use them or we can create hierarchies on our own.

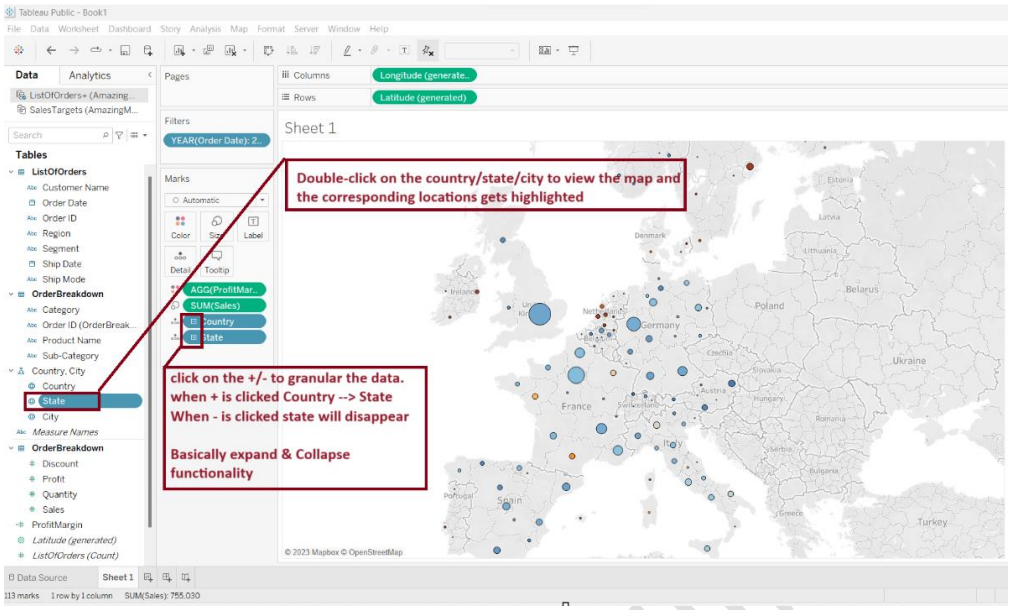


- If we want to create our hierarchy, Drag the dimension onto another dimension with which you want to create hierarchy.



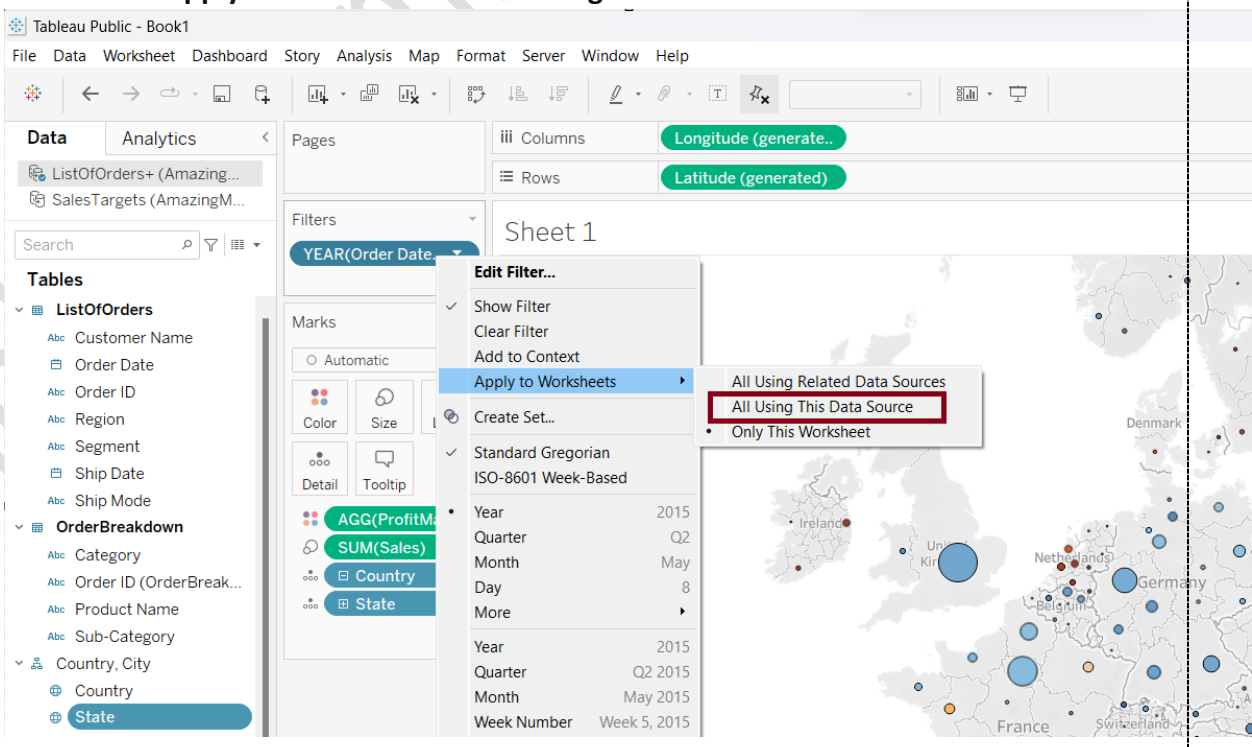
- Once we click on ok, a hierarchy gets created for country and city. Now, the state to be dragged to this hierarchy and it should come between country and city, so drag state to middle of these two in the hierarchy.



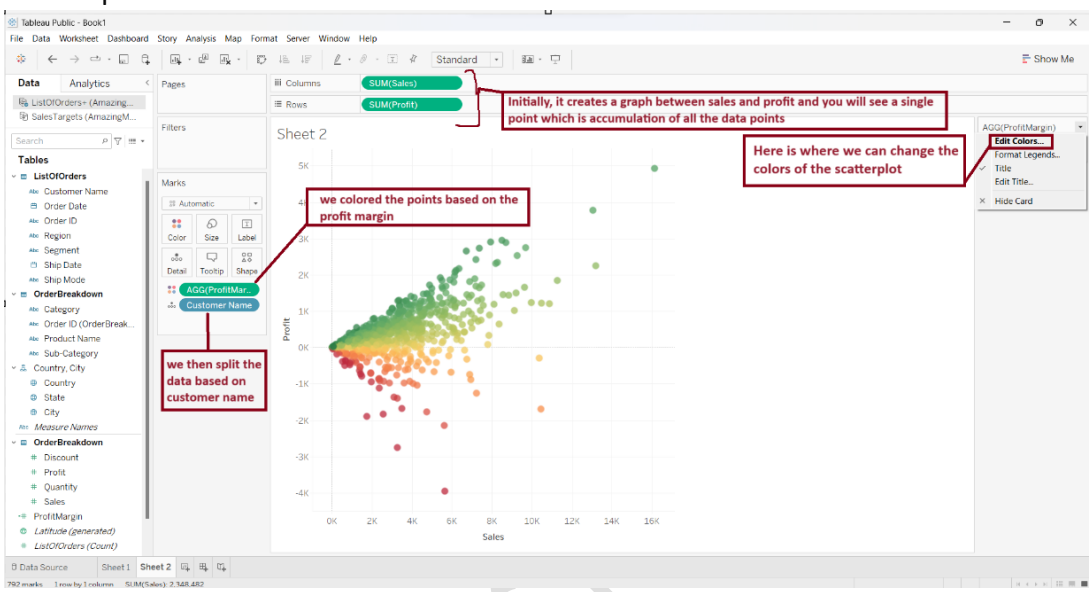


Creating a scatter plot, Applying Filters to Multiple worksheets

- Applying Filters to multiple sheets: apply filter to one sheet and then we can add the same filter to multiple worksheets using the same data source by clicking on the **filter → Apply to Worksheets → All Using This Data Source**

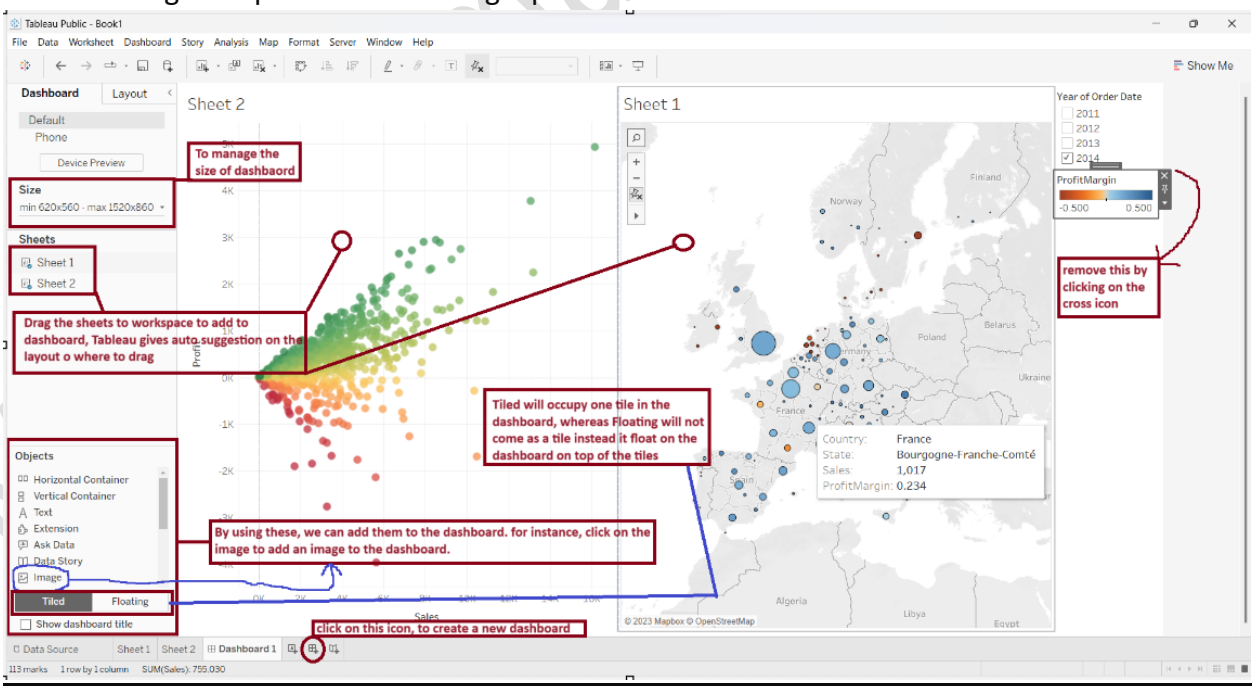


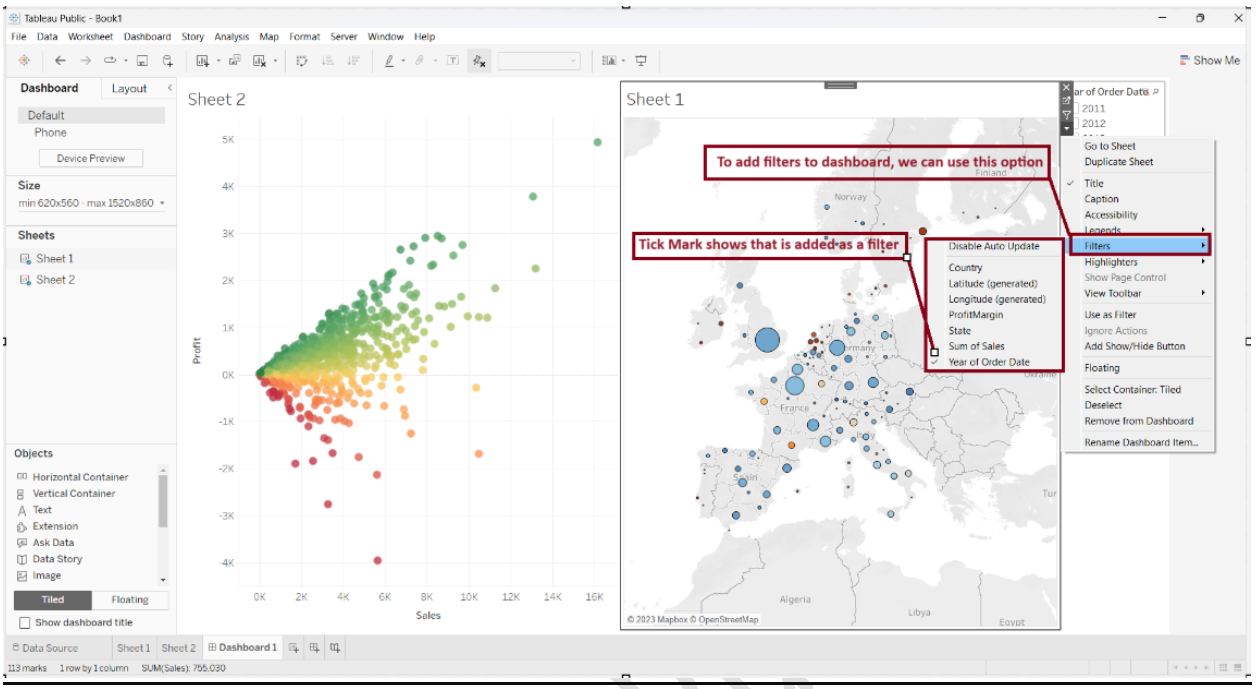
- When we create a new sheet, we can't see the filter right away, as it doesn't know whether we will be using the same data source or not. The moment we drag any of the fields to the worksheet, the filter will be added to the sheet.
- Scatter plot



Creating our First Dashboard

- Consolidating multiple sheets at a single place is Dashboard.

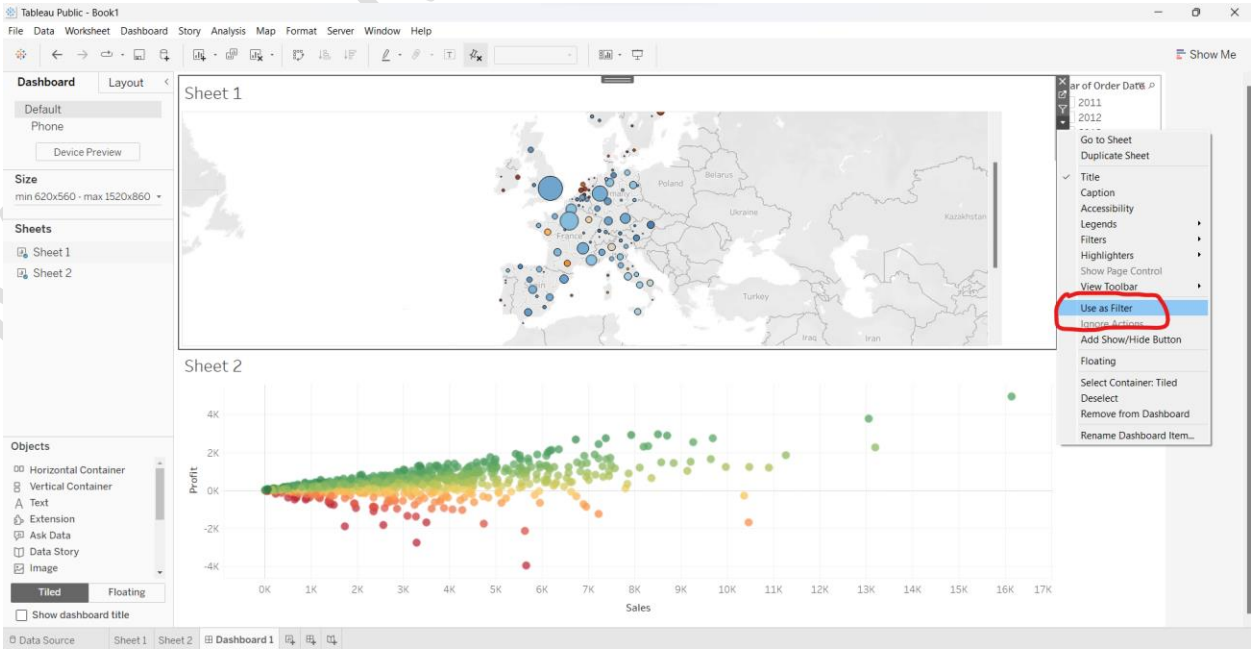


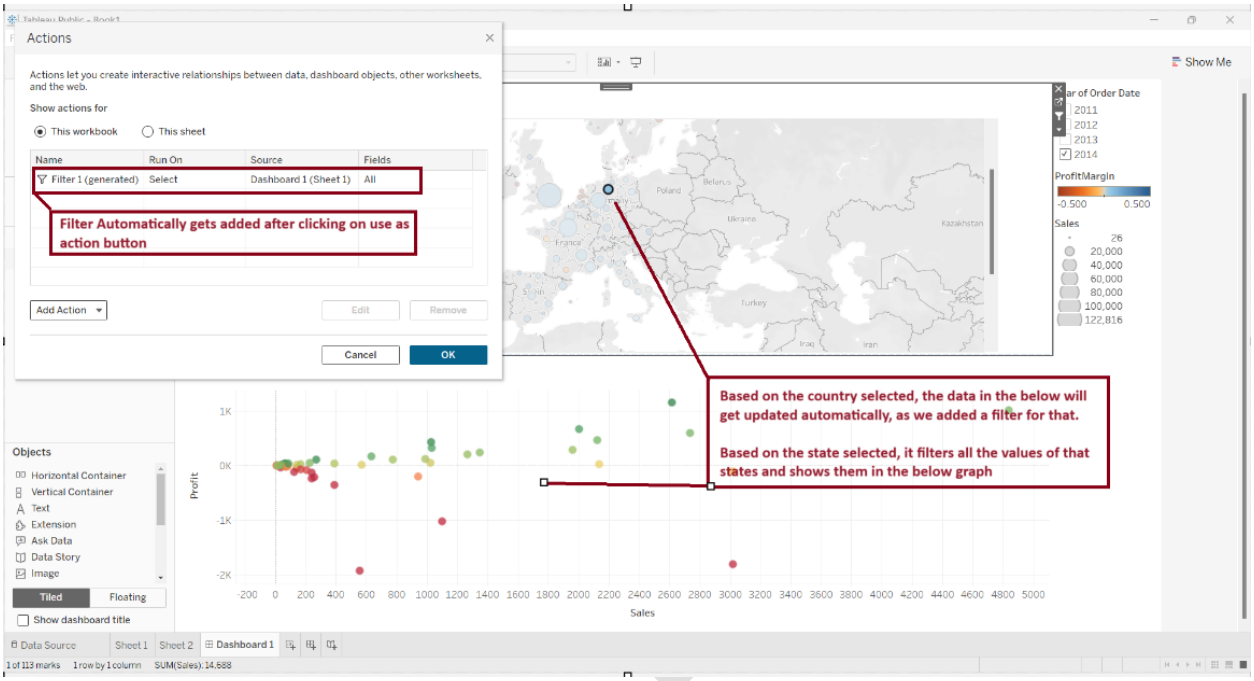


Adding an Interactive Action: Filter

- Filter can be added in two ways.

First way:

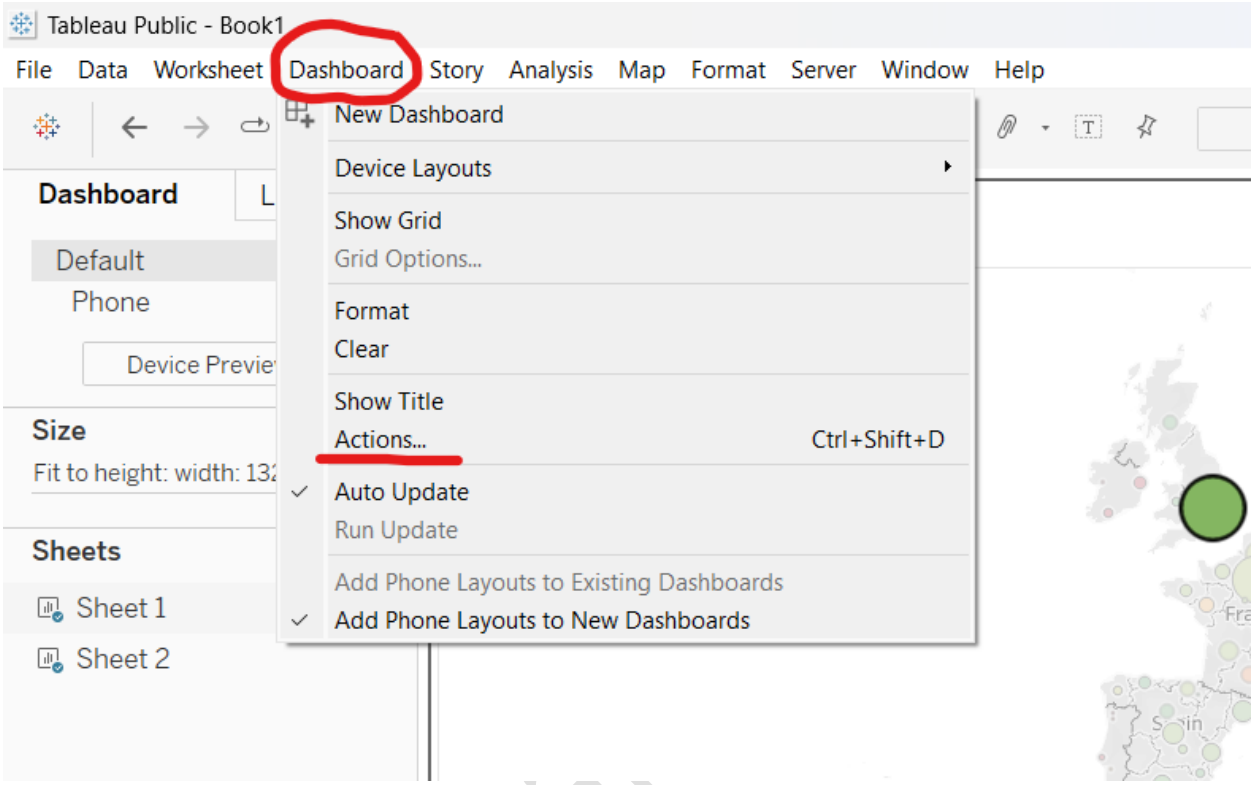


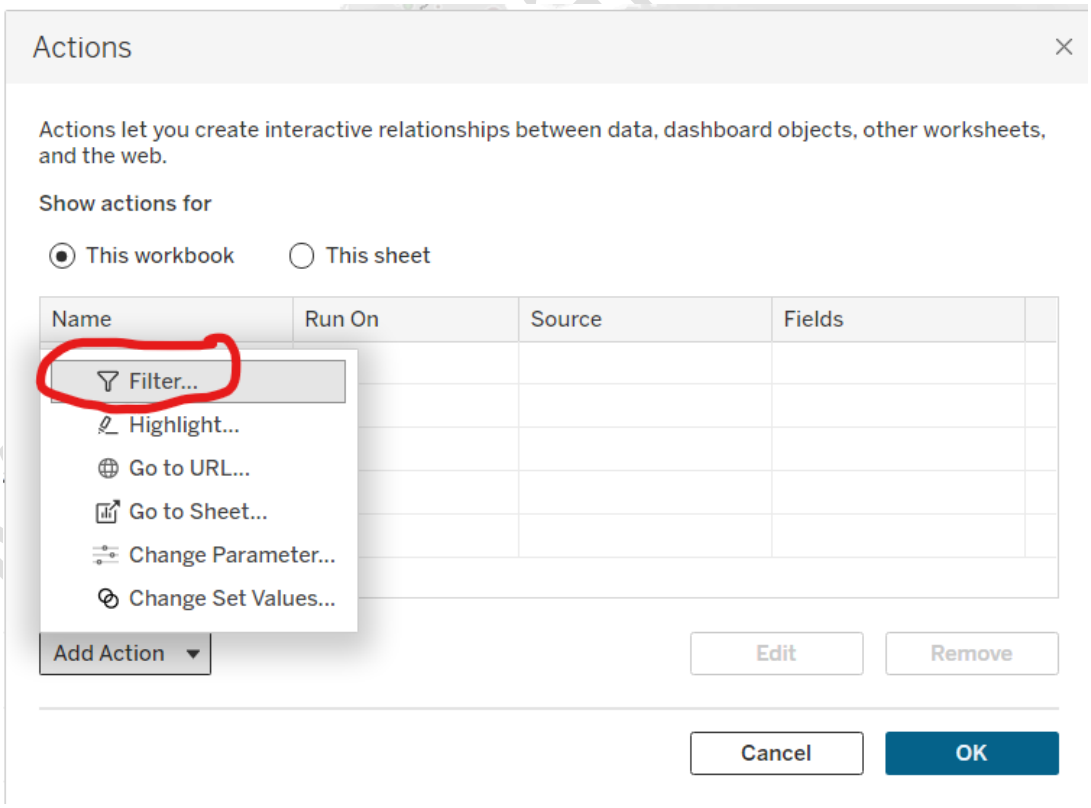


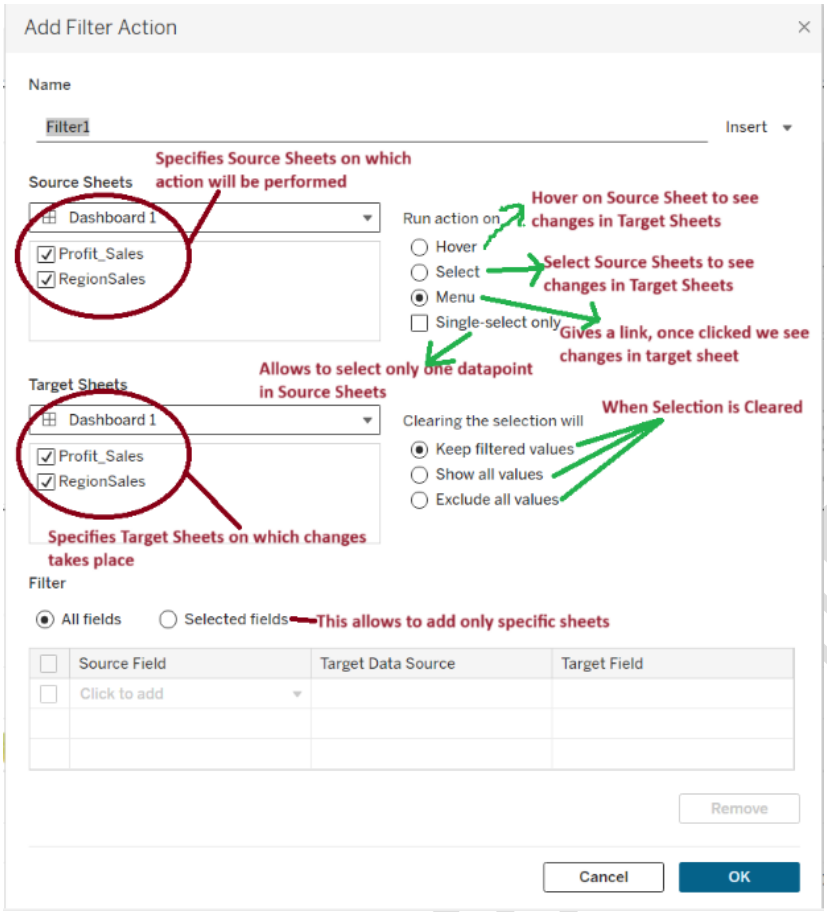
Second way of adding a filter:

Mohan Kumar P

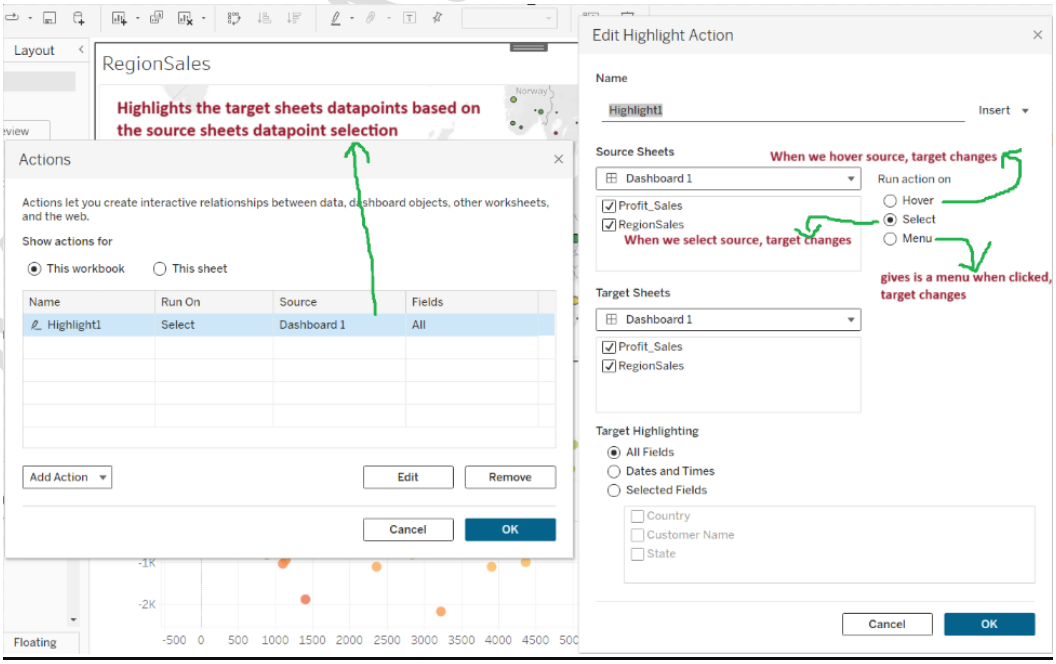
- Go to File Menu → Dashboard → Actions



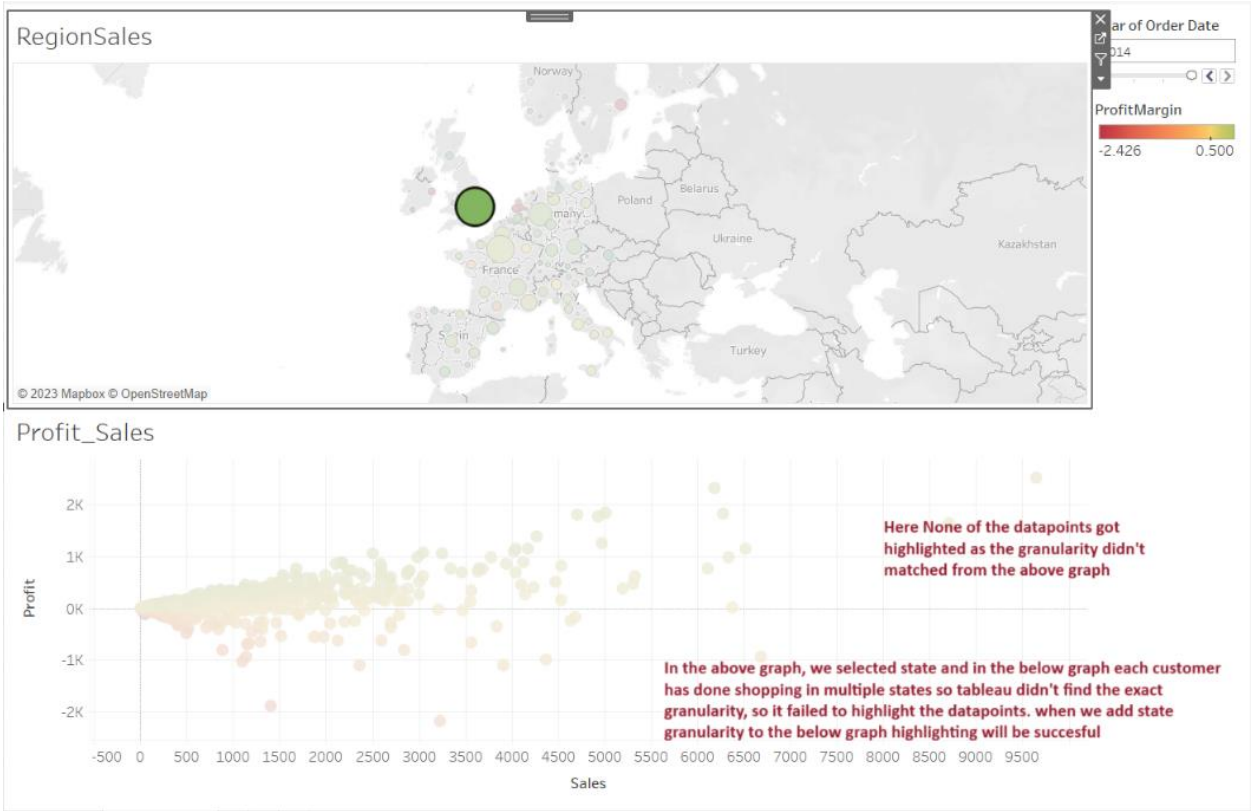




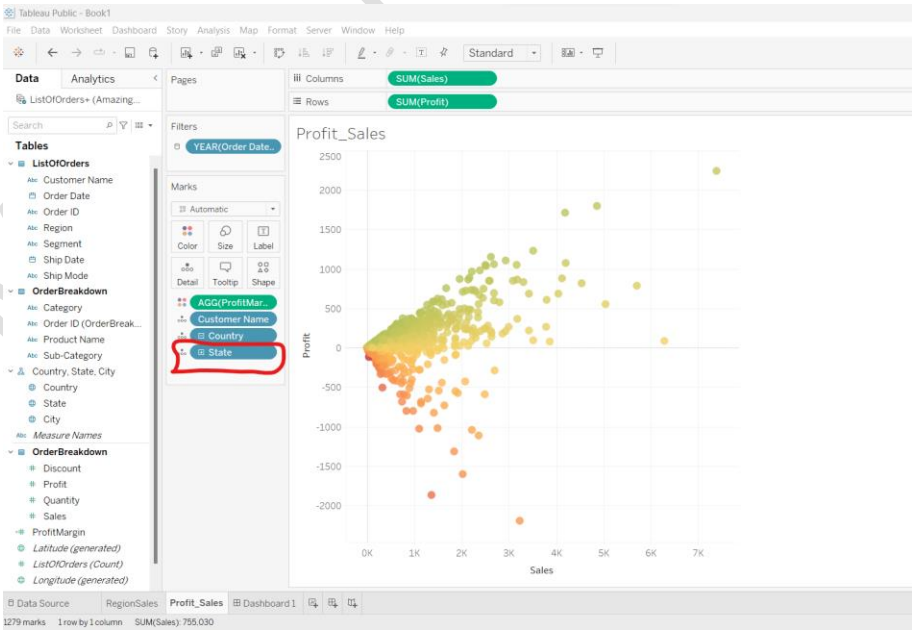
Add an interactive Action – Highlighting:



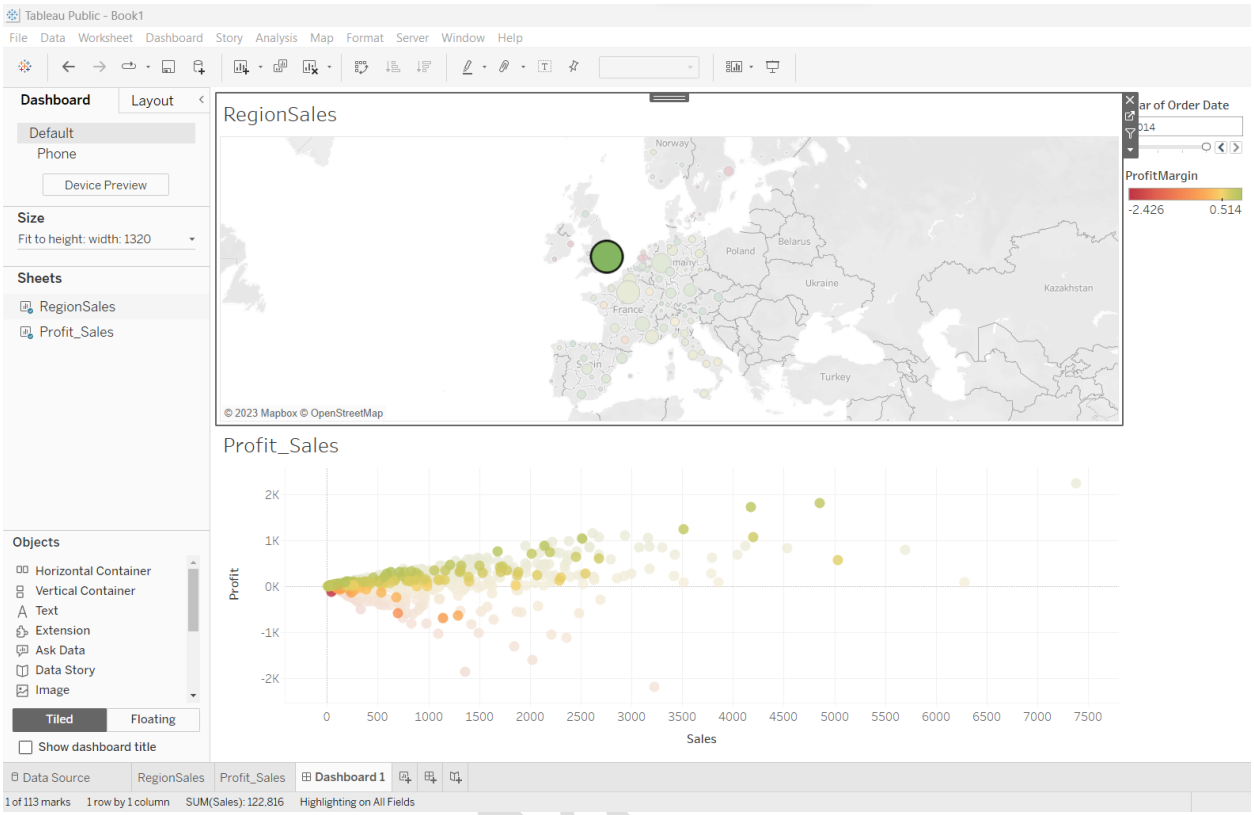
But when the granularity is not rightly selected, the points on the target will not get highlighted. As Highlight will shows up all the data and highlights some points in that whereas in Filter, it filters some points from that dataset and shows them up.



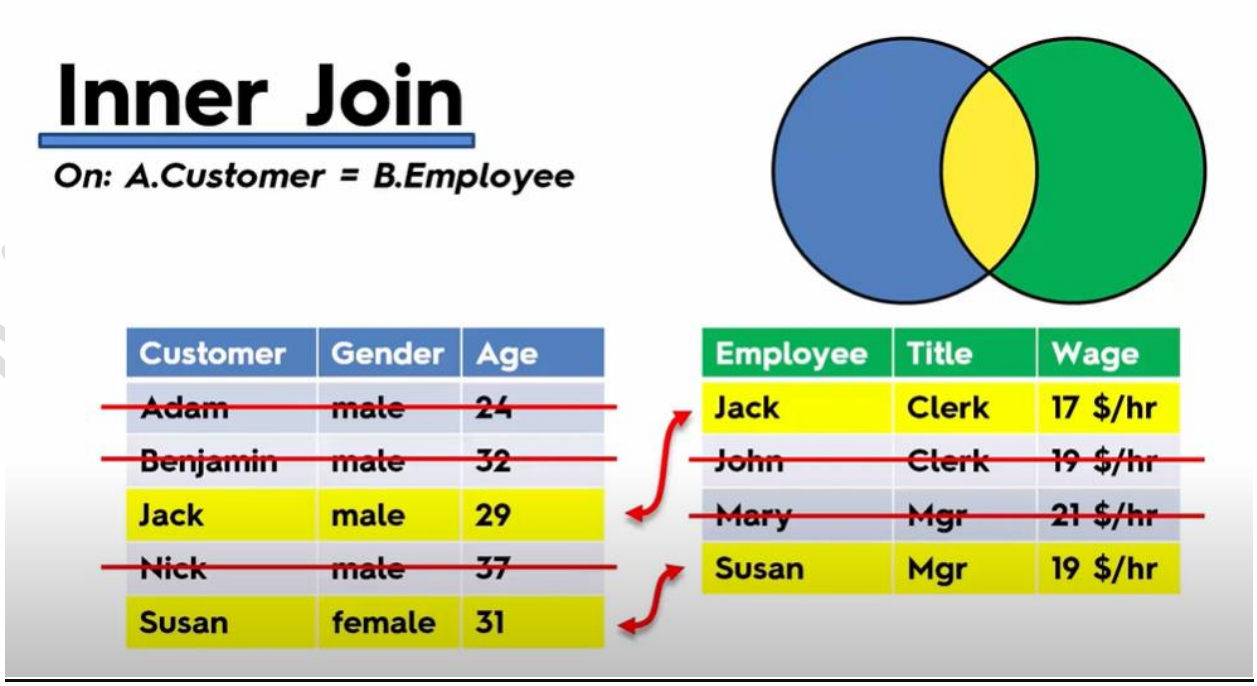
Adding State to Detail in the Profit_sales



Now, the highlight will be successful.

