

A Hands-On Guide For Beginners

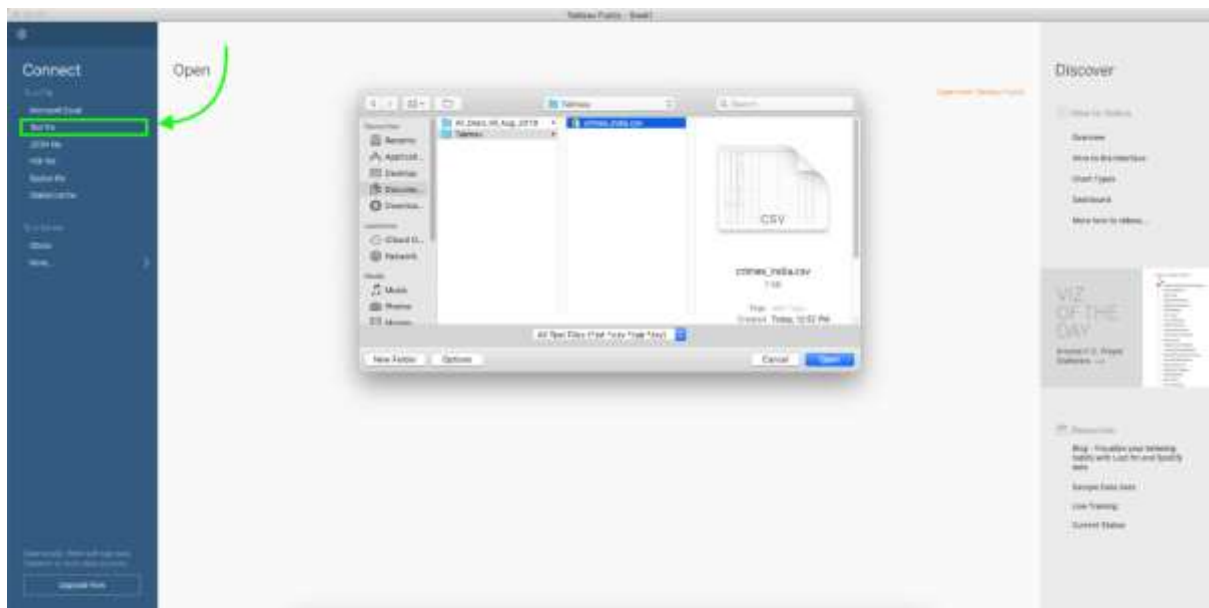
Tableau is one of the most popular tools for data visualization. As a platform that allows Data Analysts and Data Scientists to portray data effortlessly, Tableau saves a lot of time by not having to code for every pixel of the information displayed.

We will be using the state-wise crime rate in India Dataset to visualise it.

