Data Engineering Academy

Data Visualisation

Packages - Tableau

TECH TALENT ACADEMY

Session Content

What are data visualisation packages?

Tableau

Importing and Preparing Data

Visualisation

TTA

Data Visualisation Packages

Many companies offer downloadable tools that enable you to import data and to create reports and data visualisations from this data using a dedicated environment.

These software packages help users to manipulate their data at the click of a button, rather than forcing the user to write complicated code.

Pros	Cons
Mostly no code needed	Limited functionality
Simple user interface	Usually available at a cost
Easier to learn	Less powerful

TTA

What's out there?

Power BI

tableau SOFTWARE

Qlik® Sense

alteryx

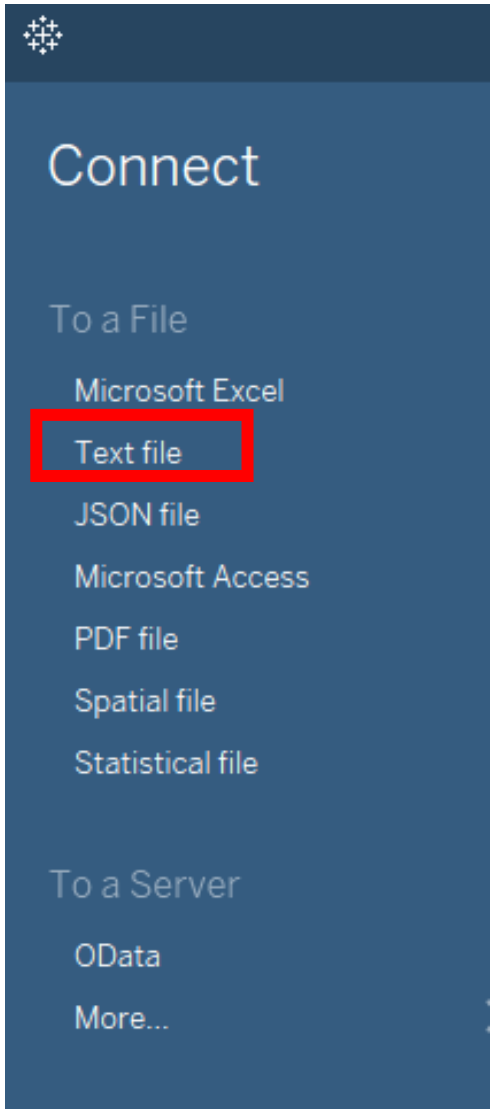
TTA |

A public version that is free to run is available, this will enable you to practice with the software for free. Please do note that some of the features are limited to premium/paid for versions.
https://public.tableau.com/en-us/s/

TTA

Importing data in Tableau

- From the desktop public version of Tableau, you are able to import data from multiple locations.
- In this example we want to import our data from a CSV flat file.
- Select Text file and find the location of the saved CSV .



TTA

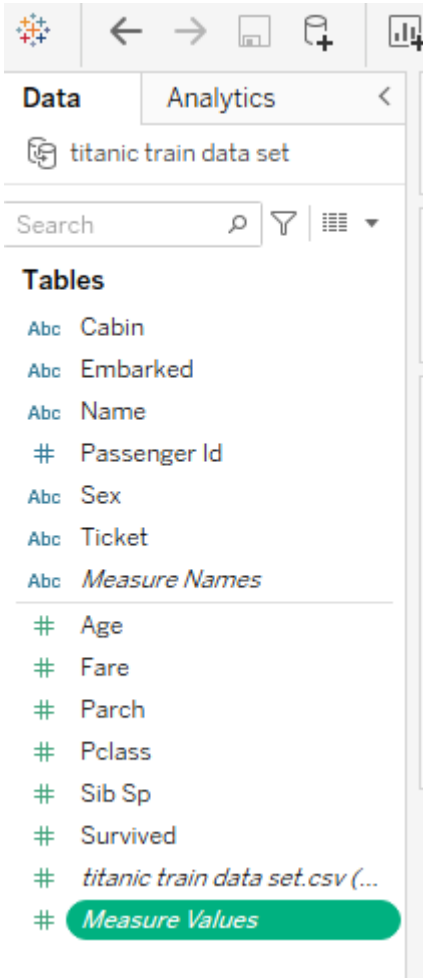
Importing data in Tableau

When data is imported, Tableau will automatically decide if the variables are dimension or measures.

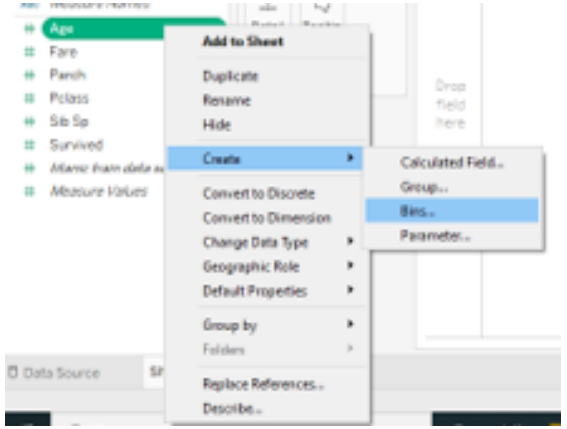
Dimensions = categories e.g. race, sex, and educational level.

Measures = data type expressed in numbers e.g. Age, height, and length.

A variable can be converted into a dimension or measure if necessary.