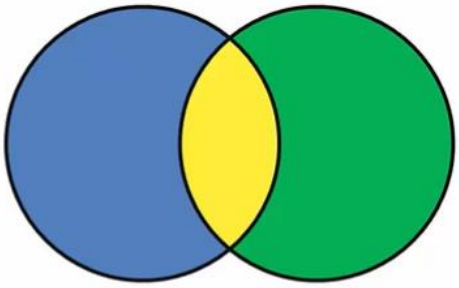


Joins



Inner Join

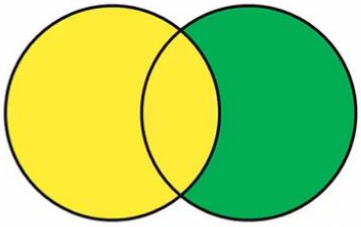
On: A.Customer = B.Employee



Customer	Gender	Age	Employee	Title	Wage
Jack	male	29	Jack	Clerk	17 $/hr
Susan	female	31	Susan	Mgr	19 $/hr

Left Outer Join

On: A.Customer = B.Employee

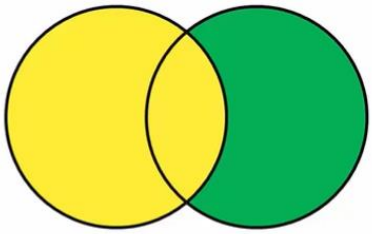


Customer	Gender	Age
Adam	male	24
Benjamin	male	32
Jack	male	29
Nick	male	37
Susan	female	31

Employee	Title	Wage
Jack	Clerk	17 $/hr
~~John~~	~~Clerk~~	~~19 $/hr~~
~~Mary~~	~~Mgr~~	~~21 $/hr~~
Susan	Mgr	19 $/hr

Left Outer Join

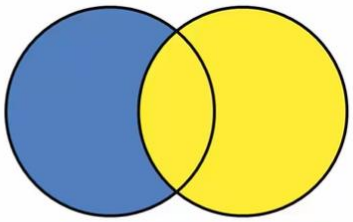
On: A.Customer = B.Employee



Customer	Gender	Age	Employee	Title	Wage
Adam	male	24			
Benjamin	male	32			
Jack	male	29	Jack	Clerk	17 $/hr
Nick	male	37			
Susan	female	31	Susan	Mgr	19 $/hr

Right Outer Join

On: A.Customer = B.Employee

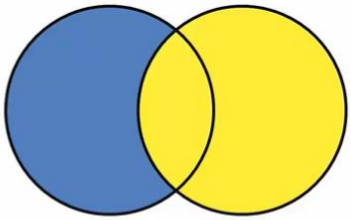


Customer	Gender	Age
~~Adam~~	~~male~~	~~24~~
~~Benjamin~~	~~male~~	~~32~~
Jack	male	29
~~Nick~~	~~male~~	~~37~~
Susan	female	31

Employee	Title	Wage
Jack	Clerk	17 $/hr
John	Clerk	19 $/hr
Mary	Mgr	21 $/hr
Susan	Mgr	19 $/hr

Right Outer Join

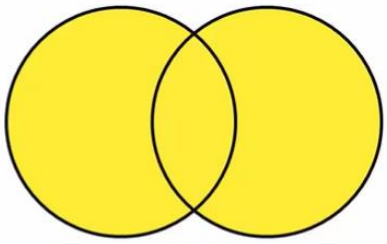
On: A.Customer = B.Employee



Customer	Gender	Age	Employee	Title	Wage
Jack	male	29	Jack	Clerk	17 $/hr
			John	Clerk	19 $/hr
			Mary	Mgr	21 $/hr
Susan	female	31	Susan	Mgr	19 $/hr

Full Outer Join

On: A.Customer = B.Employee

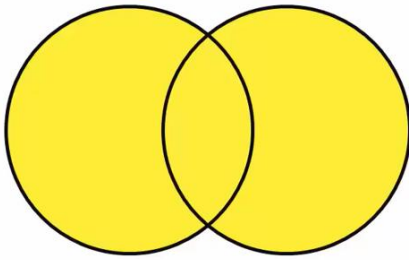


Customer	Gender	Age
Adam	male	24
Benjamin	male	32
Jack	male	29
Nick	male	37
Susan	female	31

Employee	Title	Wage
Jack	Clerk	17 $/hr
John	Clerk	19 $/hr
Mary	Mgr	21 $/hr
Susan	Mgr	19 $/hr

Full Outer Join

On: A.Customer = B.Employee

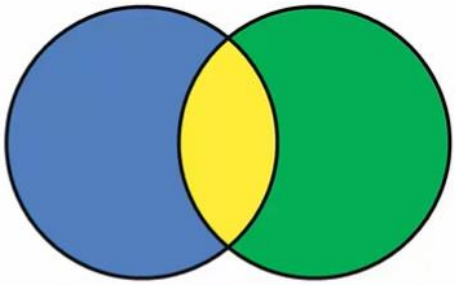


Customer	Gender	Age	Employee	Title	Wage
Adam	male	24			
Benjamin	male	32			
Jack	male	29	Jack	Clerk	17 $/hr
Nick	male	37			
Susan	female	31	Susan	Mgr	19 $/hr
			John	Clerk	19 $/hr
			Mary	Mgr	21 $/hr

Joins with Duplicate Values:

Inner Join

On: A.OrderNum = B.OrderNum

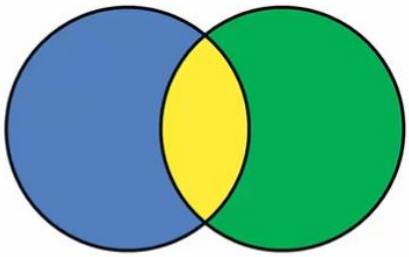


Order #	Region	Status
001	North	Unpaid
002	North	Unpaid
003	North	Paid
~~004~~	~~North~~	~~Paid~~

Order #	Item	Sales
001	Chair	$97
001	Desk	$123
002	Stapler	$8
003	Pen	$3
003	Pencil	$1
003	Eraser	$1

Inner Join

On: A.OrderNum = B.OrderNum

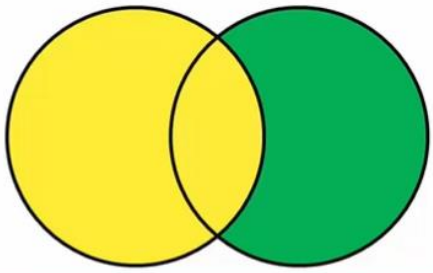


	Order #	Region	Status	Order #	Item	Sales
Duplicated rows (Table A)	001	North	Unpaid	001	Chair	$97
	001	North	Unpaid	001	Desk	$123
	002	North	Unpaid	002	Stapler	$8
Duplicated rows (Table A)	003	North	Paid	003	Pen	$3
	003	North	Paid	003	Pencil	$1
	003	North	Paid	003	Eraser	$1

Joining on Multiple fields

Left Join

On: A.OrderNum = B.OrderNum

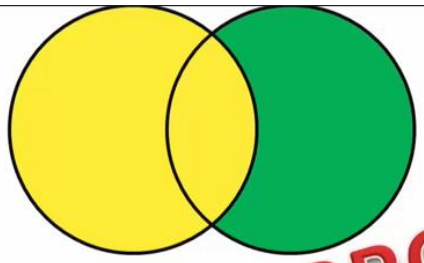


Store	Order #	Customer
North	001	Mike
North	002	Jack
South	001	Susan

Store	Order #	Item	Sales
North	001	Laptop	$999
North	001	Mouse	$49
North	002	Monitor	$155
South	001	Camera	$97

Left Join

On: A.OrderNum = B.OrderNum

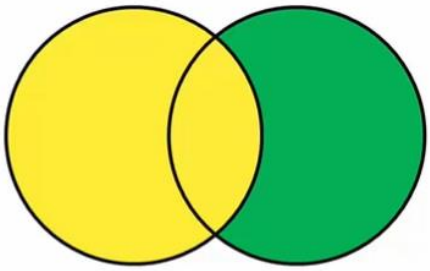


ERROR!

Store	Order #	Customer	Store	Order #	Item	Sales
North	001	Mike	North	001	Laptop	$999
North	001	Mike	North	001	Mouse	$49
North	001	Mike	South	001	Camera	$97
North	002	Jack	North	002	Monitor	$155
South	001	Susan	North	001	Laptop	$999
South	001	Susan	North	001	Mouse	$49
South	001	Susan	South	001	Camera	$97

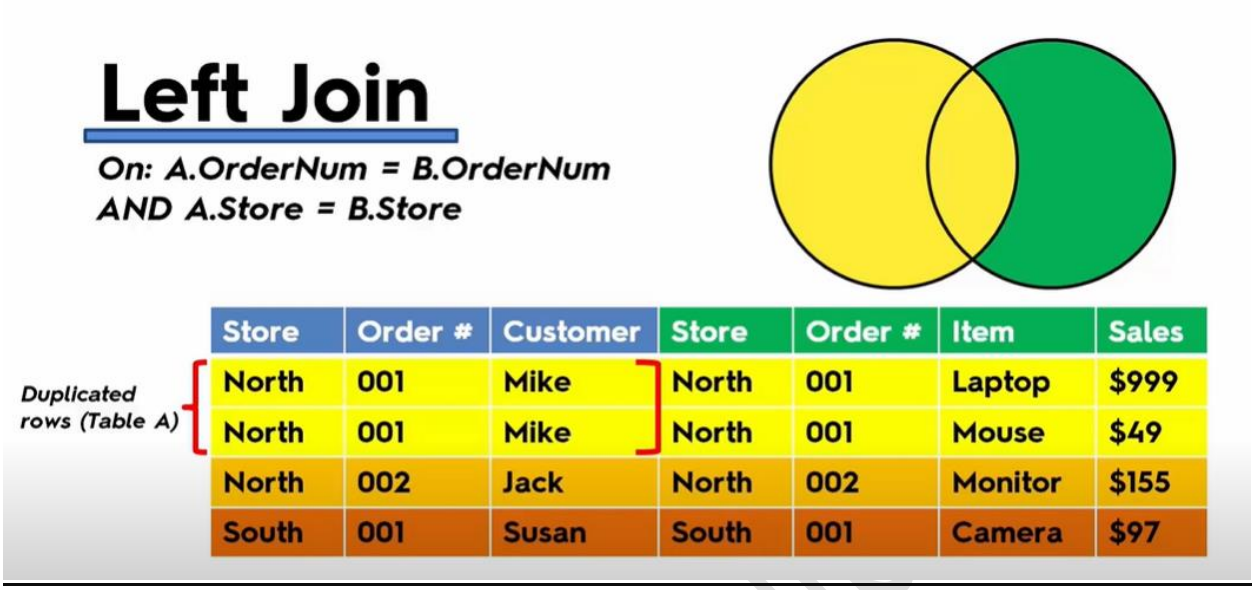
Left Join

On: A.OrderNum = B.OrderNum
AND A.Store = B.Store



Store	Order #	Customer
North	001	Mike
North	002	Jack
South	001	Susan

Store	Order #	Item	Sales
North	001	Laptop	$999
North	001	Mouse	$49
North	002	Monitor	$155
South	001	Camera	$97



Joining Data vs Blending Data in Tableau

In Tableau, "blending" and "joining" are two different concepts related to combining data from multiple tables or data sources.

Joining:

Joining is the process of combining rows from two or more tables based on a related column between them.

In Tableau, you can join tables when you have a common field (or fields) that can be used to link the data between them. For example, if you have a "Customer ID" column in both tables, you can join the tables based on this common identifier.

Tableau supports different types of joins, such as inner join, left join, right join, and outer join, which determine how the data is combined based on the matching criteria.

Blending:

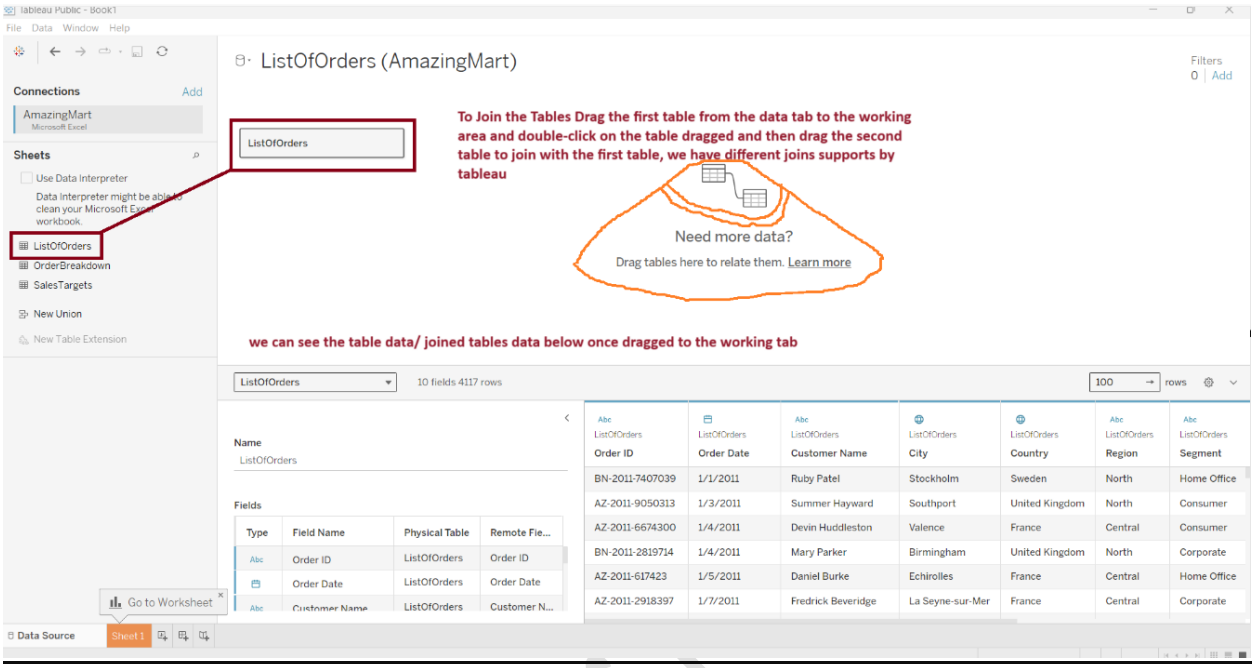
Blending is used when you want to combine data from multiple data sources in a single view on a dashboard.

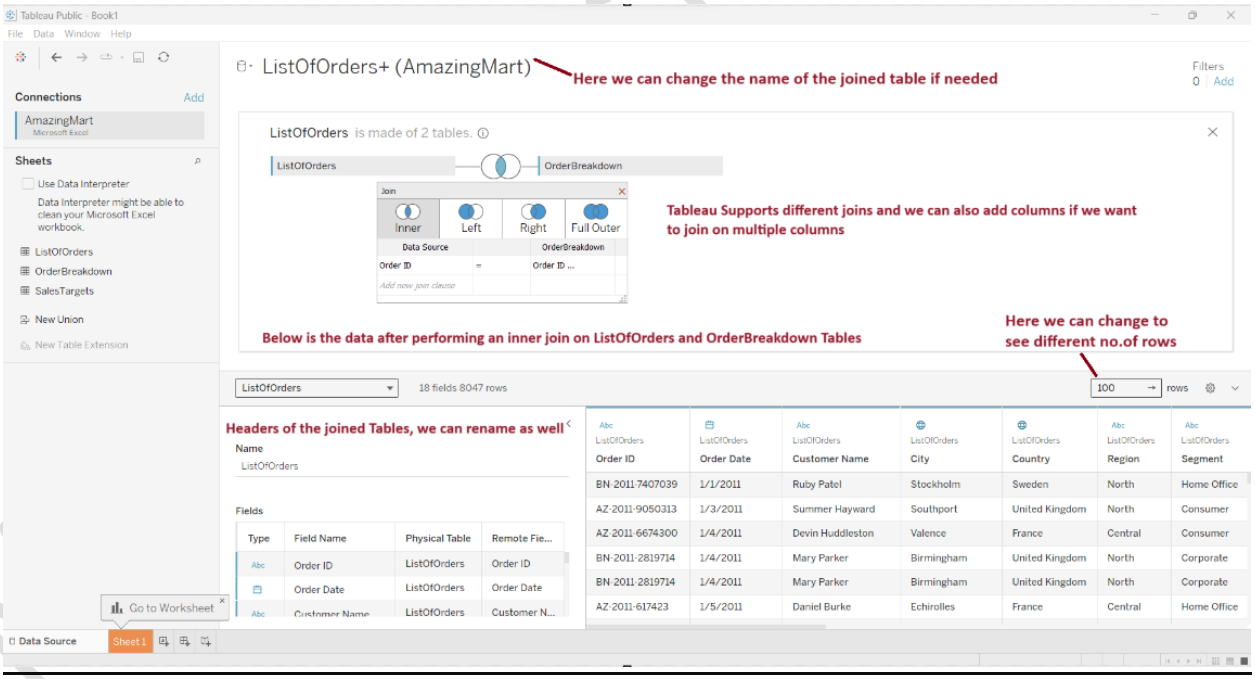
Unlike joining, blending is not performed at the data source level; instead, it's done at the visualization level. This means you can connect to multiple data sources independently and blend them in a single worksheet or dashboard.

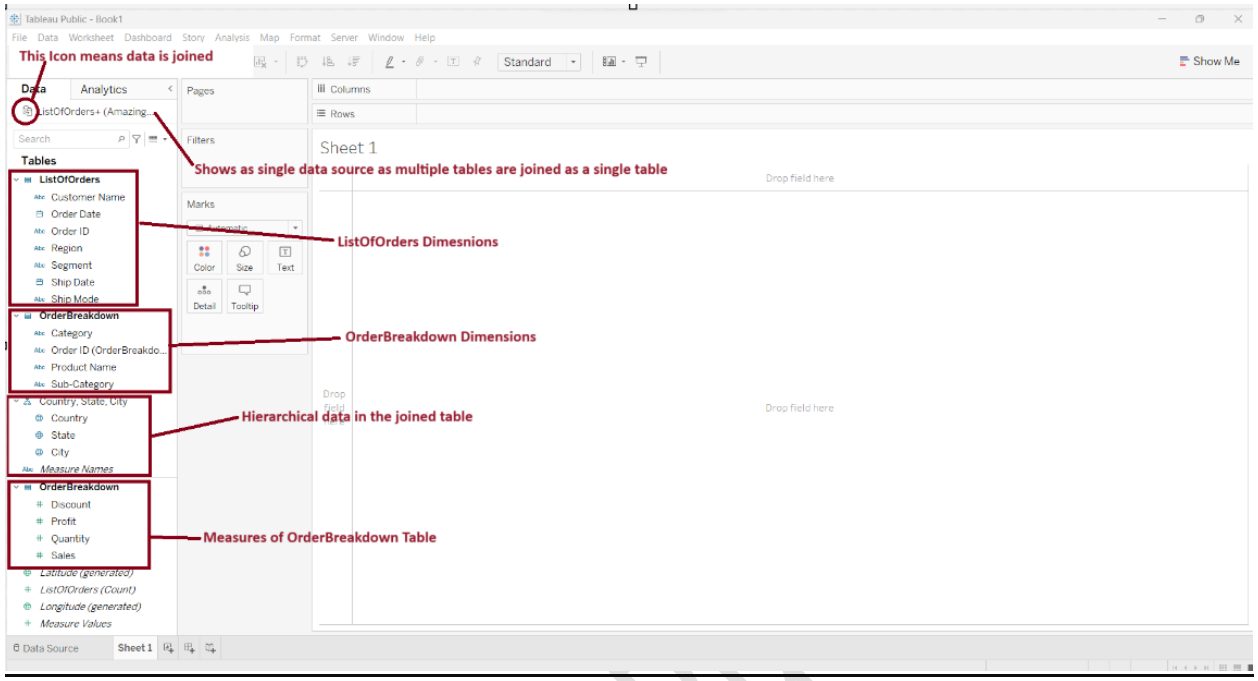
Blending is typically used when the data you want to analyze resides in separate data sources, and there is no common field for joining the tables directly. Instead, you can create relationships between the data sources at the visualization level.

Mohan Kumar Pulibanti

Joining in Tableau



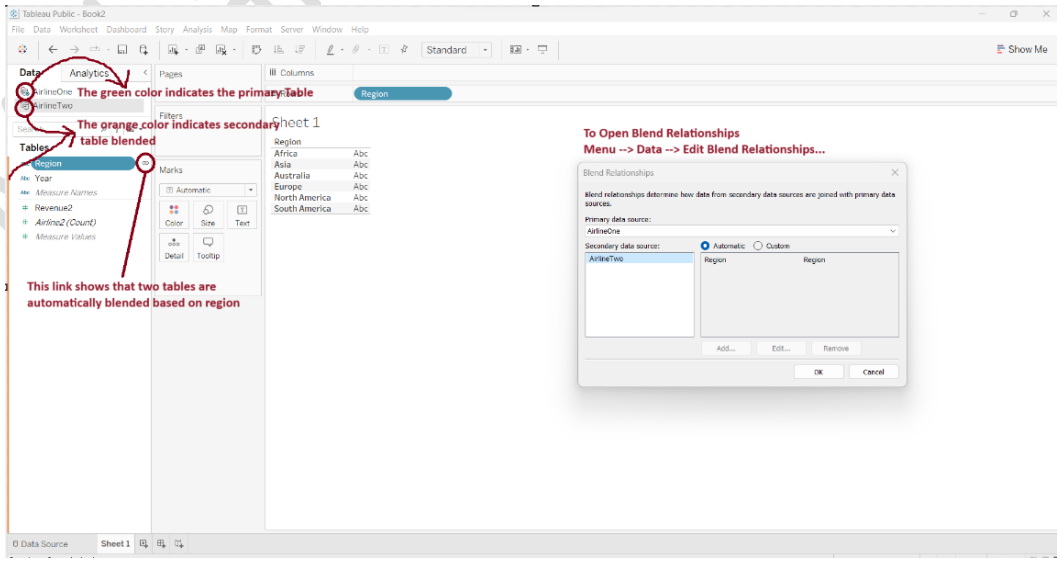


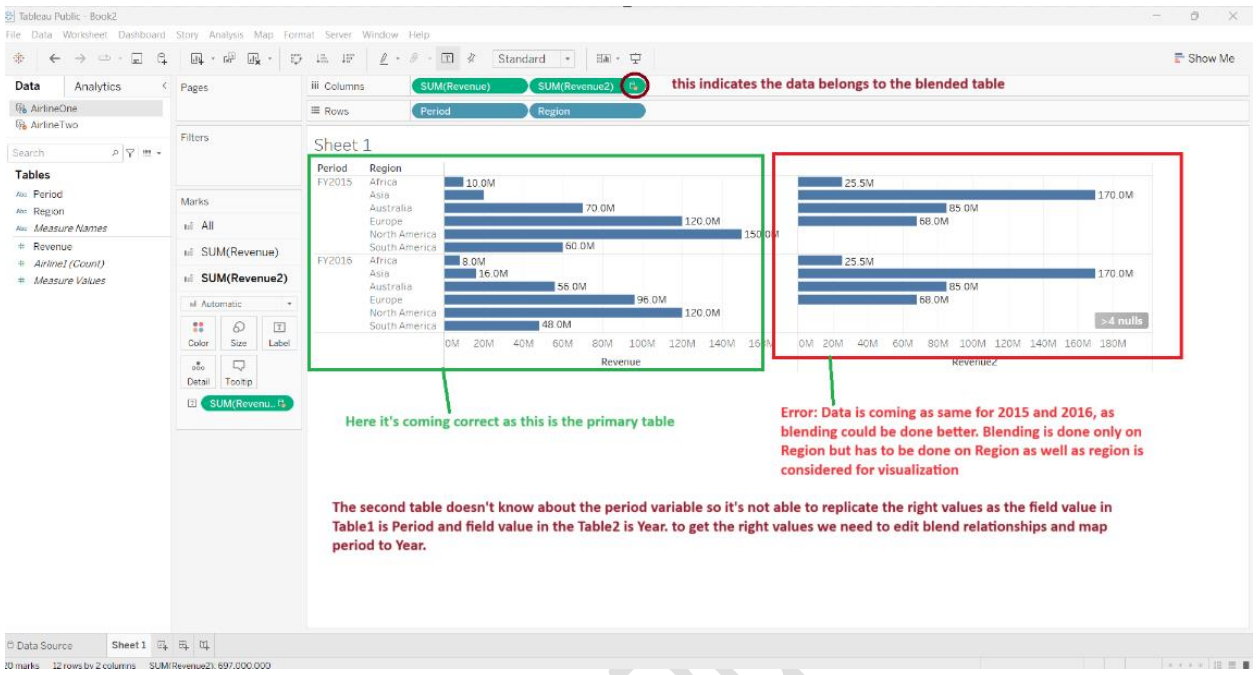


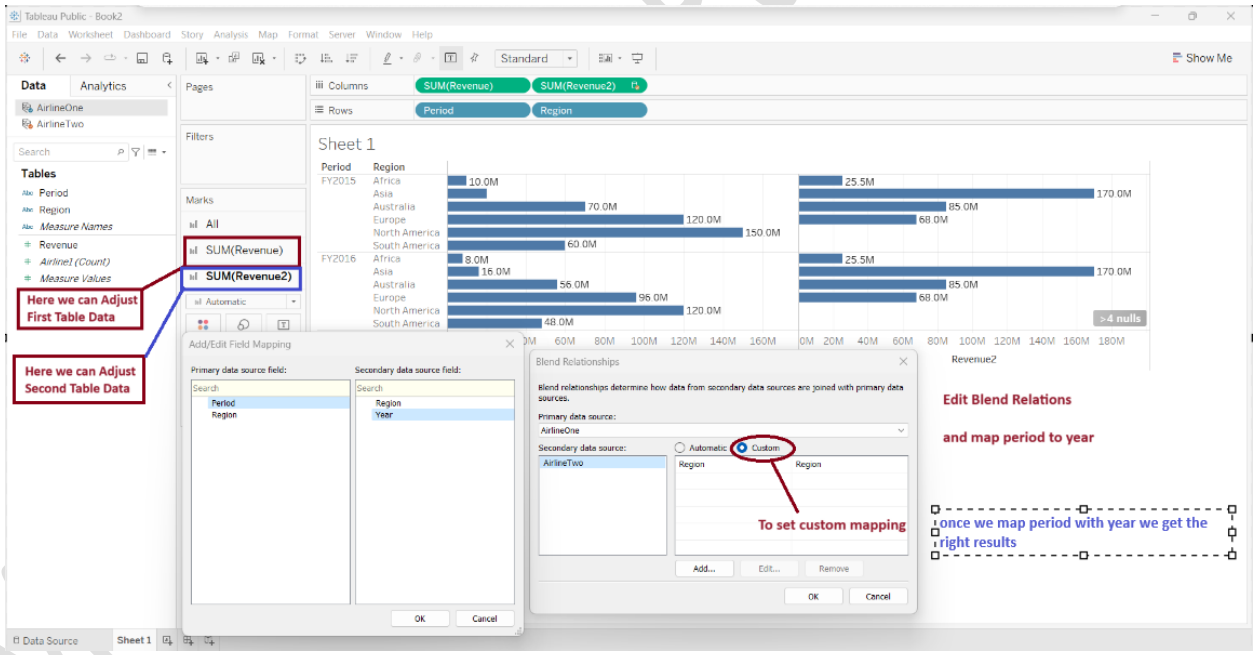
Data Blending in Tableau

Download Dataset here:
[Tableau_Data_Sets/Airline_Comparison_Data.xlsx at main · MohanPulibanti/Tableau_Data_Sets (github.com)]

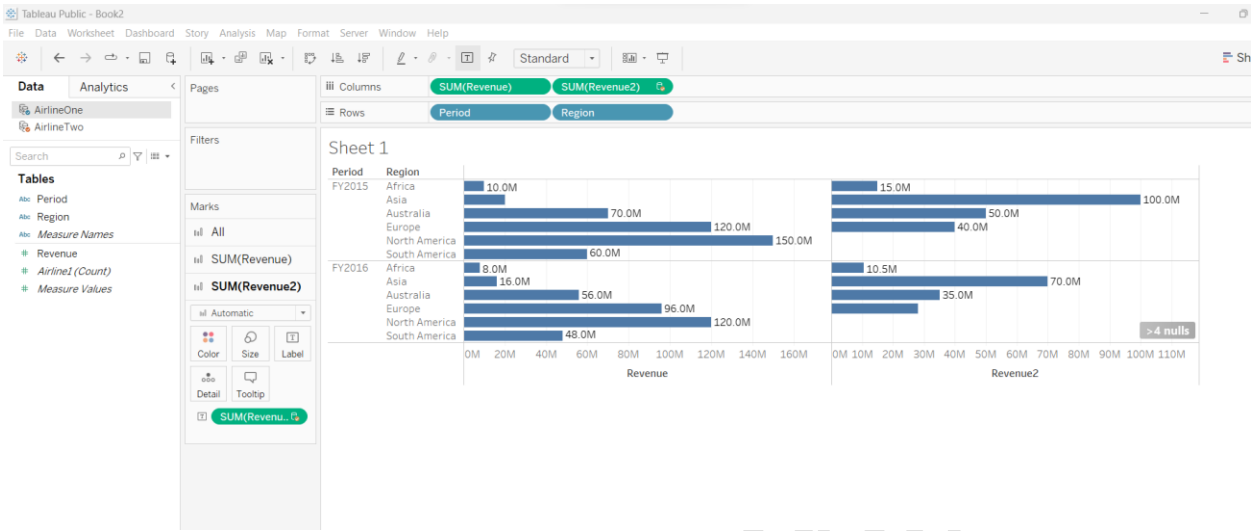
- Data Blending is done on a per-sheet basis. When blending is applied on a sheet it will not be visible on another sheet.
- From the above data source, add each sheet as a different data source, so that we get two data sources and from the first data source drag the region into the rows.



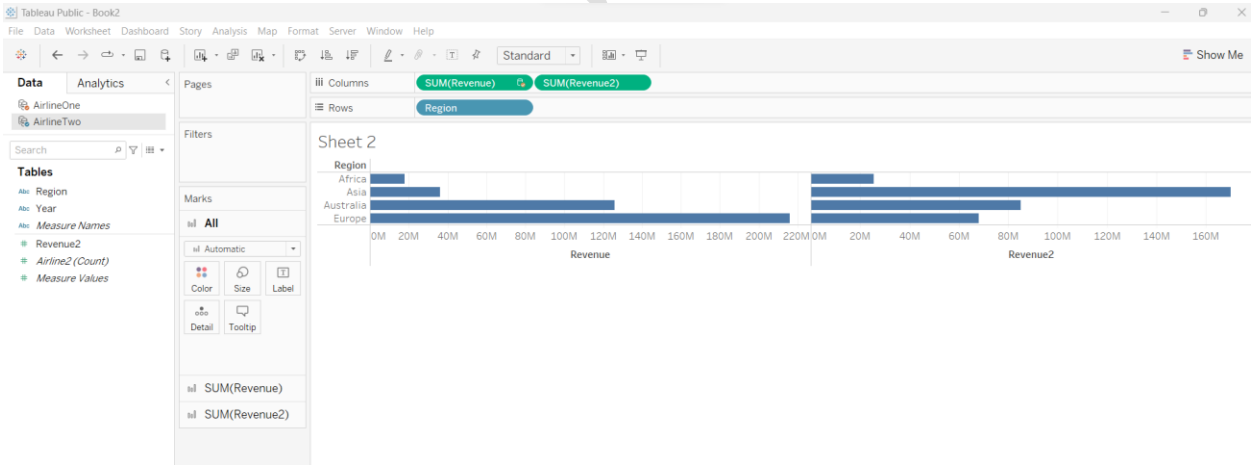




- During Blending, What tableau will do is, it creates join on the fly. It actually instead of joining the data at row level it will first send queries to each of the dataset and then it will aggregate them to the level that is required (In our case it is region) at the level of granularity of the visualization and then only it joins them.



- Data Blending is a smart Left join, in the above diagram Airline One is primary table and it has 6 countries, so left join will bring all the 6 countries whereas table two even though has 4 countries two other rows are being shown as null.
- When AirlineTwo is made as the primary table, 2 countries will be emitted from the graph which is not part of AirlineOne.

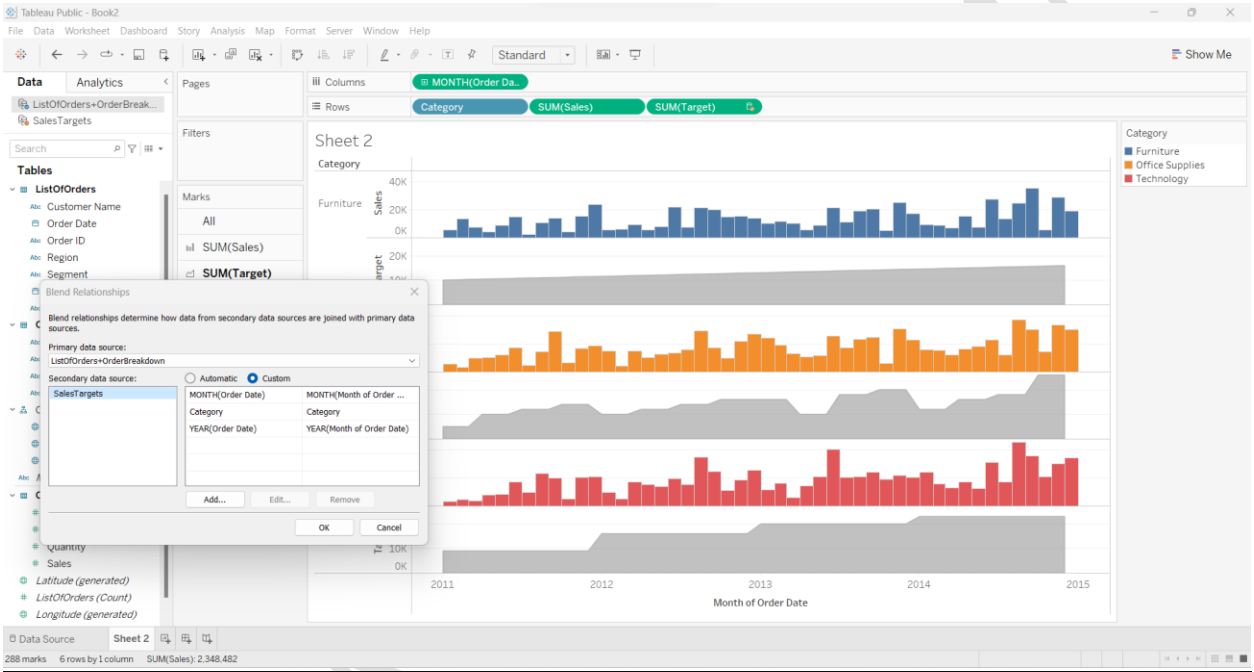


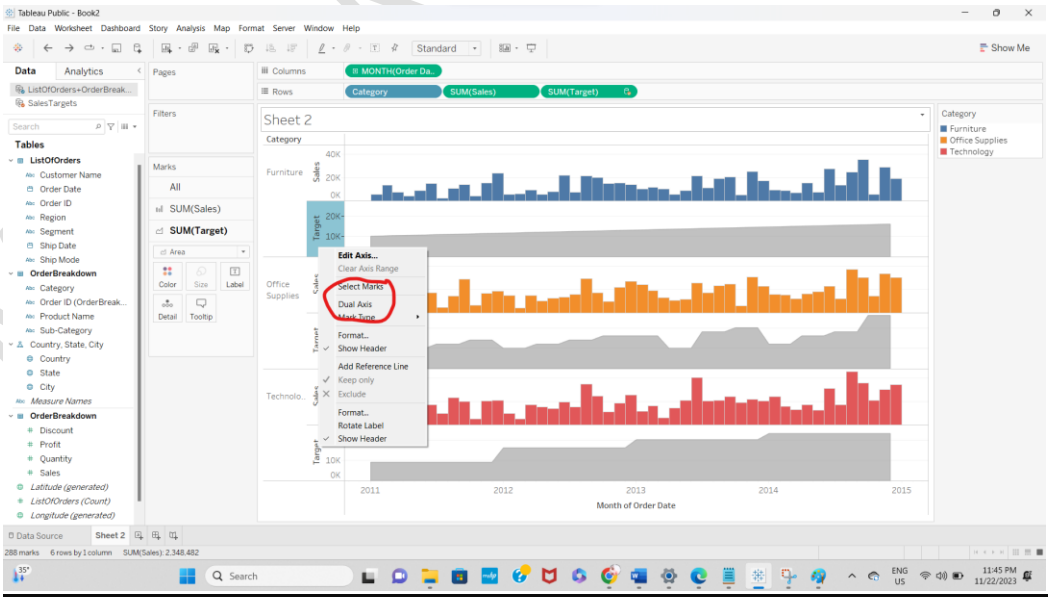
Dual Axis Chart:

Download Dataset here:

[Tableau_Data_Sets/AmazingMart.xlsx at main · MohanPulibanti/Tableau_Data_Sets (github.com)]

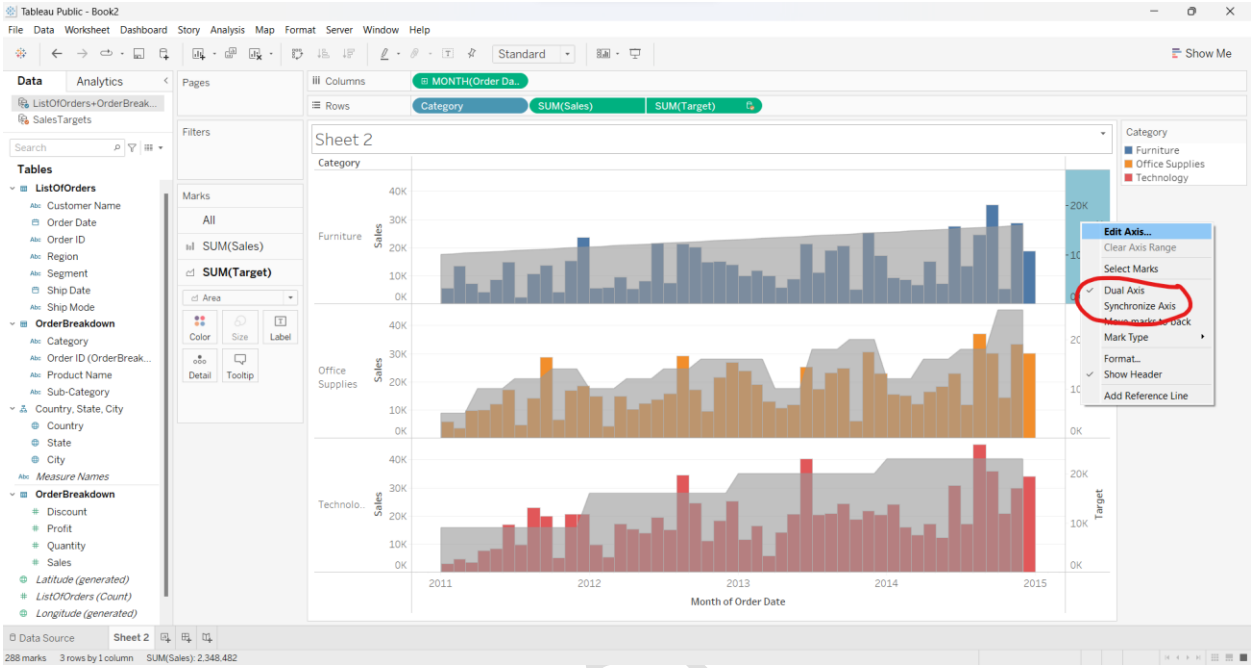
Blending is done on multiple fields to get the joining the data perfectly.