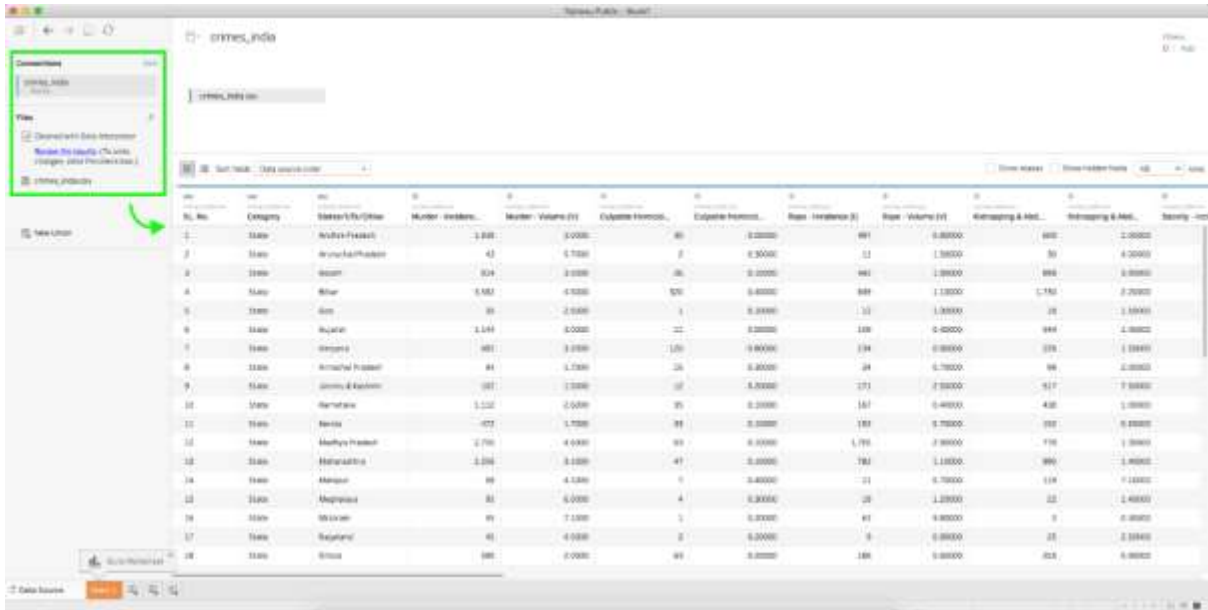
Connecting To The dataset

After you fire up the Tableau Public application, you will be greeted with a homepage that has a Connect panel on its left side. The panel will have a hand full of options to connect to a variety of data sources such as excel, text, pdf and also live database servers.

Our dataset is a .csv file. So we will click on the 'Text file' option. Select the downloaded file and click open.



Once you click open, the data will be imported into the Tableau worksheet. On the Data Source page, we will be able to see the active connections and the data which we imported.



Before we proceed to visualisation, we need to solve a small discrepancy in the data. If you observe closely, in the SL.No column we will find invalid values such as Total (States), Total (UT) and Total(All India).

