

TABLEAU WORKOUT BOOK

LEARN TABLEAU FAST: 9 EXERCISES TO BECOME
A DASHBOARD EXPERT



CARLOS PONCE

Contents

[Introduction](#)

[Before starting](#)

[Workout 1. Personal finance dashboard](#)

[Workout 2. Superstore basic sales dashboard](#)

[Workout 3. Airbnb's in New York](#)

[Workout 4. Billboard Hot 100 Analytics](#)

[Workout 5. World population dashboard](#)

[Workout 6. USA Colleges dashboard](#)

[Workout 7. FIFA 19 player rank dashboard](#)

[Workout 8. Olympics dashboard](#)

[Workout 9. YouTube trending dashboard](#)

[End](#)

Introduction

Thank you for buying this e-book!

You have taken a step on the right path into becoming a Tableau expert. Whether if you have a low grasp on what Tableau is, or if you have already used the tool for a month or two, you will learn a lot of useful things here. I hope you have a great time going through each of the workouts as I had building them. Even I got to learn a few new features in Tableau on the way, so there really is something new to learn here for everyone, no matter your skill level.

Before starting

Download all material for the workouts [here](#) and save and extract the zip on your computer.

It is highly recommended to read this e-book using the [Kindle for PC/Mac](#) app, with Tableau opened on the side. You will be prompted to do actions such as copying text from here, so this way you will avoid typing it yourself.

As a reference, you will see ratings for each workout on the next criteria:

- **Data:** table size, if there is data conditioning, calculations, and the complexity of the data.
- **Design:** the quality, detail, coloring, and layout of the graphs and dashboard(s).
- **Graphs:** the type, the amount, and the level of configuration of the graphs.
- **Interactivity:** amount and complexity of filters, parameters, and actions.

Data	■ □ □ □ □
Design	■ ■ ■ ■ □
Graphs	■ ■ ■ □ □
Interactivity	■ □ □ □ □

Tableau Desktop version **2019.4** is used throughout this e-book. I recommend you use this version as well by downloading it from [here](#). If you don't have a license, you can activate the 14-day trial, which should be enough for you to cover all the workouts.

You will see that all workouts have “Explained” and “Tip” sections. These are not part of the workout, but rather additional information that will help you learn more. The first are to explain you an important theoretical concept while the second are best practices and personal comments on dashboard development.

And last but not least: if you face any difficulty during a workout, or if an instruction is not clear for you, you can contact me through the [Facebook page for this e-book](#). I will personally help you in your problem, so you don't get stuck and are able to complete all nine of them.

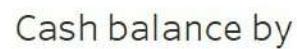
You are now ready, on to the first workout!

Workout 1. Personal finance dashboard

On this workout you will develop a basic but insightful dashboard for personal finance analysis:

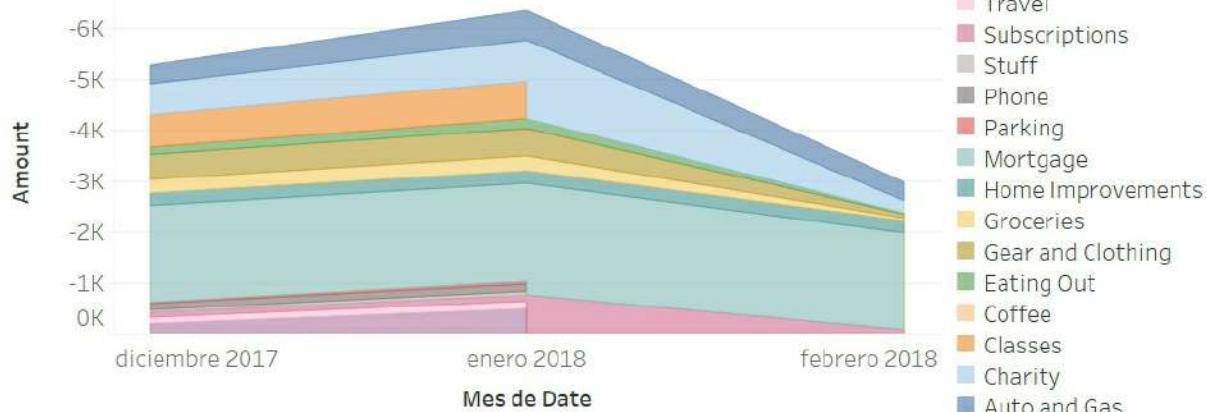
Data	■ □ □ □ □
Design	■ ■ □ □ □
Graphs	■ □ □ □ □
Interactivity	■ □ □ □ □

Personal finance dashboard



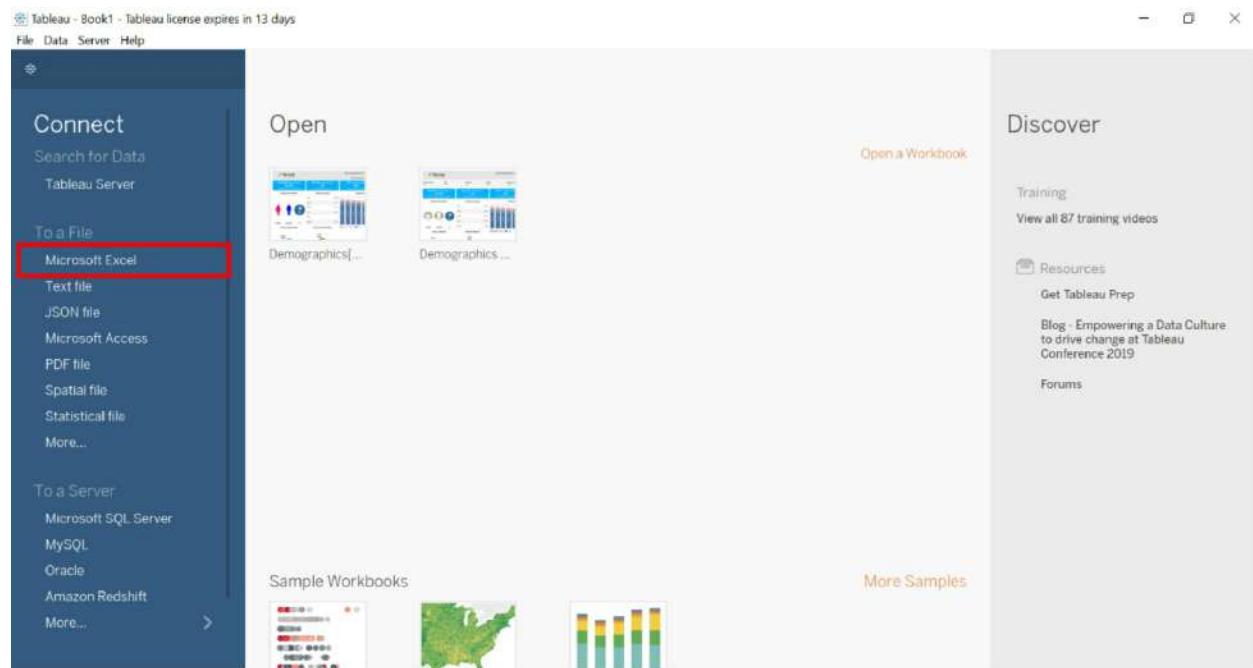
account	
<u>Account</u>	
Alaska Airlines V...	-3,253
American Express	-1,008
Family Checking	10,188
Visa Rewards	-2,288
Grand Total	3,639

Expenditures by month

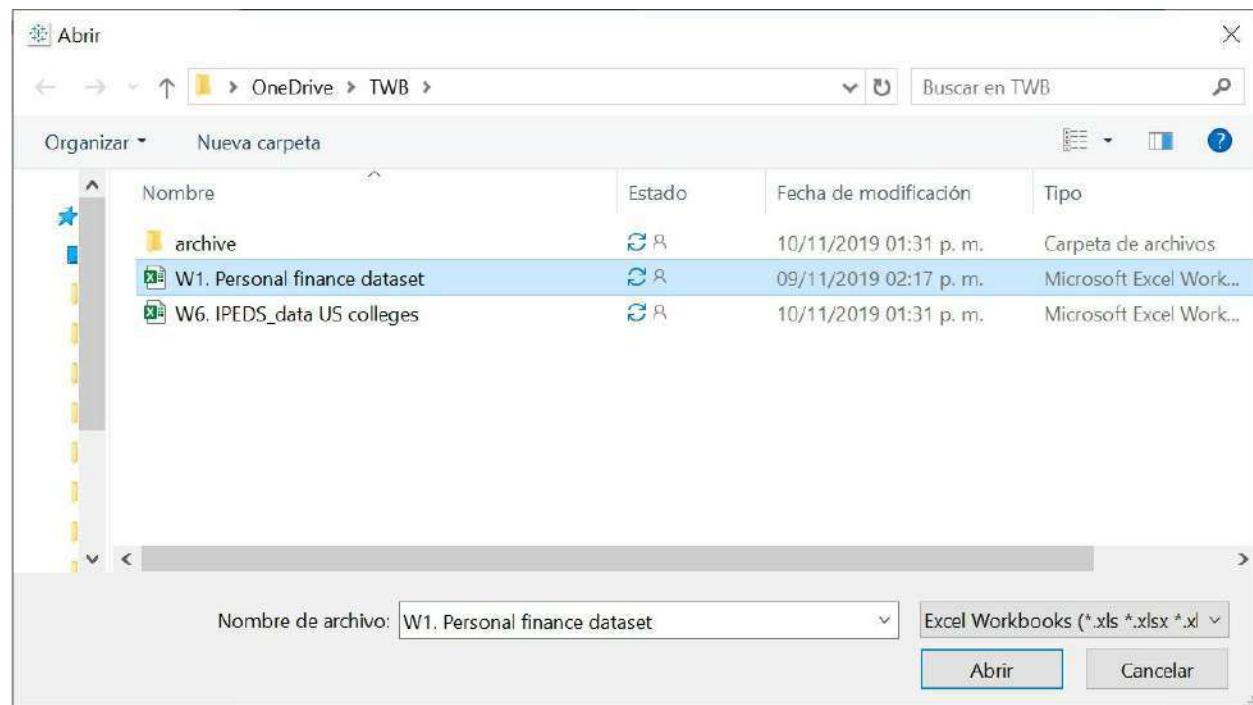


Workout Start

First step is creating a new workbook: open Tableau 2019.4, and on the left panel (called **Connect**), click **Microsoft Excel**, as highlighted in the below image:



Select the **W1 Personal finance dataset.xlsx** file from the **Materials** folder you downloaded (refer to [Before Starting](#)) and press OK.



Explained

When you open Tableau, you are presented with the startup page, divided in three sections:

- 1) **Connect:** use this pane to select a type of data source to connect to.
With Tableau, you can connect to many types of sources, from Excel or comma-separated value (.csv) files, to SQL server or a Hadoop cluster. You can even use the contents of your clipboard as a data source, which means you could copy-paste a table from any place!
- 2) **Open:** here you will find the workbooks you have previously worked with, as well as sample workbooks that showcase Tableau's features through premade dashboards.
- 3) **Discover:** where you can find Tableau tips, tutorials, trainings, resources, and news.

Now that we've gone through the first step, let's go to the data preparation.

The screenshot shows the Tableau startup interface. On the left, there is a sidebar titled 'Connections' with a single entry: 'W1. Personal finance dataset' (Microsoft Excel). Below this is a section for 'Sheets' with a checkbox for 'Use Data Interpreter'. Under 'New Union', there is a link to 'Sheet1'. The main area is titled 'Sheet1 (W1. Personal finance dataset)' and contains a data preview. The data is organized in a grid with columns labeled F1 through F10. The first few rows show transaction details like date, category, and amount. The interface includes standard window controls (minimize, maximize, close) and a toolbar at the bottom.

Explained:

This is the **Data Source** tab, where you configure your data source(s). We will not go through all the features for now, but know that it is divided in three sections:

- 1) **Connections** (left): lets you see the source(s) you connected to.

Clicking each connection, you will see below to it the different tables each contains. In this case, you see your Excel source, and below the only sheet it contains, *Sheet1*.

- 2) **Canvas** (top center): shows the name of your data source – in this case, *Sheet1 (W1. Personal finance dataset)* –, the type of connection (live or extract), and the table(s) that you are connecting to. The canvas must have at least one table, and normally it will put it automatically, but if the canvas is empty, you must add a table.
 - 3) **Preview** (bottom center): preloads a limited number of rows (as stated in the “102 rows” dropdown), and lets you see how the table is being loaded. Here you can spot issues like Tableau not correctly assigning the column names and showing only F1, F2 dimensions, the inclusion of unnecessary columns, blank rows (populated by nulls), and other more advanced settings.
-

Tableau has already dragged our table *Sheet1* to the canvas, but looking on the preview section, it seems it faced issues reading the table. If you open the original Excel file, you will see that the actual table starts from cell B3, there is a title on A1, and some comments on the right of the table.

We must make Tableau read the table correctly in order to use it. Now, you could fix this by modifying the Excel file; deleting the title and the comments, moving the table to start from A1, and so on. However, we will learn a different approach: by using **Tableau Data Interpreter**.

Click the checkbox on the **Connections** pane that says **Use Data Interpreter**, it's right above *Sheet1*. Doing so, Tableau will analyze *Sheet1* and give you a better version of the table by removing blank rows, columns, titles, and more.

After it's done, your table on the **Preview** area should look like this:

The screenshot shows the Tableau Data Interpreter interface. On the left, under 'Connections', 'W1 Personal finance dataset' is selected. Under 'Sheets', 'Sheet1' is selected. A note says 'Cleaned with Data Interpreter' with a checkbox. The main area shows a table titled 'Sheet1' with the following columns: Date, Description, Category, Amount, Account, Account #, F9, and F10. The data includes transactions like 'Method Cons...' for Charity, 'Zeitgeist Coff...' for Coffee, and various expenses from Alaska Airlines and Visa Rewards. The F9 and F10 columns contain comments such as '1. How much ...', 'null', and 'Make a line gr...'. At the bottom, there's a 'Go to Worksheet' button.

The column headers are now correct, there are no more nulls on the first line, and the first null column was removed. The Data Interpreter is not perfect, though. As a result, you see we still have the F9 and F10 columns, with the irrelevant comments that we saw before. For these, hover the mouse on the headers (F9 and F10), click the arrow on the top-right of each, and select **Hide**.

The screenshot shows a context menu for the 'F9' column header in the Tableau interface. The menu options are: Rename, Copy Values, Hide, Aliases..., Create Calculated Field..., Create Group..., Split, Custom Split..., Pivot (select multiple fields), and Describe... The 'Hide' option is highlighted with a blue selection bar.

By doing so, your Data Source will not load F9 and F10. It is a good practice to always hide the columns that are not used, this will help in performance when you work with bigger tables.

Finally, to have clear labels, rename your data source to “Expenses”, by selecting the *Sheet 1 (W1. Personal finance dataset)* title at the top.

Tableau - Book1 - Tableau license expires in 13 days

File Data Server Window Help

Connections Add

W1 Personal finance dataset Microsoft Excel

Cleaned with Data Interpreter
Review the results. (To undo changes, clear the check box.)

Sheets

Sheet1

Connection Live Extract

Filters Add

Sheet1 Sort fields Data source order ▾

Show aliases Show hidden fields 100 rows

New Union

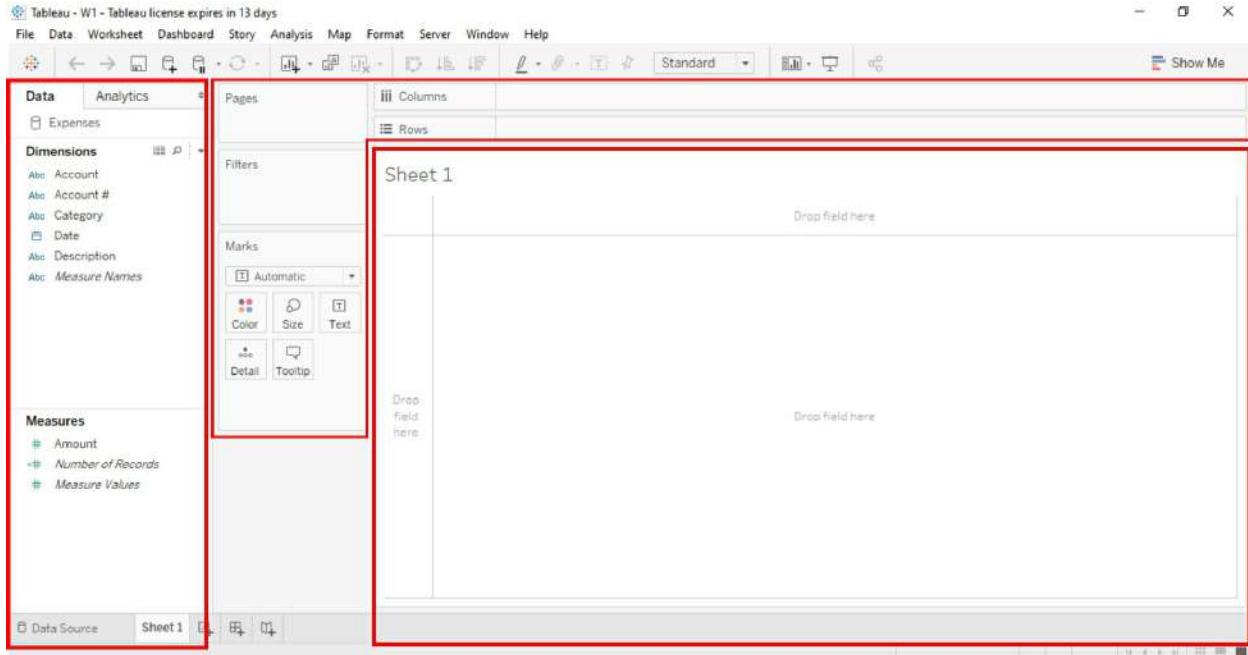
Date	Description	Category	Amount	Account	Account #
14/02/2018	Method Cons...	Charity	-250.00	Family Checkin...	xxxx3244
14/02/2018	Zeitgeist Coff...	Coffee	-10.10	Alaska Airline...	xxxx1561
13/02/2018	Shell Oil, Auto...	Auto and Gas	-38.69	American Exp...	xxxx1005
12/02/2018	Winthrop Mo...	Gear and Clot...	-39.32	Alaska Airline...	xxxx3499
11/02/2018	Toyota of Sea...	Auto and Gas	-251.38	Alaska Airline...	xxxx3499
10/02/2018	Amazon.com	Gear and Clot...	-25.81	Visa Rewards	xxxx3527
09/02/2018	Chase Bank M...	Mortgage	-1,903.00	Family Checkin...	xxxx3244
09/02/2018	Paycheck	Paycheck	5,544.00	Family Checkin...	xxxx3244

Go to Worksheet

Data Source Sheet1

The data source is now ready to be used! Click on the orange **Sheet 1** button at the bottom left, and let's start making our first graph.

After clicking Sheet 1 you will see this page:



Explained:

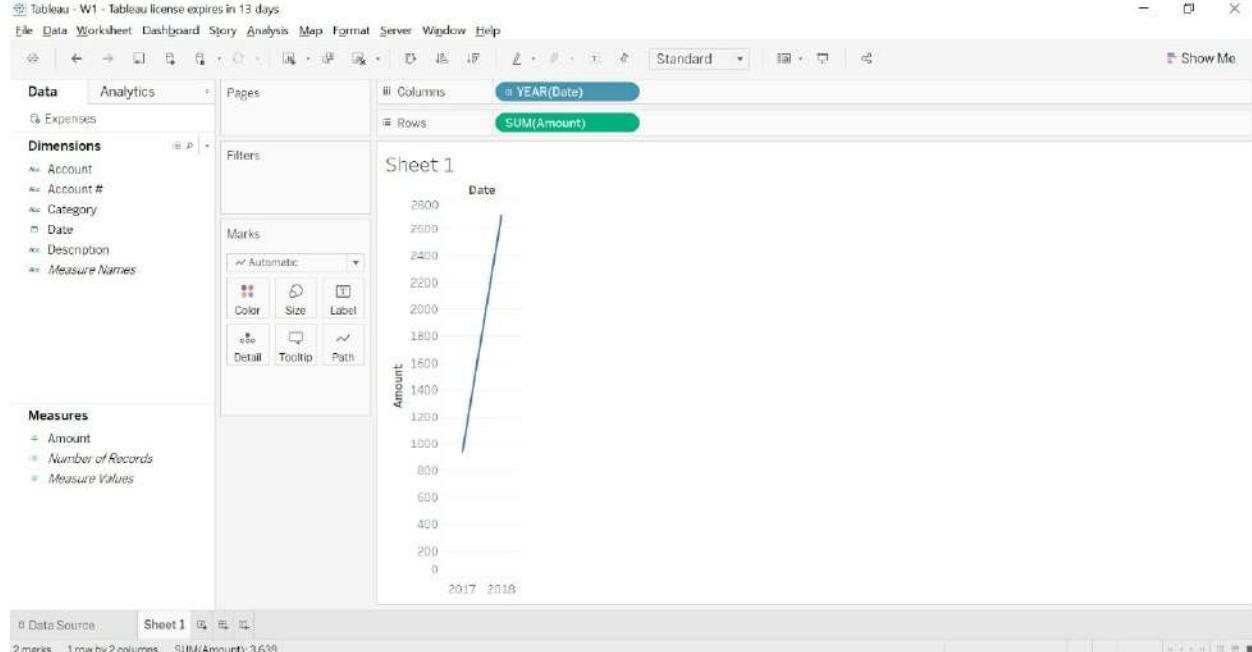
This is the Tableau workspace, the starting point for making a new graph. Let's go through the three highlighted sections:

- 1) **Data pane (left):** where you see your current data sources.
 - a. Tableau divides your columns into Dimensions and Measures.
 - b. Dimensions are the categories by which you will aggregate your Measures (example: country, customer, project). Your Measures are the amounts you want to analyze (example: sales, profit, costs).
 - c. Dimensions are commonly text-type columns, while Measures are numeric.
 - d. Tableau adds a predefined dimension called *Measure Names*, and two predefined measures, called *Number of Records* and *Measure Values*.
- 2) **Shelves (around canvas):** these are the key elements that you will use to create and customize your graph. We have the rows, columns, filters, marks, pages, and measure values shelves, but on this workout,

we will only use the first three.

- 3) **Graph canvas** (center): where the graph will be displayed. You can also modify certain properties of the graph directly from here.

Our first sheet/graph will be a line graph on which we show how our **total cash balance** is moving by day. For this, click and while holding the left click, drag the **Date** dimension from the **data pane** to the **columns shelf**. Then drag the **Amount** measure the same way to the **rows shelf**. You will get this result:



We get a line graph that shows our cash amounts for 2017 and 2018. When we dragged our two columns to the shelves, Tableau automatically made two things:

- 1) For **Date**, it decided to show only the year level, not the exact dates.
- 2) For **Amount**, it decided to aggregate it with a sum.

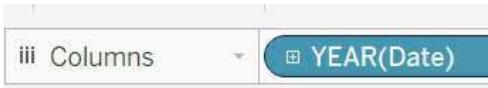
In most cases, this will be what a common user wants, but in case you don't, you can also modify these settings. For now, we will keep the **Amount** as sum, but let's open the date so that we can see how the cash fluctuated along each day.

Since we are talking about a dimension of type **Date**, Tableau allows us to **drill-down** on it, one of the key functionalities of a dashboard. To drill-down, you can do one of two things:

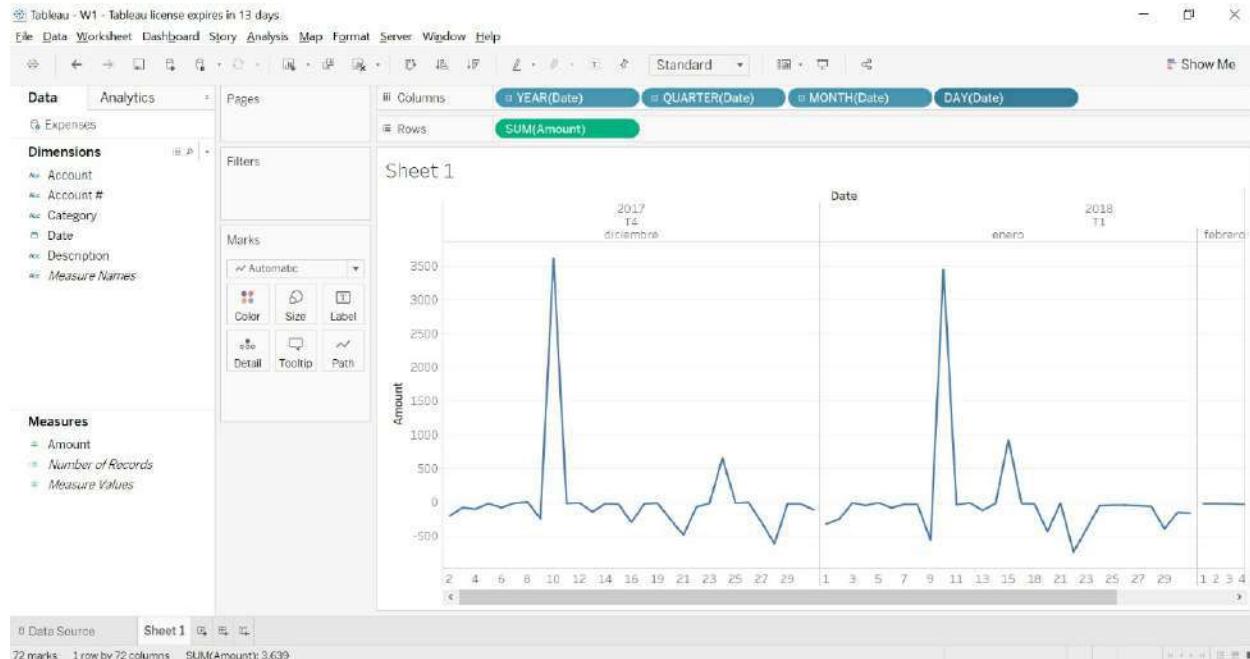
- a) Hover the mouse on the graph's X-axis (2017, 2018), and press the [+] button that appears.



- b) Press the [+] button that appears on the columns shelf inside the **YEAR(Date)** item.



Keep drilling down until you get the **DAY(Date)** pill on the **columns shelf**. You should see this:

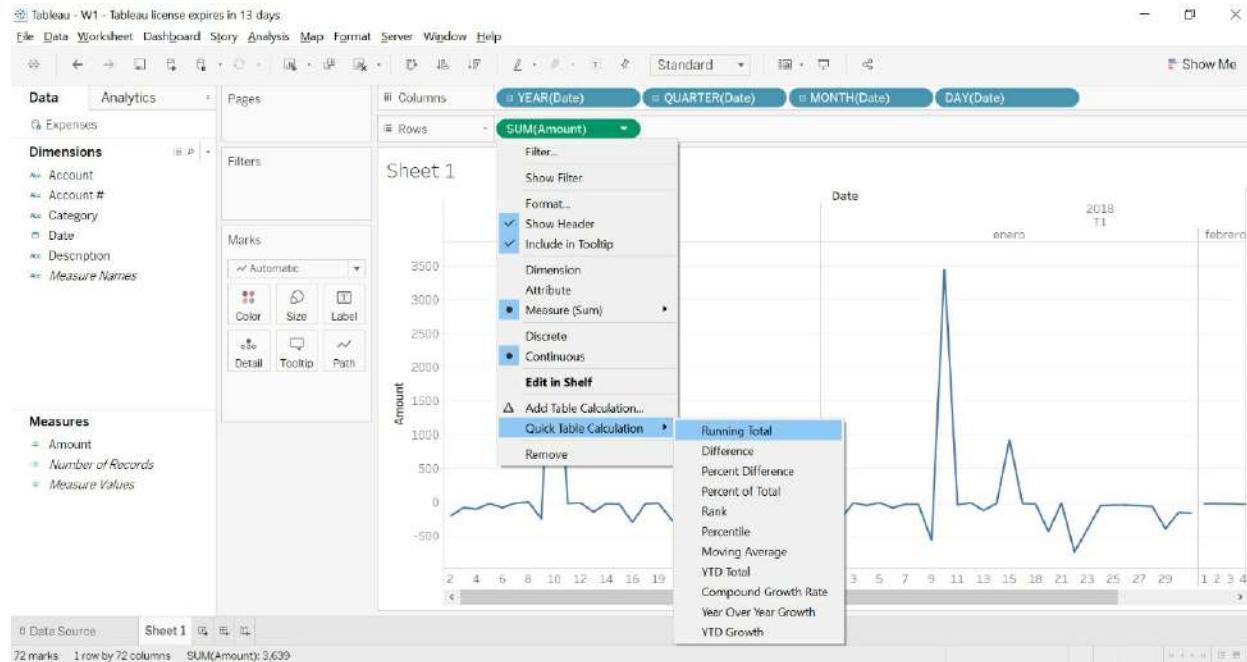


Now we see a line graph of our daily cash balance. However, note that each point in the graph indicates the sum of incomes and expenses for **only** that specific day. For example, on December 10th we got a paycheck, and the amount for that day grew close to +\$3,500. Then on December 11th the line goes back close to zero, because it's not summing the 3.5k from yesterday, nor any amount from previous days; it's only taking the expenses made on that day.

On the beginning we said we wanted a graph that shows how our **total cash balance** by days. If we just wanted to see the individual cash flow by day, then the above graph would be fine. But we want to see the accumulation of

all flows, so we need to apply a **Running Total** calculation on **SUM(Amount)**.

Hover your mouse over the **SUM(Amount)** green pill in the Rows shelf and click the arrow that appears (or right click the pill, it works both ways). Select Quick Table Calculation, and then select **Running Total**.



Now the **SUM(Amount)** pill will have a triangle, this indicates this measure is being displayed through a **table calculation**.

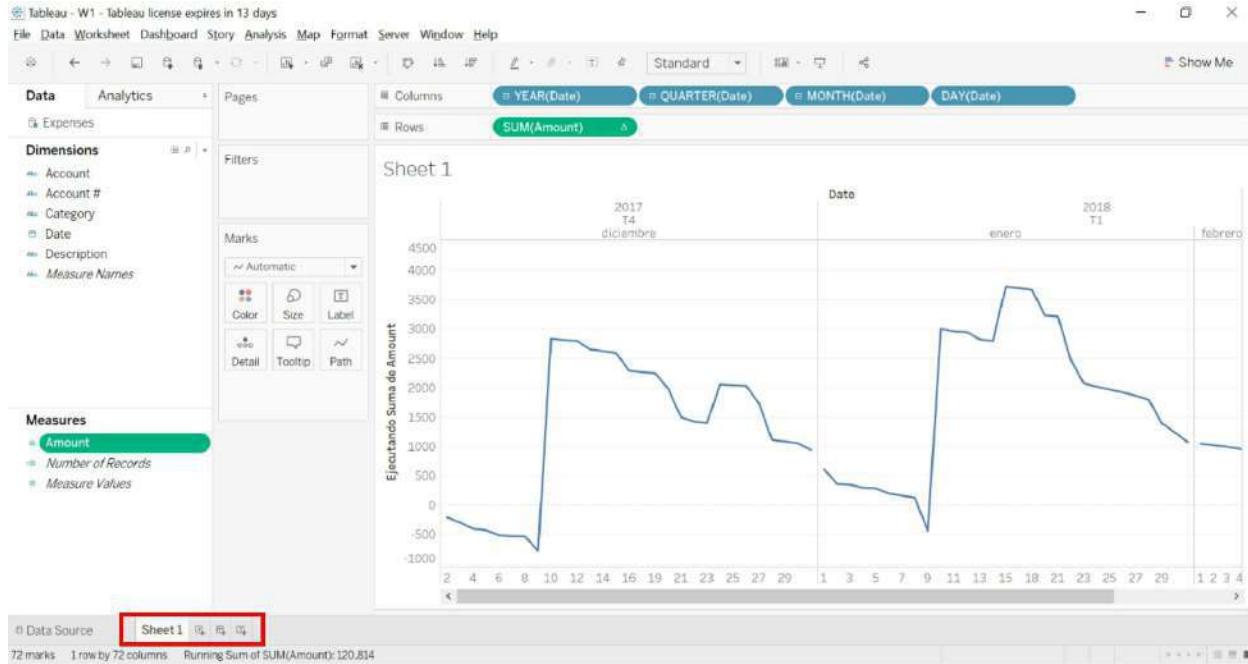
Explained:

Table calculation is a powerful feature that lets you apply complex calculations through a few clicks, so that you can modify the way your measure is presented.

Running Total is a common calculation that gradually sums a sequence of numbers, for each cell it sums its number plus the previous ones of the sequence. On our case, the sequence is the date. So if December 2nd had a balance of -\$208 and December 3rd had -\$86, with running total you will display -\$208 and -\$208+(-\$86) = -\$294, respectively.

There are other table calculations, like **Percentage of total**, **Variance difference**, **Rank**, and more. We will see some of them in future workouts.

You now have a line graph showing the **accumulated cash balance** throughout the days:

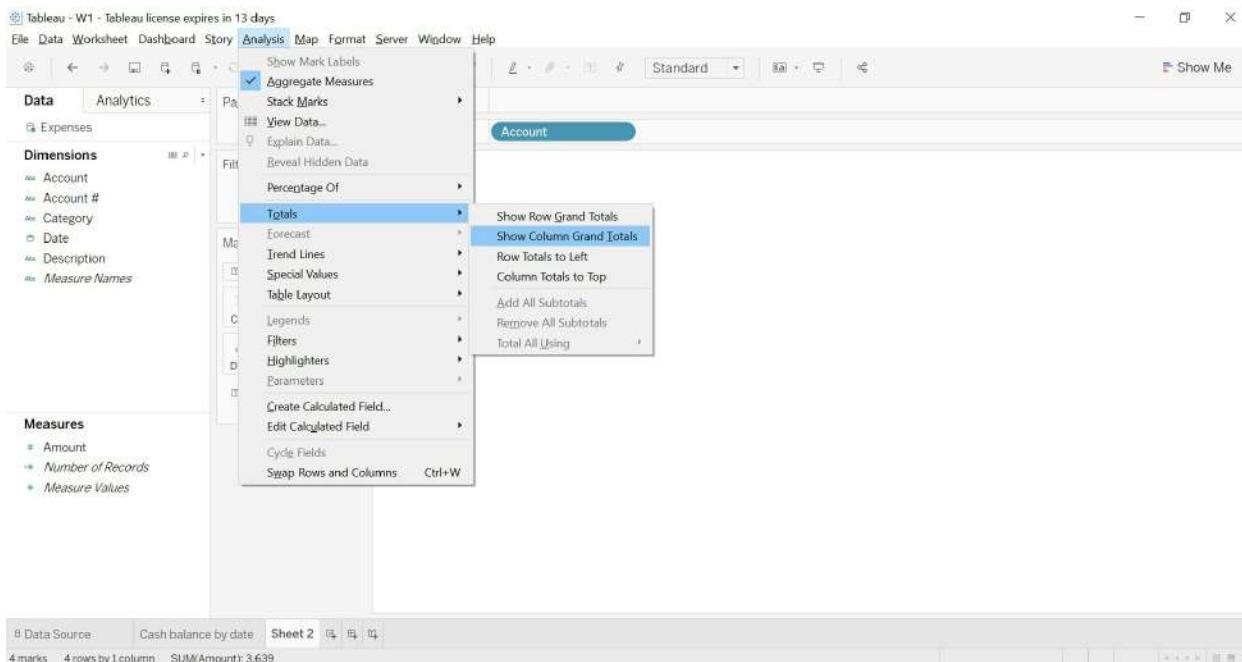


You have finished your first graph! Now we will do two more before building our final dashboard: a table with our current balance per account, and an area graph that splits our expenditures by category.

Before moving, double click the “Sheet 1” tab name at the bottom (highlighted on the previous image) and rename it to “*Cash balance by date*”. Now click the button to the right of the tab to create a **New Sheet**.



On your new worksheet, drag the **Account** dimension from the Data pane at the left to the **Rows** shelf at the top. Then drag the **Amount** measure to the **Marks** card, inside the **Label** property. Lastly, click **Analysis** at the action bar at the top, look for **Totals**, and click **Show Column Grand Totals**, as shown below:



That last toggle will have your table showing the total for all accounts. You should end up with this:

Sheet 2

Account	
Alaska Airlines V..	-3,253
American Express	-1,008
Family Checking	10,188
Visa Rewards	-2,288
Grand Total	3,639

Sum of Amount broken down by Account.

Rename your “Sheet 2” tab to “*Cash balance by account*” and create a third worksheet.

For our final graph, hold with the **right-click** the **Date** dimension and drag it to the **Columns** shelf. A window will appear, laying out all the levels of detail available. You can see the **YEAR(Date)** and **DAY(Date)** levels we used on our first graph, plus other options. Choose the first of the list, **Date**

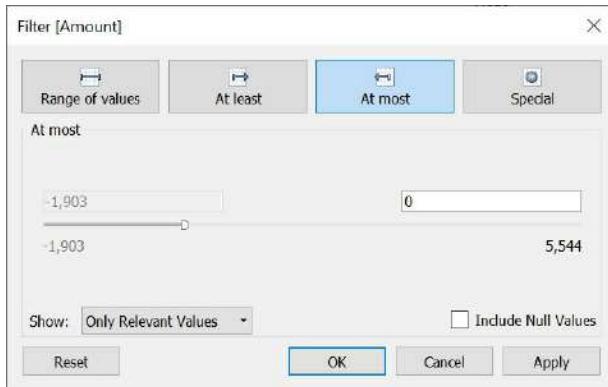
(Continuous) and press OK.



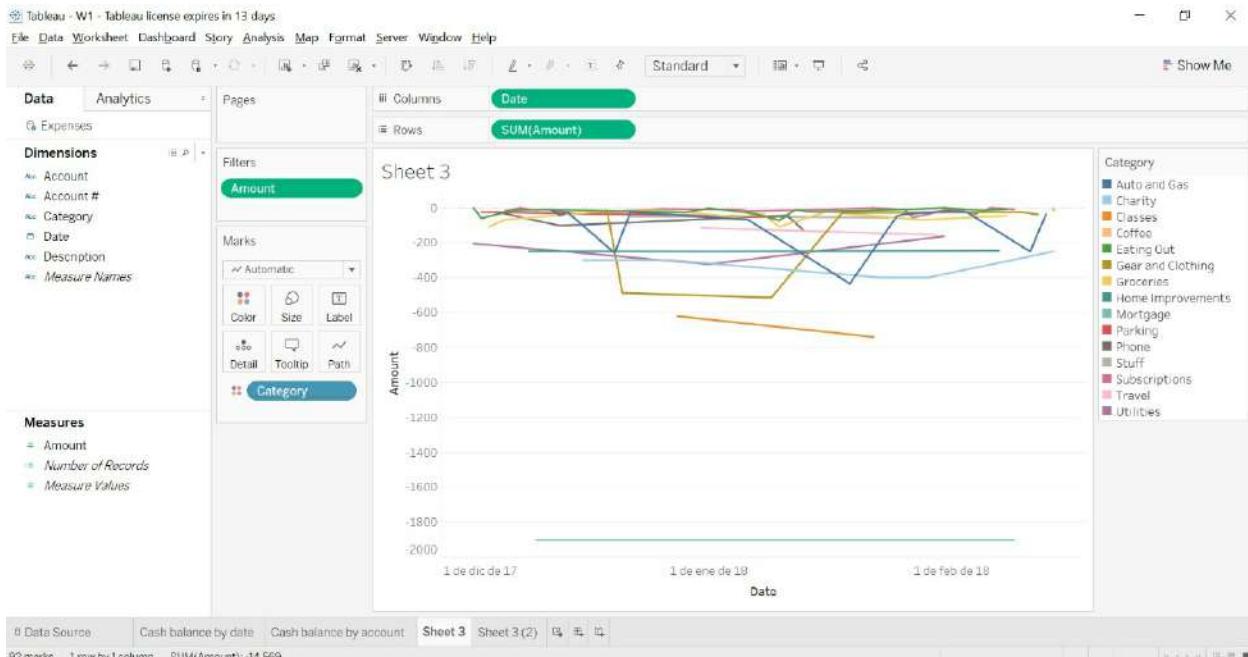
Explained

When you drag an item from the Data pane by **right-clicking** instead of left-clicking, Tableau will not automatically decide neither the level of detail for your field, nor the aggregation for your measure, as it does when you drag with the left-click. This is the recommended way to drag dates, so you don't have to change it afterwards.

Also, drag **Amount** to the **Rows** shelf and drag the **Category** dimension to the **Marks** card, inside **Color**. Lastly, since we are only focusing on the expenditures, drag **Amount** from the **Data** pane to the **Filter** shelf using **right-click**. On the appearing window, select the last option, **All Values**, press OK, and on the range that shows, select **At Most**, and put **0**. Then press OK.



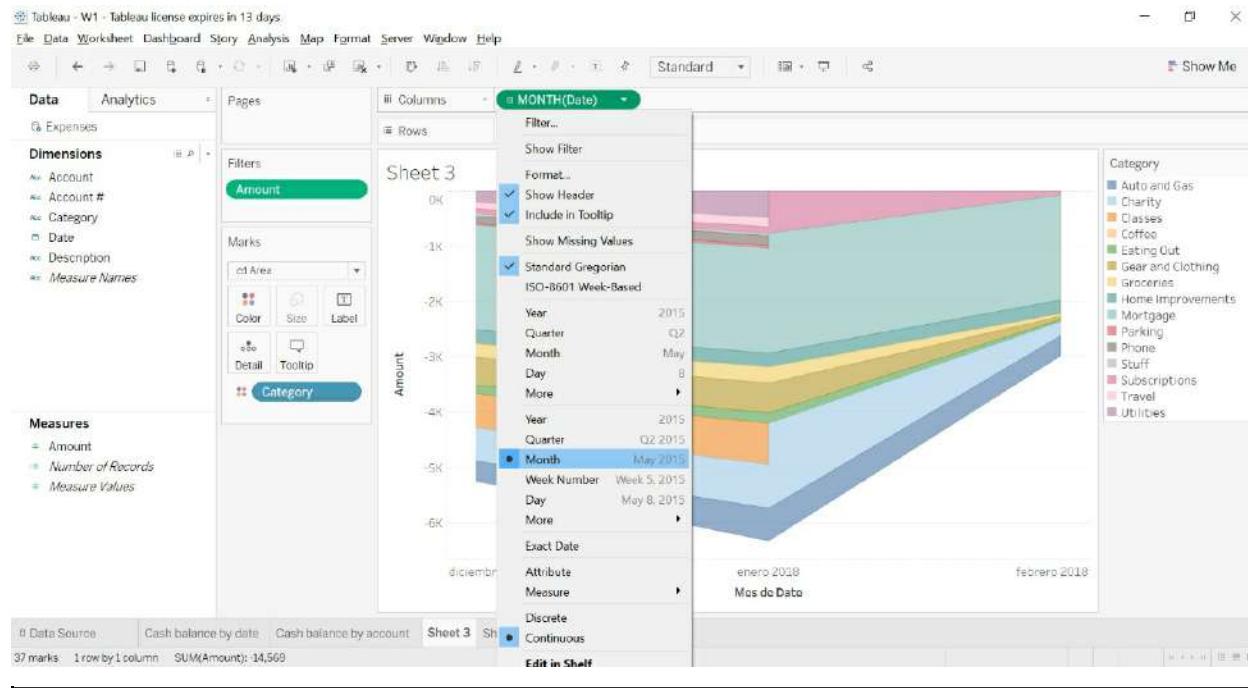
You should be able to see a graph like this:



What we have here is a graph showing expenditures, that contains one line per category. But since we have many categories, it's hard to distinguish between the line of each. For these cases, an area graph works better. To turn the line graph into an area graph, select the dropdown inside the Marks card, which says **Automatic**, and select **Area**.

The other issue is that, judging by the looks of our area graph, it is unusual that there is more than one category of expenditure on the same day. In order to make an insightful graph, we should **drill-up** to month level. To do this, hover the mouse on **MONTH(Date)**, click the arrow, and from the options that show choose **Month**. You will notice that there are two options for month (as well as for year, quarter, and day), this is because the first ones are **discrete**, while the second ones are **continuous**. Choose the continuous

version. You will see a graph like below:



Explained:

By now you should have noticed the presence of green and blue on the pills that you drag to the shelves. This refers to a property that states if the **dimension or measure** is either **discrete** or **continuous**.

For measures, discrete means it **cannot have decimals**, while it can when it's continuous. For dimensions, discrete means it's limited to a set of values and disregards hierarchies, while continuous considers the hierarchy parents (rarely used).

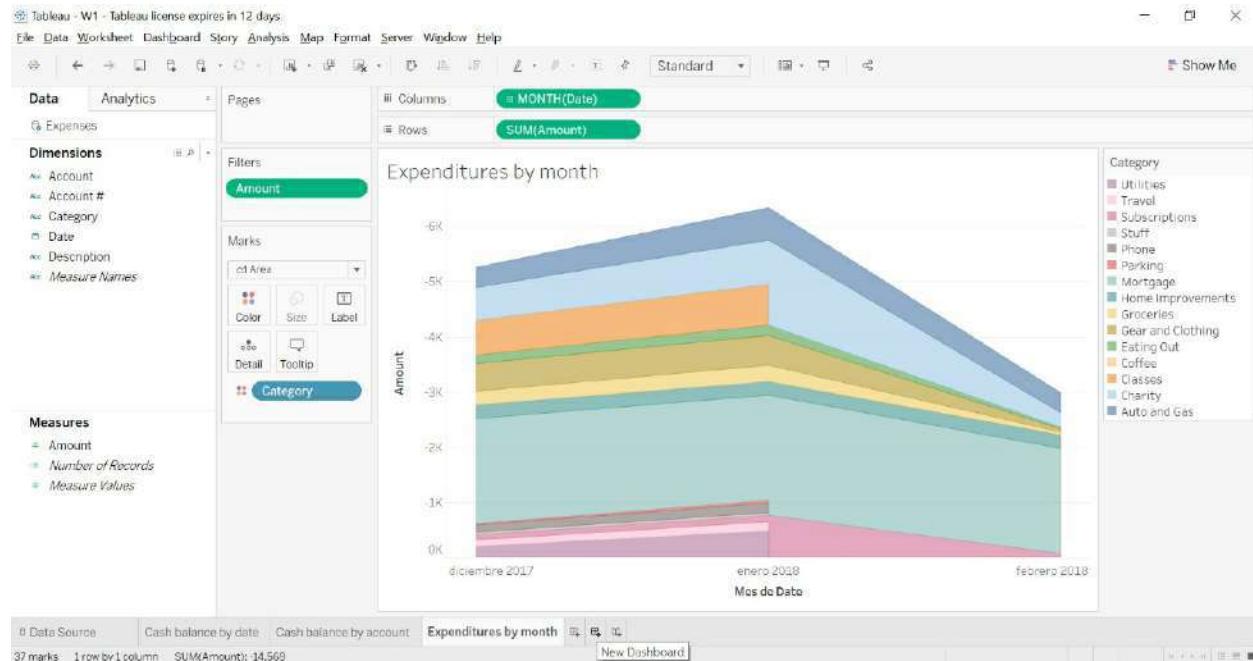
I know it is a bit confusing, fortunately you just need to know one thing: **dimensions are typically discrete (blue) and measures continuous (green)**, but sometimes you will need to make a dimension continuous, or a measure discrete. Here you can see how a discrete dimension (Date) was used as a continuous one.

Had we used the **discrete** version for MONTH(Date), you would only have seen the distinct months on your graph, **disregarding years**. Using **continuous**, you displayed the months **with the years**.

If you go back to our first graph (*Cash balance by date*) and look the Columns shelf, you will see that we used **discrete** dates. That is why we

needed to add additional pills for the year, quarter, and month levels so it shows that as well for our days pill. You could replace all these four pills with a single **continuous** DAY(Date) pill! Plus, the x-axis will look better with a single line.

For the last step, right-click on the vertical or y-axis (Amount) and select **Edit Axis**. On the **Scale** section toggle **Reversed** and close the window. Rename the sheet “*Expenditure by Month*”.

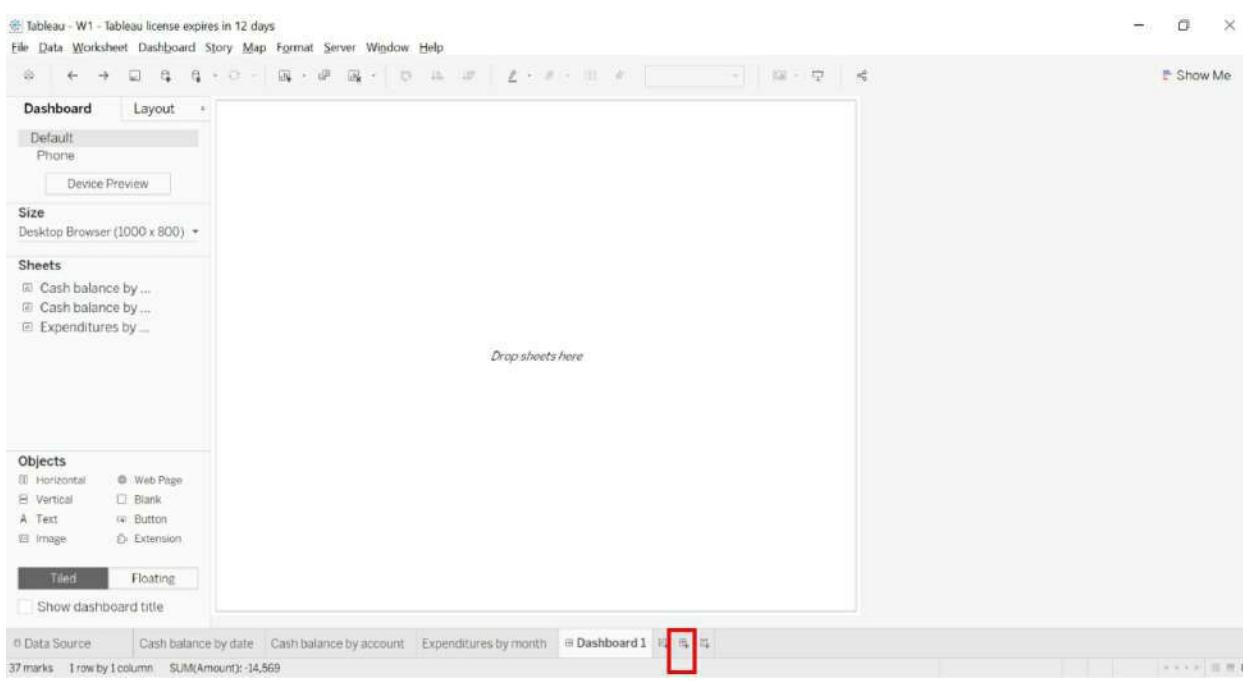


Now that we have our three graphs, it's time to make our dashboard. Click the **New Dashboard** button next to the **New Sheet** button.

Explained:

There are three types of tabs in Tableau:

- **Worksheet:** the individual graphs/tables that are added into dashboards.
- **Dashboard:** the views where users can see several worksheets laid out nicely and interact with them through filters and parameters.
- **Story:** another form of interaction where you can attach worksheets or dashboards and add annotations to them to bring more insights to your analysis.



Explained + TIP

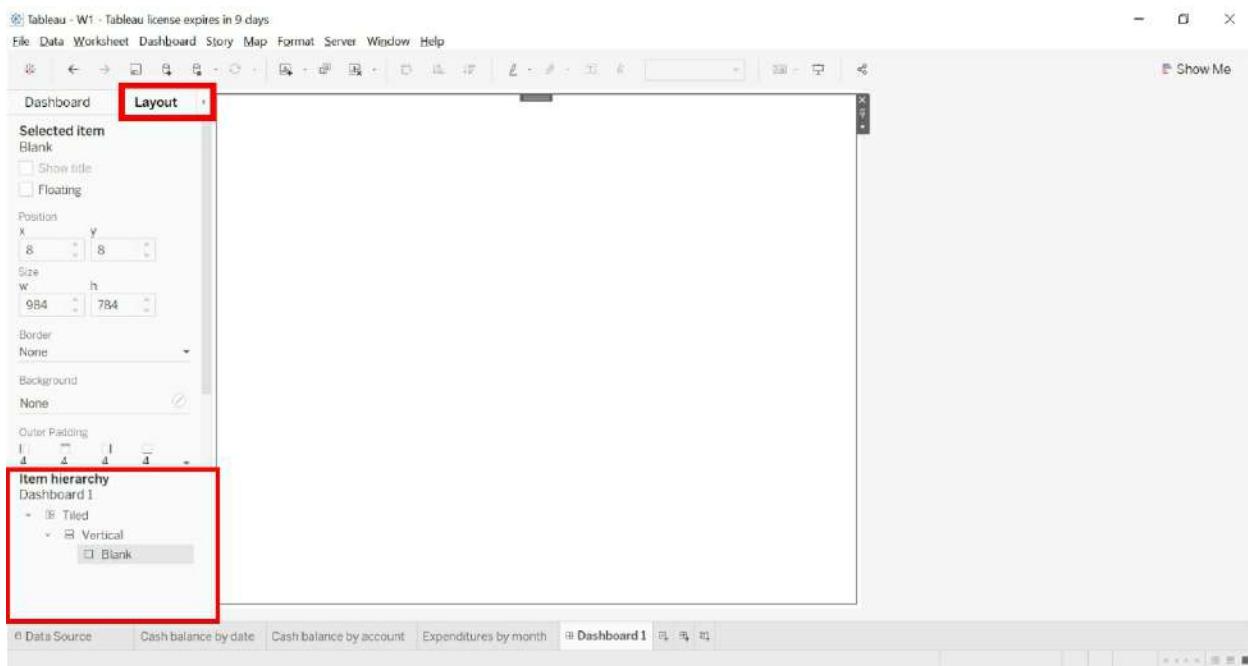
A dashboard is made up of **sheets** and **objects**, you will see both on the left pane. It can be built in two ways: tiled and floating.

- **Floating:** you place your **sheets** in exact positions in your dashboard, they can be on top of each, and they don't move when you resize your dashboard.
- **Tiled:** you use the **Horizontal**, **Vertical**, and **Blank** objects to place your sheets in relative positions, that will adapt to the size of your dashboard. Tableau manages the width, height, and proportion of elements automatically.

TIP: **Tiled** is the standard way of building dashboards, but on the workplace, you will see many dashboards are built with the **floating** layout. This mainly happens because either people self-learn how to do dashboards this way, or they have difficulty with the tiled approach. Either way, bear in mind that using floating will result in **higher rework** in the future when changes are requested (and they always are), so I recommend you stick to tiled as much as possible.