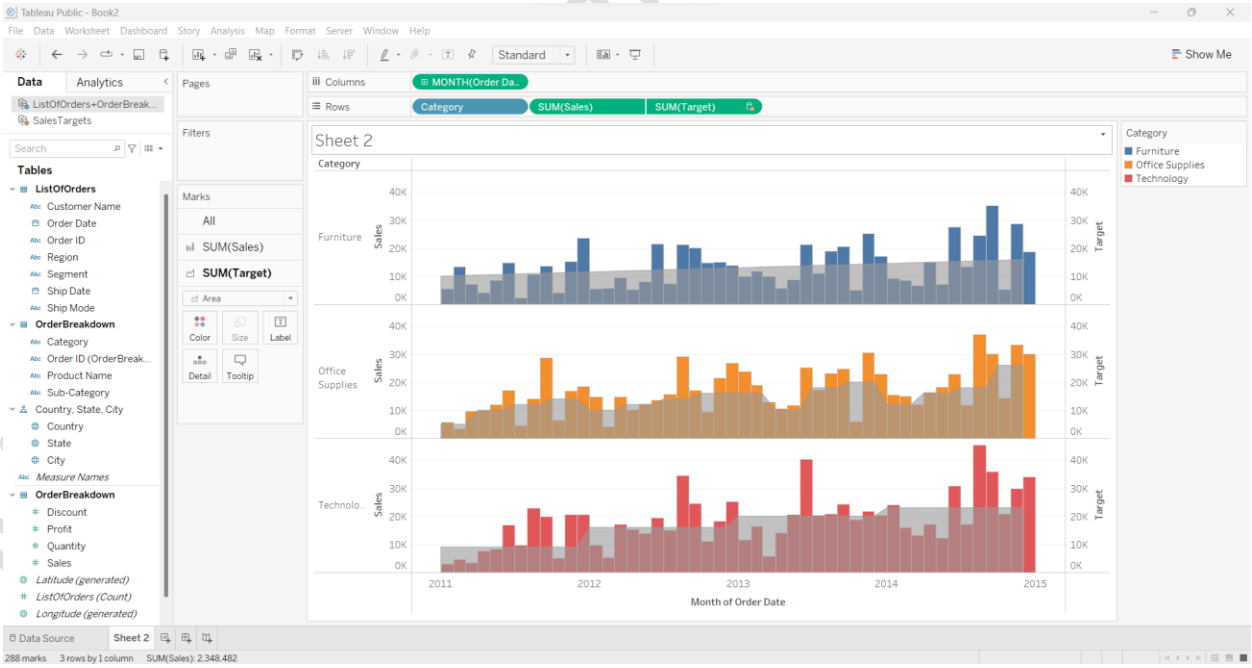


Dual Axis will create a single graph by merging two graphs which have one common axis.

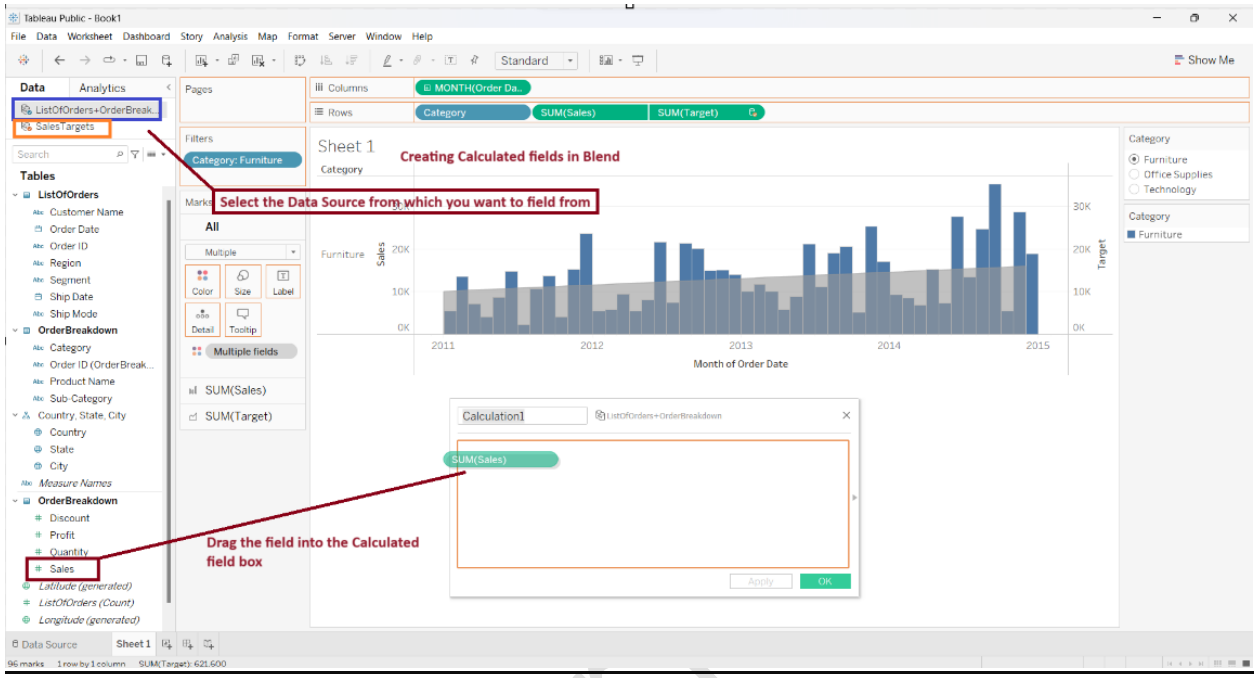
And then to sync the dimensions on both the axis, we need to click on Synchronize Axis to make the dimensions on both the axis same,

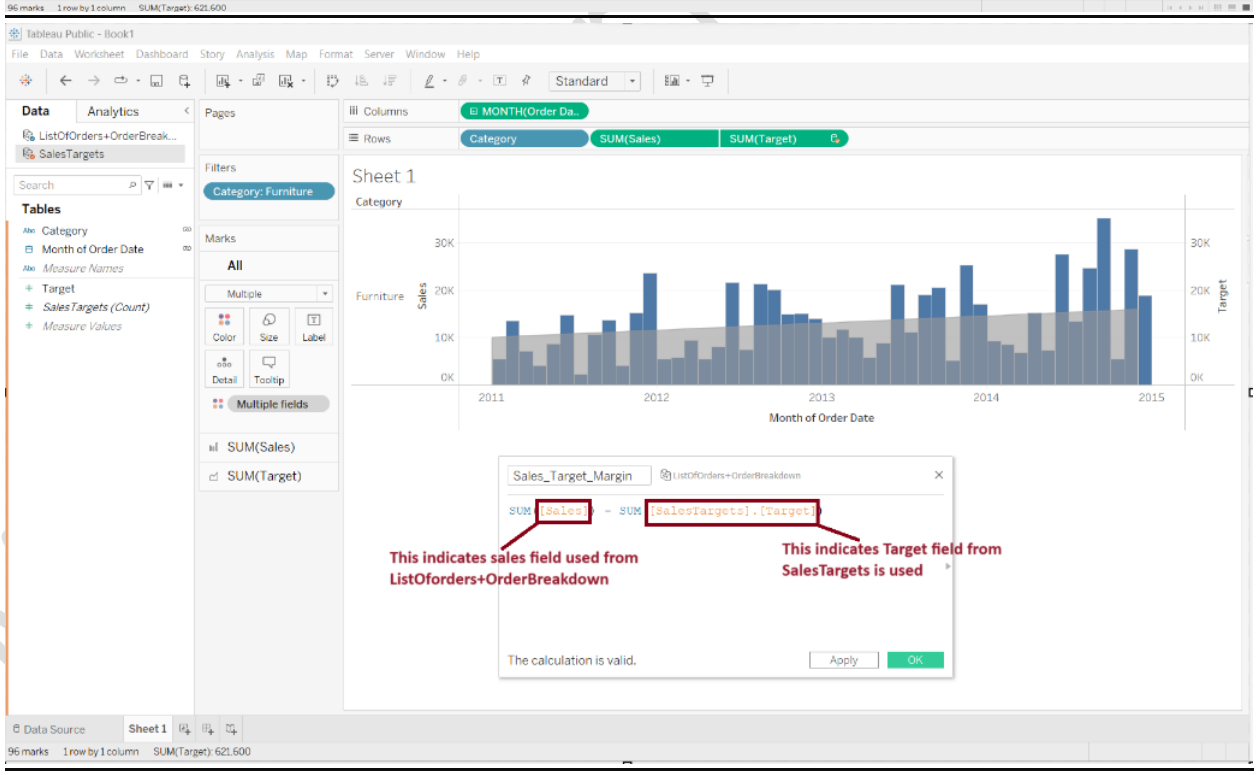


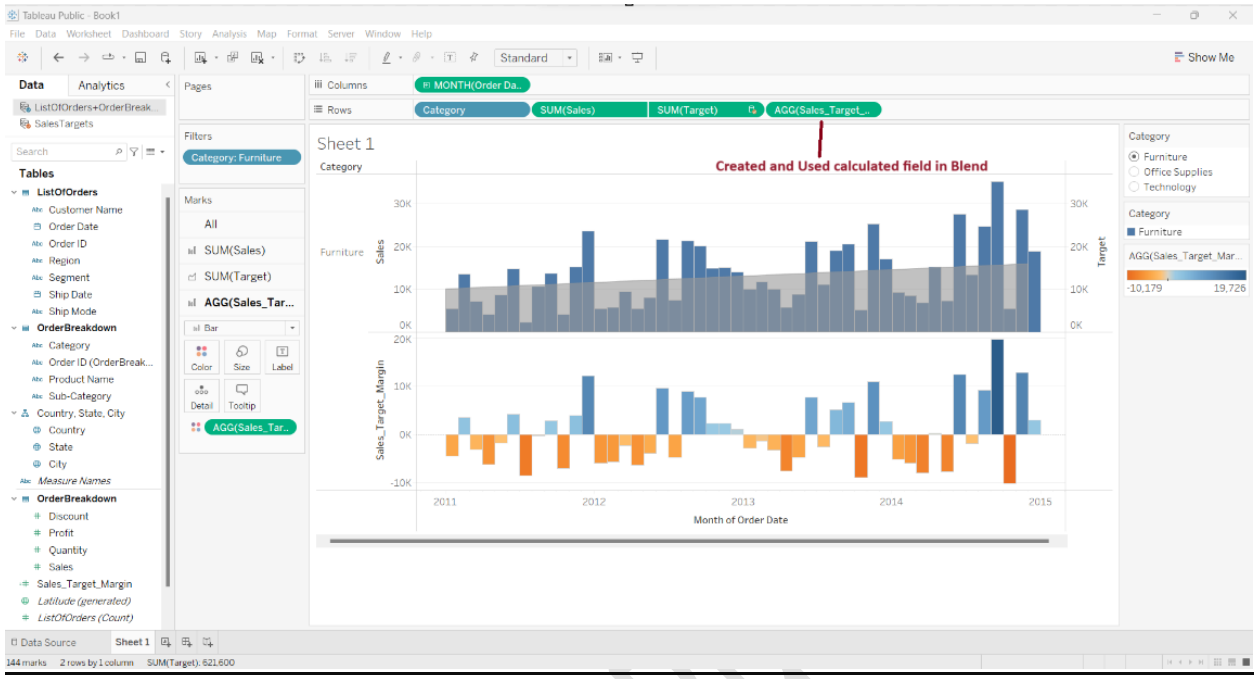
After Axis are synchronized, the graph looks as below



Creating Calculated Fields in Blend:

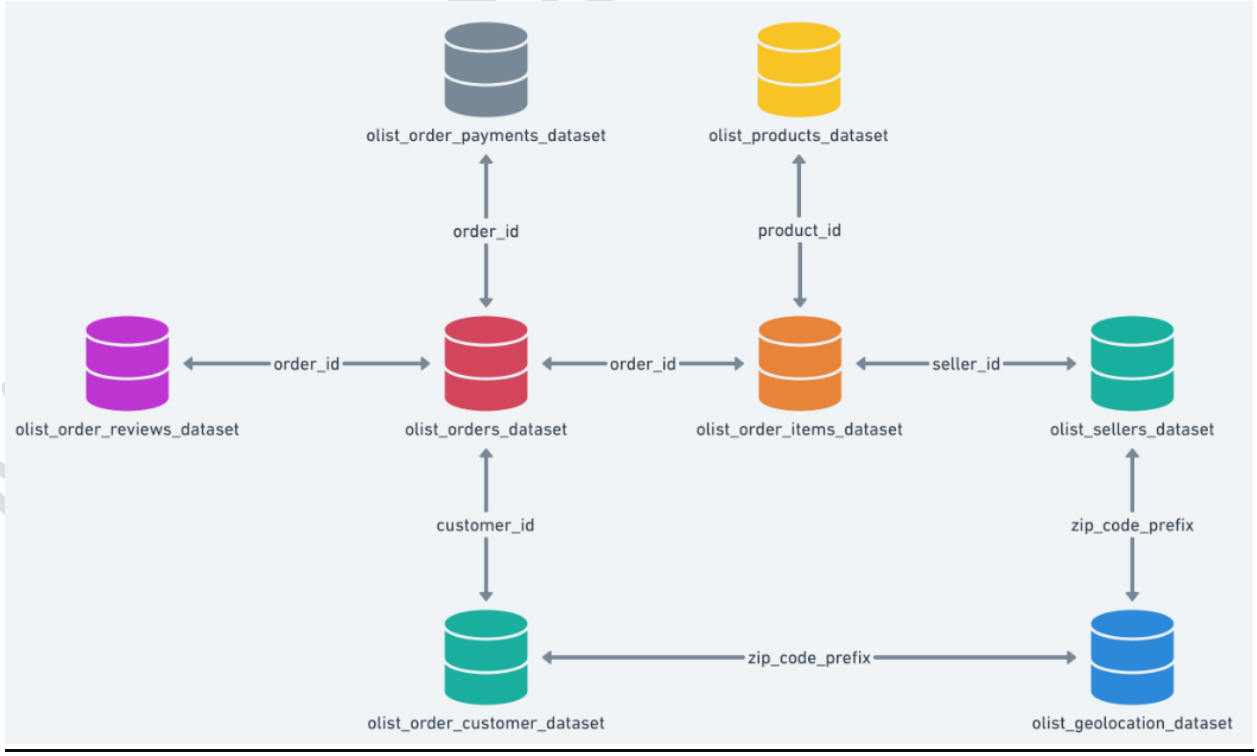






Working with Relationships in Tableau:

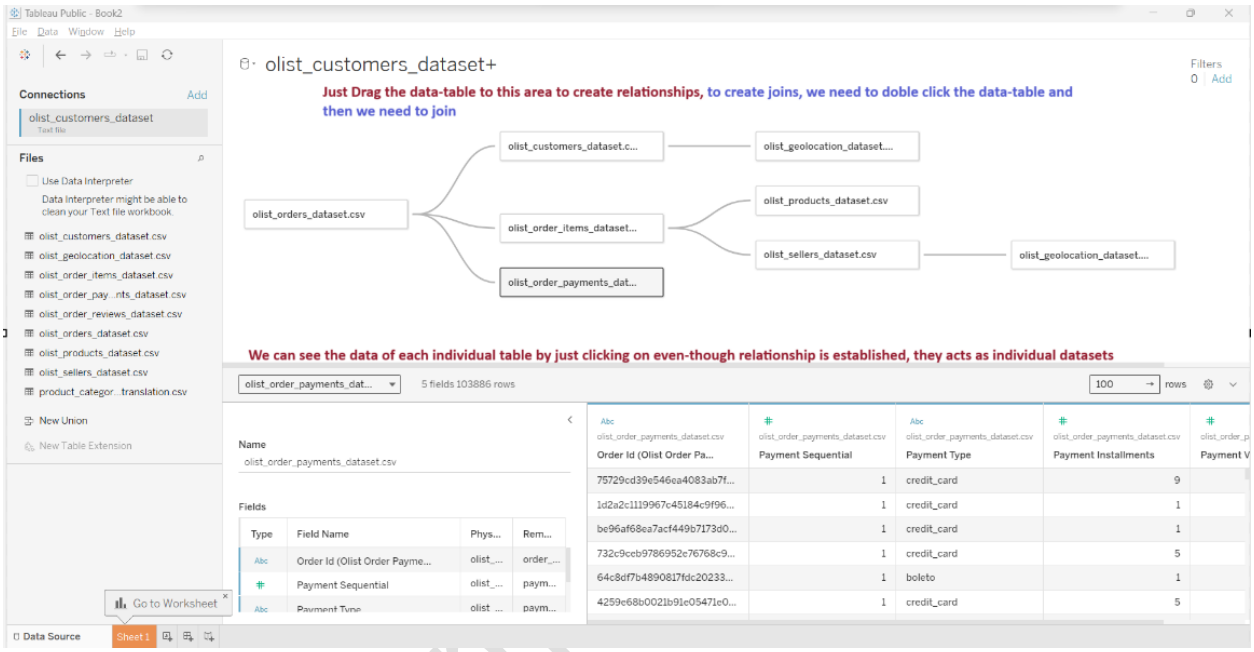
Kaggle Brazilian ecommerce Dataset and Schema:



Download Dataset here:

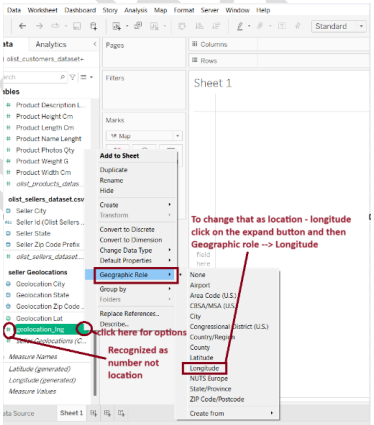
https://www.kaggle.com/datasets/olistbr/brazilian-ecommerce/

Creating Relationships in Tableau

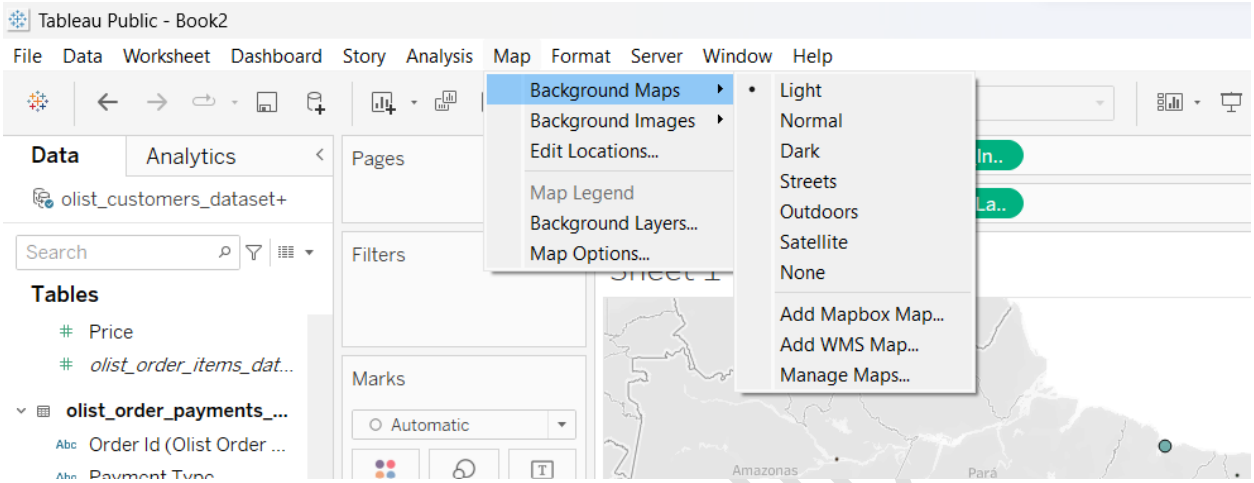


Challenge:

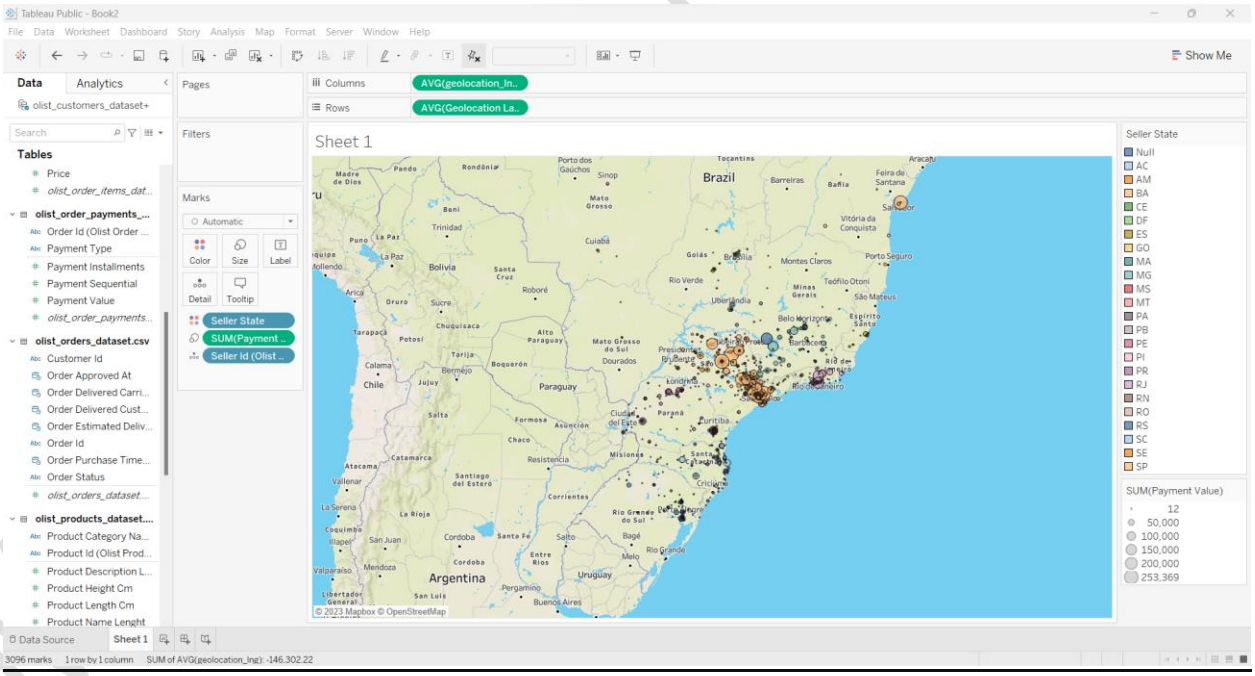
Given this dataset, Your manager has asked you to create a map showing sellers in their geolocation with bubbles to illustrate how much money in total payments they have received from customers.



If we want to change the Map background to different options, below is where we can do that



Final Output:



- we can combine relationships with Joins

  Double click on the table and join the table and comeback → this way we can join relationships to joins

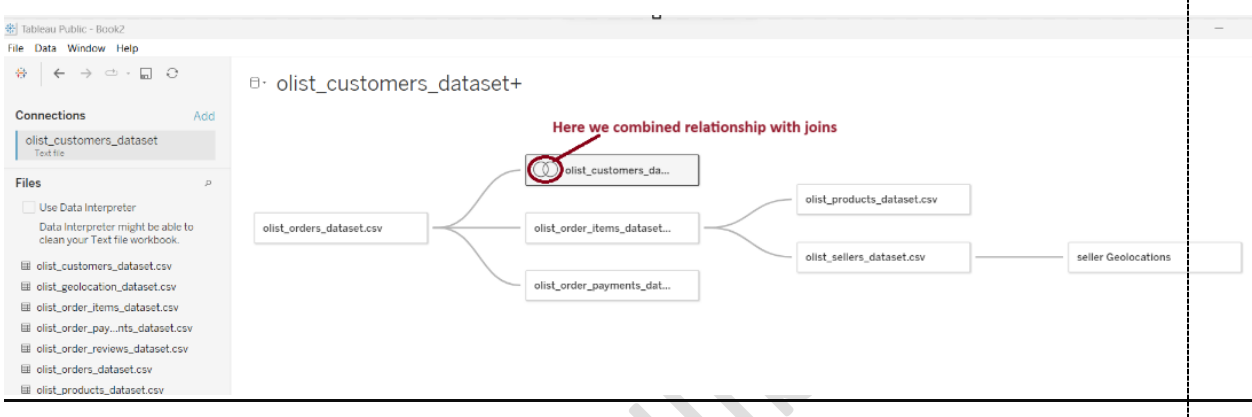
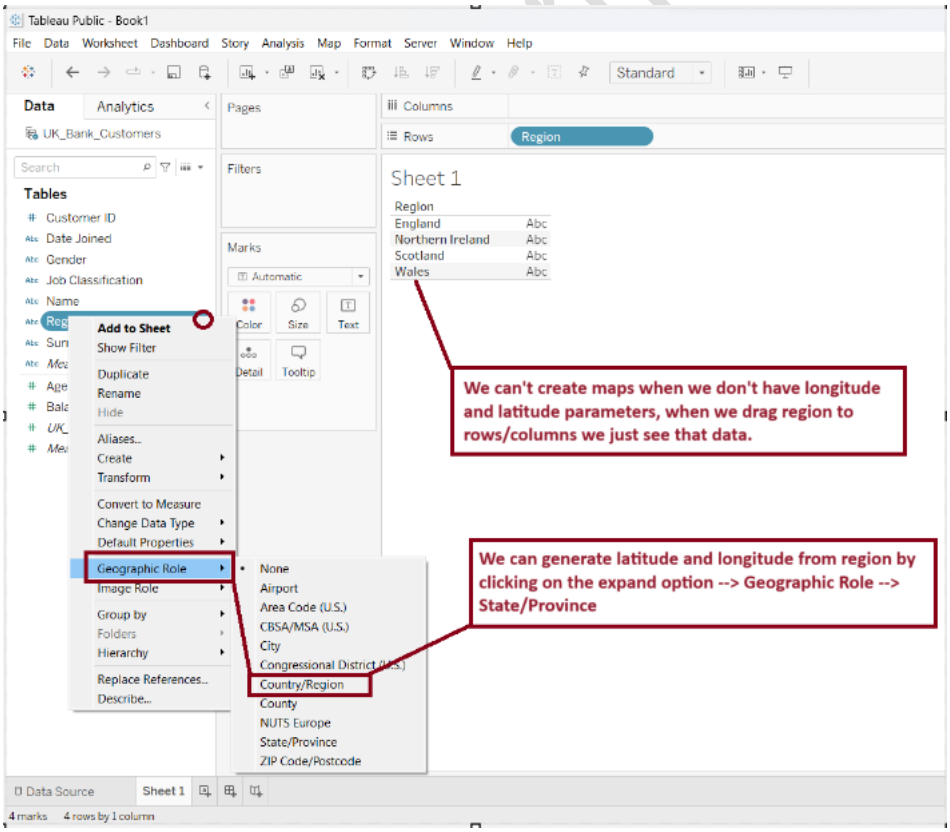
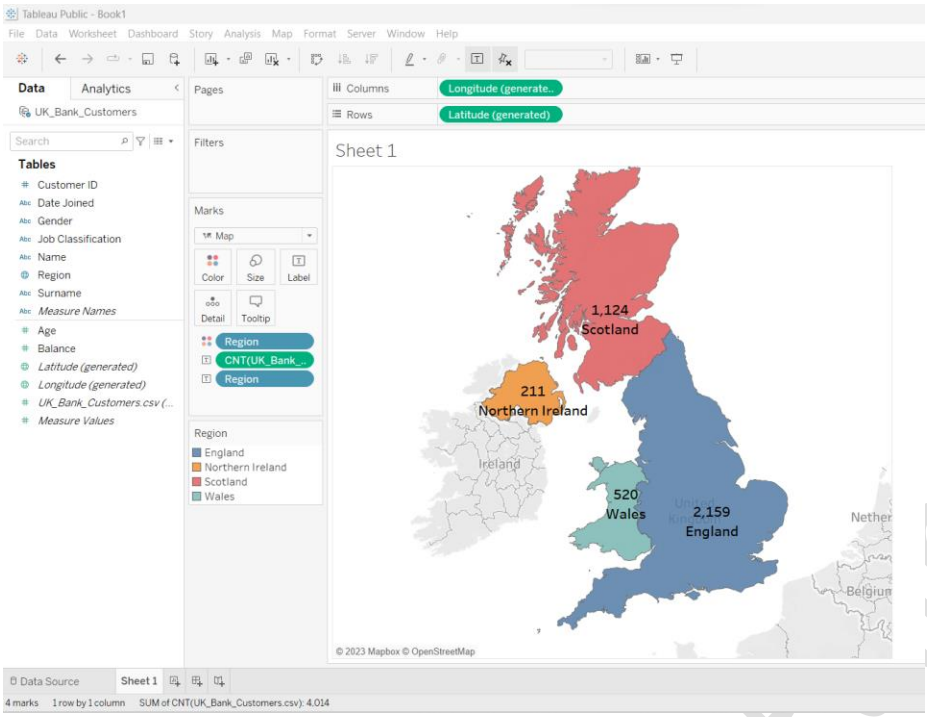


Table Calculations and Advanced Dashboards

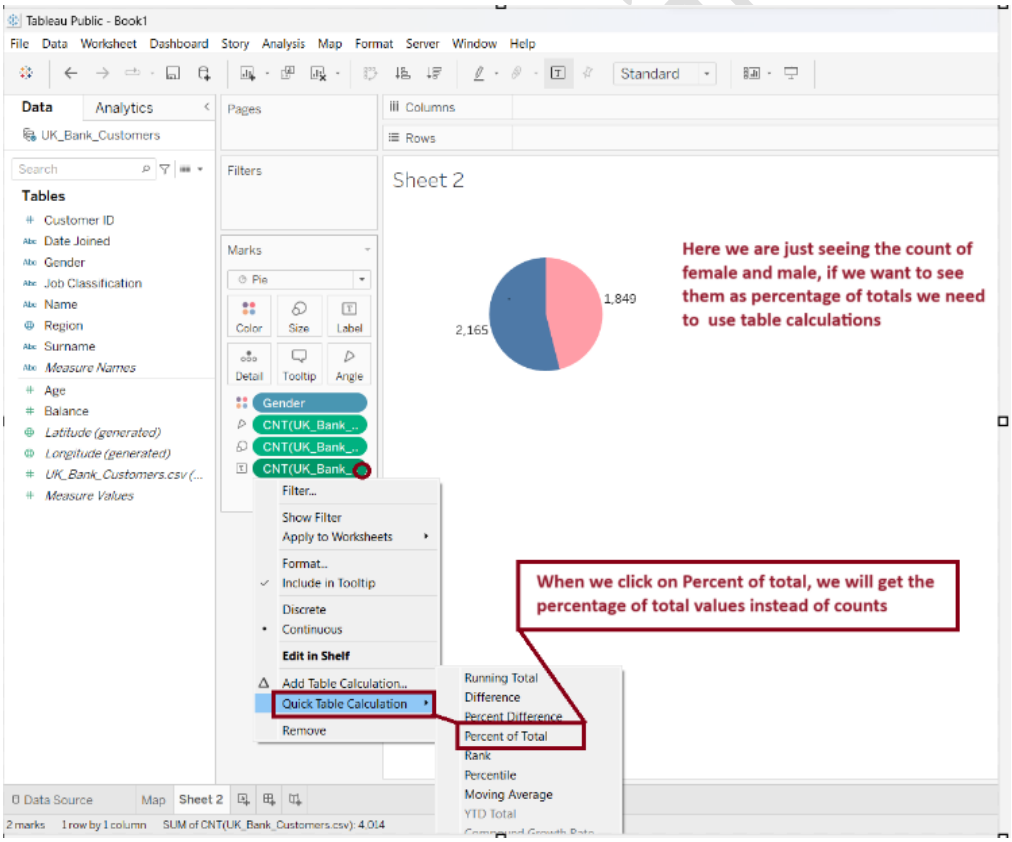
Download Dataset here:

[Tableau_Data_Sets/UK_Bank_Customers.csv at main · MohanPulibanti/Tableau_Data_Sets (github.com)]

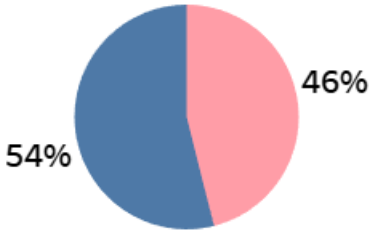




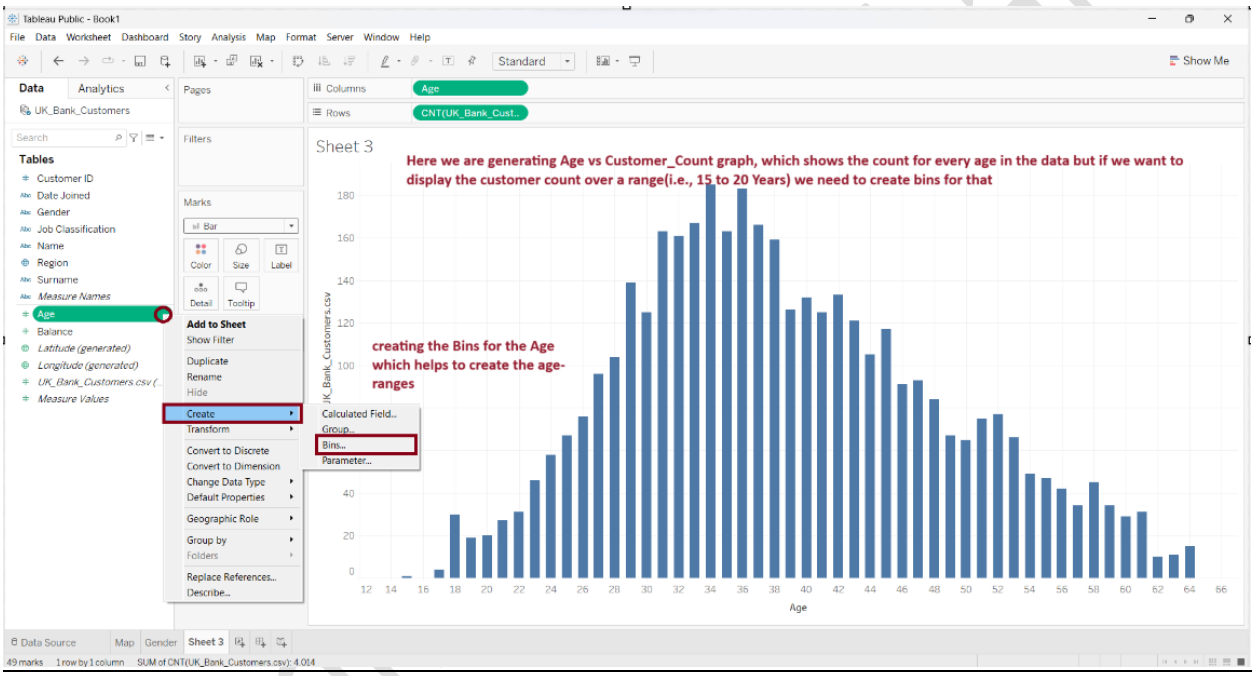
Creating Table calculations:



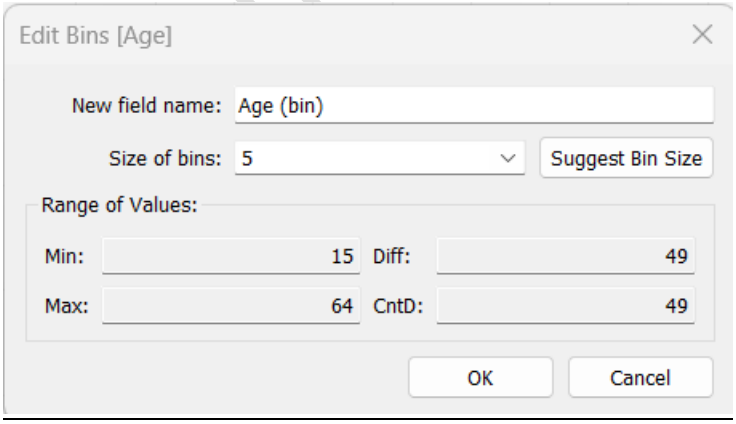
After Applying Percent of total table calculation, below is the graph

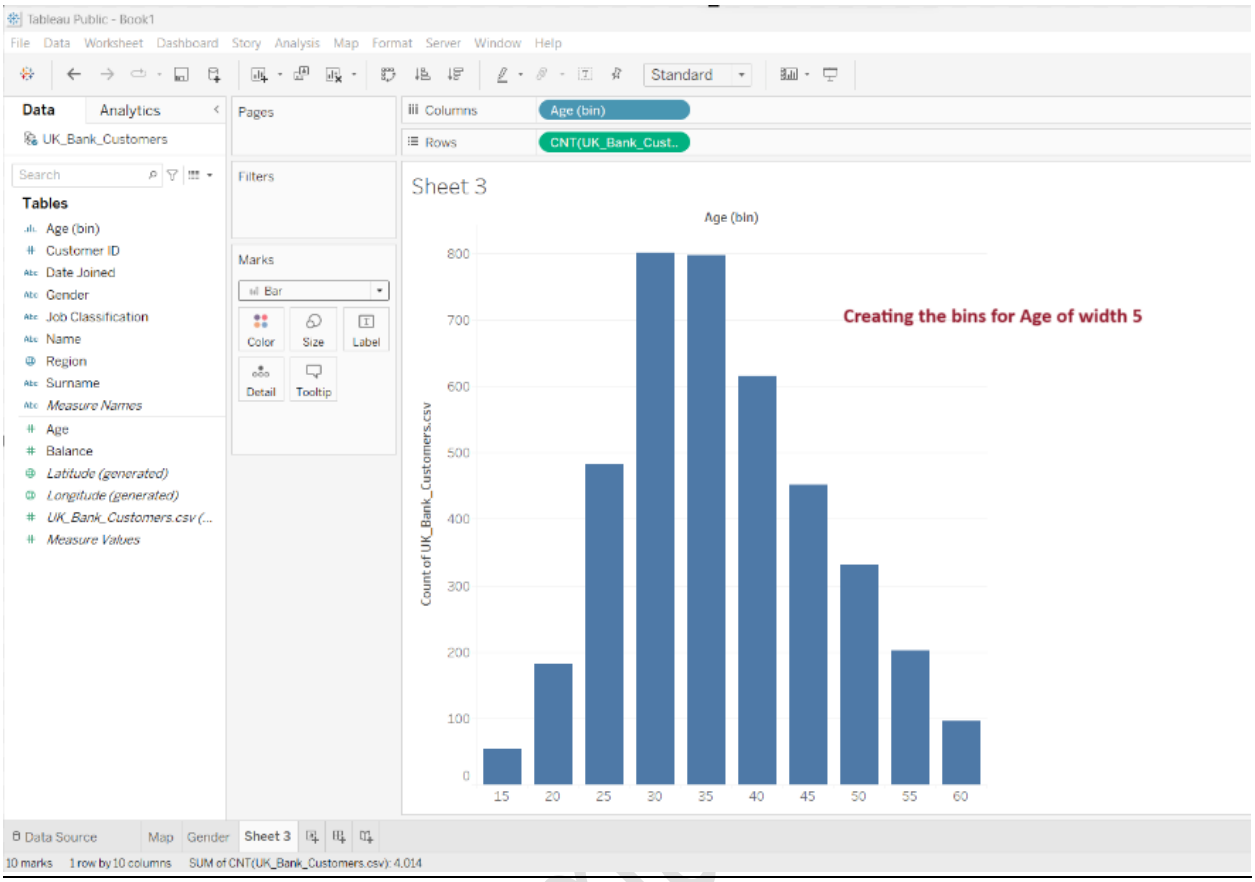


Creating Bins and Distributions for Age



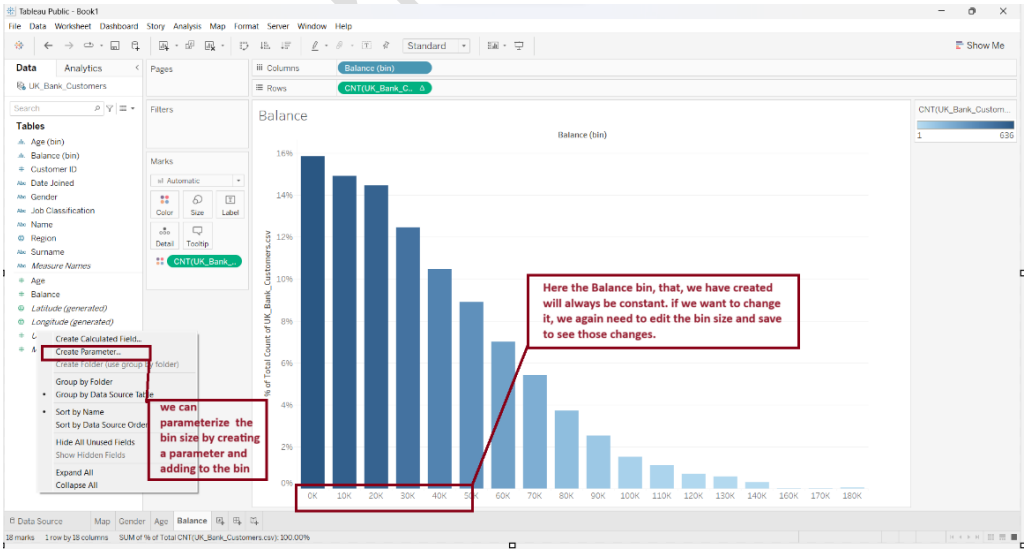
After Clicking on the Create → Bins… we get the below window where we can specify the bin width.





