

Fun with Tableau_Advanced

Rao Katie Xu

Environment: Win10 64 bit (Parallel Desktop on Mac OSX Yosemite 10.10.5)

Tableau: <http://www.tableau.com/products/desktop/download>

Project 1:

Datasets: <https://www.superdatascience.com/tableau/>

Download: [AmazingMartEU2.xlsx](#)

Aim: Find the profit margin of each city.

From the table, we can see the same Order ID with different product names, because the customer might buy several things in one order.

Order ID	Product Name	Discount	Sales	Profit	Quantity	Category	Sub-Category
IT-2011-3647632	Enermax Note Cards, Premium	0.5	\$45.00	-\$26.00	3	Office Supplies	Paper
ES-2011-4869686	Dania Corner Shelving, Traditional	0	\$854.00	\$290.00	7	Furniture	Bookcases
ES-2011-4939443	Binney & Smith Sketch Pad, Easy-Erase	0	\$140.00	\$21.00	3	Office Supplies	Art
IT-2011-2942451	Boston Markers, Easy-Erase	0.5	\$27.00	-\$22.00	2	Office Supplies	Art
IT-2011-2942451	Eldon Folders, Single Width	0.5	\$17.00	-\$1.00	2	Office Supplies	Storage

The screenshot shows two data sources in Tableau:

- ListOfOrders:** Contains columns: Order ID, Order Date, Customer Name. Data rows include BN-2014-9998478, BN-2014-9995192, BN-2014-9982544, BN-2014-9978138, BN-2014-9956954, BN-2014-9956190, BN-2014-9936235, BN-2014-9900559, BN-2014-9861831, BN-2014-9778057, BN-2014-9660071, BN-2014-9643770, BN-2014-9635866, BN-2014-9612123, BN-2014-9608893.
- OrderBreakdown:** Contains columns: Order ID, Product Name. Data rows include AZ-2011-1029887, AZ-2011-1029887, AZ-2011-107716, AZ-2011-1087704, AZ-2011-1087704, AZ-2011-114253, AZ-2011-1116129, AZ-2011-1137571, AZ-2011-1137571, AZ-2011-1174243, AZ-2011-122598, AZ-2011-1229073, AZ-2011-1229073, AZ-2011-1229073.

We notice ListOfOrders has unique Order ID and OrderBreakdown has duplicate Order ID, so we want to join two tables. We drag ListOfOrders and OrderBreakdown to the right area and a crossed circle pop up, click the circles, we can see if it is left join or inner join. Also we need to change the title to ListOfOrders + OrderBreakdown.

The screenshot shows the Tableau Join dialog box:

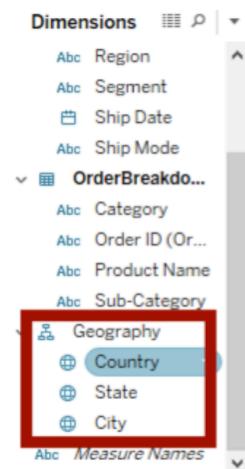
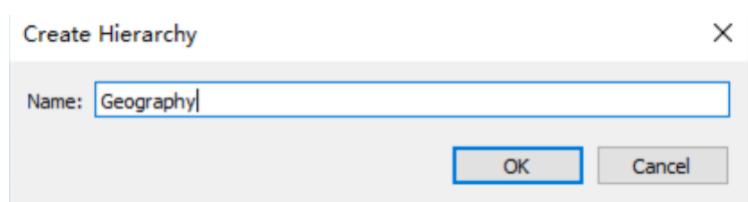
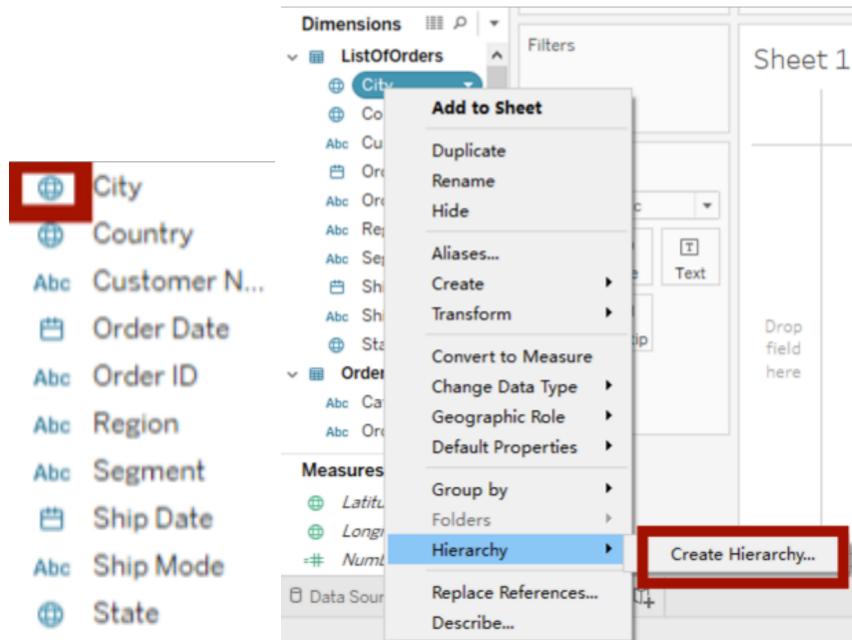
- Join Type:** Left (highlighted with a red box).
- Tables:** ListOfOrders and OrderBreakdown.
- Key:** Order ID (highlighted with a red box).
- Condition:** Order ID = Order ID (highlighted with a red box).

TWB - XML file for your Tableau Workbook, contains all the selections and layout you've made. It does not contain data. These tend to be very small.

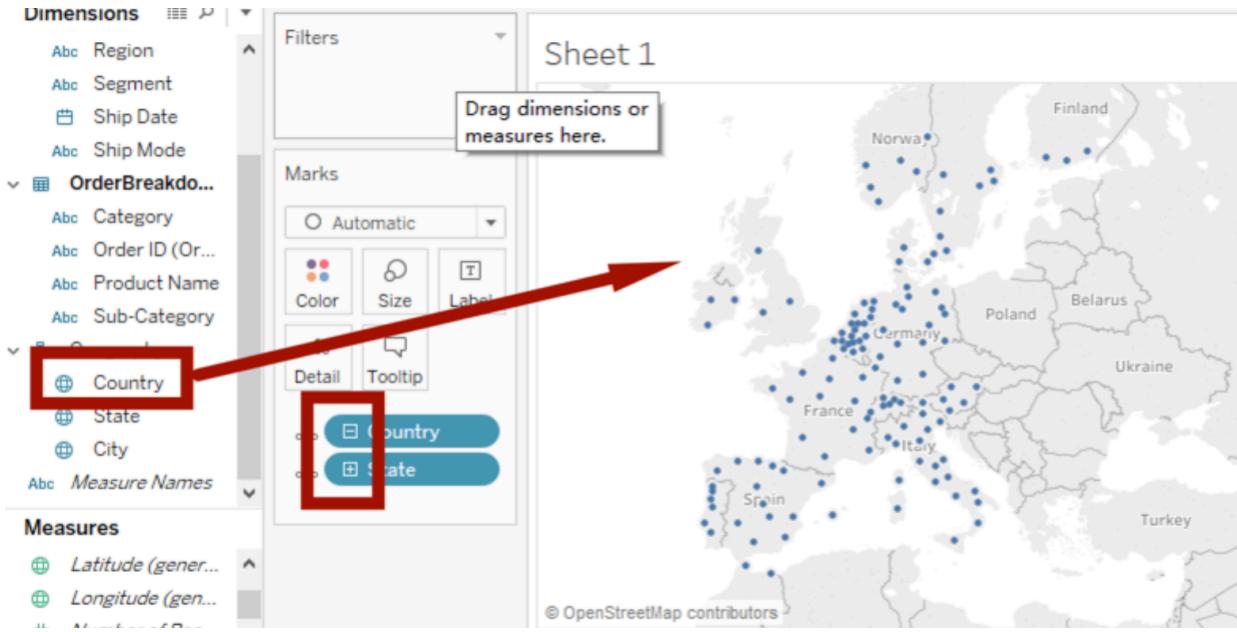
TWBX - zipped file that contains the TWB as well as data used by TWB in an extract

Now we should draw map, work with hierarchy and get profit margin.