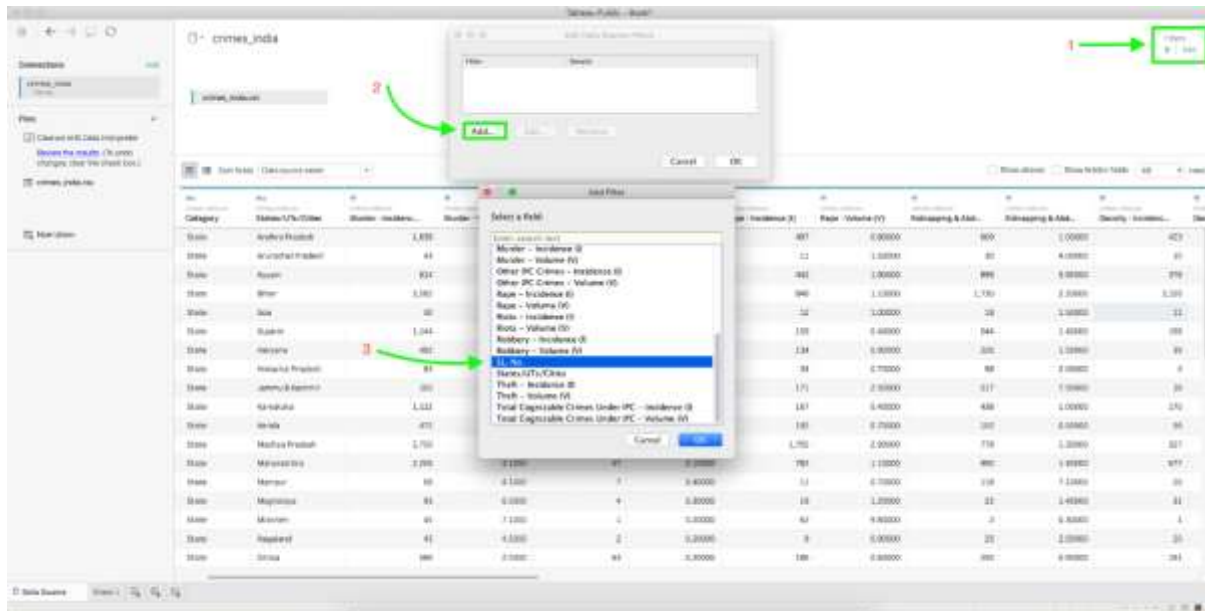
SL. No	Category	States/UTs/Cities	Murder - Incidents...	Murder - Volume (V)	Suspect Homid...	Outable Homid...	Rape - Incidents (I)	Rape - Volume (V)	Kidnapping & Abd...	Kidnapping & Abd...	Identity - In
21	State	Sikkim	6	1,000	1	0.0000	8	2,000	3	0.0000	
22	State	Tamil Nadu	1,531	2,900	22	0.0000	267	6,500	630	1.2000	
23	State	Tripura	144	6,000	4	0.0000	43	1,000	24	2.0000	
24	State	Uttar Pradesh	6,582	5,000	3,826	1.0000	1,382	1,300	3,441	2,700	
25	State	West Bengal	1,453	2,400	404	0.0000	632	1,000	886	1,400	
	Total (States)	Total (States)	26,144	6,000	3,318	0.0000	9,425	1,300	14,421	1,800	
26	UT	A & N Islands	7	8,000	1	0.0000	6	2,400	11	4.0000	
27	UT	Chandigarh	23	3,000	3	0.0000	6	6,000	27	4.0000	
28	UT	D & N Haveli	4	3,000		0.0000	2	1,600	3	2.0000	
29	UT	Daman & Diu	7	7,600		0.0000	1	1,300		0.0000	
30	UT	Delhi	311	3,000	386	2.0000	183	1,300	799	5.0000	
31	UT	Lakshadweep		0.0000		0.0000		0.0000		0.0000	
32	UT	Pondicherry	38	6,000		0.0000	7	1,000	8	1.0000	
	Total (UTs)	Total (UTs)	369	8,700	392	2.0000	124	1,900	840	6.0000	
	Total (All Ind.)	Total (All India)	26,513	6,600	3,712	0.0000	9,549	1,300	15,261	2,800	
1	City	Amritsar	81	8,000	1	0.0000	13	6,000	10	4.0000	
2	City	Bangalore	138	9,000	9	0.0000	44	1,400	86	2,700	

If we ignore these rows, it will mess up the numbers that we are going to visualise. For example, when Tableau aggregates the sum of a specific crime for the entire country, it will sum up all the rows including the ones shown above. We don't need these rows for now so we will remove them from the data temporarily.

Adding Filters

How do we remove those rows temporarily? We can filter them out using the Tableau's filters.

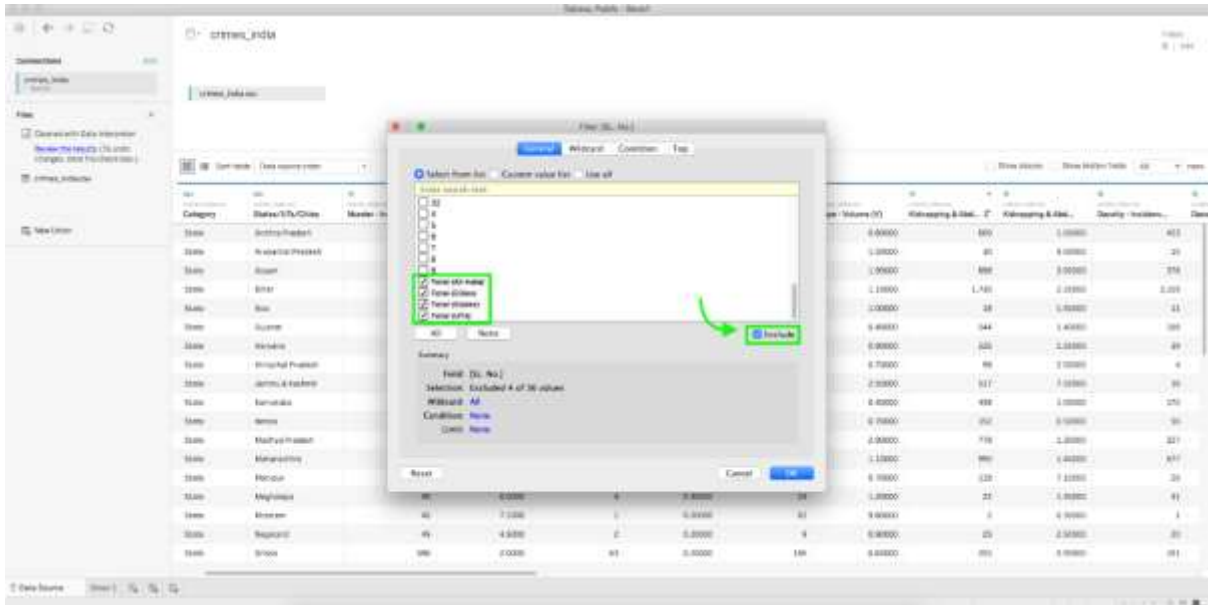
On the top right corner of the Data Source page, you will see a Filters section and an Add button. Click on the add button and a pop-up window will appear. Click on add and a new pop up will appear listing all the columns in the dataset. Select the column we want to filter data based on (SL.No.) and click OK.