We will do the **tiled** approach to build our first dashboard. Drag first a

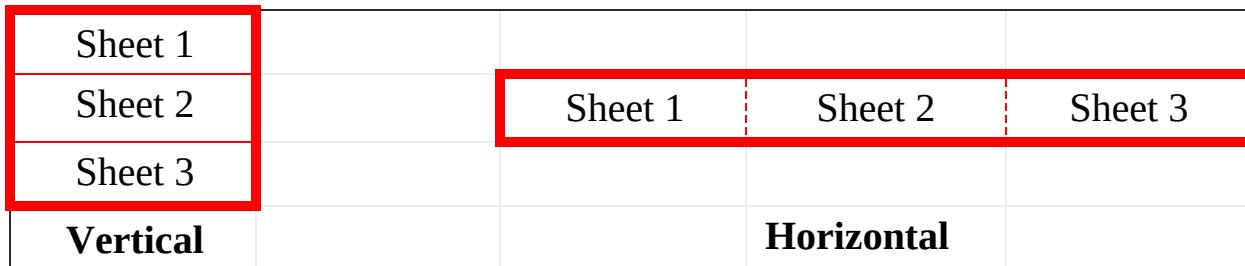
Vertical object into the dashboard canvas. Then drag a **Blank** object inside that Vertical object. Before moving to the next steps, validate that you dragged the objects correctly one inside the other: select **Layout** at the top-left (highlighted below) and open the **Item hierarchy** at the bottom. It should look like this:



Explained:

The item hierarchy is the most important aspect when doing a **Tiled** dashboard. You want to make sure you are correctly using the vertical and horizontal containers so that you don't face difficulties making future changes your dashboard.

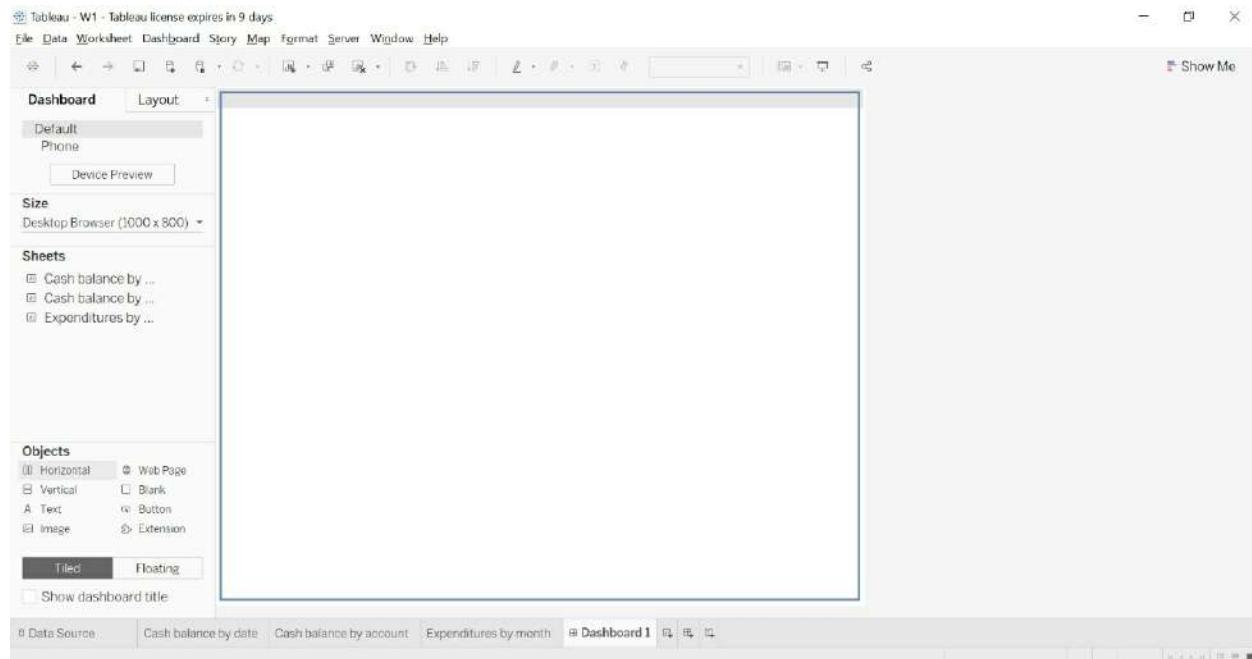
Imagine the **vertical** object is a basket that lets you put objects inside, but only one above the other. The **horizontal** object, in the other hand, is a basket that lets you put objects inside one to the side of the other.



In any space, you can add either a sheet or another container, for more complex layouts.

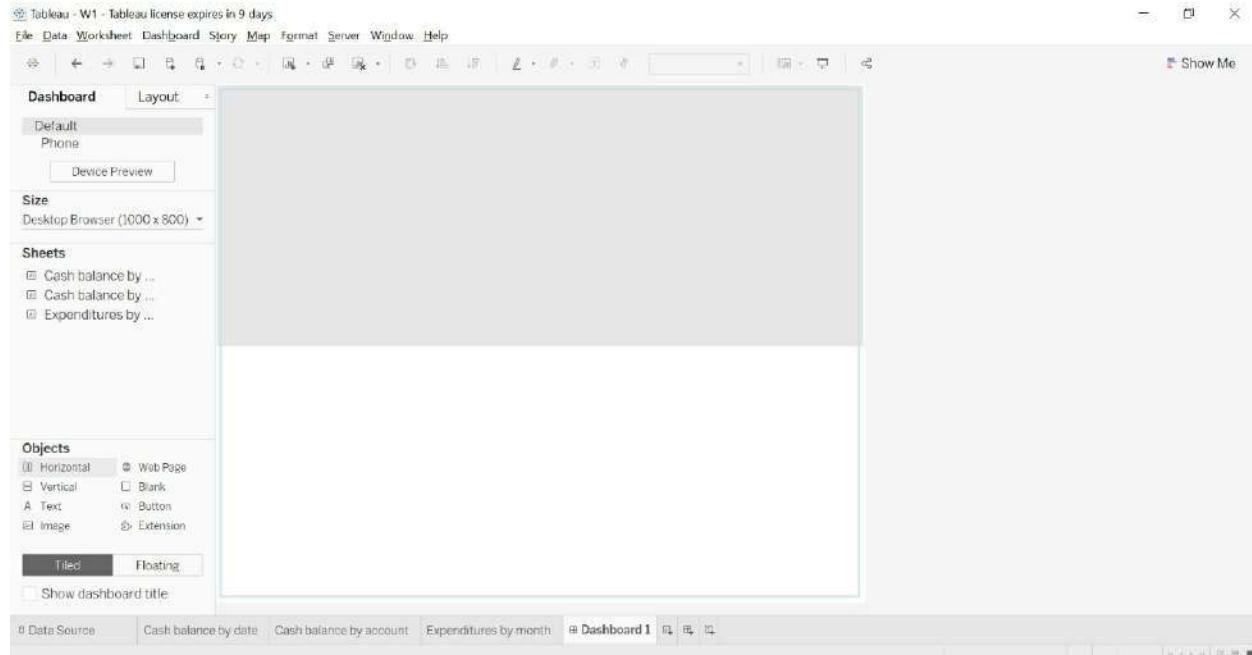
Although it's not necessary, it is a good practice that when you add a new container to your dashboard, you add a **blank** object to it, so you face less issues when trying to place your sheet in the correct space (you will understand going forward on these workouts).

Now, let's add a horizontal container above our blank object but **inside** the vertical container. Drag a **horizontal** object (go back to the **Dashboard** tab on the top left) to your canvas. Without unclicking yet, make sure that the vertical container is being highlighted and that it has a thin horizontal shade at the top, like this:

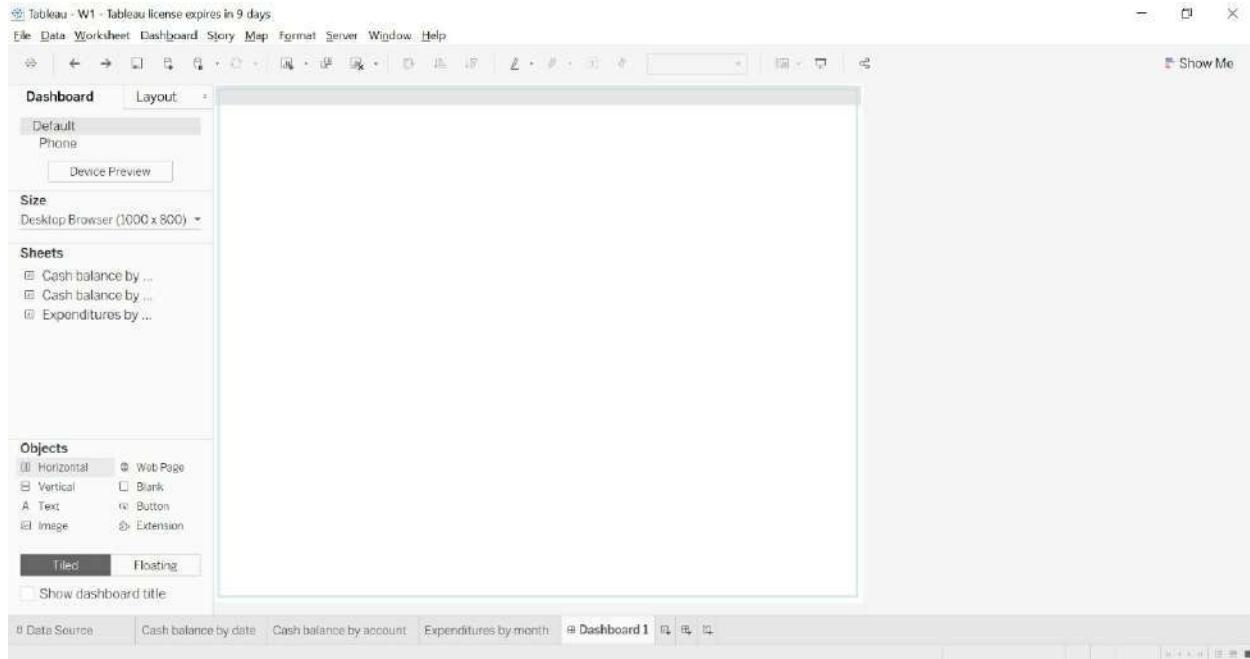


Before releasing the left click of the mouse, you should confirm that your dashboard is looking like this image; blue rectangle (the vertical container) and the thin gray line at the top. This means your object will be inside the container and placed to the top of your blank.

If you see something like below, try moving your mouse a bit down. You should not place the horizontal here, since it's **out** of the vertical container.

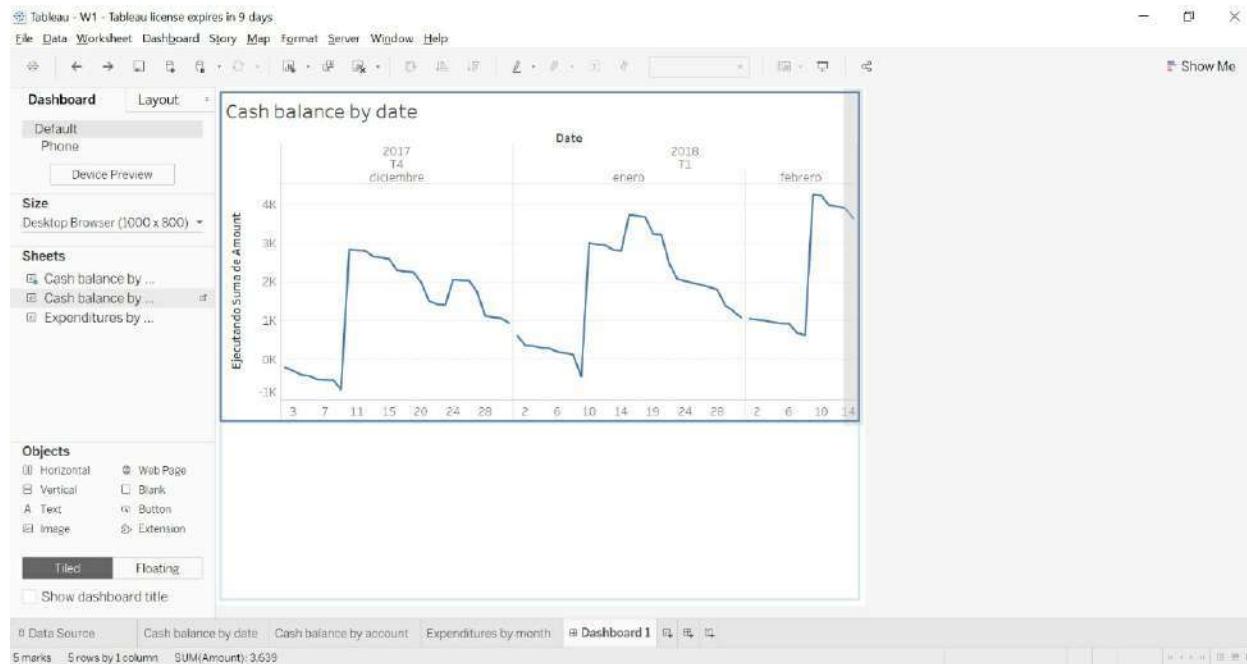


This is also a wrong place:

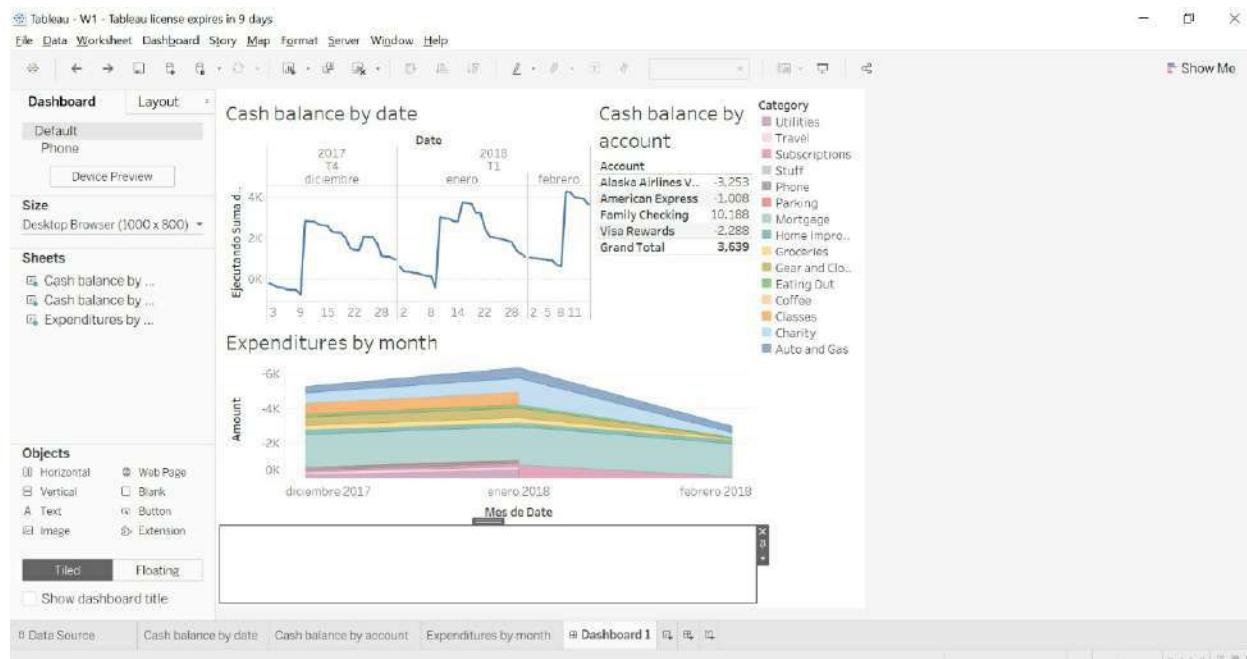


Once you find the right spot, release the mouse click. You will notice that your horizontal container border leaves a small space at the bottom; that's the blank container so it's okay. Now add your **Cash balance by date** sheet inside the horizontal container by dragging it from the left pane. This time, before you release the mouse, confirm that the whole horizontal object is grayed out (this time we want it all to be highlighted because it's the first object we will add to it).

Now drag **Cash balance by account** sheet to the right of what you just added. For this, make sure that your horizontal container has the borders highlighted, and that you have a thin vertical shade to the right of **Cash balance by date**, it's now a vertical shade instead of horizontal since you're adding an object to the right, not the top:



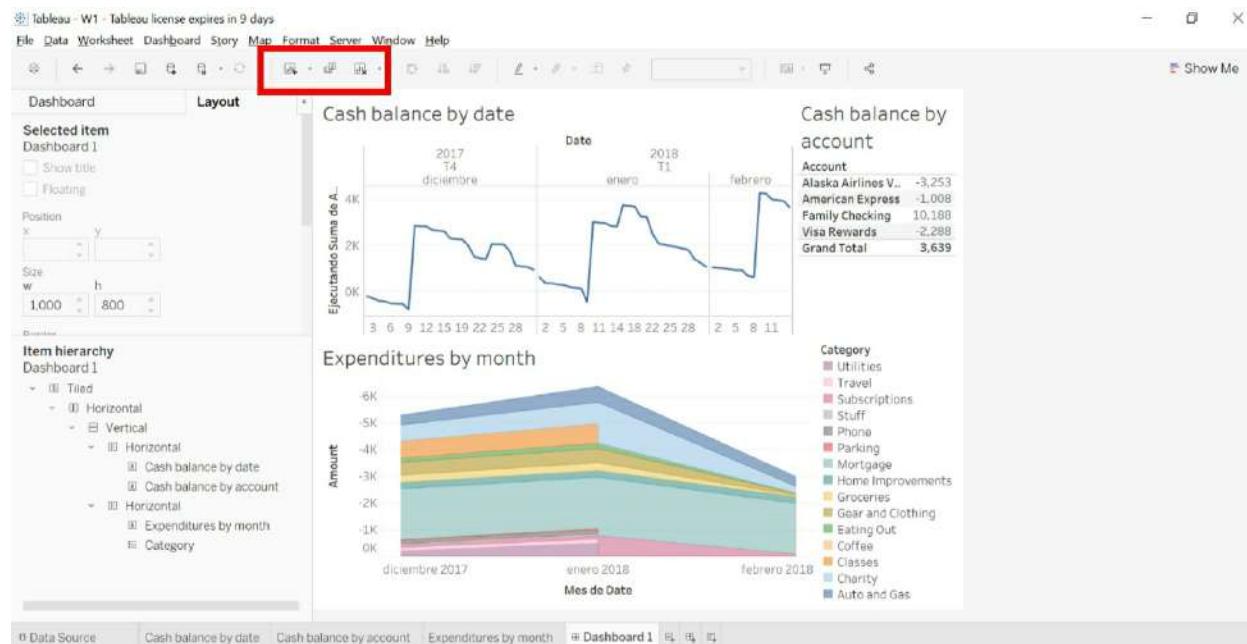
Drag the **Expenditures by month** sheet below the horizontal container but inside the vertical one, which should be highlighted (put it above the blank object). This time, since you already have two objects inside your vertical container (the horizontal container and the blank), you will see a dashed horizontal line close to the middle. Drop the sheet there, you'll be placing your sheet between the two objects. Finally, click the blank space below your sheet, and click the X to delete it.



You'll notice that the categories for **Expenditures by month** were automatically put to the right. This is the default behavior of Tableau when you have Quick Filters or legends on a sheet. However, we would prefer it to the side of the sheet it belongs to, without taking space from the first two sheets.

For this, we need to add a second horizontal container inside our vertical, on which we will have the expenditures sheet and its color legend. Drag a new **Horizontal** inside the **Vertical** object (either above or below **Expenditures by month**), then drag your existing expenditures sheet inside it and move the color legend to its right.

This is how your **item hierarchy** should look in the end, and your overall dashboard:



If you did not finish with this result, check first that you have the same containers and sheets inside the dashboard. Second, check whether if by accident a container was put in the wrong place, and try to move it to the right one based on the gray shades. Or, you can just click the empty dashboard button located on the highlighted top and try again.

To end this workout, rename *Dashboard 1* to *Personal finance dashboard*, and check the box **Show dashboard title** on the Dashboard tab of the left pane.

Congratulations, you have made your first Tableau dashboard! And by doing

so, you have learned the following topics:

- Connecting to an Excel file
- Using Data Interpreter to improve Excel read and other pre-loading tweaks
- The Tableau workspace: data pane, shelves, cards, and canvas
- Line graphs, area graphs, and tables
- Dimensions vs measures
- Discrete vs continuous elements
- Basic table calculations
- Sheets, dashboards, stories
- Floating vs tiled dashboard design
- Hands-on tiled design

You could replace the data from the source Excel file, and start using it as your personal finance dashboard. Have a break to review all the concepts you learned, and when you're ready, go to the next page for a more enterprise workout:

Workout 2. Superstore basic sales dashboard

In this workout, you will develop a sales dashboard for a company called Superstore.

Data



Design



Graphs



Interactivity



Superstore Year Results

Year of Order Date
2019

Sales by state



© 2019 Mapbox © OpenStreetMap

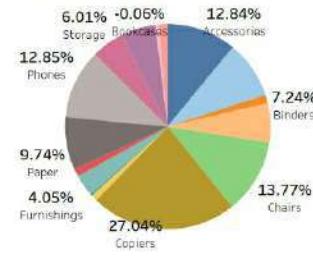
Measure Names

■ Profit ■ Sales

Sales & profit by date



Profit by sub-category



Detail by product

Product Name	Quantity	Sales	Profit
1.7 Cubic Foot Compact "Cube" Office R...	5	\$916	\$156
3-ring stapler pack	6	\$9	\$3
3.6 Cubic Foot Counter Height Office Ref...	14	\$2,711	(\$259)
3D Systems Cube Printer, 2nd Generatio...	4	\$5,200	\$1,352
3D Systems Cube Printer, 2nd Generatio...	1	\$1,040	\$104
3M Hangers With Command Adhesive	12	\$34	\$8
3M Office Air Cleaners	6	\$156	\$55
3M Organizer Strips	19	\$64	\$10
3M Polarizing Task Lamp with Clamp Ar...	8	\$1,096	\$285
6" Cubiclo Wall Clock, Black	7	\$42	\$6
9-3/4 Diameter Round Wall Clock	12	\$155	\$70
12 Colored Short Pencils	3	\$8	\$2
12-1/2 Diameter Round Wall Clock	15	\$256	\$64
14-7/8 x 11 Blue Bar Computer Printout...	8	\$327	\$127
24 Capacity Maxi Data Binder Racks, Pas...	10	\$1,769	\$189

Workout Start

Open **W2.twb** from the **Materials** folder you downloaded. Go to Data Source and observe the Tableau Extract file.

The screenshot shows the Tableau interface with the title bar "Tableau - W2 - Tableau license expires in 5 days". The left sidebar shows "Connections" with "Sample - Superstore" selected and "Table" with "Extract (Extract.Extract)". The main area is titled "Sample - Superstore" with tabs for "Connection" (Live) and "Extract". A search bar says "Extract". The data view shows a grid of order data with columns: Order ID, Order Date, Ship Date, Ship Mode, Customer, Segment, Country/Region, City, State, Postal Code, and Region. The first few rows of data are:

Order ID	Order Date	Ship Date	Ship Mode	Customer	Segment	Country/Region	City	State	Postal Code	Region
CA-2018-152...	08/11/2018	11/11/2018	Second Class	Claire Gute	Consumer	United States	Henderson	Kentucky	42420	
CA-2018-152...	08/11/2018	11/11/2018	Second Class	Claire Gute	Consumer	United States	Henderson	Kentucky	42420	
CA-2018-138...	12/06/2018	16/06/2018	Second Class	Darrin Van Huff	Corporate	United States	Los Angeles	California	90036	
US-2017-108...	11/10/2017	18/10/2017	Standard Class	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	
US-2017-108...	11/10/2017	18/10/2017	Standard Class	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	
CA-2016-115...	09/06/2016	14/06/2016	Standard Class	Brosina Hoff...	Consumer	United States	Los Angeles	California	90032	
CA-2016-115...	09/06/2016	14/06/2016	Standard Class	Brosina Hoff...	Consumer	United States	Los Angeles	California	90032	

Tip

The first step to creating a good dashboard is having a good understanding of your dataset. This is the orders data for a fictitious company called Superstore. You can see that by the columns it has. Now, the most important thing to know about your dataset is granularity; this is the level of detail on which the data is presented. On this case, we can identify it by finding which column has non-repeated or **distinct** values, and that is the **Order ID**.

On real life, you might have more difficulty finding this. The level of detail might even be made of more than one column, this will happen more when the quality of the data you're provided is not good (which, unfortunately, happens frequently).

For now, we know we have Superstore sales data by order. On later workouts we will dive on the concept of exploratory analysis, which will give you more knowledge on the dataset.

By knowing your data, you will be capable of using your creativity to develop more insightful dashboards and give more value to the users.

We will start by doing a map graph, to display sales by country. Click on the *Sheet 1* tab at the bottom.

On the blank sheet, drag the **Sales** measure to the body of the canvas, then the **State** dimension (you will find it inside the **Location** hierarchy). Click the **Show Me** button at the top-right and find the **map** graph (there are two versions, use the filled map).

The screenshot shows the Tableau interface with the following details:

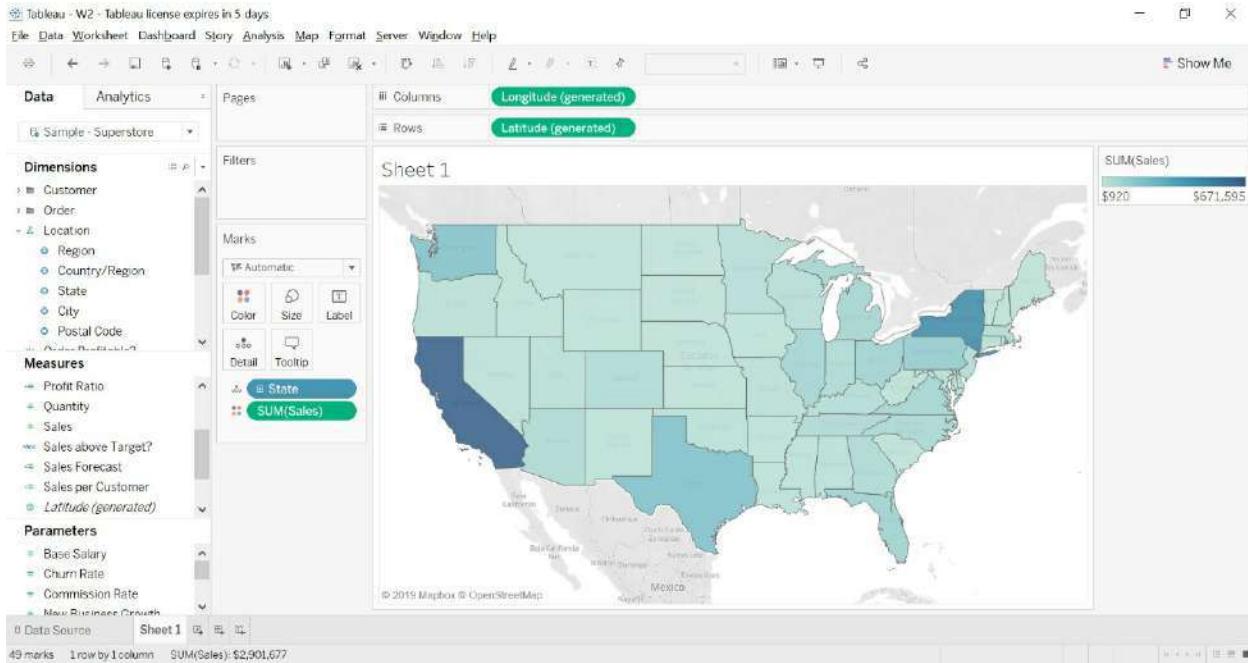
- Top Bar:** Tableau - W2 - Tableau license expires in 5 days; File, Data, Worksheet, Dashboard, Story, Analysis, Map, Format, Server, Window, Help.
- Toolbars:** Standard, View, Selection, Sort, Filter, Zoom, etc.
- Left Panel (Data Shelf):**
 - Dimensions:** Customer, Order, Location (Region, Country/Region, State, City, Postal Code).
 - Measures:** Profit Ratio, Quantity, Sales, Sales above Target?, Sales Forecast, Sales per Customer, Latitude (generated).
 - Parameters:** Base Salary, Churn Rate, Commission Rate, Max Business Growth.
- Middle Panel (Canvas):**
 - Columns:** Region (highlighted in blue).
 - Rows:** None.
 - Marks:** Automatic, Color, Size, Text, Detail, Tooltip. SUM(Sales) is selected.
- Right Panel (Show Me):** A grid of visualization icons. The "Map" icon is highlighted in orange. Below the grid, there is a section titled "For maps try" with the following requirements:
 - 1 geo = Dimension
 - 0 or more Dimensions
 - 0 or 1 Measure
- Bottom Status Bar:** 8 Data Sources, Sheet 1, 4 marks, 1 row by 4 columns, SUM(Sales): \$2,901,677.

Explained

Hierarchies work to group several dimensions as different levels of one dimension. One basic hierarchy is the **Date**, which includes the Year, Quarter, Month, and Day dimensions. We always put them from the highest level to the lowest. In this case, we only have sales for USA, so the biggest level of detail for Location is Region (Central, East, West, South). Then goes State, then City, then Postal Code.

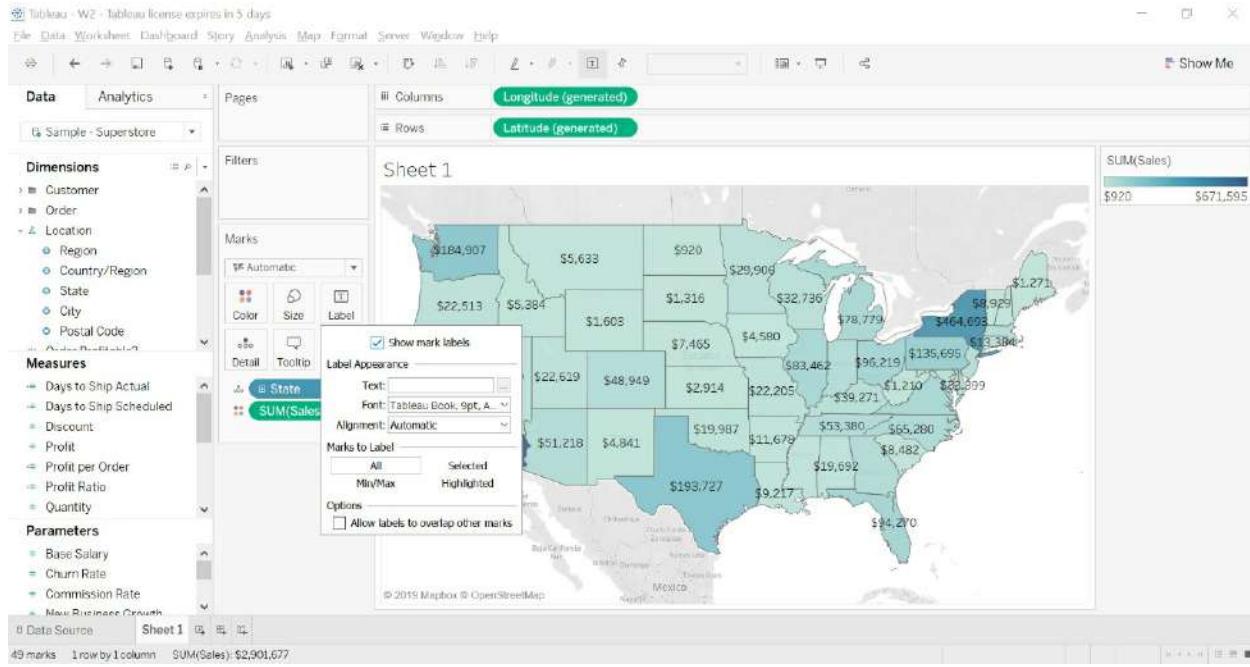
The **Show Me** button is a helper tool that shows you the different types of graphs you can make in Tableau, based on the pills (dimensions/members) you have on your canvas. It will gray out those that you cannot do, but if you want to be able to create one of those, you just have to hover the mouse and read on the bottom side what type of pills you require. In the image above, you see it says that for maps you need 1 Geometric dimension, 0 or more dimensions, and 0 or 1 measure.

Tableau can recognize standard location dimensions, it's able to convert them from the names to the **Longitude** and **Latitude** coordinates for you. After clicking the filled map, you should see something like below:



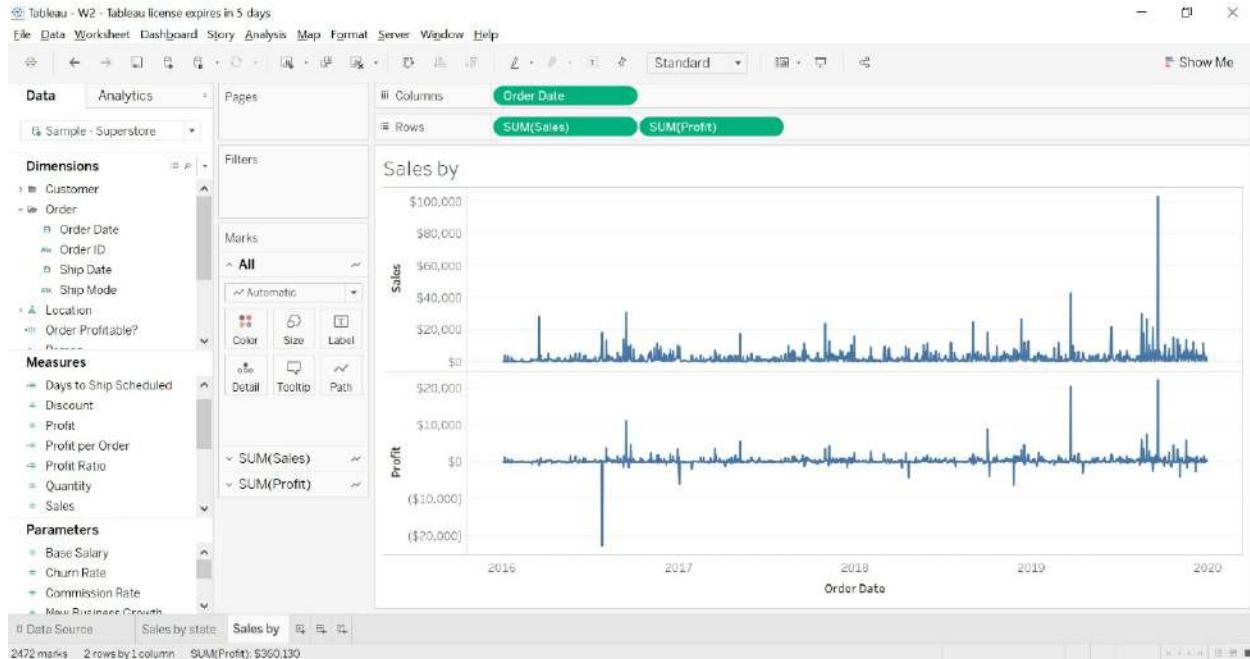
The **Show Me** detected your Geometric dimension (State) and your Measure (SUM(Sales)) and converted your original table into a filled map graph.

To finish: inside the **Marks** card, click the **Label** property. Then tick the box **Show mark labels**. This will show sales numbers for the states.

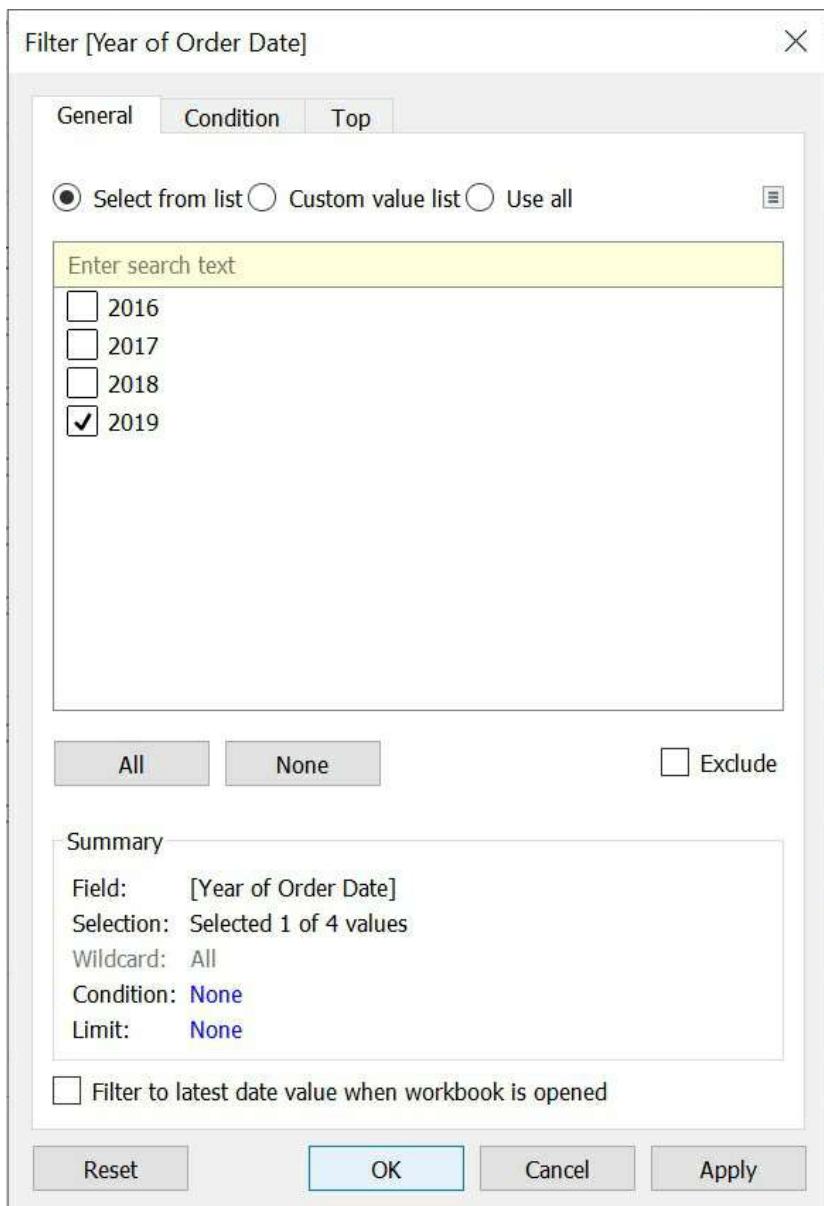


Rename this tab to *Sales by State*.

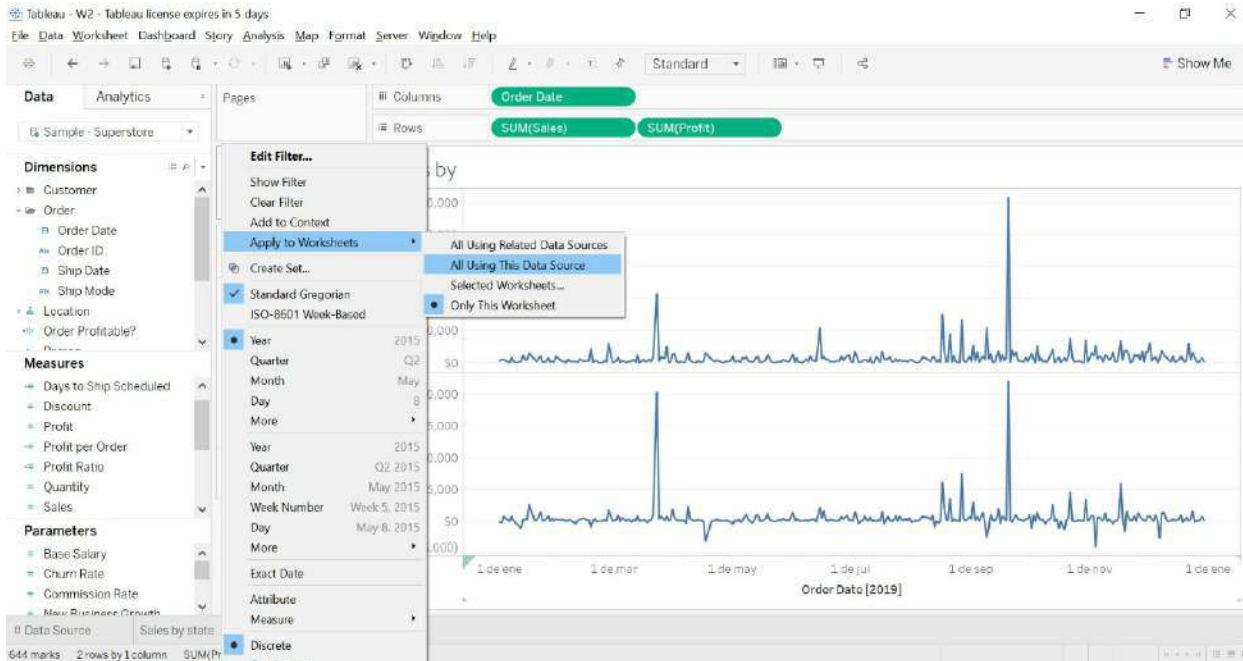
Create a new sheet, drag **Order Date** as **continuous** (green) dimension dragging with the right-click, then drag **Sales** and **Profit** on rows.



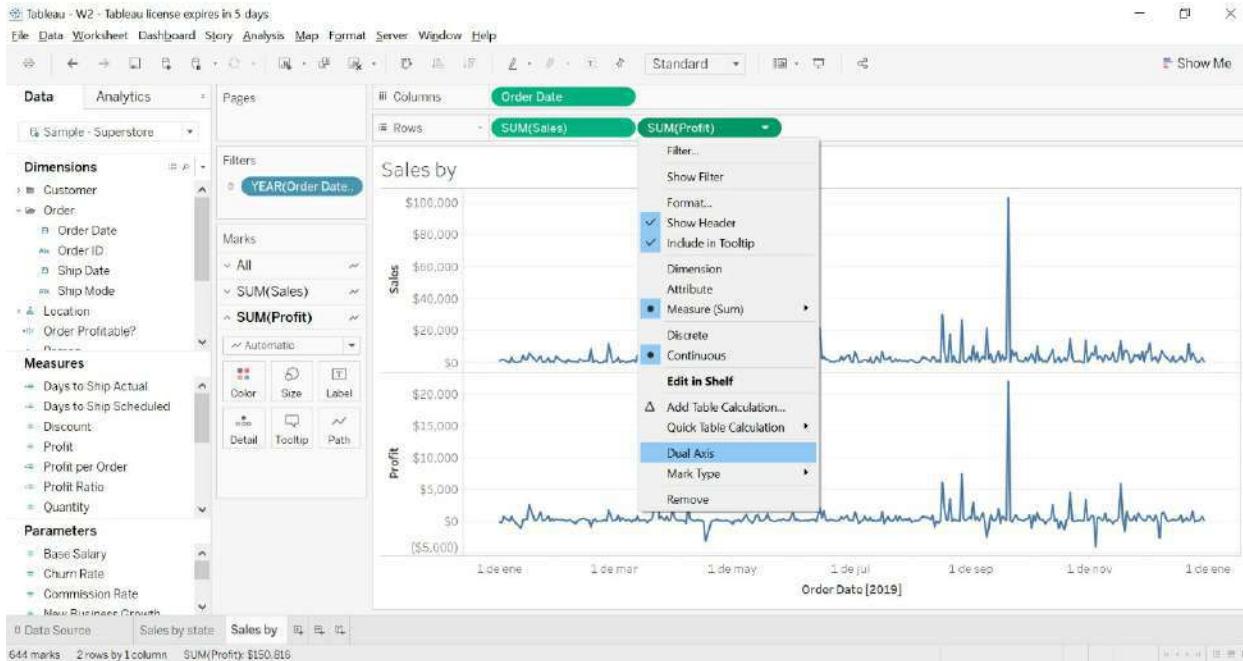
Again, drag with the right click the **Order Date** dimension, but this time put it on the **Filters** shelf. Select the **YEAR (Order Date)** option from the window, and select one year from the filter options.



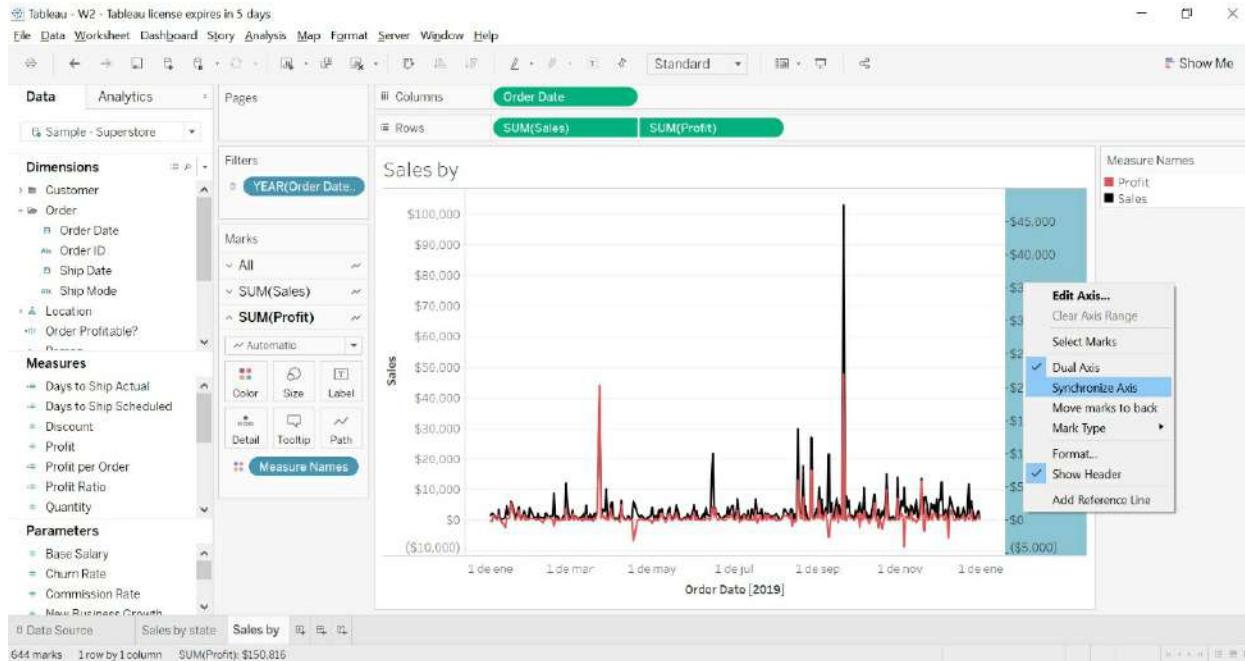
Sheet 2 is now showing sales and profit only for year 2019. For the sake of consistency, we need to have this filter also on the *Sales by state* sheet. Fortunately, we don't need to go through this step for every sheet. Hover the mouse on your new filter pill, click the arrow on the right of it, then under **Apply to Worksheets**, select **All Using This Data Source**. This will make that for present and future sheets you do, this filter will be automatically added.



Our next step will be joining those two lines onto one same graph, so we can compare one versus the other. To do it, just click on the arrow of your **second measure** of the rows shelf (in this case SUM(Profit)) and click **Dual Axis**.

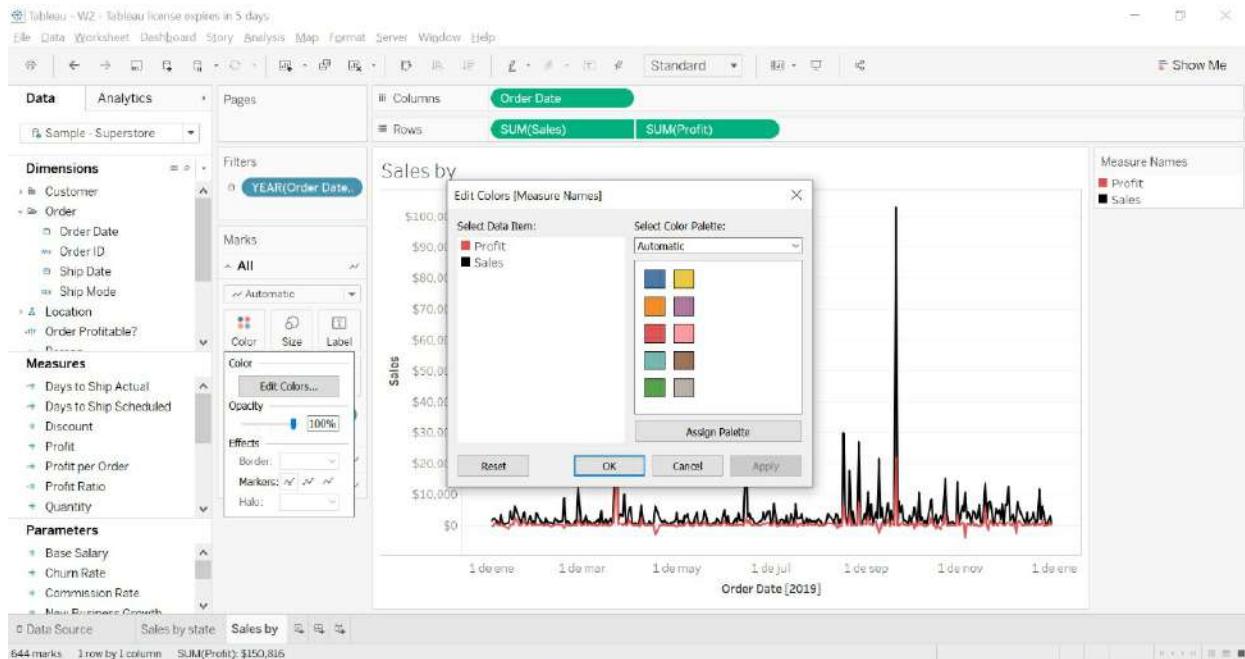


You will see the two lines together, but if you notice, we have two y-axes at different scales. This is a common omission: right click on the y-axis at the right, and toggle **Synchronize Axis**.



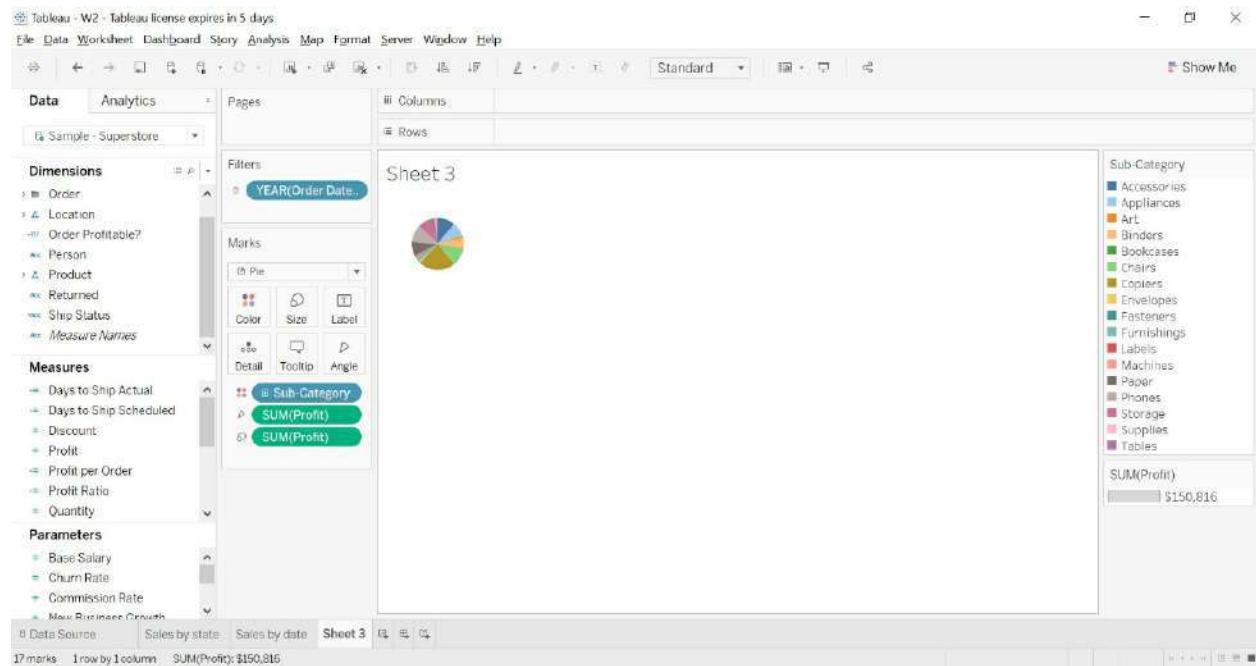
Right click that y-axis again and uncheck **Show Header** since we don't need it anymore.

Sometimes we might want to change Tableau's line color assignment. To do it, just click on the **Color** property inside the **Marks** card, click **Edit colors...**, and you can choose from a variety of color palettes on the dropdown. But for this workout, we'll just assign the default colors by clicking the **Reset** button. It will change the colors to blue and orange.



Rename the sheet to **Sales & profit by date**.

Create a third sheet and add the **Sub-Category** dimension and the **Profit** measure. Look for the **pie chart** on the **Show Me** button and click it.



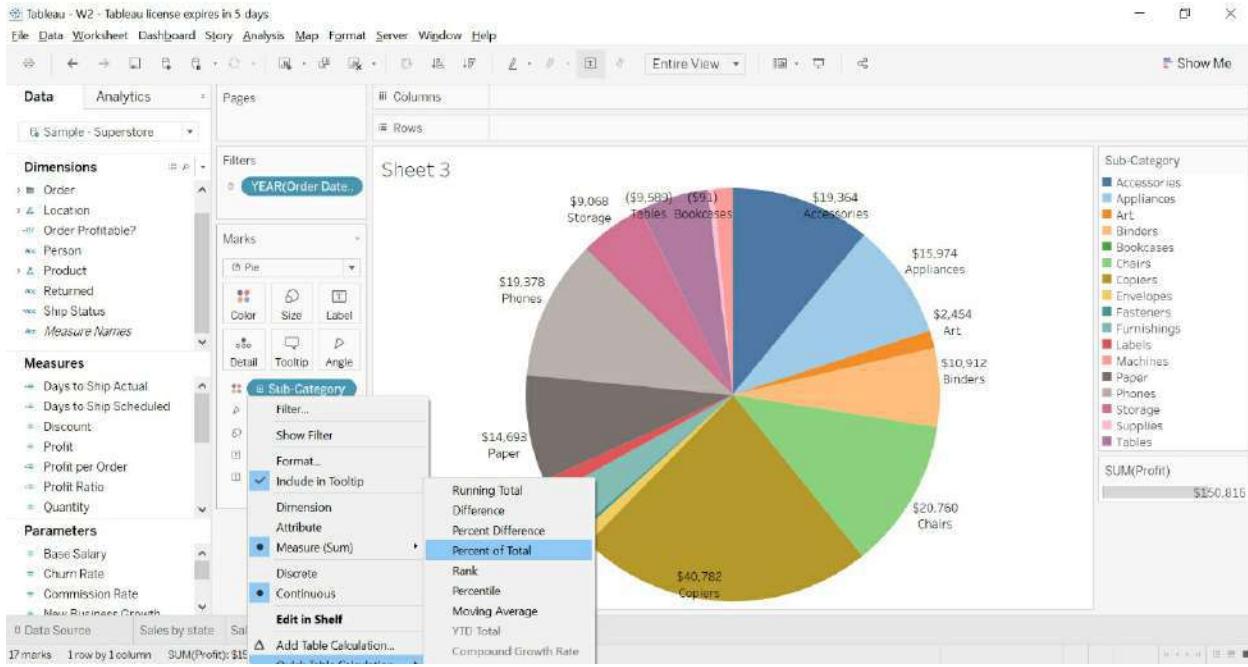
Click the **Size** property of the **Marks** card and put it on the middle line so the pie chart is bigger. Then click the dropdown on the top that says **Standard** and change it to **Entire View**.