First, remove hierarchy structure and build our own country-state-city hierarchy. Notice this **global sign** means geography, **Abc** means words, **calendar sign** means date. Right click the **City** and find **Hierarchy** -> **Create Hierarchy**, write the name **Geography**.



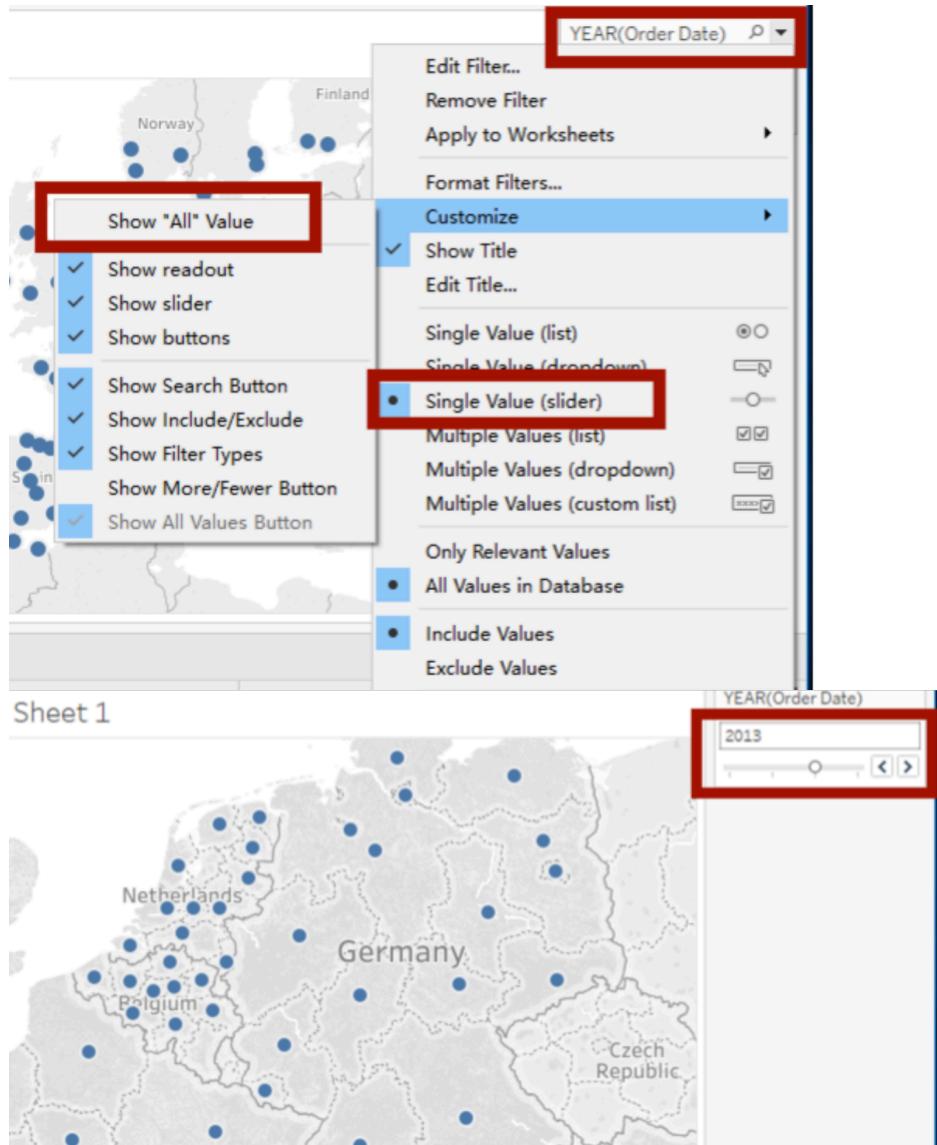
Now we have our own natural hierarchy. Now we can drag the country into our work space. The blue dots are country. And if we click the "+" sign on blue hierarchy structure, we can see state blue dots and city blue dots.



Now the map is ready, we want to see sales on 2012. We drag the **OrderDate** to the filter, then pop up the Filter Field, we click the # Years.

Click the **Year** in the **Filters**, and choose **Show Filter**, it will be more convenient.

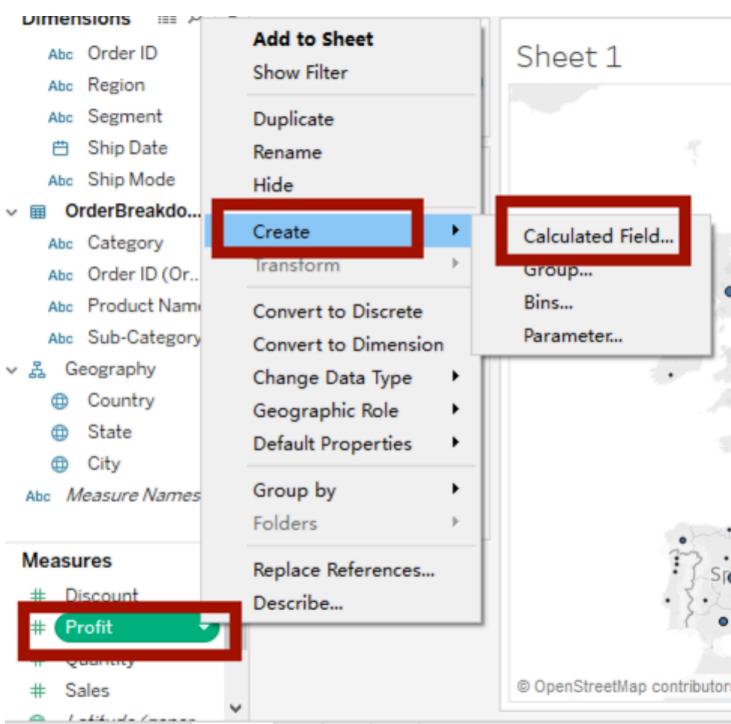
Go to the YEAR(Order Date) panel, right click, choose **Customize**-> uncheck **Show "All" Value**, and choose **Single Value(Slider)**. Now we can play the slide bar to see the changes (dots are moving) by years.



Now we want to see the sales in 2012. Drag **Sales** into **Size**. We can see blue circles are sales. Then add **Border** to **black color** to make it decent.