Once you click OK, yet another popup window appears with a list of all the values in the selected column (SL.No) and gives several filter options like General, Wildcard, Condition, Top.

- **General** lets you select specific values using the checkbox, you can either use the checkbox to select features or to exclude the checked features by checking the exclude flag at bottom right corner.
- **Wildcard** lets us add filters based on patterns in the data.
- **Condition** adds filters based on mathematical conditions or formulas.
- **Top** lets us select the Top or Bottom portion of the dataset based on conditions.

We will stick to the general filter for now. Check the **exclude** flag on and select the data we want to exclude from the dataset.



Tip: You can also use the wildcard filter to exclude all values containing the word Total.

Click OK and you will see the filter summary of how many rows have been selected from the total number of rows. Click OK again and the dataset will be refreshed.

Congrats, you have learned to add Filters in Tableau!

Let us now understand some of the most common data manipulation techniques in Tableau.

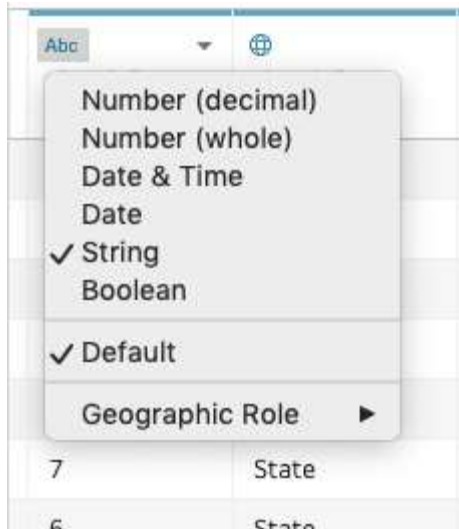
Data Types, Calculated Fields & Splits

Some of the most important techniques are within the head-band of the dataset. So let's focus on it.

SL No.	Category	Status/Type/Cities	Munier - Incidents...	Murder - Volume (Y)	Capital Homicide...	Capital Homicide...	Rape - Incidences (I)	Rape - Volume (Y)	Kidnapping & Abduct...	Kidnapping & Abduct...	Security - Inc...
1	State	Andhra Pradesh	1,328	3,900	30	0.0000	487	8,9000	608	1,0000	
2	State	Andhra Pradesh	43	5,700	2	0.3000	21	1,9000	30	4,0000	
3	State	Andhra Pradesh	814	9,900	76	0.1000	447	1,9000	894	9,9000	
4	State	Andhra Pradesh	1,302	8,500	100	0.4000	649	1,9000	1,710	2,7000	
5	State	Andhra Pradesh	30	2,600	1	0.1000	22	1,9000	18	1,9000	

- Data Types: Changing the Data types of features in Tableau is as easy as selecting from a drop-down menu. On each column header, there is an option to

change the data type. Abc specifies that currently the column is set to “String”. The globe stands for geographical data which is a speciality of Tableau that lets it identify geographical locations which will help in plotting maps. To change the data type click on the current data type (such as Abc, globe, # etc) to select a new type from the drop-down.