Now for labels, again drag **Sub-Category** and **Profit** from the Data pane, but this time put them inside the **Label** property.

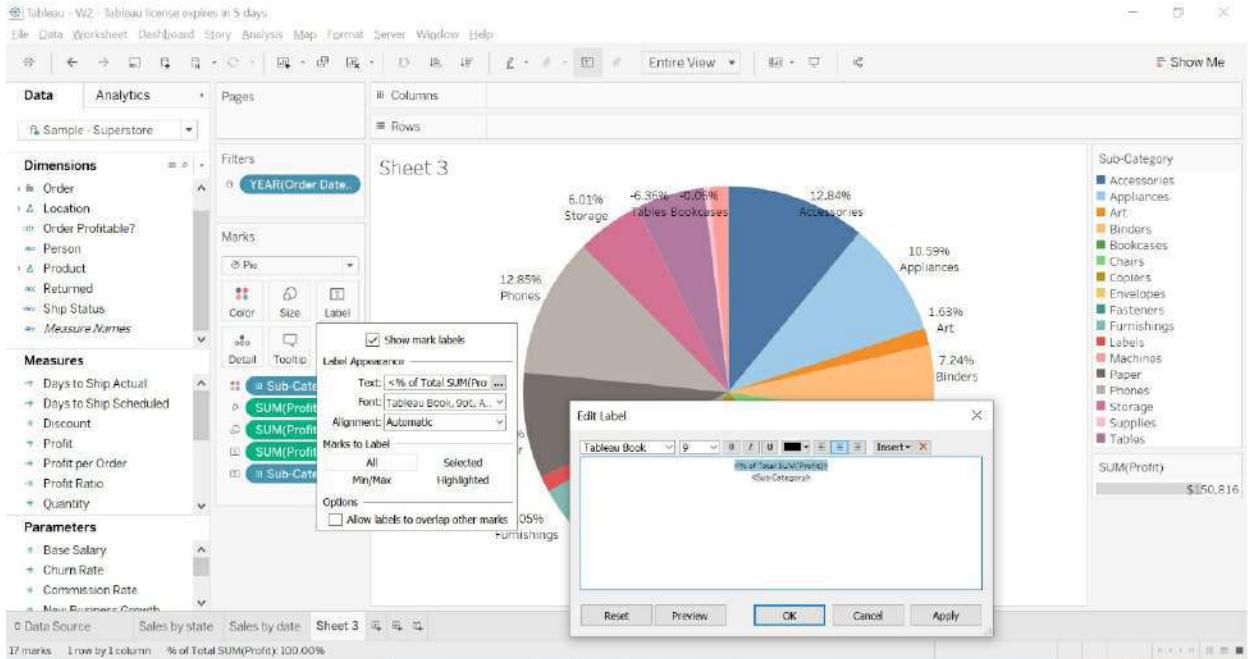
Tip

As an alternative, you can hold Ctrl and drag the Sub-Category and Profit pills inside **Marks** to Label. Ctrl works as a copy-paste function, if you did this without holding it, your original pills would disappear. You can also multi-select pills with this button.

Apply a table calculation on **Profit** to get its percentage:



Then click the **Label** property, and click the [...] button on the Text option to edit the actual label content:

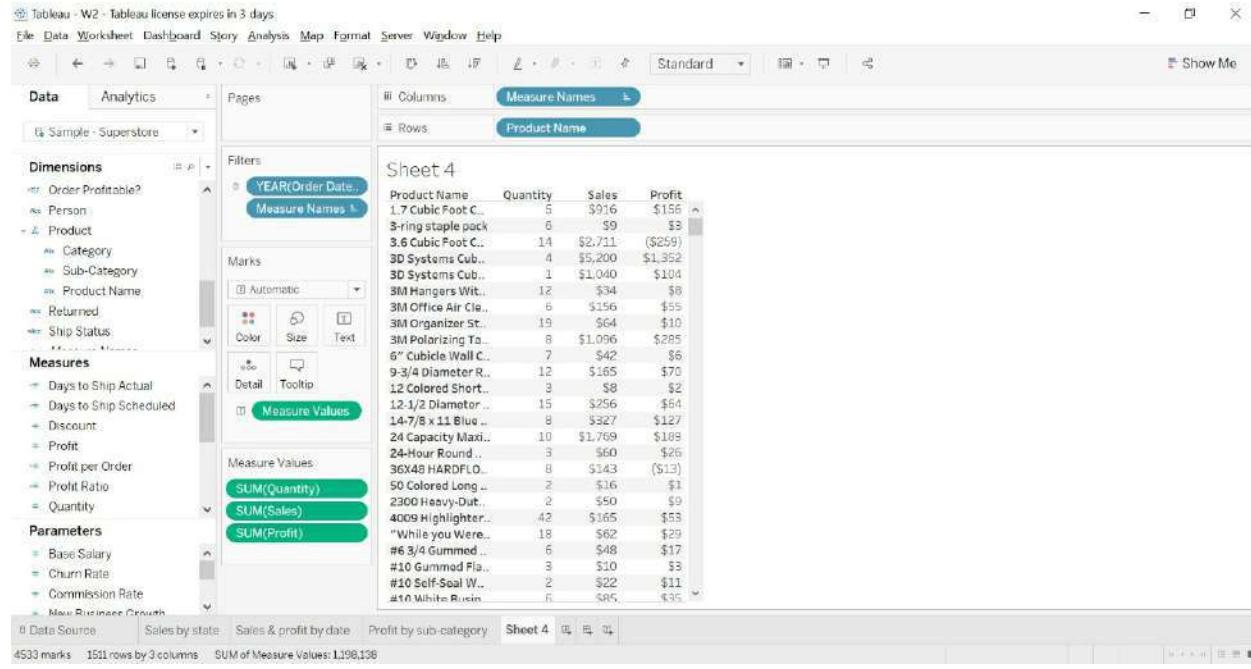


Increase font size for the first row from 9 to 12. Change the text color of the second row to a shade of gray. Then click Ok.

Rename the sheet to *Profit by Sub-Category*.

Create a new sheet, add **Product Name** to rows, then drag **Quantity** to the

canvas table, right on the ‘Abc’ cell (or you can do the usual of putting it inside the **Label** property from the **Marks** card). Then drag **Sales** over the quantity column, and finally, **Profit**.



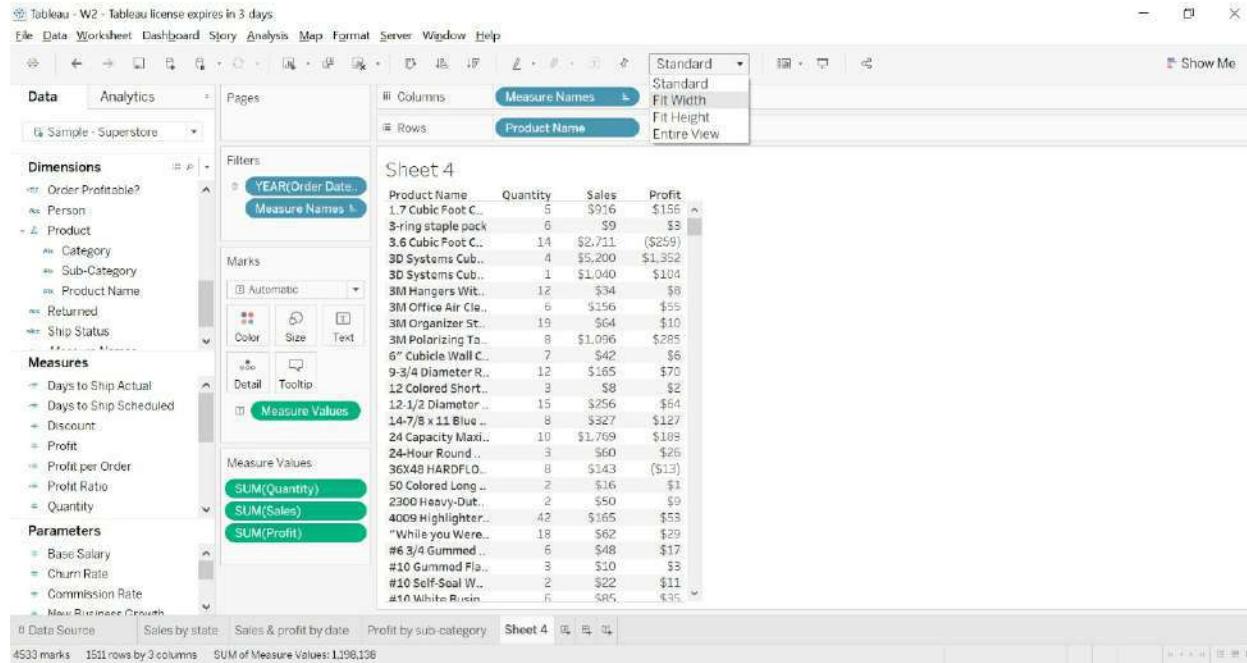
Explained

Measure Names and **Measure Values** are part of Tableau’s pre-generated dimensions and members. To be simple, we can say they work as containers for inputting more than one measure on a table. They are also useful when you, on a bar graph for example, you want to display more than two measures.

As you can see, they will appear on your shelves automatically when Tableau detects the need, so you will barely need to apply them yourself.

Notice that the global filter **YEAR (Order Date)** was automatically added, just as in the last graph.

You can adjust the width of the **Product Name** column by dragging from the right of it. However, let’s set the table to fit the width of the page, by clicking on the **Standard** dropdown. That’s the one that we previously set to **Entire View** back on the pie chart.

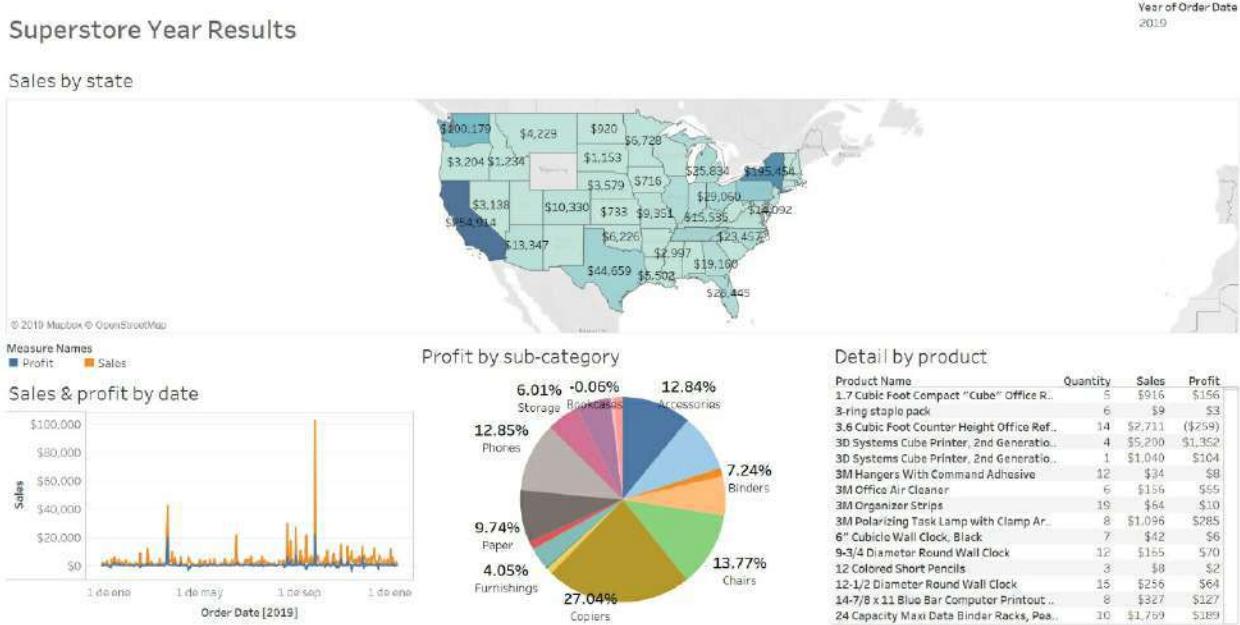


Explained

Standard sets the default size. **Fit width** stretches or tightens the graph to fit the full width of the page on which it is. **Fit height** does the same but for the page height. Finally, **Entire View** stretches or tightens the graph to fit width and height.

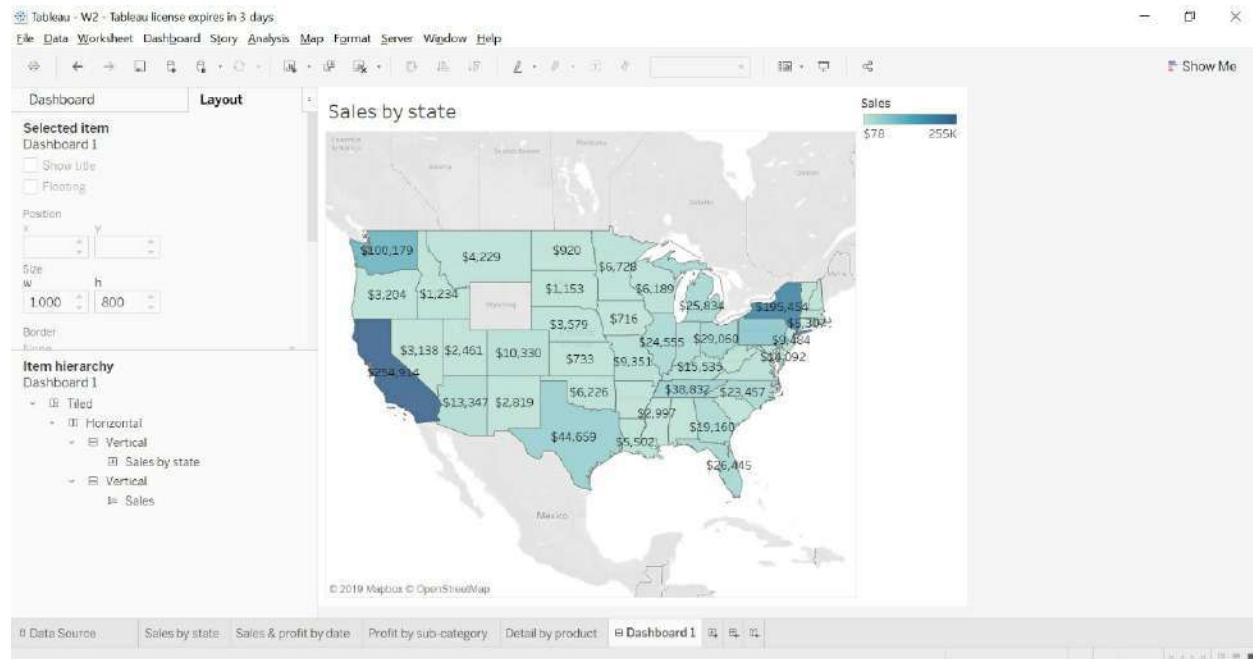
Call this sheet *Details by Product* and create a new dashboard.

We want to make a view that shows the sales by state at the top half and display the other three graphs on the bottom. If you feel like trying the **tiled** process by yourself, have below image as reference:



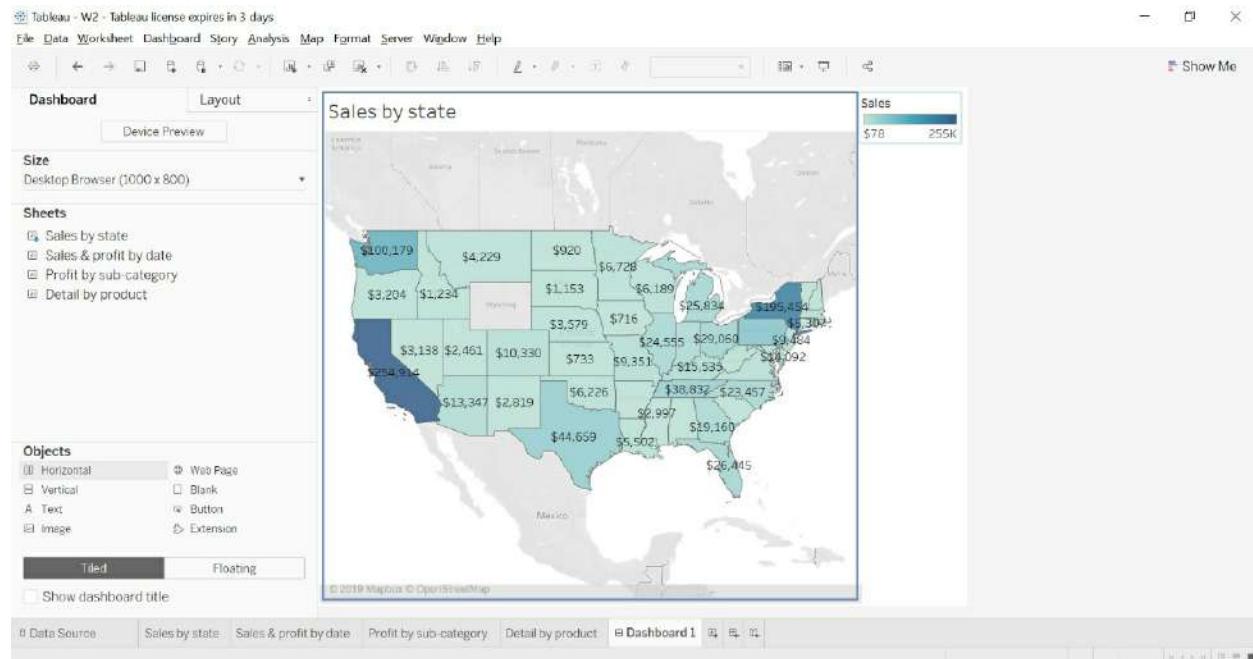
If not, continue the next steps:

With the blank dashboard, insert a **vertical** container and the *Sales by State* sheet inside it.



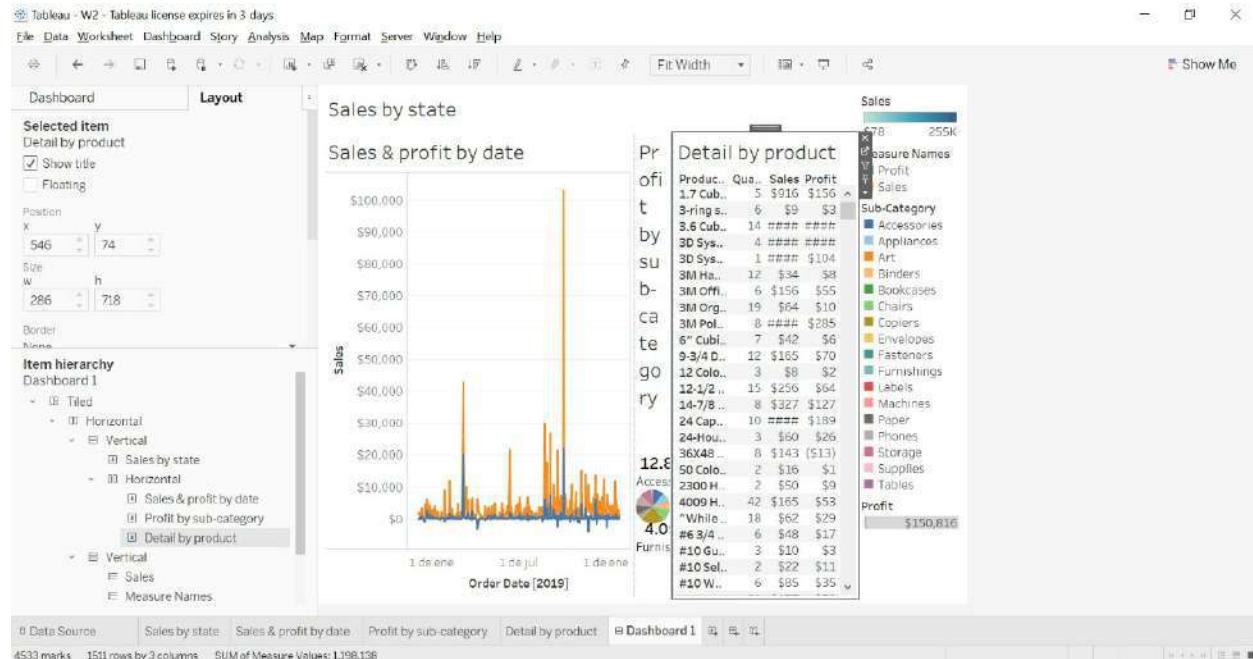
You can see that Tableau automatically generated the Vertical container on which all filters and legends will be added. Now add the **Horizontal** container below the map and remember to put it inside the vertical container by making sure it is highlighted and you see the shade at the bottom of it, like

this:



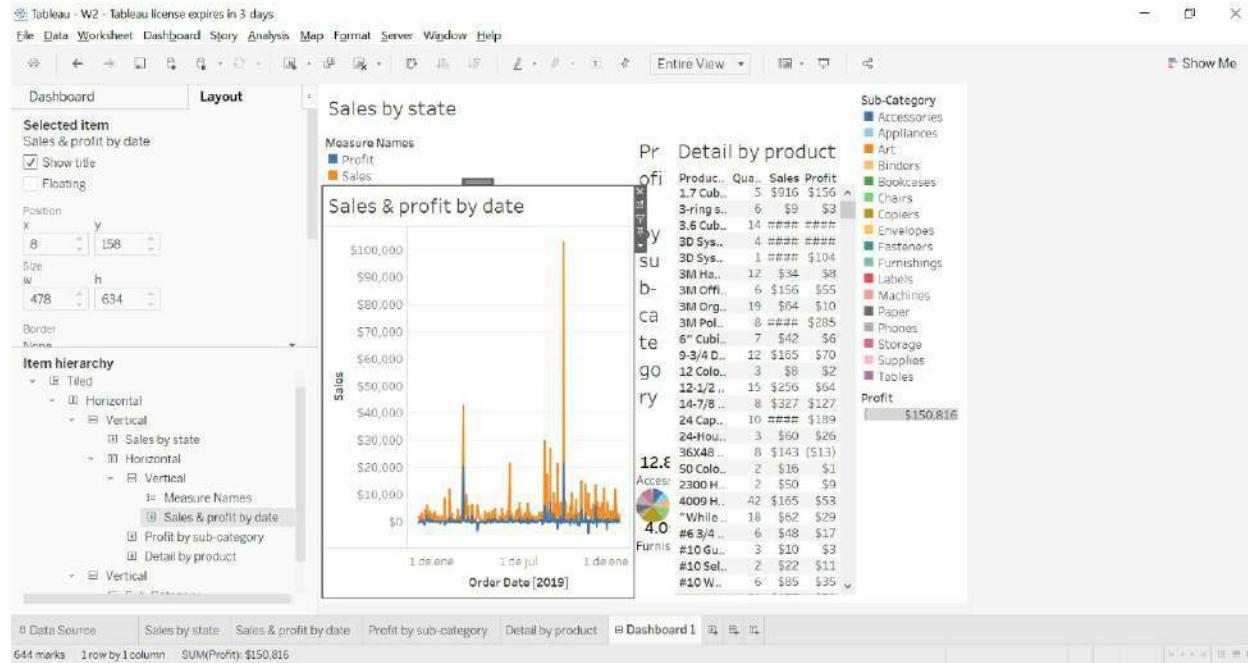
Don't worry if the size of objects moves drastically when inserting items, we will fix that at the end.

Now add the *Sales & profit by state*, *Profit by sub-category*, and the *Detail by product* graphs on the horizontal container in that order. You might end up with something messy like this:



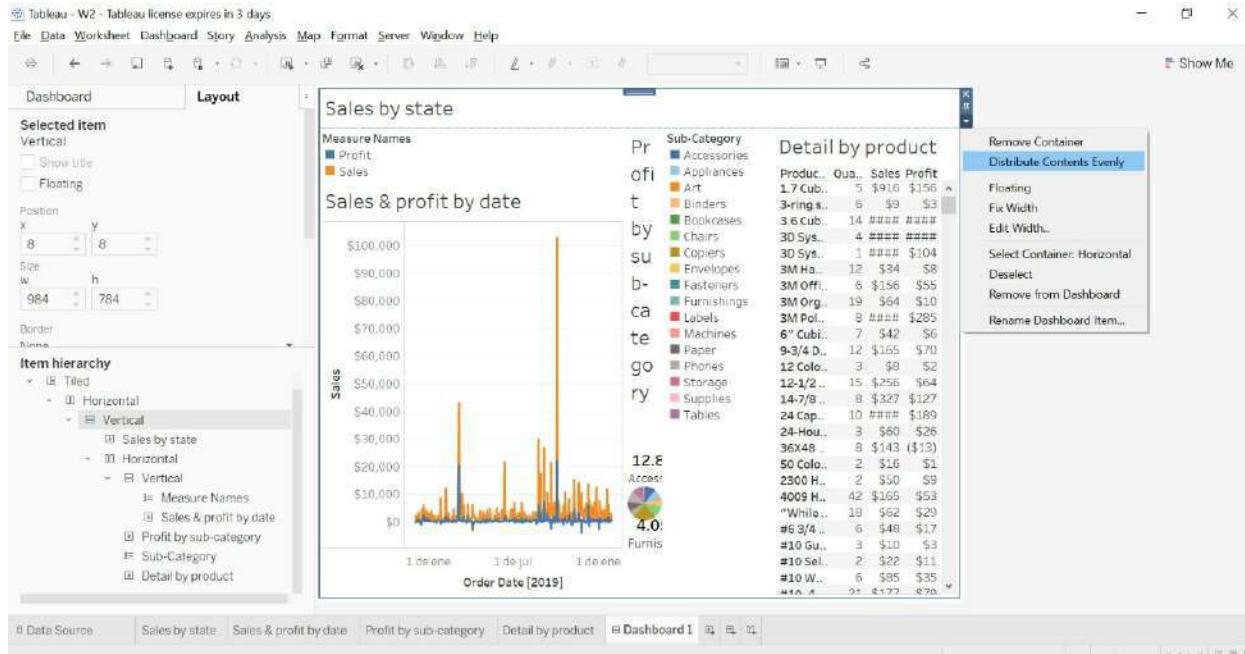
Now to clean up the layout, let's first get rid of the filters container from the

right. Click the Sales gradient at the top-right (the blue bar), and then close it to remove it from the dashboard. Next, insert a vertical container between the *Sales & profit by date* and *Profit by sub-category* graphs, and move the Measure Names legend from the top-right to this container, and move the *Sales & profit by date* graph to the bottom of it. Remember to check for the right gray shades when dragging items through containers, and keep verifying the Item hierarchy matches with the expected result below:



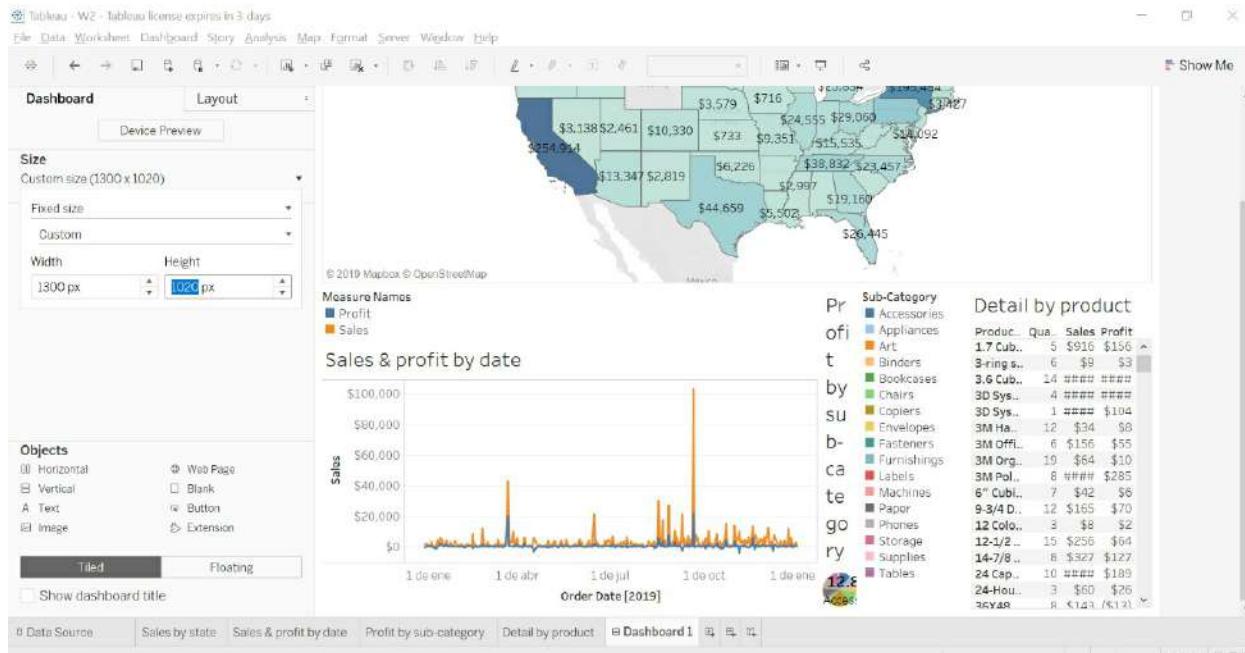
Since the **Sub-Category** legend is too big, and we already added this as part of the label, we'll remove it as well. Finally, delete the **Profit** legend, with that, your vertical filters' container should disappear.

Second step is adjusting the visibility of the dashboard elements. The containers provide functions to adjust the size of its elements. To access them, click from the **Item Hierarchy** the first **Vertical** container on the list. You will see that, once you do that, it gets highlighted on the dashboard canvas. Click the small arrow and select **Distribute Contents Evenly**:

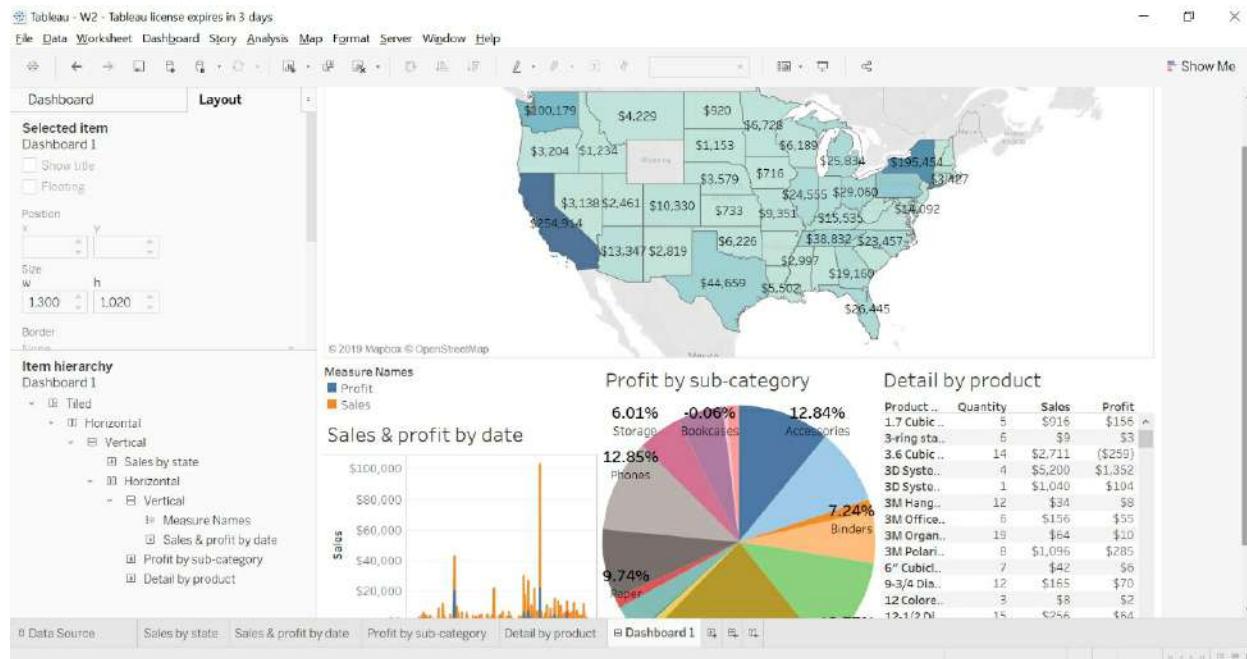


You will see that the map is visible again, or better adjusted.

Adjust the dashboard size to be 1300 x 1020 (or a size that you think can fit all dashboard elements):

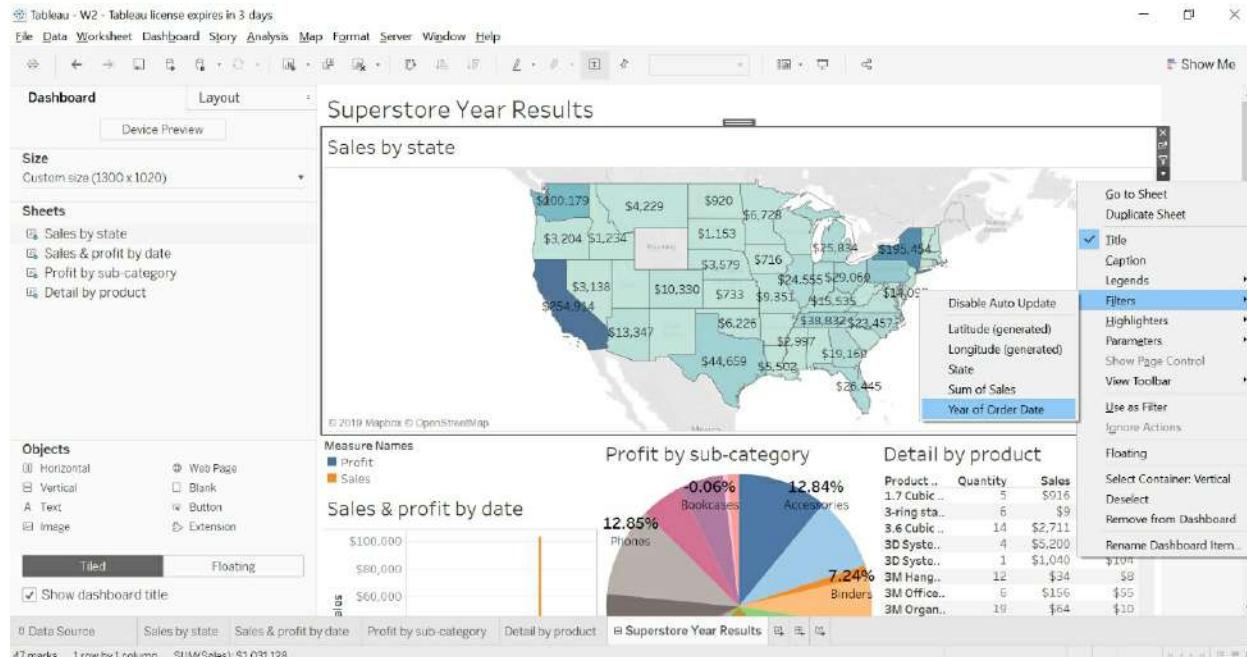


With the new size, select from the **Item Hierarchy** the **Horizontal** container with your three graphs, and again **Distribute the Contents Evenly**.



Our dashboard now has a better design! Rename it **Superstore Year Results**.

Remember the global filter we created, **YEAR (Order Date)**? Let's add it as a Quick Filter here. Click the map (or any graph), then the arrow, go to **Filters**, and select **YEAR of Order Date**:



The filter is now present in the dashboard, let's move it to the top. Add a **Horizontal** above *Sales by state*, add a **Text** element inside that says the dashboard title (just as the one that is already there) and then add the filter to

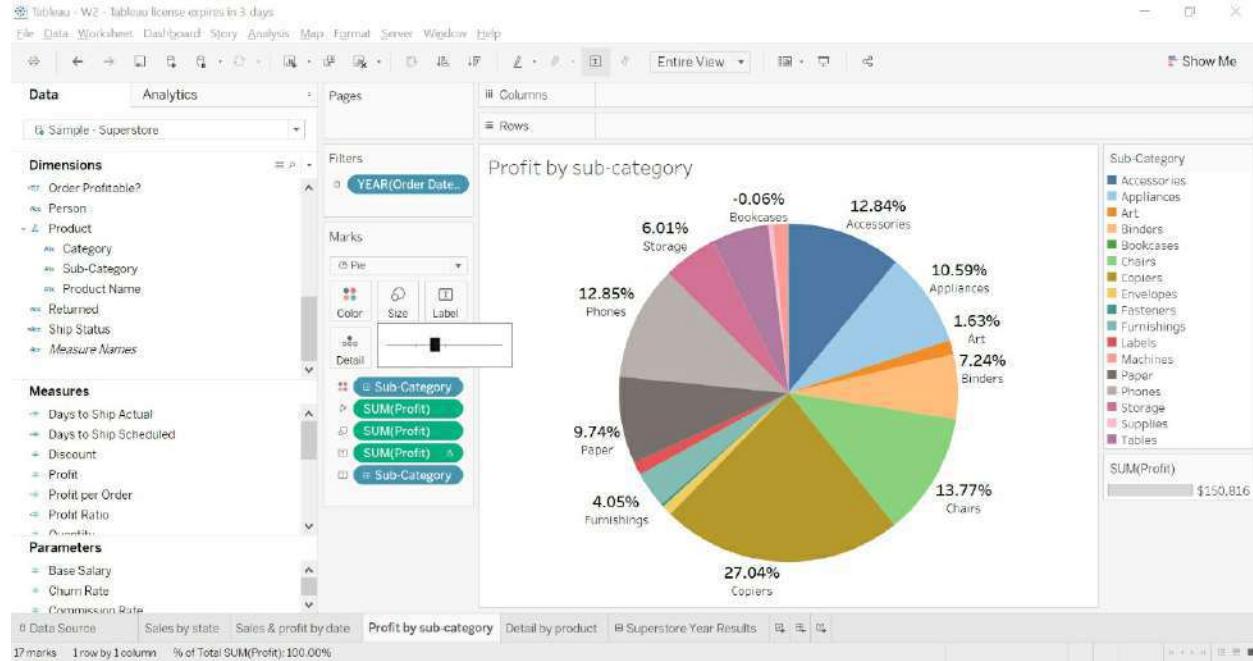
the right. Now click the filter and click the arrow that appears. Select **Single Value (slider)**, you will see that now instead of showing all possible values, it only displays a horizontal slider that lets you scroll through the years.

The screenshot shows the Tableau interface with a dashboard titled "Superstore Year Results". On the right, a context menu is open for the "Year of Order Date" filter. The "Single Value (slider)" option is highlighted, showing a horizontal slider for selecting the year. Other options like "Single Value (list)" and "Multiple Values (list)" are also visible.

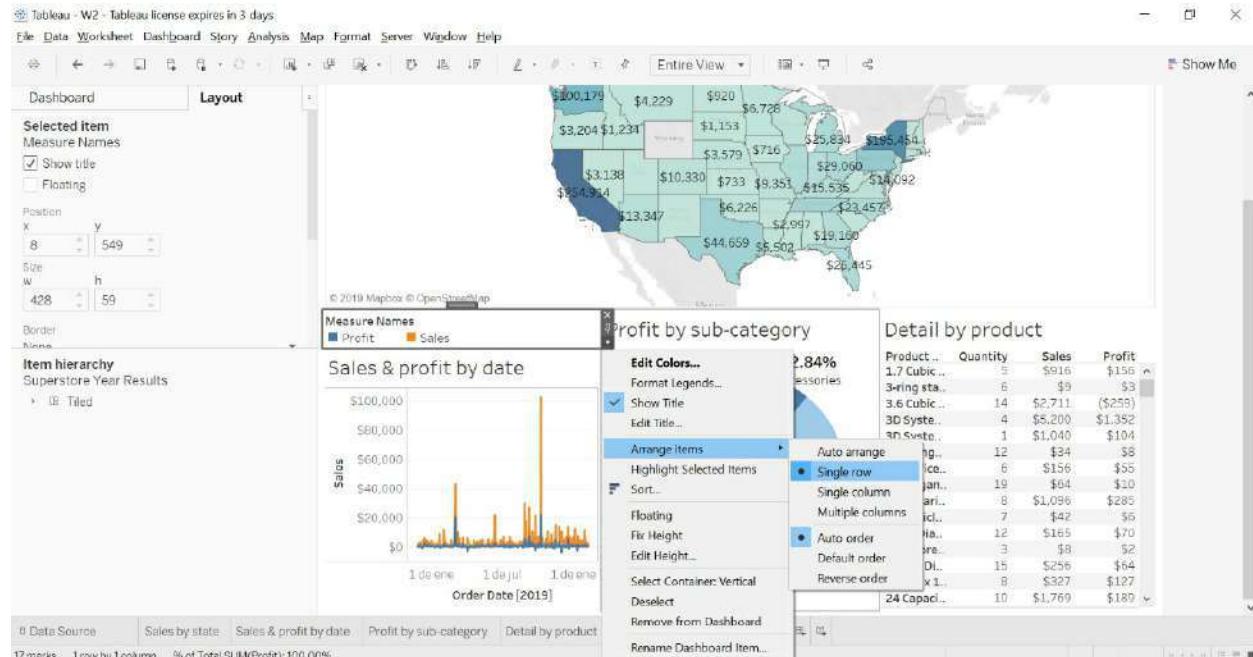
Seems like our previous setup of distributing items evenly on the **Vertical** container is making our new container too big, let's untoggle that. Now, select *Sales by state* and drag it from the bottom so it gets set a manual fixed height, and that it's visible again.

The screenshot shows the Tableau interface with the same dashboard. The "Vertical" container has been collapsed, as indicated by the "Vertical" entry in the context menu. The "Distribute Contents Evenly" option is selected in the context menu for the "Sales by state" card. The card is now much smaller and positioned above the other cards.

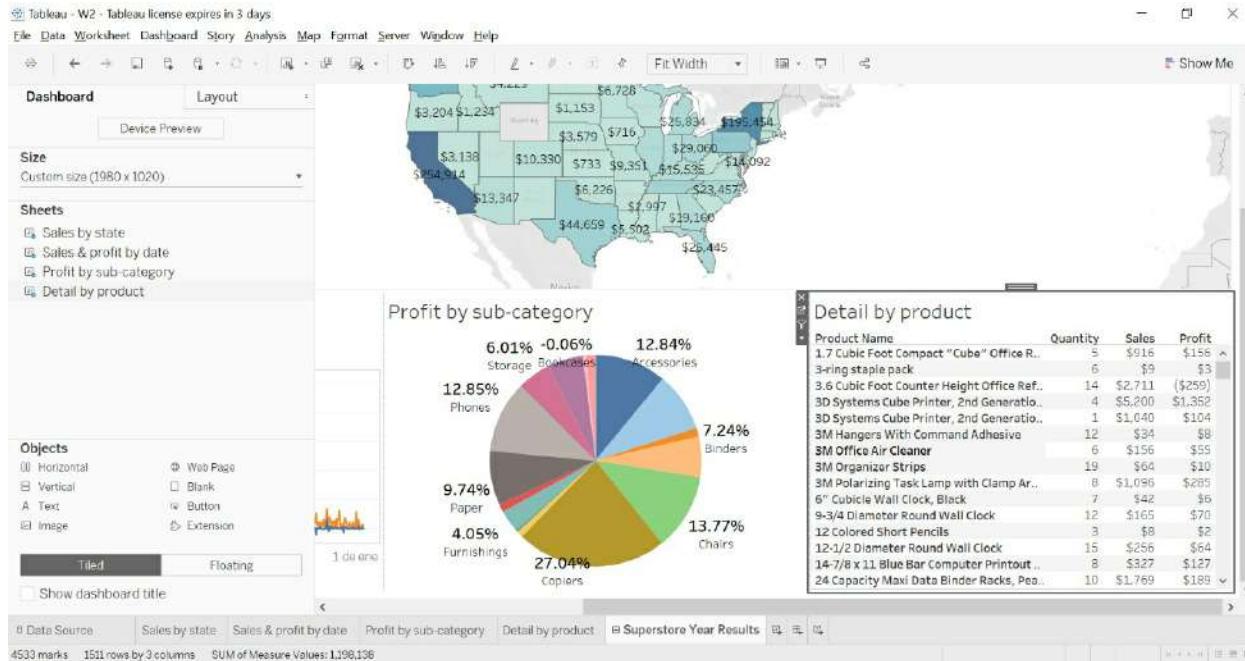
Now for the final touches on the design: go back to *Profit by sub-category* and reduce the size of the pie by selecting again the **Size** property from the **Marks** card.



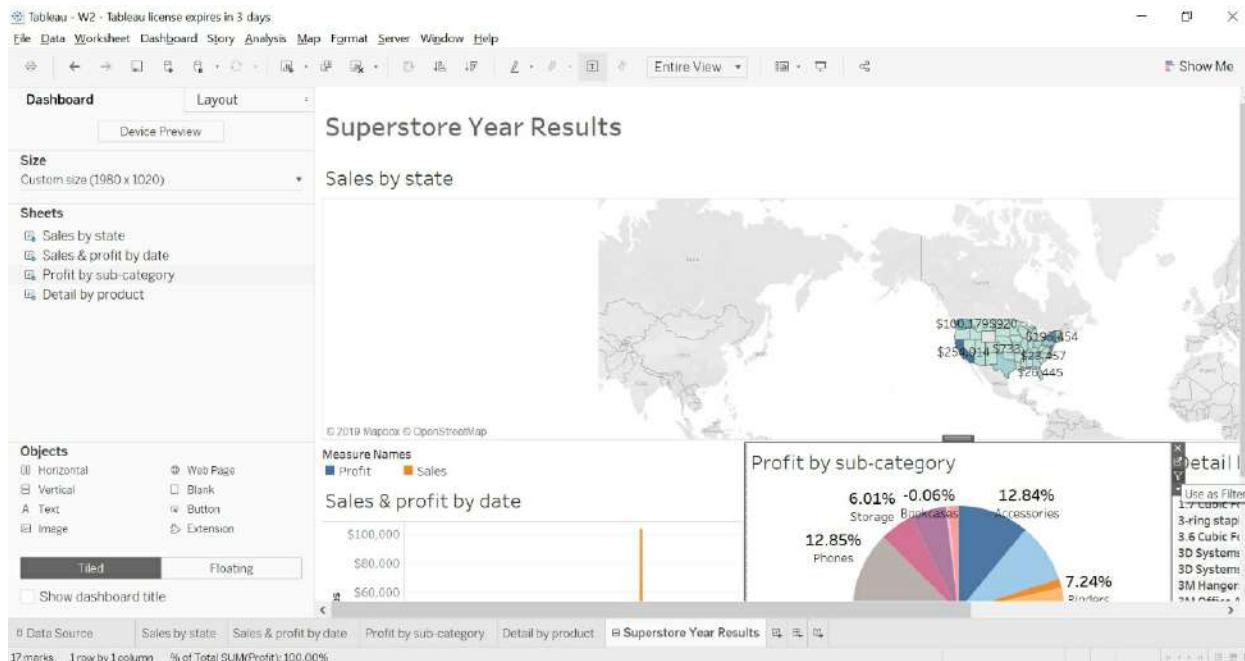
Now, back on the dashboard, select the **Measure Names** legend, and on the arrow go to **Arrange Items > Single row**:



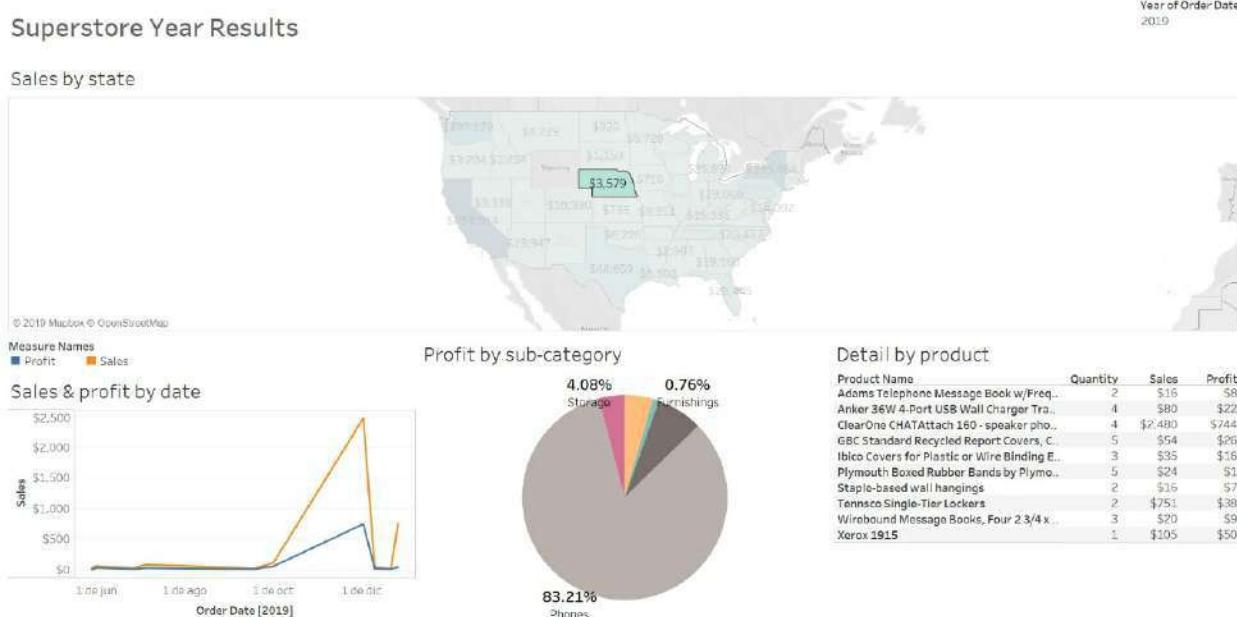
On *Detail by product*, expand the width of the product names so they're fully readable. You can also increase the width of your dashboard to 1980.



Now for *Sales by state* and *Profit by sub-category*, click the filter icon that shows above the arrow, called **Use as Filter**:



This will enable that users can select either a state from the map or a sub-category from the pie chart, and that all the other elements of the dashboard react to it as if you were applying a new filter. For example, selecting Nebraska from the map returns this:



This is one of the most powerful features of a dashboard; the ability for the user to interact with it in this intuitive way to, with one simple click, see the sales & profit for Superstore on Nebraska. Another example: selecting a product from *Detail by product* to see how it sells by states.

This is how users gain insights and value from dashboards, and as a BI developer, you can create different visuals that allow them to see their data in different ways. This is a relatively easy but applicable example of building dashboards for businesses.

Another note: if the user wanted to have the table of products sorted by quantity, sales, or profit, they can do so themselves by clicking on the column header.

Explained

Sorting is a feature that can be preconfigured by the developer or modified by the user. The first allows for more options, which include:

- Data source
- Alphabetical
- Value field
- Manual
- Nested

You can access this by clicking on the arrow of a dimension's pill and

choosing Sort.

Congratulations! That's the second workout from the list. Hope you were able to follow through with no roadblocks, on this workout we went a level deeper on graphs, as well as:

- Learning about your dataset: granularity
- Dimensional hierarchies: region – state – city
- Filled maps
- Show Me! helper
- Formatting labels
- Dual axis
- Global filters
- Formatting line graphs
- Pie graphs
- Table calculations: percentage of total
- Measure Names & Values
- Graph sizing: standard, fit width/height, entire view
- More on tiled dashboard design
- Use as Filter action
- Sorting

Have a break, recap what you just learned, and get ready for the third workout!

Workout 3. Airbnb's in New York

On this workout you will make an analysis of Airbnb's in New York.