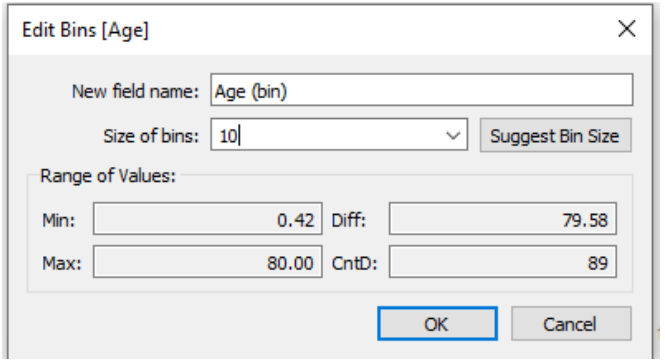


TTA

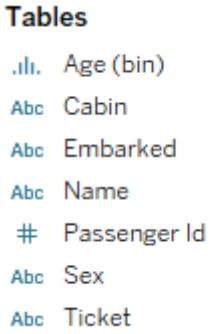


Converting a measure into a dimension

Right click on the variable name > create > bins



Change the name and the size of the bin if necessary

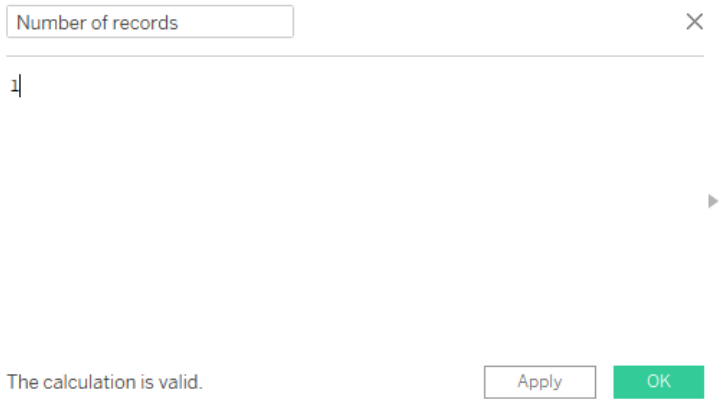


It has now been converted

TTA

Sum of Records

You can create a calculated field which will automatically calculate the SUM of Records by right clicking in the space under the measure names > create calculated field.



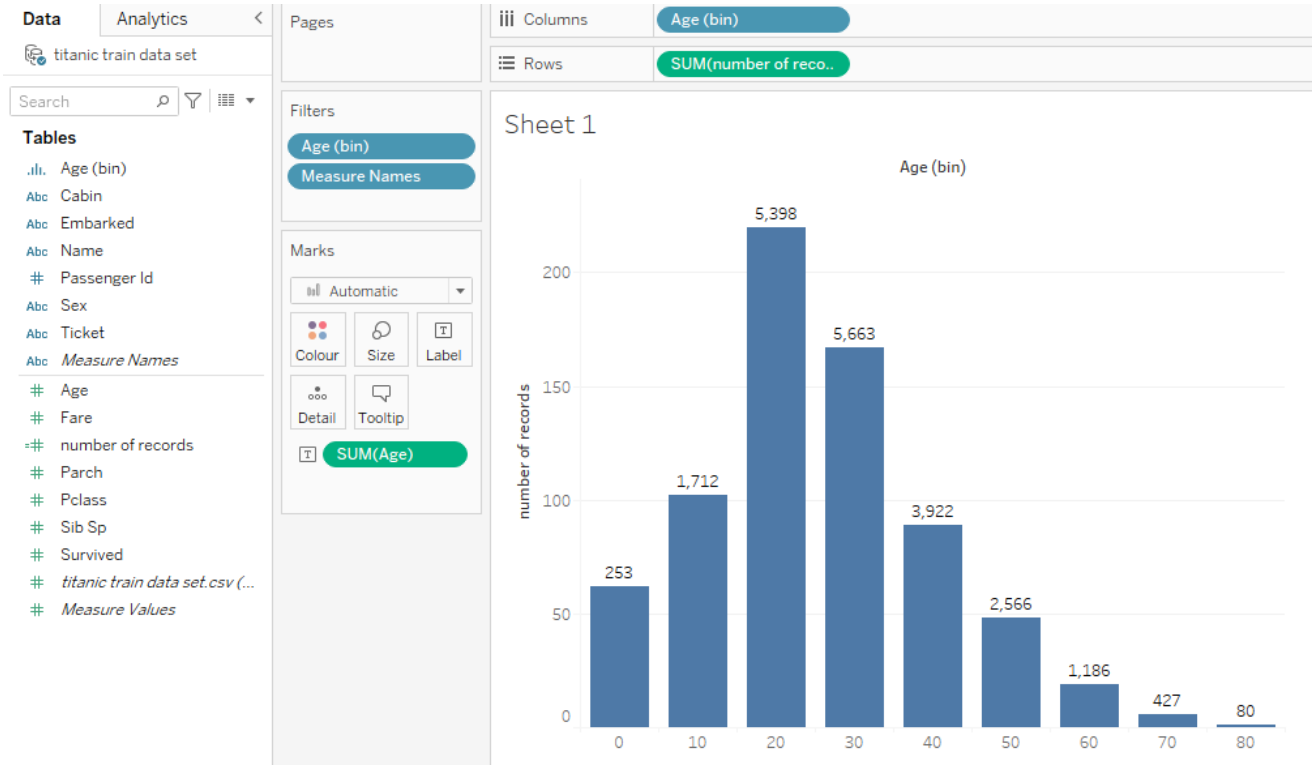
Change the name of the calculated field.

In the text box right the number 1 > click apply > click ok.

TTA

Selecting variables

Select the variables by drag and drop into the columns and rows section to create a basic visualisation.