Leveraging the power of Parameters

When we are creating bins, we are giving a specific bin size, if we want to make it as dynamic and give an option to change, we can create that using parameters.



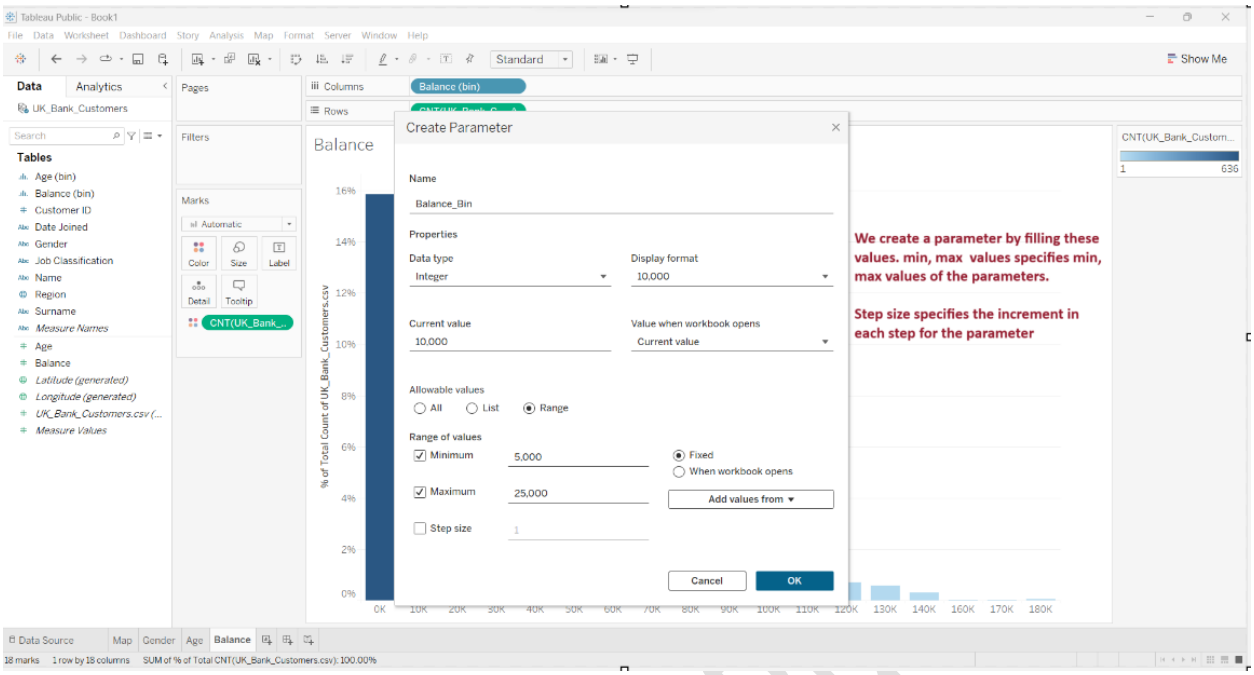


Tableau Public - Book1
Create Parameter
Name
Balance_Bin
Properties
Data type
Integer
Display format
10,000
Current value
10,000
Value when workbook opens
Current value
Allowable values
All
List
Range
Range of values
Minimum
5,000
Maximum
25,000
Step size
1
Fixed
When workbook opens
Add values from
Cancel
OK
We create a parameter by filling these values. min, max values specifies min, max values of the parameters.
Step size specifies the increment in each step for the parameter

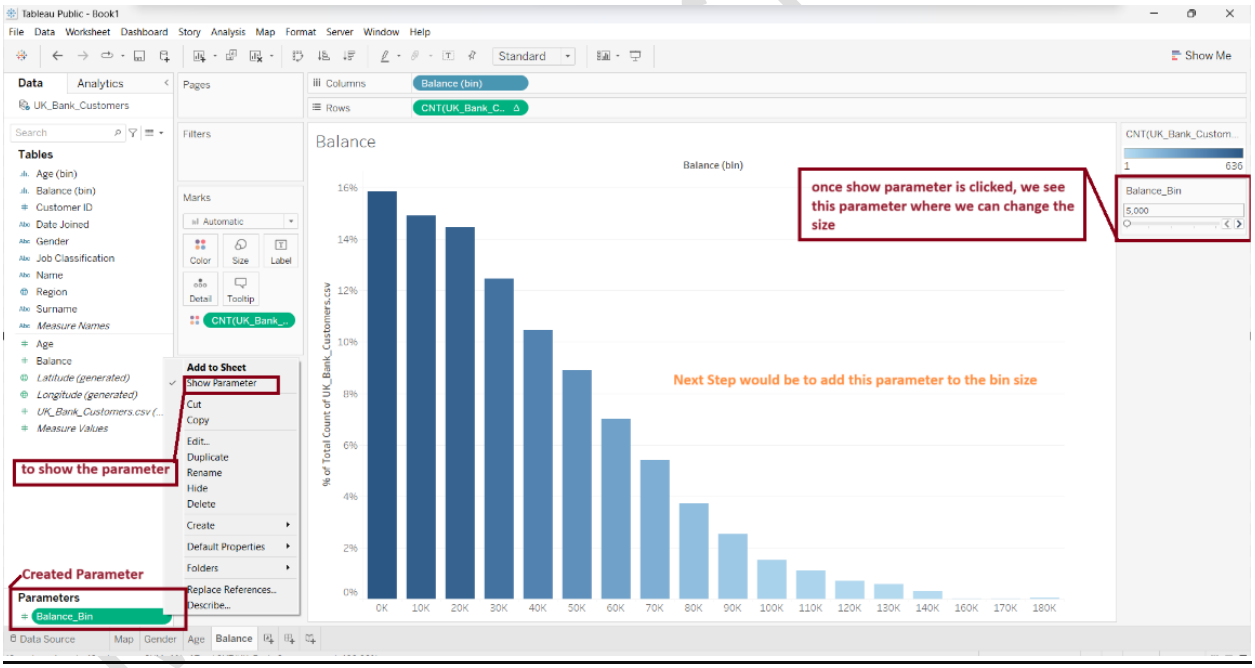


Tableau Public - Book1
once show parameter is clicked, we see this parameter where we can change the size
Balance_Bin
5,000
Next Step would be to add this parameter to the bin size
Add to Sheet
Show Parameter
to show the parameter
Created Parameter
Parameters
Balance_Bin

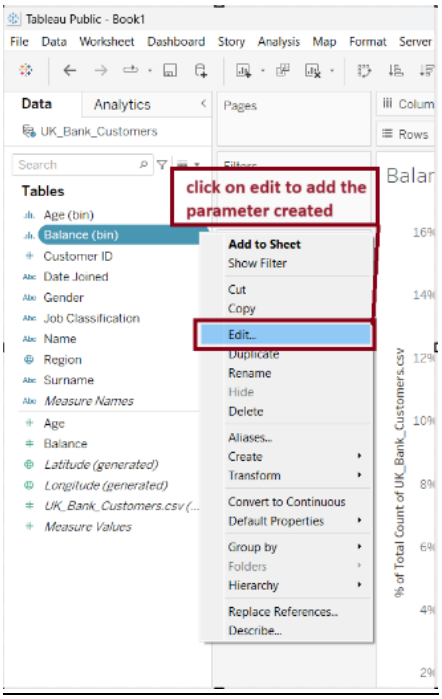


Tableau Public - Book1
File Data Worksheet Dashboard Story Analysis Map Format Server
Data Analytics
UK_Bank_Customers
Pages
Tables
click on edit to add the parameter created
Add to Sheet
Show Filter
Cut
Copy
Edit...
Duplicate
Rename
Hide
Delete
Aliases...
Create
Transform
Convert to Continuous
Default Properties
Group by
Folders
Hierarchy
Replace References...
Describe...
% of Total Count of UK_Bank_Customers.csv

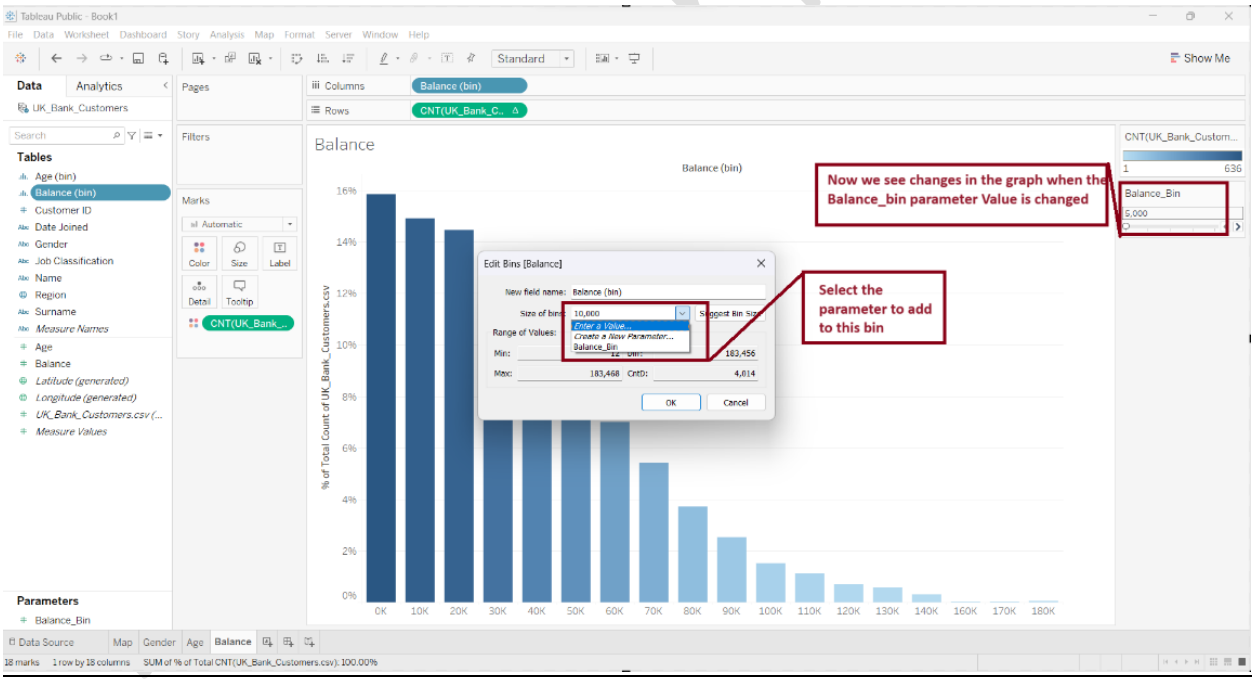
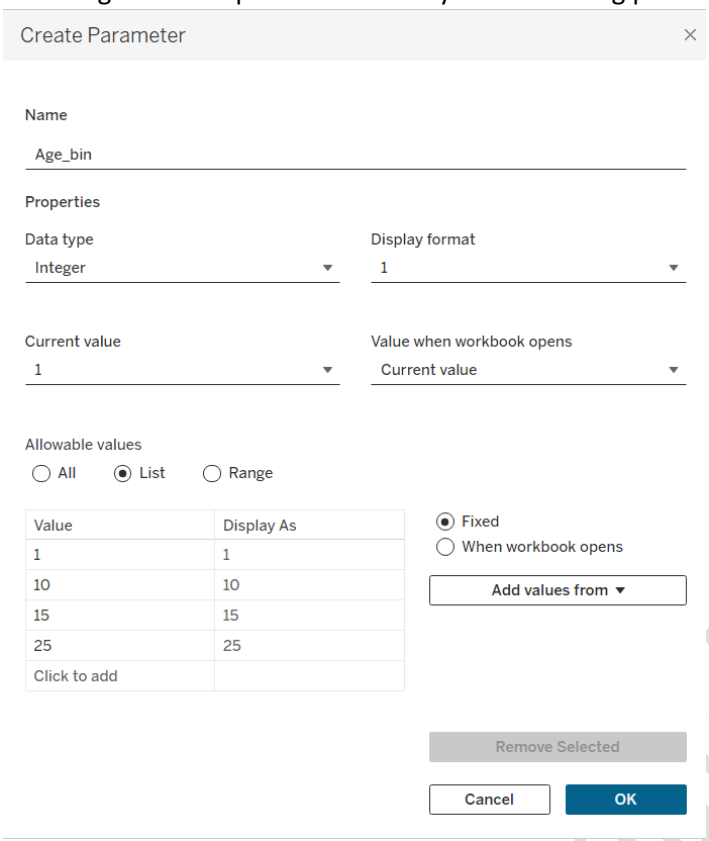
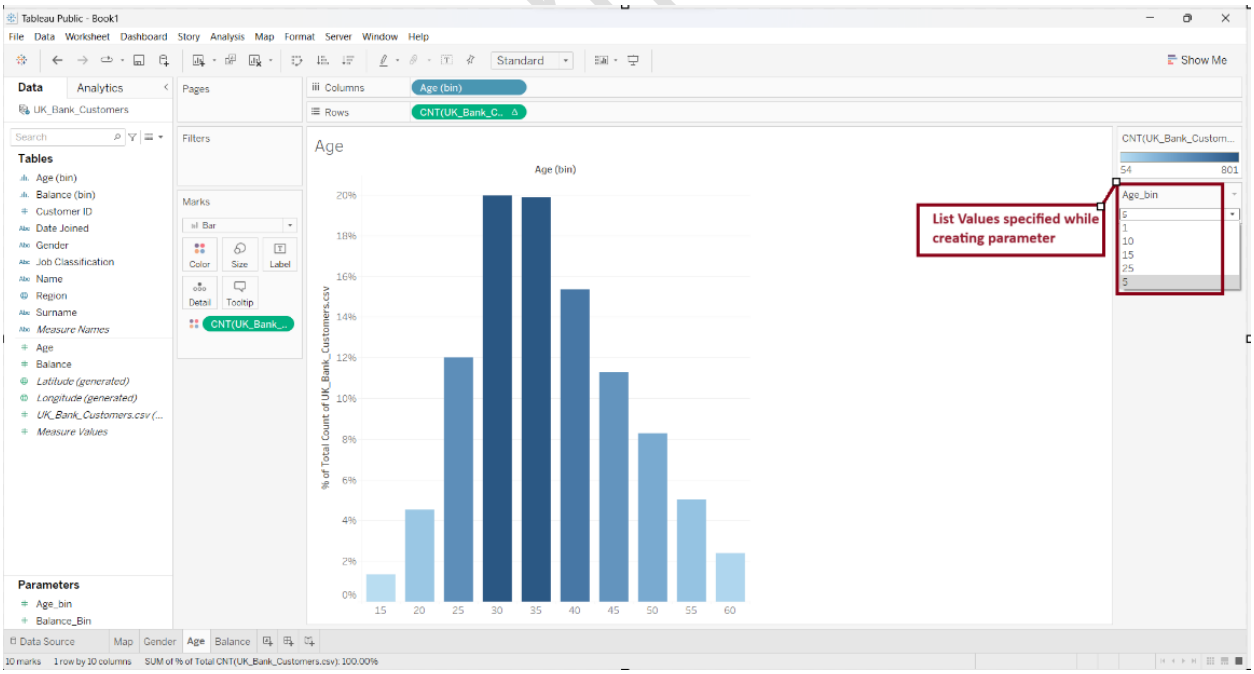
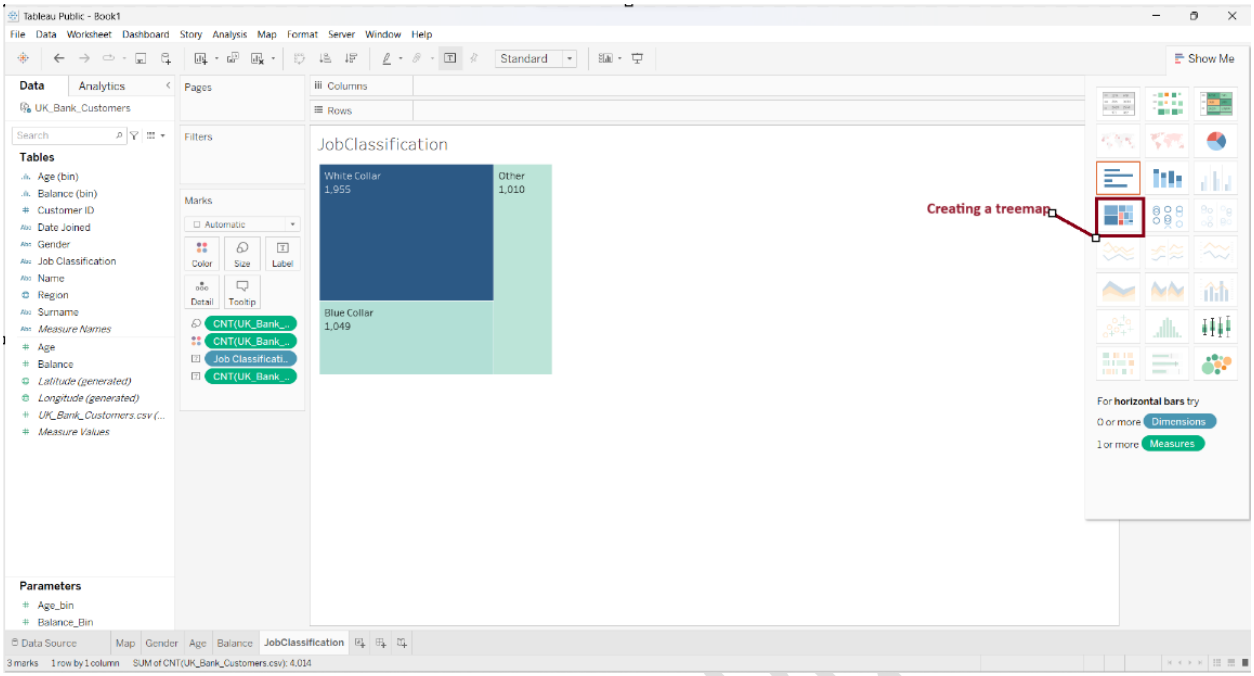


Tableau Public - Book1
Data Analytics
UK_Bank_Customers
Pages
Columns Balance (bin)
Rows CNT(UK_Bank_C...
Filters
Marks
Balance
Balance (bin)
Now we see changes in the graph when the Balance_bin parameter Value is changed
Balance_Bin
Edit Bins [Balance]
New field name: Balance (bin)
Size of bins: 10,000
Enter a Value...
Create a New Parameter...
Balance_Bin
Range of Values:
Min: 183,456
Max: 183,468
CntD: 4,014
OK
Cancel
Select the parameter to add to this bin
% of Total Count of UK_Bank_Customers.csv
Parameters
Balance_Bin
Data Source Map Gender Age Balance

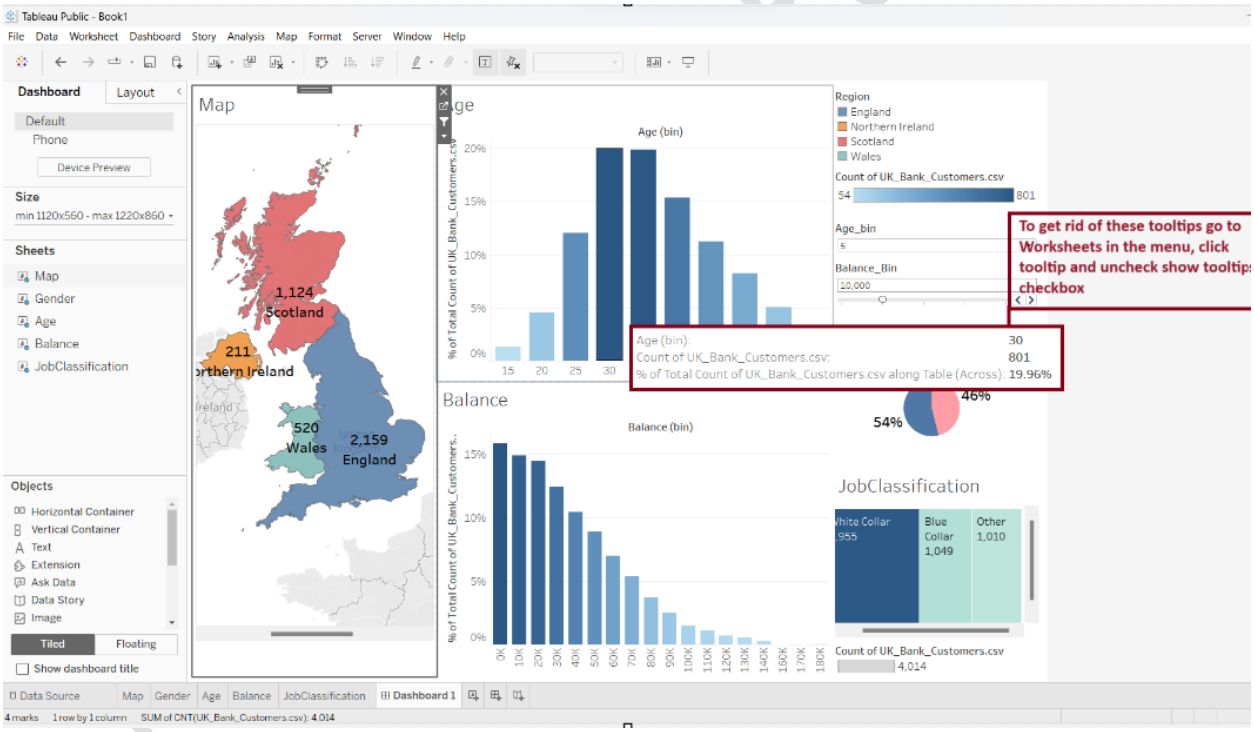
We can give some specific values only while creating parameter by using list option.



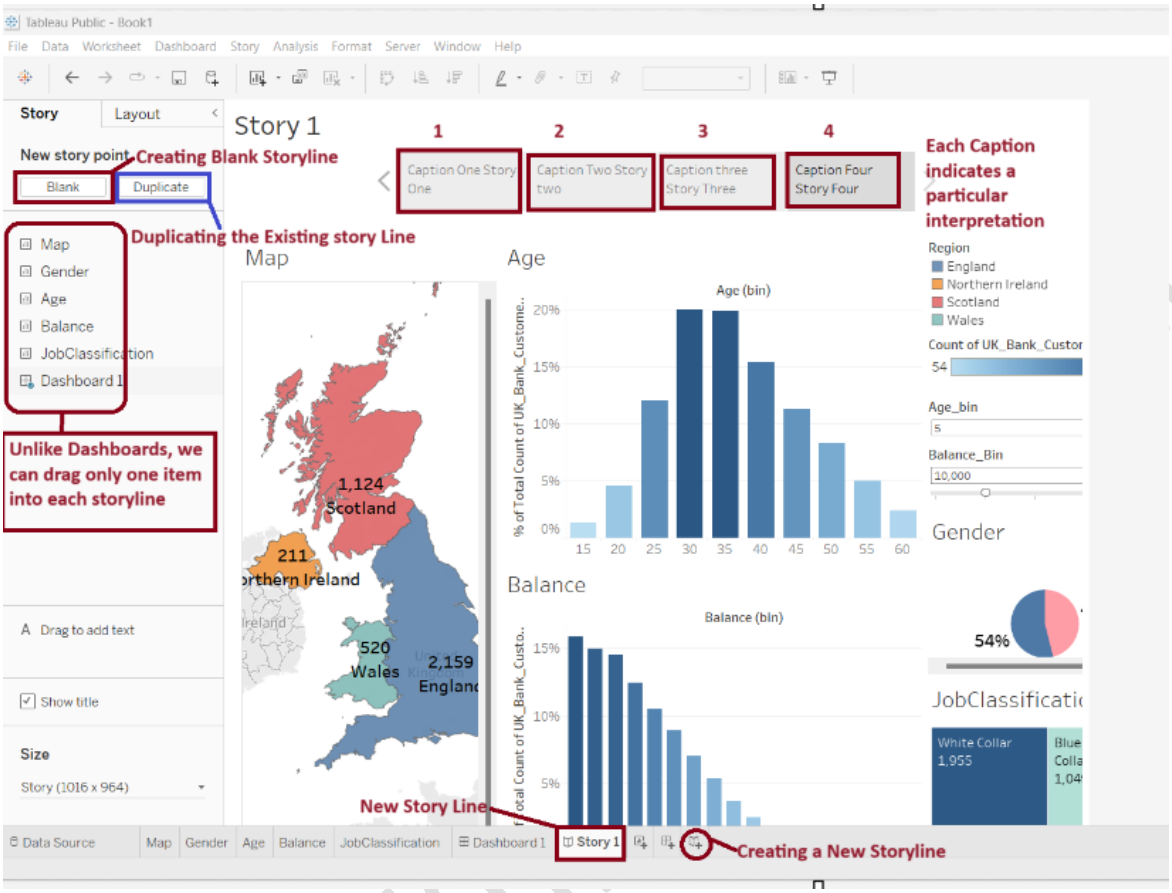




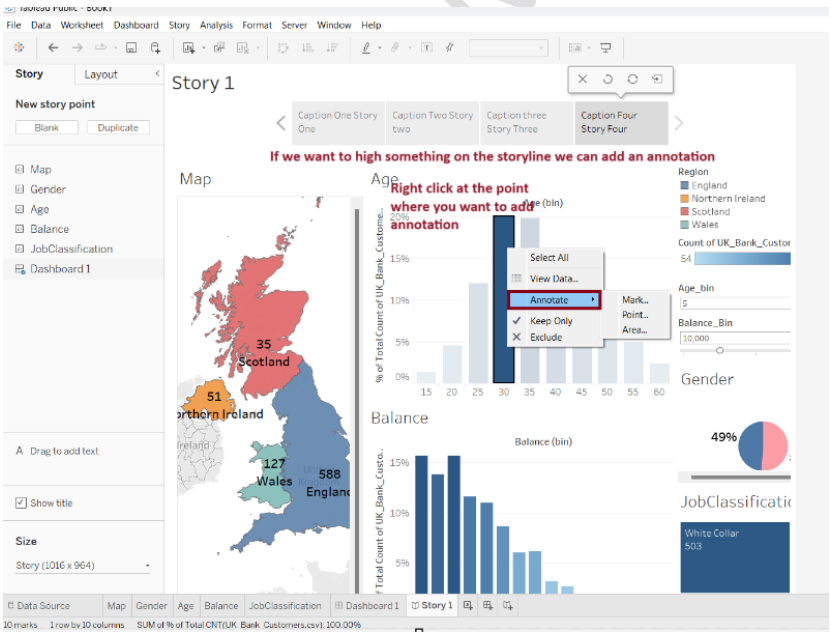
Dashboard and Applied Use as Filter for all the sheets and Uncheck Show tooltip for all the sheets



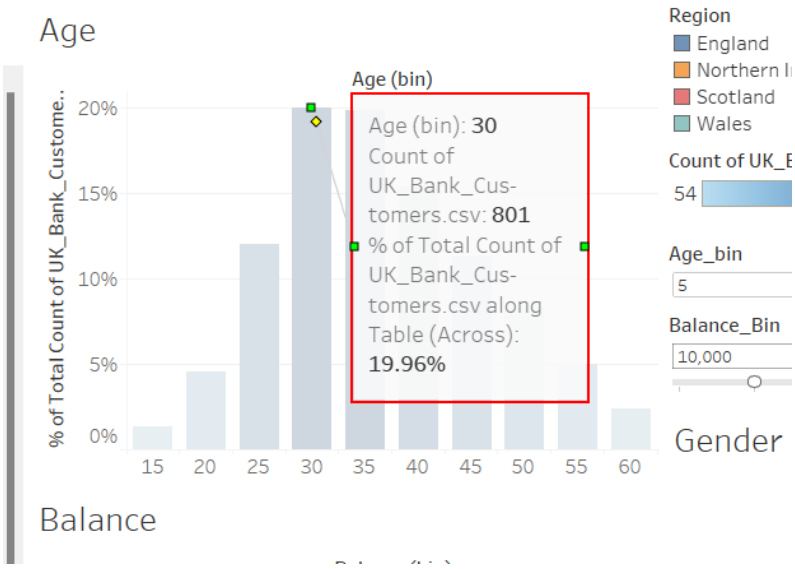
Creating a Storyline:



Adding Annotation on the storyline



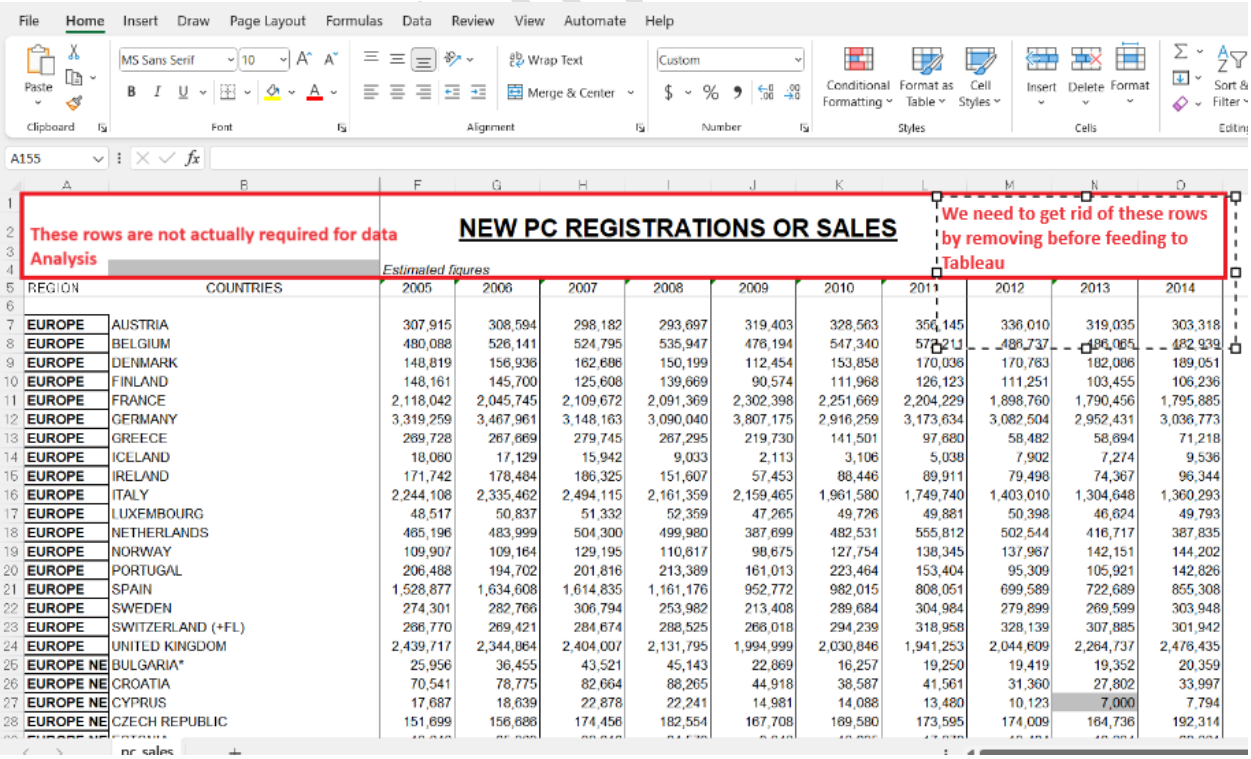
Below is how the Annotate looks like



Advanced Data Preparation:

Download Dataset here:

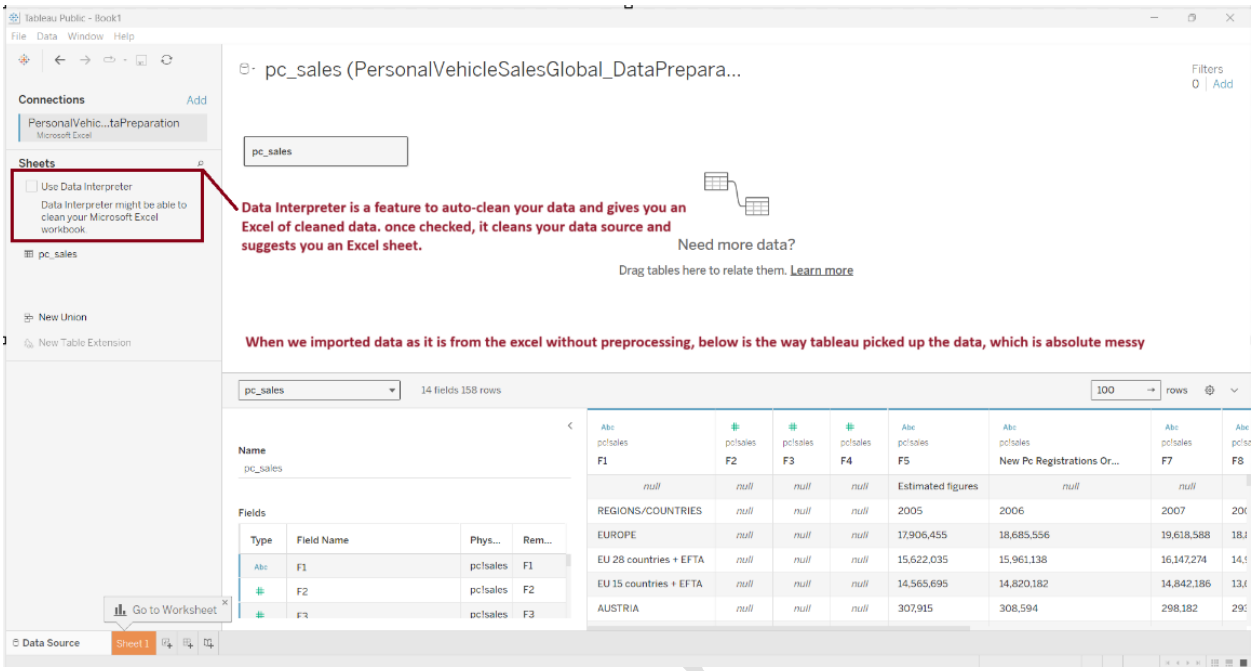
[Tableau_Data_Sets/PersonalVehicleSalesGlobal_DataPreparation.xlsx at main · MohanPulibanti/Tableau_Data_Sets (github.com)]

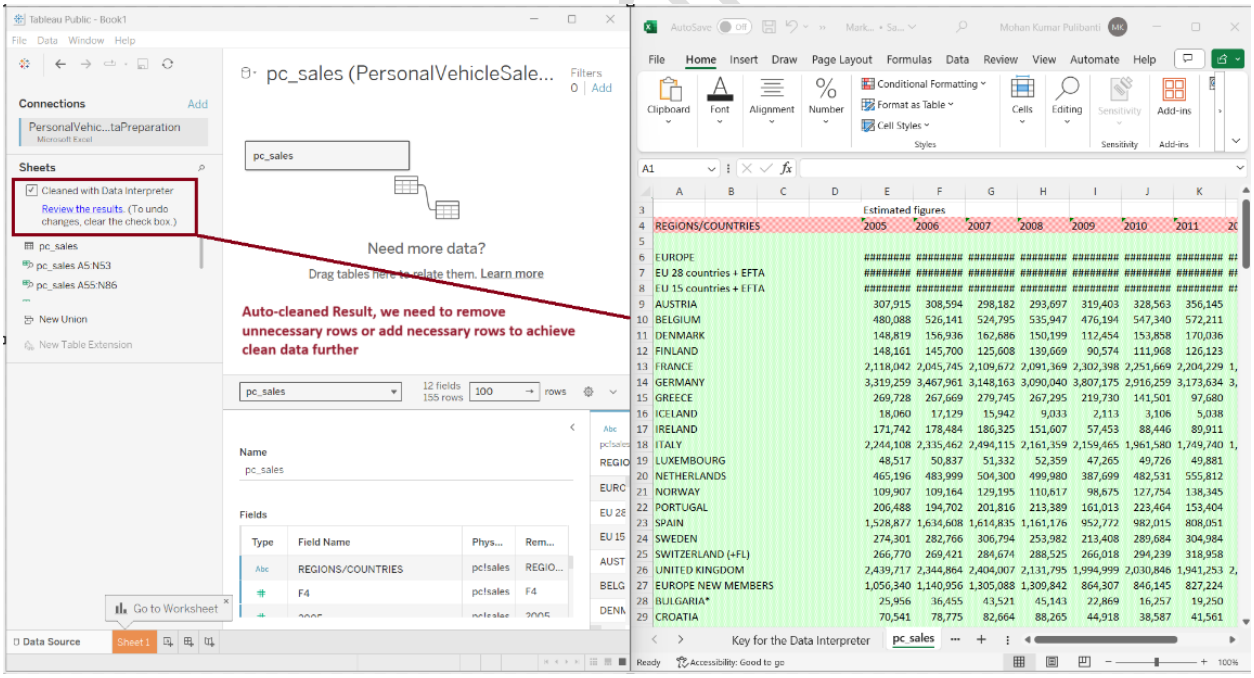


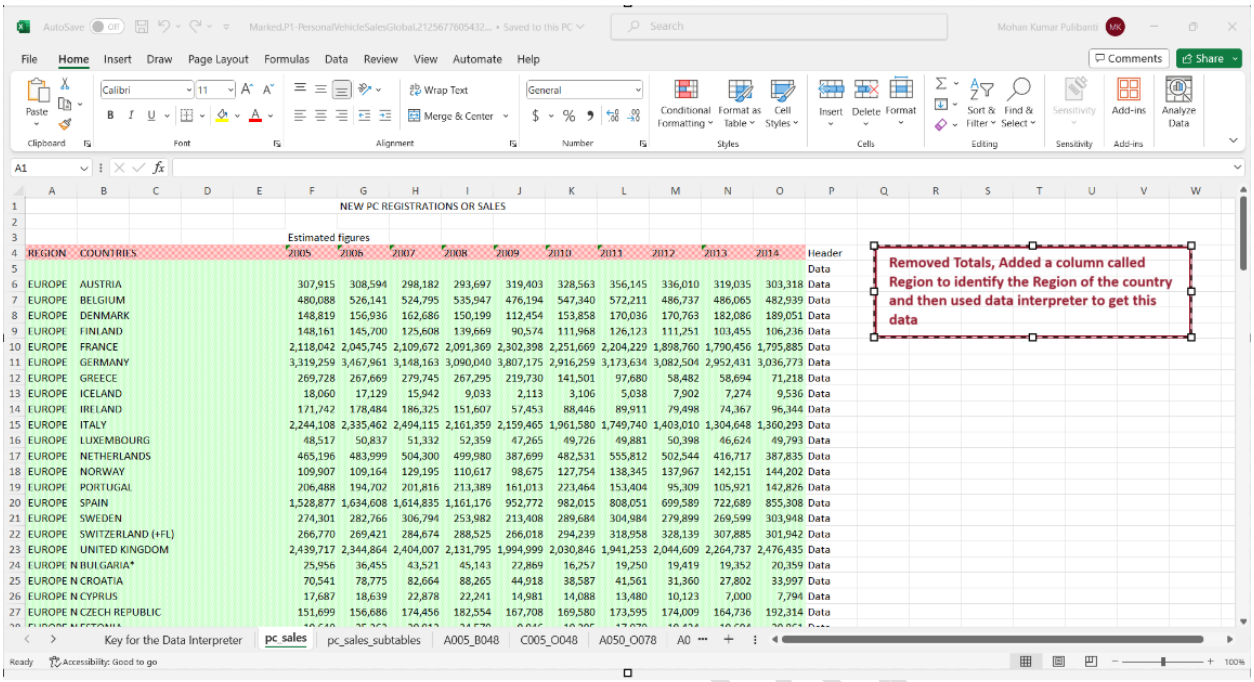
REGION	COUNTRIES	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
EUROPE	AUSTRIA	307,915	308,594	298,182	293,697	319,403	328,563	356,145	336,010	319,035	303,318
EUROPE	BELGIUM	480,088	526,141	524,795	535,947	476,194	547,340	572,211	486,737	486,065	482,939
EUROPE	DENMARK	148,819	156,936	162,686	150,199	112,454	153,858	170,036	170,763	182,086	189,051
EUROPE	FINLAND	148,161	145,700	125,608	139,669	90,574	111,968	126,123	111,251	103,455	106,236
EUROPE	FRANCE	2,118,042	2,045,745	2,109,672	2,091,369	2,302,398	2,251,669	2,204,229	1,898,760	1,790,456	1,795,885
EUROPE	GERMANY	3,319,259	3,467,961	3,148,163	3,090,040	3,807,175	2,916,259	3,173,634	3,082,504	2,952,431	3,036,773
EUROPE	GREECE	269,728	267,669	279,745	267,295	219,730	141,501	97,680	58,482	58,694	71,218
EUROPE	ICELAND	18,060	17,129	15,942	9,033	2,113	3,106	5,038	7,902	7,274	9,536
EUROPE	IRELAND	171,742	178,484	186,325	151,607	57,453	88,446	89,911	79,498	74,367	96,344
EUROPE	ITALY	2,244,108	2,335,462	2,494,115	2,161,359	2,159,465	1,961,580	1,749,740	1,403,010	1,304,648	1,360,293
EUROPE	LUXEMBOURG	48,517	50,837	51,332	52,359	47,265	49,726	49,881	50,398	46,624	49,793
EUROPE	NETHERLANDS	465,196	483,999	504,300	499,980	387,699	482,531	555,812	502,544	416,717	387,835
EUROPE	NORWAY	109,907	109,164	129,195	110,617	98,675	127,754	138,345	137,967	142,151	144,202
EUROPE	PORTUGAL	206,488	194,702	201,816	213,389	161,013	223,464	153,404	95,309	105,921	142,826
EUROPE	SPAIN	1,528,877	1,634,608	1,614,835	1,161,176	952,772	982,015	808,051	699,589	722,689	855,308
EUROPE	SWEDEN	274,301	282,766	306,794	253,982	213,408	289,684	304,984	279,899	269,599	303,948
EUROPE	SWITZERLAND (+FL)	266,770	269,421	284,674	288,525	268,018	294,239	318,958	328,139	307,885	301,942
EUROPE	UNITED KINGDOM	2,439,717	2,344,864	2,404,007	2,131,795	1,994,999	2,030,846	1,941,253	2,044,609	2,264,737	2,476,435
EUROPE NE	BULGARIA*	25,956	36,455	43,521	45,143	22,869	16,257	19,250	19,419	19,352	20,359
EUROPE NE	CROATIA	70,541	78,775	82,664	88,265	44,918	38,587	41,561	31,360	27,802	33,997
EUROPE NE	CYPRUS	17,687	18,639	22,878	22,241	14,981	14,088	13,480	10,123	7,000	7,794
EUROPE NE	CZECH REPUBLIC	151,699	156,686	174,456	182,554	167,708	169,580	173,595	174,009	164,736	192,314

Mohan Kumar Pulibanti

If the data has Empty rows or Totals, we need to get rid of all those rows as part of data preprocessing before visualization.