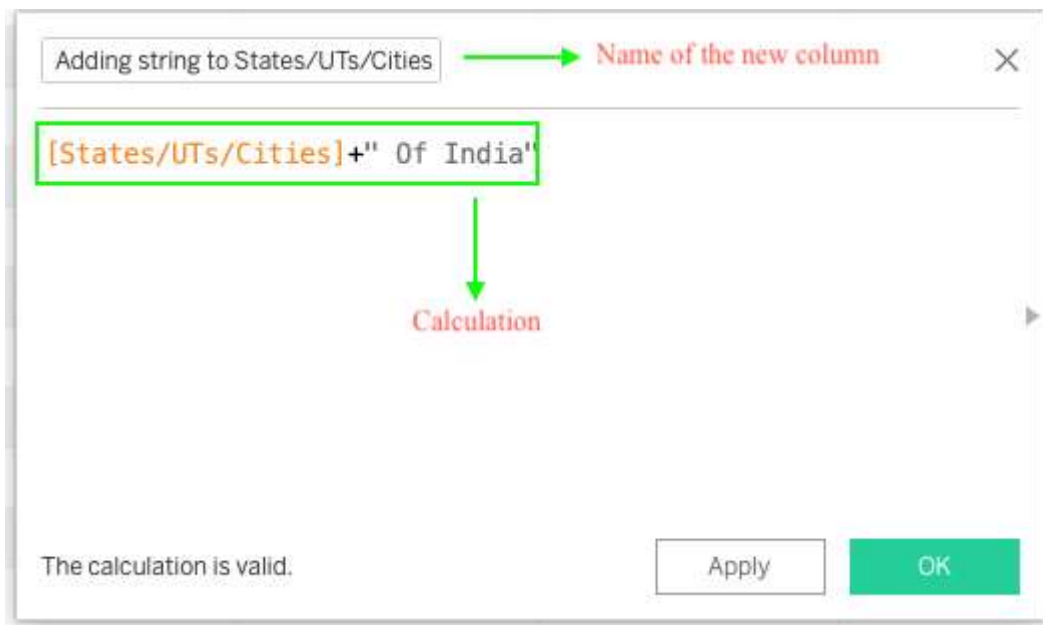
- **Calculated Fields:** Calculated Fields are very important when we are handling derivative data or Data that can be used to make calculations or derive a new feature. For example, calculating the total marks of 5 subjects represented in 5 columns.

Abc crimes_india.csv SL. No.	crimes_india.csv Category	crimes_india.csv States/UTs/Cities	#	#	crimes_india.csv r - Volume (V)	crimes_india.csv Culpable Homicid...
1	State	Andhra Pradesh			3.0000	30
2	State	Arunachal Pradesh			5.7000	2
3	State	Assam			3.5000	26
4	State	Bihar			4.5000	320
5	State	Goa			2.6000	1
6	State	Gujarat			3.0000	11
7	State	Haryana			3.2000	120
8	State	Himachal Pradesh		84	1.7000	15
9	State	Jammu & Kashmir		102	1.5000	12

To create a calculated field click on the “Create Calculated Field” option. This will open up a calculation field where you can enter the criteria or formula for the new field.

For example, let’s add a string “ Of India ” to all the values in the column States/UTs/Cities of the dataset and add it as a new column. To do this click on the drop-down from the column States/UTs/Cities ” and select create calculated

field" option. On the upcoming window add the expression for the desired calculation. See the image below.