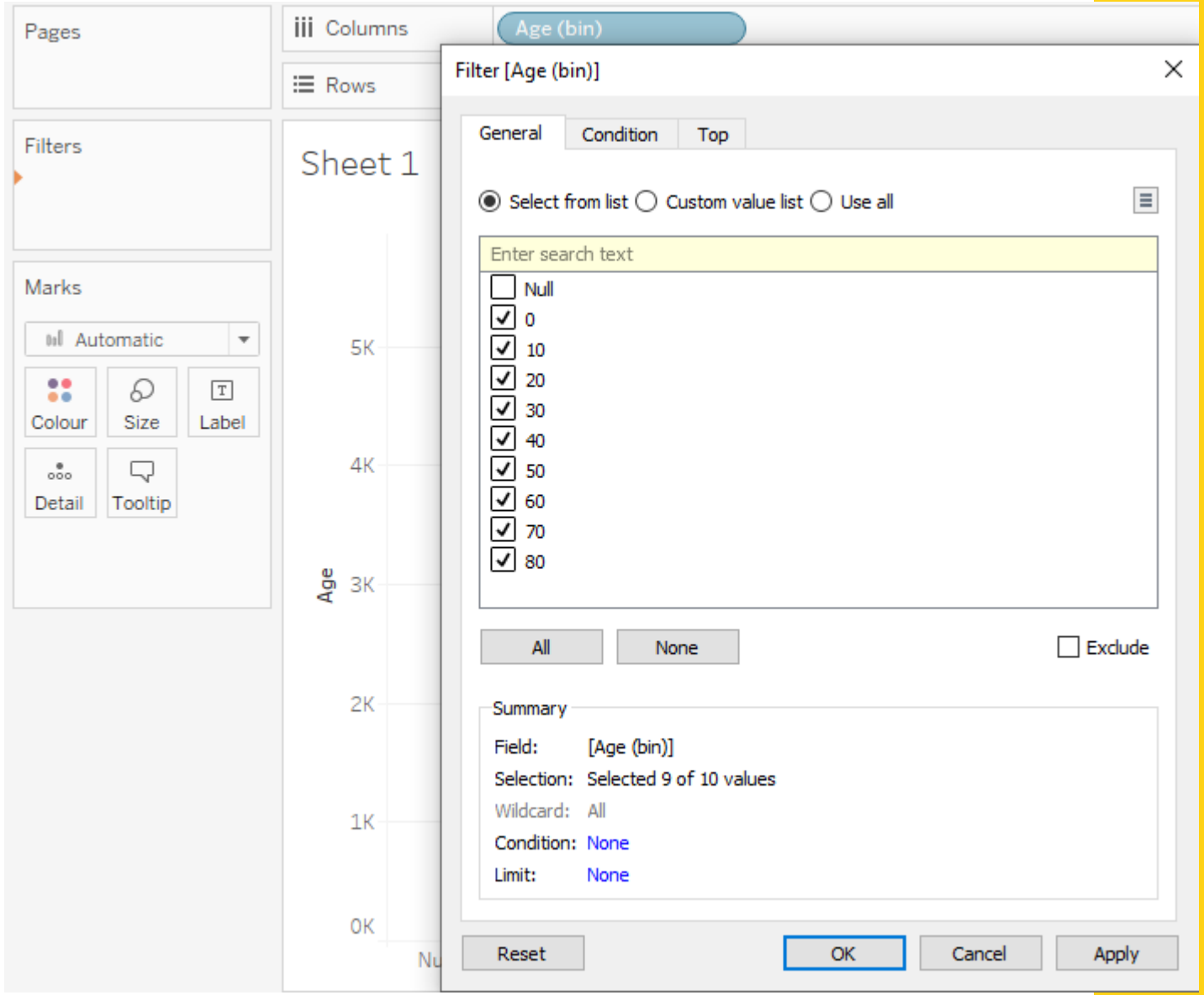
TTA

Null values

CRT and left click to drag and drop the variable name into the filter box.

From the list of variable values make sure the Null box is NOT selected.

This is will ensure Null values from this bin will not be included.