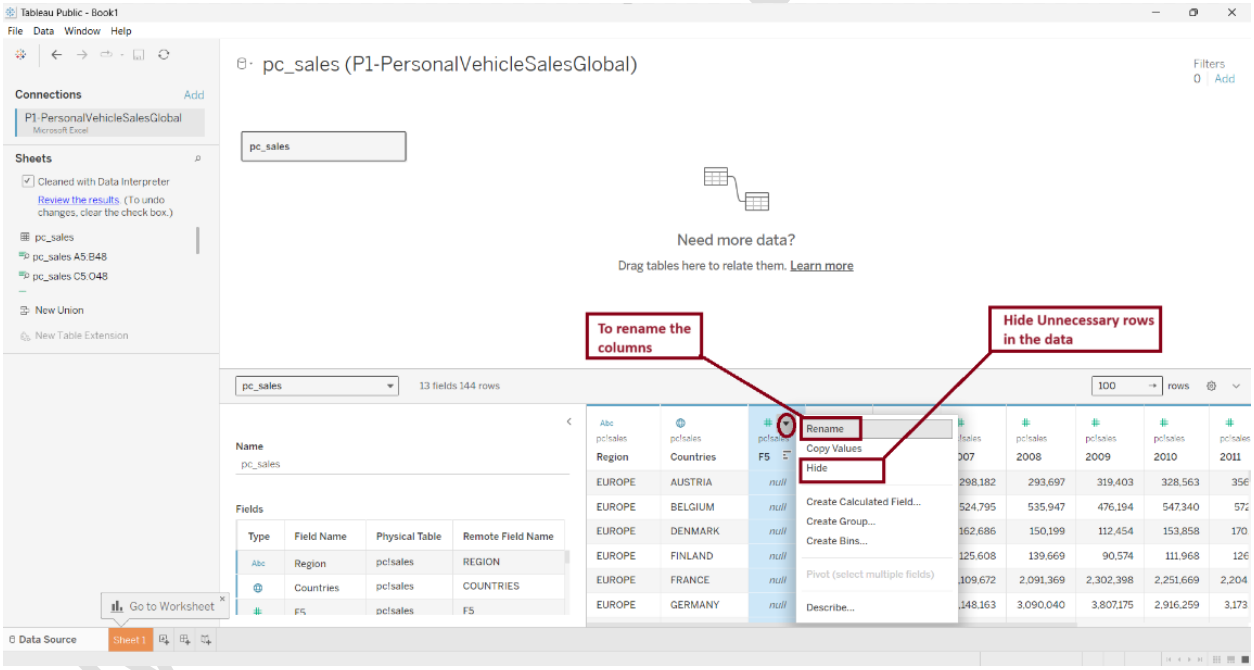




NEW PC REGISTRATIONS OR SALES

REGION	COUNTRIES	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Header
		Estimated figures										Data
EUROPE	AUSTRIA	307,915	308,594	298,182	293,697	319,403	328,563	356,145	336,010	319,035	303,318	Data
EUROPE	BELGIUM	480,088	526,141	524,795	535,947	476,194	547,340	572,211	486,737	486,065	482,939	Data
EUROPE	DENMARK	148,819	156,936	162,686	150,199	112,454	153,858	170,036	170,763	182,086	189,051	Data
EUROPE	FINLAND	148,161	145,700	125,608	139,669	90,574	111,968	126,123	111,251	103,455	106,236	Data
EUROPE	FRANCE	2,118,042	2,045,745	2,109,672	2,091,369	2,302,398	2,251,669	2,204,229	1,898,760	1,790,456	1,795,885	Data
EUROPE	GERMANY	3,319,259	3,467,961	3,148,163	3,090,040	3,807,175	2,916,259	3,173,634	3,082,504	2,952,431	3,036,773	Data
EUROPE	GREECE	269,728	267,669	279,745	267,295	219,730	141,501	97,680	58,482	58,694	71,218	Data
EUROPE	ICELAND	18,060	17,129	15,942	9,033	2,113	3,106	5,038	7,902	7,274	9,536	Data
EUROPE	IRELAND	171,742	178,484	186,325	151,607	57,453	88,446	89,911	79,498	74,367	96,344	Data
EUROPE	ITALY	2,244,108	2,335,462	2,494,115	2,161,359	2,159,465	1,961,580	1,749,740	1,403,010	1,304,648	1,360,293	Data
EUROPE	LUXEMBOURG	48,517	50,837	51,332	52,359	47,265	49,726	49,881	50,398	46,624	49,793	Data
EUROPE	NETHERLANDS	465,196	483,999	504,300	499,980	387,699	482,531	555,812	502,544	416,717	387,835	Data
EUROPE	NORWAY	109,907	109,164	129,195	110,617	98,675	127,754	138,345	137,967	142,151	144,202	Data
EUROPE	PORTUGAL	206,488	194,702	201,816	213,389	161,013	223,464	153,404	95,309	105,921	142,826	Data
EUROPE	SPAIN	1,528,877	1,634,608	1,614,835	1,161,176	952,772	982,015	808,051	699,589	722,689	855,308	Data
EUROPE	SWEDEN	274,301	282,766	306,794	253,982	213,408	289,684	304,984	279,899	269,599	303,948	Data
EUROPE	SWITZERLAND (+FL)	266,770	269,421	284,674	288,525	266,018	294,239	318,958	328,139	307,885	301,942	Data
EUROPE	UNITED KINGDOM	2,439,717	2,344,864	2,404,007	2,131,795	1,994,999	2,030,846	1,941,253	2,044,609	2,264,737	2,476,435	Data
EUROPE N	BULGARIA*	25,956	36,455	43,521	45,143	22,869	16,257	19,250	19,419	19,352	20,359	Data
EUROPE N	CROATIA	70,541	78,775	82,664	88,265	44,918	38,587	41,561	31,360	27,802	33,997	Data
EUROPE N	CYPRUS	17,687	18,639	22,878	22,241	14,981	14,088	13,480	10,123	7,000	7,794	Data
EUROPE N	CZECH REPUBLIC	151,699	156,686	174,456	182,554	167,708	169,580	173,595	174,009	164,736	192,314	Data

Removed Totals, Added a column called Region to identify the Region of the country and then used data interpreter to get this data



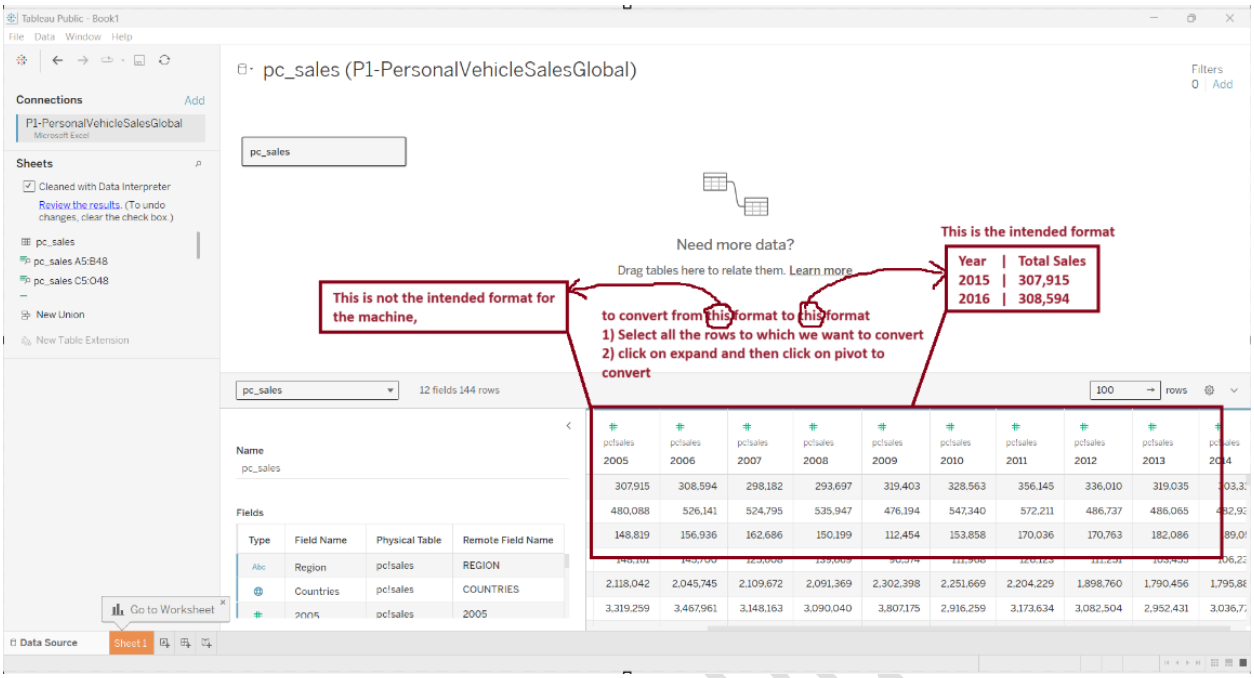


Tableau Public - Book1
File Data Window Help
pc_sales (P1-PersonalVehicleSalesGlobal)
Filters
0 Add
Connections
Add
P1-PersonalVehicleSalesGlobal
Microsoft Excel
Sheets
Cleaned with Data Interpreter
Review the results. (To undo changes, clear the check box.)
pc_sales
pc_sales A5:B48
pc_sales C5:O48
New Union
New Table Extension
pc_sales
Need more data?
Drag tables here to relate them. Learn more
This is the intended format
Year | Total Sales
2015 | 307,915
2016 | 308,594
This is not the intended format for the machine,
to convert from this format to this format
1) Select all the rows to which we want to convert
2) click on expand and then click on pivot to convert
12 fields 144 rows
100 rows
Name
pc_sales
Fields
Type
Field Name
Physical Table
Remote Field Name
Region
pclsales
REGION
Countries
pclsales
COUNTRIES
2005
pclsales
2005
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Go to Worksheet
Data Source
Sheet 1

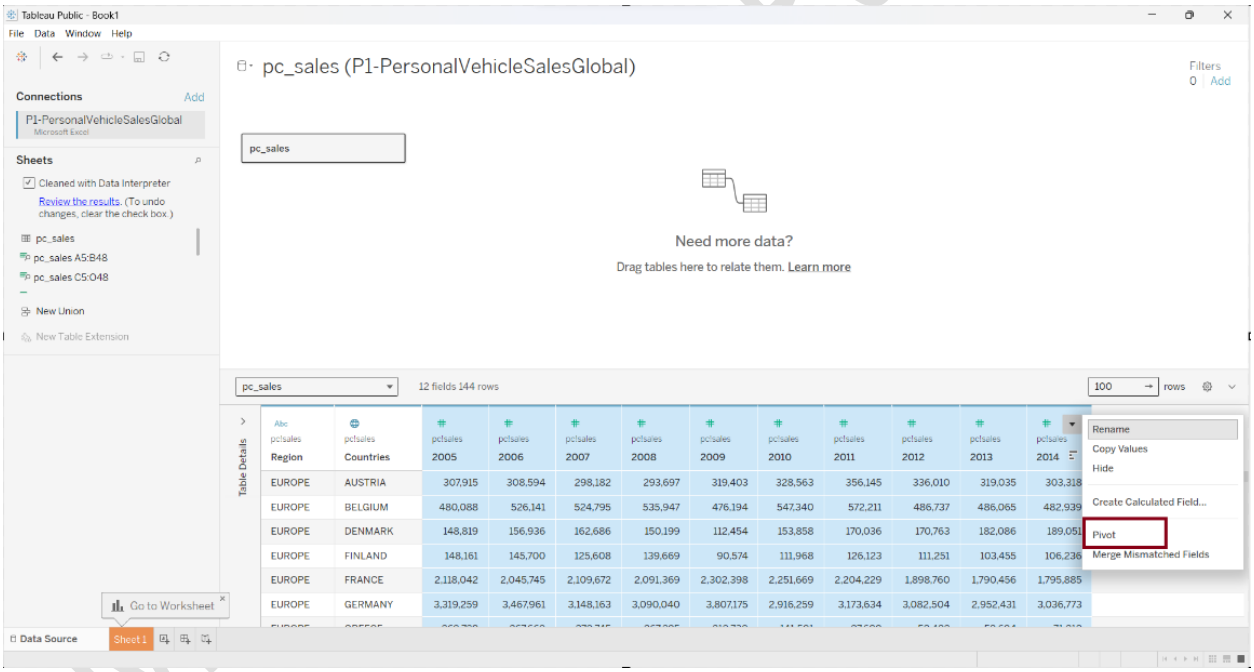


Tableau Public - Book1
File Data Window Help
pc_sales (P1-PersonalVehicleSalesGlobal)
Filters
0 Add
Connections
Add
P1-PersonalVehicleSalesGlobal
Microsoft Excel
Sheets
Cleaned with Data Interpreter
Review the results. (To undo changes, clear the check box.)
pc_sales
pc_sales A5:B48
pc_sales C5:O48
New Union
New Table Extension
pc_sales
Need more data?
Drag tables here to relate them. Learn more
12 fields 144 rows
100 rows
Table Details
Region
Countries
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
EUROPE AUSTRIA 307,915 308,594 298,182 293,697 319,403 328,563 356,145 336,010 319,035 303,318
EUROPE BELGIUM 480,088 526,141 524,795 535,947 476,194 547,340 572,211 486,737 486,065 482,939
EUROPE DENMARK 148,819 156,936 162,686 150,199 112,454 153,858 170,036 170,763 182,086 189,051
EUROPE FINLAND 148,161 145,700 125,608 139,669 90,574 111,968 126,123 111,251 103,455 106,236
EUROPE FRANCE 2,118,042 2,045,745 2,109,672 2,091,369 2,302,398 2,251,669 2,204,229 1,898,760 1,790,456 1,795,885
EUROPE GERMANY 3,319,259 3,467,961 3,148,163 3,090,040 3,807,175 2,916,259 3,173,634 3,082,504 2,952,431 3,036,773
Rename
Copy Values
Hide
Create Calculated Field...
Pivot
Merge Mismatched Fields
Go to Worksheet
Data Source
Sheet 1

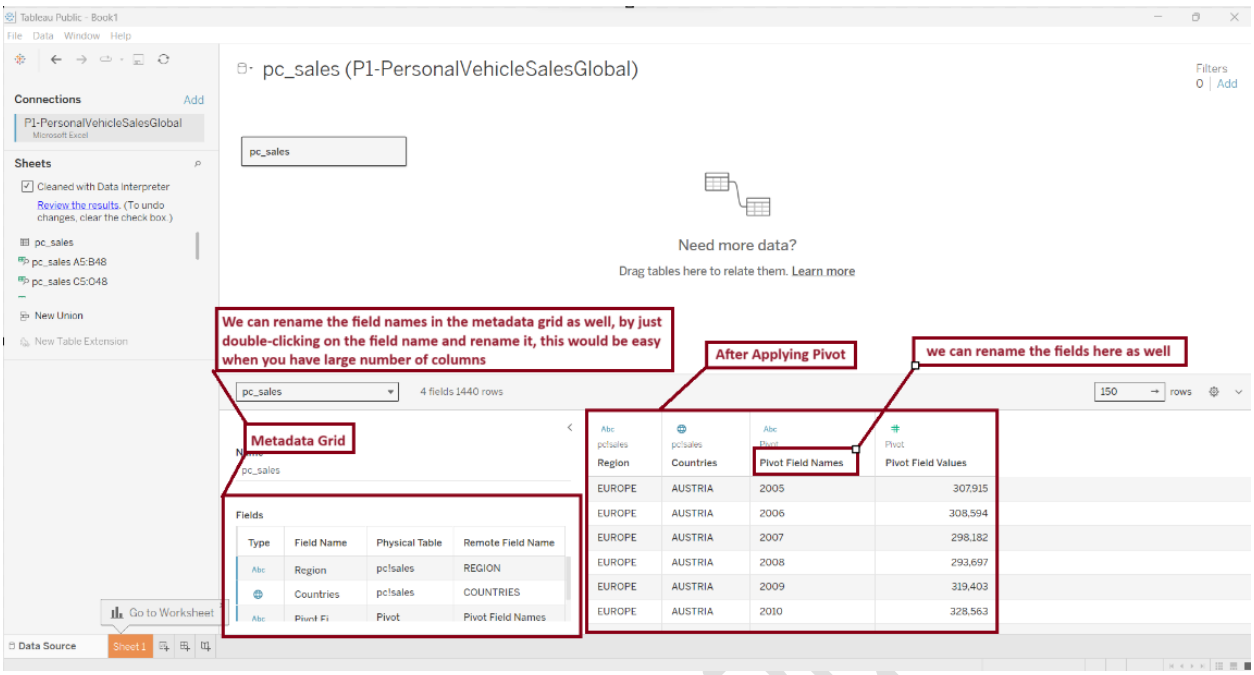


Tableau Public - Book1
File Data Window Help
pc_sales (P1-PersonalVehicleSalesGlobal)
Connections Add
P1-PersonalVehicleSalesGlobal
Sheets
Cleaned with Data Interpreter
Review the results. (To undo changes, clear the check box.)
pc_sales
pc_sales A5:B48
pc_sales C5:O48
New Union
New Table Extension
Need more data?
Drag tables here to relate them. Learn more
We can rename the field names in the metadata grid as well, by just double-clicking on the field name and rename it, this would be easy when you have large number of columns
After Applying Pivot
we can rename the fields here as well
4 fields 1440 rows
Metadata Grid
Fields
Type Field Name Physical Table Remote Field Name
Region pclsales REGION
Countries pclsales COUNTRIES
Pivot Field Names
Pivot Field Values
EUROPE AUSTRIA 2005 307,915
EUROPE AUSTRIA 2006 308,594
EUROPE AUSTRIA 2007 298,182
EUROPE AUSTRIA 2008 293,697
EUROPE AUSTRIA 2009 319,403
EUROPE AUSTRIA 2010 328,563
Go to Worksheet
Data Source
Sheet 1

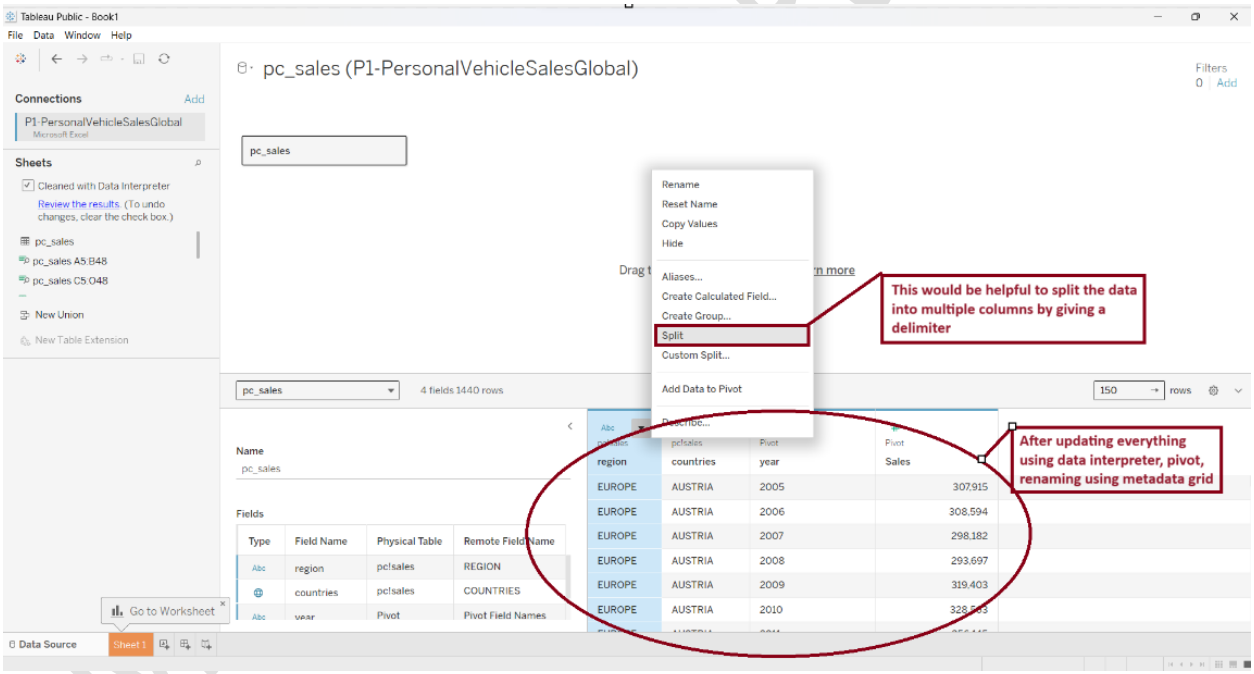
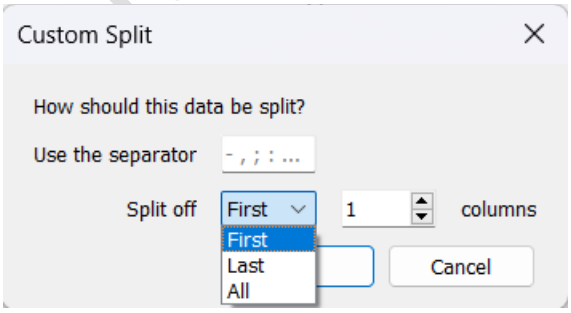
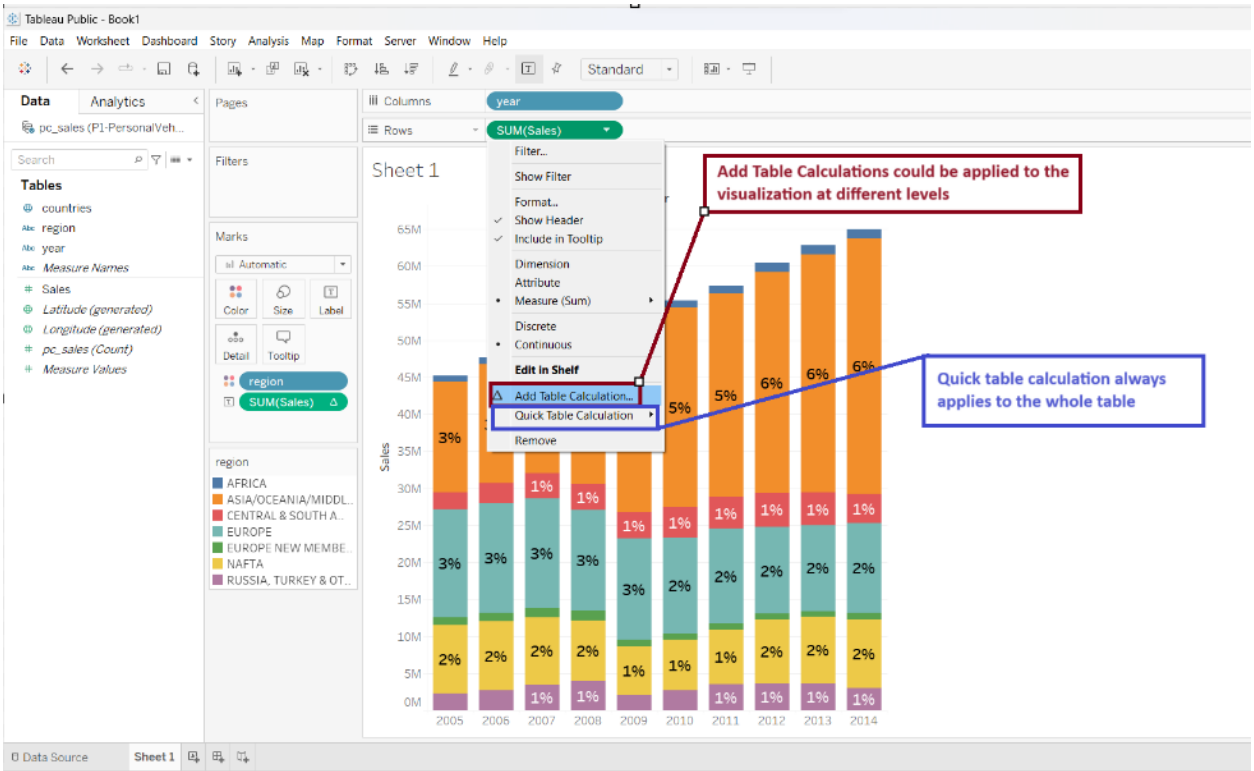
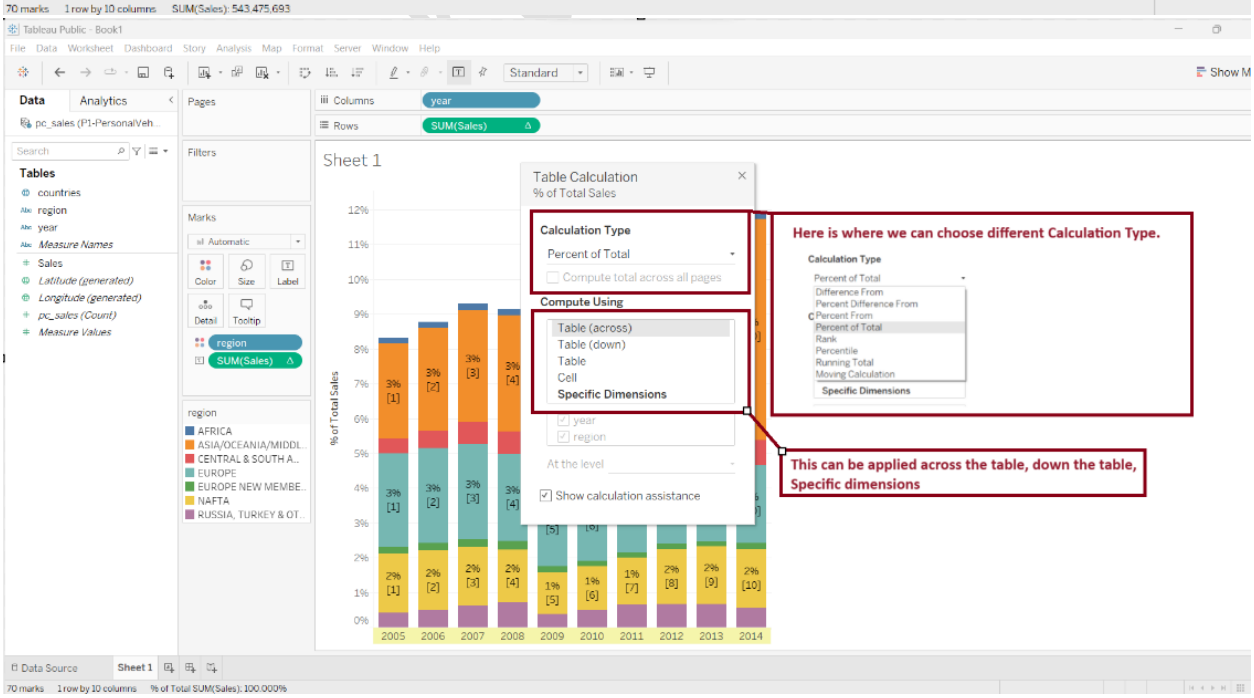


Tableau Public - Book1
File Data Window Help
pc_sales (P1-PersonalVehicleSalesGlobal)
Rename
Reset Name
Copy Values
Hide
Aliases...
Create Calculated Field...
Create Group...
Split
Custom Split...
Add Data to Pivot
This would be helpful to split the data into multiple columns by giving a delimiter
4 fields 1440 rows
region countries year Sales
EUROPE AUSTRIA 2005 307,915
EUROPE AUSTRIA 2006 308,594
EUROPE AUSTRIA 2007 298,182
EUROPE AUSTRIA 2008 293,697
EUROPE AUSTRIA 2009 319,403
EUROPE AUSTRIA 2010
After updating everything using data interpreter, pivot, renaming using metadata grid
Go to Worksheet
Data Source
Sheet 1

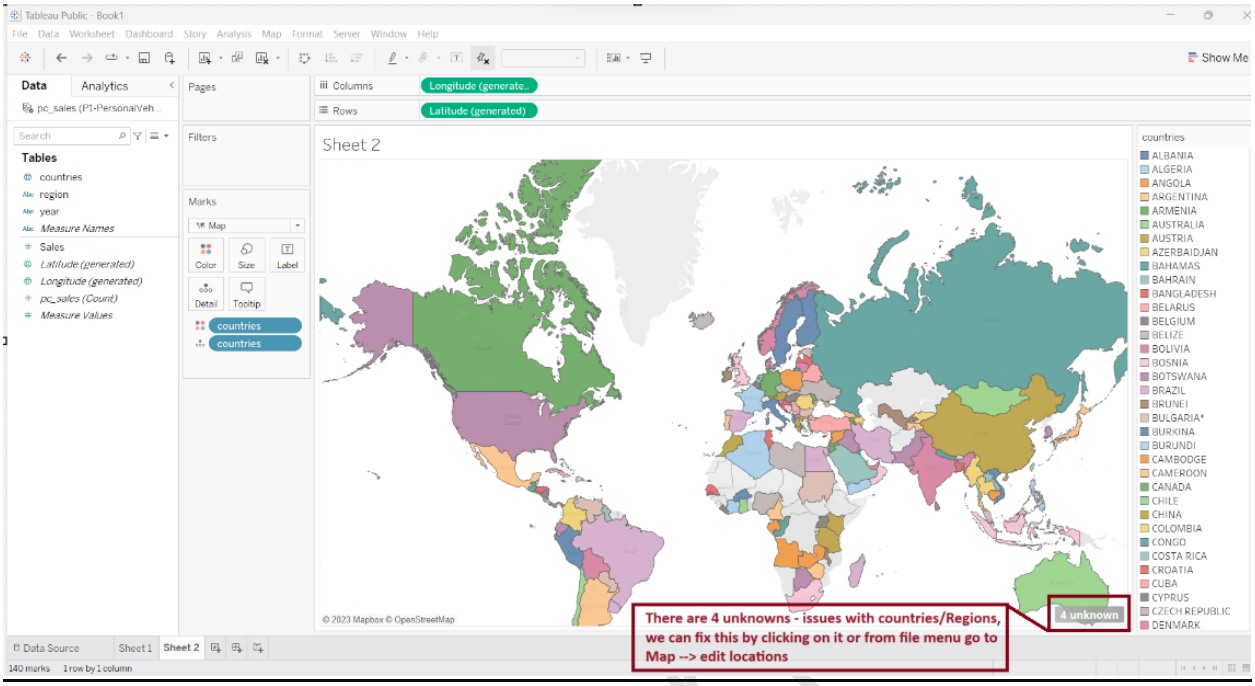

Custom Split
How should this data be split?
Use the separator
Split off
First
1
columns
First
Last
All
Cancel

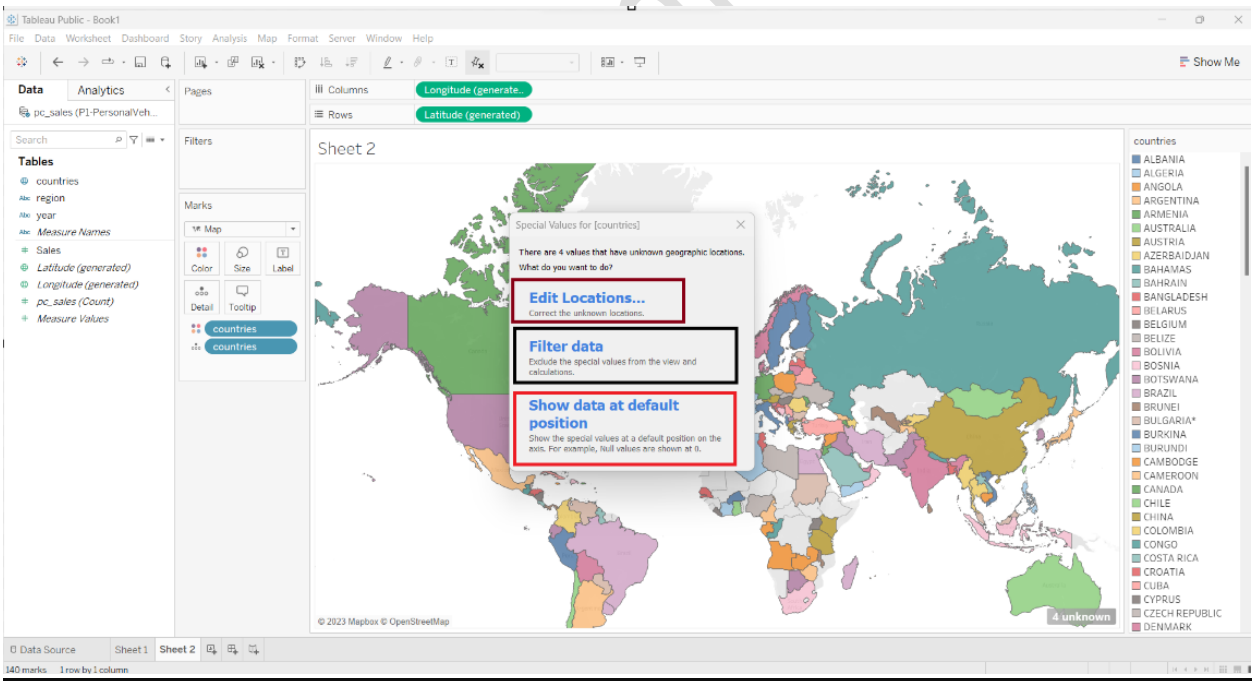
Separator → Specifies what is the delimiter
How to Split off
First → splits the first words. (it divides the number of columns specified from the first)
Last → Splits the last words (it divides the number of columns specified from the last)
All → Divides all the columns