Data



Design



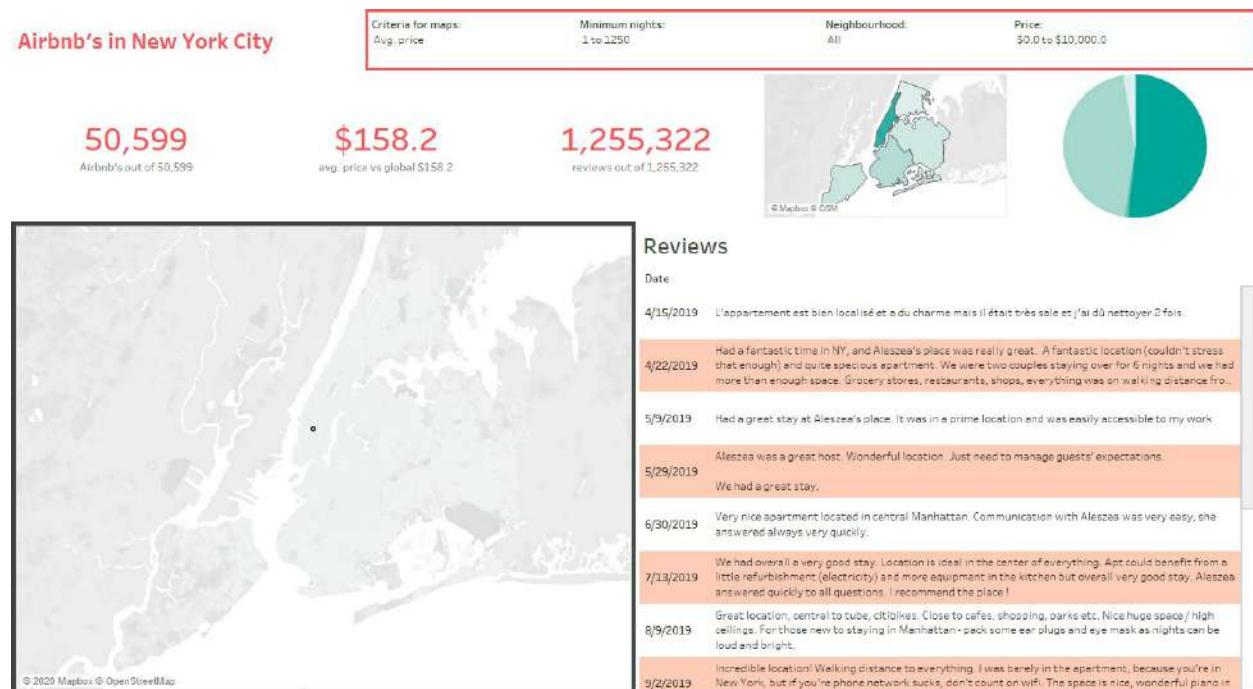
Graphs



Interactivity



Airbnb's in New York City



Workout Start

This time we will connect to two different tables, W3 Airbnb listings.csv and W3 Airbnb reviews.csv. Open both files in Excel or your preferred program and have a look at both:

File	Home	Insert	Page Layout	Formulas	Data	Review	View	Developer	Smart View	Help	DV Desktop	Search	Ponce, Carlos (GE Renewable Energy)							
Q1	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Id	name	host_id	host_name	neighbourhood	neighbourhood_group	latitude	longitude	room_type	price	minimum_nights	number_of_reviews	last_review	reviews_per_month	calculated_availability_365	host_since	host_id	host_name	host_since	host_id
2	2550	Skyline Midtown	2845	Jennifer	Manhattan	Midtown	40.75362	-73.98377	Entire home	225	10	48	11/4/2019	0.39	1	1	2550	Skyline Midtown	2845	Jennifer
3	3831	Cool Entire 6	4869	Lisayloanna	Brooklyn	Clinton Hill	40.68514	-73.95976	Entire home	89	1	295	11/22/2019	4.67	1	1	3831	Cool Entire 6	4869	Lisayloanna
4	5059	Large Cozy 1	7232	Chris	Manhattan	Murray Hill	40.74767	-73.9575	Entire home	200	3	78	10/13/2019	0.6	1	19	5059	Large Cozy 1	7232	Chris
5	5112	BilletsArtSpa	7356	Garon	Brooklyn	Bedford-Stuyvesant	40.69888	-73.95956	Private room	60	29	49	10/5/2017	0.88	1	365	5112	BilletsArtSpa	7356	Garon
6	5178	Large Furnished	6967	Shurmin	Manhattan	Hell's Kitchen	40.76489	-73.98493	Private room	79	2	454	11/21/2019	3.52	1	242	5178	Large Furnished	6967	Shurmin
7	5208	Cute Clean 6	7490	Maryann	Manhattan	Upper West Side	40.80136	-73.96723	Private room	79	2	118	7/21/2017	0.85	1	0	5208	Cute Clean 6	7490	Maryann
8	5238	Cute & Cozy	7549	Ben	Manhattan	Chinatown	40.71344	-73.99037	Entire home	150	1	161	7/29/2019	1.29	4	1	5238	Cute & Cozy	7549	Ben
9	5441	Central Man	7585	Kate	Manhattan	Hell's Kitchen	40.76076	-73.98667	Private room	85	3	204	11/20/2019	1.56	1	10	5441	Central Man	7585	Kate
10	5803	Lovely Room	9741	Laurie	Brooklyn	South Slope	40.65629	-73.99779	Private room	89	4	175	10/23/2019	1.35	3	0	5803	Lovely Room	9741	Laurie
11	6094	West Village	13975	Alyne	Manhattan	West Village	40.7553	-74.00525	Entire home	120	180	27	10/31/2018	0.22	1	271	6094	West Village	13975	Alyne
12	6848	Only 2 stops	15991	Allison & Irina	Brooklyn	Williamsburg	40.70887	-73.95523	Entire home	140	2	170	11/22/2019	1.33	1	0	6848	Only 2 stops	15991	Allison & Irina
13	7057	Perfected for Y	17371	Jane	Brooklyn	Fort Greene	40.65939	-73.97183	Entire home	199	1	213	12/4/2019	1.79	1	0	7057	Perfected for Y	17371	Jane
14	7322	Chelse Perla	18946	Dotti	Manhattan	Chelsea	40.74192	-73.95501	Private room	120	1	296	11/16/2019	2.32	1	26	7322	Chelse Perla	18946	Dotti
15	7726	Hip Historic	20550	Adam And C	Brooklyn	Crown Heights	40.67592	-73.94654	Entire home	99	3	75	11/30/2019	4.43	1	63	7726	Hip Historic	20550	Adam And C
16	7750	Huge 2 Br Up	17985	Sing	Manhattan	East Harlem	40.79685	-73.94872	Entire home	190	7	0			2	350	7750	Huge 2 Br Up	17985	Sing
17	7801	Sweet and S	21207	Chaya	Brooklyn	Williamsburg	40.71842	-73.95718	Entire home	299	3	9	12/28/2011	0.07	1	1	7801	Sweet and S	21207	Chaya
18	8024	CBD CityBed	22485	Usel	Brooklyn	Park Slope	40.68069	-73.97078	Private room	115	1	148	10/27/2019	1.15	4	0	8024	CBD CityBed	22485	Usel
19	8025	CBD Helps H	22486	Usel	Brooklyn	Park Slope	40.67985	-73.97978	Private room	80	1	41	11/20/2019	0.37	4	150	8025	CBD Helps H	22486	Usel
20	8110	CBD Helps H	22486	Usel	Brooklyn	Park Slope	40.68001	-73.97865	Private room	32	2	77	10/4/2019	0.64	4	0	8110	CBD Helps H	22486	Usel
21	8450	MAISON DES	25183	Nathally	Brooklyn	Bedford-Stuyvesant	40.68271	-73.94028	Entire home	120	2	105	11/17/2019	0.84	2	296	8450	MAISON DES	25183	Nathally
22	8505	Sunny Bedr	25125	Gregory	Brooklyn	Windsor Terrace	40.65599	-73.97519	Private room	60	1	23	10/14/2019	1.22	2	78	8505	Sunny Bedr	25125	Gregory
23	8700	Magnifique !	26934	Claude & So	Manhattan	Inwood	40.86754	-73.92639	Private room	80	4	0			1	0	8700	Magnifique !	26934	Claude & So
24	9518	SPACIOUS, L	31374	Shon	Manhattan	Inwood	40.86842	-73.92108	Private room	44	3	114	11/11/2019	1.12	1	253	9518	SPACIOUS, L	31374	Shon
25	9657	Modern 1 BR	21304	Dana	Manhattan	East Village	40.73292	-73.98542	Entire home	180	10	33	11/11/2019	0.26	1	86	9657	Modern 1 BR	21304	Dana
26	9668	front room/	32294	Sameer Or	Manhattan	Harlem	40.82245	-73.95104	Private room	50	3	256	11/3/2019	2.07	2	361	9668	front room/	32294	Sameer Or
27	9704	Spectious 1 b	32045	Teri	Manhattan	Harlem	40.81205	-73.95466	Private room	52	2	97	11/9/2019	1.45	1	326	9704	Spectious 1 b	32045	Teri
28	9782	Loft in Willies	32169	Andreas	Brooklyn	Greenpoint	40.72219	-73.97302	Private room	55	4	212	11/18/2019	1.71	3	331	9782	Loft in Willies	32169	Andreas
29	9783	back room/E	32294	Sameer Or	Manhattan	Harlem	40.8213	-73.95318	Private room	50	3	280	10/50/2019	2.37	2	1	9783	back room/E	32294	Sameer Or
30	10452	Large B&B S	35935	Angela	Brooklyn	Bedford-Stuyvesant	40.68331	-73.95457	Private room	70	1	75	10/12/2019	0.64	2	251	10452	Large B&B S	35935	Angela
31	10962	Lovely room	57441	Laura	Brooklyn	South Slope	40.68685	-73.98787	Private room	89	4	177	11/2/2019	1.43	3	349	10962	Lovely room	57441	Laura
32	11174	HUGE Loft! C	40729	Madek	Brooklyn	Bushwick	40.70457	-73.91869	Entire home	98	2	70	11/28/2019	0.56	1	8	11174	HUGE Loft! C	40729	Madek
33	11708	Cute apt in a	44145	Tyrome	Brooklyn	Bushwick	40.70838	-73.92745	Entire home	85	2	253	11/24/2019	2.06	2	43	11708	Cute apt in a	44145	Tyrome
34	11943	Country spec	45445	Harriet	Brooklyn	Flatbush	40.67027	-73.96327	Private room	150	1	0			1	365	11943	Country spec	45445	Harriet
35	12048	Lovely EastSide	7549	Ben	Manhattan	Lower East Side	40.71401	-73.98817	Shared room	40	1	219	11/13/2019	1.78	4	0	12048	Lovely EastSide	7549	Ben
36	12192	ENJOY Down	46979	Edward	Manhattan	East Village	40.729	-73.98199	Private room	68	2	258	11/13/2019	2.1	2	62	12192	ENJOY Down	46979	Edward

The first thing we see is that the listings table has a **granularity** by the Id

column. We can see this refers to each Airbnb present in New York City. The reviews table has a granularity by review, and it also includes the listing_id, we can see that this will be our link between the two tables: **Id** from listings.csv and **listing_id** from reviews.csv.

Connect to a Text file and look for **W3 Airbnb listings.csv** from the materials you downloaded. You should end up with a window like this (ignore the differences in table names):

The screenshot shows the Tableau Data Source interface with the 'listings' worksheet selected. The 'listings.csv' file is loaded. The data preview table has the following structure:

#	Airbnb ID	Airbnb Name	Airbnb Host ID	Airbnb Host Name	Airbnb Neighbourhood Group	Airbnb Neighbourhood	Airbnb Latitude	Airbnb Longitude	Airbnb Room Type	Airbnb Price	Airbnb Minimum Nights	Airbnb Number of Reviews
2595	Skylit Midtown Castle	2845	Jennifer	Manhattan	Midtown	40.753620	-73.983770	Entire home/apt	225	10		
3831	Cozy Entire Floor of B...	4868	Lisa Roxanne	Brooklyn	Clinton Hill	40.685140	-73.955780	Entire home/apt	89	1		
5099	Large Cozy 1BR Apt...	7322	Chris	Manhattan	Murray Hill	40.747670	-73.975000	Entire home/apt	200	3		
5121	BlissArtsSpace!	7356	Gérón	Brooklyn	Bedford-Stuyvesant	40.695880	-73.955960	Private room	60	29		
5178	Large Furnished Roo...	8967	Shunichi	Manhattan	Hells Kitchen	40.764890	-73.984880	Private room	79	2		
5203	Cozy Clean Guest Roo...	7490	MaryEllen	Manhattan	Upper West Side	40.801780	-73.967230	Private room	79	2		
5238	Cute & Cozy Lower Ea...	7549	Ben	Manhattan	Chinatown	40.713440	-73.990370	Entire home/apt	150	1		
5441	Central Manhattan N...	7988	Kate	Manhattan	Hells Kitchen	40.760760	-73.988670	Private room	85	3		
5803	Lovely Room 1, Bedr...	8744	Laurie	Brooklyn	South Slope	40.688290	-73.987790	Private room	89	4		
6090	West Village Nest - Su...	11975	Aline	Manhattan	West Village	40.735300	-74.005250	Entire home/apt	120	180		
6948	Only 2 steps to Manha...	15991	Allen & Irina	Brooklyn	Williamsburg	40.706370	-73.953820	Entire home/apt	140	2		
7097	Perfect for Your Par...	17571	Jane	Brooklyn	Fort Greene	40.691690	-73.971850	Entire home/apt	199	1		

We will now proceed to do a **data join**, one of the two ways to connect different tables. You should see on the left pane the rest of the files on your Materials folder, look for **W3 Airbnb reviews.csv** and drag it to the right of listings.csv inside the canvas. It will then look like below:

Now click on the two dotted circles between both tables, make sure the join is set to **Inner**, then select on the first dropdown the **Id** column and select **Listing Id** on the second dropdown. With this, you will be setting the column that is the link between the two tables.

Explained

Tableau has two ways of combining different tables: data join and data blending.

Data join lets you add the columns of one table (secondary table) to another one (primary table). It works by

- 1) Specifying a **link** with which both tables will connect.
- 2) Specifying if it's inner, left, right, or full outer join.

The **link** should be the column with distinct values or a combination of columns that return distinct values for both tables. Therefore, the link is usually the table **Id**, as we saw on our listings example. It's important to make sure it's distinct, because if not, you will have duplicate numbers.

On **join types**:

Inner	Only rows where a link value/combination is found on both tables.
Left	Inner + the rows on the primary table not found on the secondary.
Right	Inner + the rows on the secondary table not found on the primary.
Full outer	Includes all rows, regardless if matched or not.

An example of a case where you need data join is when you have two tables, both by year-country level, in one you have historic population data and in the other you have historic GDP data. On that case, your data join link would be the **Year** and the **Country** columns, and that combination will return distinct values for both tables.

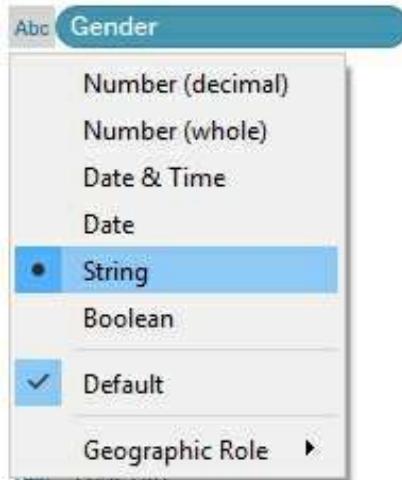
IMPORTANT: There are cases when data join is not suited for the job, and you must use the alternative data blending, we will explain such cases later.

Did you get an error? That might happen, and it's ok. If you hover the mouse over the circles, you will see that the error is related to **data types**.

The screenshot shows the Tableau Data Source pane. On the left, there's a sidebar with 'Connections' (listing 'listings'), 'Files' (listing 'listings.csv', 'neighbourhoods.csv', 'reviews.csv', 'reviews_1.csv'), and 'New Union'. The main area shows two connections: 'listings.csv' and 'reviews_1.csv'. A join line connects them on the 'listing_id' field. A warning message at the top right says: 'Type mismatch between join fields [listings.csv].[id] [Number (whole)] and [reviews_1.csv].[listing_id] [String].'. Below the connections, there's a table of fields with icons indicating their data types: 'listing_id' (blue pill, ABC), 'Name' (green pill, #), 'Host Id' (blue pill, ABC), 'Host Name' (green pill, #), 'Neighbourhood Gr...' (blue pill, ABC), 'Neighbourhood' (green pill, #), 'Latitude' (green pill, T/F), 'Longitude' (green pill, T/F), 'Room Type' (blue pill, ABC), 'Price' (green pill, #), 'Minimum Nights' (green pill, #), and 'Number Of Re...' (green pill, #).

Explained

If you go to a previous workout and look on the data pane, you will see that our dimensions and measures are not only classified by discrete or continuous (if they're blue or green pills), but they also have an icon to the left.



The 'Abc', the '#', the 'T|F', and the date and world icons represent the different **data types** Tableau uses to identify columns based on their contents. This is automatically done, and since on the previous two workouts it had been done correctly there was no need to change it.

- Number (decimal): allows numbers with decimals like 1.1, 20.99, 100.0001.
 - Number (whole): allows numbers without decimals like 1, 20, 100.
 - Date & Time: for date-time fields like ‘01/01/98 23:59:59.995’.
 - Date: for date fields like ‘01/01/1900’.
 - String: for any text field.
 - Boolean: only allows 1 or 0, which is the equivalent to TRUE or FALSE.
 - Geographic: allows for any value that is related to anything geographic, such as country, city, ZIP code, and more.
-

Now to fix that, we must change the data type of one of the two columns we used as link. Below the canvas you can see the table with the result of the data join, which is currently empty because of the error. But if you go search the **Id** column you will see that it’s set as **numeric** type (#), while **Listing Id** is set as **string** (Abc).

The screenshot shows the Tableau Data Source interface. On the left, the 'Connections' pane lists 'listings' and 'reviews'. The main area shows a join between 'listings.csv' and 'reviews_1.csv'. A join line connects the 'listing_id' field from 'listings.csv' to the 'listing_id' field in 'reviews_1.csv'. The join type is set to 'inner'. The 'listing_id' fields are currently defined as string type (Abc). The 'reviews_1.csv' connection shows fields like 'listing_id', 'reviewer_id', 'date', 'reviewer_name', 'comments', and 'rating'. The 'listings.csv' connection shows fields like 'id', 'name', 'host_id', 'neighbourhood', 'neighbourhood_group', 'latitude', 'longitude', 'availability_365', and 'price'.

Change **Listing Id** from string to numeric by clicking the ‘Abc’ icon and selecting **Numeric (whole)**. Now the error will go, and your table will be populated. Also, change the data join to be **left** instead of **inner** so we don’t miss any Airbnb without review.

The screenshot shows the Tableau Data Source view. A join is established between 'listings.csv' and 'reviews_1.csv'. The 'reviews_1.csv' connection is selected, indicated by a blue border around its row in the connections list. The resulting table has 1,000 rows.

Have a look at the resulting table and notice something. I will give you one clue: it has to do with the measures and the data join we just did. You can remove the reviews table from the canvas and press Ctrl + Z to compare the end table so you can see what happens with our number of rows...

The screenshot shows the Tableau Data Source view with a context menu open over the joined table. The 'Remove' option is highlighted in blue. The menu also includes 'Field names are in first row', 'Generate field names automatically', 'Text File Properties...', 'Convert to Union...', and 'Duplicates'.

Have you seen it yet? If not, open Sheet 1, and to show the price for a random Airbnb on a table. Move **Name** to filters, pick any value, then move **Name** to

Rows and **Price** to Text. Now compare that price with that in the original csv file. Do you see what is happening now? The price is duplicating times the number of reviews of the Airbnb!

This is why, to use data joins correctly, it is crucial that you understand the concept of **granularity** of your tables, and that you make sure that the columns you are using for the link actually give **unique combinations** of values per row, on both your primary and secondary table.

When the rule above does not apply, you can use the **data blending** approach.

Explained

Data blending is the second approach for connecting different tables.

Unlike the data join, this does not require that your primary and secondary table have the same granularity. It adds other limitations, though, as not being able to sort a table, locking aggregation functions like COUNTD (count distinct), and making calculations based on fields from both tables more complicated.

As an additional note, there also exists **data union**, which we won't cover on this workout, but it's a way to combine tables that have the exact same columns. This is useful when you have many Excel files of your monthly sales, and each one is for a month. You'd like to have them all combined as a data source, so that's what you use to do so instead of making a new Excel file and copy-pasting every month.

Let's use data blending for our example: remove the reviews table from the canvas, click the cylinder icon and click New Data Source.

The screenshot shows the Tableau Data Source interface for the 'listings' connection. The top navigation bar includes 'File', 'Data', 'Server', 'Window', and 'Help'. Below the navigation is a 'Connections' section with 'listings' selected. A 'New Data Source' button is highlighted. The main area displays a preview of the 'listings.csv' data with columns: 'id', 'name', 'host_id', 'host_name', 'neighbourhood_group', 'neighbourhood', 'latitude', 'longitude', 'room_type', 'price', 'minimum_nights', and 'availability_30'. The preview shows 1000 rows of data. At the bottom, there are tabs for 'Data Source' (selected), 'Sheet 1', and other sheet icons.

Add there your reviews table. This time also click on the “Extract” checkbox of the top-right. You should have a result close to this:

The screenshot shows the Tableau Data Source interface for the 'reviews' connection. The top navigation bar includes 'File', 'Data', 'Server', 'Window', and 'Help'. Below the navigation is a 'Connections' section with 'reviews_1' selected. A 'New Data Source' button is present. The main area displays a preview of the 'reviews_1.csv' data with columns: 'listing_id', 'reviewer_id', 'date', 'comments', 'f7', 'fb', and 'f9'. The preview shows 1000 rows of data. At the bottom, there are tabs for 'Data Source' (selected), 'Sheet 1', and other sheet icons. The 'Connection' section at the top right shows 'Extract' checked.

Explained

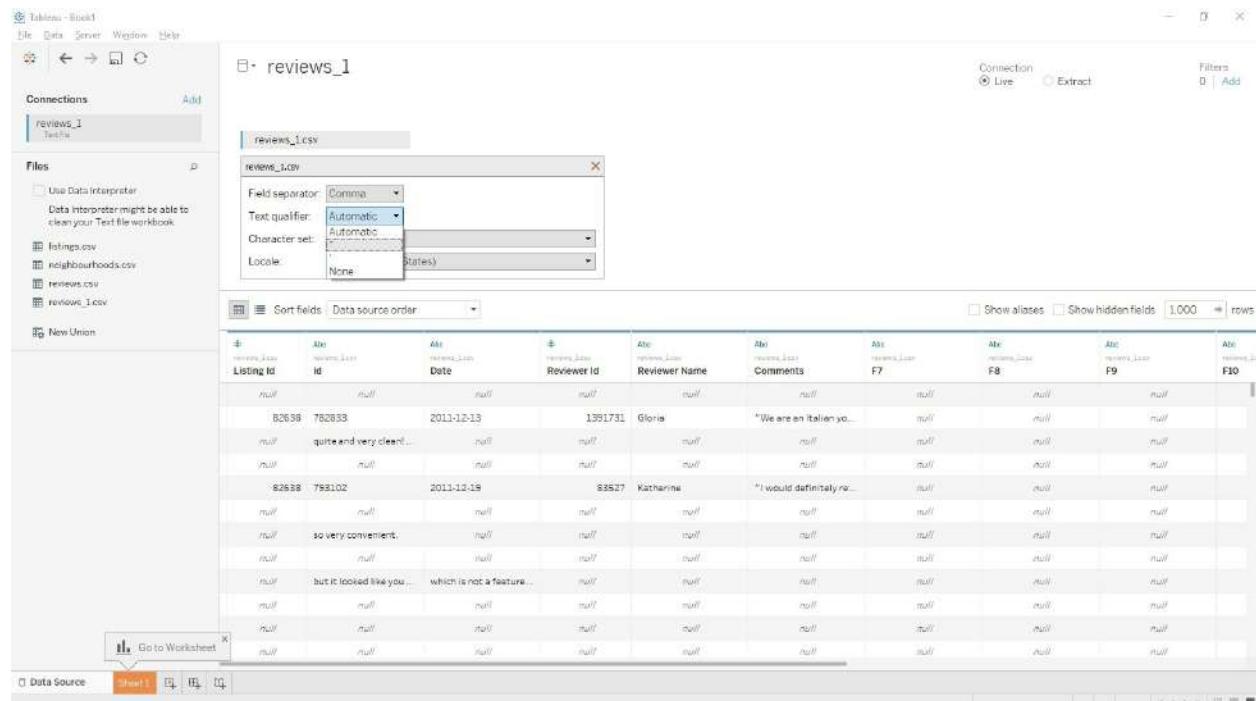
Data extracts are useful to have faster data sources and overall faster

dashboards when working with big tables. When you see that working with your dashboard becomes slow, try changing the connection to be extract instead of live.

What it does is that it takes a snapshot of your data source and stores it in a friendlier format called **.hyper**. Since the source files we're working with now are fixed, this means no change. But if you were connected to a database being constantly updated, this would mean you have to refresh the extract every time you want to see the latest entries.

Notice a new error on the reviews source; the end table has a lot of nulls. It seems like it's messing up with the comments column, it's cutting them and moving the parts. This is because Tableau is reading every comma inside the comments themselves and defining them as the value for the next column (because csv files are comma-separated value).

To fix this, hover your mouse on the reviews table in the canvas, click the arrow on the right, and click Text File Properties. Change the text qualifier to be “ instead of automatic.



The screenshot shows the Tableau Data Source editor. On the left, there's a sidebar with 'Connections' (selected), 'Files' (listing 'listings.csv', 'neighbourhoods.csv', 'reviews.csv', 'review_1.csv'), and 'New Union'. The main area shows a table named 'reviews_1' with 10 columns: 'listing_id', 'listing_reviewer_id', 'listing_date', 'reviewer_id', 'reviewer_name', 'comments', and F7 through F10. A tooltip 'Go to Worksheet' is visible over the table. A modal window titled 'reviews_1.csv' is open, showing 'Field separator: Comma', 'Text qualifier: Automatic', 'Character set: Automatic', and 'Locale: None (United States)'. At the bottom of the modal, there are buttons for 'OK' and 'Cancel'.

You will now see that the reviews table is fixed. Time to start with our first graph!

The screenshot shows the Tableau Data Source interface. At the top, there are tabs for 'Data' and 'Analytics'. Below the tabs, the 'listings' source is selected. The interface includes sections for 'Dimensions' (listing ID, listing name, date, etc.) and 'Measures' (number of records, measure values). On the right, there's a 'Sheet 1' workspace with columns and rows defined, and a 'Marks' card with options for color, size, and text.

The data blending is not set up yet. Click the listings source on the top-left, then from the Action Bar click Data > Edit Relationships.

The screenshot shows the 'Relationships' dialog box. It has a primary data source 'listings' and a secondary data source 'reviews_1'. Both are joined on the 'Id' field. There are buttons for 'Add...', 'Edit...', and 'Remove' at the bottom.

You should have on the first dropdown your listings table. Now click the Custom checkbox, delete the existing line with Ids, click the Add button, and add a new relationship like this:

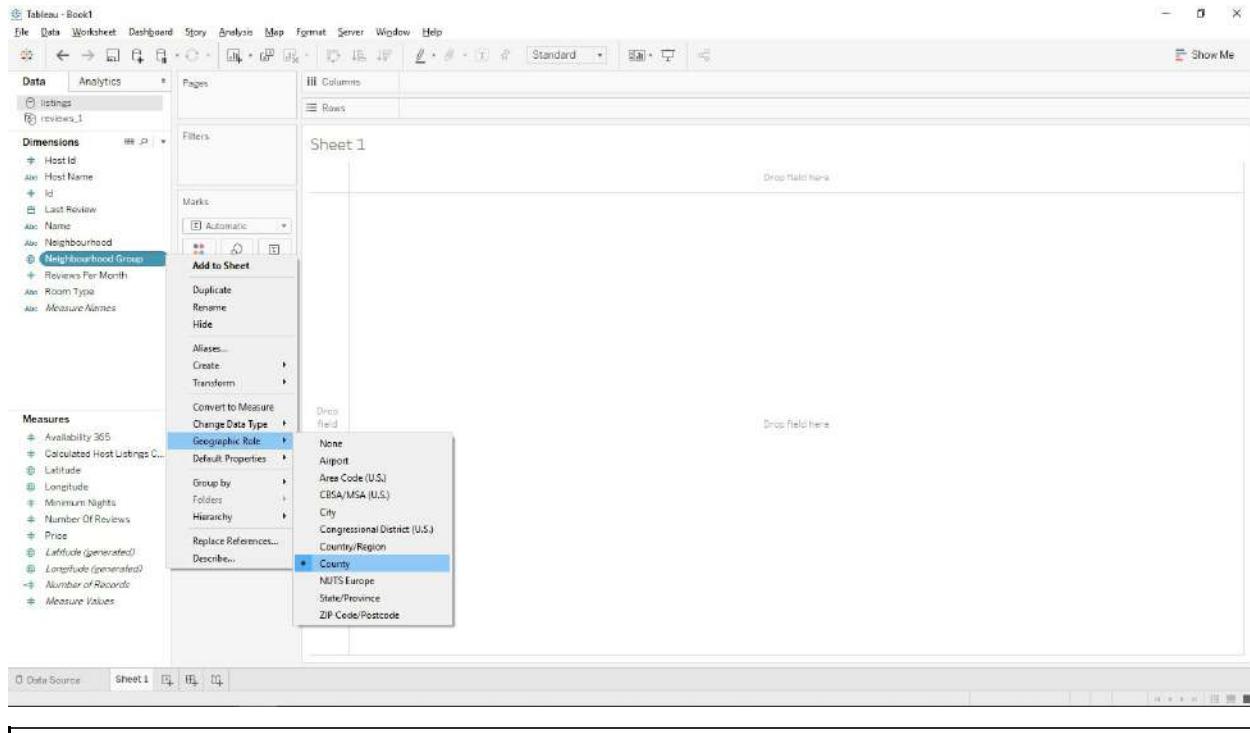
The screenshot shows the Tableau Data Source - Edit Field Mapping dialog. The Primary data source field is set to 'listing' and the Secondary data source field is set to 'review'. Both dropdown menus show a list of fields from their respective tables, including 'Id', 'Last Review', 'Name', 'Neighbourhood', 'Neighbourhood Group', 'Reviews Per Month', 'Room Type', and 'Room Type'. The 'Id' field is selected in both dropdowns. The 'OK' button is visible at the bottom right of the dialog.

The relationship between the two tables will be listings' **Id** and reviews' **Listing Id** columns. You can now click OK on all windows.

Tip

Even though you see a column with the same name on two tables, it does not necessarily mean that it's for the same thing. Be sure to not link columns of different nature, for example, Id for listings and Id for reviews, as this will give you trouble when using your data blending.

Now, change the data type of the **Neighbourhood Group** dimension on listings to be of Geographical – County. Right click it, look for Geographic Role, and select County as in below:



Explained

Geographic is one of the data types we discussed in a previous segment. What it does is translate your text or numeric values of a column into latitude and longitude coordinates, so you can plot values on a map. It can return coordinates for a city, a ZIP code, a country, and more, and it applies worldwide!

Drag Neighbourhood Group to the Rows shelf, open the Show Me pane, and select Filled Map.

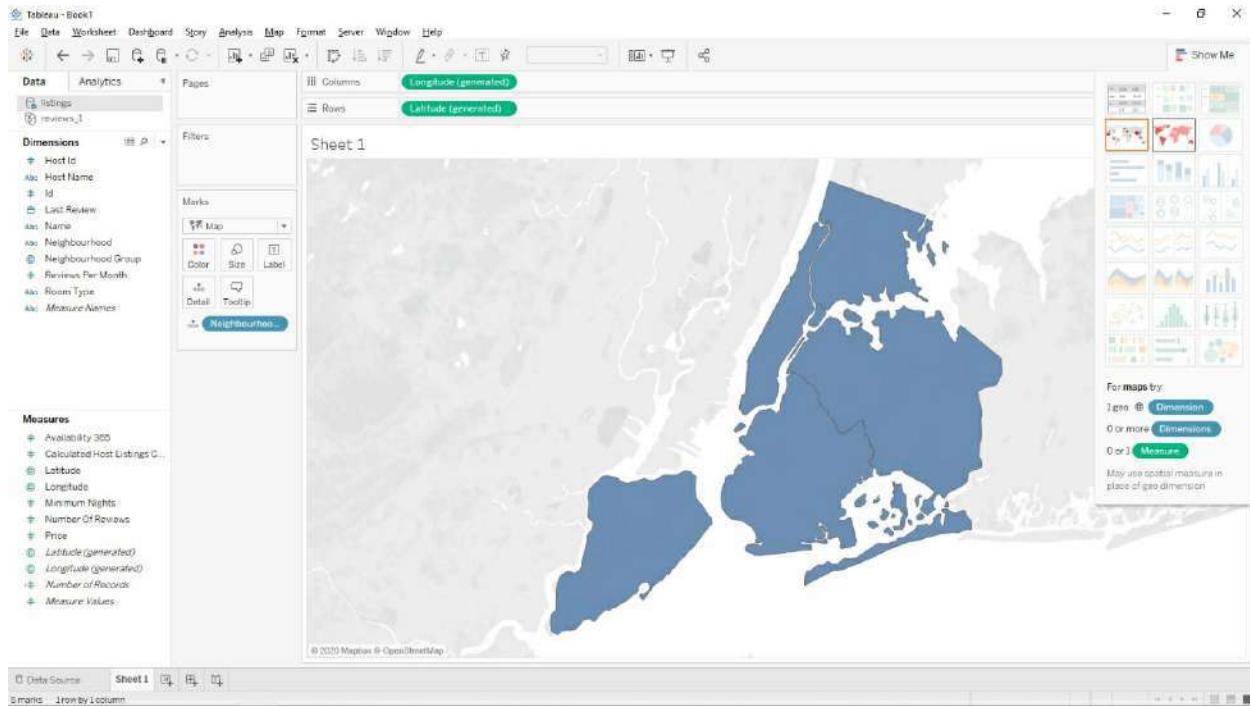
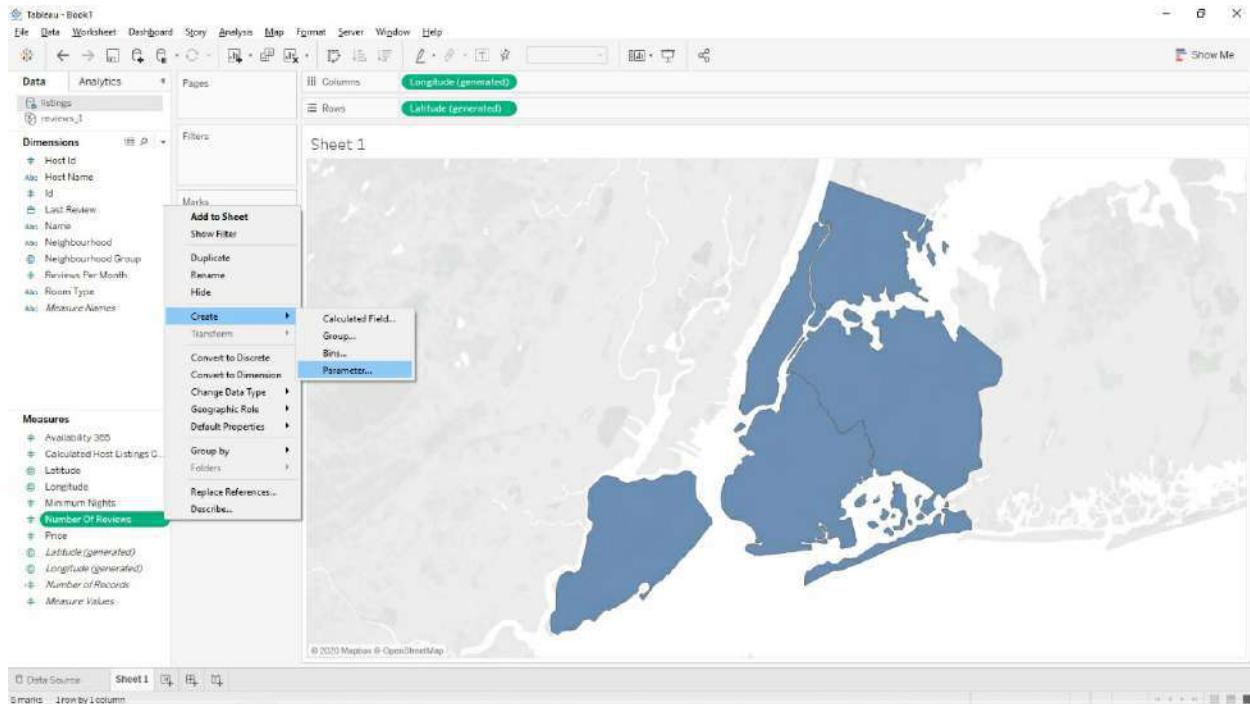


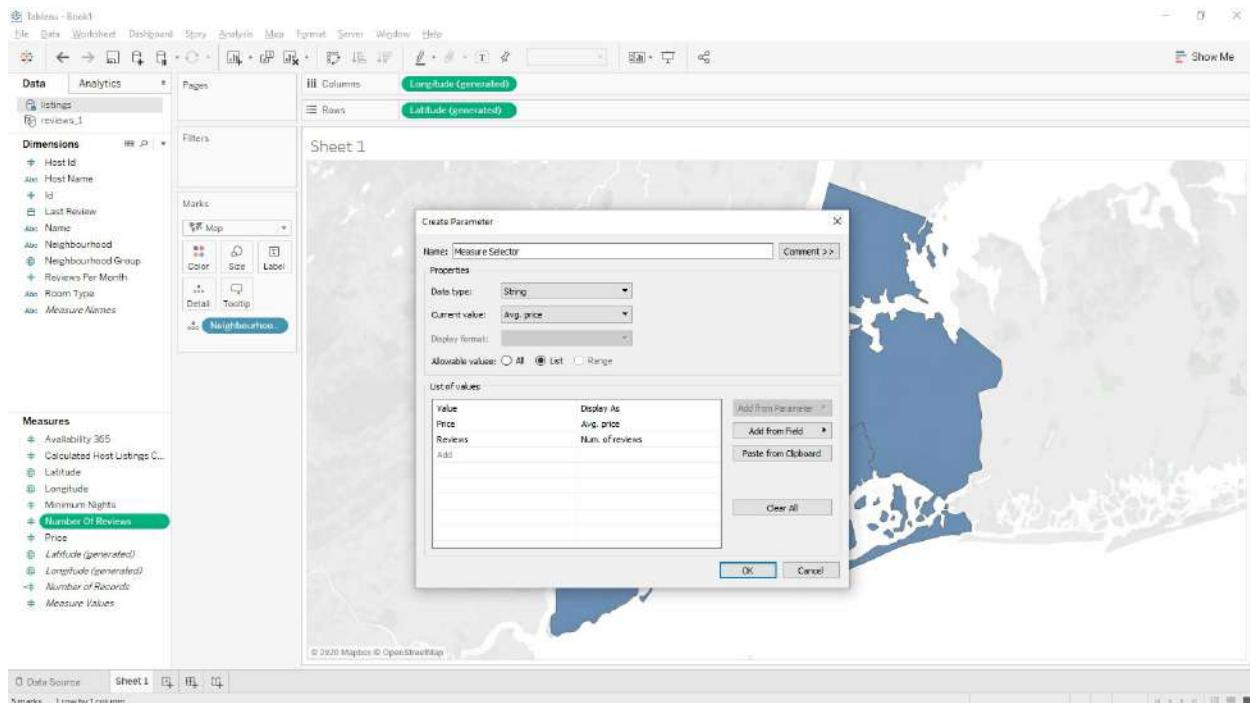
Tableau moved your dimension to the Detail property in Marks, and placed Longitude and Latitude measures (automatically generated) on the Columns and Rows shelf respectively.

We want this map by county to be colored depending on a measure. On our Measures we have **Number of reviews** and **Price** as good options, so why not instead allow the user to select between which of the two he/she wants to see? The next process will teach you how to do that:

First, create a new parameter.



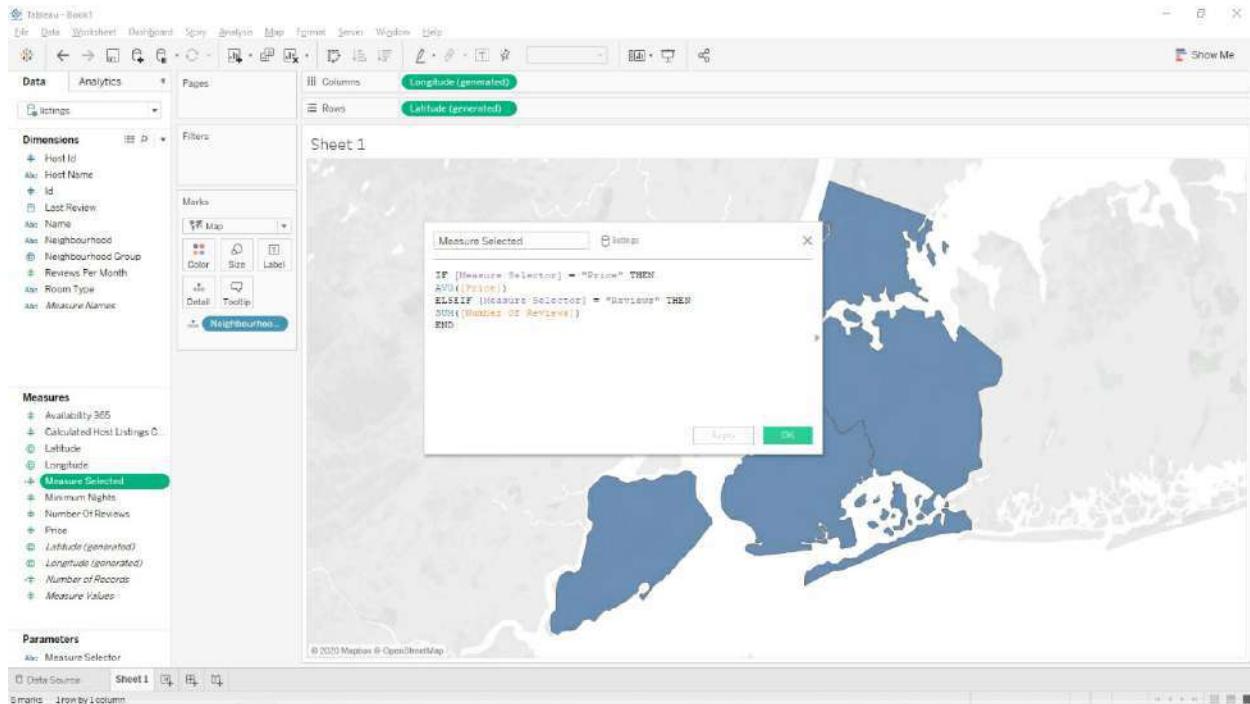
Call your new parameter **Measure Selector**, set it of Data type String, under Allowable values select List, and fill the list of values with two options: Price and Reviews. You see you can modify the text that will be displayed on the parameter dropdown so that it's different from the value itself (in case you want to make it more descriptive). Set display values “Avg. price” and “Num. of reviews”. Finally, set the Current value as Avg. price if it's not already set.



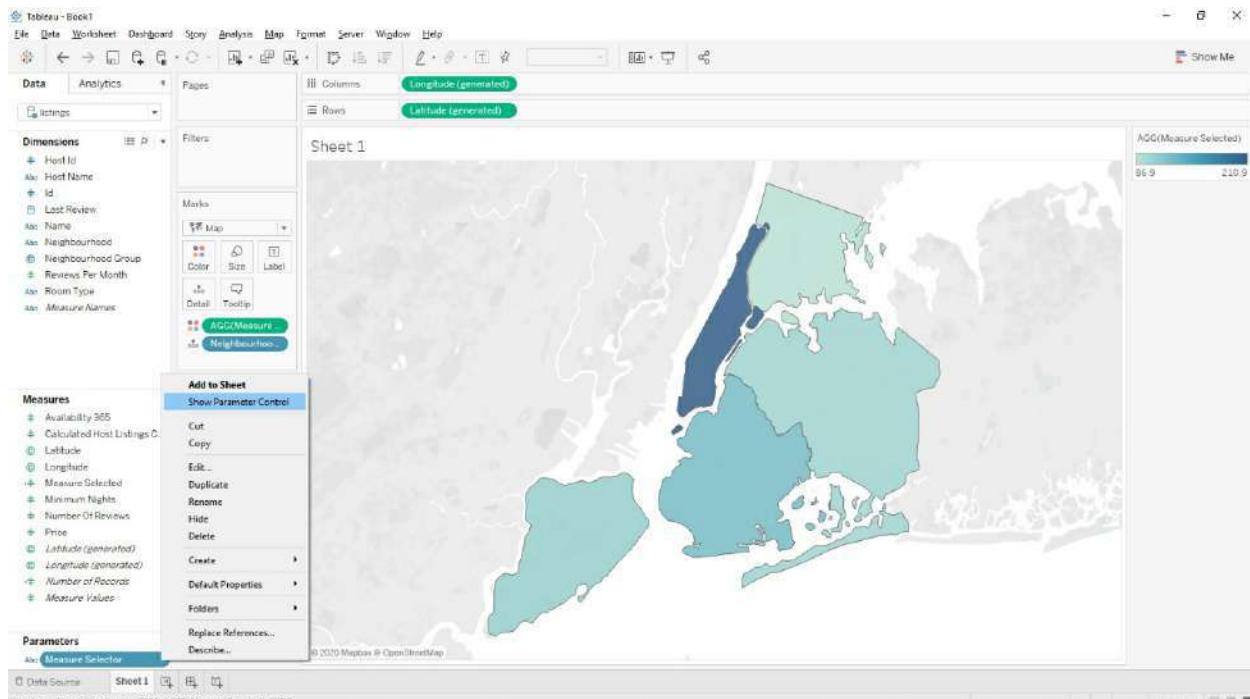
Now create a calculated field.

The screenshot shows a Tableau worksheet titled 'Sheet 1' displaying a map of New York City. The map features blue-shaded regions representing different neighborhoods. In the top-left corner, the Tableau interface is visible with the menu bar (File, Data, Worksheet, Dashboard, Story, Analysis, Map, Format, Server, Window, Help) and various toolbars. The data pane on the left lists dimensions and measures. Under 'Dimensions', items like Host Id, Host Name, Id, Last Review, Name, Neighborhood, Neighbourhood Group, Reviews Per Month, Room Type, and Measure Alarms are listed. Under 'Measures', items like Availability 365, Calculated Host Listings G., Latitude, Longitude, and Minimum Nights are listed, with 'Minimum Nights' currently selected. A context menu is open over the 'Minimum Nights' item, with the 'Create' option highlighted. The 'Create' submenu includes options like 'Calculated Field...', 'Transform', 'Group...', 'Convert to Discrete', 'Convert to Dimension', 'Change Data Type', 'Geographic Role', 'Default Properties', 'Group by', 'Folders', 'Replace References...', and 'Describe...'. At the bottom of the screen, the status bar shows 'Data Source: Sheet 1' and 'Marks: 1 row by 1 column'.

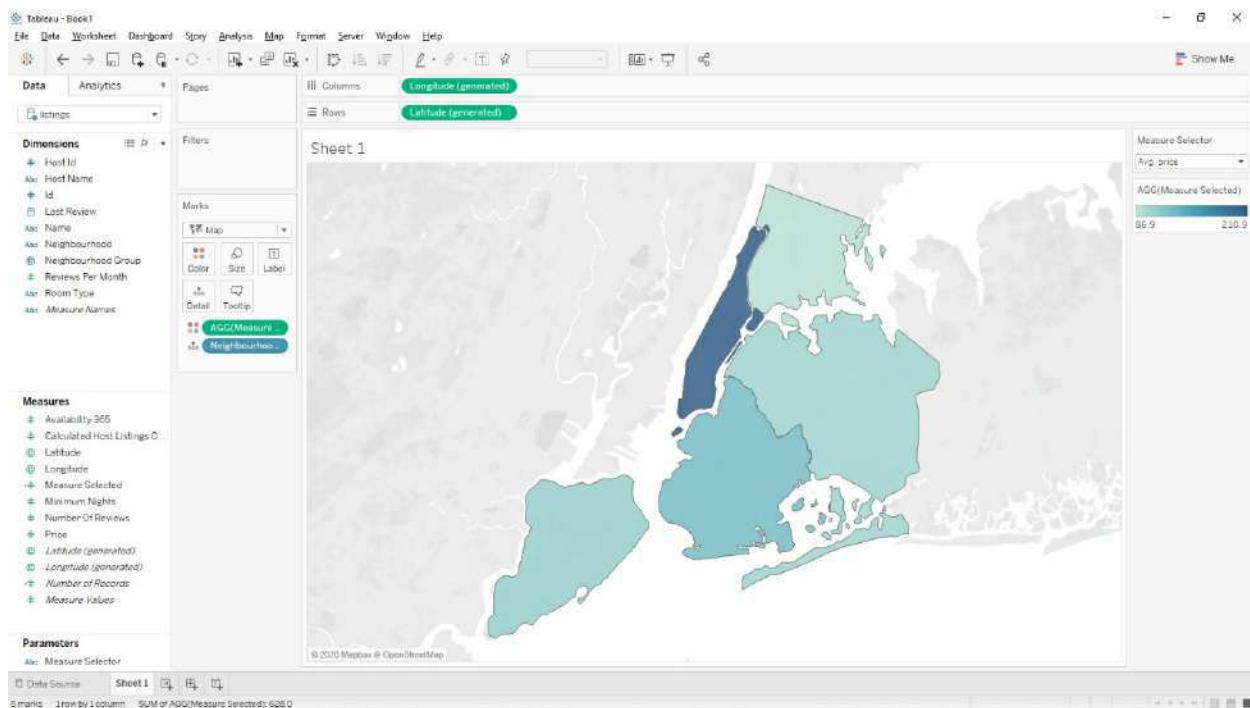
Call your calculated field “Measure Selected” and formulate it so that if Measure Selector is equal to Price then be the average price, else if the Measure Selector is equal to Reviews then show the sum of number of reviews. Notice that we’re using the values and not the display values in the conditional for the parameter, this will always be the case. Try to make the formula yourself, if you can’t then copy it from the image below:



To use your new fields, right click Measure Selector and Show Parameter Control. Now drag Measure Selected to the Colors property in Marks.

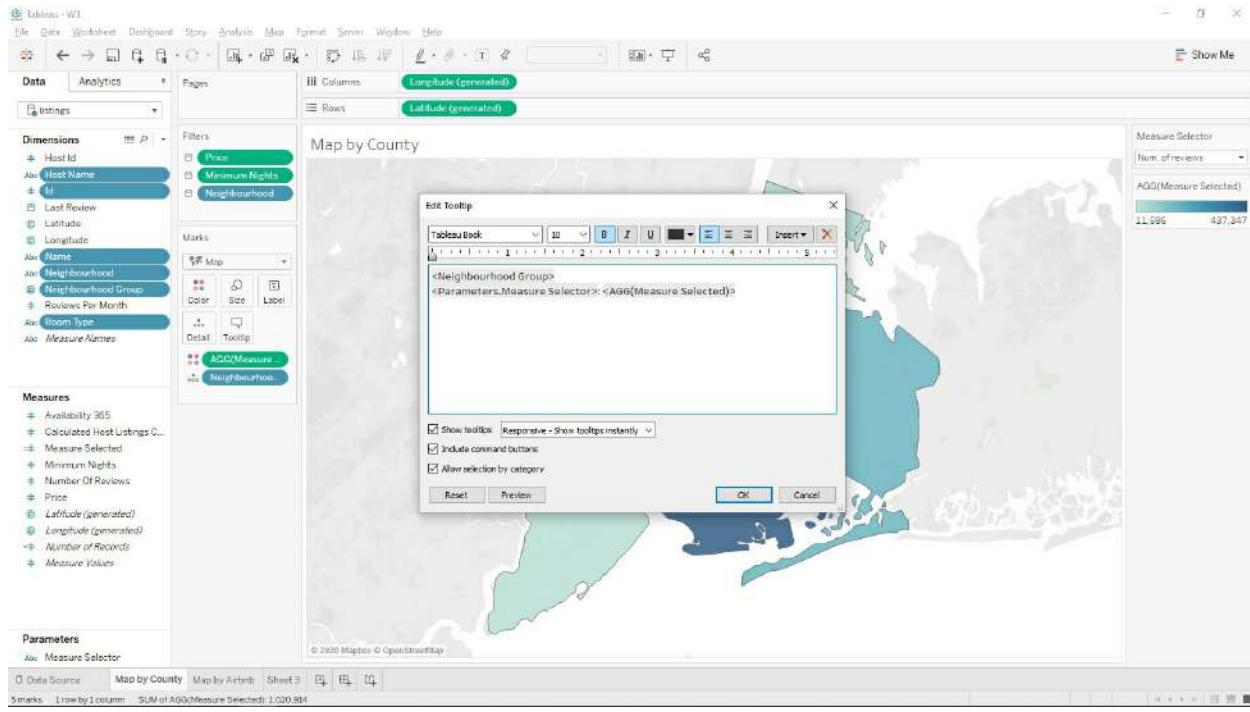


You now have a functional parameter that lets the user change the measure in use. Change the value of Measure Selector and see how Measure Selected reacts to the change and colors your map in a different way.



We can see that in price, Manhattan is the most expensive zone in average. And based on the number of reviews, Brooklyn is the most frequented county (also called “borough”) from the five.

The visual element can be improved. Click on the Tooltip property in Marks and edit the tooltip content so it looks like below:

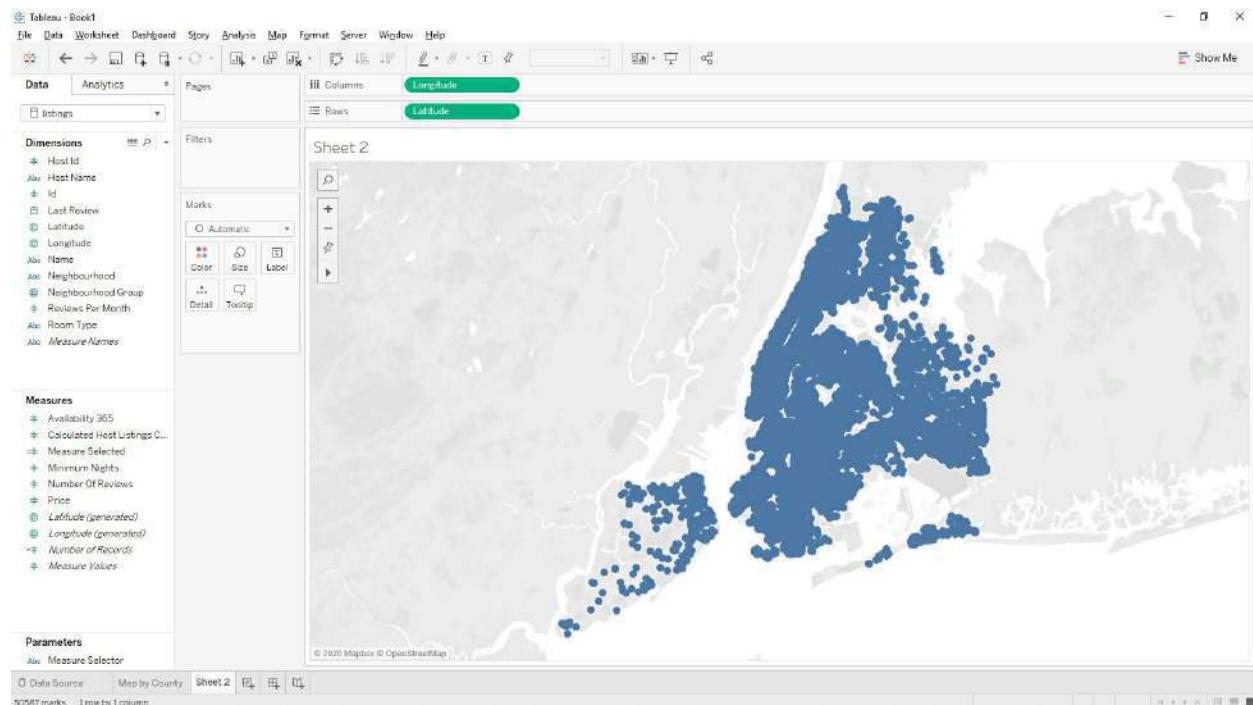


Explained

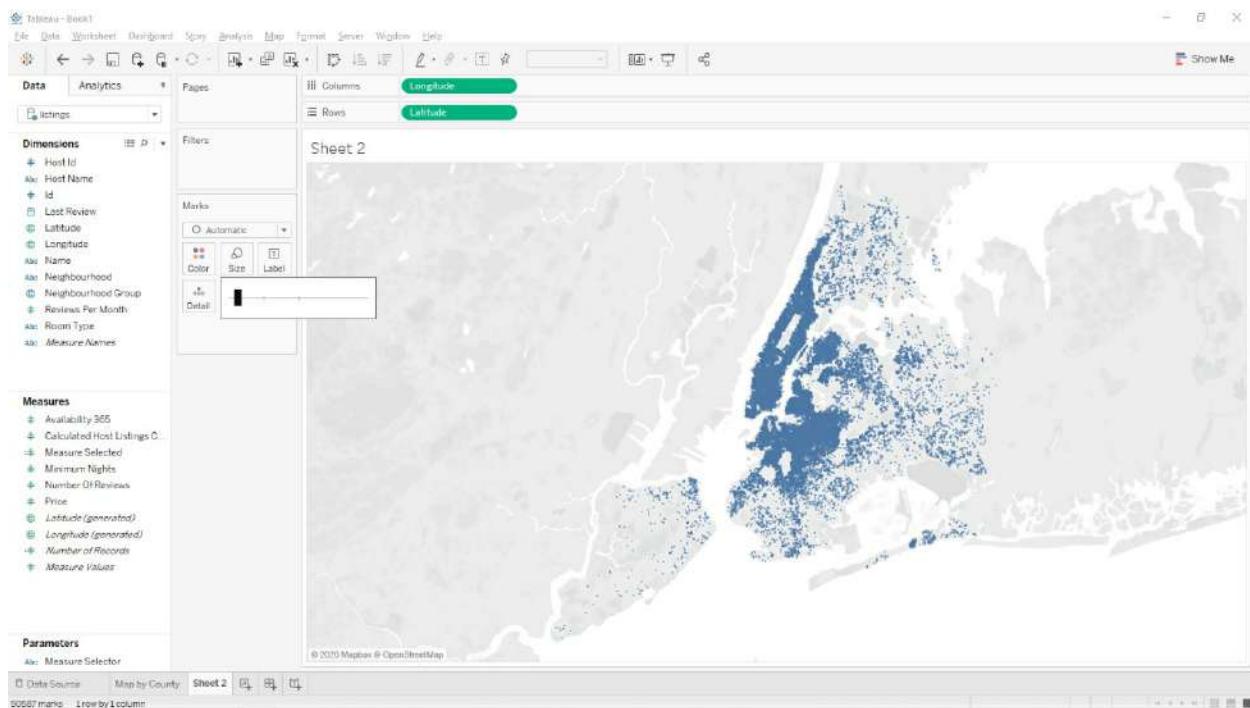
The tooltip is what appears when you hover the mouse over a number, a bar, a shape, or a populated place in the map. It serves to give the user more details on the point being analyzed, you can add any dimension, parameter, custom text, you can even add a graph inside the tooltip!