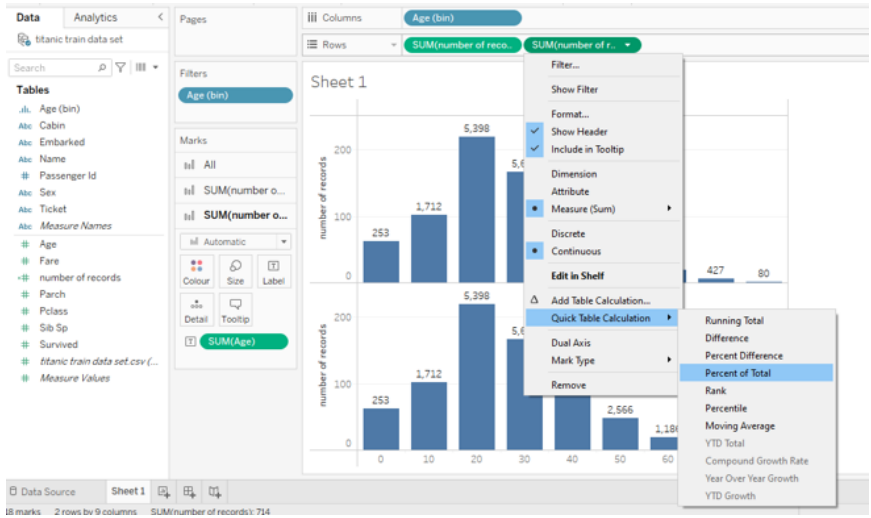


TTA

Adding percentages

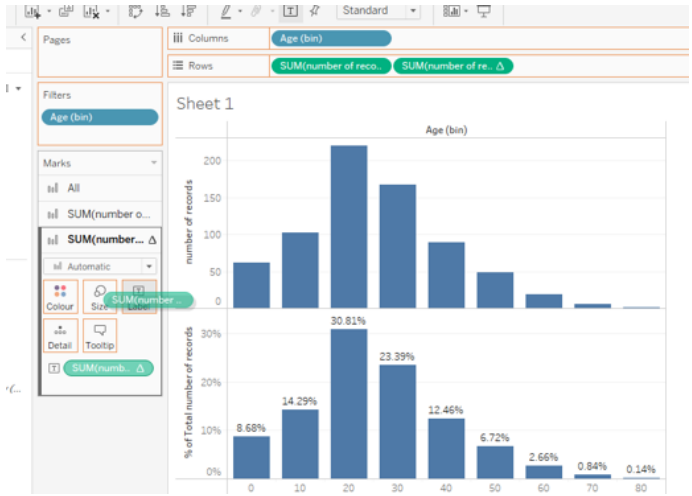
CTRL > left click > drag and drop SUM(number of records) into the rows section

Click on the drop down box > quick table calculation > percent total



CTRL > left click > drag and drop SUM(number of records) into labels

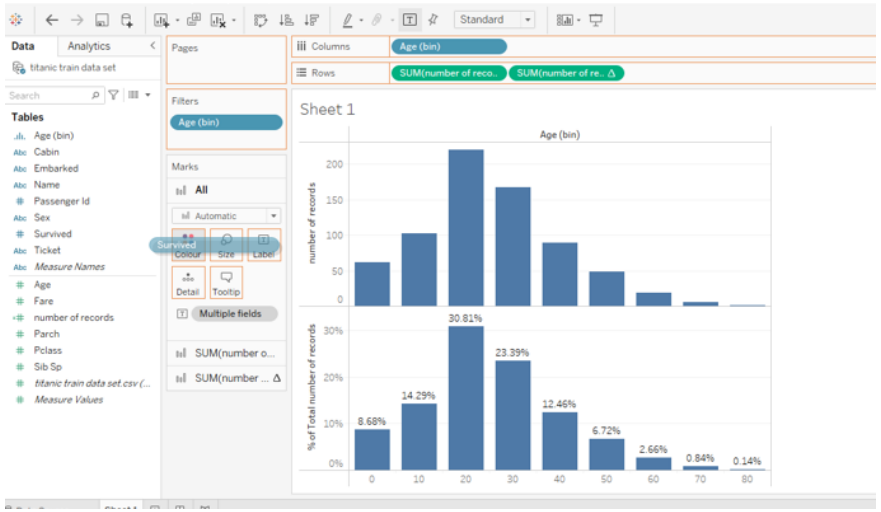
Percentage will now above each bin

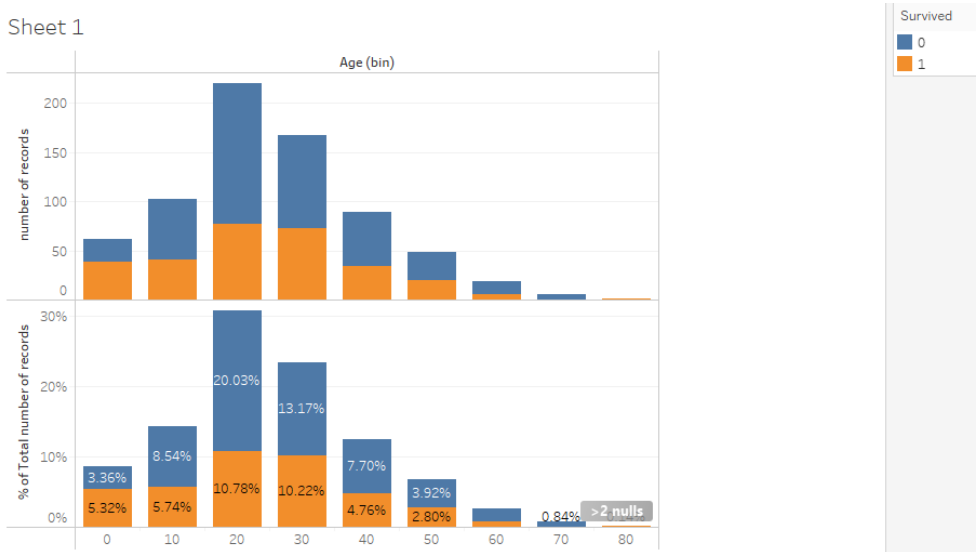


TTA

Adding in a second variable

CRTL > left click > drag and drop survived into colour





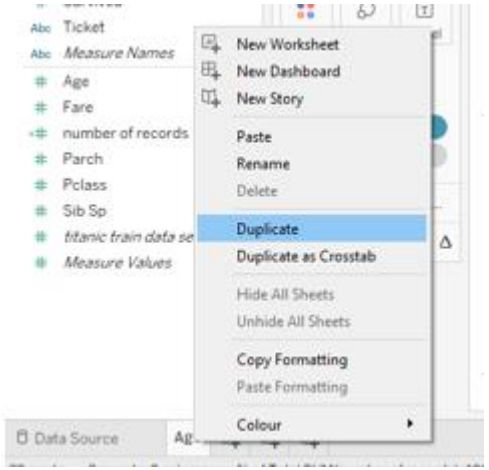
A key will appear in the top right hand corner to indicate whether a person survived