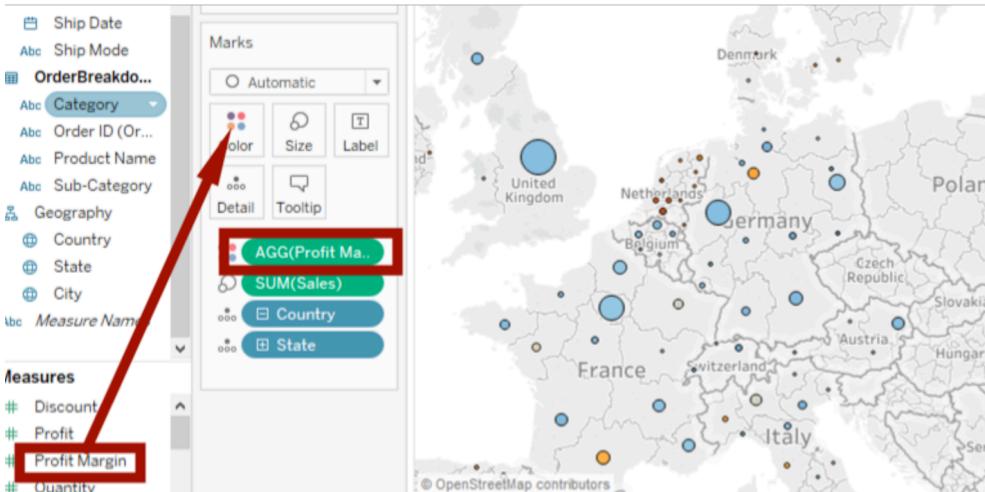
How to find the profit margin? Right click Profit -> Create -> Calculated Field



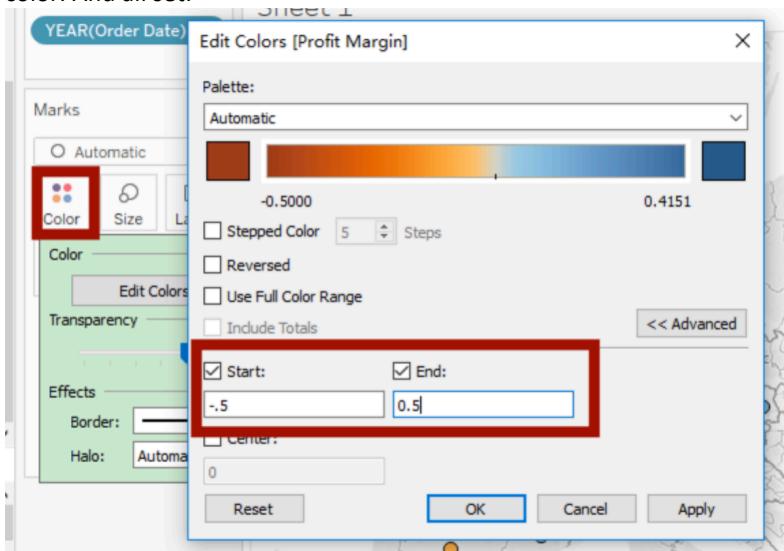
Put name Profit Margin, and the formula $\text{SUM}([\text{Profit}])/\text{SUM}([\text{Sales}])$

The screenshot shows the 'Calculated Field' dialog box. The 'Name' field contains 'Profit Margin'. The 'Formula' field contains the formula $\text{SUM}([\text{Profit}])/\text{SUM}([\text{Sales}])$. At the bottom, a message says 'The calculation is valid.' There are 'Apply' and 'OK' buttons at the bottom right.

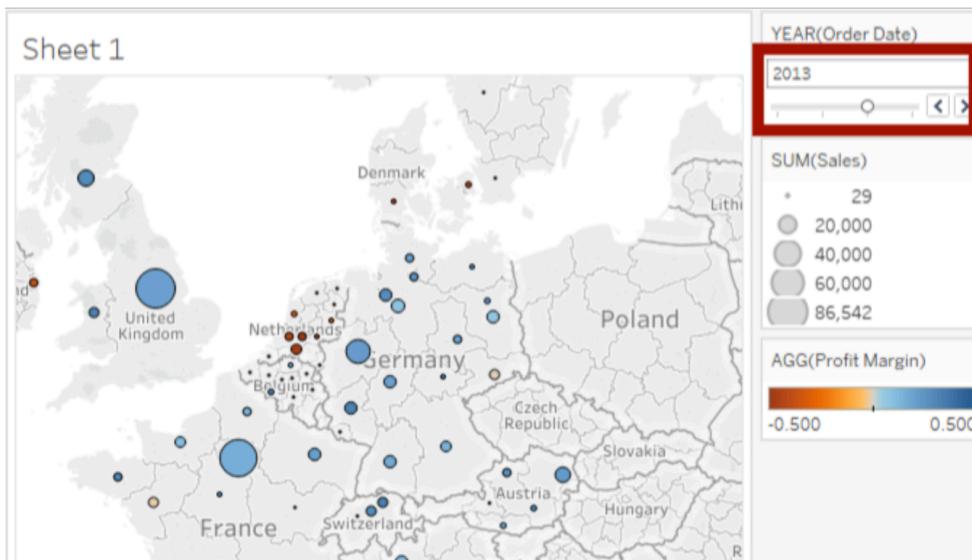
Drag Profit Margin to the Color, now we can see AGG(Profit Margin) which means aggregate Profit Margin.



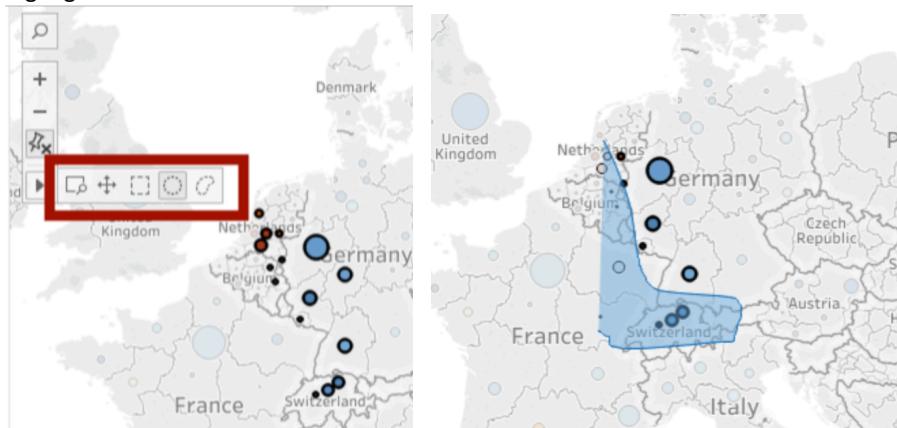
We can change **color**, just click **Color**, and click **Edit Colors->check Start and End** and put **-0.5** and **0.5** to limit the color. And all set.



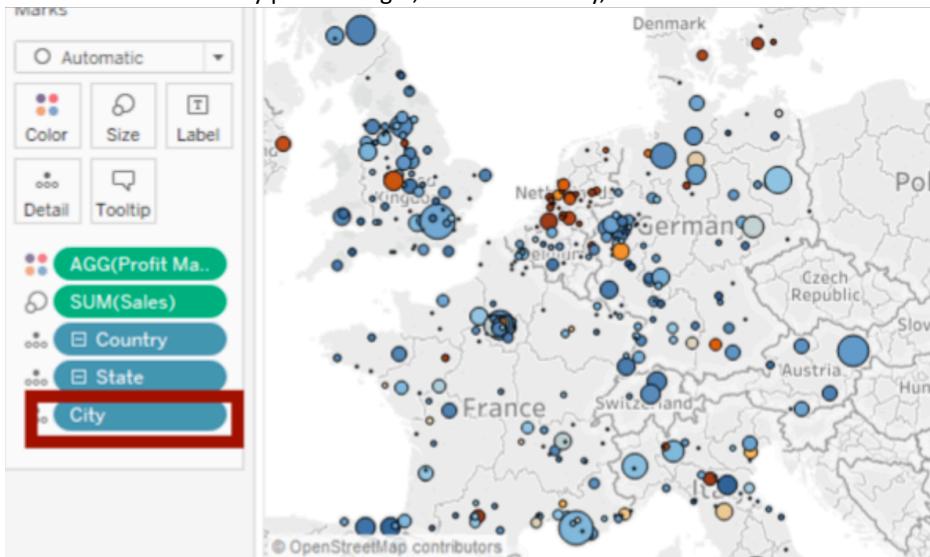
Now we can play with year bar to see the profit margin in 2013 or 2014.