Fixing Geographical Data Errors





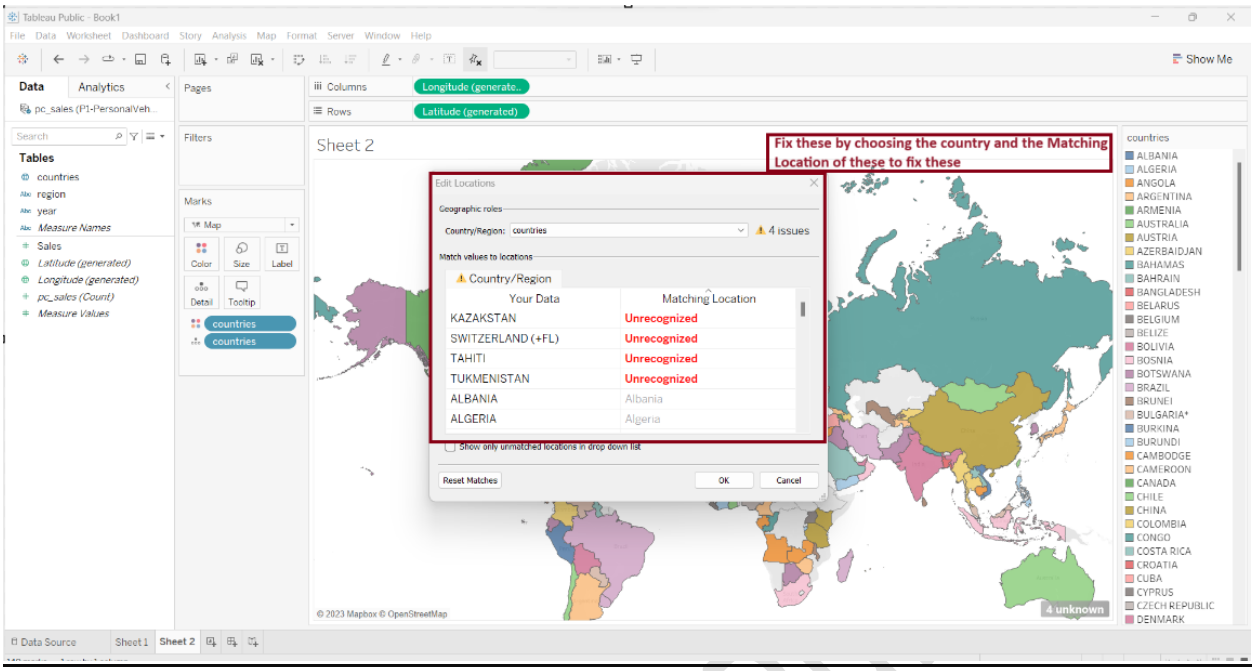


Tableau Public - Book1
Fix these by choosing the country and the Matching Location of these to fix these
Edit Locations
Geographic roles
Country/Region: countries
4 issues
Match values to locations
Country/Region
Your Data
Matching Location
KAZAKSTAN
Unrecognized
SWITZERLAND (+FL)
Unrecognized
TAHITI
Unrecognized
TUKMENISTAN
Unrecognized
ALBANIA
Albania
ALGERIA
Algeria
Reset Matches
OK
Cancel
4 unknown

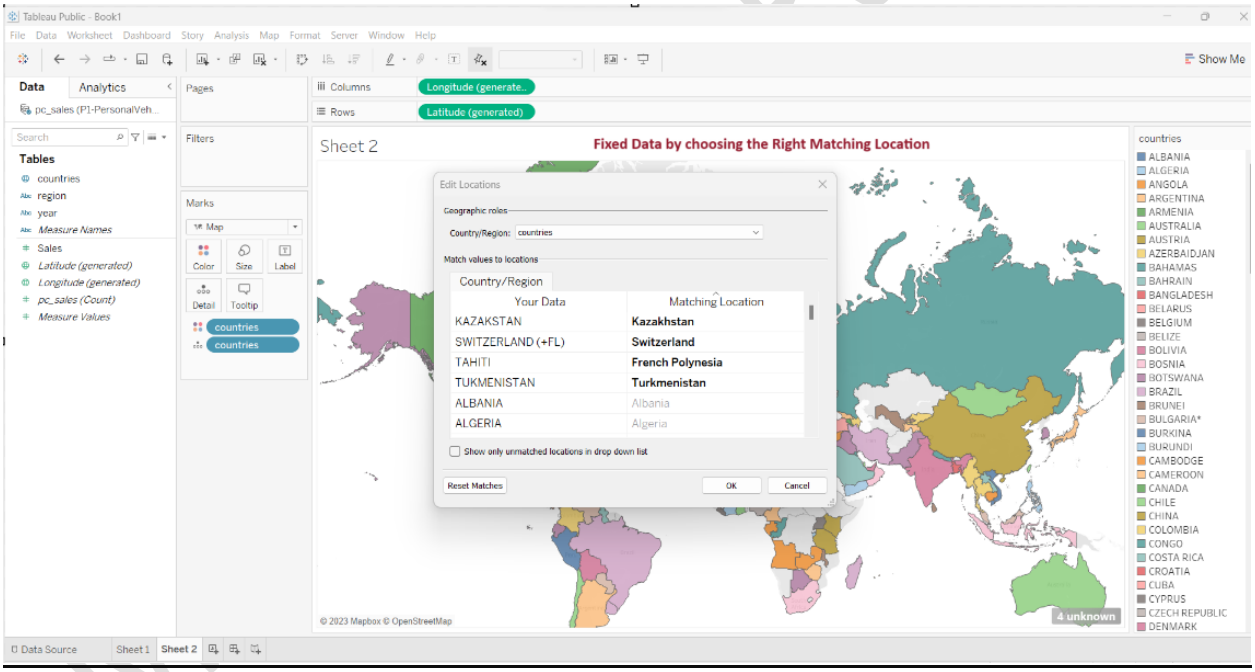
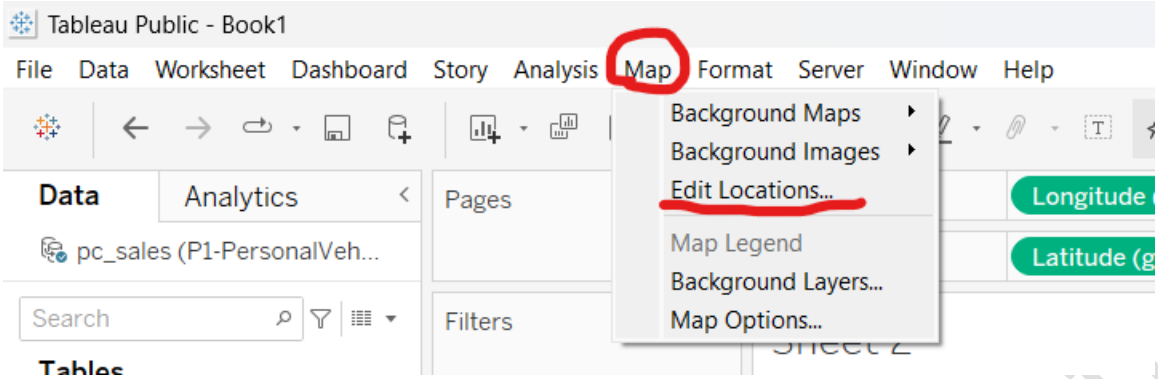


Tableau Public - Book1
Fixed Data by choosing the Right Matching Location
Edit Locations
Geographic roles
Country/Region: countries
Match values to locations
Country/Region
Your Data
Matching Location
KAZAKSTAN
Kazakhstan
SWITZERLAND (+FL)
Switzerland
TAHITI
French Polynesia
TUKMENISTAN
Turkmenistan
ALBANIA
Albania
ALGERIA
Algeria
Reset Matches
OK
Cancel
4 unknown

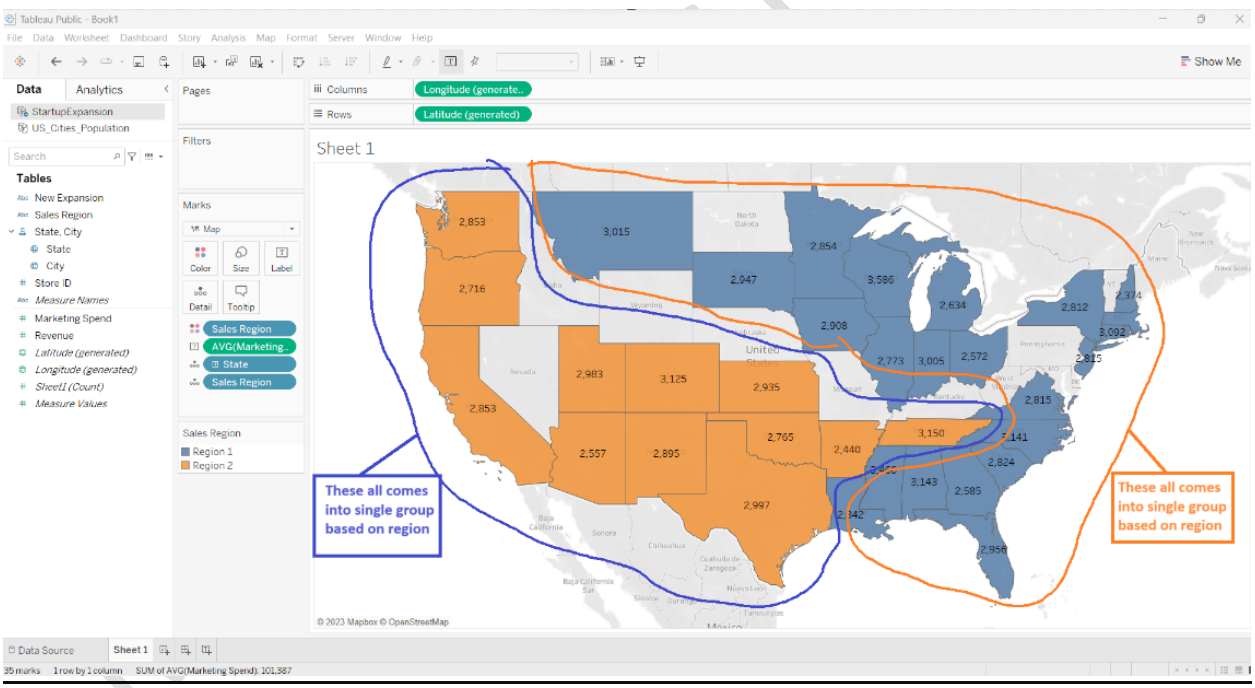


Grouping (or) Creating Custom Territories

Download Dataset Here:

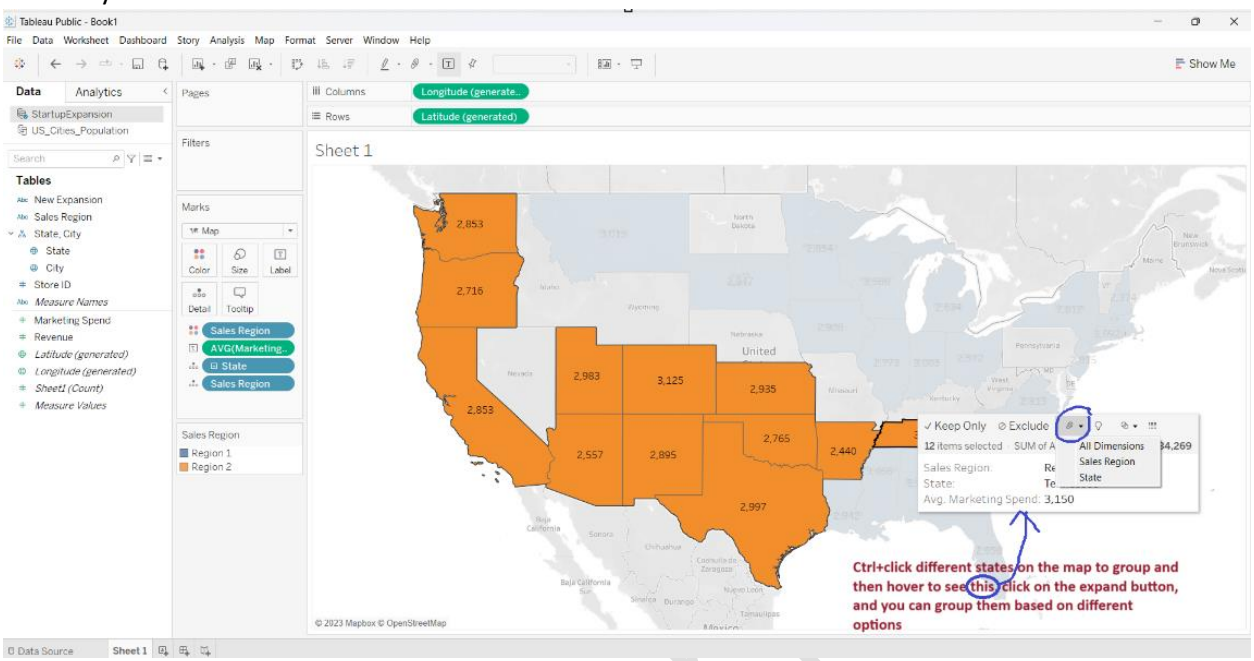
[Tableau_Data_Sets/StartupExpansion.xlsx at main · MohanPulibanti/Tableau_Data_Sets (github.com)]

[Tableau_Data_Sets/US_Cities_Population.csv at main · MohanPulibanti/Tableau_Data_Sets (github.com)]

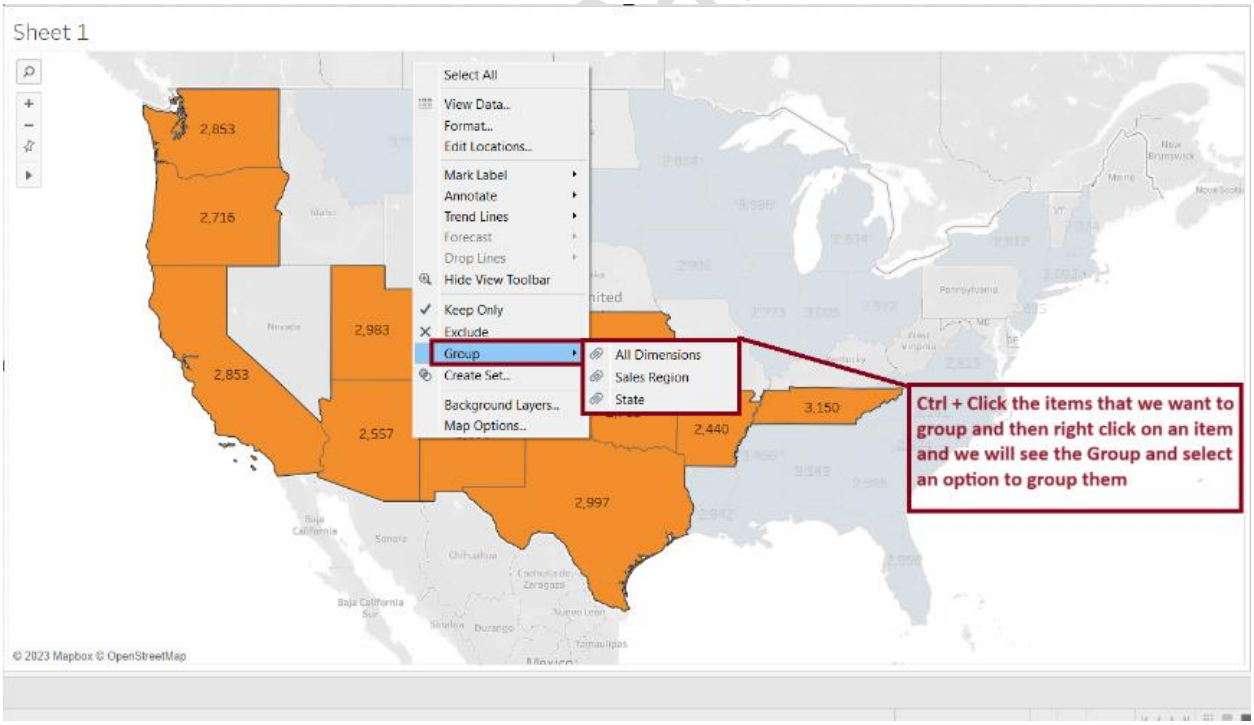


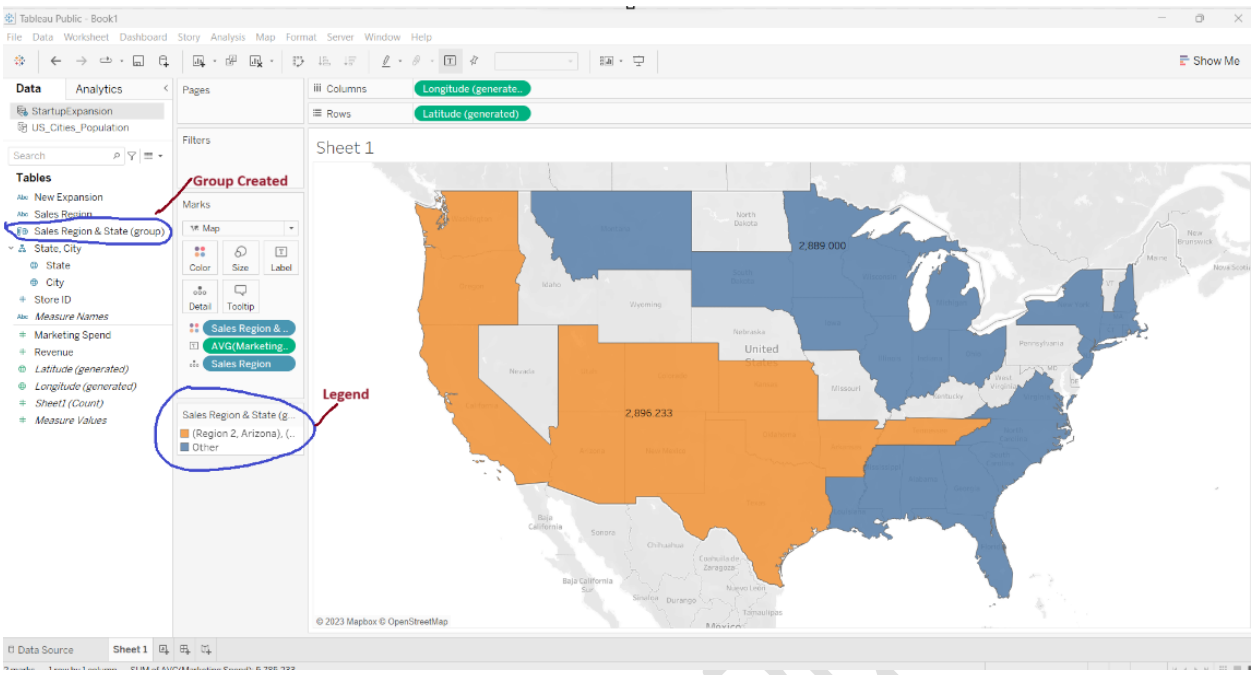
Grouping can be done in two ways:

First Way:

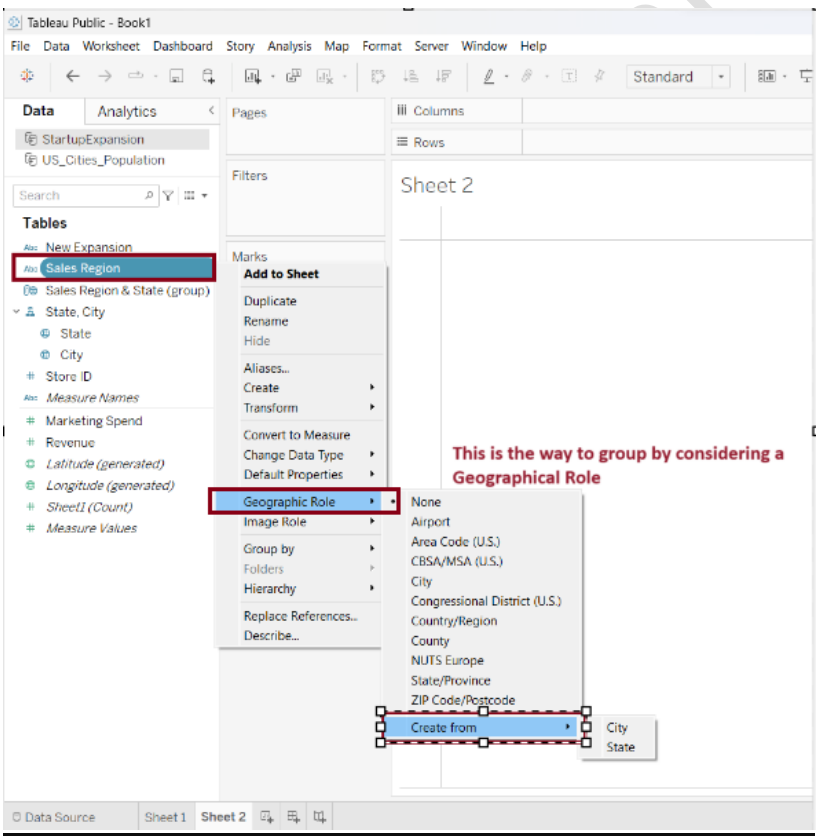


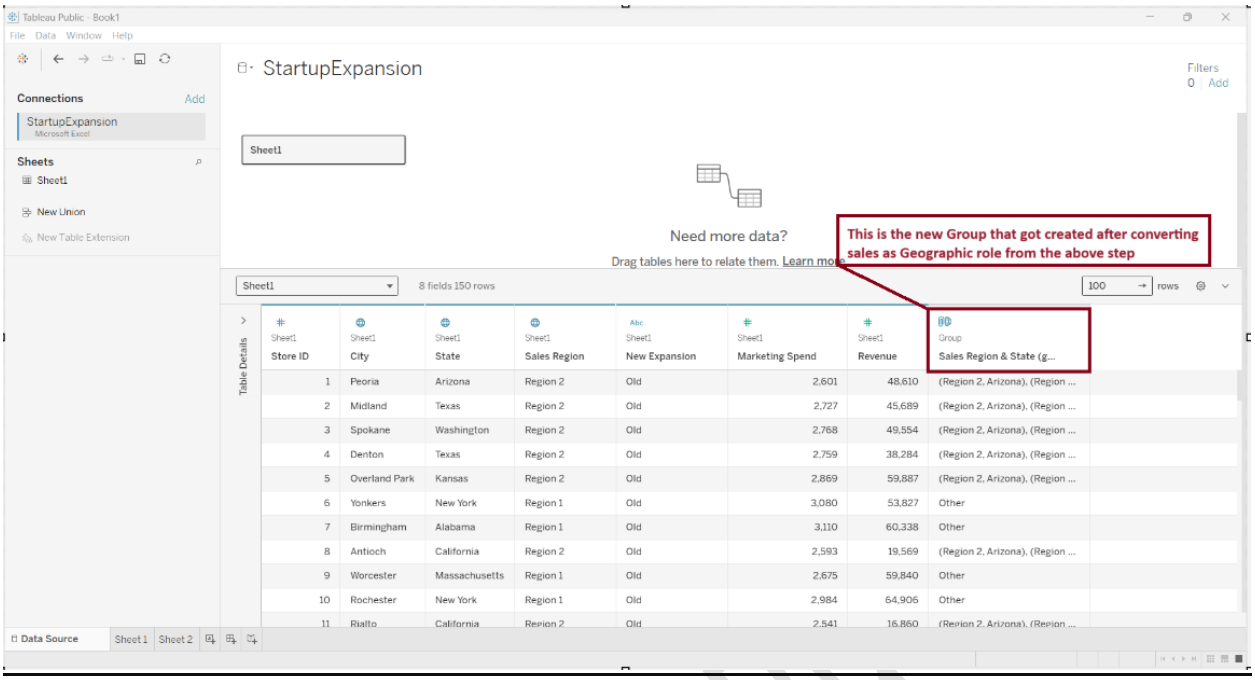
Second Way:



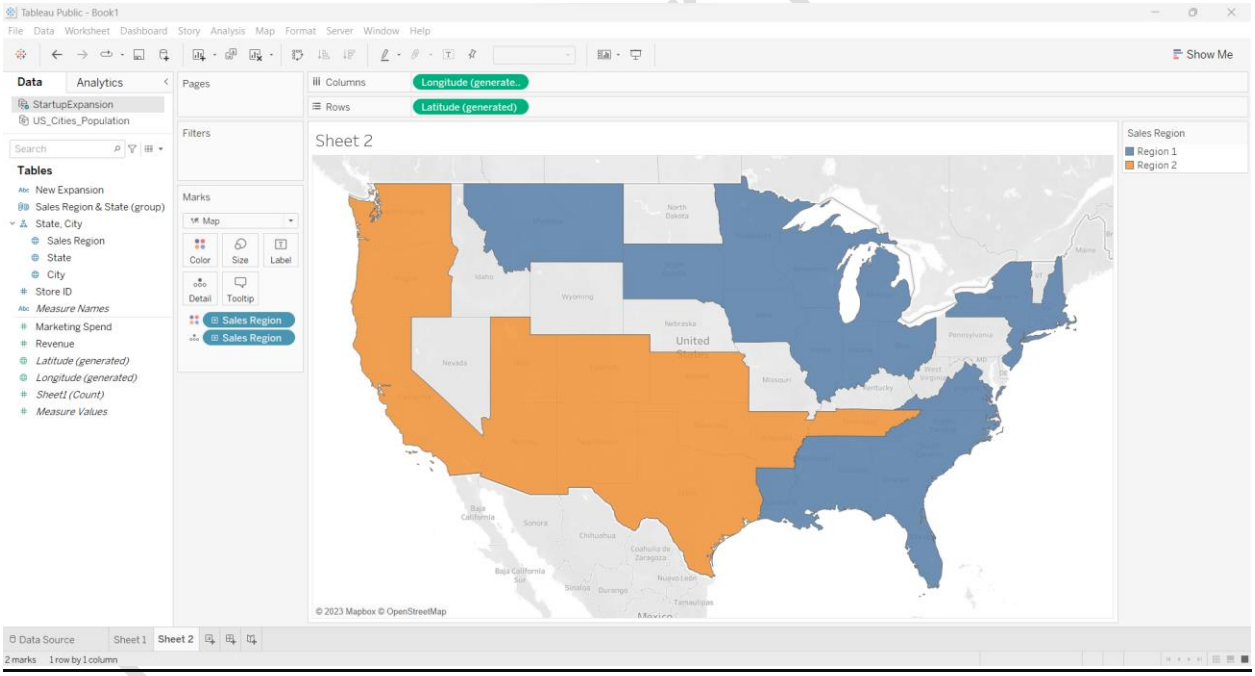


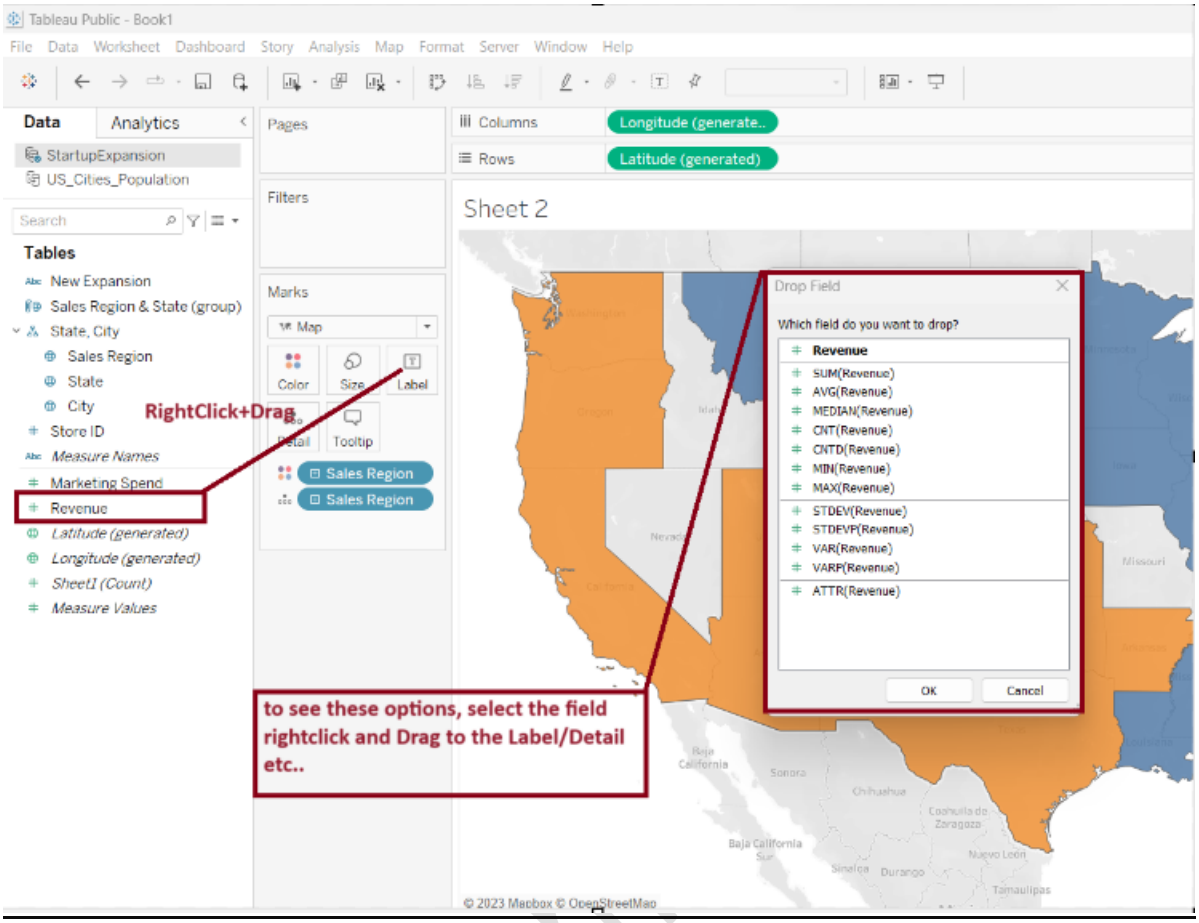
Custom Territories via Geographic Roles



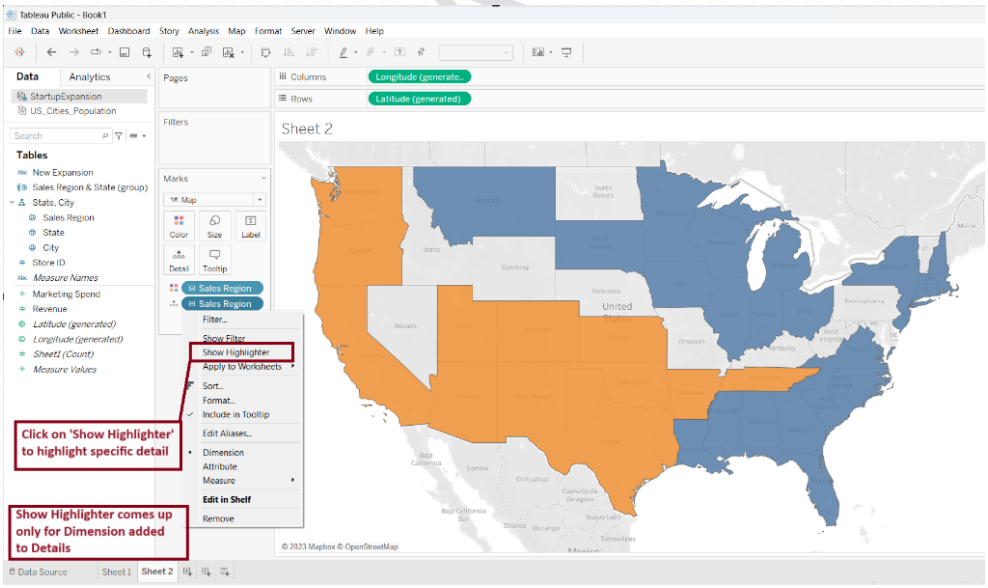


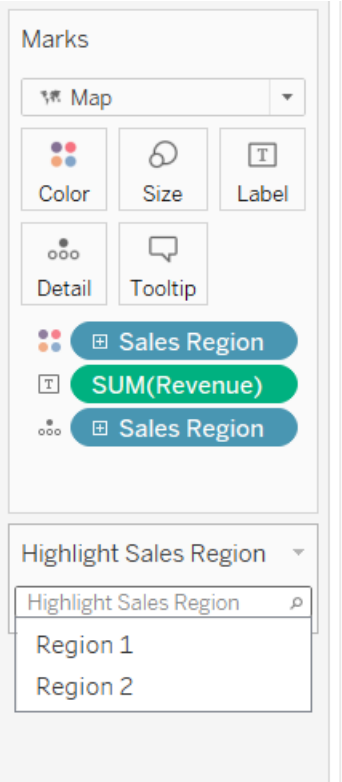
By Dragging the sales region to the Working tab, we see as below.