0 = No, 1 = Yes

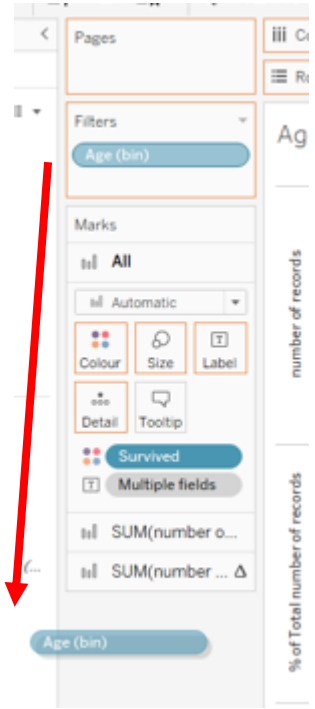
TTA

Reusing worksheets

Right click on the worksheet name > duplicate



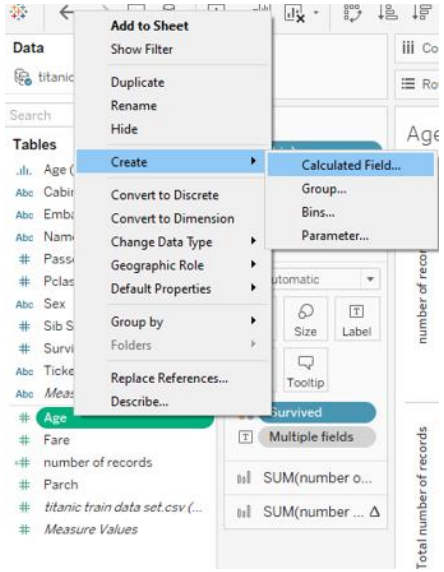
Drag out the bins you no longer need from the filter and column and replace with a new variable

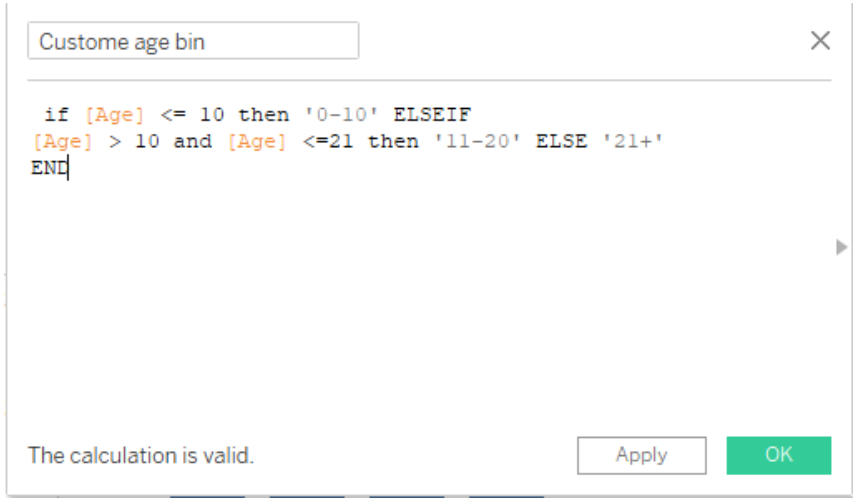


TTA

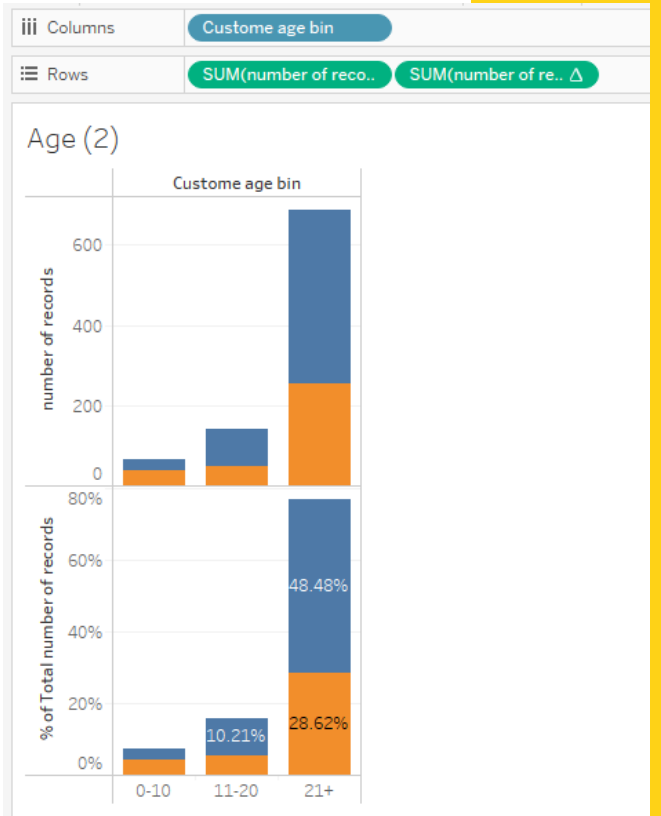
Custom bins

Right click on bin name > create > calculated field





Write an IF statement to set the parameters

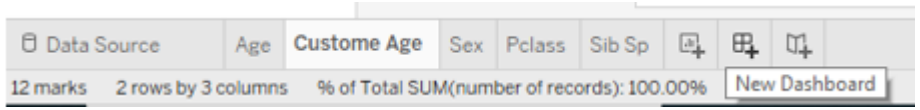


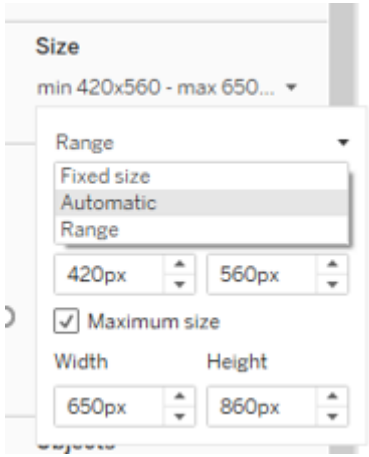
Drag and drop the new customised bin into the column section