If you want your filter to have all values as in this example, you must select **Use all**, don't **Select**

Now let's do a map by Airbnb. On a new sheet, convert the Latitude and Longitude columns to Dimensions, but keep them as Continuous. Now drag longitude to columns and latitude to rows.

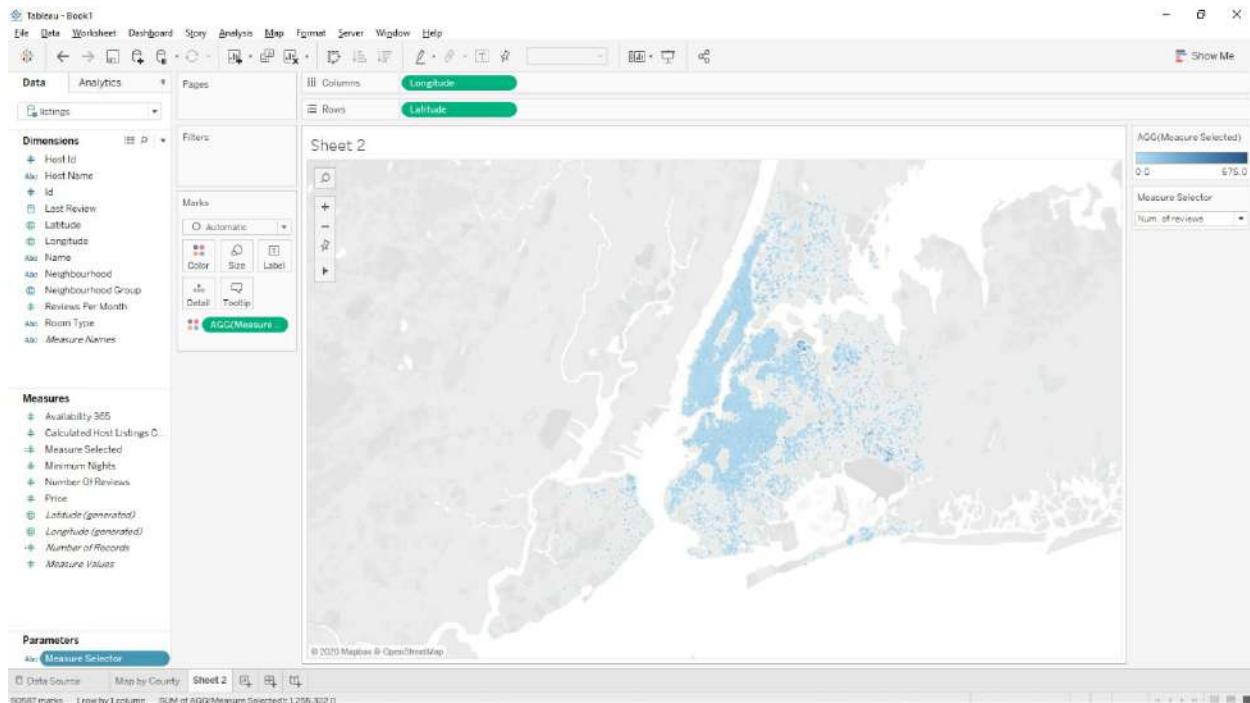


Each point on the graph represents an Airbnb listing, since Longitude and Latitude are present for each row of the listings table, meaning for each Airbnb. There's too many, though, and the points are hard to distinguish. Click the Size property in Marks and lower the size of the points.



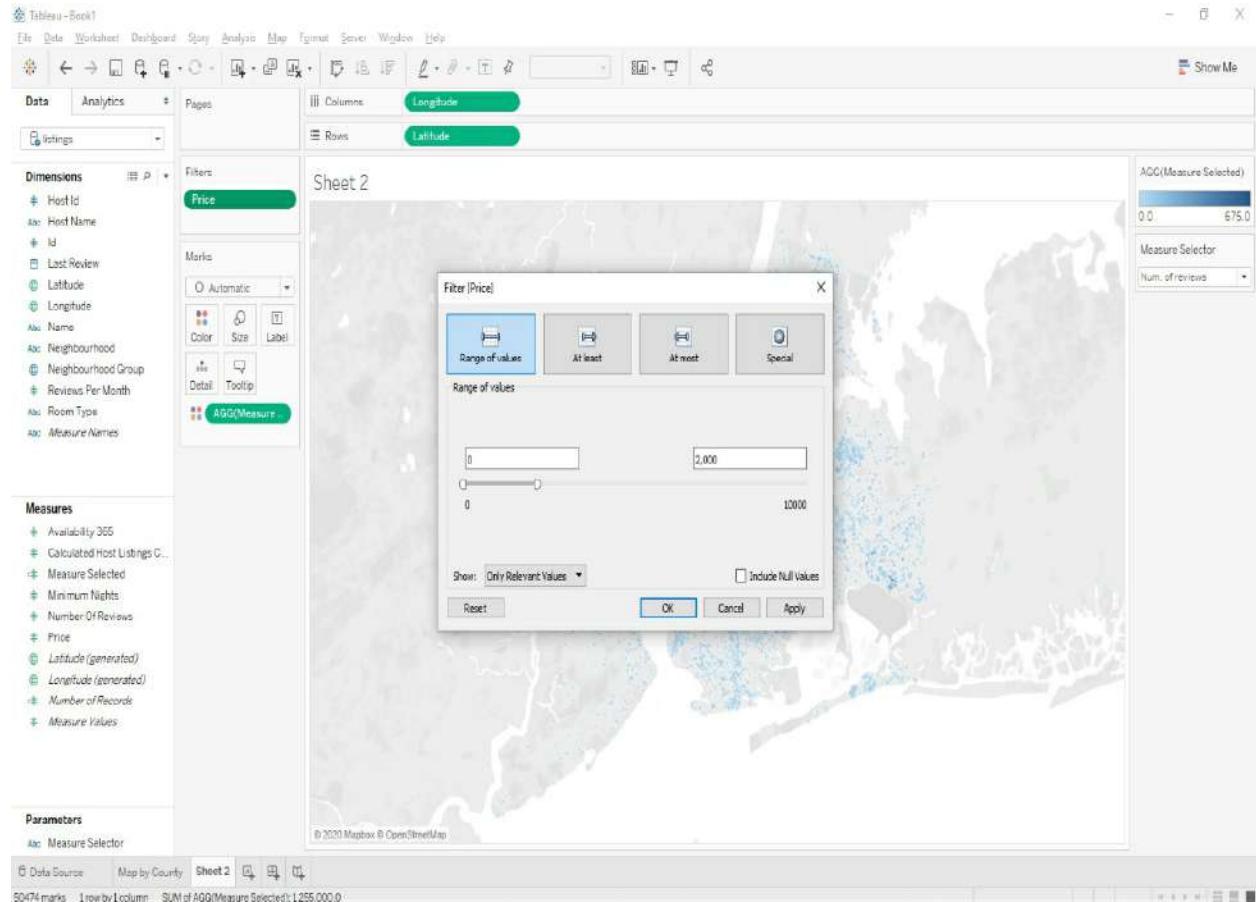
Now drag **Measure Selected** to the Color property. It will look with a much brighter color, that's because there are some outliers in either number of reviews or price (meaning that a few Airbnb's have much more reviews or a much higher price than the city's average).

Don't forget to also show **Measure Selector** on this sheet, so you can also play here with it.



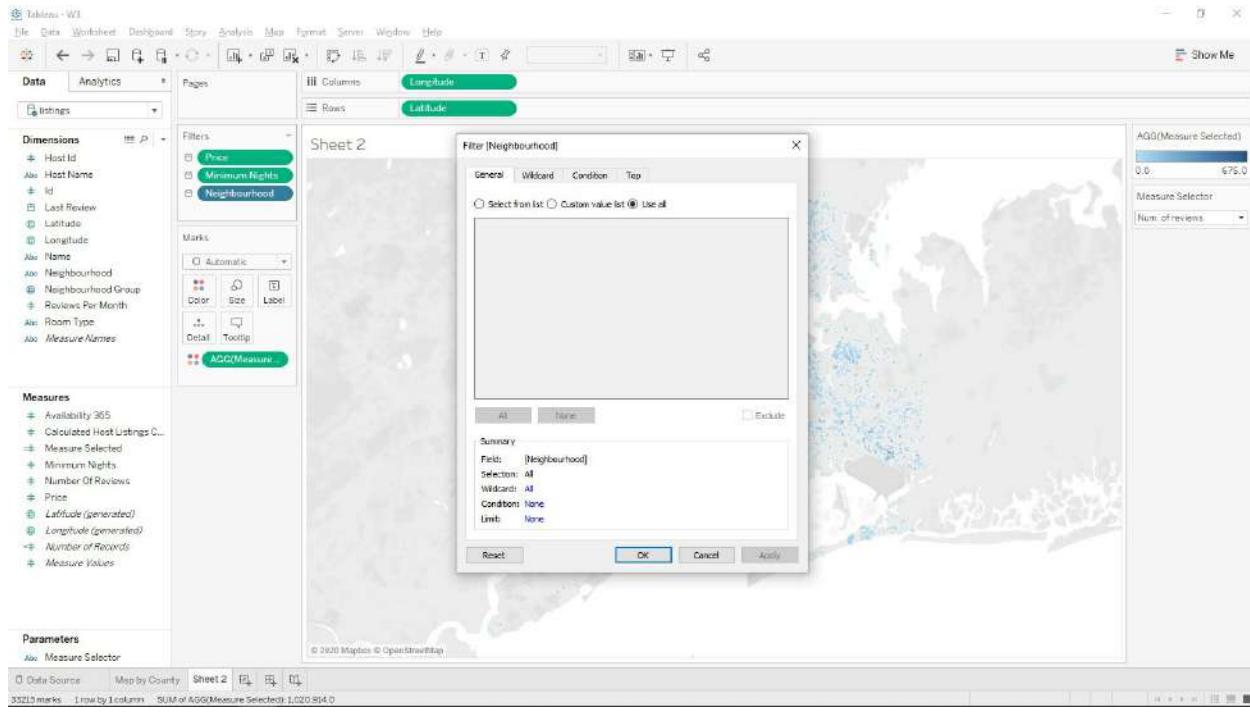
It is still hard to visualize all the dots in the map. That's why we need filters so the user can show only the relevant ones by his/her criteria.

First, drag Price to Filters. Set the filter to be in a range of values, from 0 to 2,000.



Add **Minimum Nights** as filter in the same way, you can set any maximum value for it.

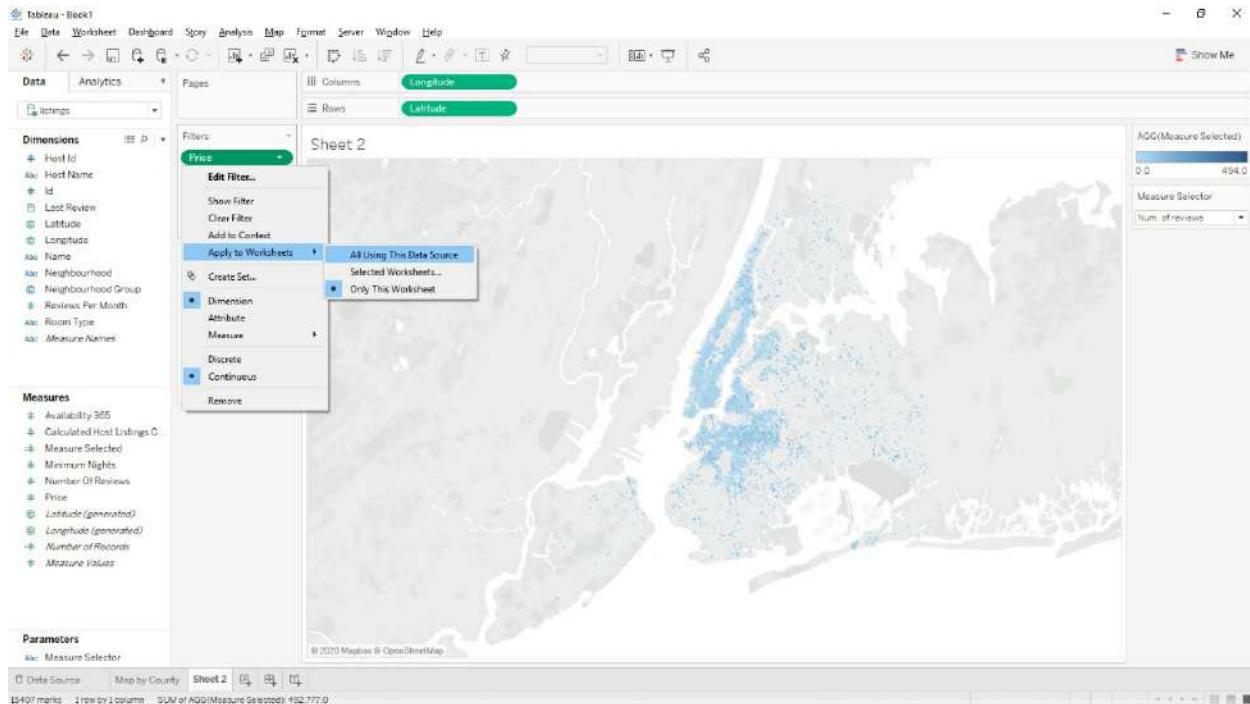
Now add **Neighbourhood** as filter, let it have all values selected, by clicking the Use all on the top:



Tip

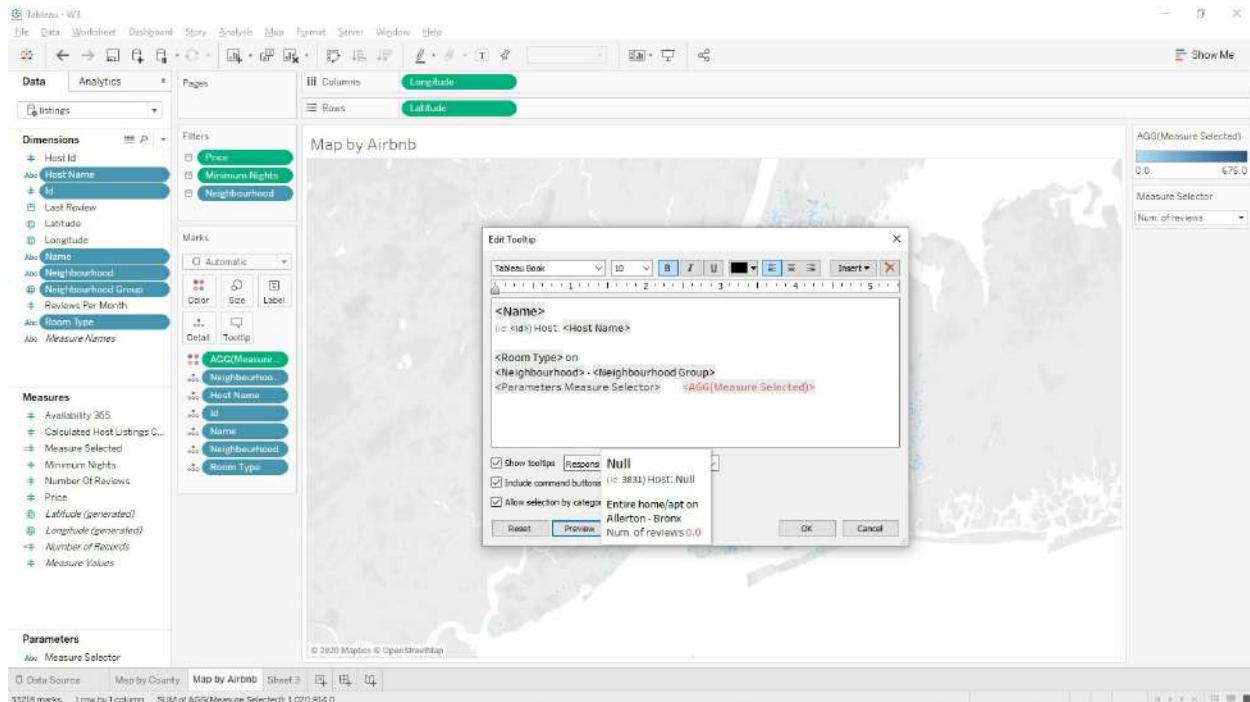
If you want your filter to have all values as in this example, you must select **Use all**, don't **Select from list** and select all items from the list. If you do that, your filter will exclude any new values that come to that column in the future, and you will never know until you enter the filter.

Make all these four filters global as we did on Workout 2. Superstore basic sales dashboard.



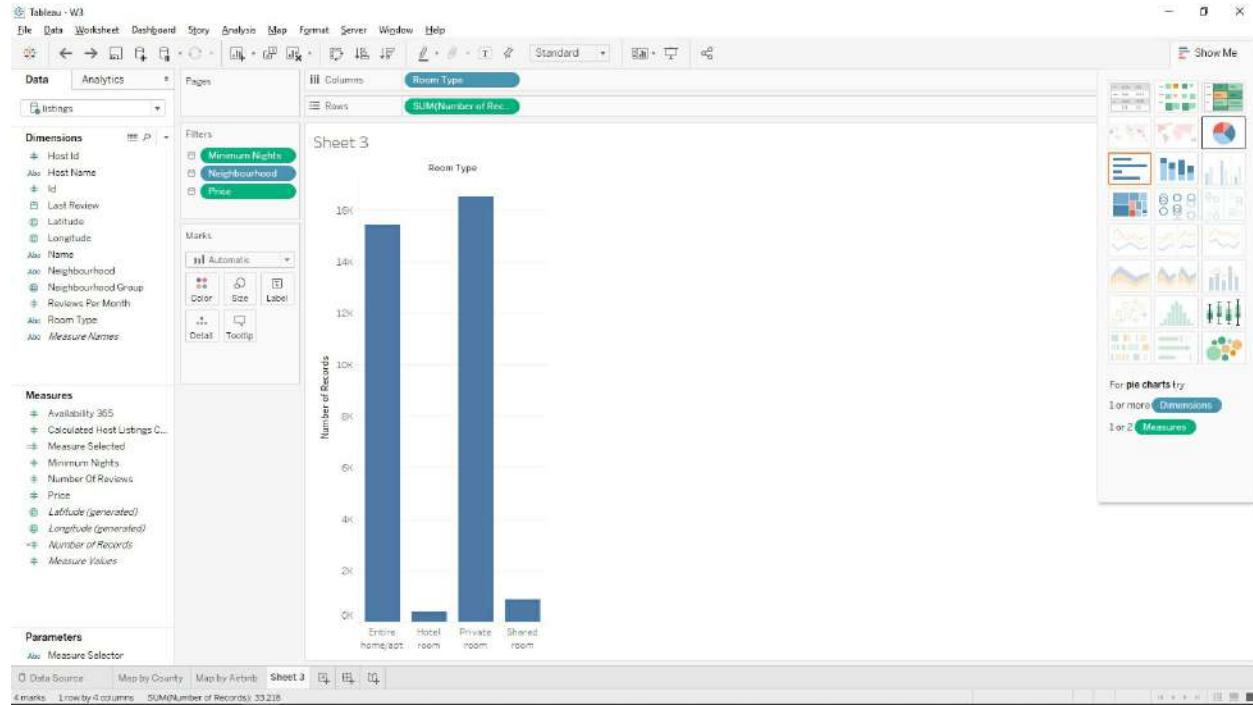
Rename the sheet to Map by Airbnb.

Edit the tooltip on this sheet as well, but this time let's add more detail. Add Neighbourhood Group, Host Name, Id, Name, Neighbourhood, and Room Type to the Detail property in Marks, and then click the Tooltip property to edit it, making it look like this:

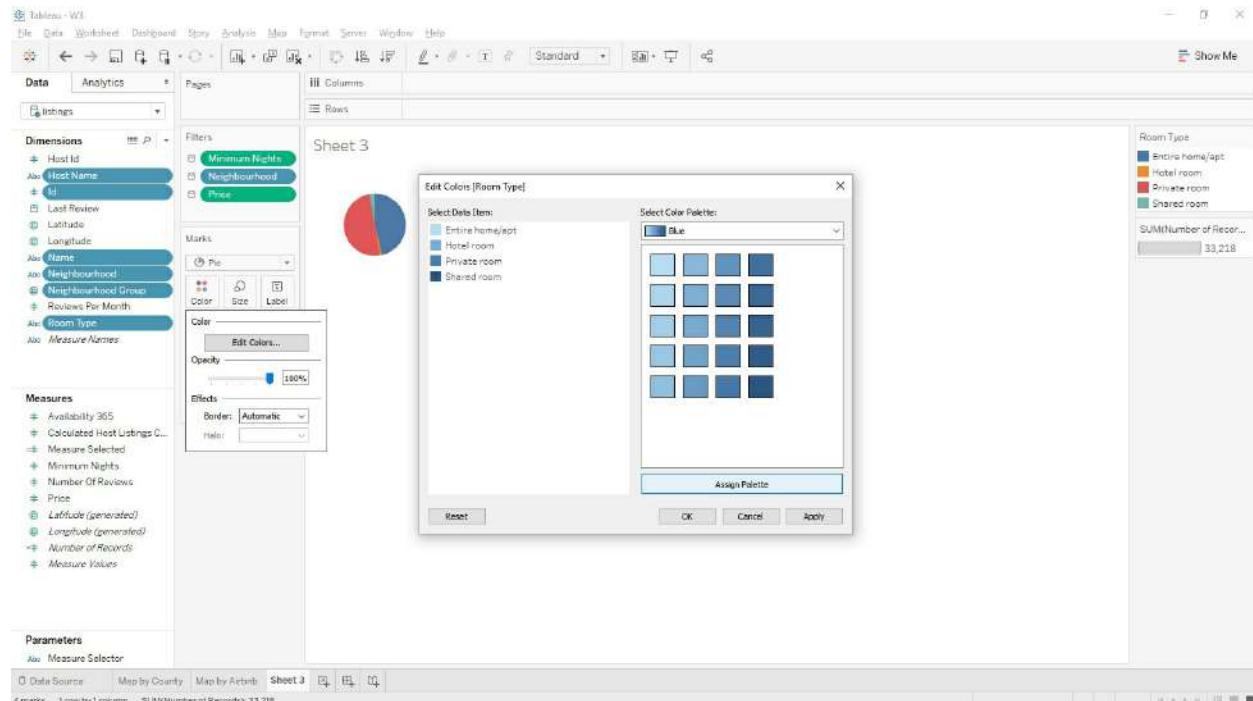


You can click the Preview button to see how your tooltip will look.

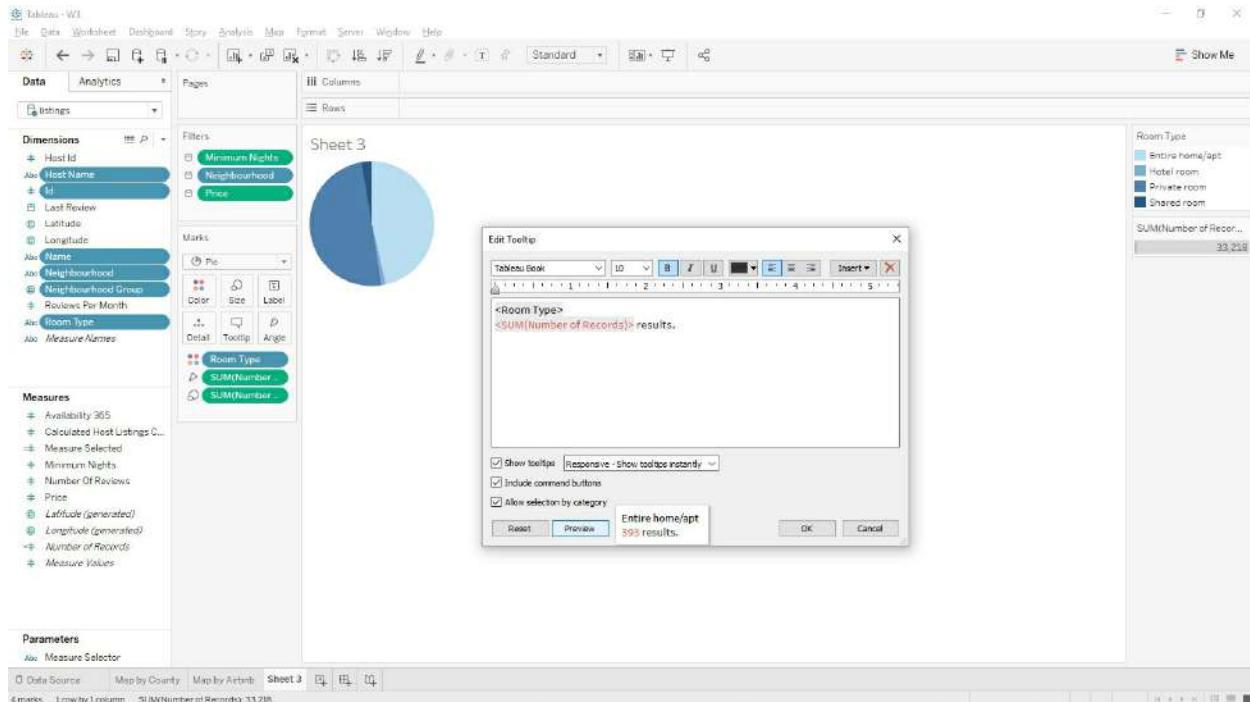
Make a new sheet, drag **Room Type** to columns and **Number of Records** to rows. Then, on the Show Me, click the pie chart.



Now edit the colors of the pie to use the blue color palette.



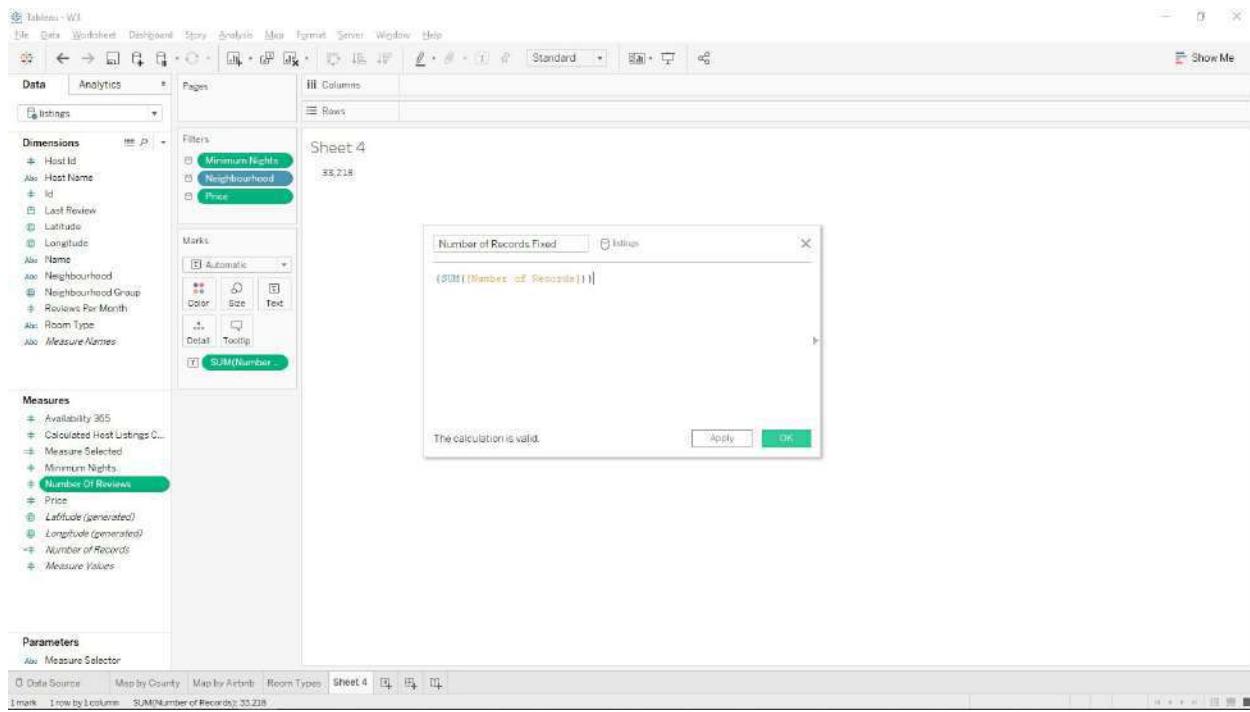
Increase the pie chart size to be half, and edit the tooltip to look like below:



Rename the sheet to Room Types.

Create another sheet, add **Number of Records** to the table content (inside the “Abc” cell), and create a new calculated field called Number of Records Fixed. Add the next formula:

{SUM([Number of Records])}

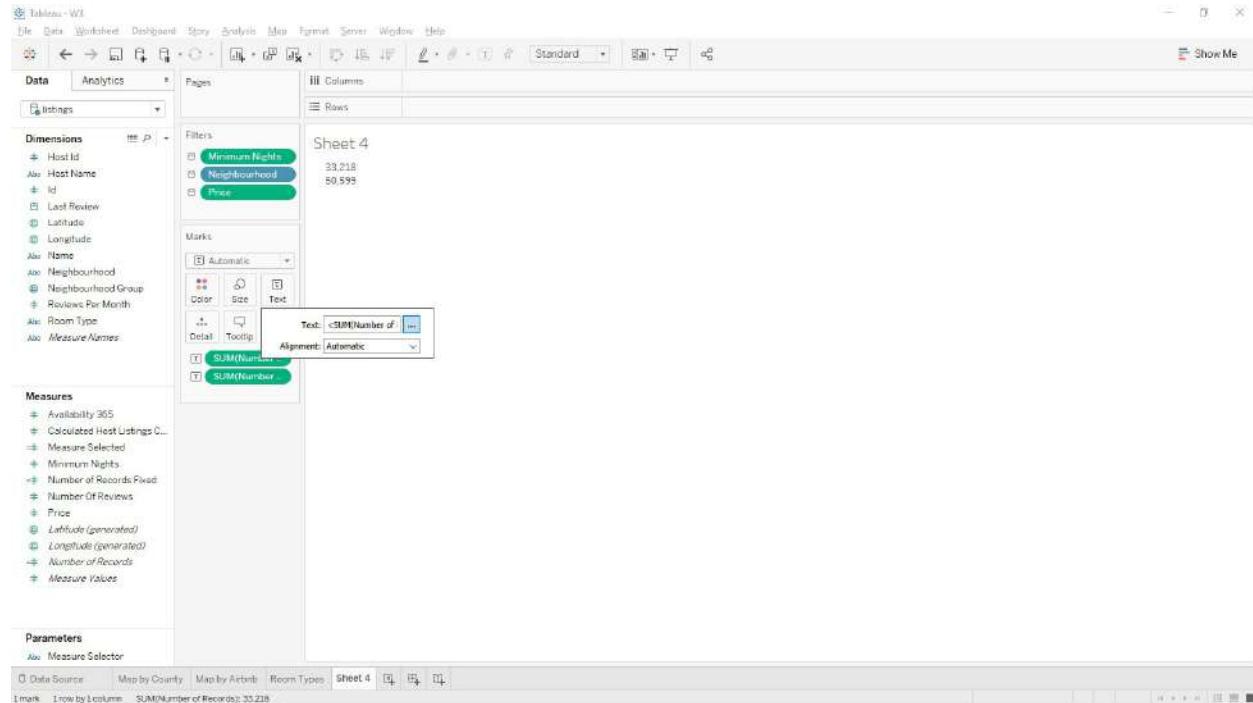


Explained

You might have noticed the brackets “{}” outside the formula for Number of Rows Fixed. What these do is make the calculation ignore all dimensions that are used in the sheet, and that includes the filters. This means that it will always display the total number of rows, regardless if you’re only filtering by Airbnb’s in a certain neighbourhood, or if you’re using a dimension on the rows/columns shelves.

This is part of Level of Detail (LOD) calculations, one of the most advanced concepts of Tableau. Putting brackets outside your calculation makes it a **FIXED** calculation. We will later see the other two types of LOD’s: **INCLUDE** and **EXCLUDE**.

Drag your new field to the Text property in Marks, then click the Text property and the [...] button to edit it:



Make the text look like below. As a reference, use font sizes 18 and 9 for **SUM (Number of Records)** and **SUM(Number of Records Fixed)**, respectively. Don’t miss the “Airbnb’s out of “ label on the gray line.

The screenshot shows the Tableau desktop interface with the following details:

- Sheet 4:** Displays the value "33,218" with the subtitle "Airbnb's out of 50,599".
- Labels:** A tooltip labeled "<SUM(Number of Records)>" is displayed.
- Edit Label Dialog:** An open dialog titled "Edit Label" shows the formula "<SUM(Number of Records)>".
- Marks Card:** Shows the aggregation type as "Sum".
- Dimensions:** Host Id, Host Name, Id, Last Review, Latitude, Longitude, Name, Neighbourhood, Neighbourhood Group, Reviews Per Month, Room Type, Measure Names.
- Measures:** Availability 365, Calculated Host Listings C..., Measure Selected, Minimum Nights, Number of Records Fixed, Number of Reviews, Price, Latitude (generated), Longitude (generated), Number of Records, Measure Values.
- Parameters:** An "Add Measure Selector" button is present.
- Bottom Navigation:** Data Source, Map by County, Map by Attrb, Room Types, Sheet 4, etc.

Rename the sheet to NumAirbnbs and create a new one. We will make other two sheets with the same logic, only that it will be to show the average price and the number of reviews. Let's start with the price sheet.

Since we will not display **Prices** aggregated as sum but in average, edit the default aggregation like this:

The screenshot shows the Tableau desktop interface with the following details:

- Sheet 5:** A blank sheet with a placeholder "Drop field here".
- Context Menu:** The "Price" measure in the Measures list has a context menu open, specifically the "Default Properties" section under "Aggregation".
- Aggregation Options:** The "Average" option is selected from the list.
- Dimensions:** Host Id, Host Name, Id, Last Review, Latitude, Longitude, Name, Neighbourhood, Neighbourhood Group, Reviews Per Month, Room Type, Measure Names.
- Measures:** Availability 365, Calculated Host Listings C..., Measure Selected, Minimum Nights, Number of Records Fixed, Number of Reviews, Price, Latitude (generated), Longitude (generated), Number of Records, Measure Values.
- Bottom Navigation:** Data Source, Map by County, Map by Attrb, Room Types, NumAirbnbs, etc.

Now when you drag with the left click the measure and put it on a table, instead of having $\text{SUM}(\text{Price})$ you will by default have $\text{AVG}(\text{Price})$. If you were showing $\text{SUM}(\text{Price})$ in previous sheets it might change those to AVG as well, so I recommend you do this from the beginning, or review your previous sheets.

Create a new variable called Price Fixed which is equal to “ $\{\text{AVG}([\text{Price}])\}$ ”. Drag to the Text property, and again edit text like below:

The screenshot shows the Tableau Data Editor interface. On the left, the Data pane lists dimensions like Host Id, Host Name, Id, Last Review, Latitude, Longitude, Name, Neighbourhood, Neighbourhood Group, Reviews Per Month, Room Type, and Measure Alarms. The Measures pane lists Availability 365, Calculated Host Listings C..., Measure Selected, Minimum Nights, Number of Records Fixed, Number of Reviews, Price, Price Fixed, Latitude (generated), Longitude (generated), Number of Records, and Measure Values. The Parameters pane shows a Measure Selector. In the center, the 'Edit Label' dialog is open, showing the formula <AVG(Price)> and a preview of 141.5. The preview also includes the note 'avg price vs global <SUM(Price Fixed)>'. At the bottom of the dialog are 'Reset', 'Preview', 'OK', 'Cancel', and 'Apply' buttons. The top of the screen shows the Tableau ribbon with File, Data, Worksheet, Dashboard, Story, Analysis, Me, Format, Server, Window, Help, and a Show Me button.

Since it's a price, we should give it a better format, like include the dollar sign. You can do this in two ways: you can change the format for the Price and Price Fixed measures on the sheet only, or you can change their default format property, so it applies for present and future sheets. Let's do the second option, right click Price, go to Default Properties, and Number Format.

The screenshot shows the Tableau Data Editor interface. A context menu is open over the 'Price' measure in the Dimensions pane. The menu path 'Default Properties > Number Format...' is highlighted. Other options in the submenu include 'Comment...', 'Color...', 'Aggregation...', and 'Total using...'. The main menu bar at the top includes 'File', 'Data', 'Worksheet', 'Dashboard', 'Story', 'Analysis', 'Map', 'Format', 'Server', 'Window', and 'Help'. The 'Format' tab is selected. The bottom status bar shows '1 mark 1 row by 1 column SUM of AVG[Price]: 141.5'.

Select Number (Custom), set 1 decimal place, set negative values to show as “()”, keep display units in none, add “\$” as prefix, and lastly, toggle the Include thousands separators checkbox.

The screenshot shows the 'Default Number Format [Price]' dialog box. In the 'Number (Custom)' section, the 'Decimal places:' dropdown is set to 1. The 'Negative values:' dropdown shows '\$(1129)'. Under 'Display Units:', 'None' is selected. In the 'Prefix / Suffix:' section, there is a '\$' symbol in the first input field. The 'Include thousands separators' checkbox is checked. The 'OK' button is visible at the bottom right of the dialog.

Now we have a better-looking number.

Tip

Change the default aggregation and the number format properties, instead of doing it inside the sheet. This will save you a lot of time when a user asks you to change the format of a measure you have in many sheets.

When you change the format for a measure by right-clicking it from a shelf and going to Format, you are overriding the default setting only on that sheet. This can be useful at times, but if you overuse it, you will consume more time when asked for changes.

Another tip: if you have many sheets, and you do need to show a measure like price in different formats, try duplicating the measure and applying the second format on it as default.

The screenshot shows the Tableau Data Editor interface with the following details:

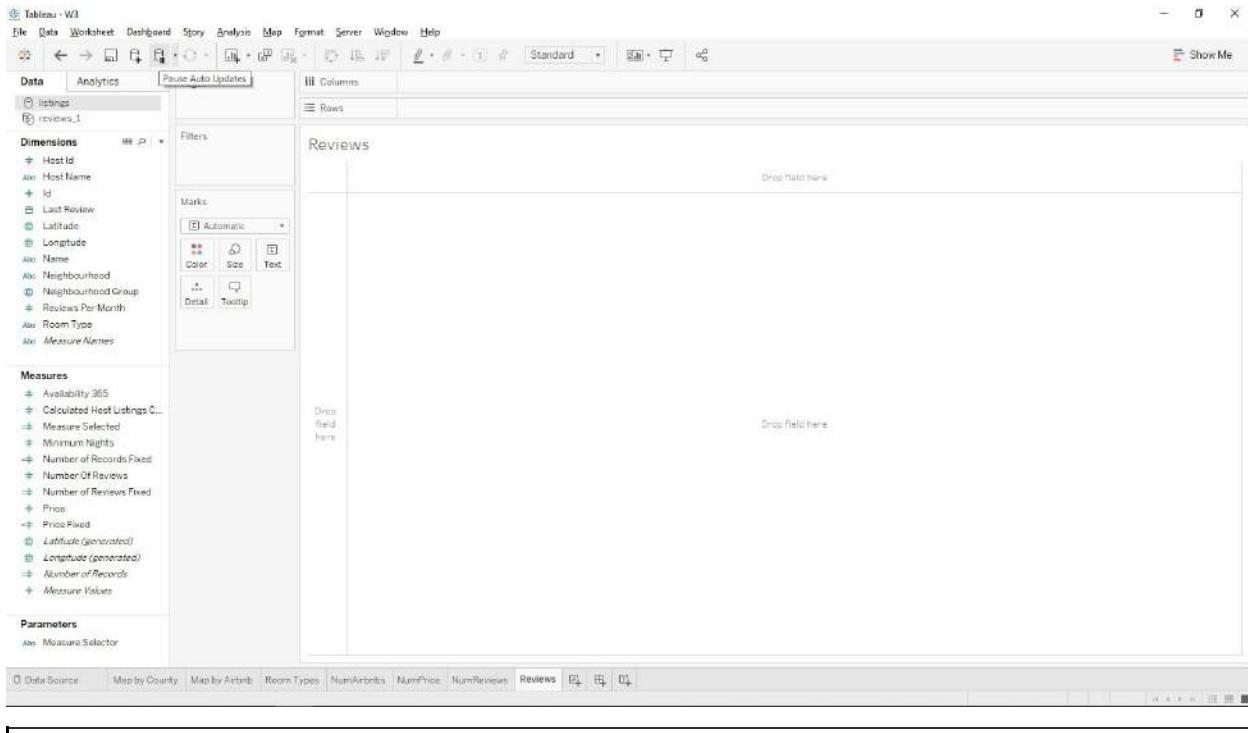
- Top Bar:** File, Data, Worksheet, Dashboard, Story, Analysis, Map, Format, Server, Window, Help.
- Left Sidebar (Dimensions):** Host Id, Host Name, Id, Last Review, Latitude, Longitude, Name, Neighborhood, Neighborhood Group, Reviews Per Month, Room Type, Measure Names.
- Left Sidebar (Measures):** Availability 365, Calculated Host Listings 0, Measure Selected, Minimum Nights, Number of Records Fixed, Number of Reviews, Price, Price Fixed, Latitude (generated), Longitude (generated), Number of Records, Measure Values.
- Left Sidebar (Parameters):** Measure Selector.
- Filters:** Minimum Nights, Neighborhood, Price (selected).
- Marks:** Automatic, Color, Size, Text, Detail, Tooltip, AVG(Price) (selected), SUM(Price Fix).
- Sheet Content:** A single text value '\$141.5' displayed prominently.
- Bottom Status Bar:** Data Source, Map by County, Map by Arbitr., Room Types, NumArbitrs, NumPrice, 1 mark, 1 row by 1 column, SUM of AVG(Price): \$141.5.

Rename the sheet to NumPrice. Now repeat the process for **Number of Reviews**. Create Fixed measure, drag to sheet, edit text, and check the formatting.

The screenshot shows the Tableau Data Editor interface. On the left, the Dimensions and Measures panes are visible, listing various fields like Host Id, Host Name, and Number of Reviews. In the center, a 'NumReviews' sheet is open. A floating window titled 'Number of Reviews Fixed' contains the formula <SUM([Number Of Reviews])>. Below the formula, a tooltip says 'The calculation is valid.' At the bottom right of the formula window are 'Apply' and 'OK' buttons.

This screenshot shows the same Tableau interface after the calculated field has been applied. The 'Price' filter is selected in the filters pane. An 'Edit Label' dialog box is open, displaying the formula <SUM(Number Of Reviews)> and the result '1,020,914'. A tooltip below the result states 'reviews out of 1,255,322'. The bottom status bar shows the mark count as '1 mark - Row by 1 column - SUM(Number Of Reviews): 1,020,914'.

Create a new sheet called Reviews. Now we will play with our second data source, for which, before doing anything, you should click the Pause Auto Updates button on the top left, to the side of the save button.



Tip

When making big tables or playing with big data, it's recommended to Pause Auto Updates. If you don't, you will work very slowly because Tableau will automatically update your table by each change you do. Each update will take many seconds, and it's not efficient.

Move to the reviews.csv data source and drag **Date** as attribute and **Comments** to Rows (you do the first one by dragging with right click and selecting **Date (discrete)**). Now go back to the listings data source from the top left, check the chain symbol on the Id dimension, drag the dimension to the Filters shelf, and select 2595 (so you only show comments for one random Airbnb and not the full list).

Resume Auto Updates, after some seconds you will see a table like below:

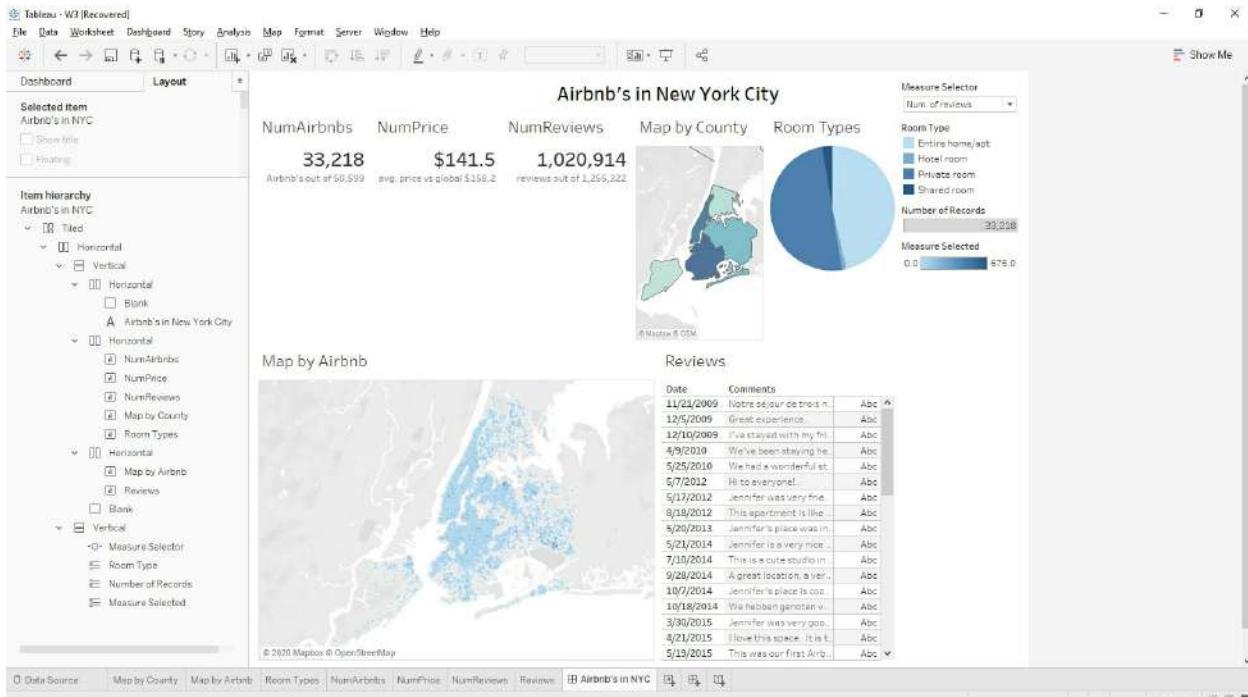
The screenshot shows a Tableau dashboard titled "Tableau - W3 [Recovered]". The interface includes a top navigation bar with File, Data, Worksheet, Dashboard, Story, Analysis, Map, Format, Server, Window, Help, and a Show Me button. Below the navigation is a toolbar with various icons. On the left, there are three main panels: "Dimensions" (listing "Comments", "Date", "Id", "Listing Id", "Reviewer Id", "Reviewer Name", and "Measure Alarms"), "Measures" (listing "Number of Records" and "Measure Values"), and "Parameters" (listing "Measure Selector"). The main workspace is titled "Sheet 7" and contains a data grid. The grid has two columns: "Date" and "Comments". The data consists of 48 rows, each containing a date and a comment in French or English. The comments describe various stays and experiences at different Airbnb listings.

Date	Comments
11/21/2009	Notre séjour de trois nuit...
12/5/2009	Great experience.
12/10/2009	I've stayed with my friend...
4/9/2010	We've been staying here f...
5/25/2010	We had a wonderful stay.
5/7/2012	Hi everyone!
5/17/2012	Jennifer was very friendly.
6/19/2012	This apartment is like a re...
5/20/2013	Jennifer's place was in a g...
5/21/2014	Jennifer is a very nice ho...
7/10/2014	This is a cute studio in a v...
9/28/2014	A great location, a very co...
10/7/2014	Jennifer's place is comfort...
10/18/2014	We hebben genoten van ons...
3/30/2015	Jennifer was very good at ...
4/21/2015	I love this space. It is tru...
5/19/2015	This was our first Airbnb.
9/21/2015	Great location, in the mid...
9/28/2015	Jennifer was without dou...
4/11/2016	Jennifer was an unbelieva...
7/22/2017	The Midtown Castle was a...
9/16/2017	Tiene una ubicación excel...
10/16/2017	Jennifer is a gracious host.
10/29/2017	Jennifer was super friendly.
11/4/2017	Jennifer's space was such...
12/3/2017	This is a nice apartment b...
12/19/2017	Beautiful and stylish spot.
12/31/2017	Location location! Location...
1/1/2018	Great place. Would defin...
4/29/2018	We really enjoyed our sta...
5/18/2018	This apartment is convenie...

Now to the final step: making the dashboard.

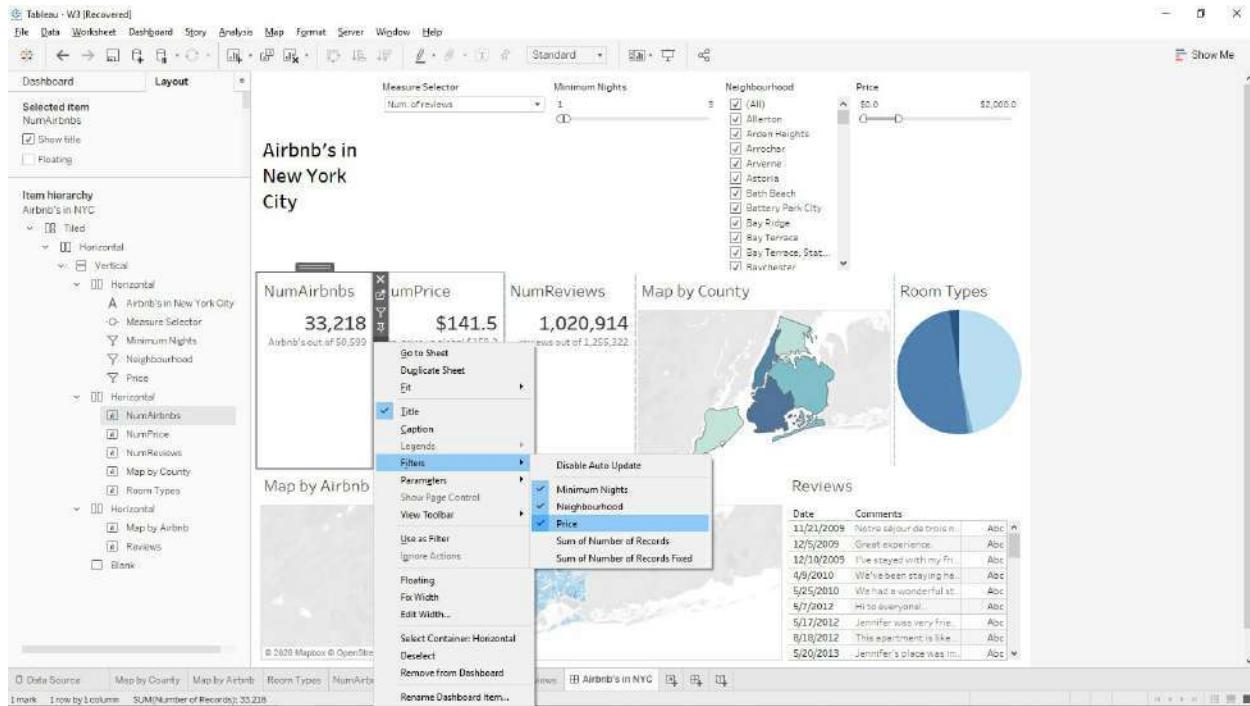
1. Add Vertical (1).
2. Inside Vertical (1), add Blank.
3. Inside Vertical (1) and above Blank, add Horizontal (1).
4. Inside Horizontal (1), add a new Blank and a Text that says, “Airbnb’s in New York City”.
5. Inside Vertical (1) and below Horizontal (1), add Horizontal (2).
6. Inside Horizontal (2), add the next sheets in the same order: NumAirbnbs, NumPrice, NumReviews, Map by County, Room Types.
7. Inside Vertical (1) and below Horizontal (2), add Horizontal (3).
8. Inside Horizontal (3), add Map by Airbnb and Reviews.