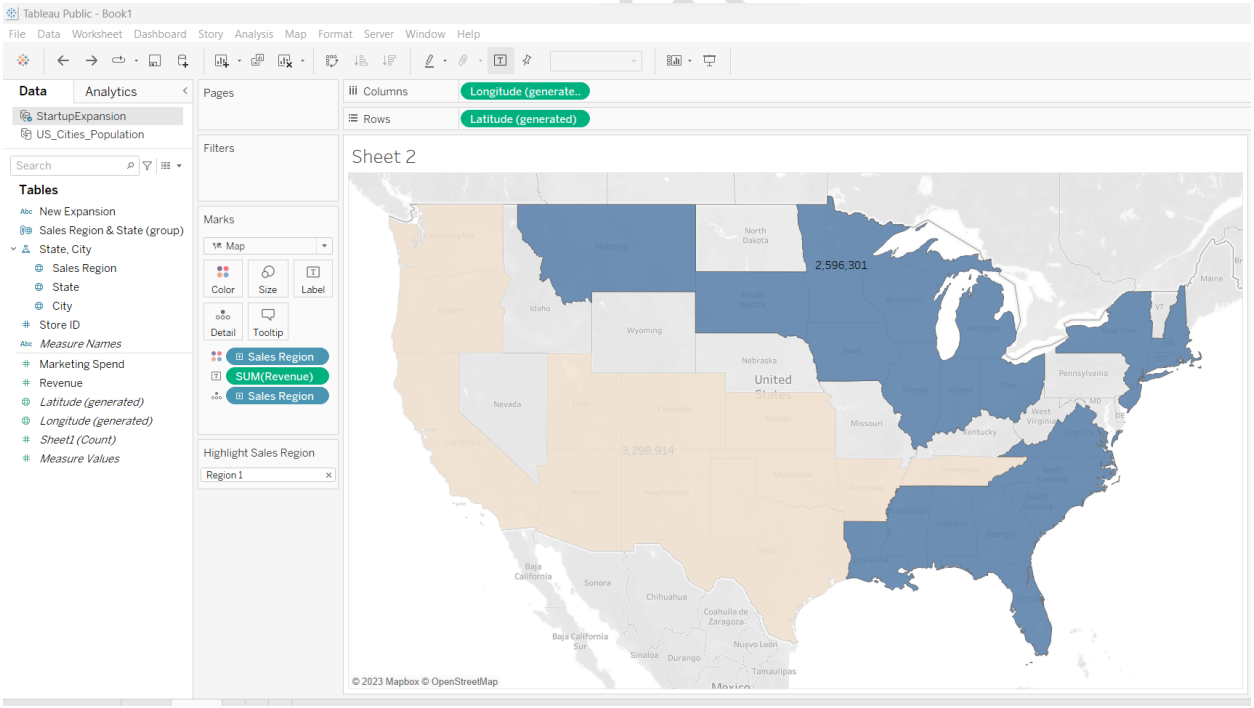




Adding a Highlighter

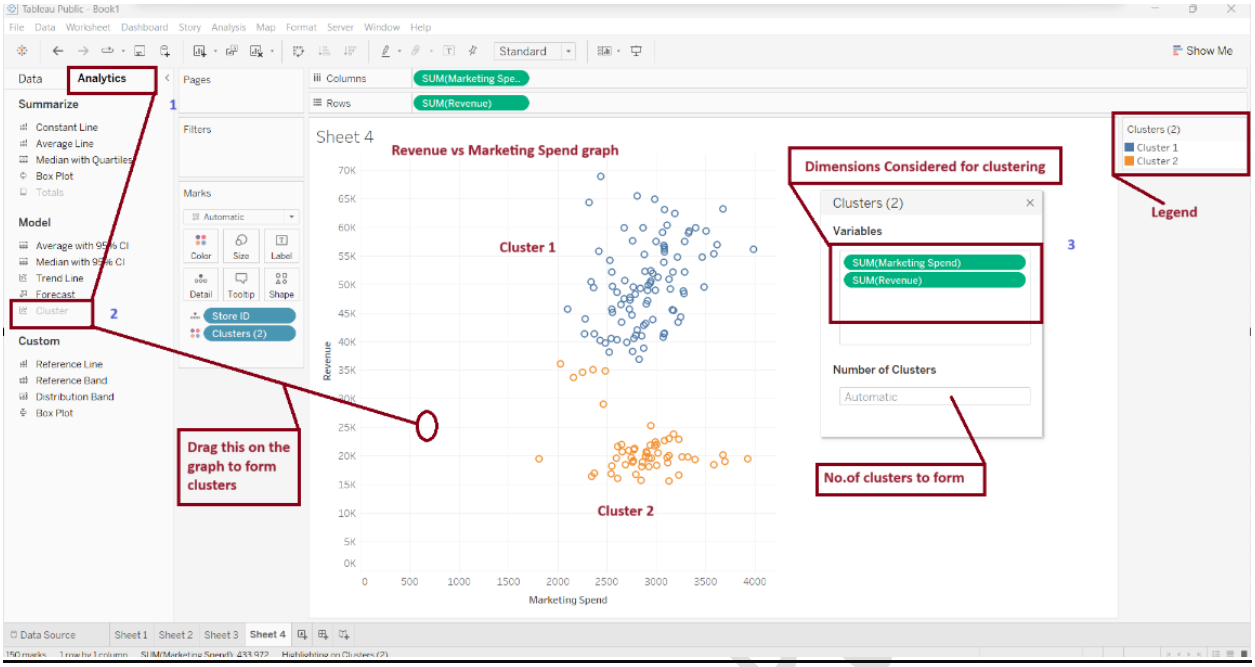




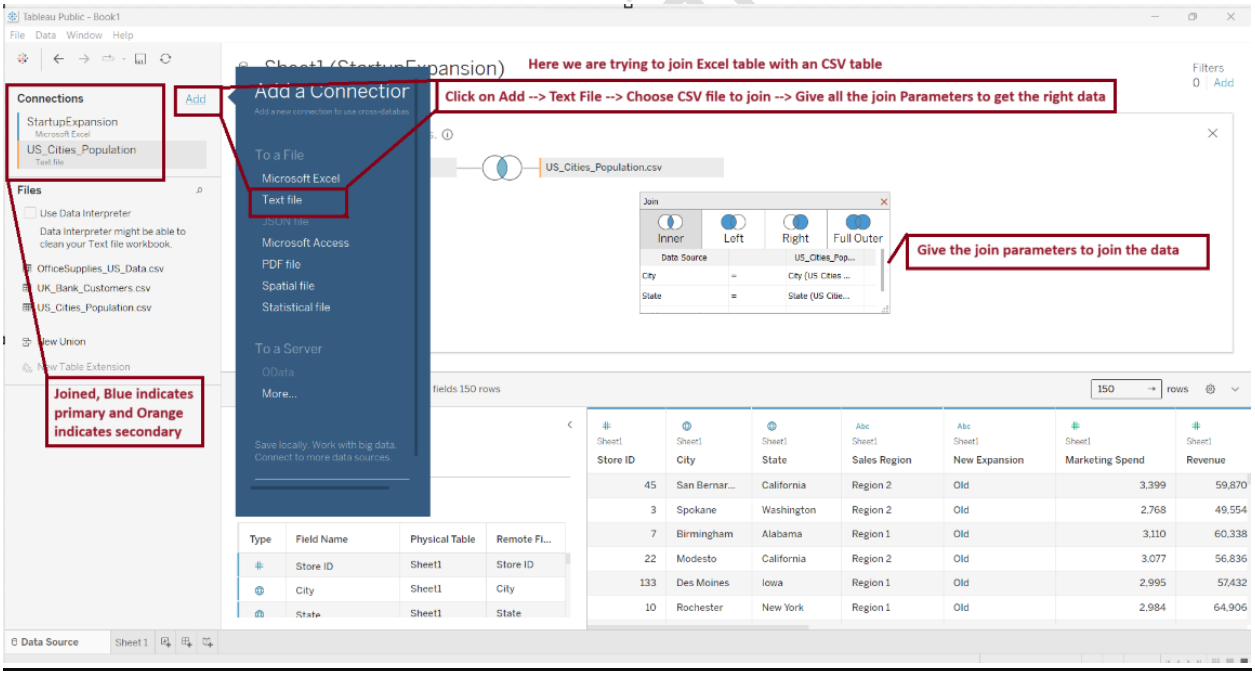


Mohan Kumar Pulibanti

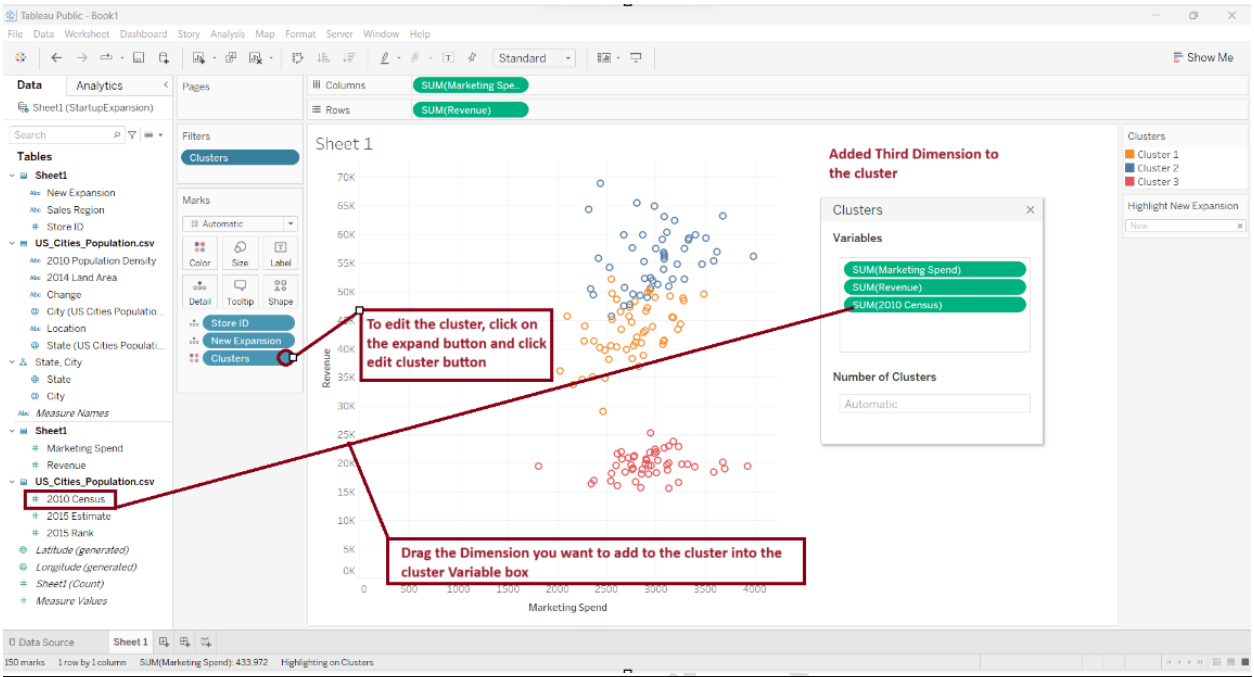
Clustering in Tableau



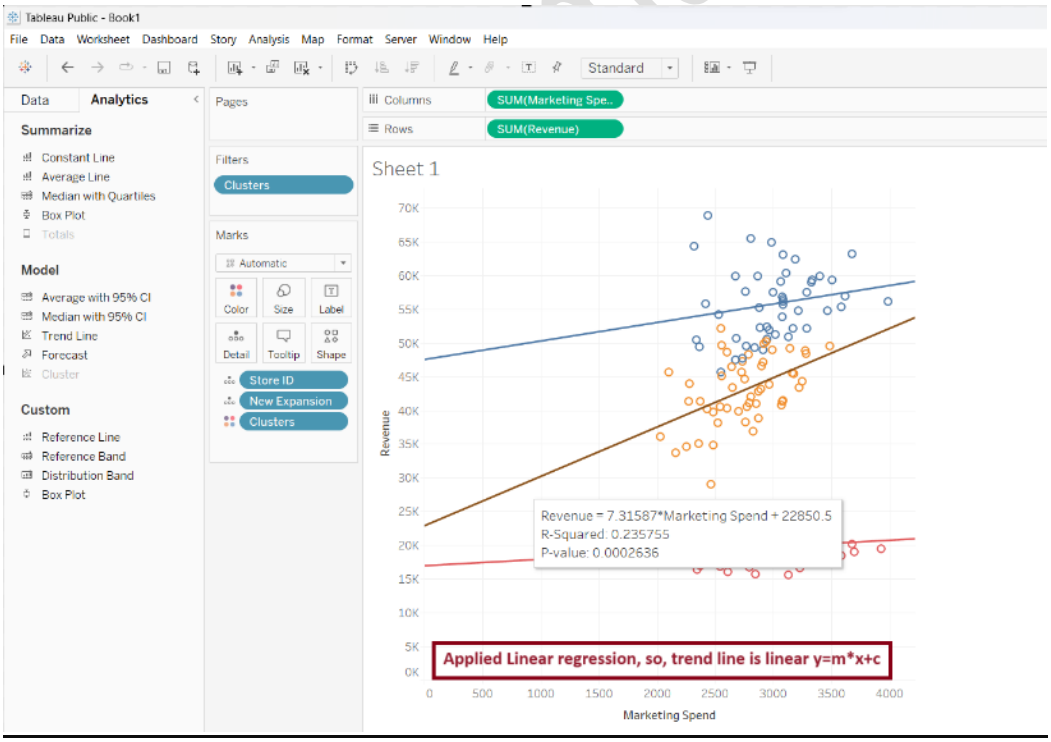
Cross-Database joins

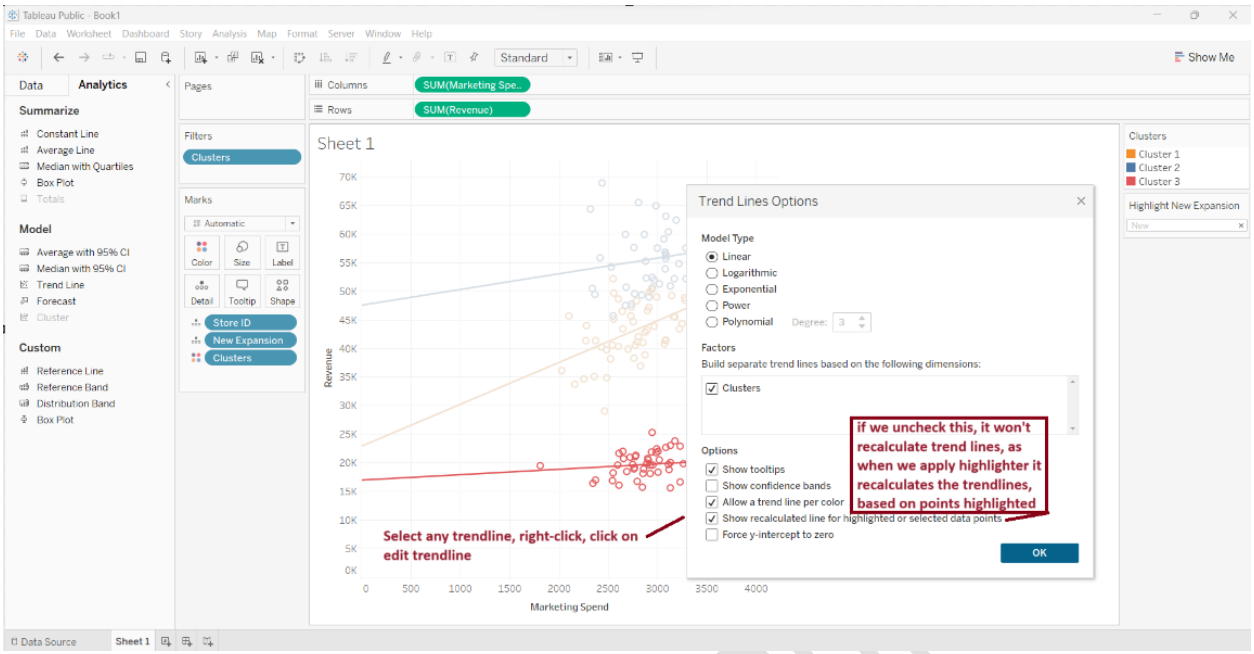


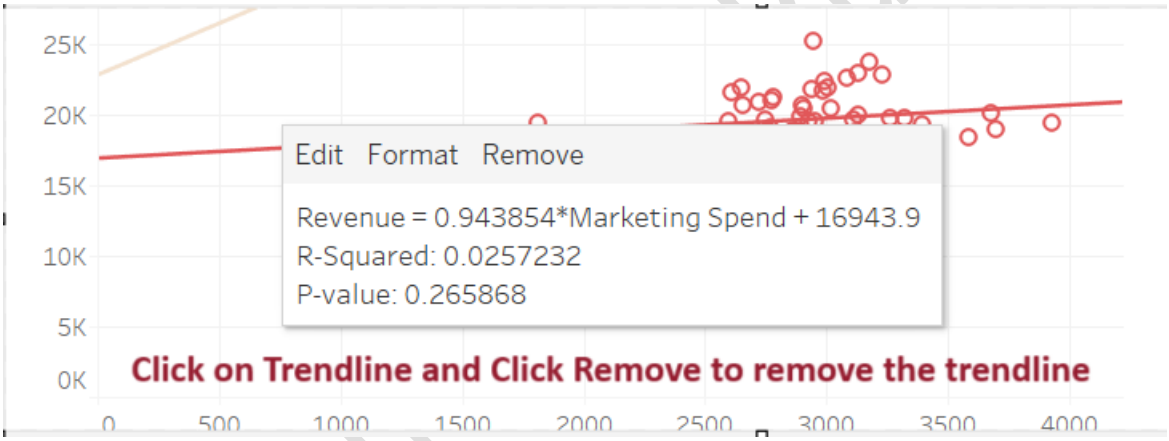
Modeling With Clusters:



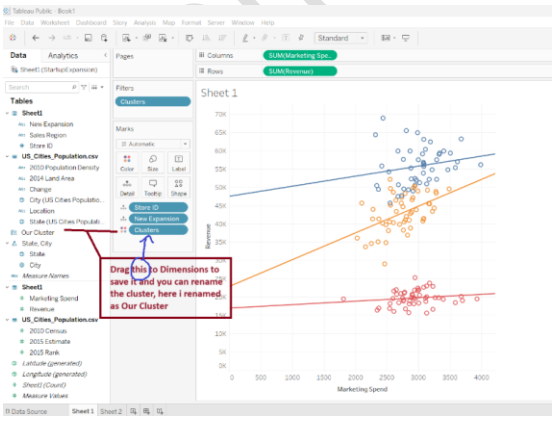
Adding trendlines



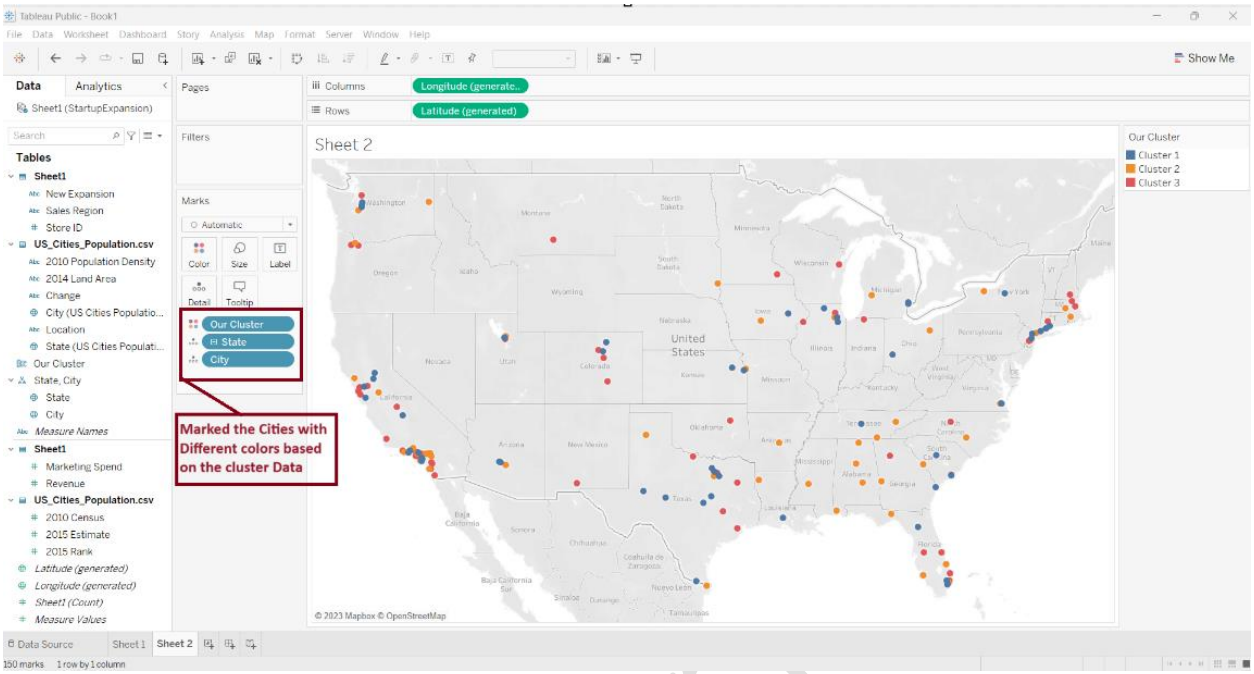




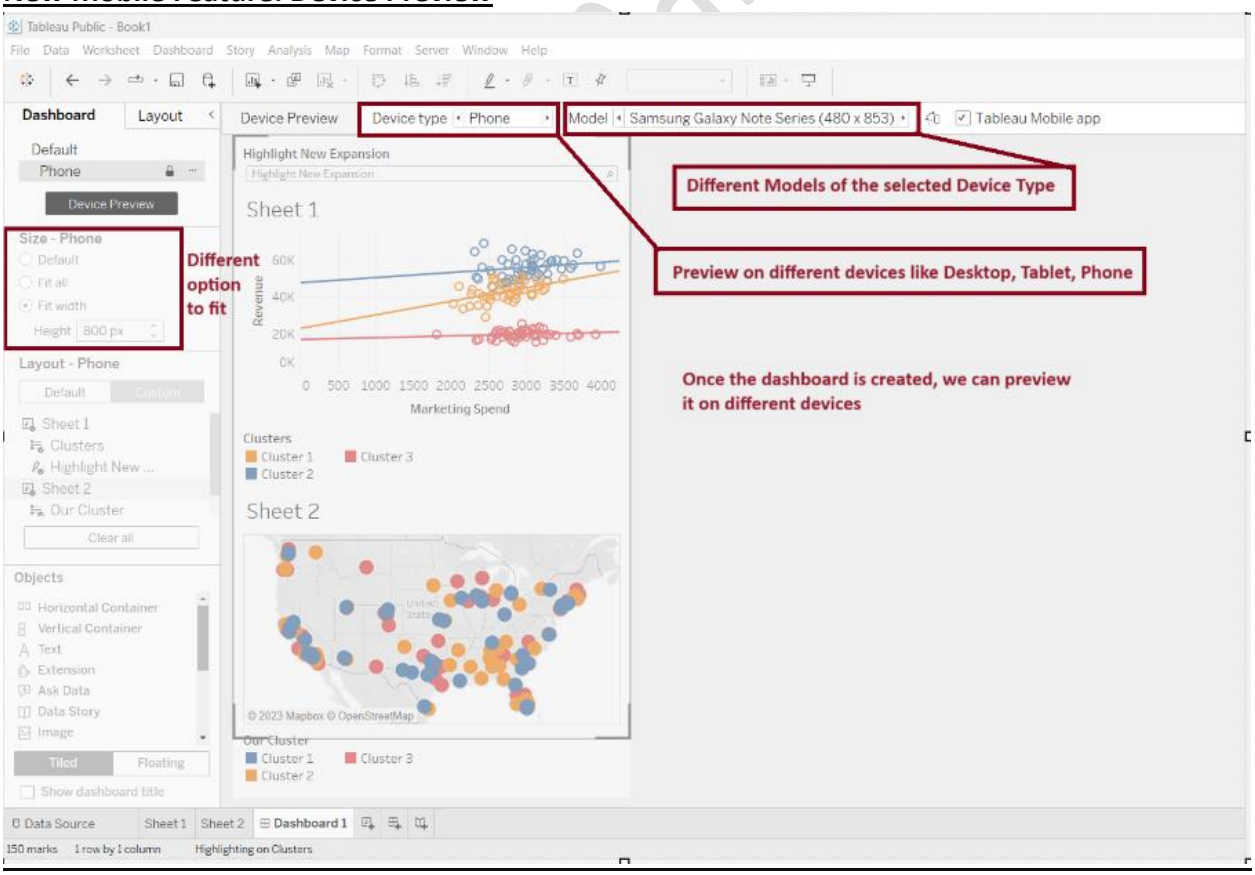
Saving Clusters



Using the Saved Clusters in Another sheet, the Saved Cluster would be Just like a Group



New Mobile Feature: Device Preview

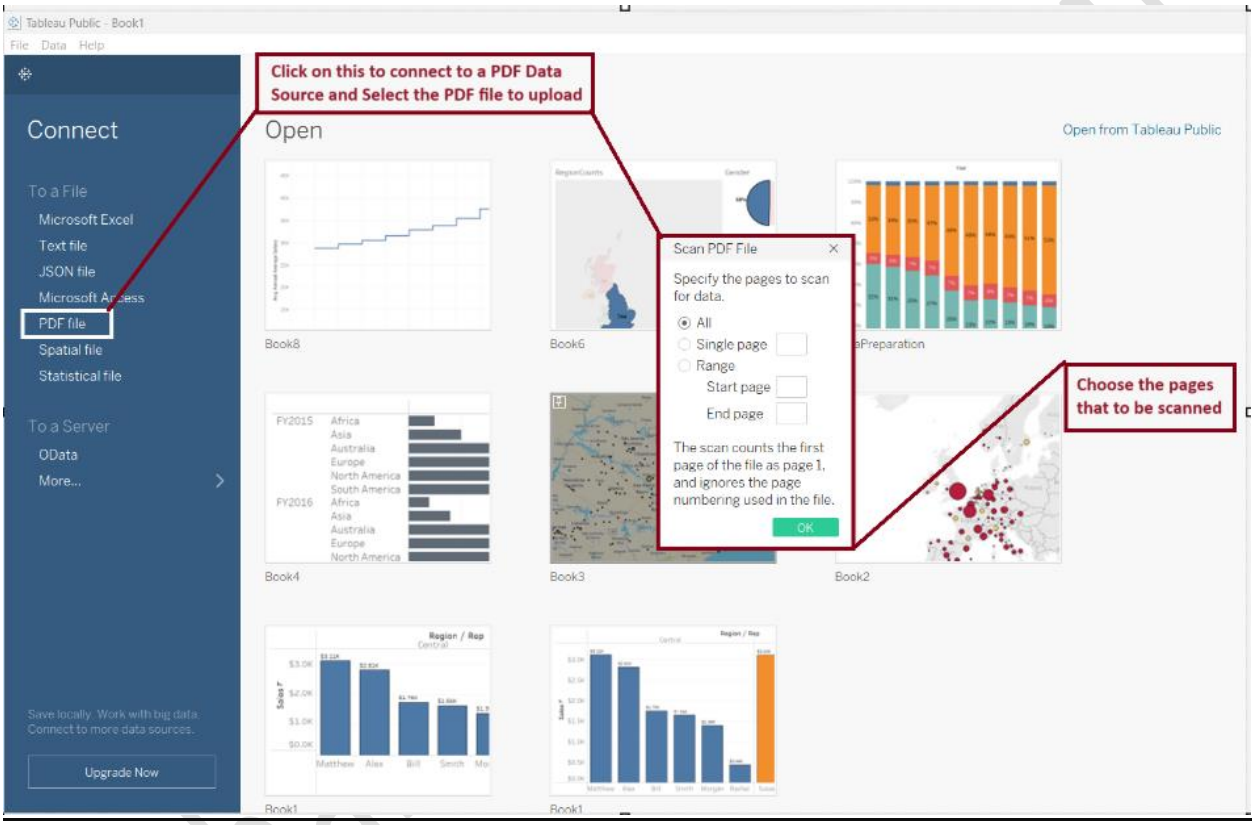


Data From PDF Files:

Download PDF Dataset from here:

[Tableau_Data_Sets/PDFDataSource_NYC_Crime.pdf at main · MohanPulibanti/Tableau_Data_Sets (github.com)]

Connecting to a PDF Data Source



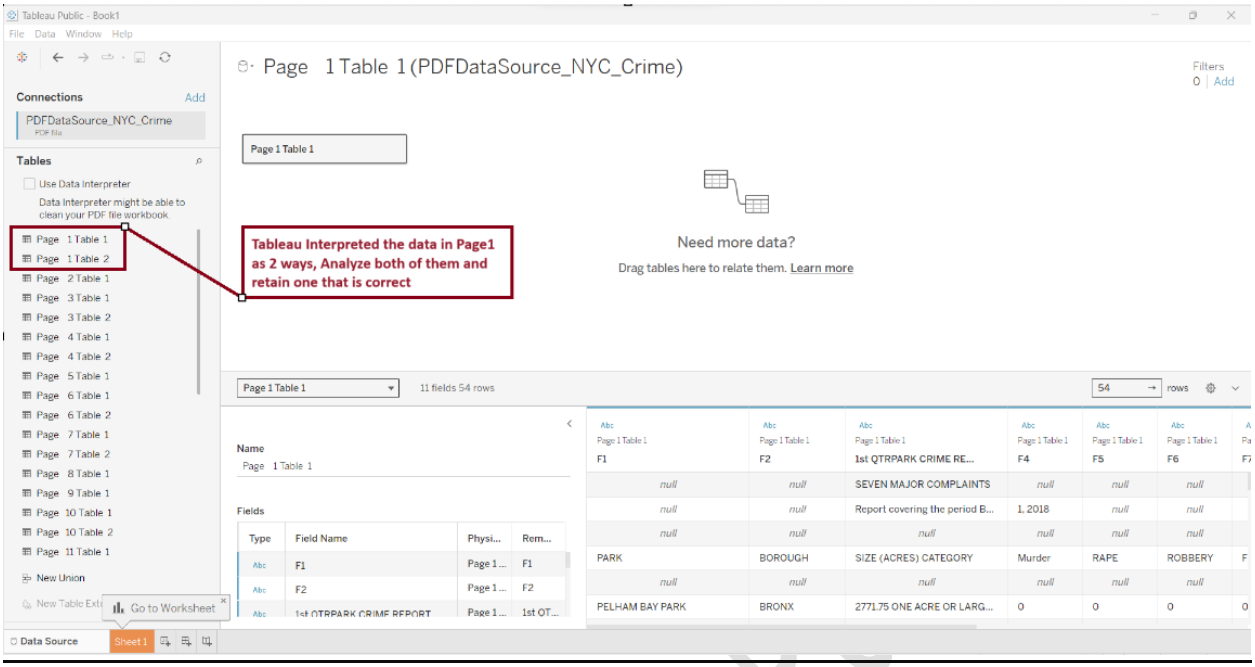
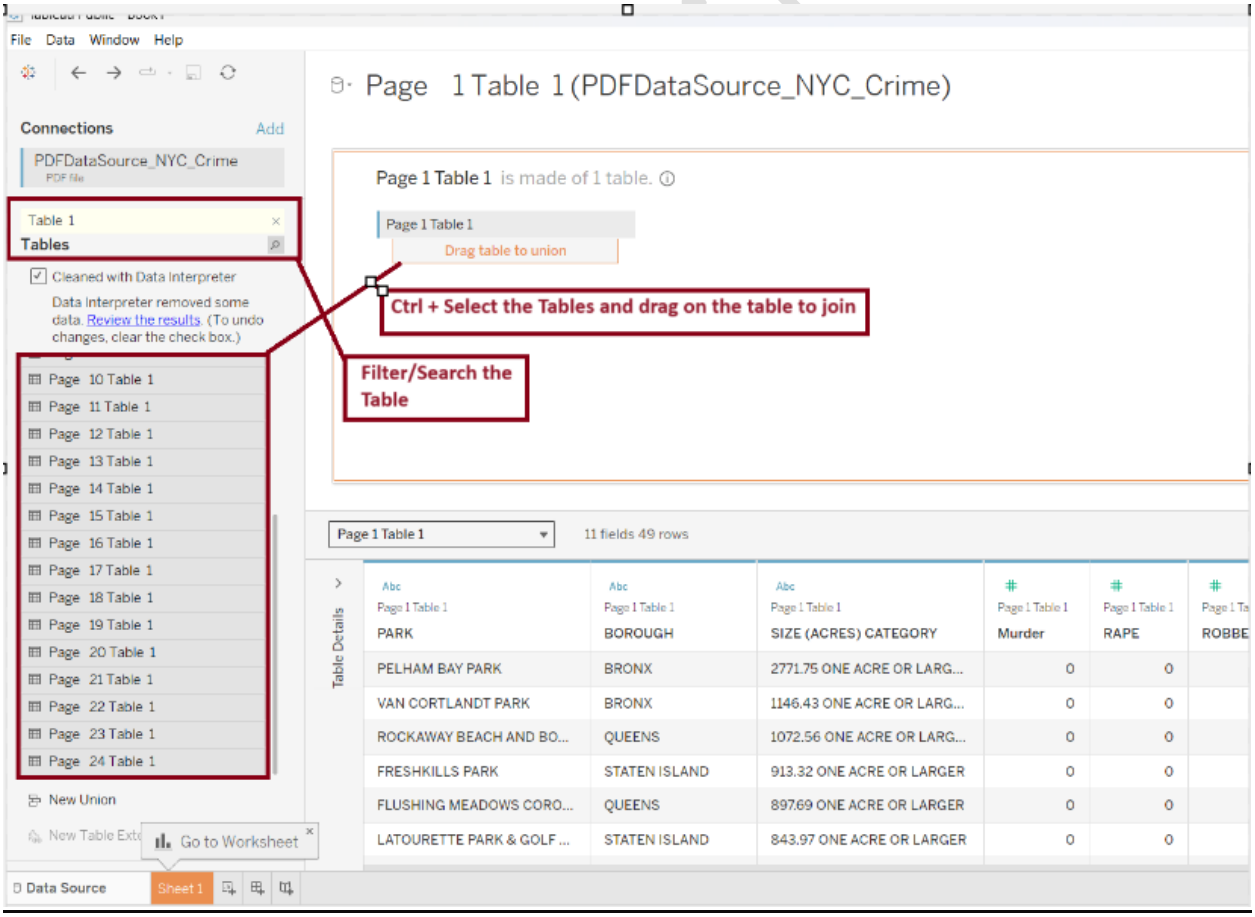
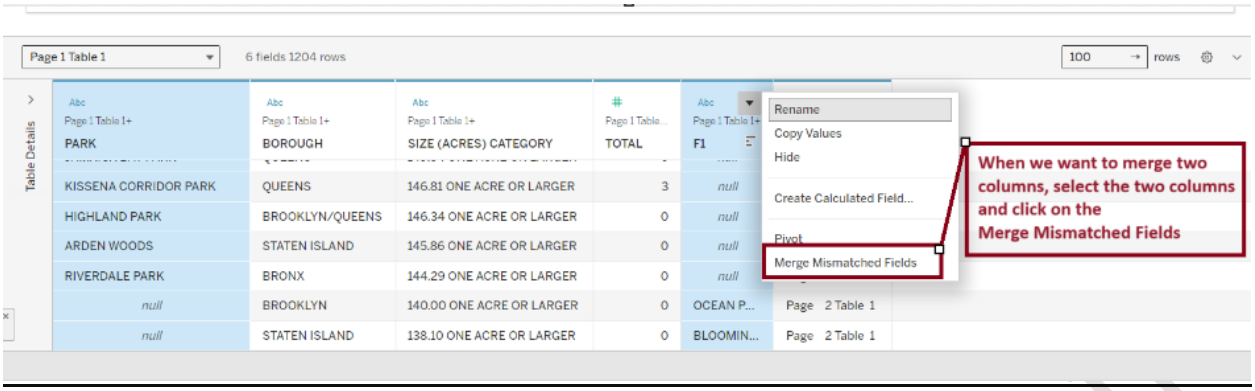


Tableau Public - Book1
File Data Window Help
Connections Add
PDFDataSource_NYC_Crime
PDF file
Tables
Use Data Interpreter
Data Interpreter might be able to clean your PDF file workbook.
Page 1 Table 1
Page 1 Table 2
Page 2 Table 1
Page 3 Table 1
Page 3 Table 2
Page 4 Table 1
Page 4 Table 2
Page 5 Table 1
Page 6 Table 1
Page 6 Table 2
Page 7 Table 1
Page 7 Table 2
Page 8 Table 1
Page 9 Table 1
Page 10 Table 1
Page 10 Table 2
Page 11 Table 1
New Union
Go to Worksheet
Data Source
Sheet 1
Page 1 Table 1 (PDFDataSource_NYC_Crime)
Filters
0 | Add
Page 1 Table 1
Need more data?
Drag tables here to relate them. Learn more
Tableau Interpreted the data in Page1 as 2 ways, Analyze both of them and retain one that is correct
11 fields 54 rows
54 rows
Name
Page 1 Table 1
Fields
Type
Field Name
Physi...
Rem...
F1
F2
1st QTRPARK CRIME REPORT
Page 1 Table 1
F1
F2
1st QTRPARK CRIME RE...
F4
F5
F6
null
SEVEN MAJOR COMPLAINTS
Report covering the period B...
1, 2018
PARK
BOROUGH
SIZE (ACRES) CATEGORY
Murder
RAPE
ROBBERY
PELHAM BAY PARK
BRONX
2771.75 ONE ACRE OR LARG...
0


File Data Window Help
Connections Add
PDFDataSource_NYC_Crime
PDF file
Table 1
Tables
Cleaned with Data Interpreter
Data Interpreter removed some data. Review the results. (To undo changes, clear the check box.)
Page 10 Table 1
Page 11 Table 1
Page 12 Table 1
Page 13 Table 1
Page 14 Table 1
Page 15 Table 1
Page 16 Table 1
Page 17 Table 1
Page 18 Table 1
Page 19 Table 1
Page 20 Table 1
Page 21 Table 1
Page 22 Table 1
Page 23 Table 1
Page 24 Table 1
New Union
Go to Worksheet
Data Source
Sheet 1
Page 1 Table 1 (PDFDataSource_NYC_Crime)
Page 1 Table 1 is made of 1 table.
Page 1 Table 1
Drag table to union
Ctrl + Select the Tables and drag on the table to join
Filter/Search the Table
Page 1 Table 1
11 fields 49 rows
Table Details
PARK
BOROUGH
SIZE (ACRES) CATEGORY
Murder
RAPE
PELHAM BAY PARK
BRONX
2771.75 ONE ACRE OR LARG...
0
VAN CORTLANDT PARK
BRONX
1146.43 ONE ACRE OR LARG...
ROCKAWAY BEACH AND BO...
QUEENS
1072.56 ONE ACRE OR LARG...
FRESHKILLS PARK
STATEN ISLAND
913.32 ONE ACRE OR LARGER
FLUSHING MEADOWS CORO...
QUEENS
897.69 ONE ACRE OR LARGER
LATOURETTE PARK & GOLF ...
STATEN ISLAND
843.97 ONE ACRE OR LARGER


Page 1 Table 1
6 fields 1204 rows
100 rows
Table Details
PARK
BOROUGH
SIZE (ACRES) CATEGORY
TOTAL
F1
KISSENA CORRIDOR PARK
QUEENS
146.81 ONE ACRE OR LARGER
HIGHLAND PARK
BROOKLYN/QUEENS
146.34 ONE ACRE OR LARGER
ARDEN WOODS
STATEN ISLAND
145.86 ONE ACRE OR LARGER
RIVERDALE PARK
BRONX
144.29 ONE ACRE OR LARGER
BROOKLYN
140.00 ONE ACRE OR LARGER
STATEN ISLAND
138.10 ONE ACRE OR LARGER
Rename
Copy Values
Hide
Create Calculated Field...
Pivot
Merge Mismatched Fields
When we want to merge two columns, select the two columns and click on the Merge Mismatched Fields

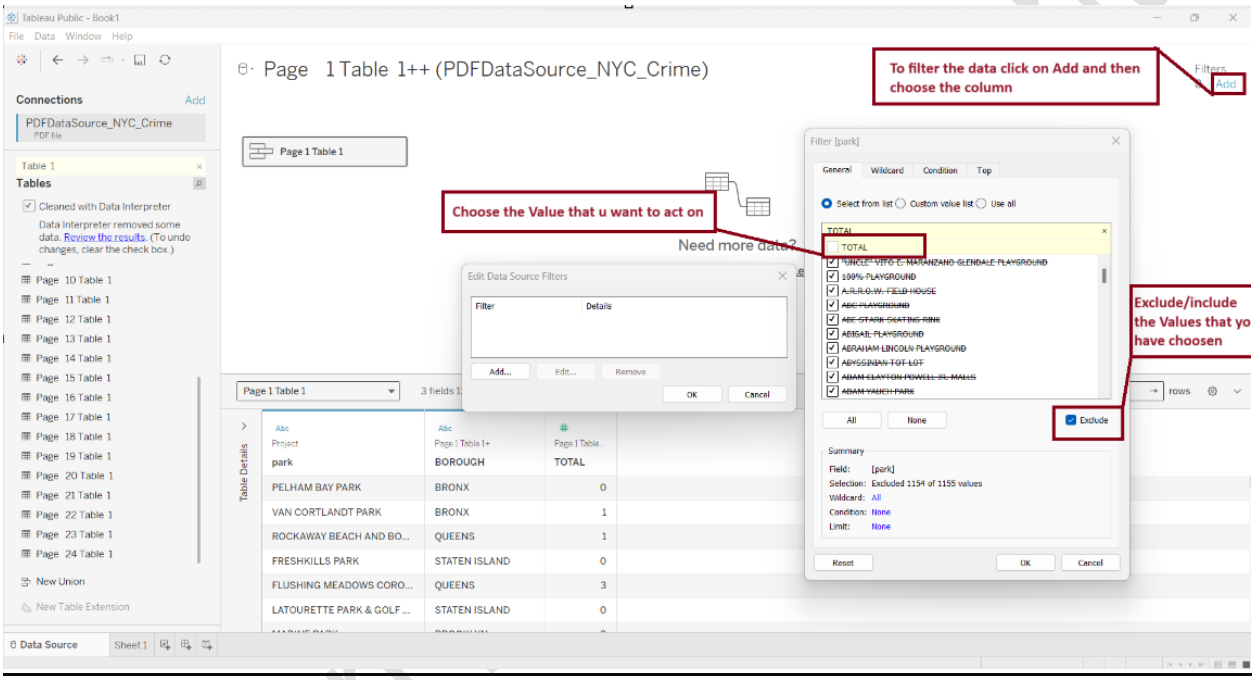
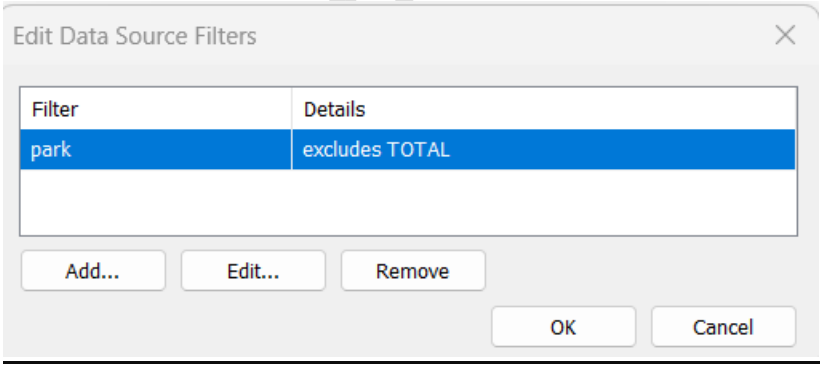


Tableau Public - Book1
File Data Window Help
Page 1 Table 1++ (PDFDataSource_NYC_Crime)
To filter the data click on Add and then choose the column
Filters
Add
Connections
Add
PDFDataSource_NYC_Crime
Tables
Page 1 Table 1
Choose the Value that u want to act on
Need more data?
Edit Data Source Filters
Filter
Details
Add...
Edit...
Remove
OK
Cancel
Filter [park]
General
Wildcard
Condition
Top
Select from list
Custom value list
Use all
TOTAL
Exclude/include the Values that yo have choosen
All
None
Exclude
Summary
Field: [park]
Selection: Excluded 1154 of 1155 values
Wildcard: All
Condition: None
Limit: None
Reset
OK
Cancel
Page 1 Table 1
park
BOROUGH
TOTAL
PELHAM BAY PARK
BRONX
VAN CORTLANDT PARK
BRONX
ROCKAWAY BEACH AND BO...
QUEENS
FRESHKILLS PARK
STATEN ISLAND
FLUSHING MEADOWS CORO...
QUEENS
LATOURETTE PARK & GOLF ...
STATEN ISLAND
Data Source
Sheet 1


Edit Data Source Filters
Filter
Details
park
excludes TOTAL
Add...
Edit...
Remove
OK
Cancel