Once you click apply, you will find a new column in your dataset like what's shown below:

SL. No.	Category	States/UTs/Cities	Adding string to States/UTs/Cities	Murder - Incidenc...
1	State	Andhra Pradesh	Andhra Pradesh Of India	1,838
2	State	Arunachal Pradesh	Arunachal Pradesh Of India	43
3	State	Assam	Assam Of India	814
4	State	Bihar	Bihar Of India	3,582
5	State	Goa	Goa Of India	30
6	State	Gujarat	Gujarat Of India	1,144
7	State	Haryana	Haryana Of India	482
8	State	Himachal Pradesh	Himachal Pradesh Of India	84
9	State	Jammu & Kashmir	Jammu & Kashmir Of India	102
10	State	Karnataka	Karnataka Of India	1,112
11	State	Kerala	Kerala Of India	472
12	State	Madhya Pradesh	Madhya Pradesh Of India	2,750

Try doing some mathematical operations on the numerical features or columns!

- Splits: We can split the values in one column as different features. For example, let's split the newly generated calculated field with the word "Of". To do this select the drop-down menu of the column and select "Custom split" option. On the upcoming window choose the separator and choose whether to keep both sides of the split or either of it.

The screenshot shows a data table with columns for state names and murder volume. A dialog box titled "Custom Split" is open, allowing the user to define how a column should be split. The dialog box contains the following elements:

- Question: "How should this data be split?"
- Separator: A text input field containing "Of".
- Split off: A dropdown menu currently set to "All".
- Number of columns: A numeric input field set to "1".
- Buttons: "Cancel" and "OK".

State	Murder - Incidence	Murder - Volume
Andhra Pradesh Of India		1,838
Arunachal Pradesh Of India		
Assam Of India		
Bihar Of India		
Goa Of India		
Gujarat Of India		
Haryana Of India		
Himachal Pradesh Of India		84
Jammu & Kashmir Of India		102
Karnataka Of India		1,112
Kerala Of India		472
Madhya Pradesh Of India		2,750

Once you click ok you will find 2 new columns at the end of the dataset as shown below.

# crimes_india.csv Total Cognizable C...	# crimes_india.csv Total Cognizable C...	=Abc Calculation Adding string to S...	=Abc Calculation Adding string to S...
94,279	156.400	Andhra Pradesh	India
1,355	179.800	Arunachal Pradesh	India
37,244	161.800	Assam	India
107,245	134.100	Bihar	India
4,551	388.800	Goa	India
84,714	219.400	Gujarat	India
22,918	149.900	Haryana	India
6,479	134.600	Himachal Pradesh	India
19,158	277.800	Jammu & Kashmir	India
79,456	187.300	Karnataka	India
56,806	199.500	Kerala	India
181,103	302.800	Madhya Pradesh	India
184,596	258.800	Maharashtra	India

We will now move on to visualisation. Delete all the three columns generated in the above examples by selecting the drop-down and clicking delete option.