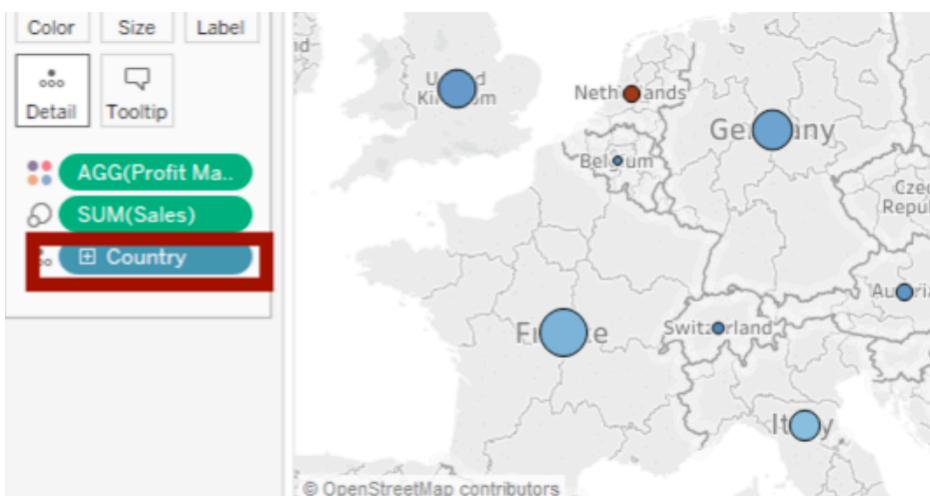
And click arrow to highlight specific area, you can choose **radial selection** or **lasso selection** tool to select area to highlight.



if we want to check city profit margin, we can click City, then more data are shown on map,



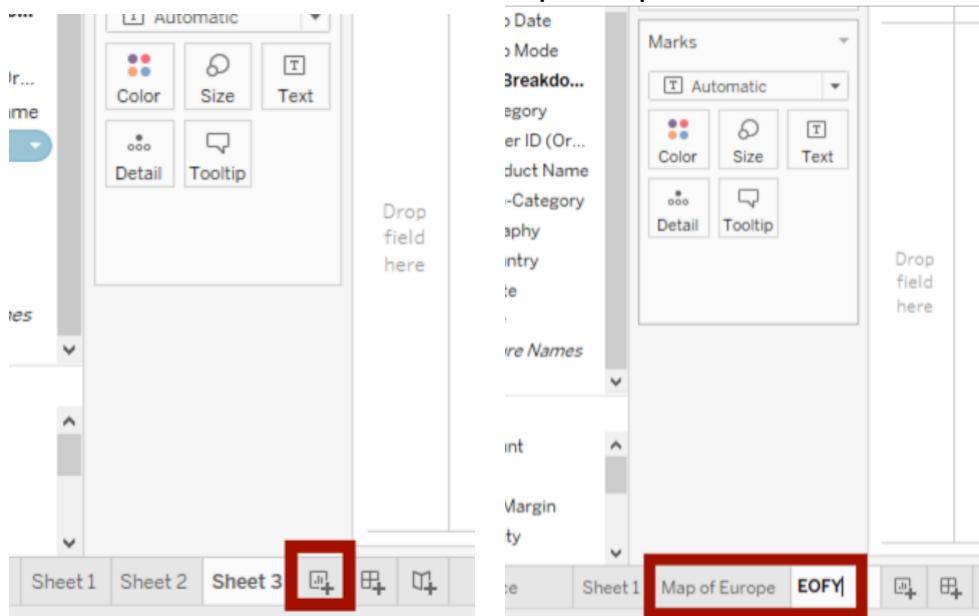
or only Country.



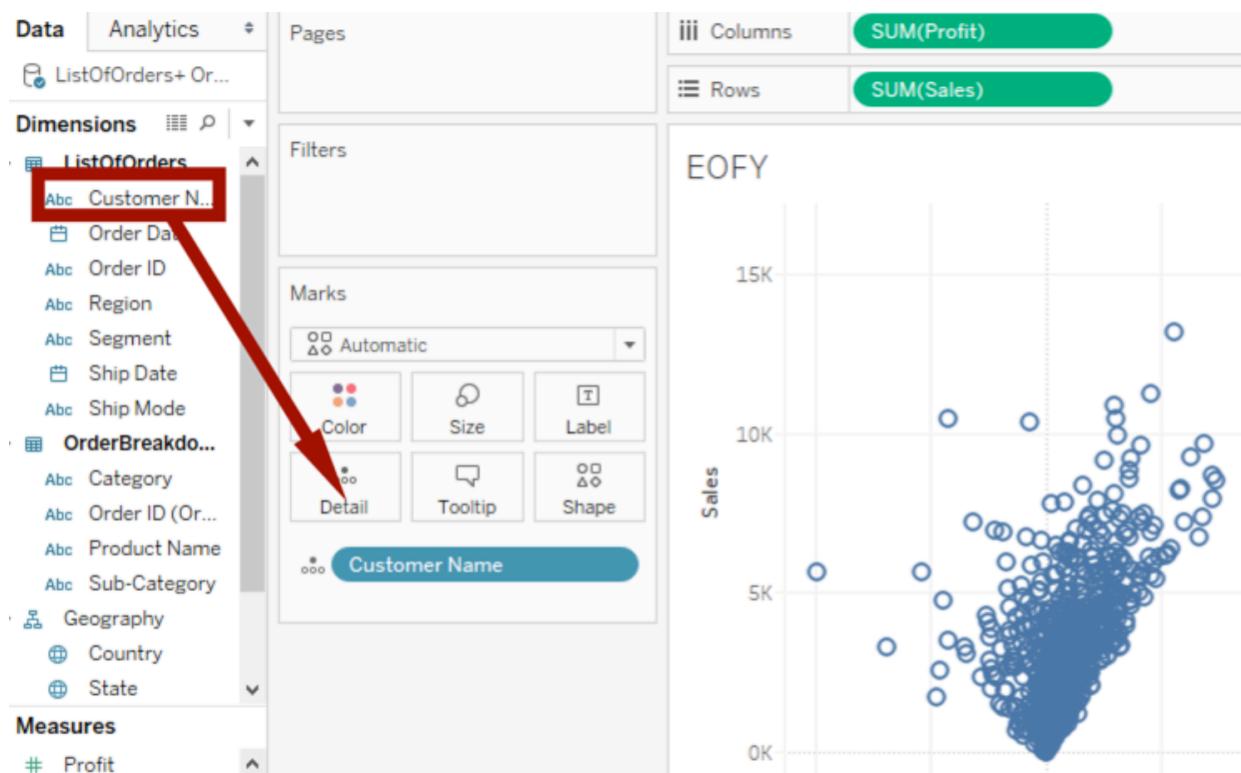
Project 2

Scatter plot, applying filters to multiple worksheets

First we create more worksheets and name it **Map of Europe** and **EOFY**



On EOFY worksheet, drag **Profit** to Columns, **Sales** to Rows. We find only one dot in the graph, now we need to drag **Customer Name** to the **Detail**. Now we can see the scatter plot.



Now go back to Sheet1, right click **YEARS** in the Filters, choose **Apply to Worksheets-> All Using This Date Source**, then all worksheets will share the same data source. If we go to EOFY worksheet, it will show the same data source.

Dashboard Story Analytics

Show Filter Clear Filter Add to Context

Apply to Worksheets

Create Set...