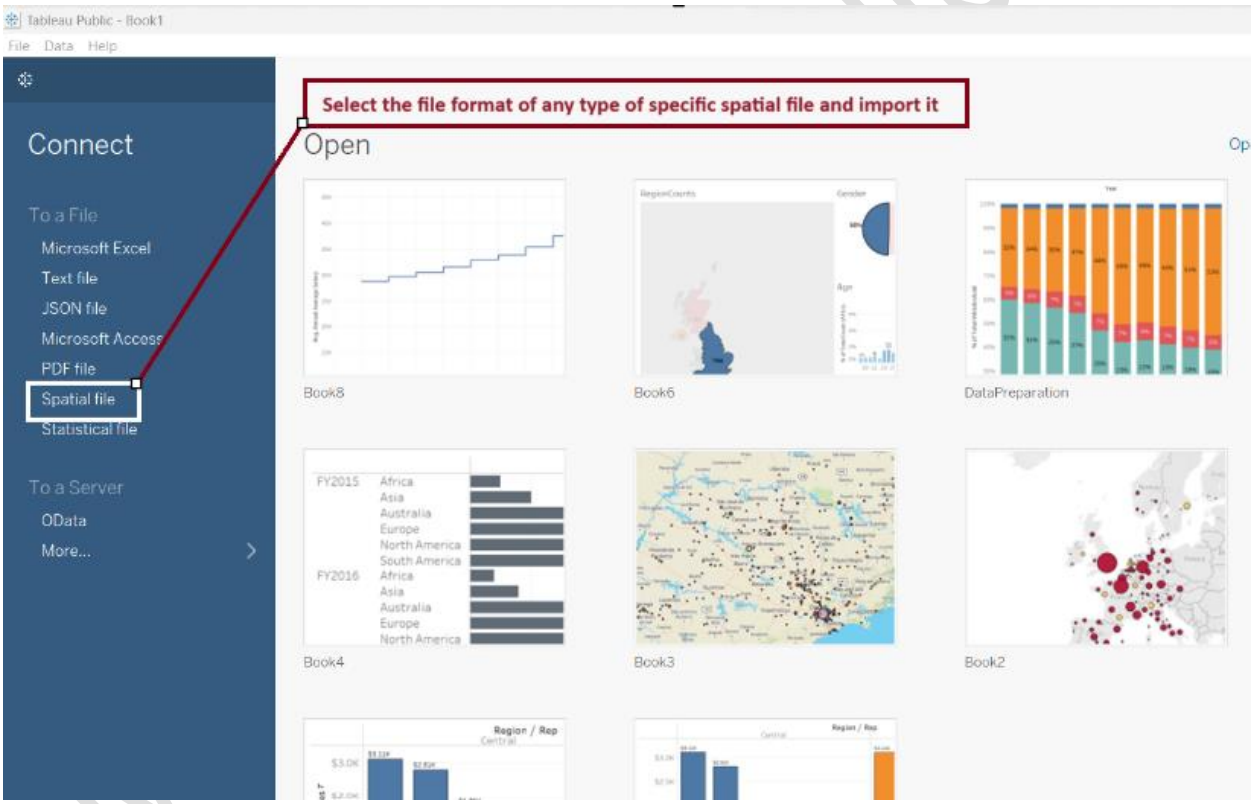
Connecting to Spatial Files

Tableau Allows us to connect to different Spatial Files

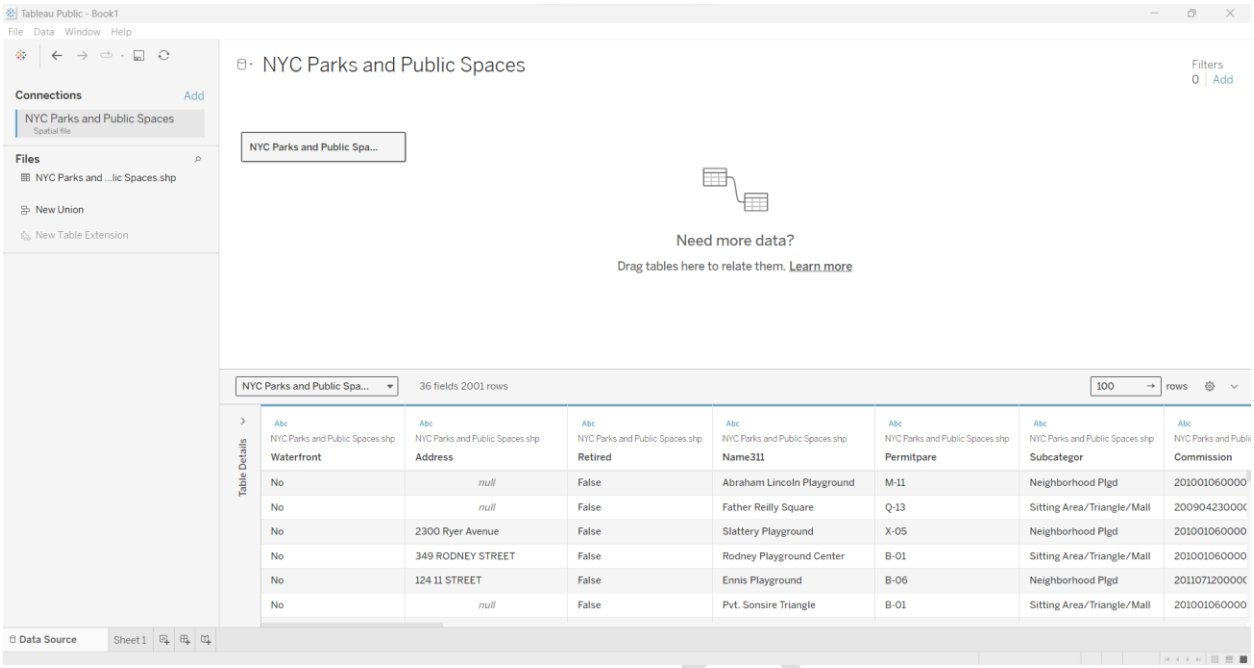
1) Shapefiles (.shp, .shx, .dbf) → Mostly widely used spatial files and we must have .shp, .shx, .dbf files in the package.
2) MapInfo tables (.TAB, .DAT, .MAP, .ID)
3) KML (Keyhole Markup Language) .kml
4) GeoJSON (.geojson)

Download your datset here:

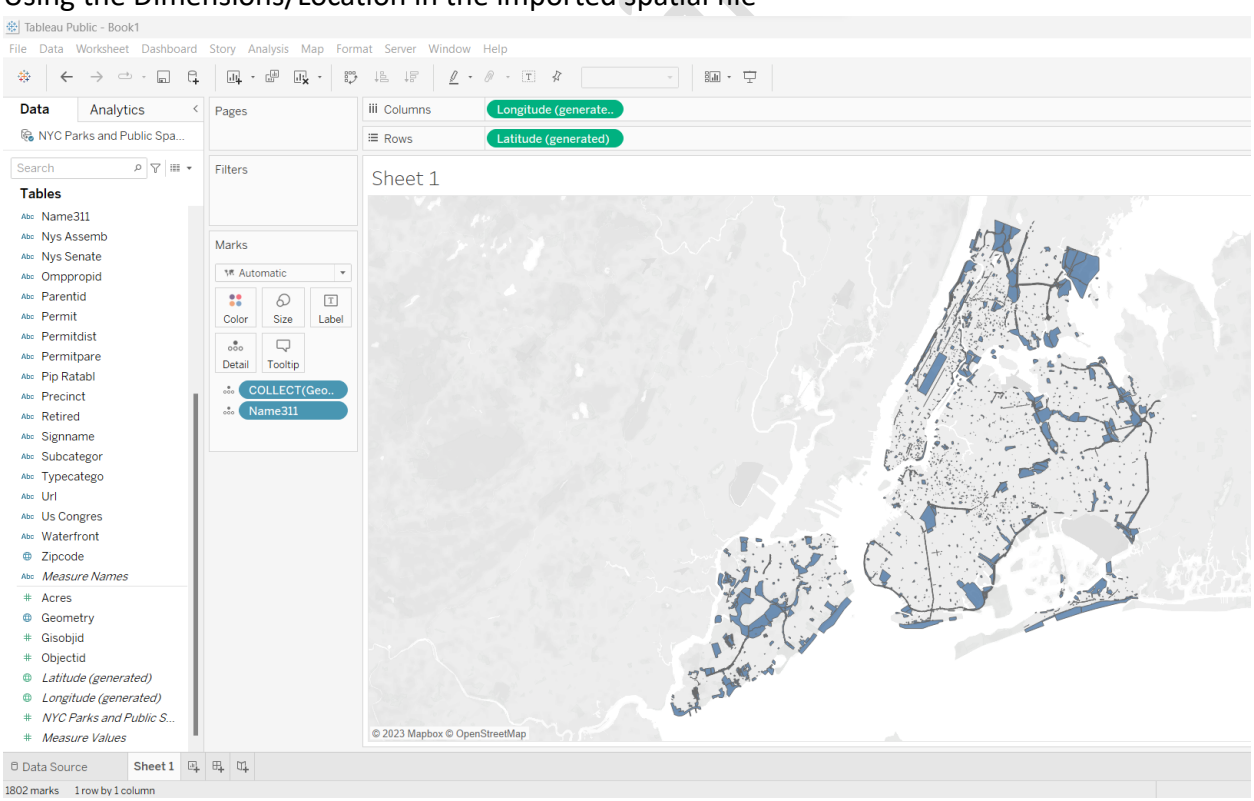
[Tableau_Data_Sets/Spatial_Files_Datasets at main · MohanPulibanti/Tableau_Data_Sets (github.com)]



Imported Spatial file:



Using the Dimensions/Location in the imported spatial file



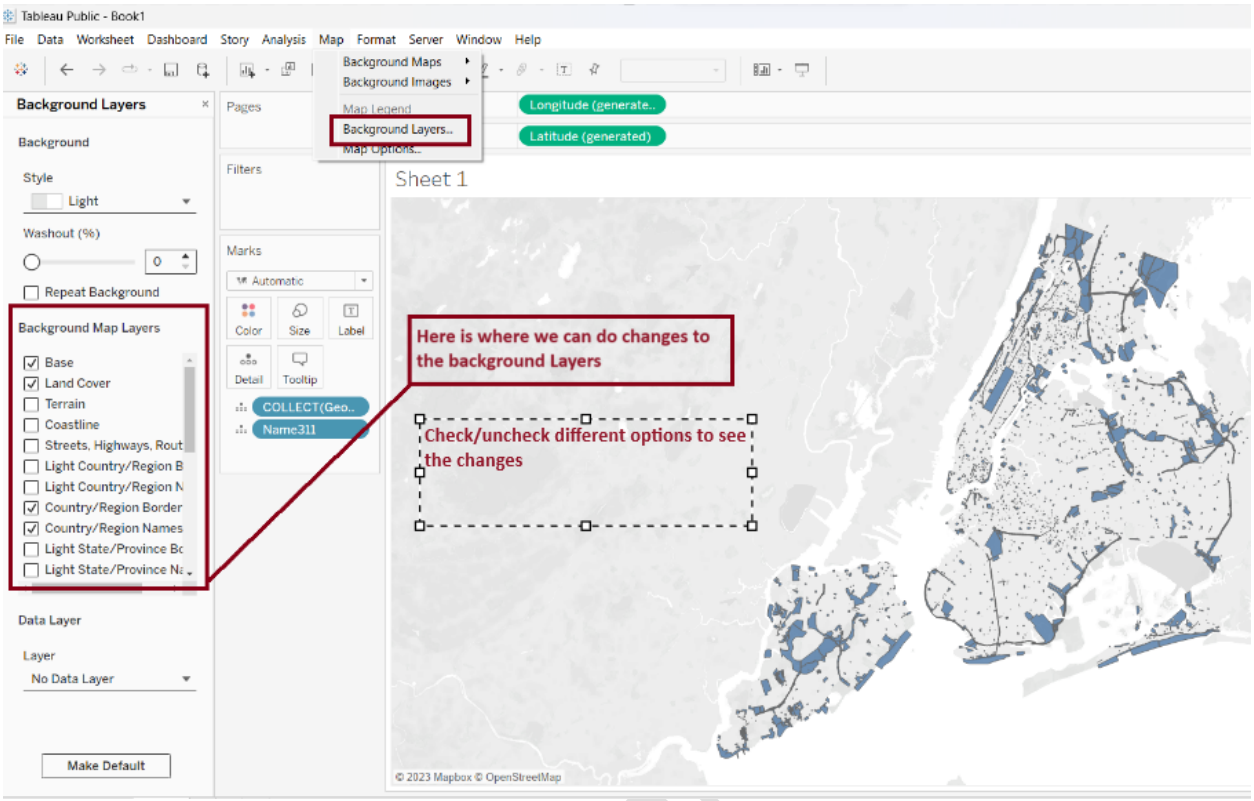


Tableau Public - Book1
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Background Maps
Background Images
Map Legend
Background Layers...
Map Options...
Background Layers
Background
Style
Light
Washout (%)
Repeat Background
Background Map Layers
Base
Land Cover
Terrain
Coastline
Streets, Highways, Rout
Light Country/Region B
Light Country/Region N
Country/Region Border
Country/Region Names
Light State/Province Bc
Light State/Province Na
Data Layer
Layer
No Data Layer
Make Default
Pages
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
COLLECT(Geo..
Name311
Longitude (generate..
Latitude (generated)
Sheet 1
Here is where we can do changes to the background Layers
Check/uncheck different options to see the changes
© 2023 Mapbox © OpenStreetMap

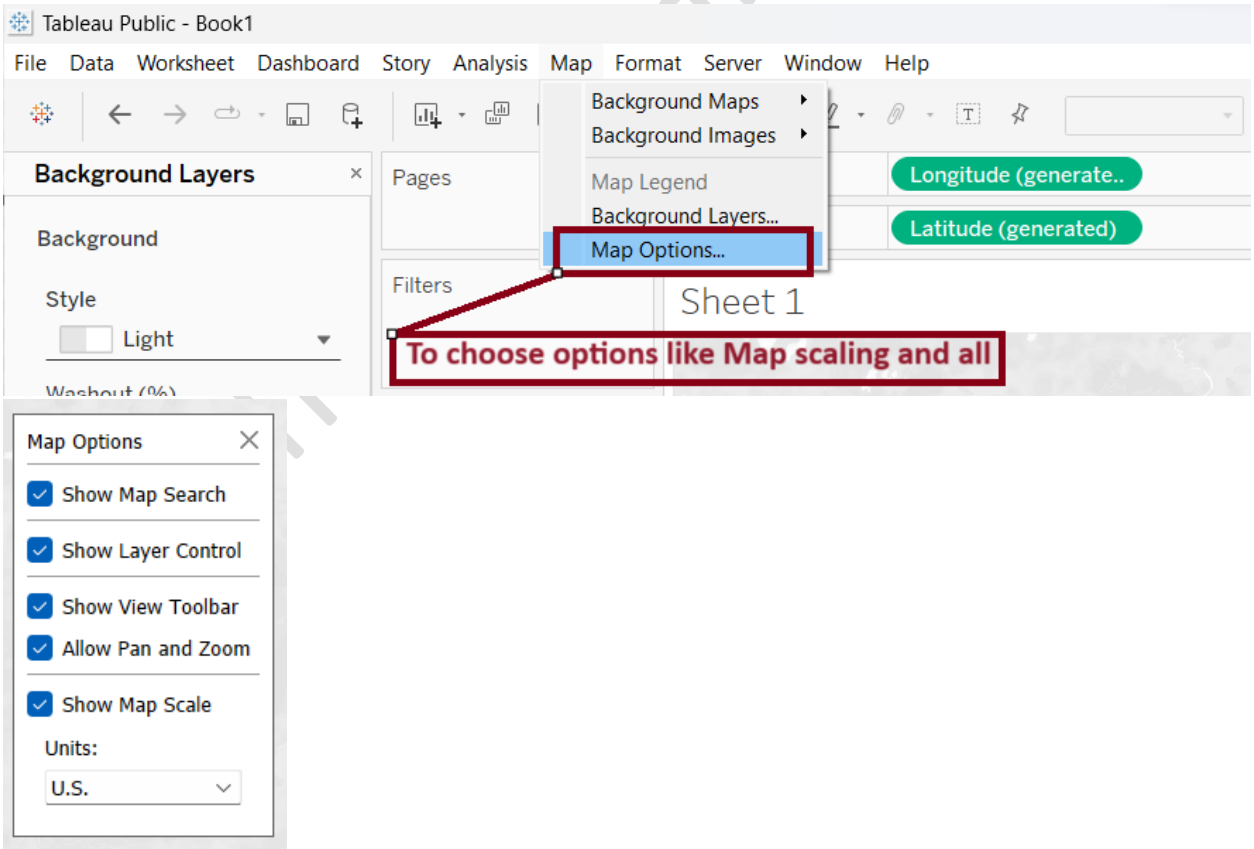
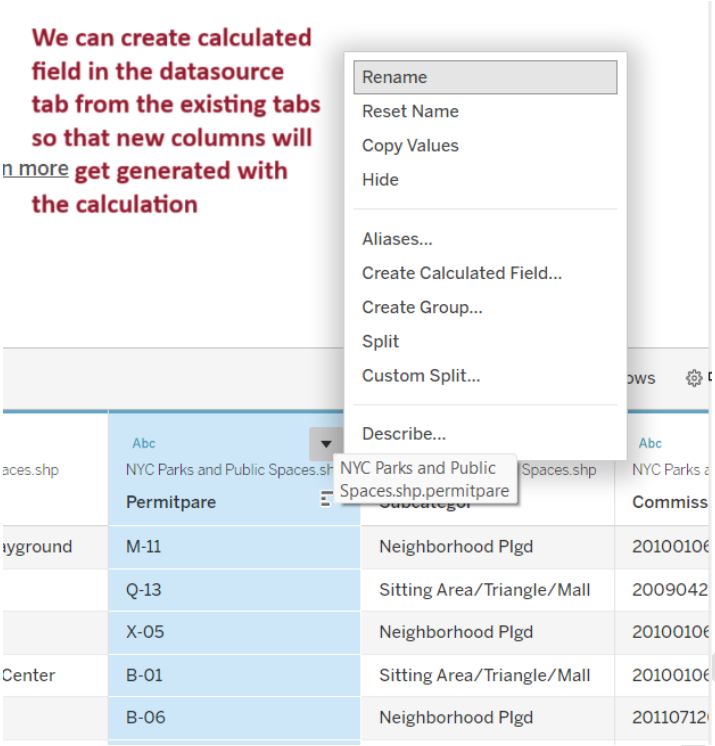


Tableau Public - Book1
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Background Maps
Background Images
Map Legend
Background Layers...
Map Options...
Background Layers
Background
Style
Light
Pages
Filters
Longitude (generate..
Latitude (generated)
Sheet 1
To choose options like Map scaling and all
Map Options
Show Map Search
Show Layer Control
Show View Toolbar
Allow Pan and Zoom
Show Map Scale
Units:
U.S.


We can create calculated field in the datasource tab from the existing tabs so that new columns will get generated with the calculation
Rename
Reset Name
Copy Values
Hide
Aliases...
Create Calculated Field...
Create Group...
Split
Custom Split...
Describe...
NYC Parks and Public Spaces.shp.permitpare
Permitpare
Commiss
M-11
Q-13
X-05
B-01
B-06
Neighborhood Plgd
Sitting Area/Triangle/Mall

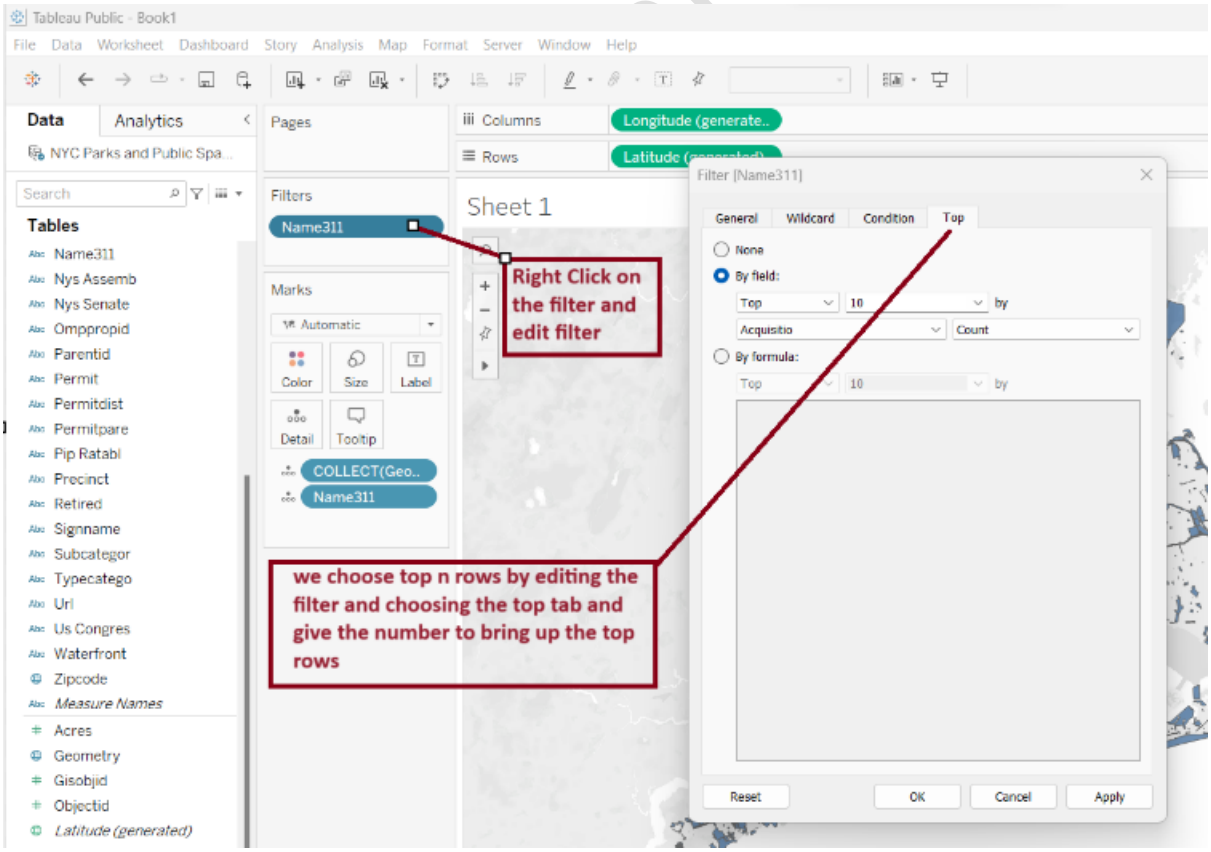


Tableau Public - Book1
Data
Analytics
Pages
Filters
Name311
Marks
Columns
Rows
Longitude (generate..
Sheet 1
Right Click on the filter and edit filter
Filter [Name311]
General
Wildcard
Condition
Top
None
By field:
By formula:
Reset
OK
Cancel
Apply
we choose top n rows by editing the filter and choosing the top tab and give the number to bring up the top rows

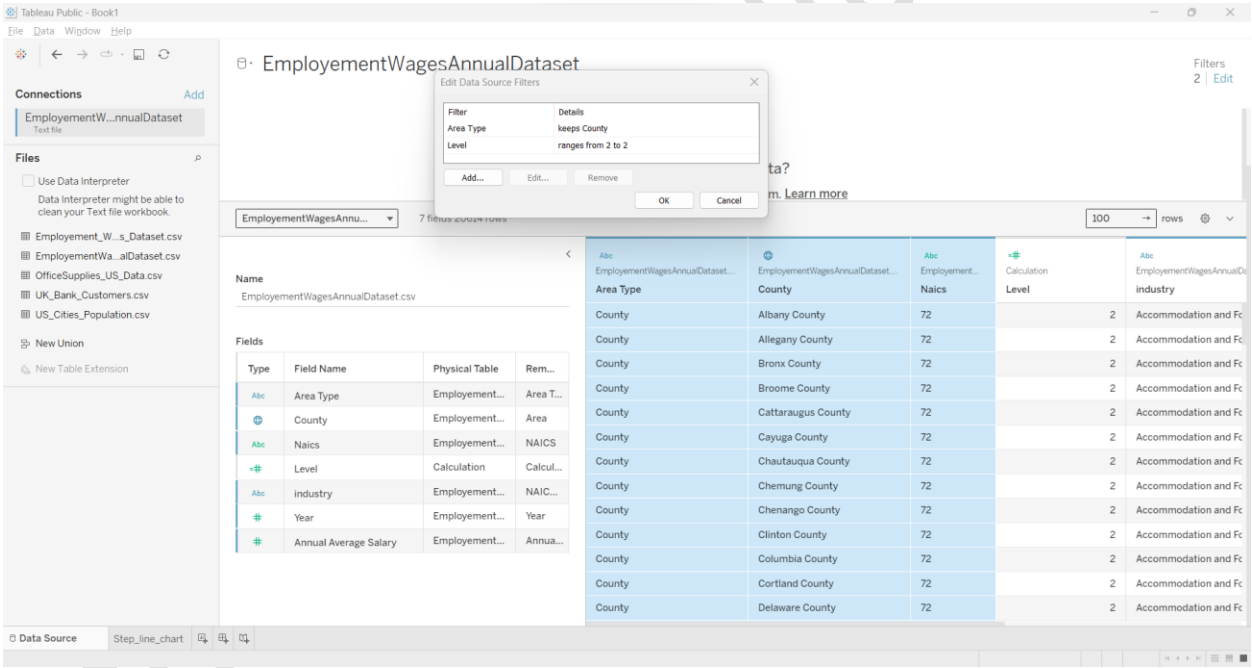
Visualization in ToolTip

Download Dataset here:

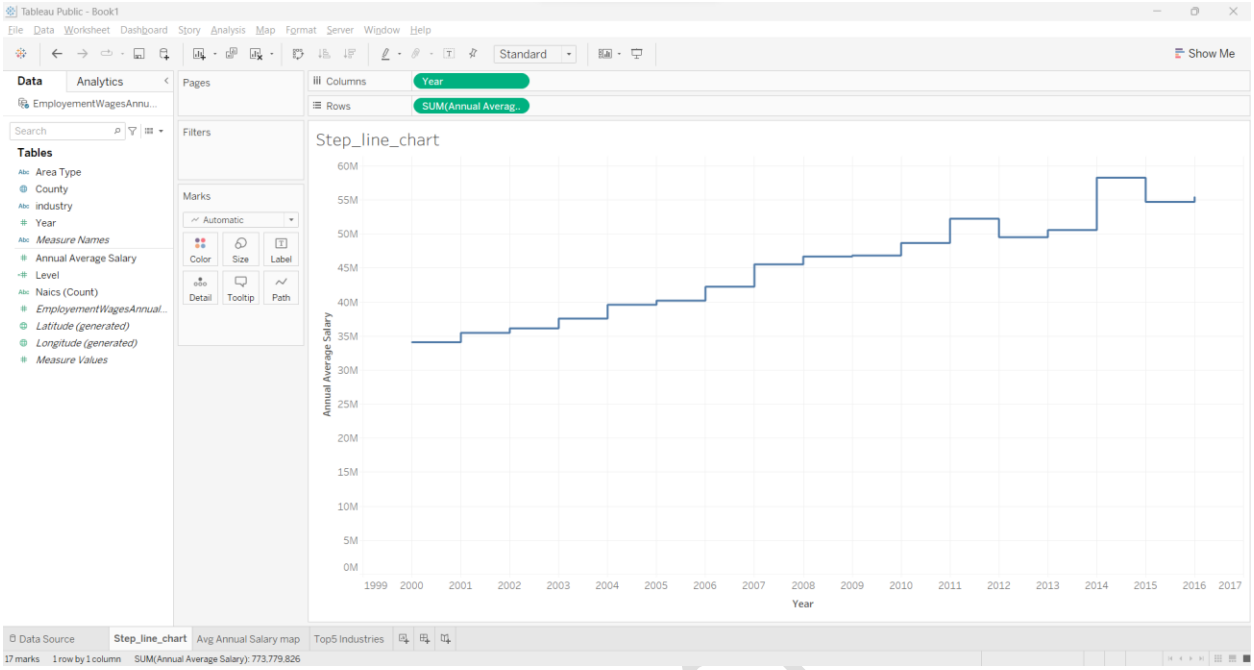
[Tableau_Data_Sets/EmployementWagesAnnualDataset.csv at main · MohanPulibanti/Tableau_Data_Sets (github.com)]

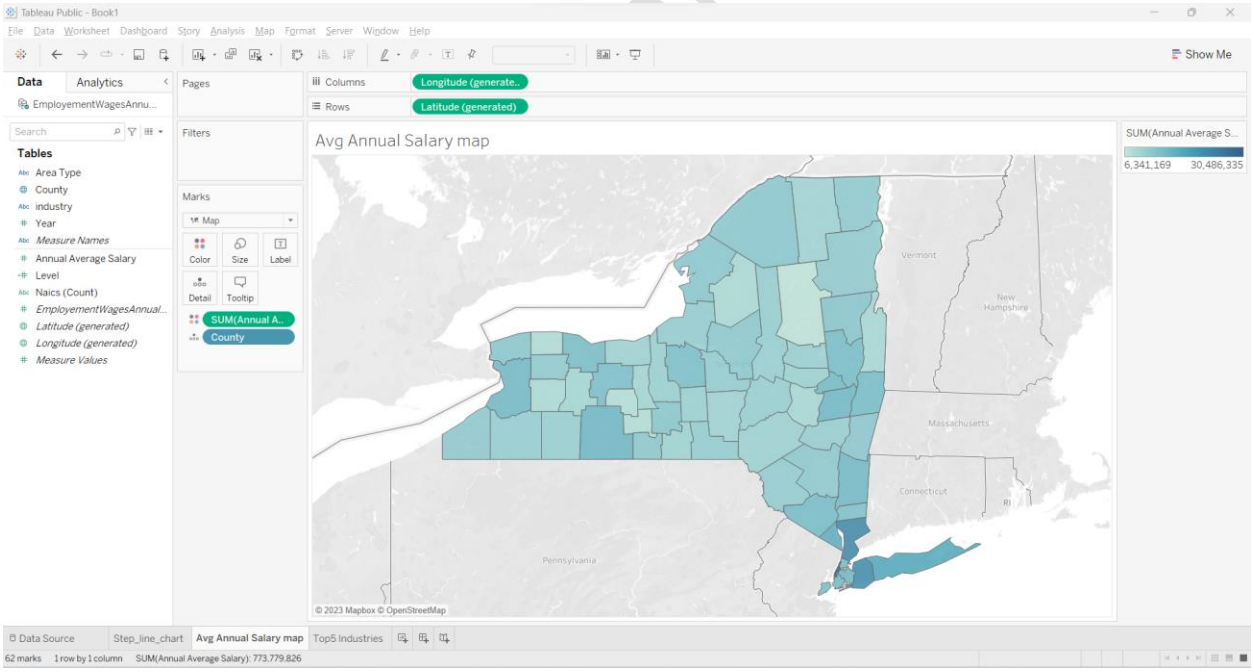
Data Preprocessing Steps:

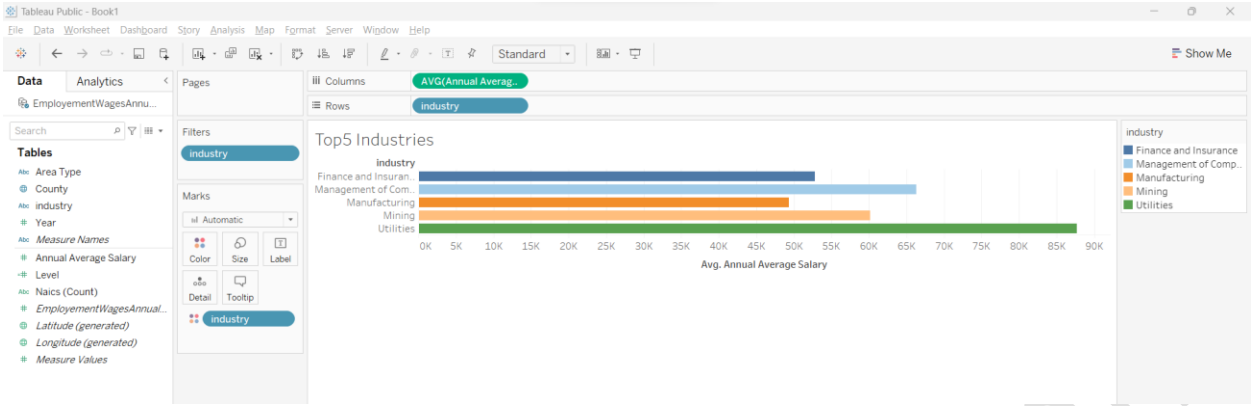
1) Filter Area Type with County
2) Filter the data where Naics length is 2. (convert Naics to string and then Create Calculated Filed named Level on Naics as LEN(Naics) and then add filter on the Level with range 2)



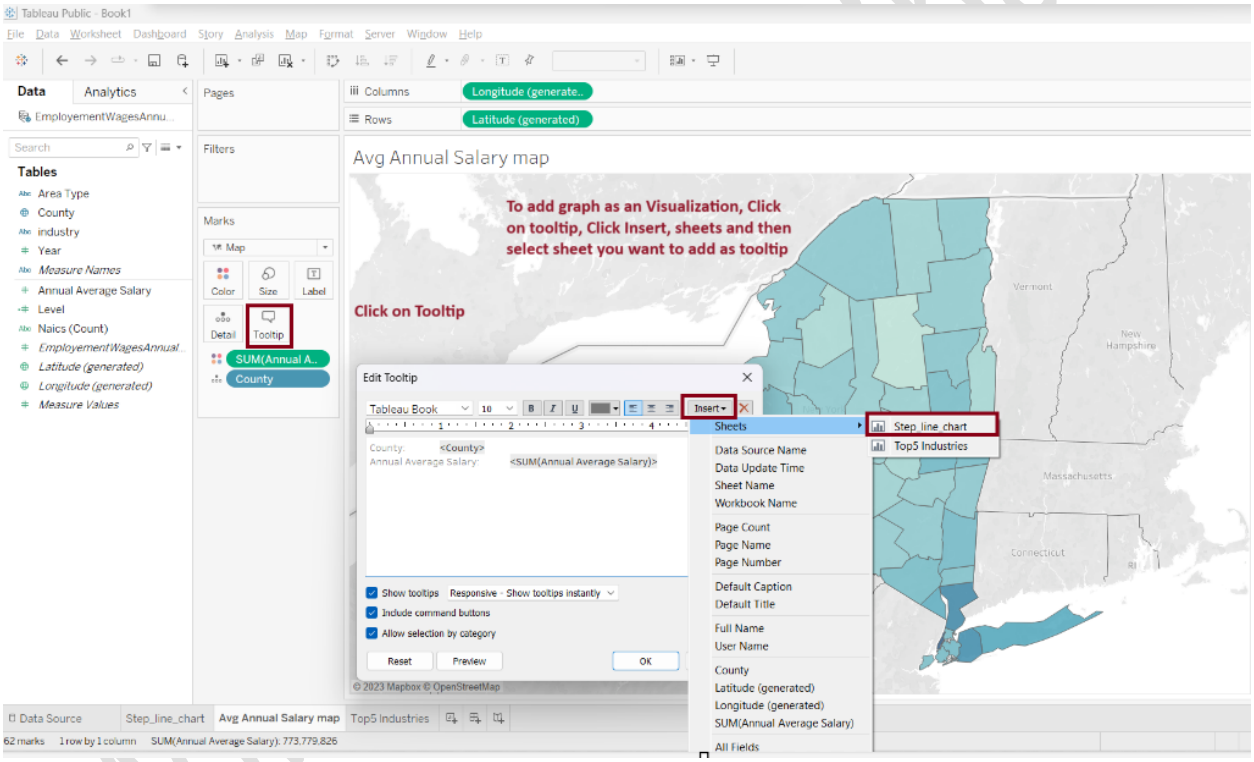
Created Three different Viz from the above dataset as below

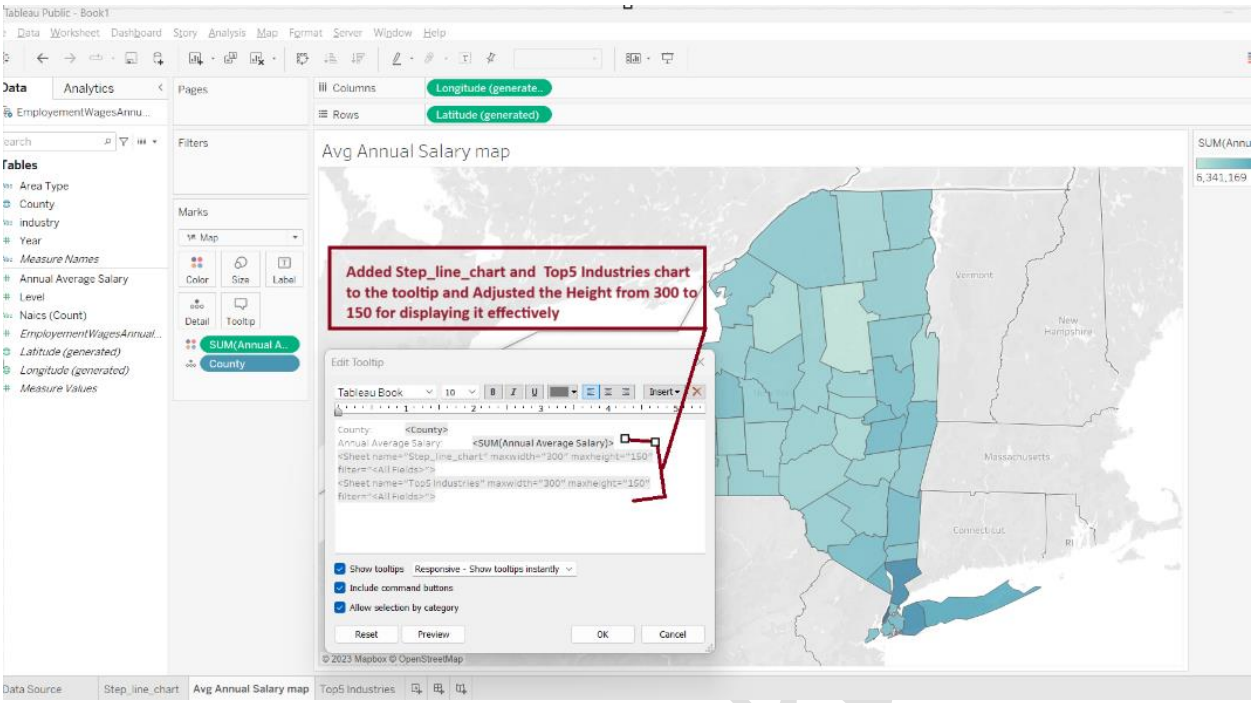




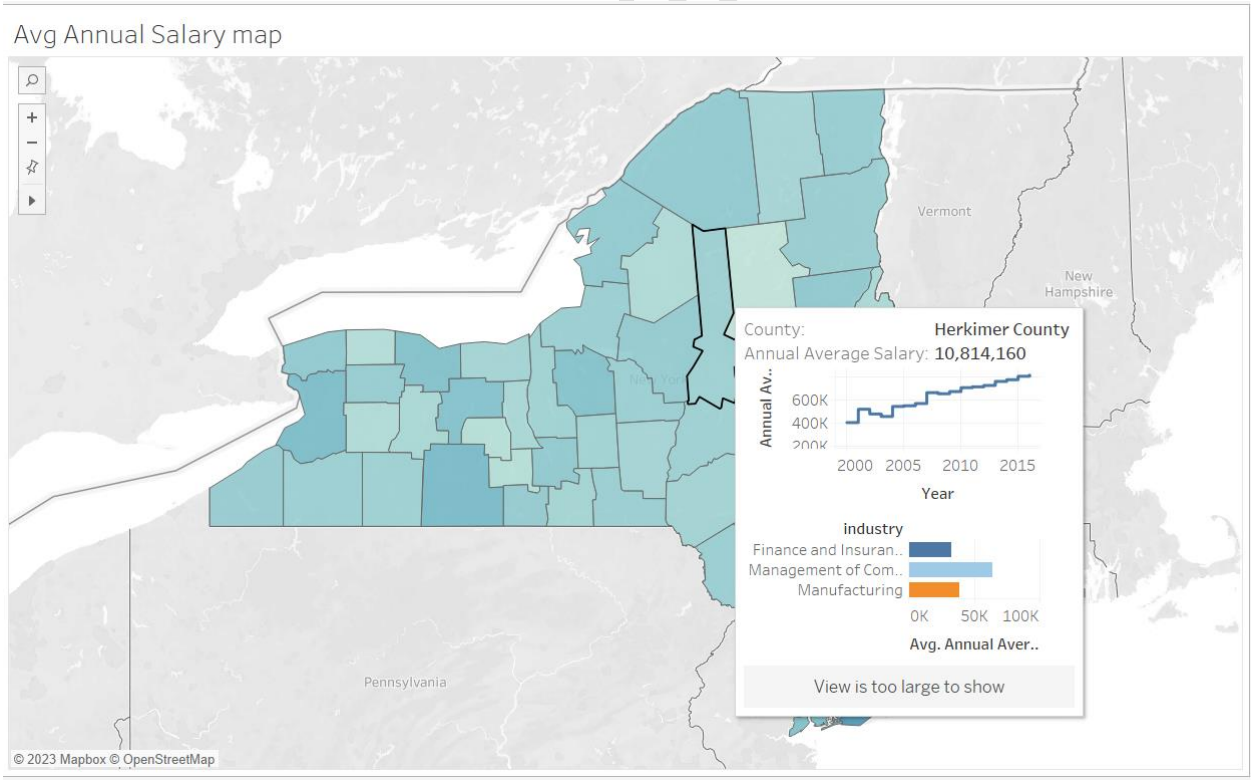


Now Adding Graph1 and Graph3 as tooptips Visualizations to graph2





In the below image, we can see sheets popping in the tooltip



Mohan Kumar Pulibanti