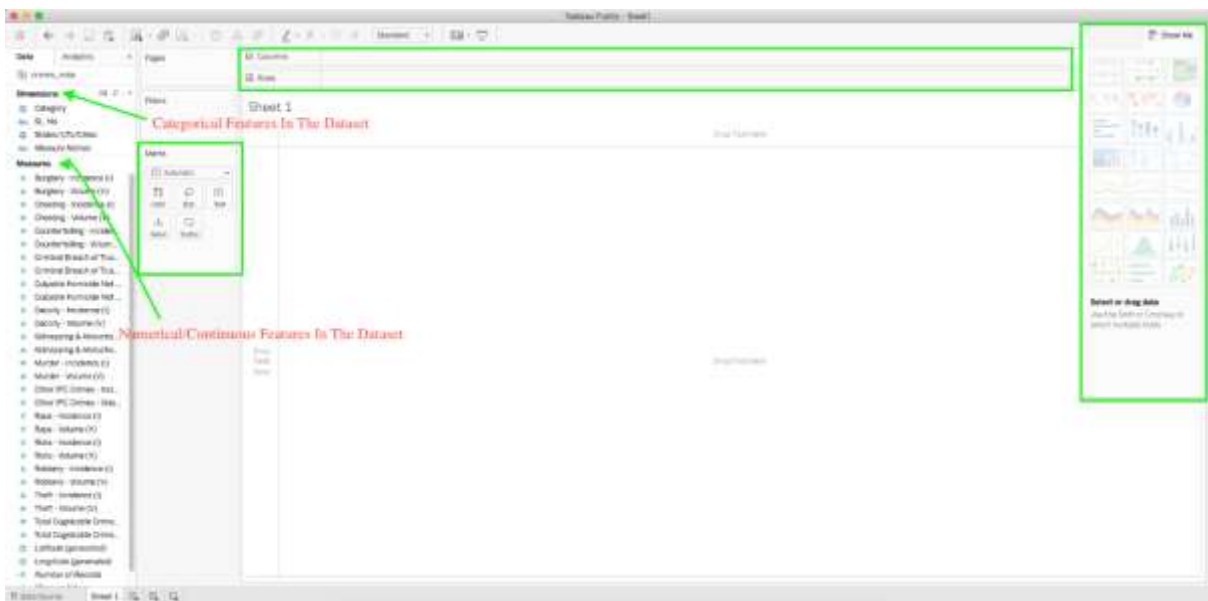
Visualising Crimes In India

To open your visualisation worksheet click on the sheet 1 tab at the bottom left corner.

State	Category	State/Type/Offense	Number - incidents (I)	Murder - Volume (V)	Substance Offense -	Substance Offense -	Rape - Volume (V)	Rape - Volume (V)	Kidnapping & Hostage -	Kidnapping & Hostage -
1	State	Armed Robbery	2,208	0.0000	89	0.0000	497	0.0000	607	1.0000
1	State	Armed Robbery	47	0.7500	2	0.0000	11	1.0000	53	4.0000
1	State	Arson	114	0.0000	26	0.0000	442	1.0000	868	0.0000
4	State	Arson	9,183	4.0000	931	0.0000	369	1.0000	1,739	0.2000
5	State	Sex	30	0.0000	1	0.0000	22	1.0000	18	1.0000
6	State	Sexual	1,344	0.0000	41	0.0000	309	0.0000	144	1.0000
7	State	Sexual	403	0.0000	101	0.0000	234	0.0000	228	1.0000
8	State	Armed Robbery	84	1.7500	15	0.0000	34	0.7500	14	0.0000
8	State	Armed Robbery	302	1.0000	22	0.0000	171	1.0000	107	7.0000
10	State	Domestic	1,117	0.0000	22	0.0000	127	0.0000	438	1.0000
11	State	Arson	919	1.7500	49	0.0000	199	0.0000	101	0.0000
12	State	Armed Robbery	2,701	4.0000	43	0.0000	1,705	0.0000	778	1.0000
13	State	Armed Robbery	3,200	0.0000	47	0.0000	302	1.0000	896	1.0000
14	State	Sexual	68	0.0000	7	0.0000	11	0.7500	118	0.0000
16	State	Sexual	91	0.0000	4	0.0000	10	1.0000	22	1.0000
16	State	Sexual	68	1.0000	1	0.0000	12	0.0000	1	0.0000
17	State	Sexual	49	0.0000	2	0.0000	9	0.0000	25	0.0000
18	State	Sexual	388	1.0000	48	0.0000	305	0.0000	454	0.0000

Tableau is a highly interactive GUI tool for plotting graphs and creating a beautiful visualisation, which means there is zero coding required. However, it also means that we are required to interact with the user interface to do our work. Creating visualisations in Tableau is basically all about dragging and dropping stuff. What you must know is what to drop and where you have to drop it.

Pay close attention to the image below and identify all of the highlighted parts because we will be using them more often while creating visualisations. Also, we will be mentioning these fields with their respective label or tag or name as <label or tag or name> Tab (eg. Marks Tab, Rows Tab, Columns Tab from the image)



Dimensions & Measures

Dimensions and Measures are Tableau's way of distinguishing Categorical and Numerical features from the dataset.