TTA

Creating a Dashboard

Select New Dashboard tab at the bottom of the page



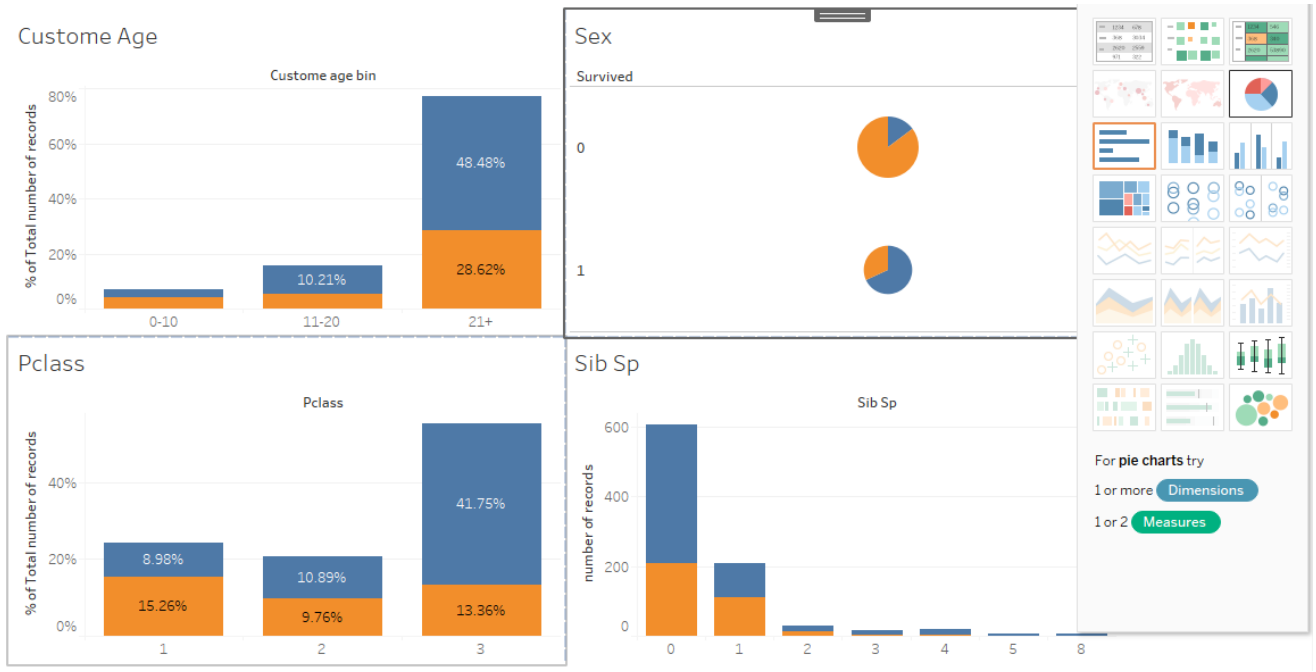


Change size to automatic

Double click on the sheets you want to view on your dashboard

TTA

Creating a Dashboard

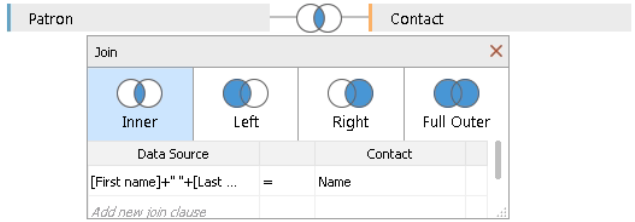


Here you can change the size of each individual graphs by dragging the corners of each graph and change the type of individual graphs.

TTA

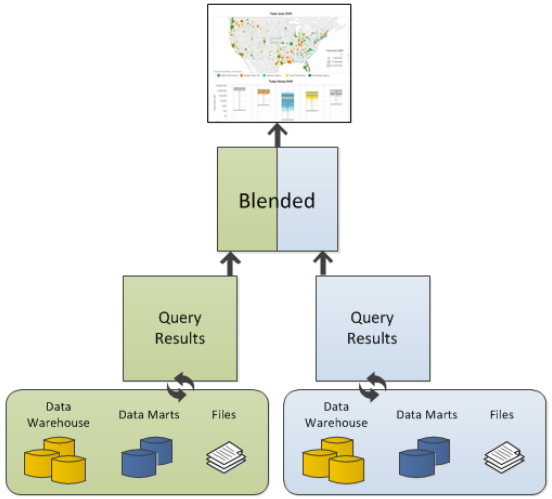
Joins, Relationships and Blends

Joins - join data from multiple sources or join data from different tables in a single source.



Relationships connecting lines created between the logical tables in your data source. Some people affectionately call relationships "noodles", but we usually refer to them as "relationships" in our help documentation.

Blends - unlike relationships or joins, never truly combine the data. Instead, blends query each data source independently, the results are aggregated to the appropriate level.