Year 2015
Quarter Q2
Month May
Day 8
More

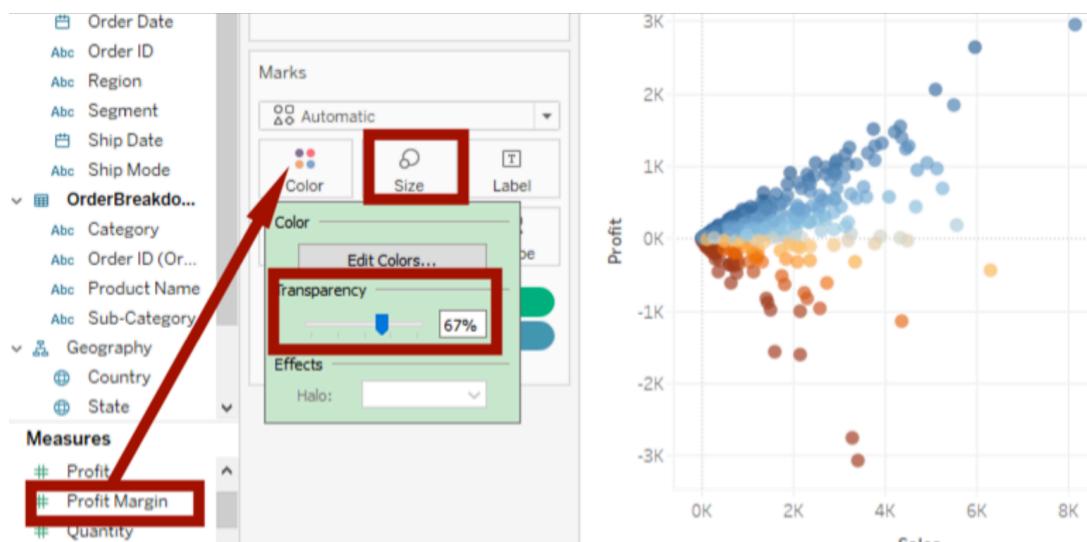
Pages

Filters

Marks

YEAR(Order Date..)

Now we are making the scatter plot prettier. Drag **Profit Margin** to the color, add some **transparency**, change **size** and **shape**.



Now we need to put several worksheets in the same window, so when we need to analyze them we can see both graphs together. Click the **grid sign** below and we can see the new dashboard show up with worksheets on the left panel.

Sheets

- Sheet 1
- Map of Europe
- EOFY

Objects

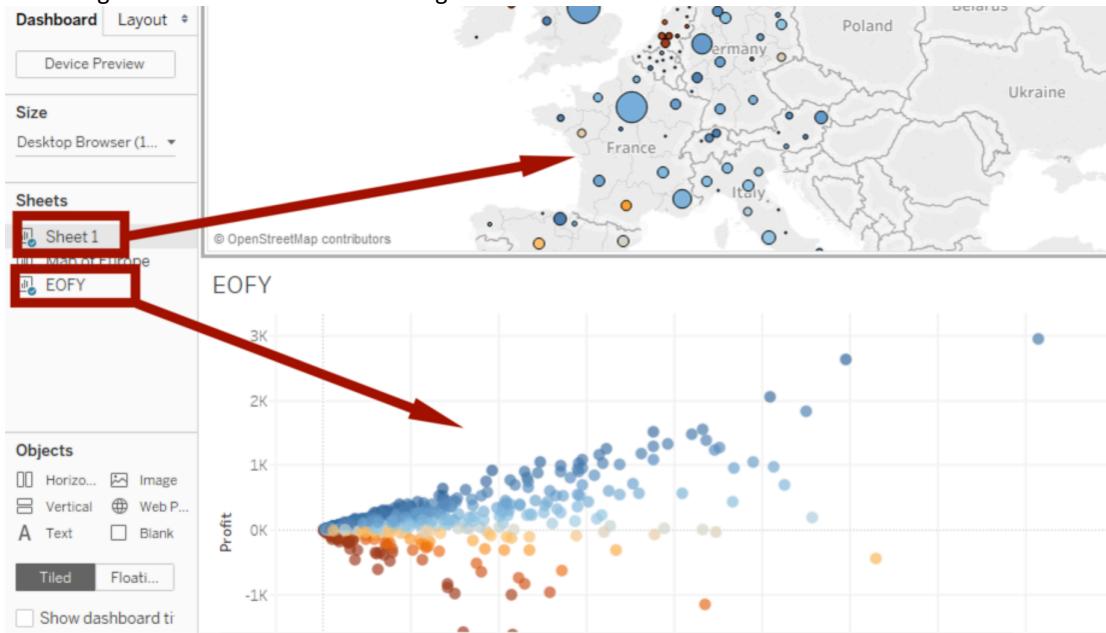
- Horizontal Image
- Vertical Web Page
- Text
- Blank

Tiled Floating

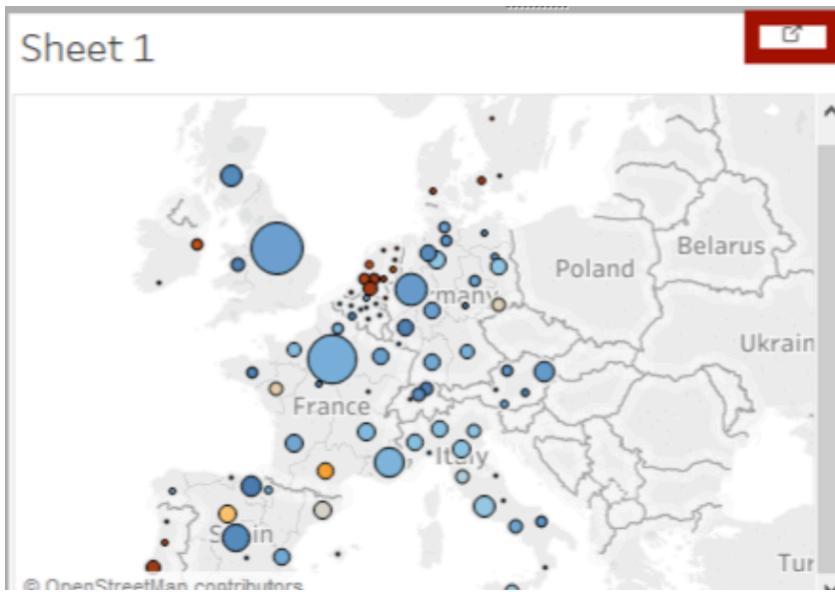
Show dashboard tiles

Data Source Sheet 1 Map of Europe EOFY Dashboard 1

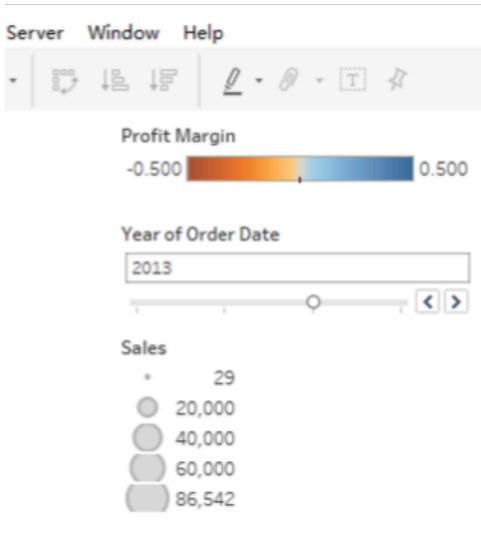
Now drag the **sheet1** and **EOFY** to the right area