- Dimensions are categorical features which are responsible for the different dimensions or axis in a graph.
- Measures are the continuous values that represent a datapoint which is plotted along an axis.

1. Bar Chart For State/UT/City Wise Burglary Incidents In India

Steps to Plot:

1. Drag and drop the Category feature from the Dimensions tab to the Columns area.
2. Drag and drop the Burglary Incidence feature from the Measures tab to the Rows area.
3. To add colour to the graphs, drag the feature Category (or Burglary Incidence) to the Color button in the Marks Tab.
4. Add label to the bars by dragging the feature Burglary Incidence into the Label button in the Marks Tab.

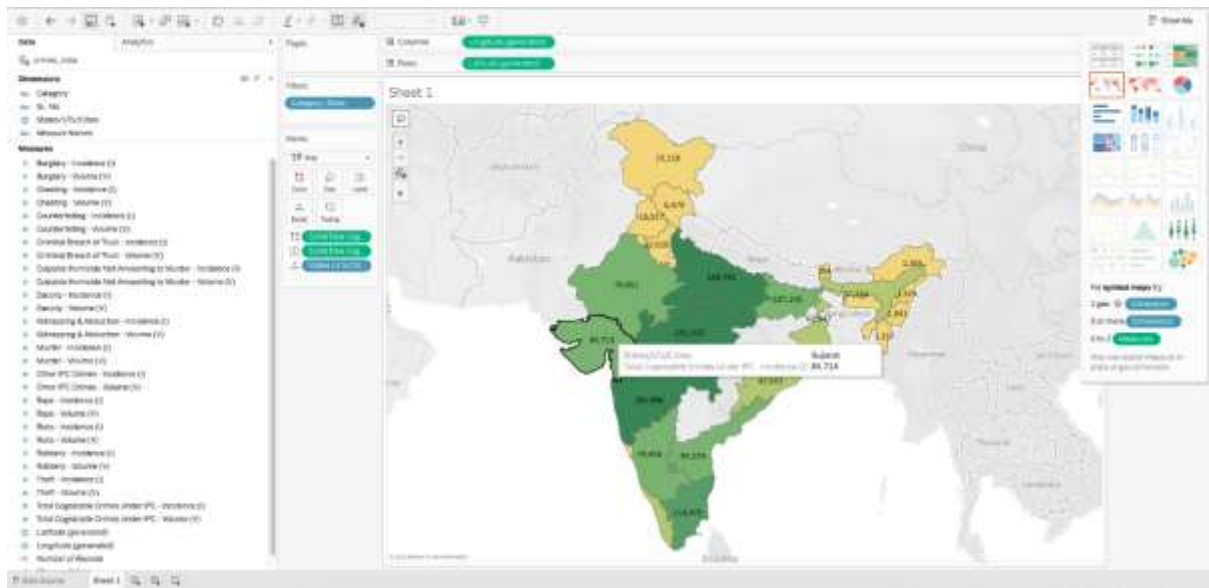


Note:

Tableau automatically aggregates data as a sum. This can be changed by selecting different criteria for aggregation from the drop-down menu of the selected measures.

2. Bubbles Chart For Most Rape Incidents In India State/UT/City Wise

Steps to Plot:



A lot of data still remain unexplored and same goes for Tableau’s visualisation options. Tableau offers a variety of plots which can be used to make really attractive visualisations.

Try out different plots and choose the one that you think is both beautiful as well as informative.

Saving Your Progress

Unfortunately, Tableau Public does not allow users to save a local copy of the workbook that they made. Instead, you can save your work in Tableau cloud which is a public gallery for Tableau visualisations. You can save your visualisations in Tableau Public with File -> Save to Tableau Public option.

Closing Note

Tableau is an invaluable addition to a Data Scientist’s tool kit. Being on top of all the visualisation tools available at present Tableau is still growing both in popularity and in coming up with better and easier ways to demystify data.