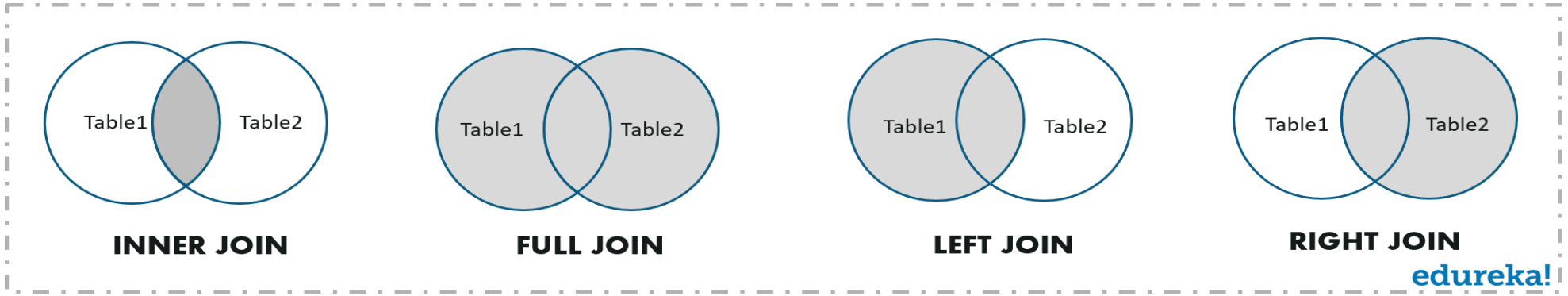


TTA

Types of Joins

JOINS in Tableau are commands which are used to combine rows from two or more tables, based on a related column between those tables. There are predominantly used when a user is trying to extract data from tables which have one-to-many or many-to-many relationships between them.

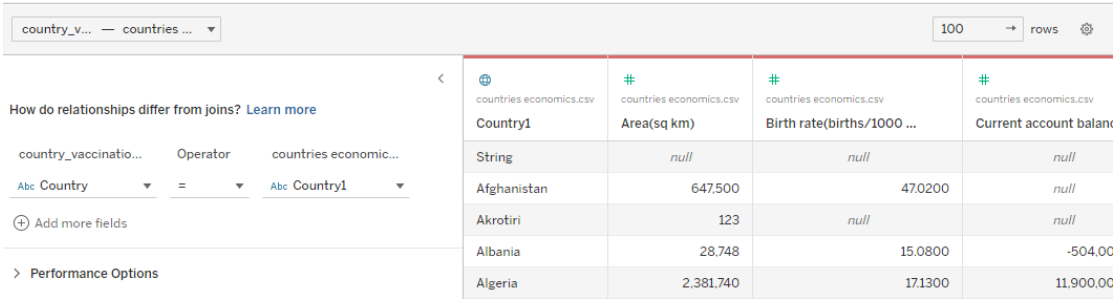
- **(INNER) JOIN**: Returns records that have matching values in both tables
- **FULL (OUTER) JOIN**: Returns all records when there is a match in either left or right table
- **LEFT (OUTER) JOIN**: Returns all records from the left table, and the matched records from the right table
- **RIGHT (OUTER) JOIN**: Returns all records from the right table, and the matched records from the left table

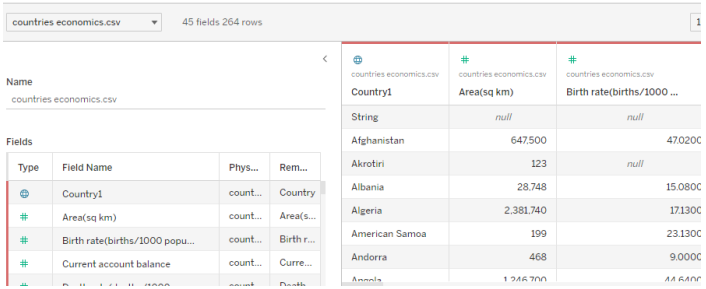


TTA

Creating a Relationship

Drag and drop the tables. Tableau will automatically detect the relationship or you can choose the relationship if necessary.