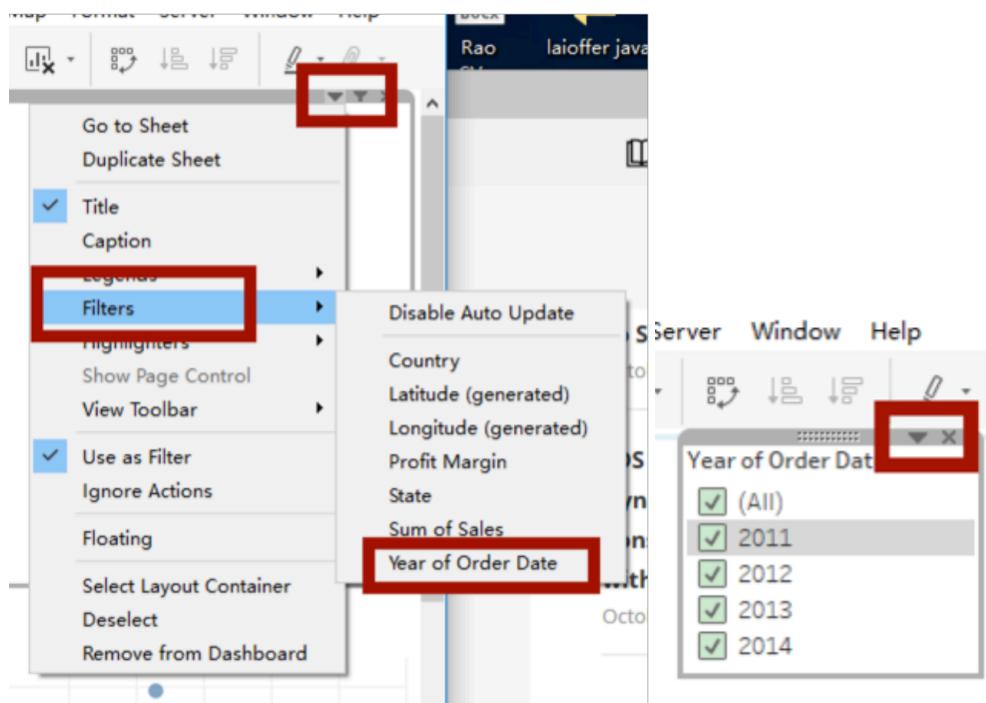
Since the map is too big, we can click little arrow on the right top of the map and adjust the size of the map.



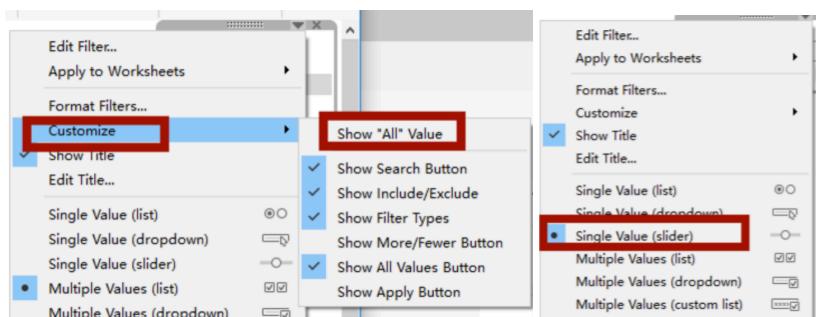
The following is the right side panel, so we can adjust the year and see the sales.



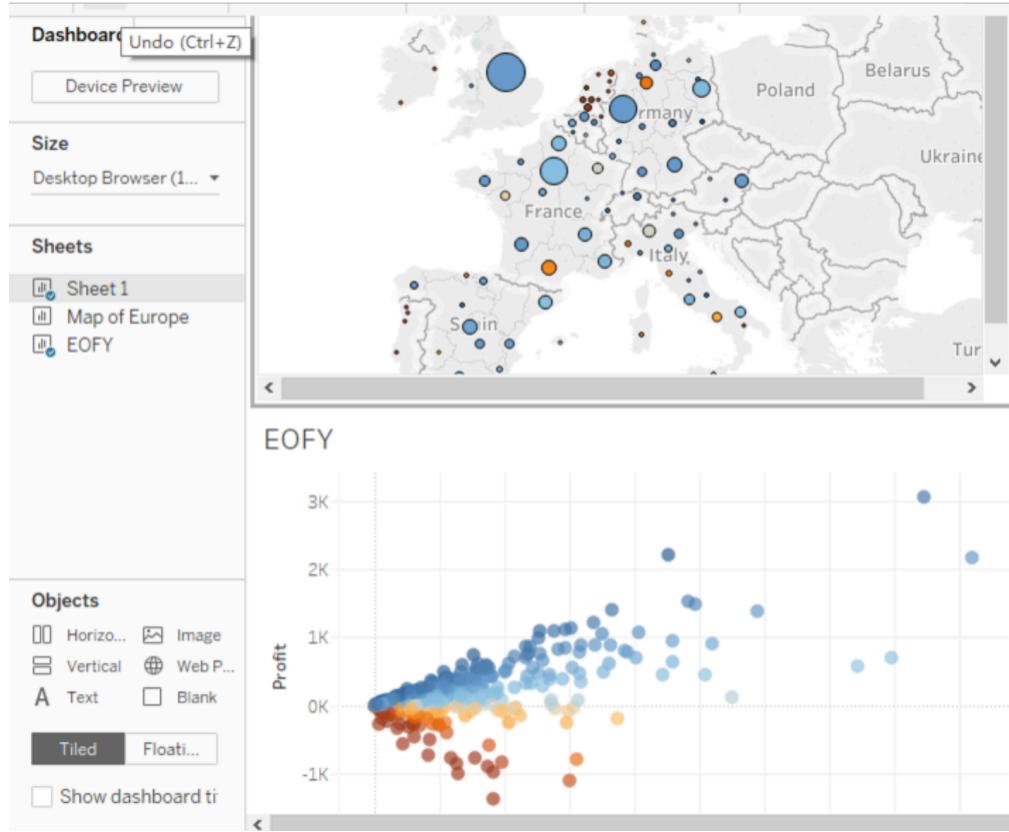
However, what if we do not have this panel, how to get this? We click the tiny triangle on the right top corner of the panel and choose **Filters-> Year of Order Date**, now the **Year of Order Date** popped up.



Still click the tiny triangle on the right top corner, we can see choose **Customize-uncheck show "All" Values**, and also choose **Single Value(Slider)**. When we slide year, both graphs will change by years.



Now we have the dashboard for all the graphs



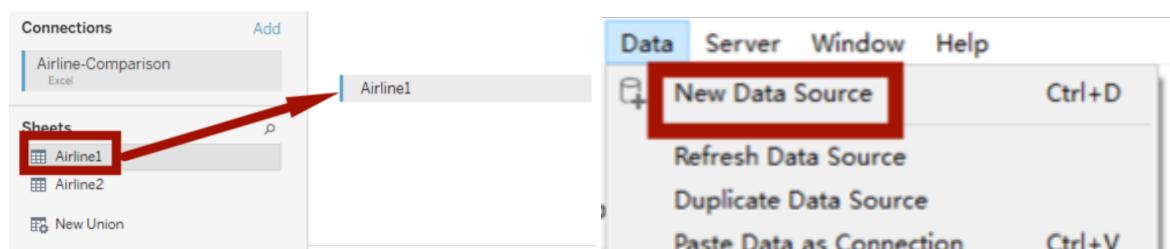
Project 3:

Aim: Use Tableau blending (left join) to find Joined Flight

Datasets: <https://www.superdatascience.com/tableau/>

Download: [Airline Comparison.xlsx](#)

First, we drag the Airline1 to the workspace, then we click Data on the top menu, choose New Data Source, and still choose the same file **Airline Comparison.xlsx**, drag the Airline2 to the workspace.