With these steps, you should end up with:



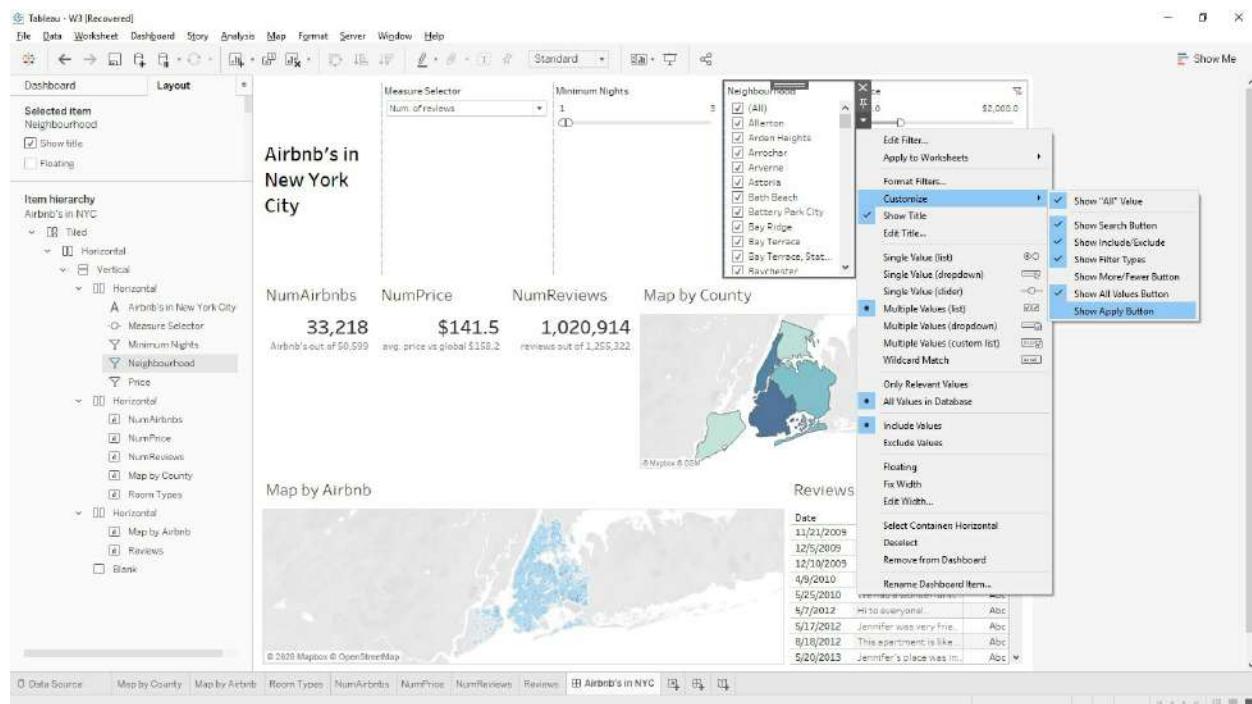
Now:

1. Move the Measure Selector from the Filters container to inside Horizontal (1), to the right of your text.
2. Remove the Blank inside Horizontal (1).
3. Remove the other three legends from the right (Room Type, Number of Records, Measure Selected).
4. Show all global filters by clicking any sheet in the dashboard, clicking the small arrow that appears, going to Filters, then toggling each of the three (Minimum Nights, Neighbourhood, Price). They will probably appear inside Horizontal (1). If not, move them there.

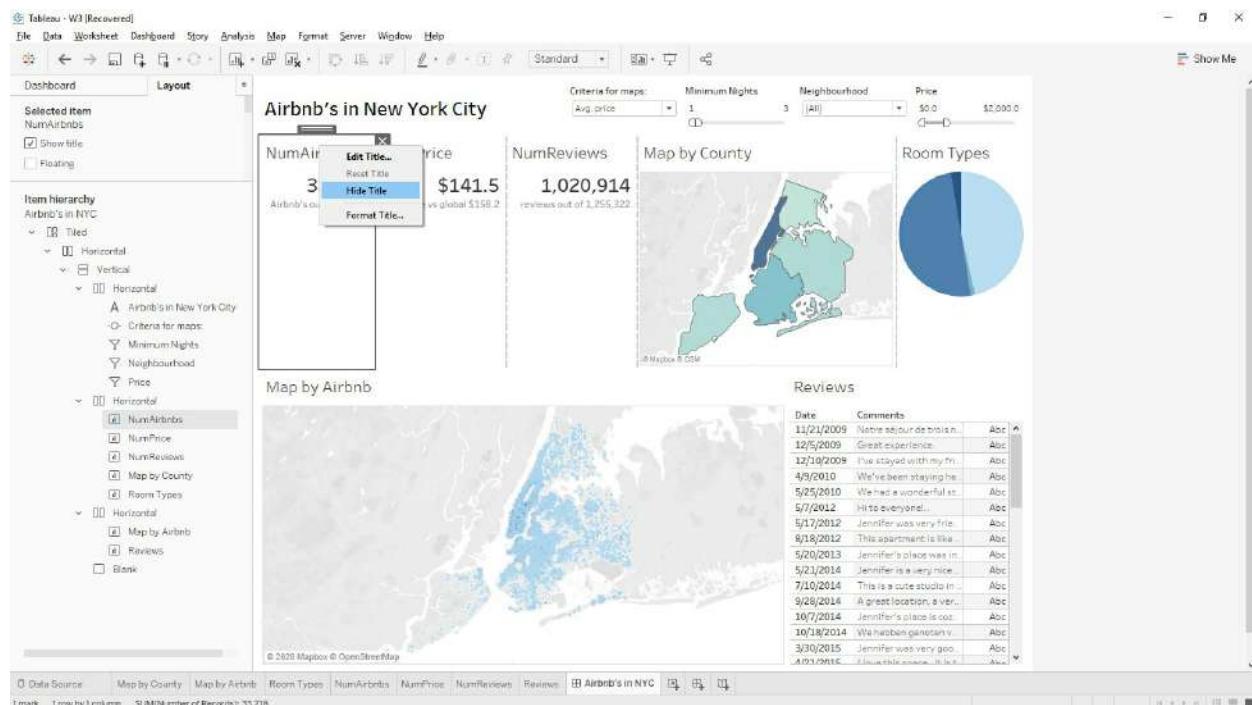


5. Format parameters and filters:

- Measure Selector: rename to “Criteria for Maps” by double clicking the title.
- Neighbourhood: change filter type to **Multiple Values (dropdown)** and **Show Apply Button** going in Customize:



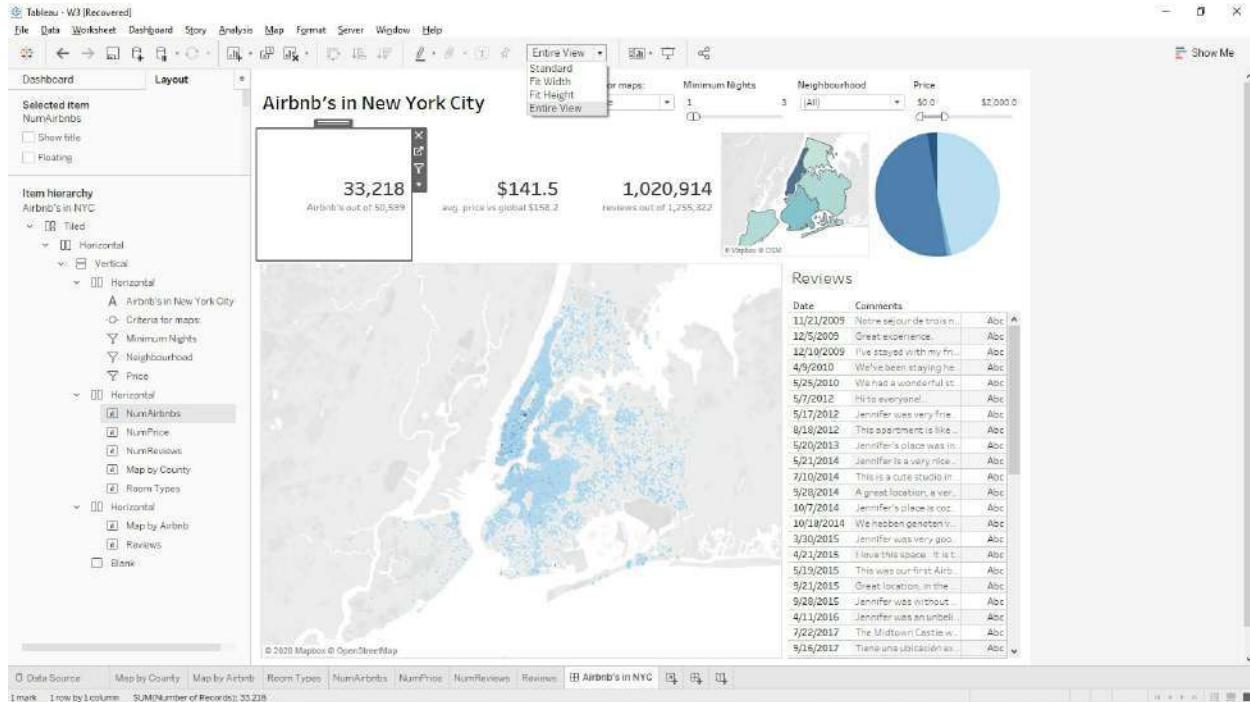
6. Toggle the **Use as Filter** button for *Map by County* and *Room Types*.
7. Hide all the sheet headers except for *Reviews* by right clicking them and selecting **Hide Title**:



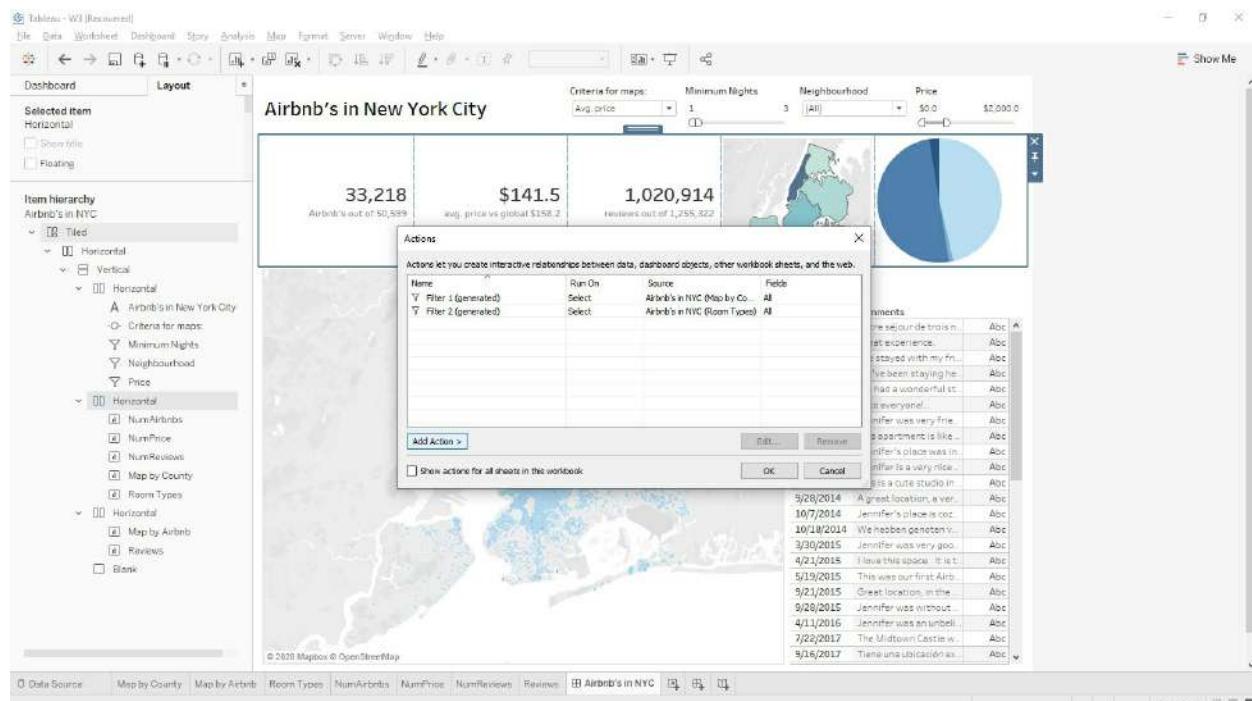
8. Adjust the height of Horizontal (2), so it's close to the height of

the *Room Types* pie graph.

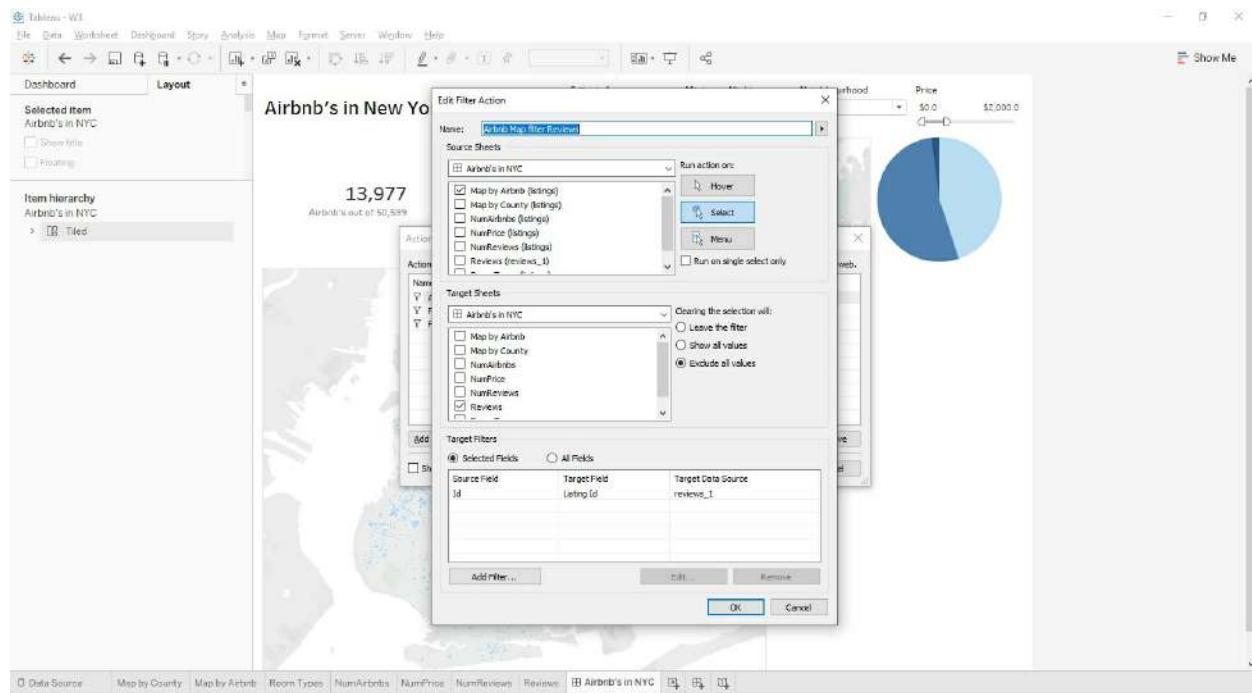
9. Adjust the width of *Map by County* so it's close to the width of the *Room Types* pie graph.
10. Select Horizontal (2) from the Item Hierarchy pane, click the arrow and select **Distribute Contents Evenly**.
11. For each of your *Num* graphs and for *Room Type*, select **Entire View** instead of Standard:



Open Dashboard > Actions > New Action > Filter. Notice how you already have two actions; these are the two graphs for which you clicked **Use as Filter (Map by County, Room Types)**. Click **Add Action**.



Call it **Airbnb Map filter Reviews**, make as **Source Sheets** only **Map by Airbnb**s, have it run on **Select**, not on Action, make as **Target Sheets** only **Reviews**, make it exclude all values when selection is cleared. Finally, set as target filters **Listing Id** from reviews.csv. Below these changes visually:

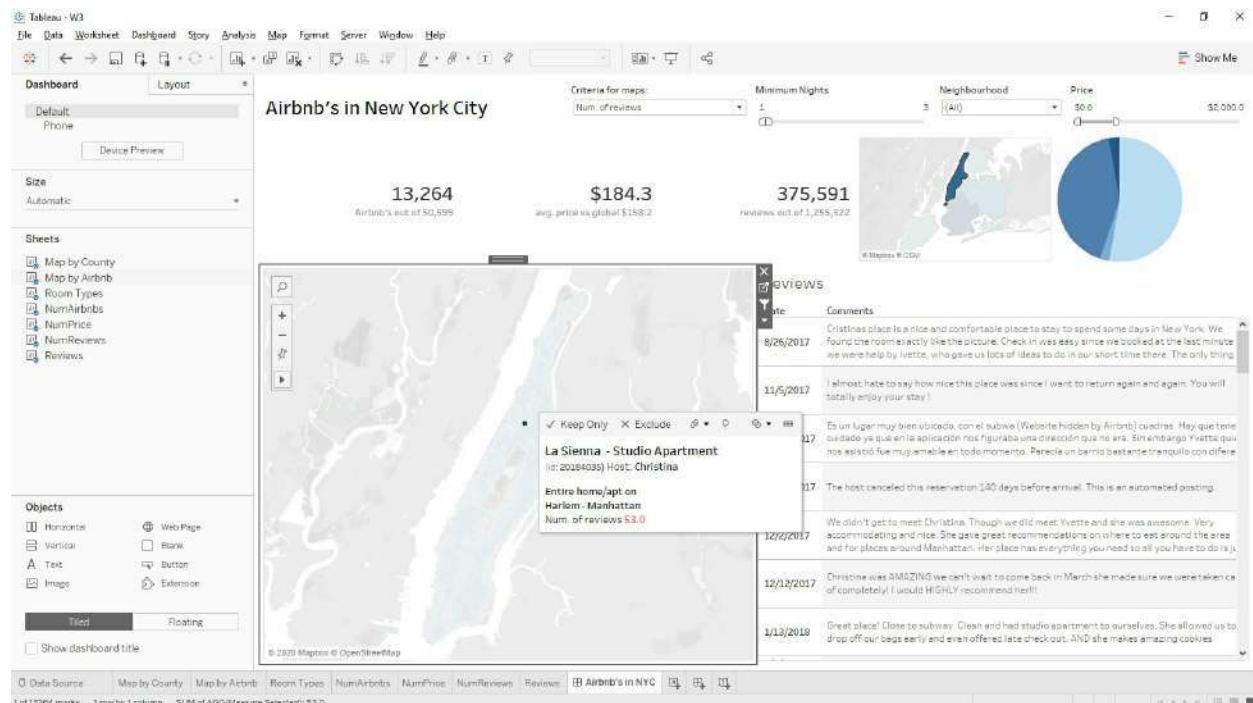


With this, you can go to **Reviews** sheet and delete the previous Id filter we had set (Id = 2595).

Delete the blank item at the bottom of Vertical (1) and **Distribute Items Evenly** for Horizontal (3), the one with the map and the reviews. Set *Reviews* to fit **Entire Width**, drag the **Comments** column width to be over the “Abc” column, and drag any row inside the table from the bottom to increase the row height.

Finally, set the dashboard **Size** to **Automatic** by clicking the option from the pane at the left.

Test your dashboard and see how the reviews table refreshes each time you select a different Airbnb from the map.



The dashboard is ready for use! Now let's add some additional formatting so you can also learn the basics of design.

Since we're making a dashboard that shows information on Airbnb, we would like to have it match the Airbnb app's design. If we search on the Internet what is the Airbnb's color palette (meaning what is the set of colors that they use in their user interface), we find this:

Rausch: #FF5A5F

Babu: #00A699

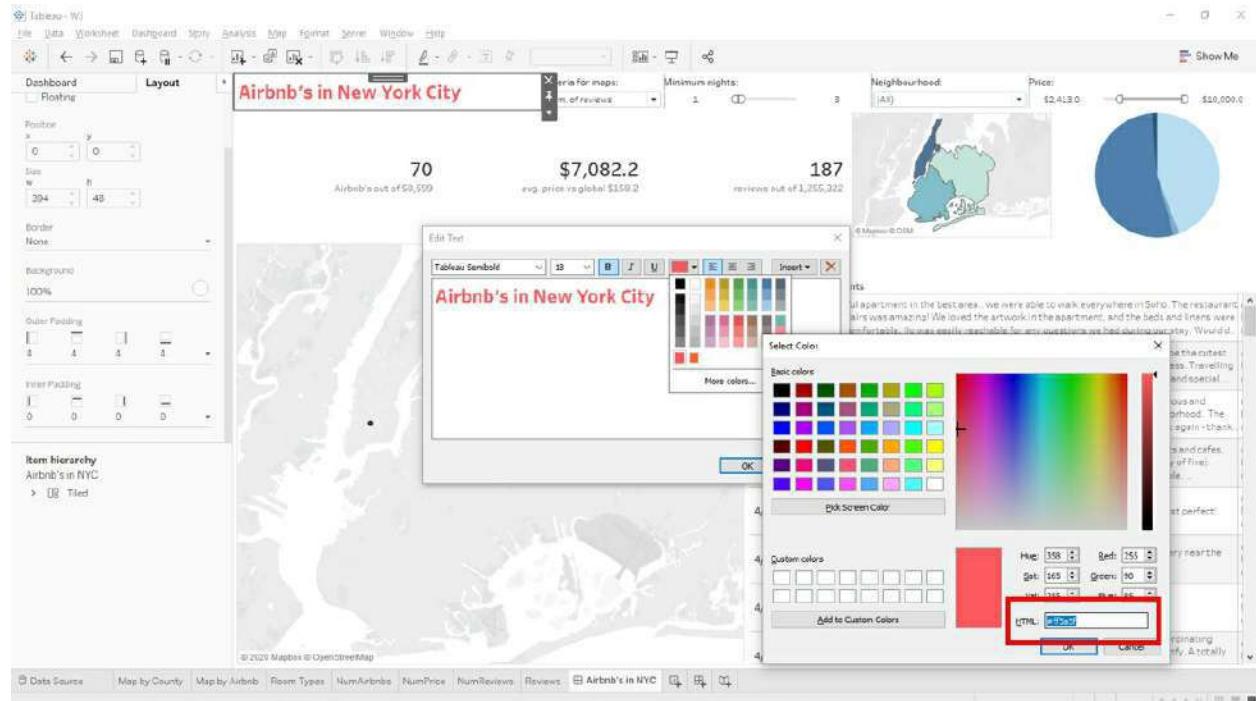
Arches: #FC642D

Hof: #484848

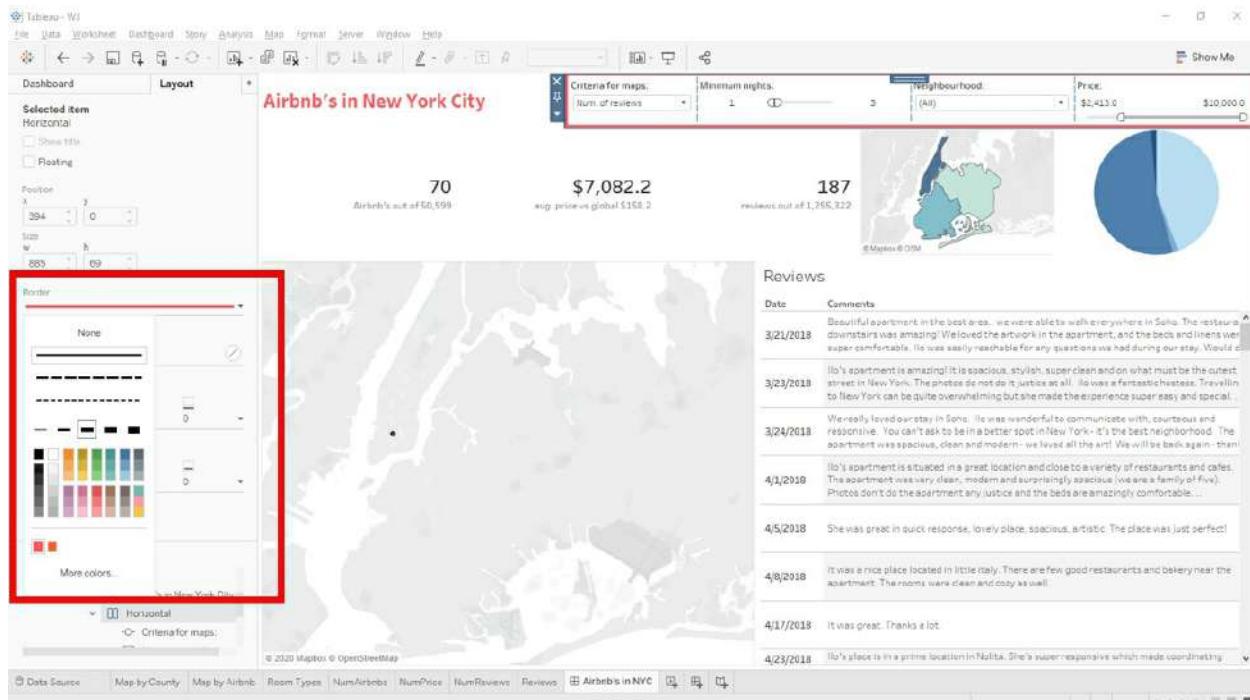
Foggy: #767676

We will use this color palette to give our dashboard a better design.

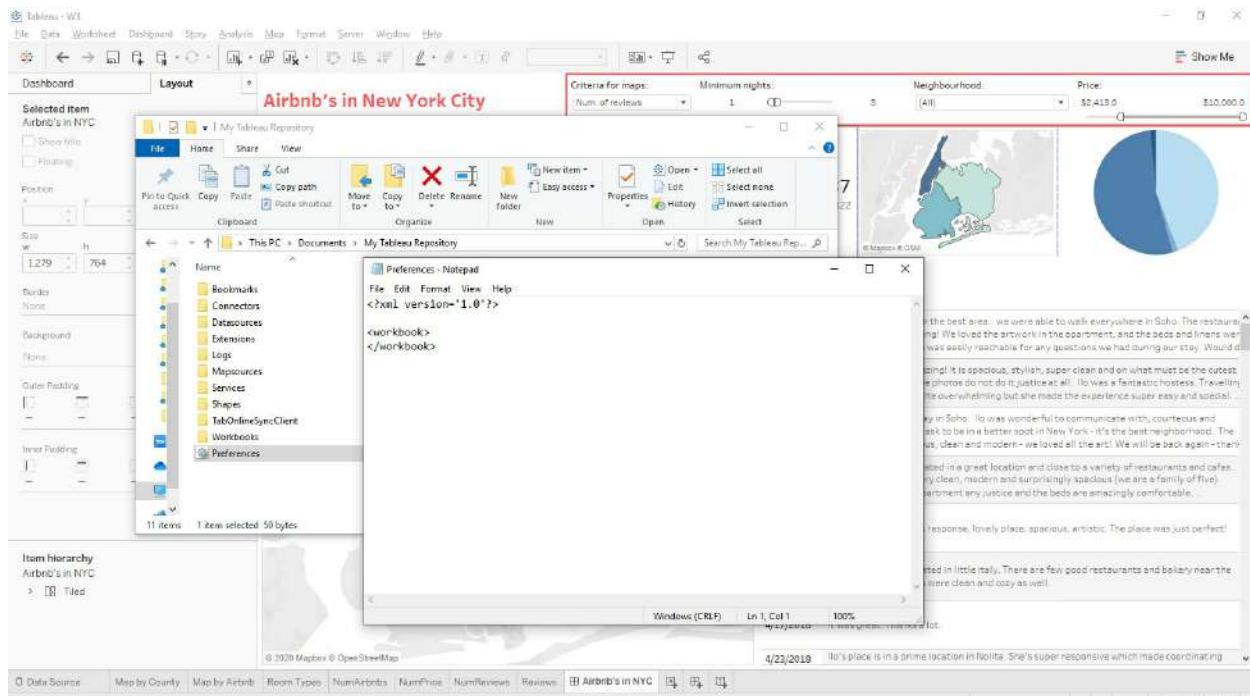
Start with the title, double click it and change Font Style to Tableau Semibold and change color to Rausch. You can copy the color HEX code shared above and paste it on the HTML textbox as in the below image:



Add a new Horizontal inside Horizontal (1), insert all four filters, and apply a **Rausch**-colored border to it using the Layout pane on the left.



Now we will make custom color palettes for the maps and the pie chart.
Open your Documents folder, go to My Tableau Repository, then open the Preferences file with Notepad.



Replace all the code with the below:

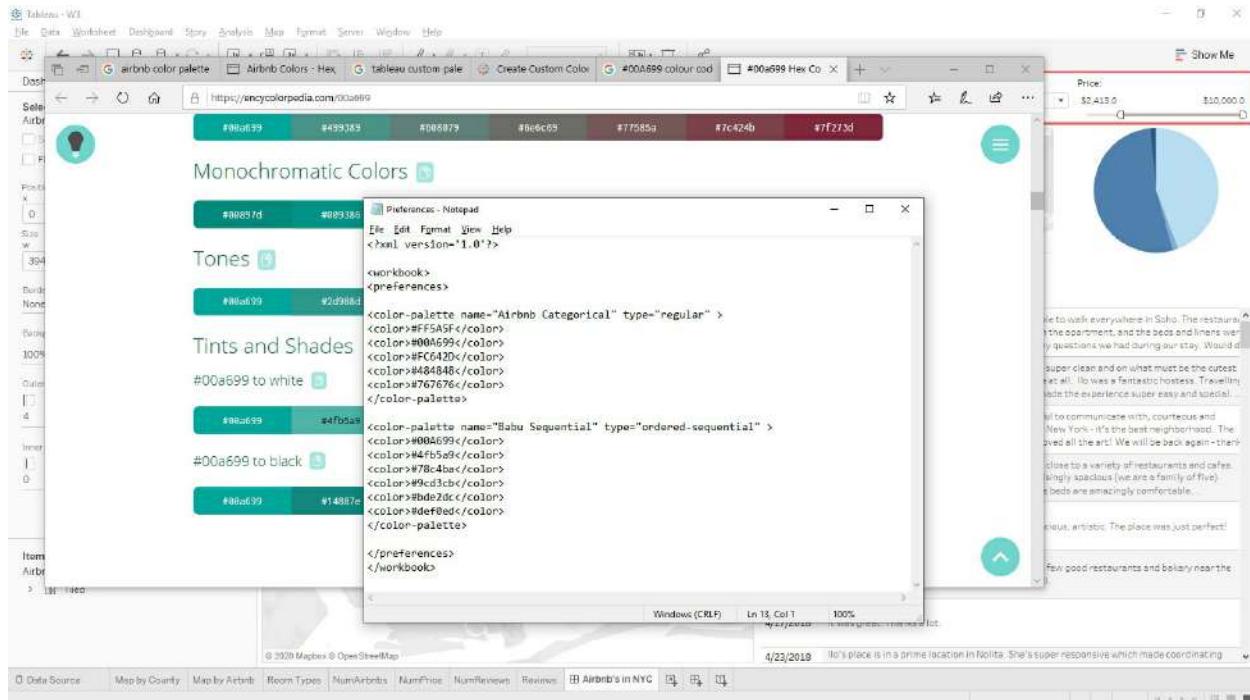
```
<?xml version='1.0'?>
```

```
<workbook>
<preferences>

<color-palette name="Airbnb Categorical" type="regular" >
<color>#FF5A5F</color>
<color>#00A699</color>
<color>#FC642D</color>
<color>#484848</color>
<color>#767676</color>
</color-palette>

<color-palette name="Babu Sequential" type="ordered-sequential" >
<color>#00A699</color>
<color>#4fb5a9</color>
<color>#78c4ba</color>
<color>#9cd3cb</color>
<color>#bde2dc</color>
<color>#def0ed</color>
</color-palette>

</preferences>
</workbook>
```

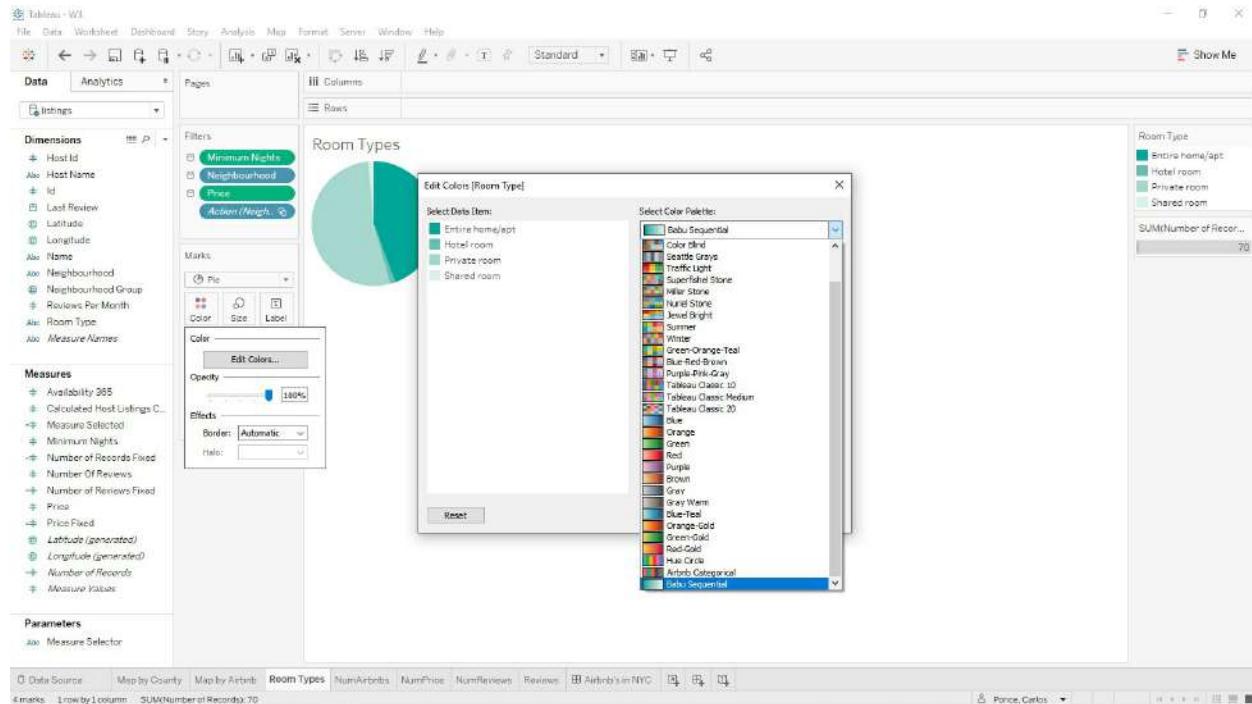


This code is generating two sets of color palettes; one called **Airbnb Categorical** and one called **Babu Sequential**. There are two types of color palettes; the ones that use different colors altogether (categorical), and the ones that use one color in different gradients (sequential). Airbnb Categorical is using the Airbnb color palette we already discussed earlier. Babu Sequential is using a new palette based on the monochromatic colors of the Babu color that I found on a quick search.

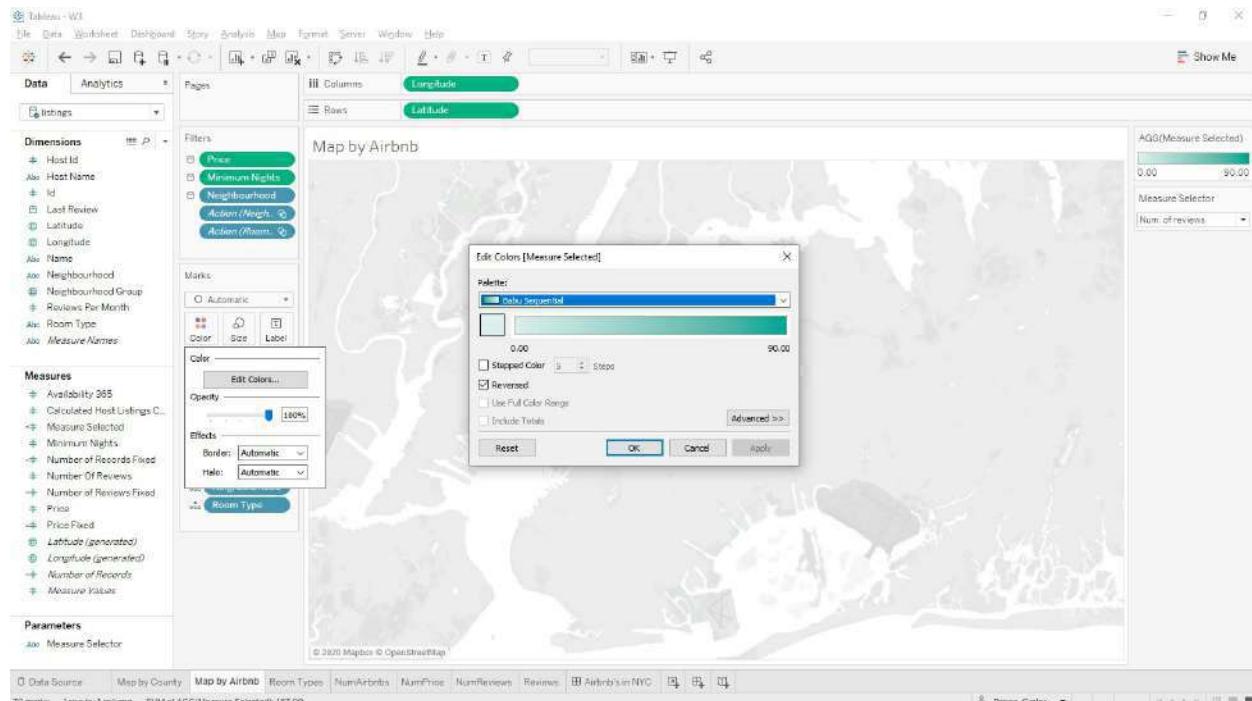
Tip

Always look on the web for nice color palettes! It's much better than selecting them yourself, unless you have some background on design.

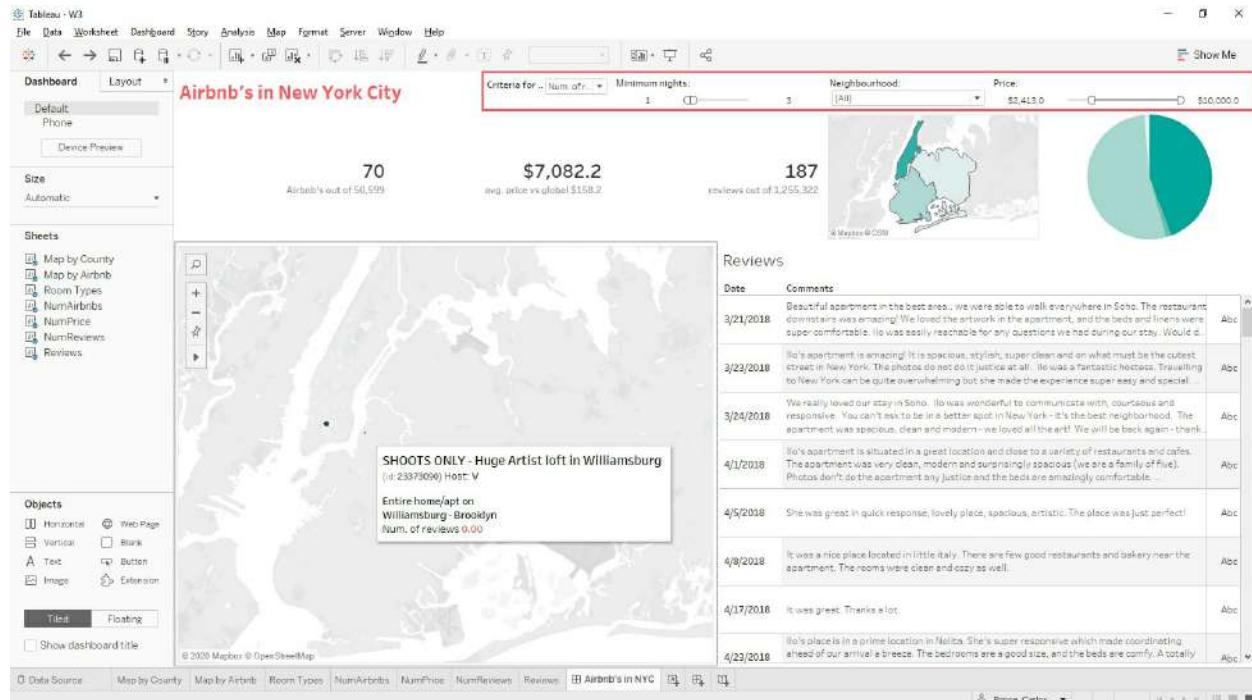
Save the Preferences file, and restart Tableau. Go to Room Types, edit colors, and look for our two new color palettes at the bottom.



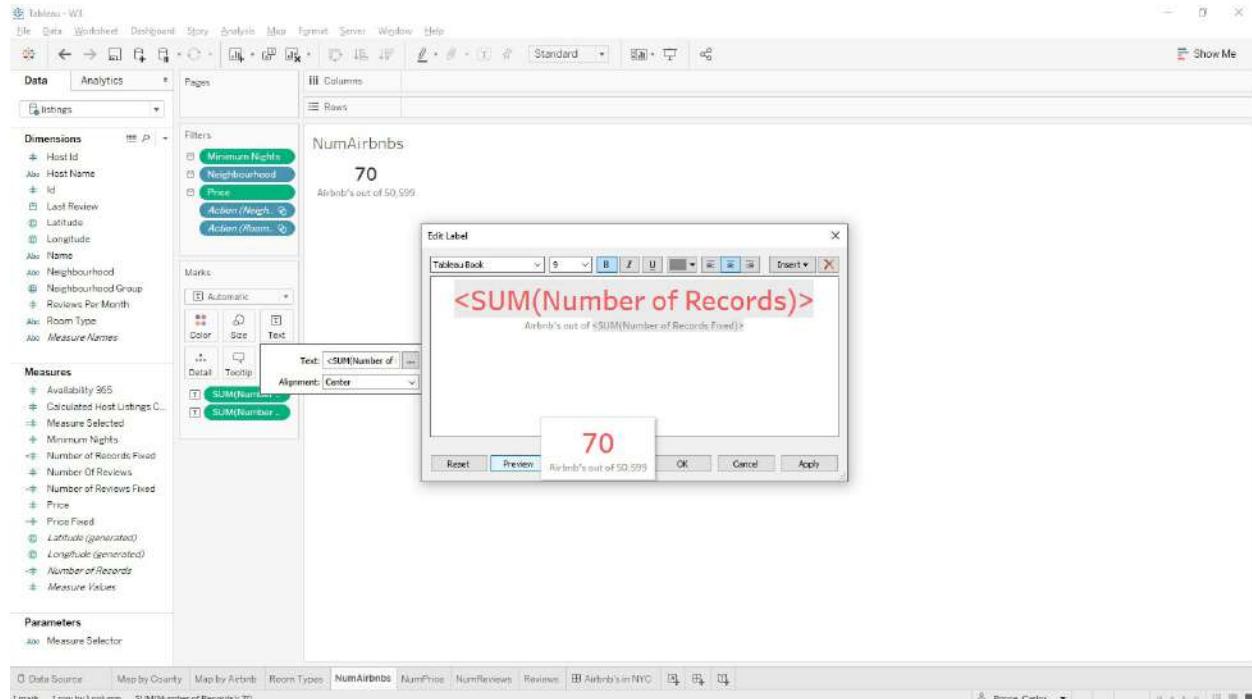
Select **Babu Sequential** for *Room Types* and repeat for *Map by County* and for *Map by Airbnb*.



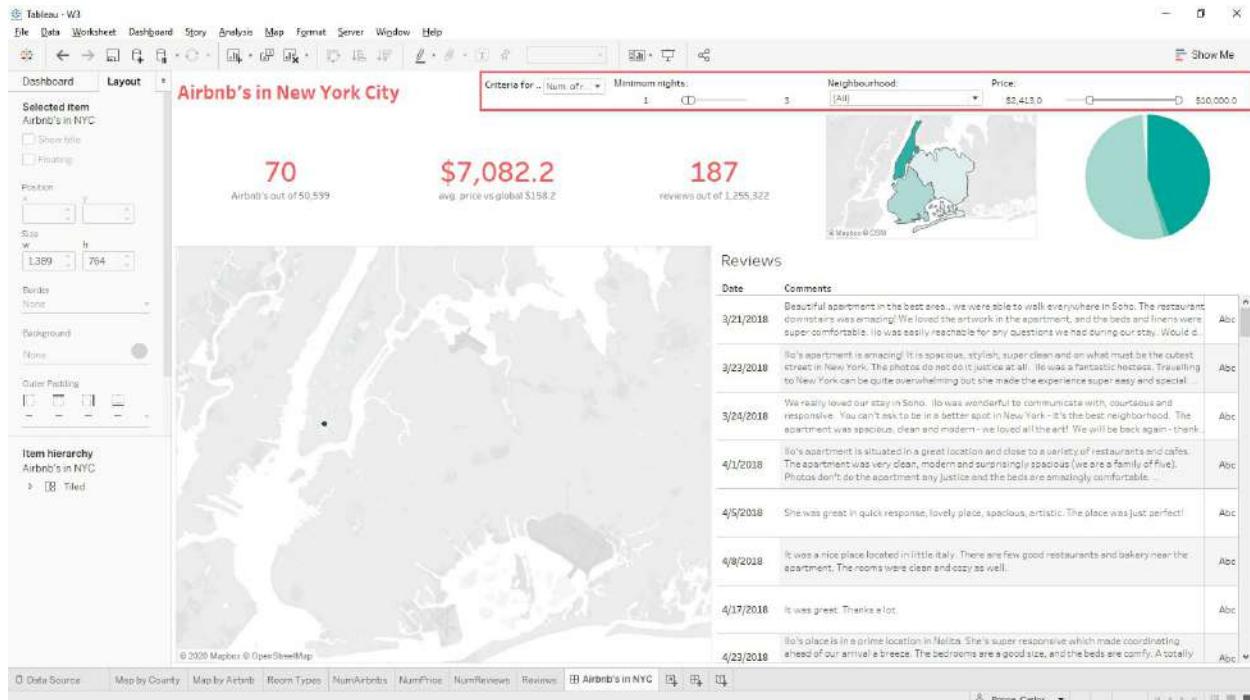
The dots will look a bit too bright because we have big outliers, but that is okay. This process should make the dashboard look like this now:



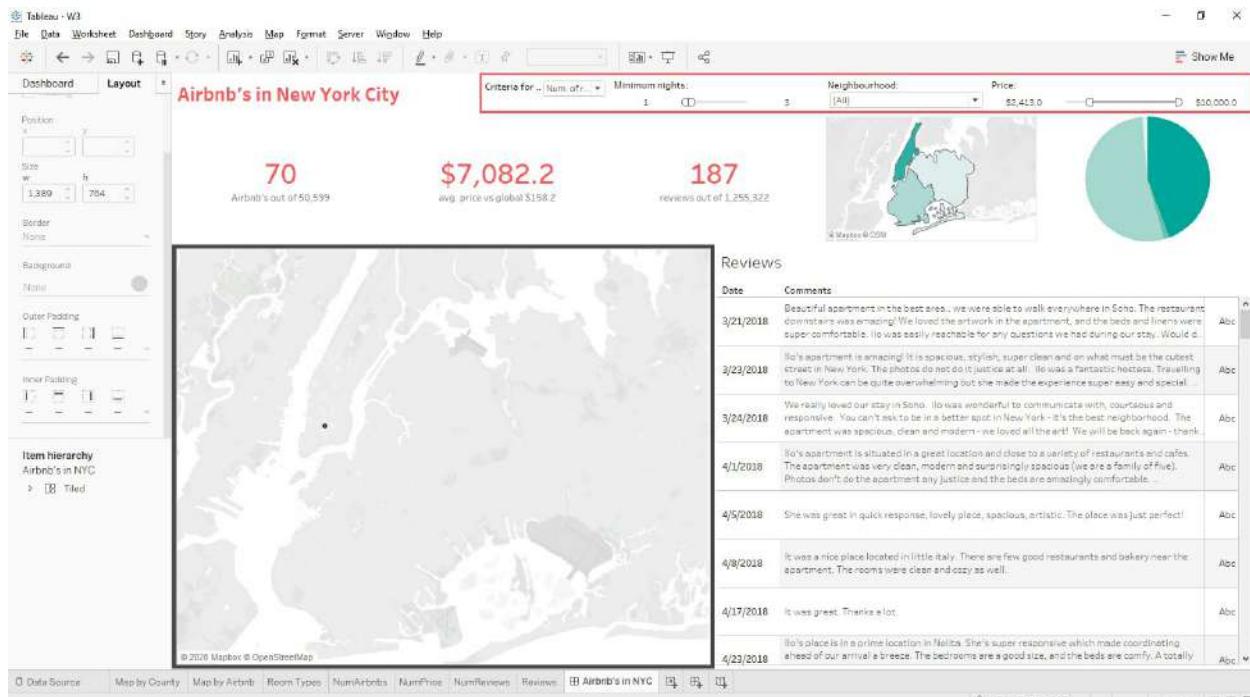
Let's format the *Num* sheets now. For NumAirbnbs, NumPrice, and NumReviews, edit the text label so the first number is color Rausch, font size 26, and the bottom will stay as we already formatted it.



Now the dashboard looks like this:

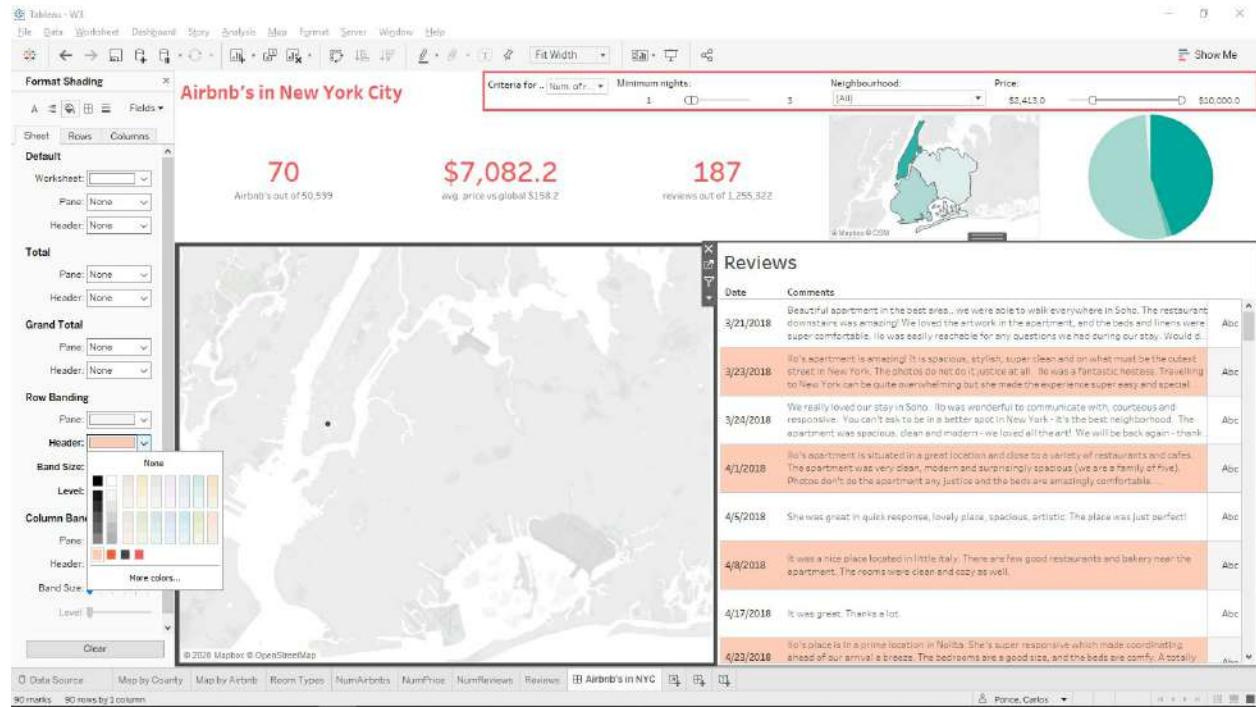


Click the Map by Airbnb element and on the Layout pane, set outer padding to 0 and add a border of the widest option colored Hof.

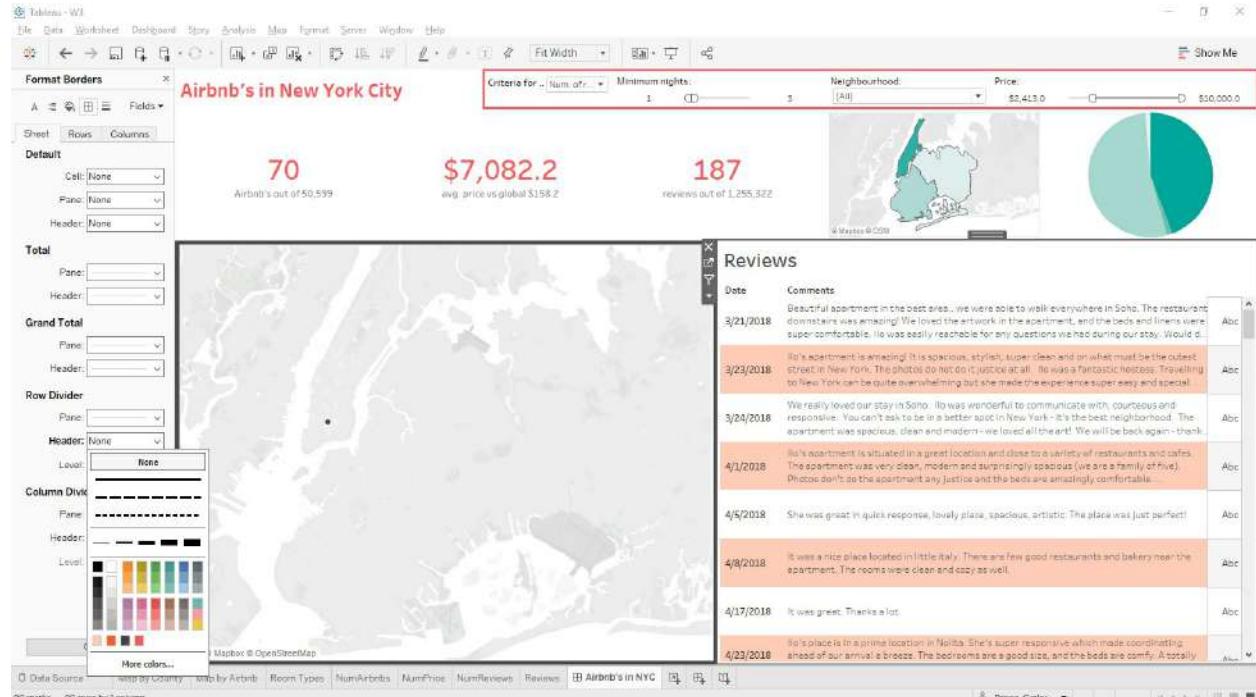


Now format the *Reviews* graph. Click the table, select Format > Shading on the Action bar at the top. From the left pane and on the Sheet tab, look for the **Row Banding** section, and on the **Header** and **Pane** dropdowns choose Rausch or look for #FFCDB7. Move the Level slider to the left, so the color

also shows on the **Date** column, not only on **Comments**.



Now go to **Format Borders** (click the fourth icon on the Formatting pane), look for the **Row Divider** section, and on the **Header** and **Pane** dropdowns click **None** so it does not show gray lines as row dividers.



Go to the *Reviews* sheet, move **Comments** from the Rows shelf to the Text

property. The **Pane** formatting you did earlier made that the comments have formatting even though they're not part of the header anymore (meaning they're not on the rows/columns shelves).

This screenshot shows a Tableau interface with a 'Format Borders' panel open on the left. The 'Comments' shelf is selected. The 'Reviews' field is highlighted with a blue border. The 'Marks' card shows 'Automatic' as the type, with options for Color, Size, and Text. Below the marks are three comment cards, each with a date and a review text. The reviews are styled with orange backgrounds and white text.