Double click on the name of the table to view the data is just in that table

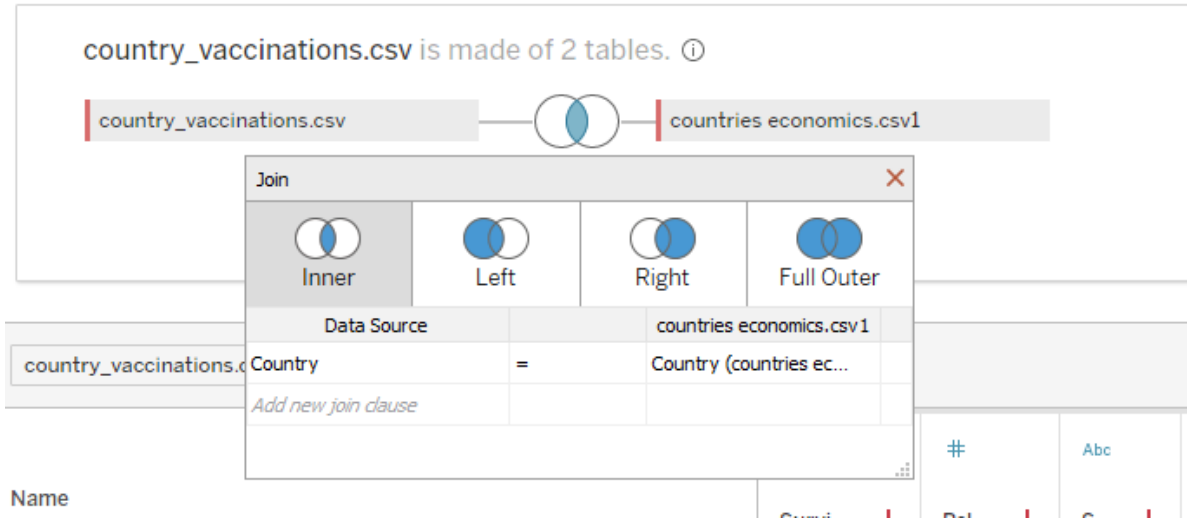
TTA

Joining tables

Double click on the name of first table > drag out the second table

Select the name of the columns on which you would like to join

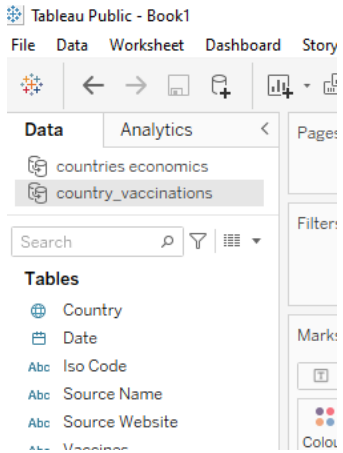
Click on the picture to specify what type of join you would like



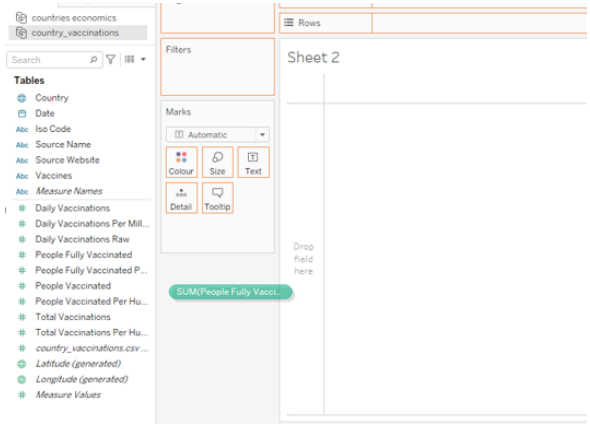
TTA

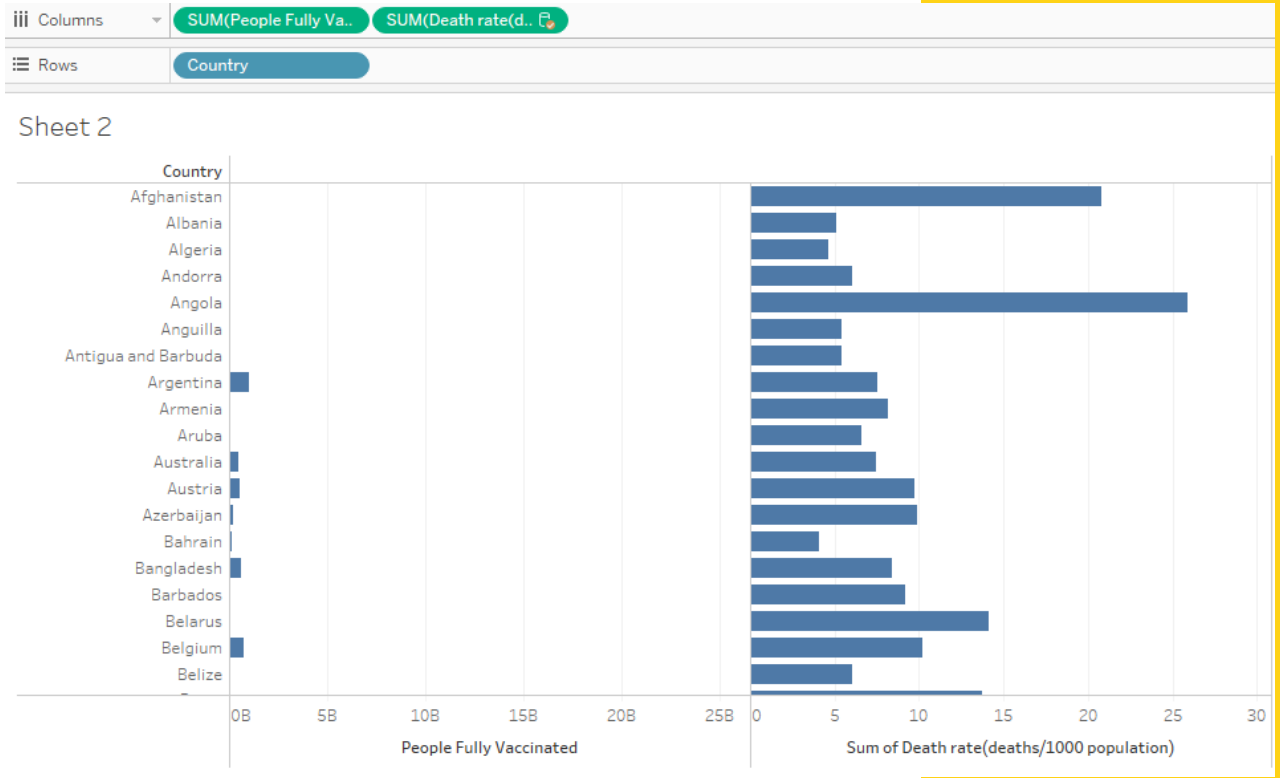
Blending data

Connect to a data source in a new worksheet, add in a second data source CRTL D > Text file > select name of data source



From the data sources, drag and drop chosen variables onto the worksheet





TTA

Home Learning Tasks

Create Your Own Dashboard on Tableau

Download Tableau and create a dashboard of your own!

Tableau – https://public.tableau.com/en-us/s/

Feel free to use a dataset of your choosing. You might have your own or you are welcome to search Kaggle for appropriate datasets:

https://www.kaggle.com/datasets

TTA


TECH TALENT
ACADEMY