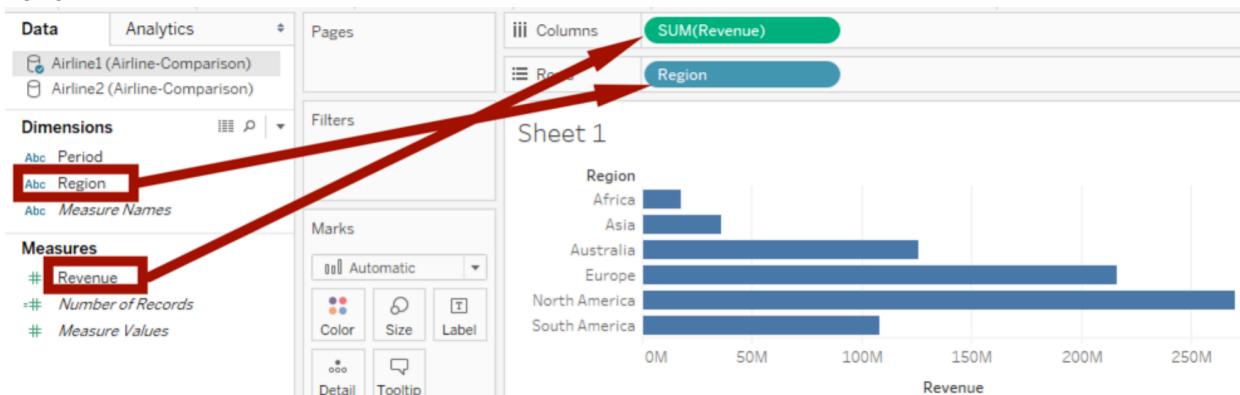


The screenshot shows the Tableau Connections pane. At the top, there's a 'Connections' section with an 'Add' button. Below it is a 'Sheets' section containing 'Airline1' and 'Airline2'. A red arrow points from the 'Airline2' entry in the 'Sheets' list to its corresponding entry in the 'Connections' pane.

Now go back to the **sheet1** tab on the bottom, we will see two data source **Airline1**, **Airline2** both show up under the Data.

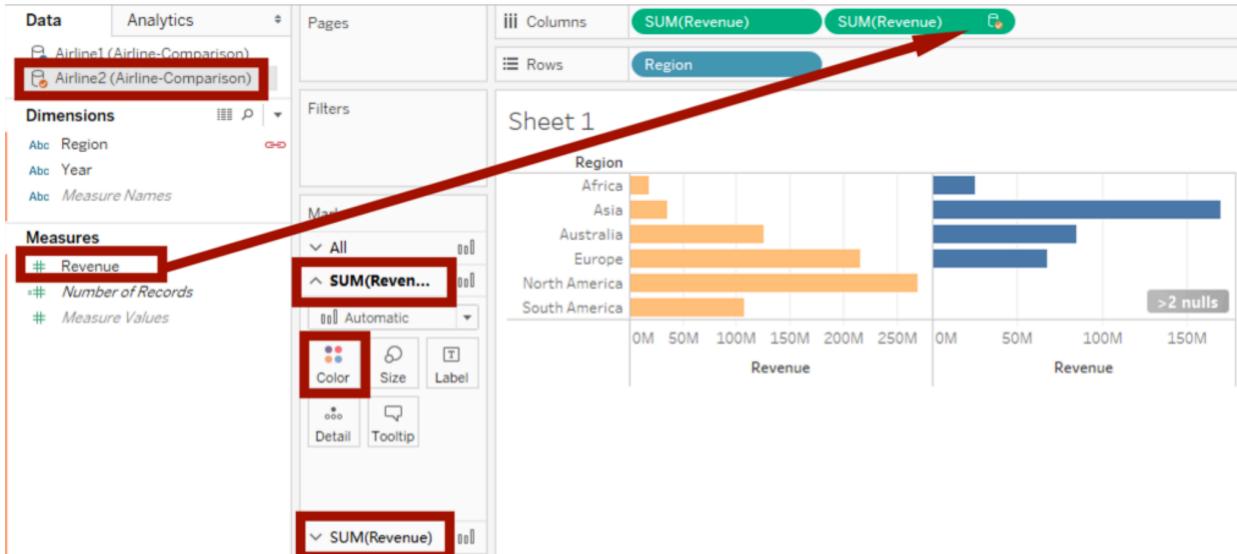
The screenshot shows the Tableau interface with the 'Data' shelf selected. It lists 'Airline1 (Airline-Comparison)' and 'Airline2 (Airline-Comparison)'. Both items are highlighted with a red box. Below the shelf, the 'Dimensions' and 'Measures' sections are visible.

Now we're going to see revenue split by regions. Click Airline1, and just drag **Revenue** to **Columns**, **Region** to **Rows**.

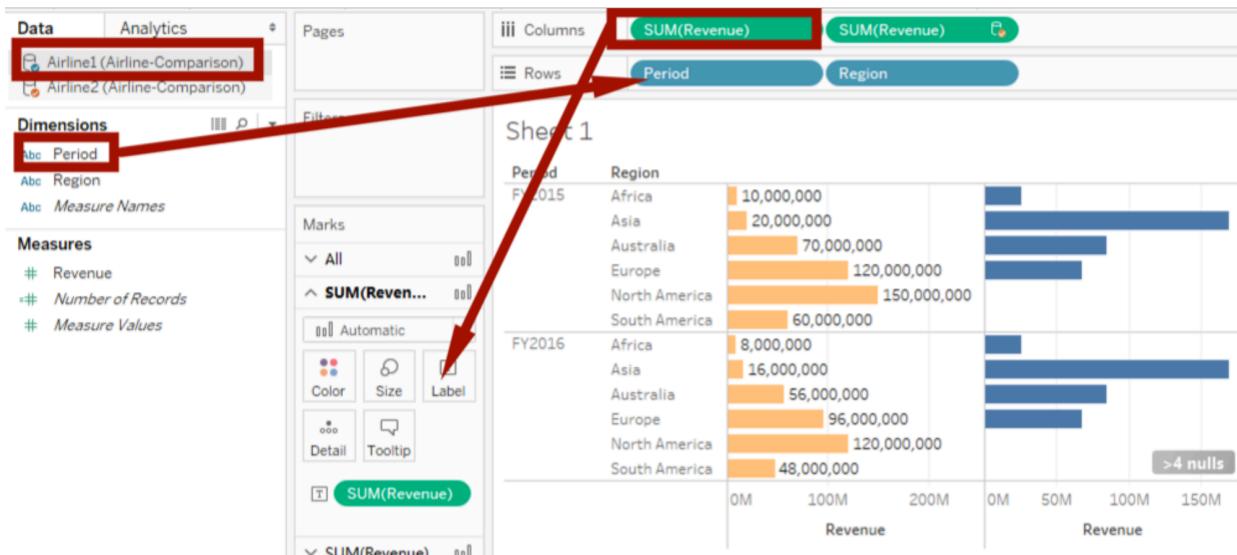


Then click **Airline2** to switch data source (We will notice there are orange lines on the left). We can drag the **Revenue** to the **Columns**, too. Now we have two bar graphs one of which shows “>2 nulls”.

Also in the **Marks** panel, we can see **SUM(Revenue)** for the left bar graph and another one is for the right bar graph. Just use Color to change the color to make it prettier.



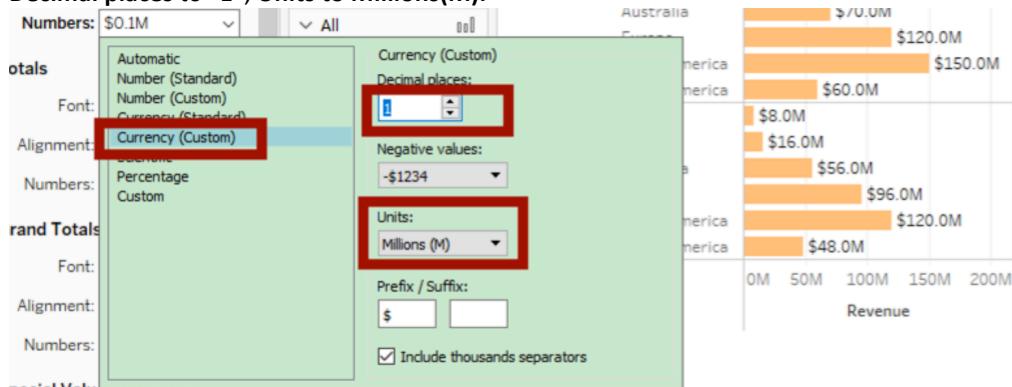
Now we need to go back to **Airline1** and drag **Period** to the **Rows**. (If we forget to go back to Airline1 and still in Airline2, then there is no Period, only Years. When you try to drag Years to Rows. Null will show up.) And then we want to see the revenues. Just **Ctrl+SUM(Revenue)**, and drag it to the **Labels**



Go to Marks->SUM(Revenue), right click, choose Format, change Numbers to Currency(Standard), now we can see the dollar sign.