Date	Review Text
3/21/2018	Beautiful apartment in the best area... we were able to walk everywhere in Soho. The restaurant downstairs was amazing! We loved the artwork in the apartment, and the beds and linens were super comfortable. Illo was easily reachable for any questions we had during our stay. Illo's apartment is amazing! It is spacious, stylish, super clean and on what must be the cutest street in New York. The photos do not do it justice at all. Illo was a fantastic hostess. Travelling to New York can be quite overwhelming but she made the experience super easy and special.
3/23/2018	We really loved our stay in Soho. Illo was wonderful to communicate with, courteous and responsive. You can't ask to be in a better spot in New York - it's the best neighborhood. The apartment was spacious, clean and modern - we loved all the art! We will be back again - than...
4/1/2018	Illo's apartment is situated in a great location and close to a variety of restaurants and cafes. The apartment was very clean, modern and surprisingly spacious (we are a family of five). Photos don't do the apartment any justice and the beds are amazingly comfortable.
4/5/2018	She was great in quick response, lovely place, spacious, artistic. The place was just perfect!
4/6/2018	It was a nice place located in little Italy. There are few good restaurants and bakery near the apartment. The rooms were clean and cozy as well.
4/17/2018	It was great. Thanks a lot.
4/23/2018	Illo's place is in a prime location in Nolita. She's super responsive which made coordinating ahead of our arrival a breeze. The bedrooms are a good size, and the beds are comfy. A totally pleasant stay overall.
4/29/2018	If you're considering Illo's place, book now!! The location is perfect and the apartment was clean and stylish. Illo was very responsive and a joy to work with. Five stars in all categories. I will be staying here next time I visit NYC. There was a little sidewalk art market both weekend morn...
5/6/2018	Belägenheten är perfekt. Mitt i det här, men alltgeväl rymligt/rörligt. Lite söt med mängd plats till 4
5/14/2018	

This screenshot shows a Tableau dashboard titled 'Airbnb's in New York City'. The dashboard includes a map of New York City with a green overlay, a pie chart showing price distribution, and several summary statistics: '70' (Airbnb's out of 50,599), '\$7,082.2' (avg. price vs global \$158.2), and '187' (reviews out of 1,255,322). The 'Reviews' shelf is selected in the 'Format Borders' panel, and the 'Comments' card is highlighted with a blue border. The reviews are styled with orange backgrounds and white text.

Date	Review Text
3/21/2018	Beautiful apartment in the best area... we were able to walk everywhere in Soho. The restaurant downstairs was amazing! We loved the artwork in the apartment, and the beds and linens were super comfortable. Illo was easily reachable for any questions we had during our stay. Would definitely stay again.
3/23/2018	Illo's apartment is amazing! It is spacious, stylish, super clean and on what must be the cutest street in New York. The photos do not do it justice at all. Illo was a fantastic hostess. Travelling to New York can be quite overwhelming but she made the experience super easy and special. We loved it!!
4/1/2018	We really loved our stay in Soho. Illo was wonderful to communicate with, courteous and responsive. You can't ask to be in a better spot in New York - it's the best neighborhood. The apartment was spacious, clean and modern - we loved all the art! We will be back again - thank you!
4/5/2018	Illo's apartment is situated in a great location and close to a variety of restaurants and cafes. The apartment was very clean, modern and surprisingly spacious (we are a family of five). Photos don't do the apartment any justice and the beds are amazingly comfortable.
4/6/2018	She was great in quick response, lovely place, spacious, artistic. The place was just perfect!
4/8/2018	It was a nice place located in little Italy. There are few good restaurants and bakery near the apartment. The rooms were clean and cozy as well.
4/17/2018	It was great. Thanks a lot.
4/23/2018	Illo's place is in a prime location in Nolita. She's super responsive which made coordinating ahead of our arrival a breeze. The bedrooms are a good size, and the beds are comfy. A totally pleasant stay overall.

Explained:

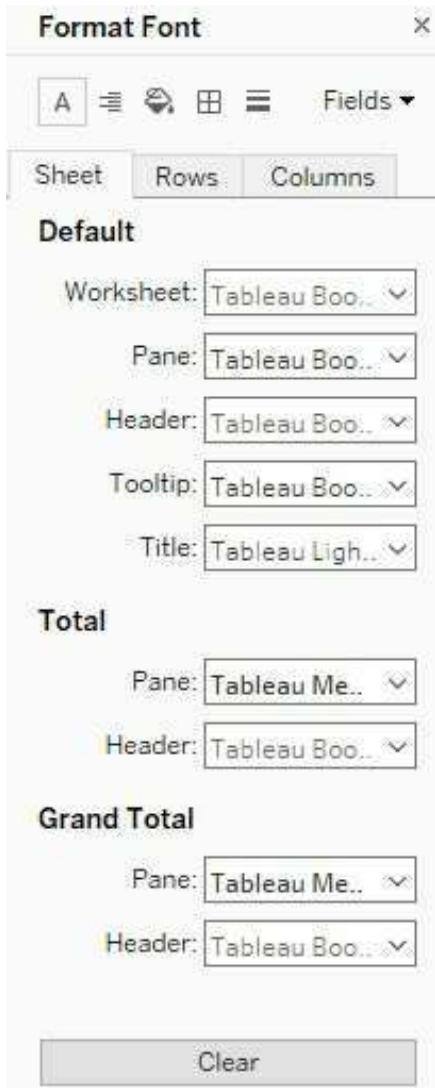
The Format button on the Action bar is used to improving design and formatting your dashboards. You can apply a default format for the whole **workbook** by going into Format > Workbook. You will be able to apply default fonts for your worksheets, your tooltips, the worksheet/dashboard/story titles, and grid lines.

You can also apply default format for **dashboards** or **stories** by going into Format > Dashboard or Format > Story. You can go now to see which options you have available for both.

Finally, you can format at a **sheet** level, which is what we did on this workout. You do that by going to Format and then choosing any of the five options:

- Font
- Alignment
- Shading
- Borders
- Lines

This is an example of how the **Format** pane looks when looking on Font:



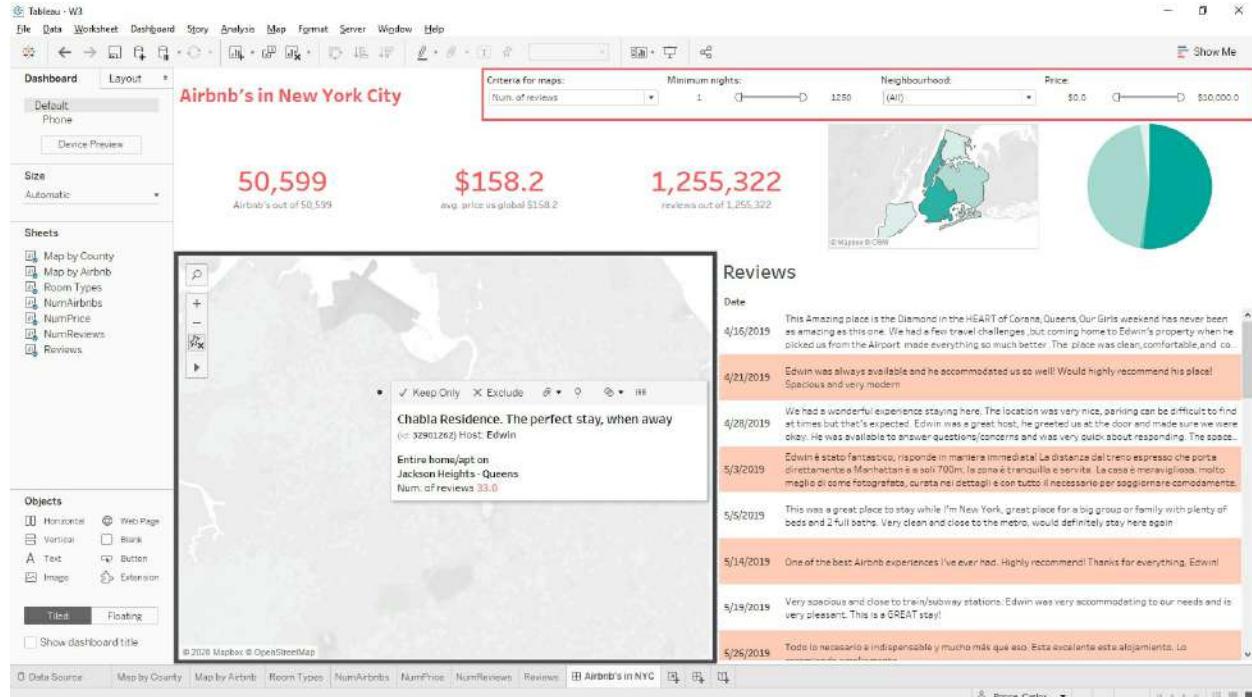
On the top you have the other four options to switch through them. To the right where it says **Fields**, you can select a specific dimension/measure on your sheet to apply formatting just to that one (I recommend you only apply format to all fields). Below you have three tabs: **Sheet**, **Rows**, and **Columns** (I recommend you only use Sheet). Then inside you will have **Default**, **Total**, and **Grand Total**. The first is the most common one you will change. If your table shows subtotals, you will also use Total. If your table shows Grand Totals, then you will also need to format the Grand Total section.

On **Shading** and **Borders**, you will have additional configuration for your row/column dividers and banding.

Finally, the Clear button lets you move back to the default settings for the specific tab you're viewing. Helpful when you get lost on all the changes you

did and want to roll back.

This is the ending result:



Congratulations for your third dashboard! Here you can see how dashboards can be helpful in many different situations. I guarantee you that any Airbnb user will find this very helpful. If you wish, you can go to the Inside Airbnb data site and download more data like this but from another city you would like to visit: <http://insideairbnb.com/get-the-data.html>

- Data join
- Data blending
- Data union
- The importance of knowing granularities to correctly join tables
- Types of join: inner, left, right, full outer
- Data types
- Data extracts
- Text File Properties
- Geographical Roles
- Parameters
- Create a Measure Selector for displaying different measures on a graph

- Editing Tooltips
- Level-of-detail calculations: Fixed
- Editing Text Labels – and formatting
- Default Properties
- Customizing – and formatting - filters in dashboards
- Dashboard actions
- Design concepts: color palettes, borders, spacing, resizing
- Custom color palettes
- The Format pane

You might want to recap and make sure you're understanding all these concepts. It's not a bad idea to try to recreate the dashboard all by yourself without reading the steps.

Each workout considers you already master what was taught in the previous ones, so the steps will be gradually less explained. With that said, we move on to the fourth workout:

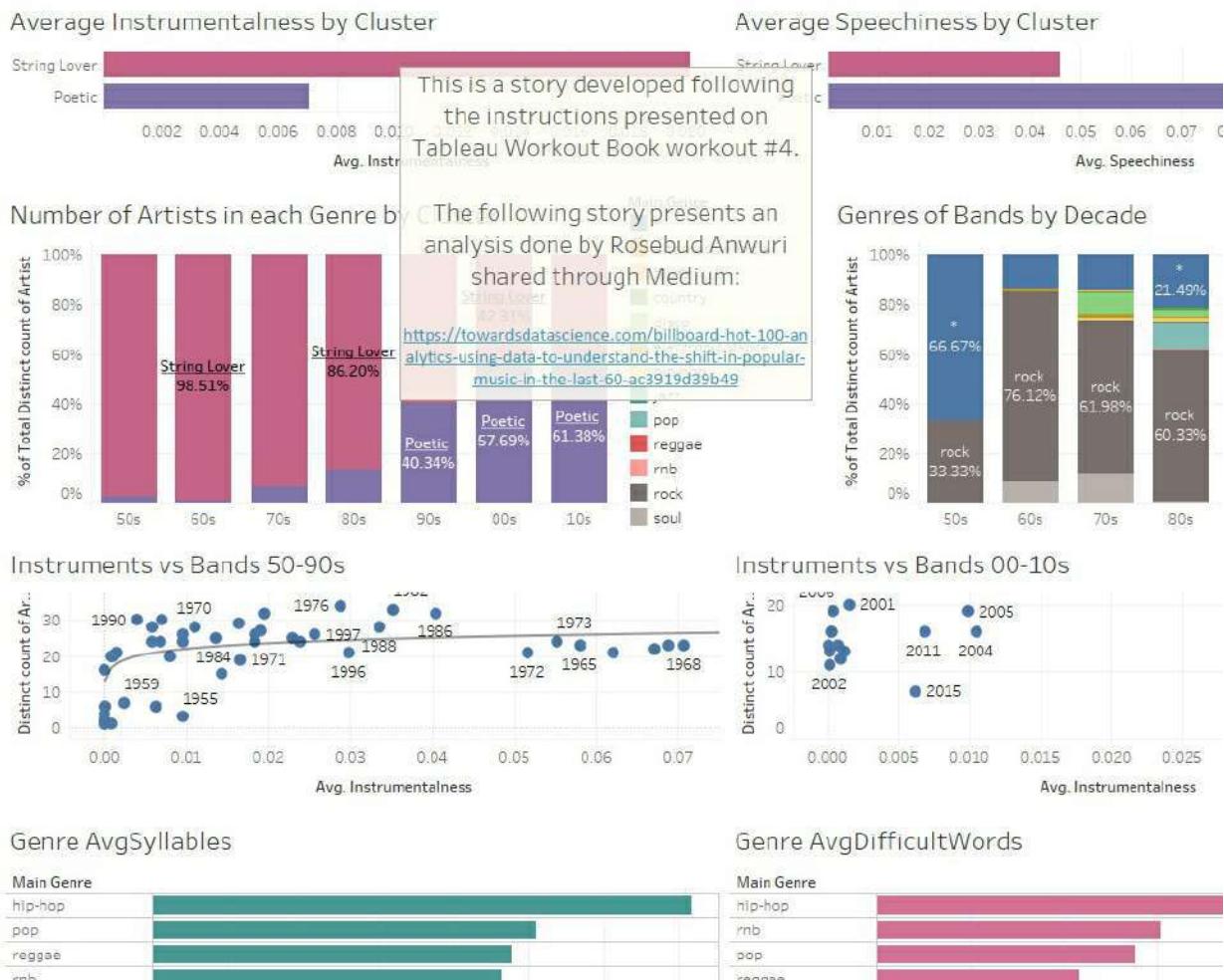
Workout 4. Billboard Hot 100 Analytics

On this workout you will present the results of an analysis done by Rosebud Anwuri on Billboard's Hot 100 songs throughout the years. If you wish to read her full analysis, you can check her article on Medium:

<https://towardsdatascience.com/billboard-hot-100-analytics-using-data-to-understand-the-shift-in-popular-music-in-the-last-60-ac3919d39b49>.



Billboard Hot 100 Analytics



Workout Start

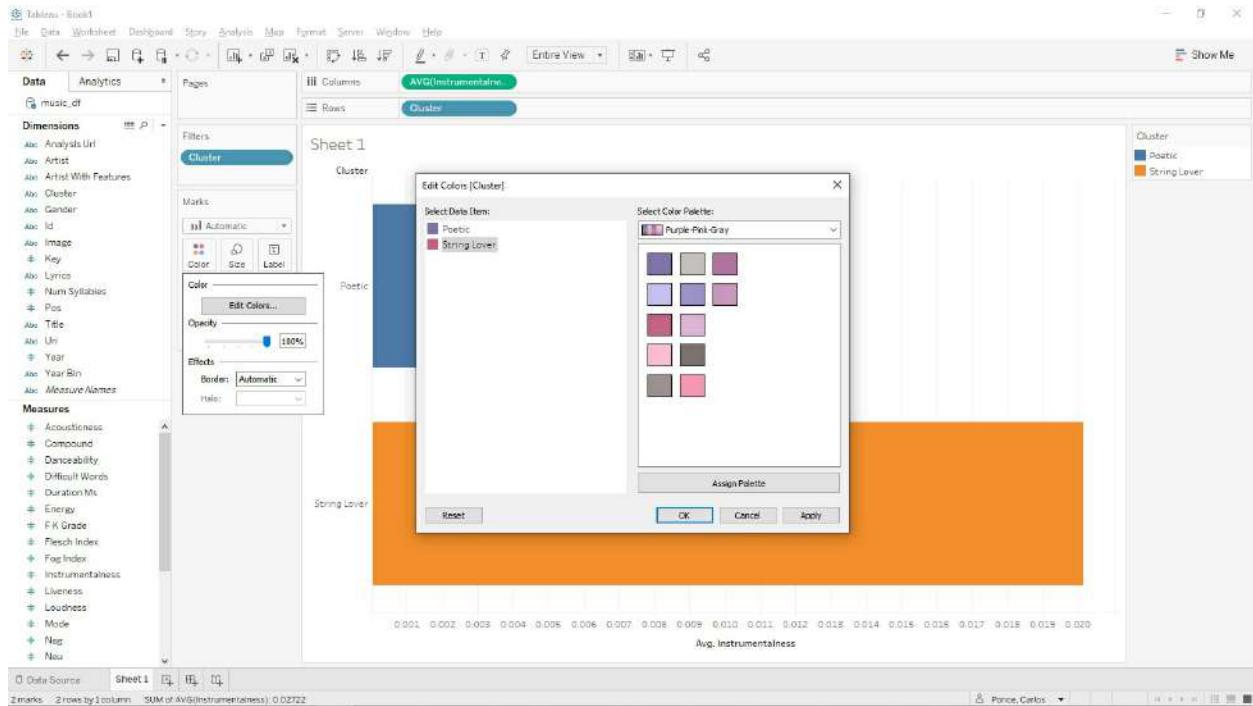
Connect to W4 Music dataset.csv, and again change the text qualifier to [“] accessing the **Text File Properties** from the canvas, so Tableau reads the table correctly.

The screenshot shows the Tableau Data Source interface. At the top, there's a navigation bar with File, Data, Server, Window, Help, and a Connections section where 'music_df' is selected. Below the navigation is a 'Connections' pane listing various CSV files like 'lyrics.csv', 'music_df.csv', etc. A 'File' dropdown menu is open, showing options like 'Use Data Interpreter' and 'Extract'. On the right, there are 'Connection' and 'Filters' sections. The main area displays a preview of the 'music_df.csv' data with columns such as 'Title', 'Year', 'Flesch Index', 'Num Words', and 'Artist'. A tooltip for 'Go to Worksheet' is visible at the bottom left of the preview area.

To start, make the following sheet:

The screenshot shows the Tableau Worksheet interface. The top navigation bar includes File, Data, Worksheet, Dashboard, Story, Analysis, Map, Format, Server, Window, Help, and a 'Show Me' button. The left sidebar contains a 'Dimensions' section with fields like 'Cluster', 'Artist', 'Title', 'Year', and 'Instrumentalness'. A 'Measures' section lists 'Accoutness', 'Compound', 'Danceability', 'Difficult Words', 'Duration Min', 'Energy', 'FK Grade', 'Flesch Index', 'Fog Index', 'Instrumentalness', 'Liveness', 'Loudness', 'Mode', 'Neg', and 'Pos'. The main workspace shows a horizontal bar chart titled 'Sheet 1' with 'Cluster' on the X-axis and 'Instrumentalness' on the Y-axis. The bars are colored by 'Cluster'. A tooltip for 'String Love' is shown, indicating an average instrumentalness of 0.02722. The bottom status bar shows 'Marks: 2 rows by 1 column SUM OF AVG(Instrumentalness): 0.02722'.

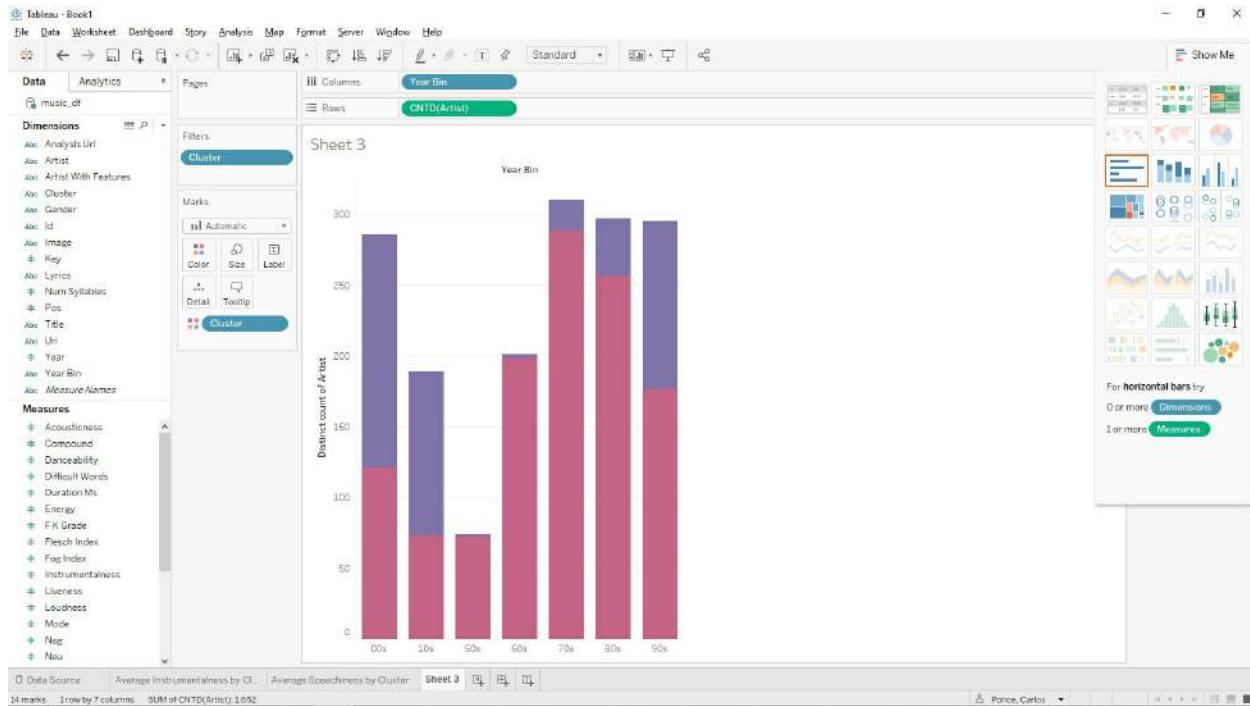
Drag **Cluster** and **Instrumentalness**, change the aggregation of the second from sum to average, and filter out the “Null” Cluster. Set the graph size to Entire View and drag **Cluster** to color (Ctrl + drag it from Rows so it stays there too). For colors, pick up a couple of them from the Purple-Pink-Gray color palette, like below:



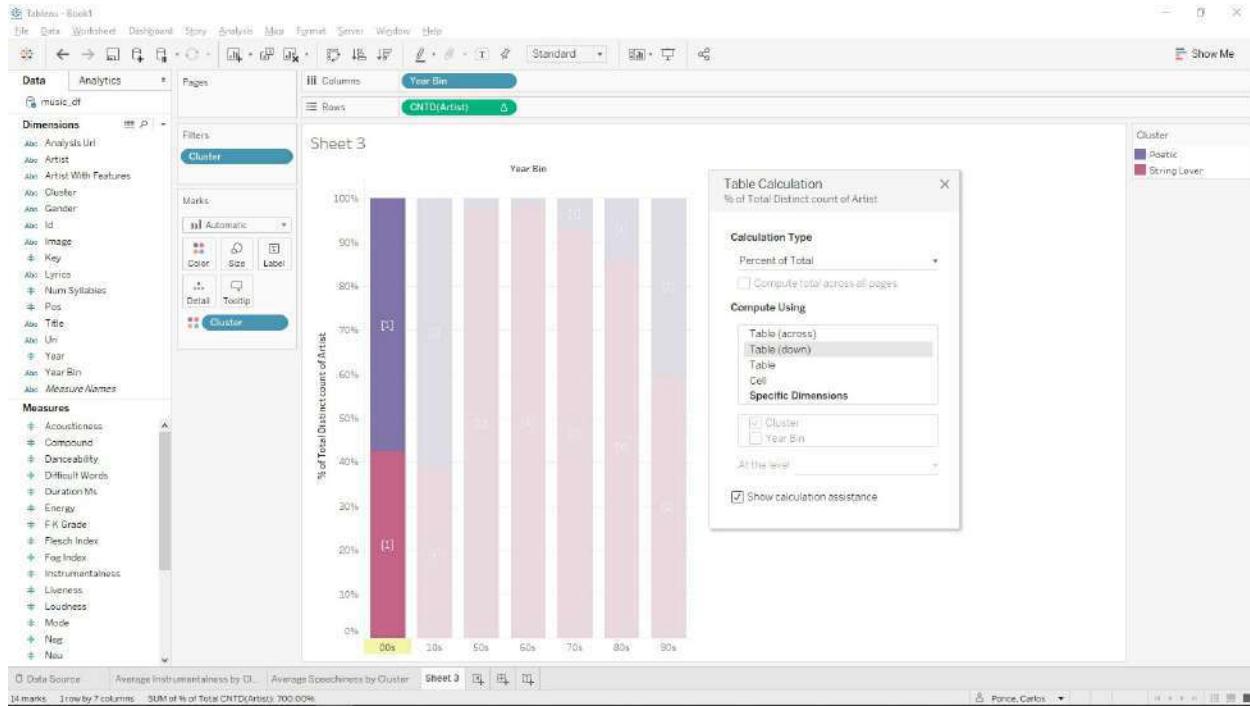
Drag the “String Lover” bar to the top by dragging the label, so it’s manually sorted. **Hide Field Labels for Rows** by right clicking the small “Cluster” header, right below where it says Sheet 1.

Call the sheet “Average Instrumentalness by Cluster”. Duplicate it and replace AVG(Instrumentalness) to **AVG(Speechiness)** on Columns. Also rename the sheet to “Average Speechiness by Cluster”.

Make another blank sheet, add **Cluster** to the Color property. Again, filter out the Null cluster. Drag **Year Bin** to columns and add **COUNTD(Artist)** to Rows with right-click.



Add table calculation Percentage of total, then Edit Calculation so it's Table (down) instead of (across). You should see that it's always giving the 100% total by bar:

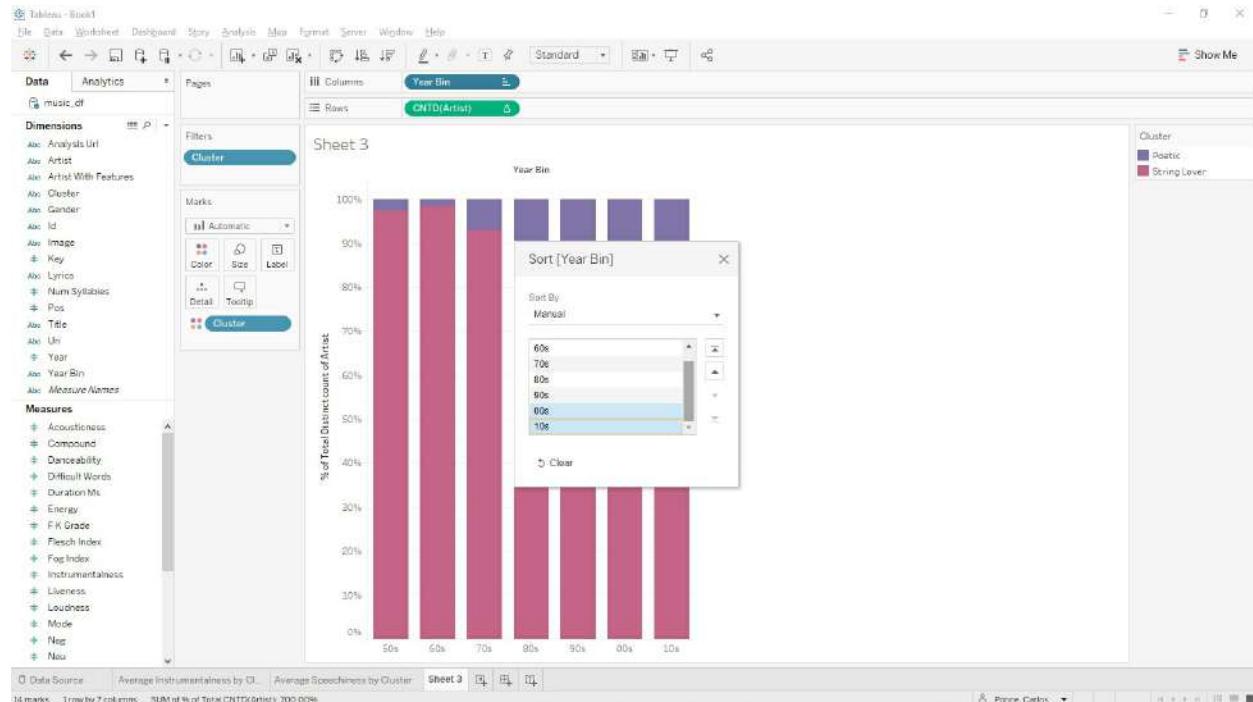


Explained:

The **Compute Using** on a table calculation lets you know the reference that will be taken to make the calculation. Commonly, you will either use Table (across) and Table (down). The first means that it will consider all the elements or cells across your table (horizontally). The second means that it will consider them vertically on your table. Try to switch the compute on the last step so you see how the 100% is based on the total for Poetic and String Lovers (meaning the Cluster) when you put it on Table (across), and based on the total for the Year Bin when you use Table (down).

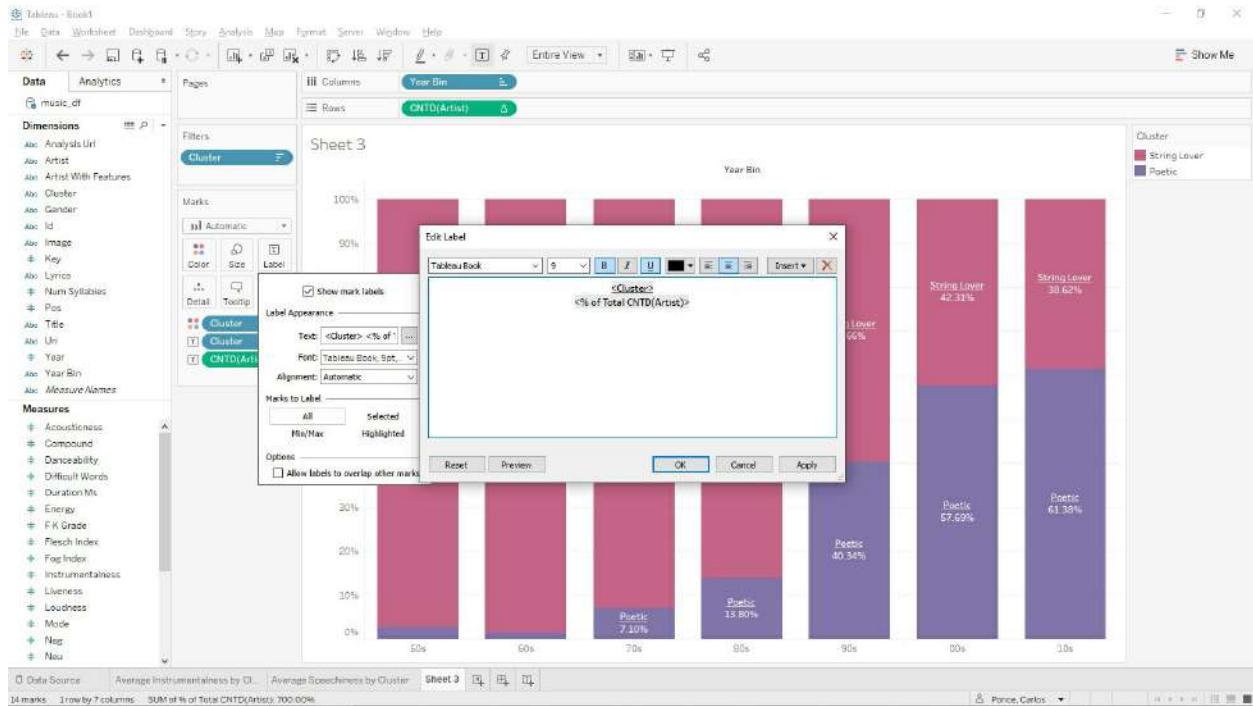
There will be advanced cases when you can compute using only specific dimensions, but we will cover such cases in later workouts.

Right-click the **Year Bin** pill, click Sort, set to manual, and drag **00s** and **10s** to the bottom of the list (they're 2000 and 2010, not 1900 and 1910 so they should go at end).



Also, apply sorting on the **Cluster** pill, just invert the data source order to descending so “String Lovers” keeps being the one at the top.

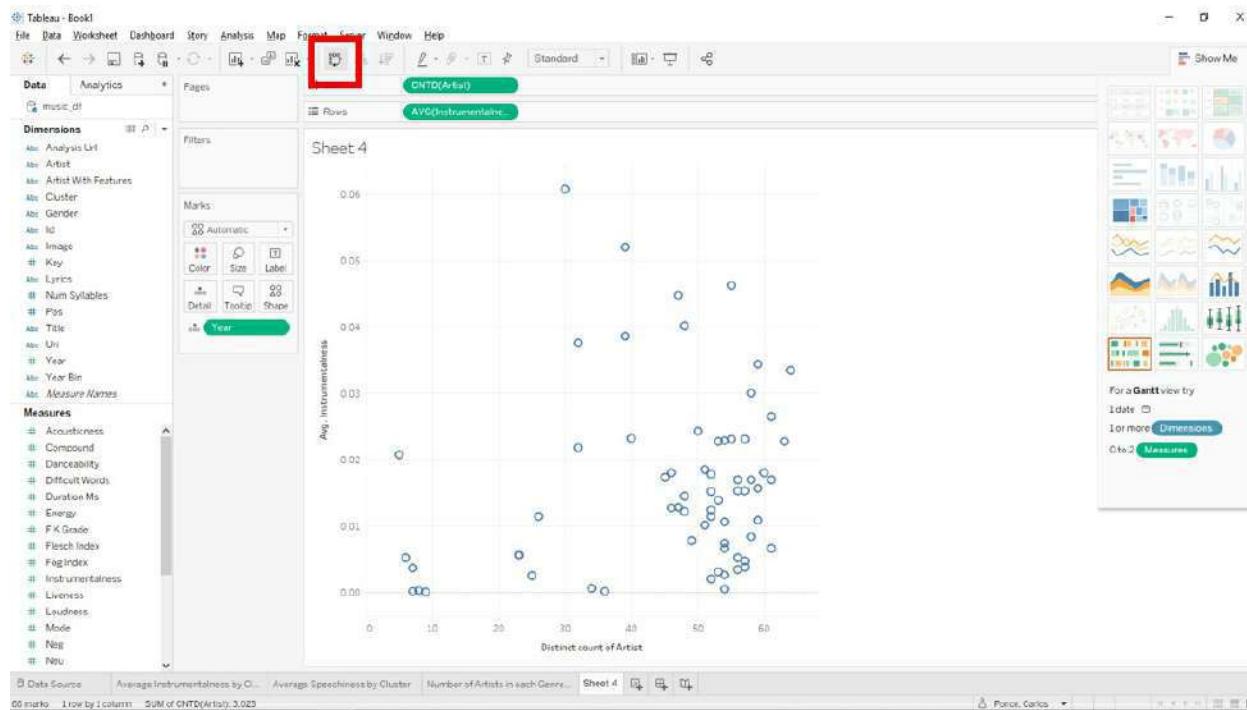
With the Ctrl key held, drag **Cluster** from Color and **CNTD(Artist)** from Rows to the Label property on Marks, so you now have two pills of each. Set the graph to fit Entire View. Edit the label to have it all bold, underlining the cluster name as shown below:



Hide Field Labels for Columns by right clicking the small **Year Bin** header title right above the 80s bar.

Rename the sheet to *Number of Artists in each Genre by Cluster* and make a new sheet.

Now drag **AVG(Instrumentalness)** to rows, **CNTD(Artists)** to columns, and **Year** to the Detail property in Marks. Then, press **Swap Rows and Columns**, the button is just above the Columns shelf. You will get a scatter plot like below:

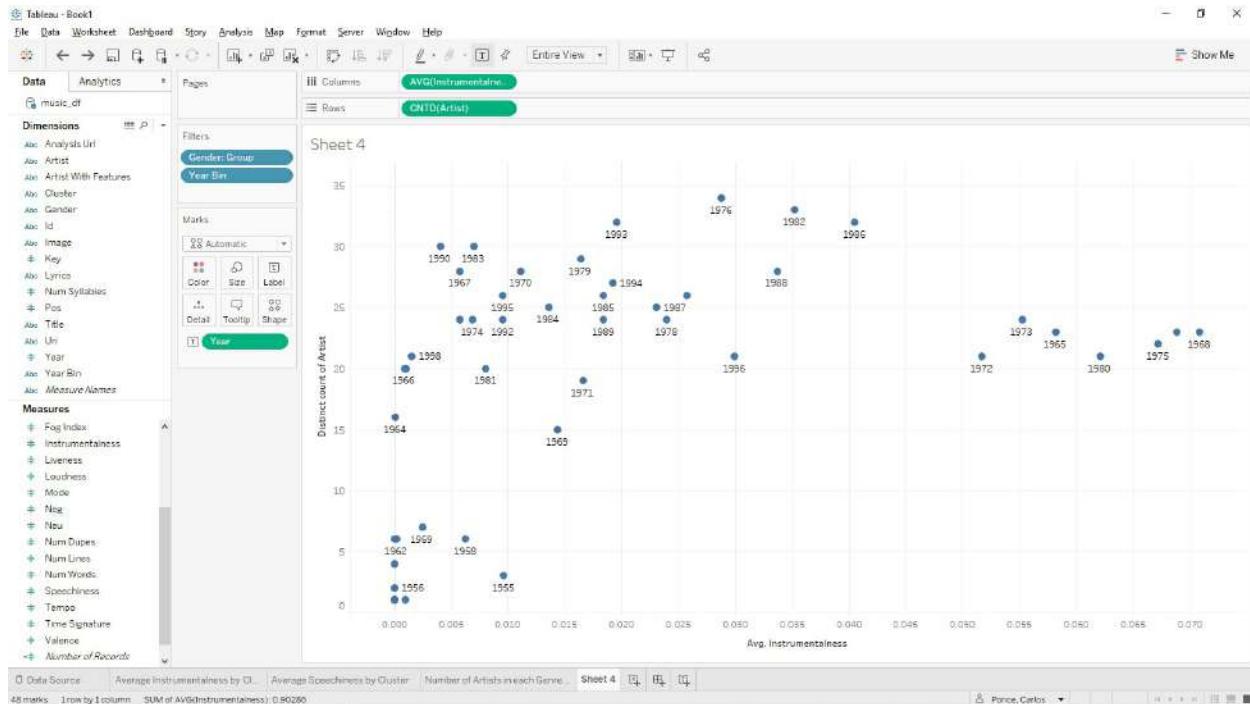


CNTD(Artist) should now be on the Y axis.

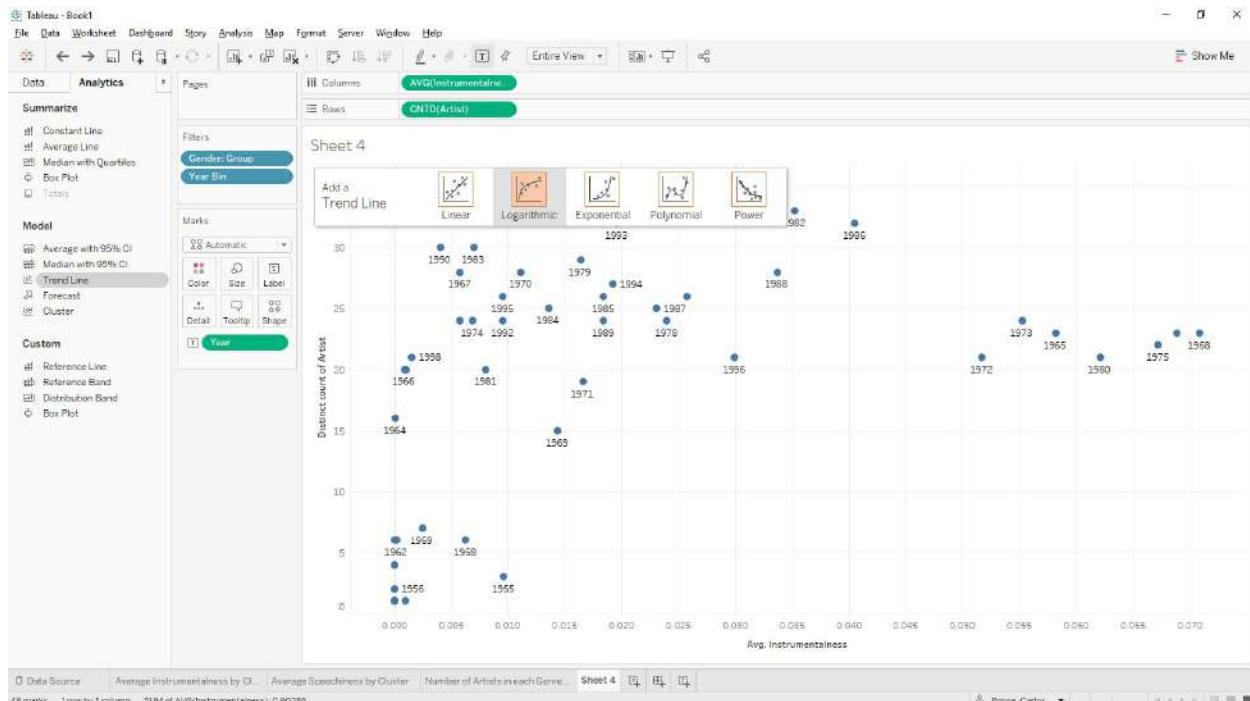
Explained:

Scatter plots are helpful to analyze correlations between two measures. You can do many analyses like finding clusters, or building a model based on the points, which can be done by Tableau in a few clicks (will show briefly an example).

Click the Shape property and select the filled circle. Now drag **Year** to Label and fit this graph to Entire View. Add two filters: **Gender = Group** and **Year Bin 50s to 90s** (we don't want to show 00s and 10s here).

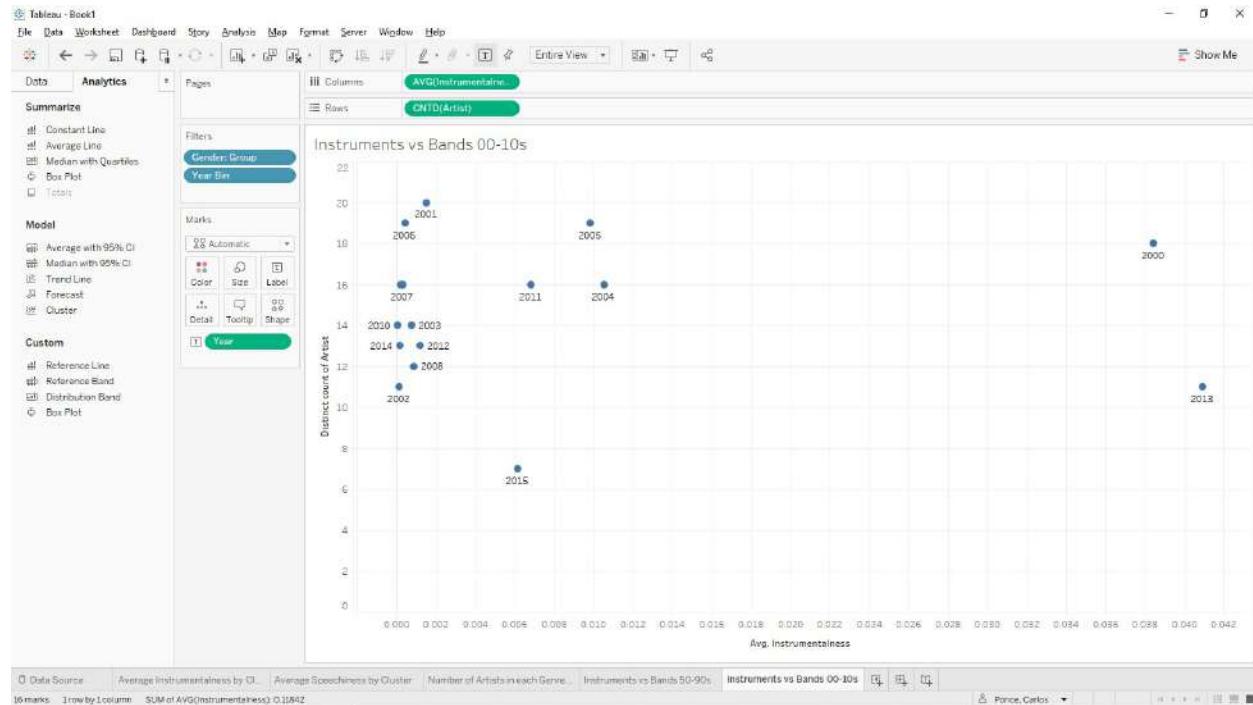


Go to the **Analytics** tab from the data pane on the left, under Model drag Trend Line to your graph, and release it on Logarithmic so we get that kind of trend line.



Name this sheet *Instruments vs Bands 50-90s*. Right-click the sheet tab, select **Duplicate**, and change the Year filter to be equal to **00s - 10s**. Also, disappear the trend line from this duplicate sheet by right clicking it and

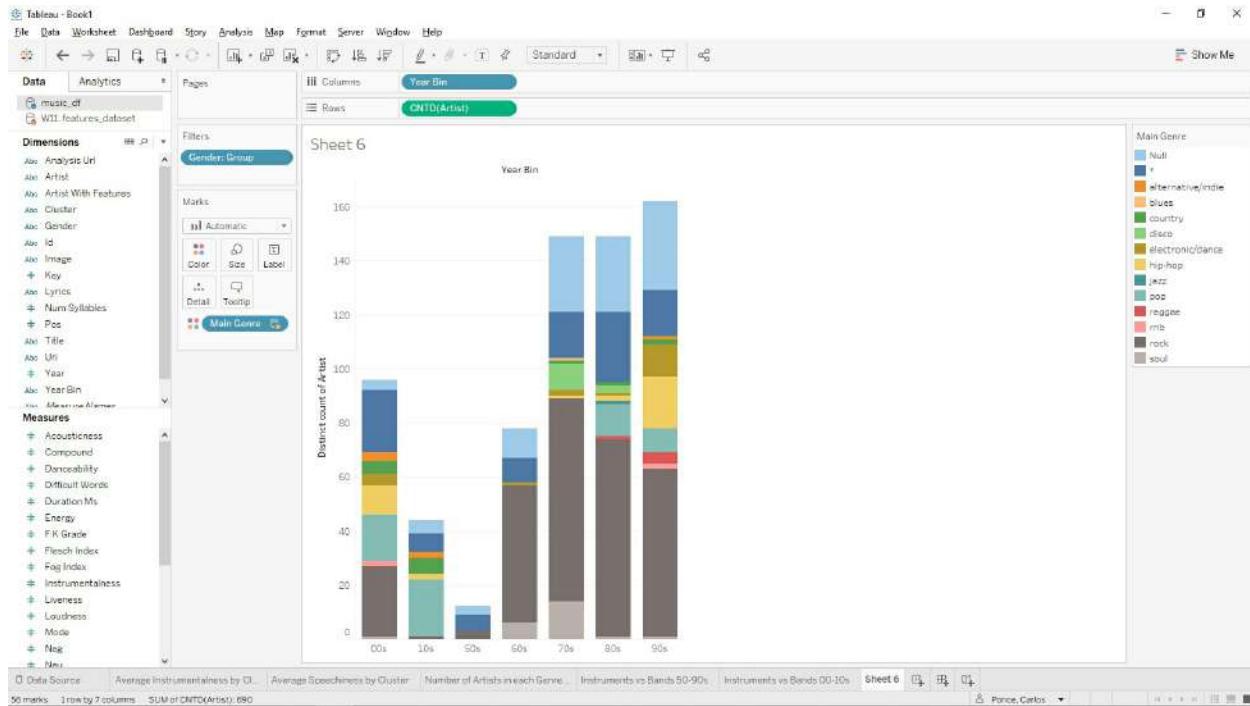
untoggling Show Trendline.



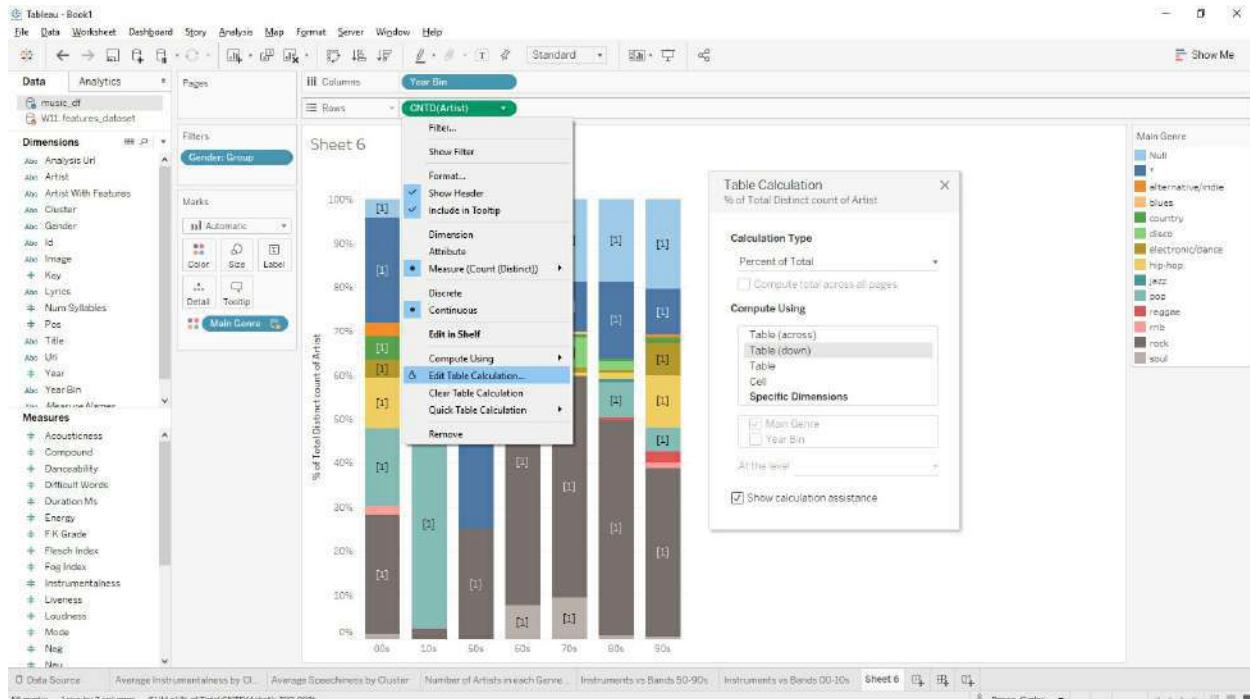
If you observe the 50s – 90s scatter plot, you can see there is a trend that on the years there were more bands, the instrumentalness was higher, and the less bands the less instrumentalness. Also, the instrumentalness was higher back when bands were popular in the 50s – 90s. Now, you see in the 00s and 10s that there are less bands, and therefore instrumentalness is also lower (mainly because we now see more singer bands like Fifth Harmony and One Direction).

For our next graphs, we'll need to add another data source. Click the **New Data Source** button to the right of the **Save** button, look for W4 Features dataset.csv. No need to change **Text File Properties** here, the null cells are normal.

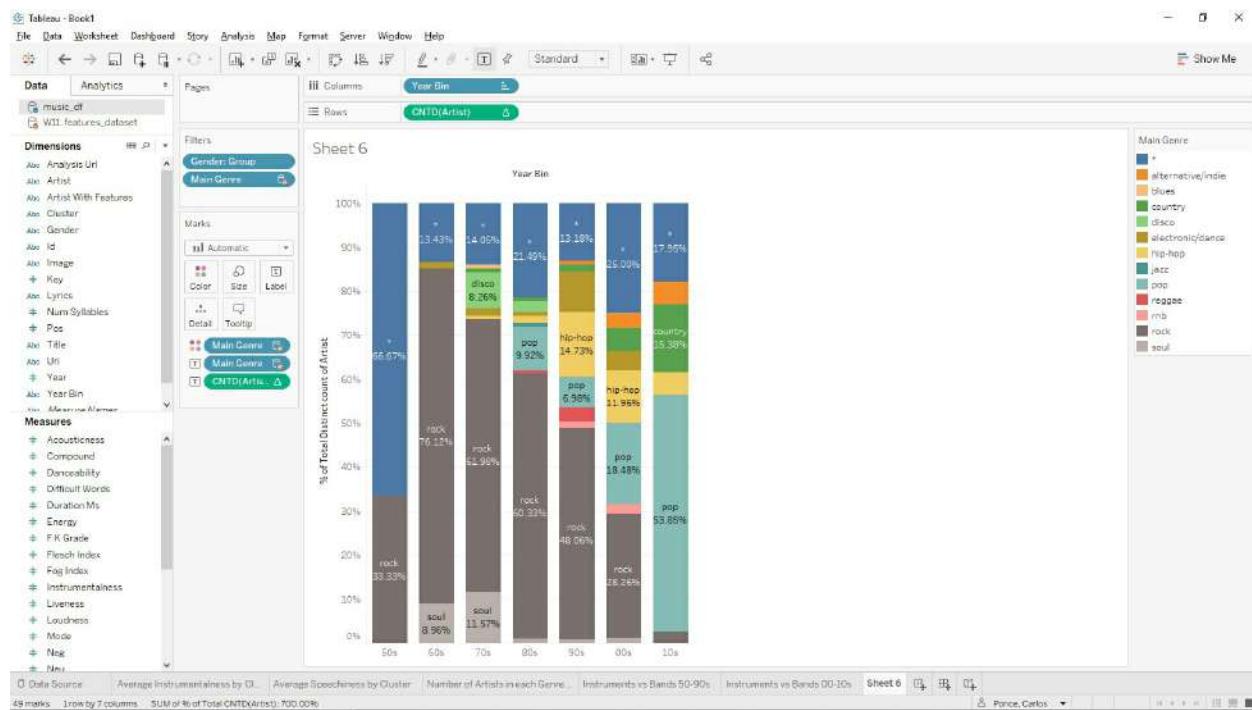
Make a new sheet, add **Year Bin** on Columns, a filter **Gender = Group**, **CNTD(Artist)** on rows, and now from the new data source take **Main Genre** and put it on the Colors property. If it shows you a warning, click Ok and toggle the chain on the **Artist** dimension. That means that your data blending is now linked by Artist on this sheet.



Again, set the **CNTD(Artist)** to be a **% of total calculation** to be **Table (down)** instead of (across).