The screenshot shows the Tableau interface with a bar chart titled 'Region'. The 'Marks' shelf on the left has several items: 'Default', 'Totals', 'Grand Totals', and 'Special Values (eg. NULL)'. A context menu is open over one of the bars, with the 'Format...' option highlighted. Other options in the menu include 'Include in Tooltip', 'Dimension', 'Attribute', 'Measure (Sum)', 'Discrete', 'Continuous', 'Edit in Shelf', 'Add Table Calculation...', 'Quick Table Calculation', and 'Remove'. The chart area shows revenue values for different regions, with the first bar labeled '\$10,000,000.00'.

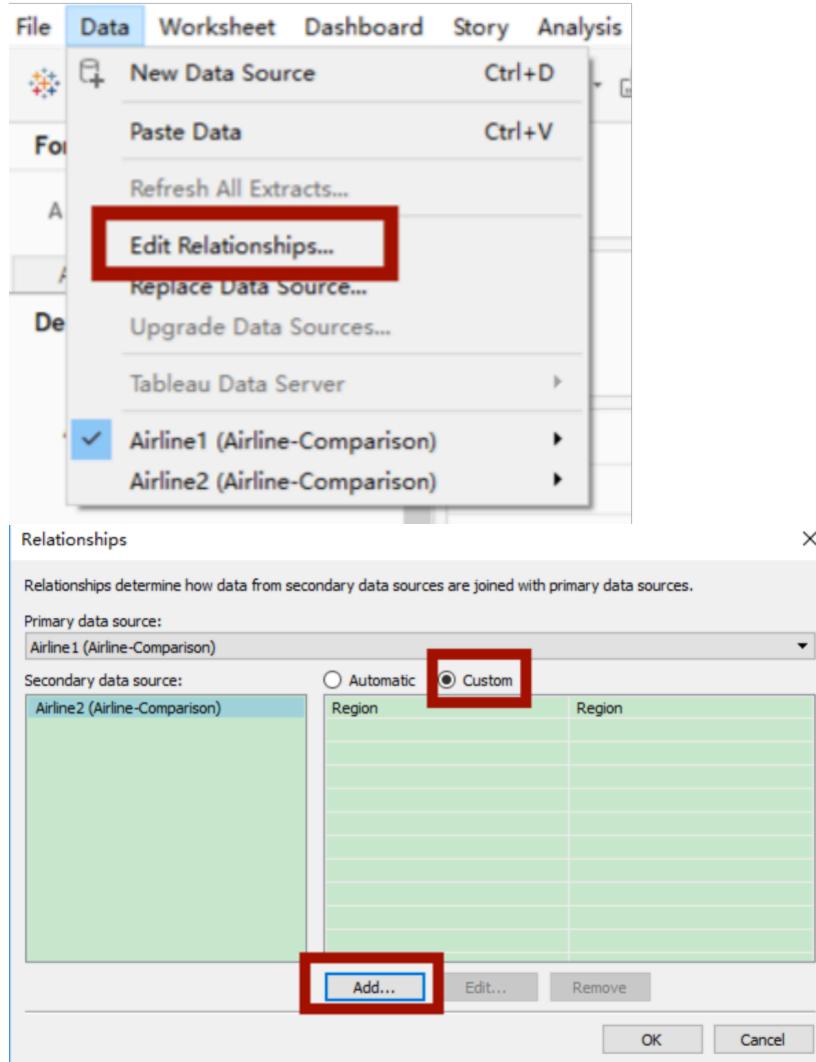
Now we can even change the label more professional. Just change **Numbers to Currency(Custom)**, change **Decimal places** to "1", **Units to Millions(M)**.



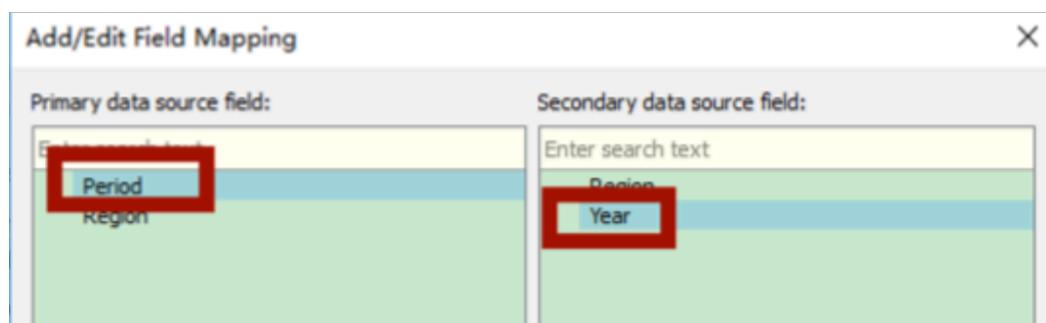
However, we notice Airline2's graph is identical. How come? Because two tables don't know how to connect.



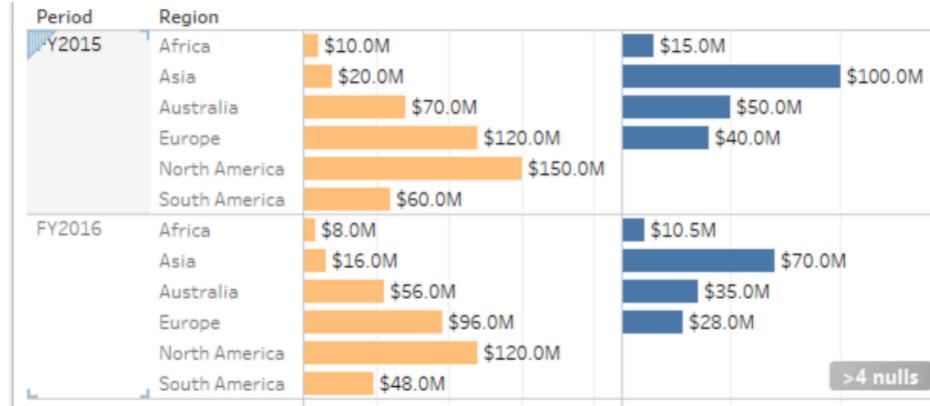
Go click **Data** on the menu and choose **Edit Relationships**, click **Custom** and click **Add**.



Choose **Period** from Primary data source field and **Year** from Secondary data source field.



Now we can see the graph changed. We can see Tableau has smart blending



The only thing is the first data source will be the left table. If you choose Airline2 as the first data source, then Airline will be the left table for blending. (Blue one is the first chosen table. Left: Airline1. Right: Airline2)

Data	Analytics
Airline1 (Airline-Comparison) Airline2 (Airline-Comparison)	
Dimensions	
Abc Period Abc Region Abc Measure Names	
Measures	

vs.

Data	Analytics
	Airline1 (Airline-Comparison) Airline2 (Airline-Comparison)
Dimensions	
	Abc Period (c/o) Abc Region (c/o) Abc Measure Names

Join vs Blend:

Use **Joins** when:

- Combining data at row level

Use **Blends** when:

- Datasources have different level of granularity
- Datasource come from different system