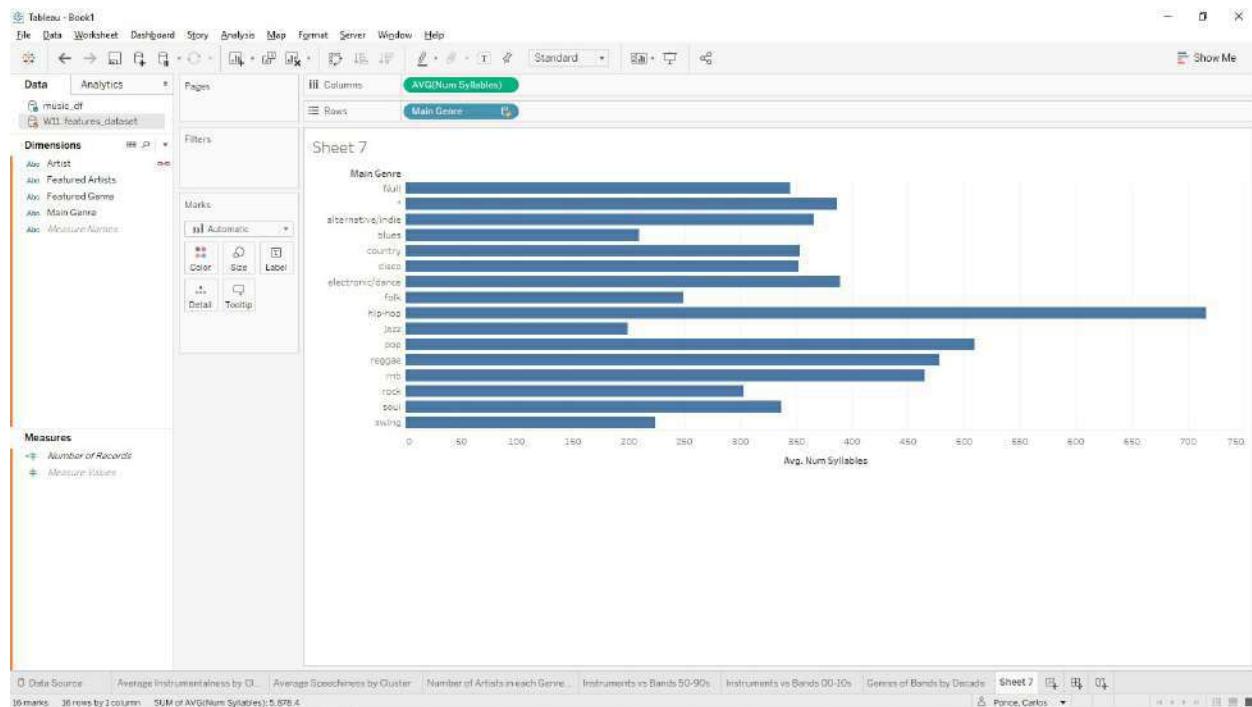


And again, sort **Year Bin** so 00s and 10s are to the right. Hold Ctrl and drag **Main Genre** and **CNTD(Artist)** to Label. Add a filter to exclude Null from **Main Genre**.



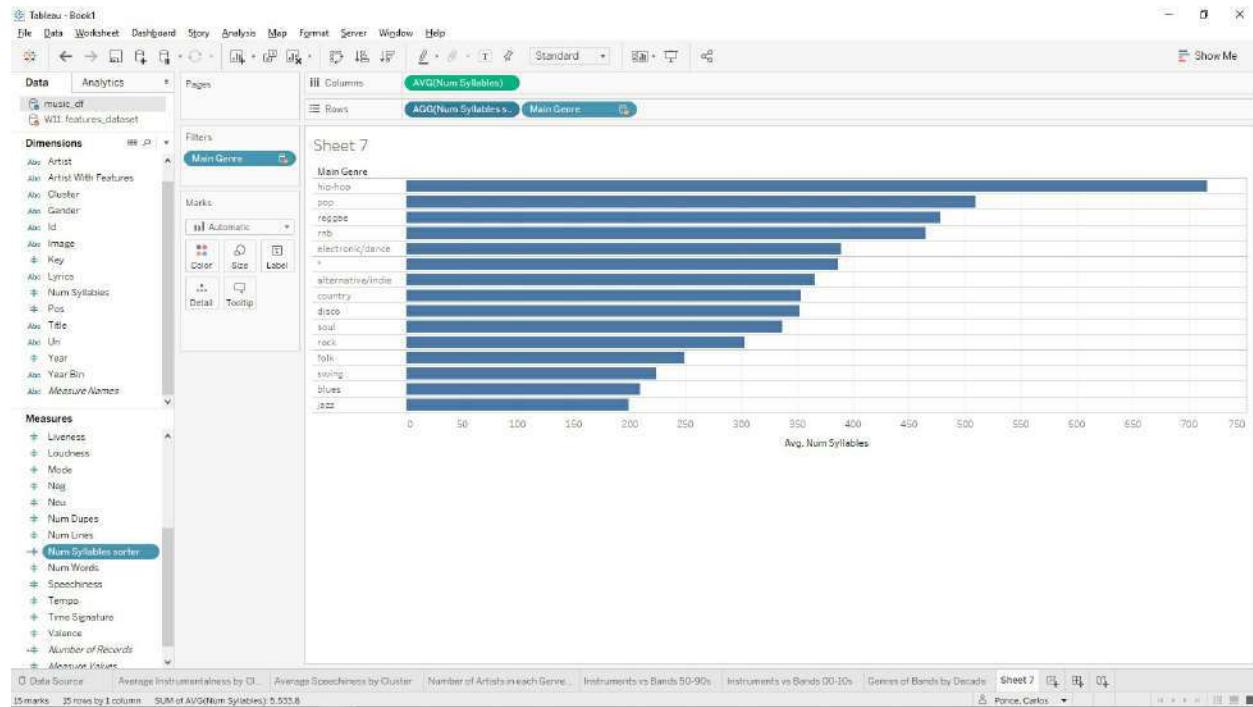
Remove the **Year Bin** header title, fit to **Entire View** and rename to *Genres of Bands by Decade*.

Add a new sheet, drag **Num Syllables** from dimensions to the Columns shelf with right click and select the continuous (green) **AVG (Num Syllables)**. Then go back to the features source (from the top-left) and put **Main Genre** in the Rows shelf.

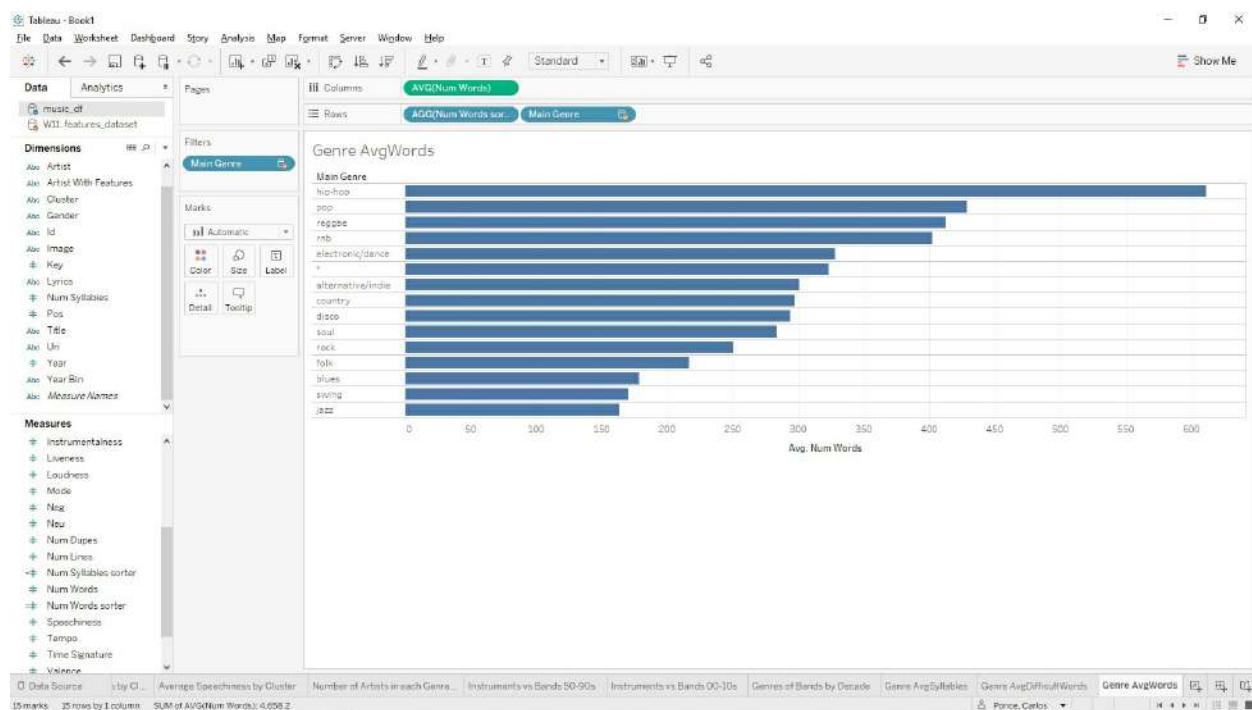
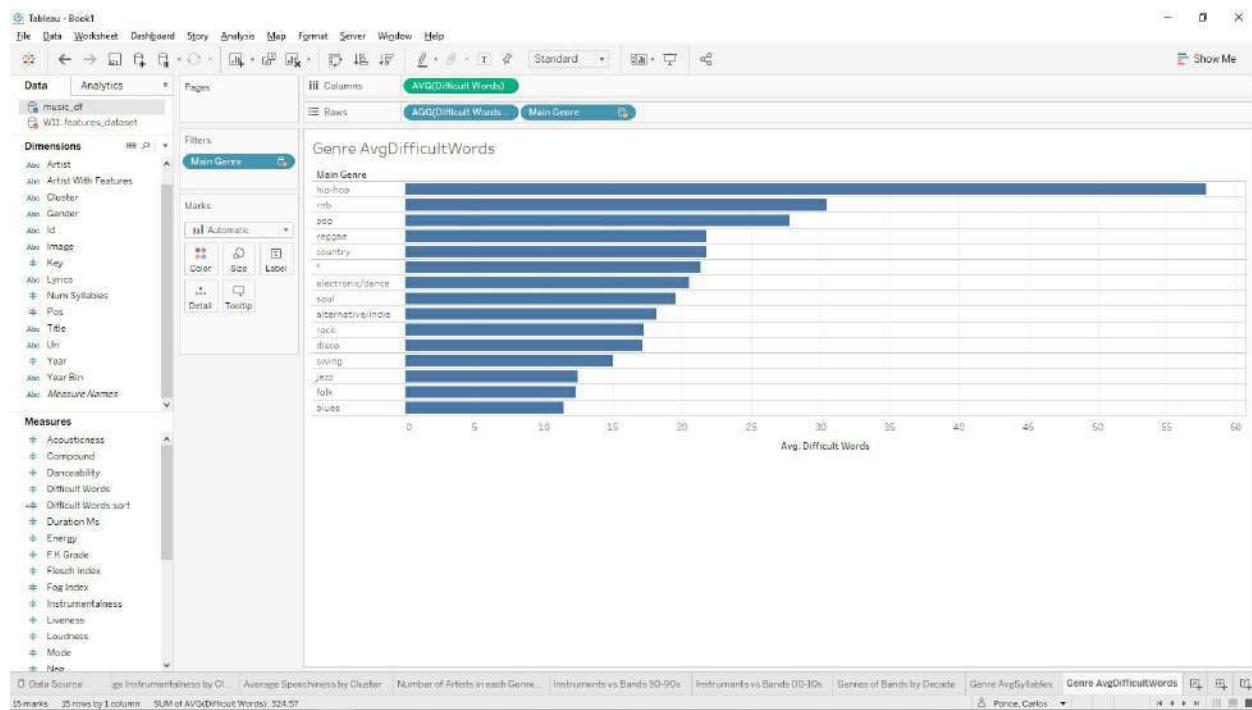


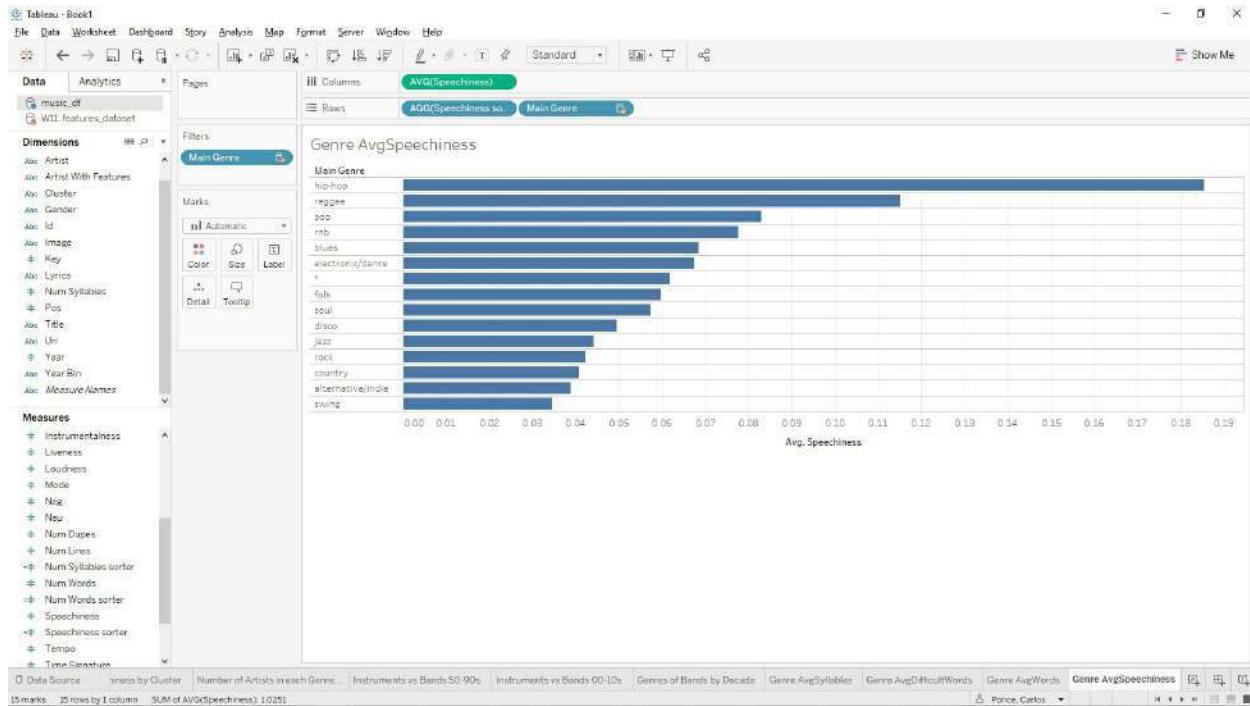
Exclude the null genre, and try to sort by descending on **AVG (Num Syllables)**. You will find out **you can't**, because you're using data blending (this is one of its limitations). If you used a data join, you would be able to sort, but you can't since the features table has repeated artists. So, in this situation you can use a trick as follows:

1. Create new calculated value called **Num Syllables Sorter**.
2. Make it equal to **-AVG ([Num Syllables])** (the negative of average syllables).
3. Convert it to discrete (blue) right-clicking it and choosing **Convert to Discrete**.
4. Drag it to the Rows shelf, right before the **Main Genre** pill.
5. Right-click **Num Syllables Sorter** and untoggle **Show header**.



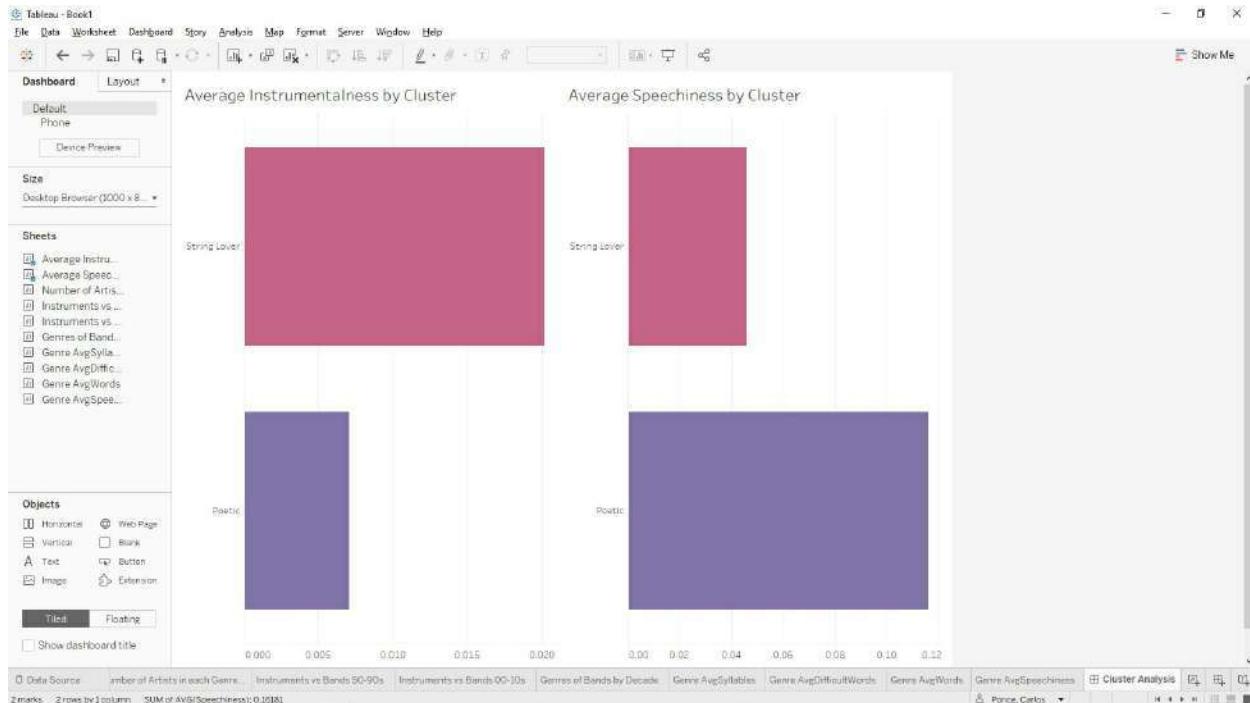
Name this sheet **Genre AvgSyllables**. Duplicate this sheet and repeat those five steps for **AVG (Difficult Words)**, **AVG(Num Words)**, and **AVG(Speechiness)**. Below the images of each of the three sheets, so you can confirm you repeated the process correctly:



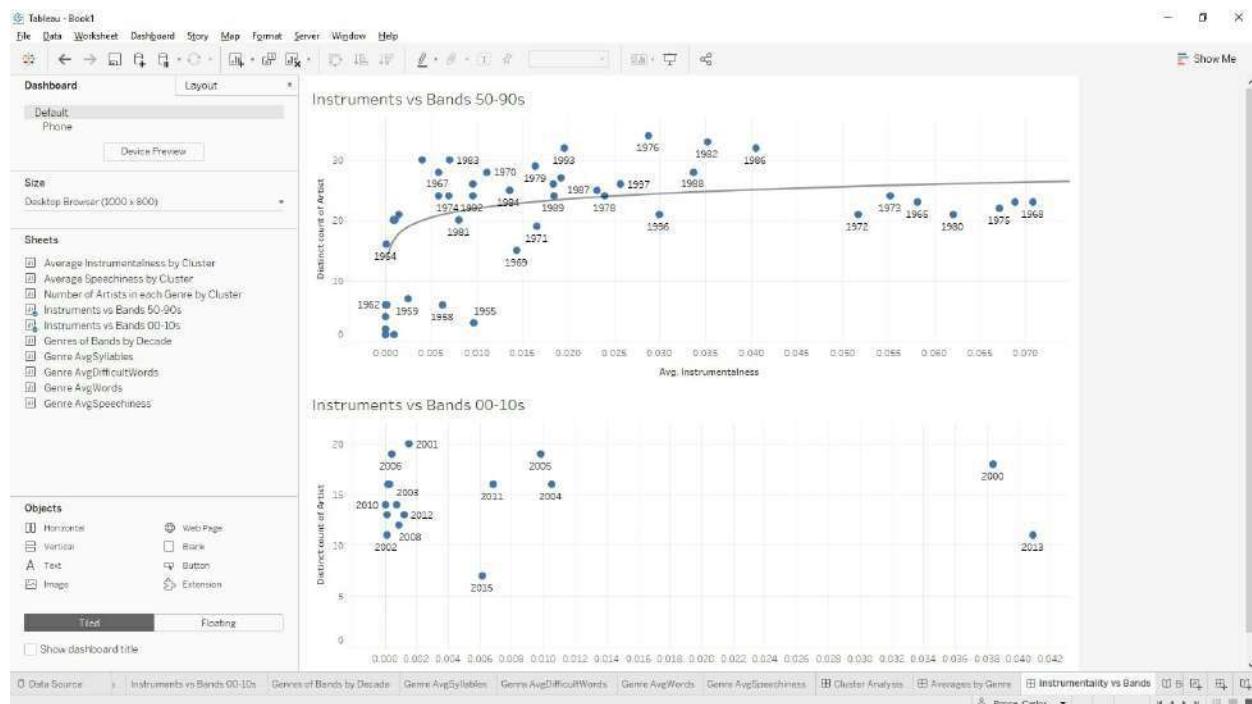


Don't forget to name the tabs as well!

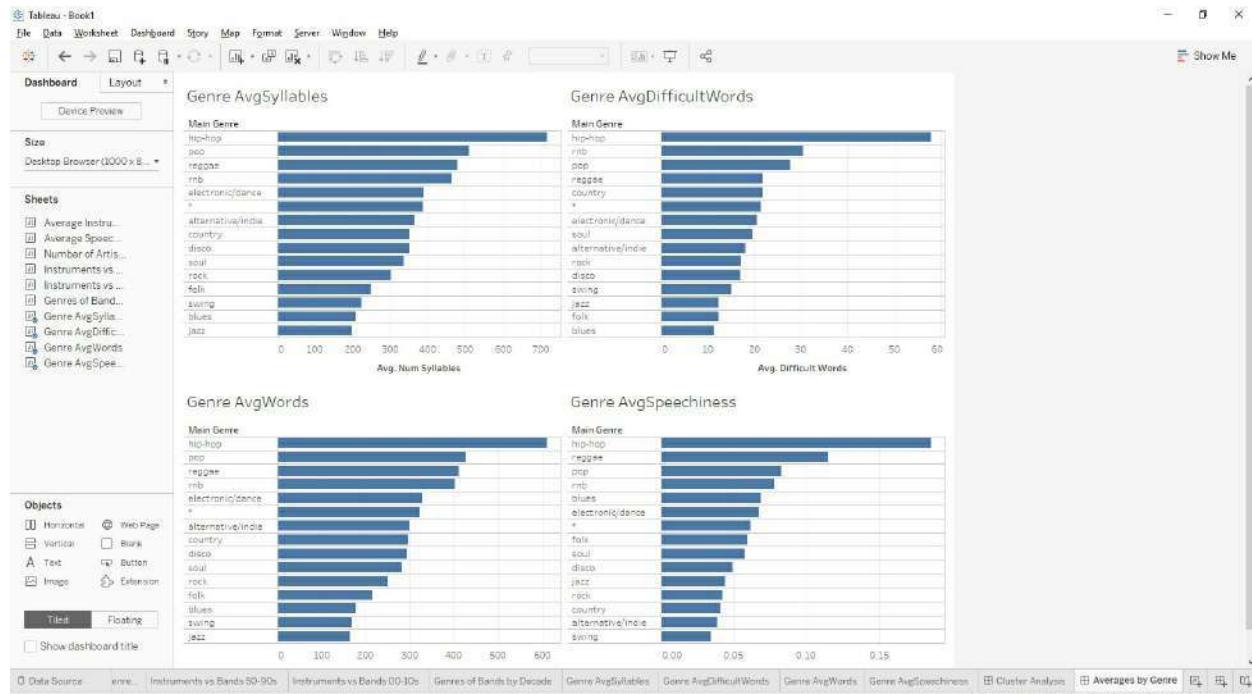
Now it's time to build some simple dashboards. Create a new one and drag *Average Instrumentalness*... and *Average Speechiness*... as shown below:



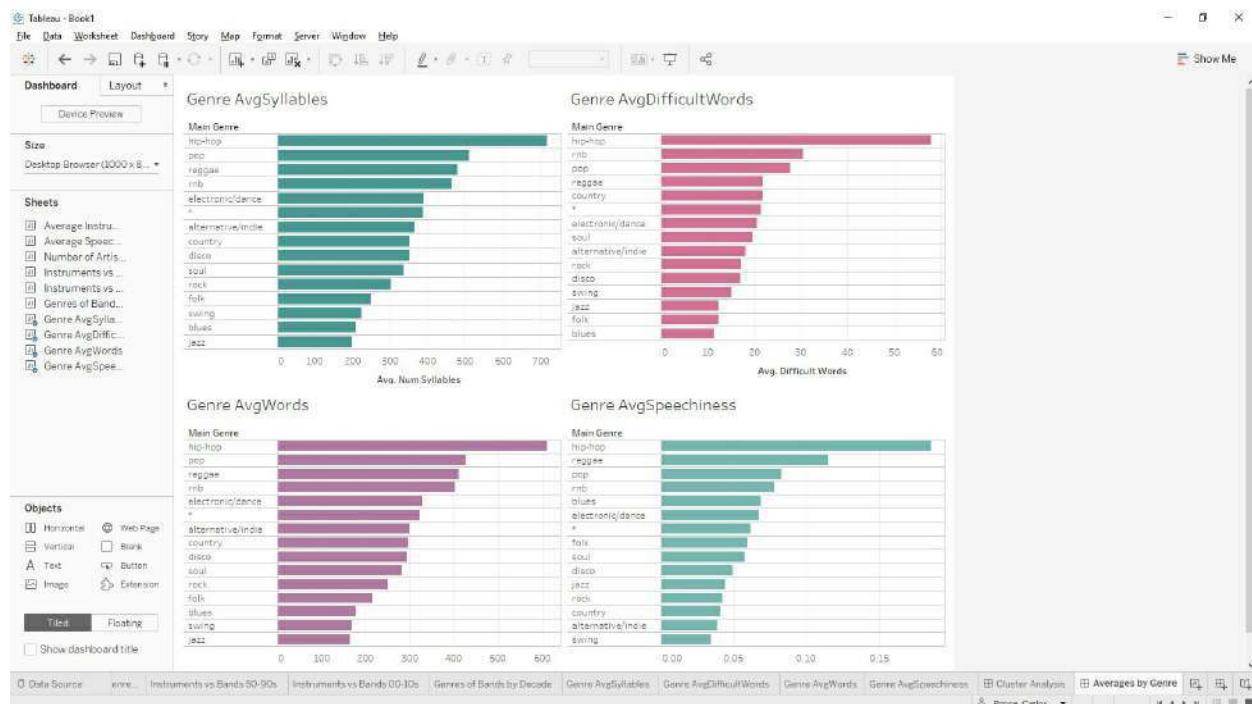
Call it *Cluster Analysis*. Now create a second dashboard and add the two *Instruments vs Bands* sheets.



Call this one *Instrumentality vs Bands*. Now for the third dashboard add the four *Genre Avg...* sheets, as shown here (use the vertical/horizontal containers as you wish):

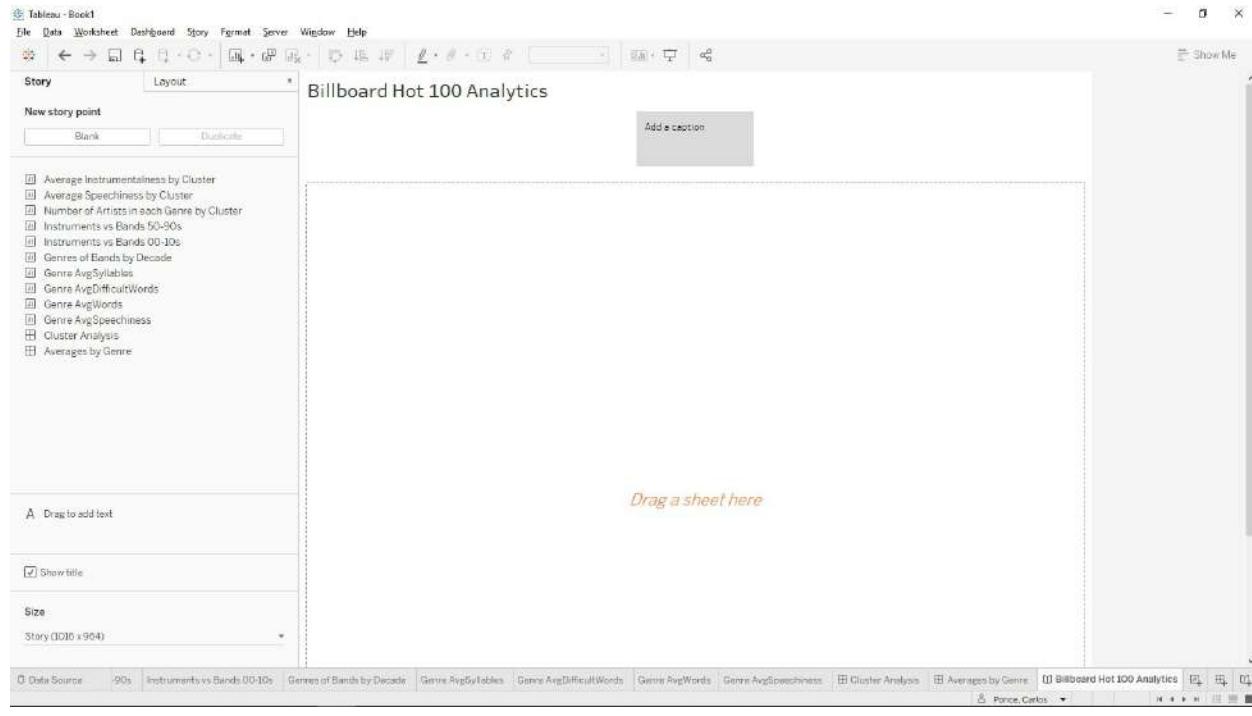


Set the four objects to fit Entire View right from the dashboard. Also, add some color to each so it looks nicer, as below:



Name that Averages by Genre.

Time for the last tab (finally). Click the icon next to the New Dashboard button; this is the **New Story** button. Call your story *Billboard Hot 100 Analytics*:



Explained

If you go back to the first workout, right before we made our first dashboard, we touched the topic of stories. Here is the line:

Story: another form of interaction where you can attach worksheets and dashboards, and you can add annotations to them to bring more insights to your analysis.

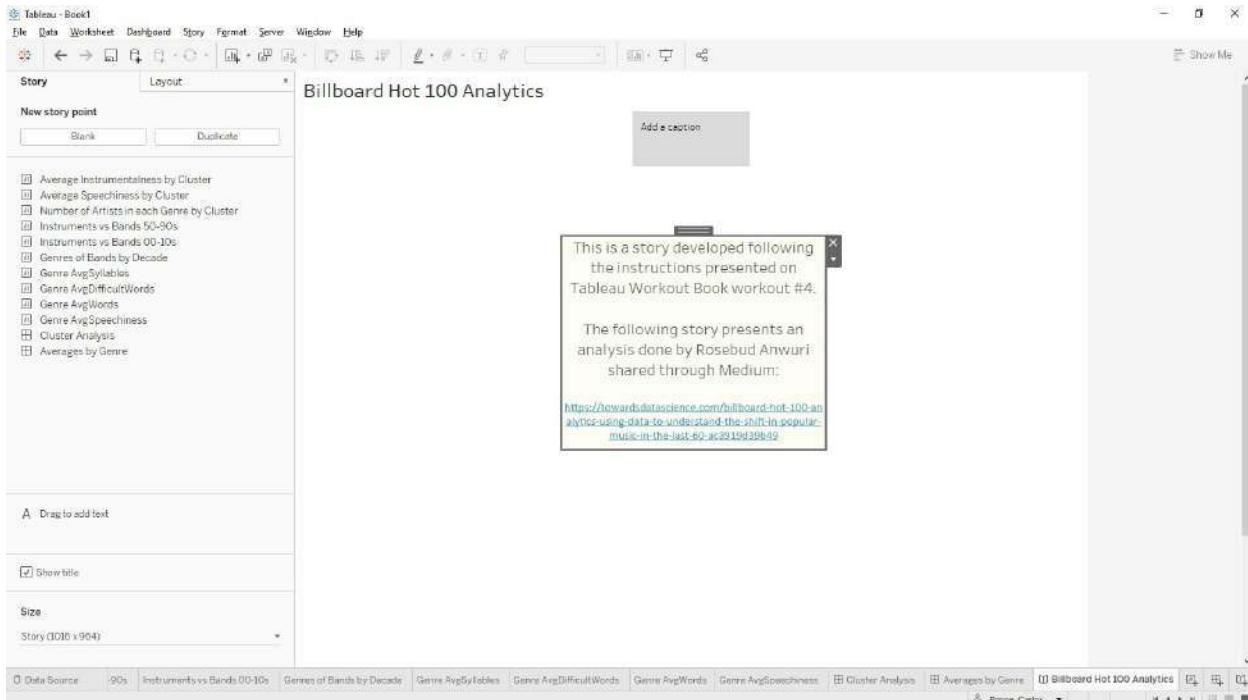
Stories are not commonly used in the workplace, specially when you talk of reporting and not data exploration or analysis (also because people are more familiar with dashboards). But it is a powerful way to explain a data discovery, or to go through an insightful analysis, because you can explain in a detailed manner and present in a linear way; you can write the story you want to tell from your data.

Take the **Drag to add text** icon on the bottom-left to the story canvas and paste this text:

This is a story developed following the instructions presented on Tableau
Workout Book workout #4.

The following story presents an analysis done by Rosebud Anwuri shared through Medium:

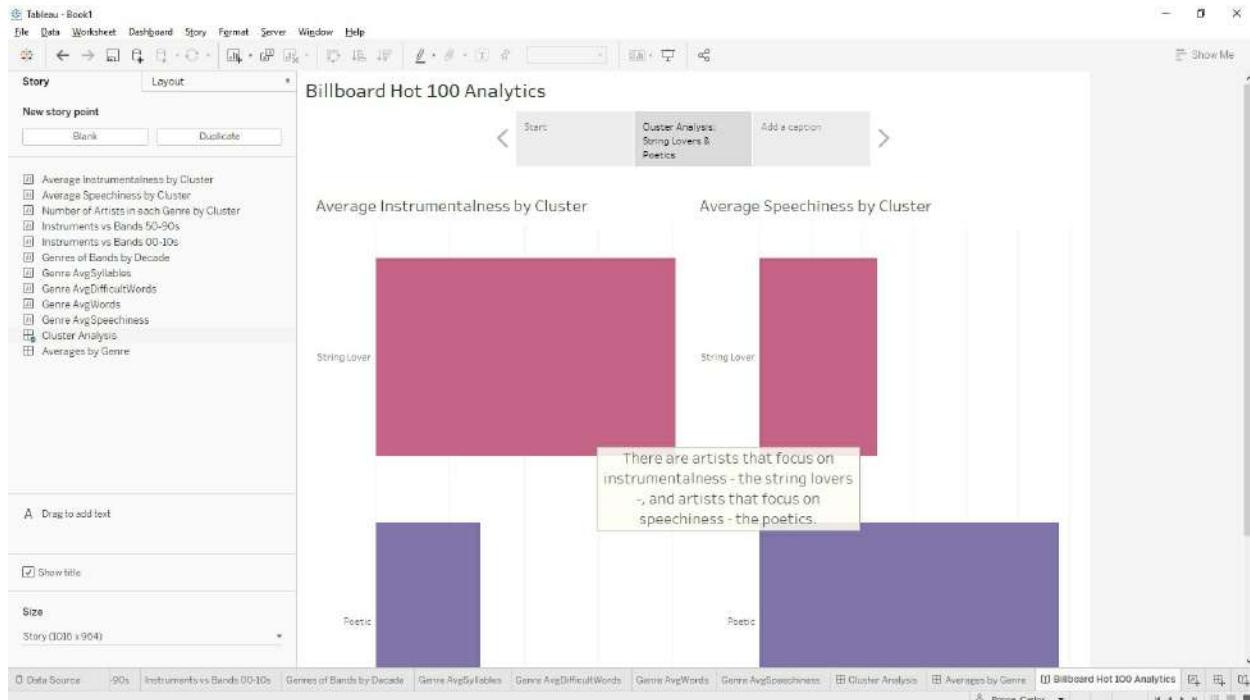
<https://towardsdatascience.com/billboard-hot-100-analytics-using-data-to-understand-the-shift-in-popular-music-in-the-last-60-ac3919d39b49>



Click on the **Add a caption** gray button at the top, and type “Start”.

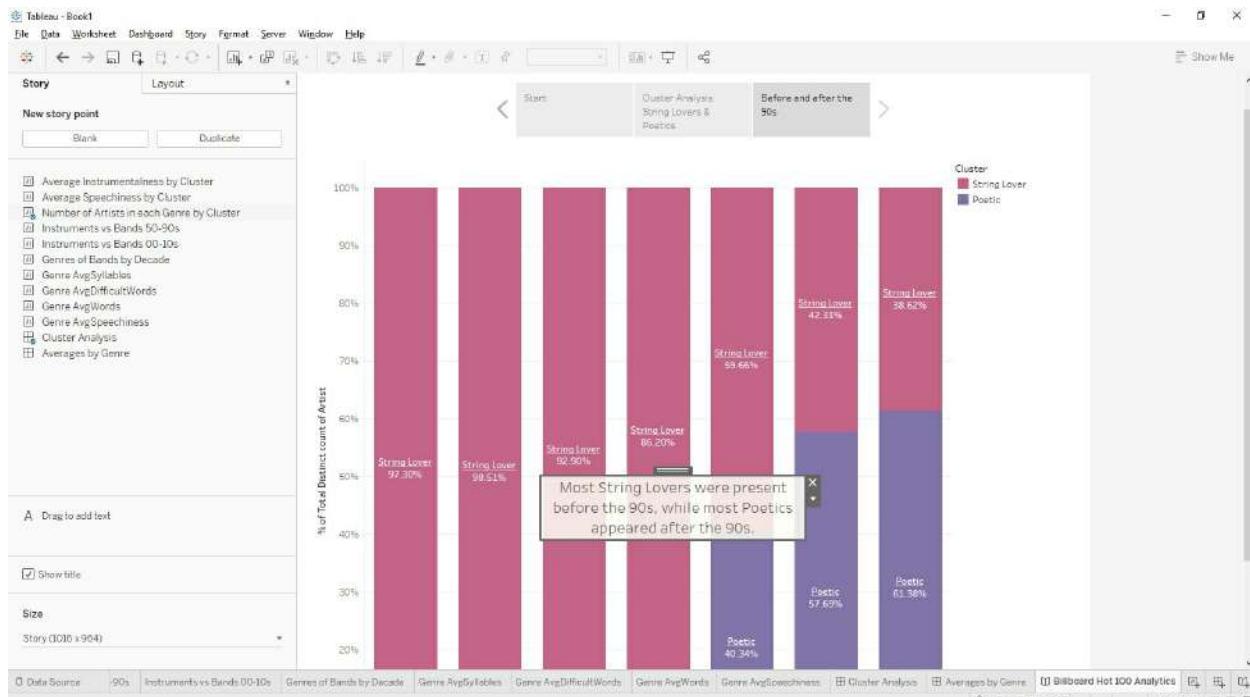
Click the **Blank** button on the top left to make a new tab on your story and drag your **Cluster Analysis** dashboard. Then drag another text to the center and type: “There are artists that focus on instrumentalness - the “String Lovers” -, and artists that focus on speechiness - the “Poetics”.”

Name your caption “Cluster Analysis: String Lovers & Poetics”.



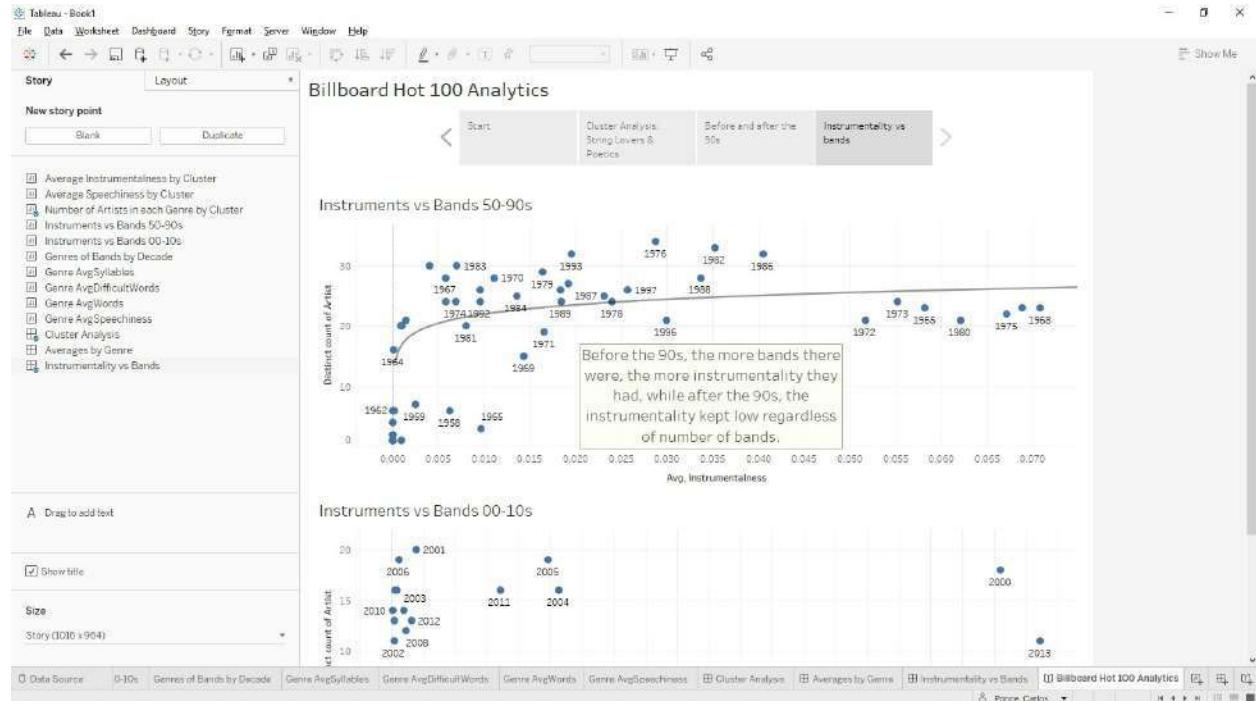
Again, click **Blank**, drag **Number of Artists in each Gender by Cluster**, a text and type: “Most String Lovers were present before the 90s, while most Poetics appeared after the 90s”.

On your caption add: “Before and after the 90s”.

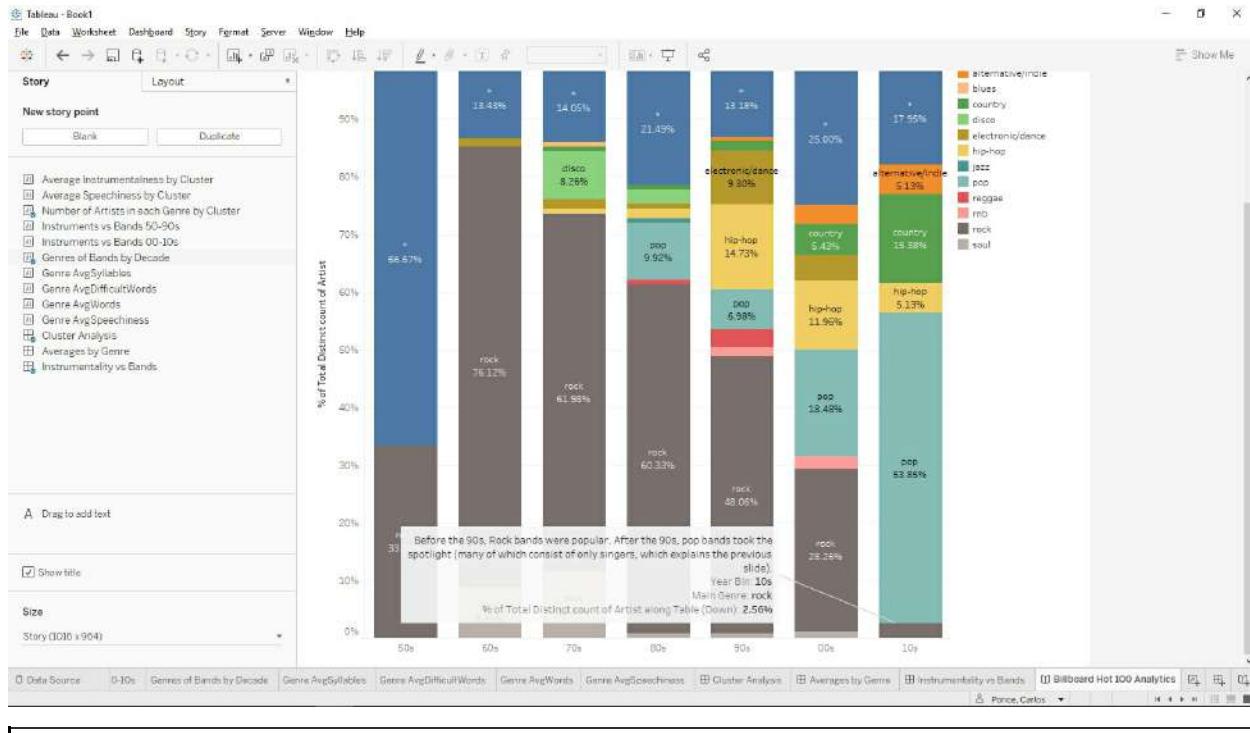


Another **Blank**, drag **Instrumentality vs bands dashboard**, and add text:

“Before the 90s, the more bands there were, the more instrumentality they had, while after the 90s, the instrumentality kept low regardless of number of bands”. Caption this “Instrumentality vs bands”.



Add the fifth **Blank** tab, drag **Genres of bands by decade**, and instead of adding the text element, right-click the rock portion of the 10s bar, select Annotate > **Mark**, and type above the pre-added text: *“Before the 90s, Rock bands were popular. After the 90s, pop bands took the spotlight (many of which consist of only singers, which explains the previous slide)”*. Caption to “Genres of bands by decade”.



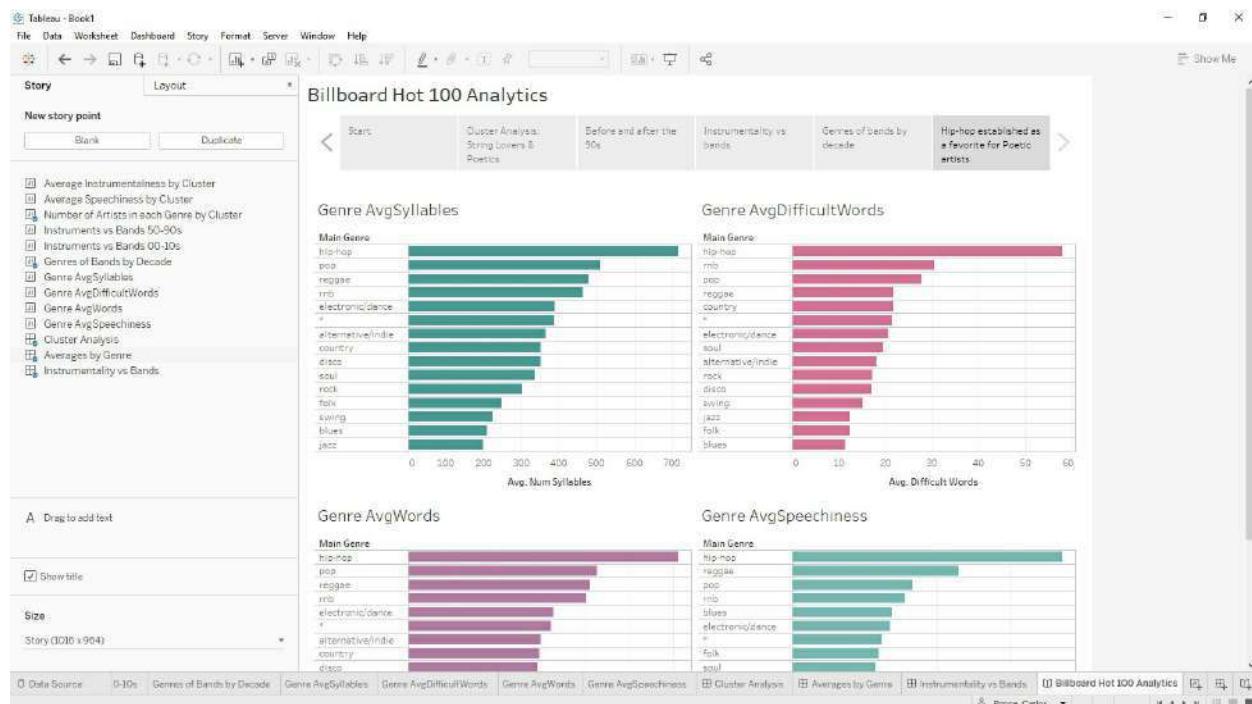
Explained:

There are three types of annotations in stories:

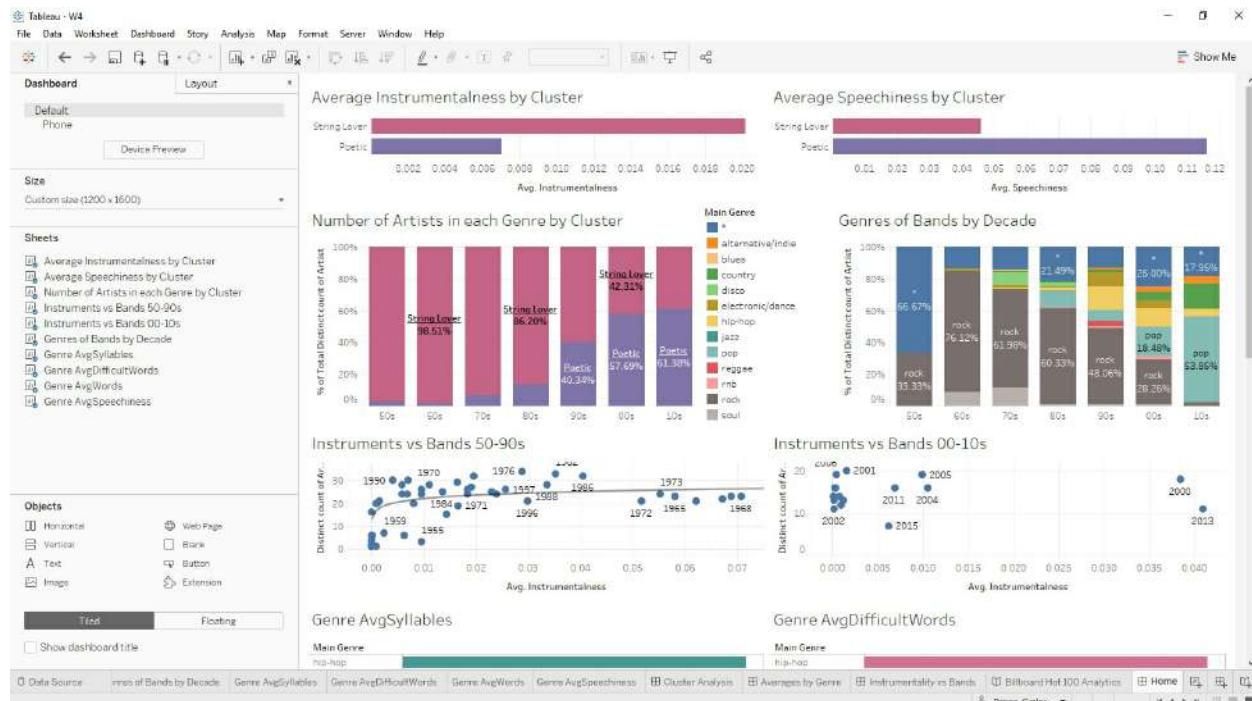
- Mark: you annotate a specific point in your graph, marked by an arrow, and you fix to it.
- Point: you annotate a specific place in your graph, marked by an arrow, you can move it anywhere.
- Area: you annotate a specific area in your graph, marked by a box.

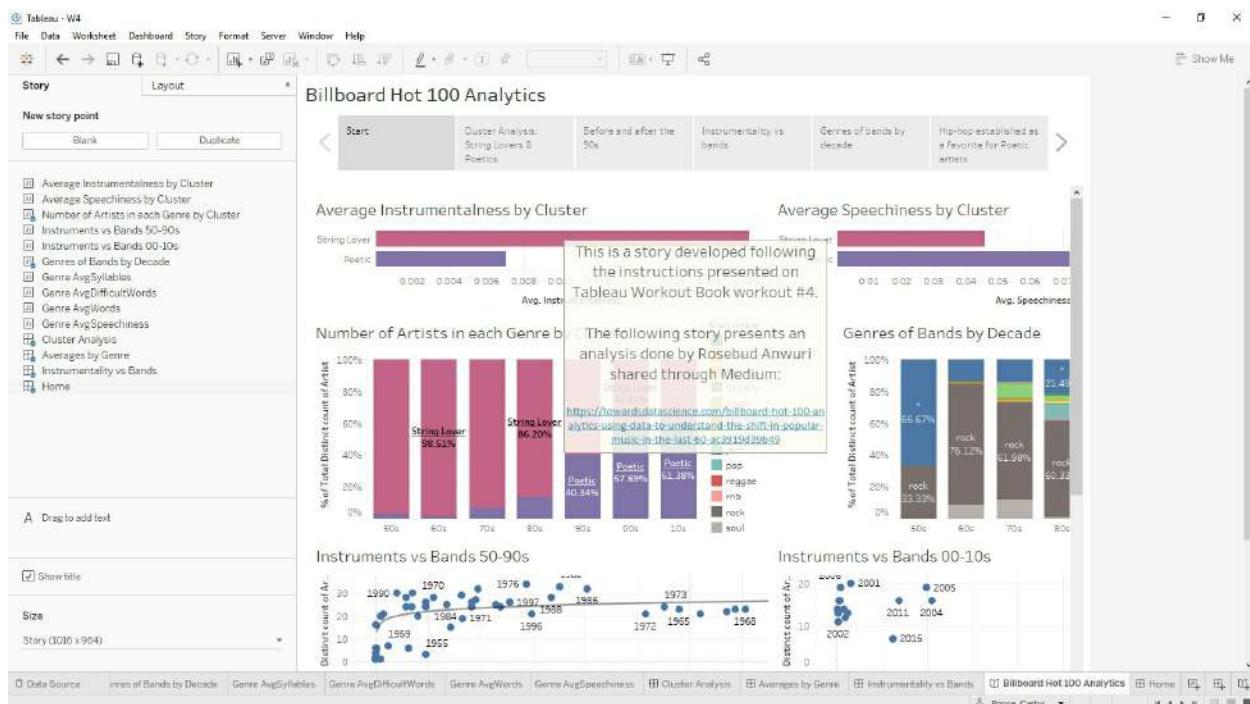
You can also format the annotations by right-clicking them.

Finally, make a slide with the **Averages by Genre dashboard** and set as caption “Hip-hop established as a favorite for Poetic artists”.



To give it a nicer portrait, make a dashboard called **Home** with all the sheets on it, and put it on the first story slide.





There you go, fourth dashboard finished!

By now, you should now have all the fundamentals covered.

You can develop **sheets**, **dashboards**, and **stories**, the three elements of a Tableau workbook. You have the basics of **formatting and design** needed to make appealing and responsive dashboards. And you know how to **connect properly to data**, validate, adjust settings and determine if join or blend should be used.

Here's what you learned on workout 4:

- Hiding unnecessary field labels
- Table calculations: compute logic
- Scatter plots
- Analytics tab: trendlines
- How to sort when using data blending
- Stories and annotations

The topics learned will be less as we go through the workouts, and for more specific scenarios (like the case of sorting with data blending). Again, be sure to recap the topics you didn't fully grasped, because we will increase the pace on the next exercises.

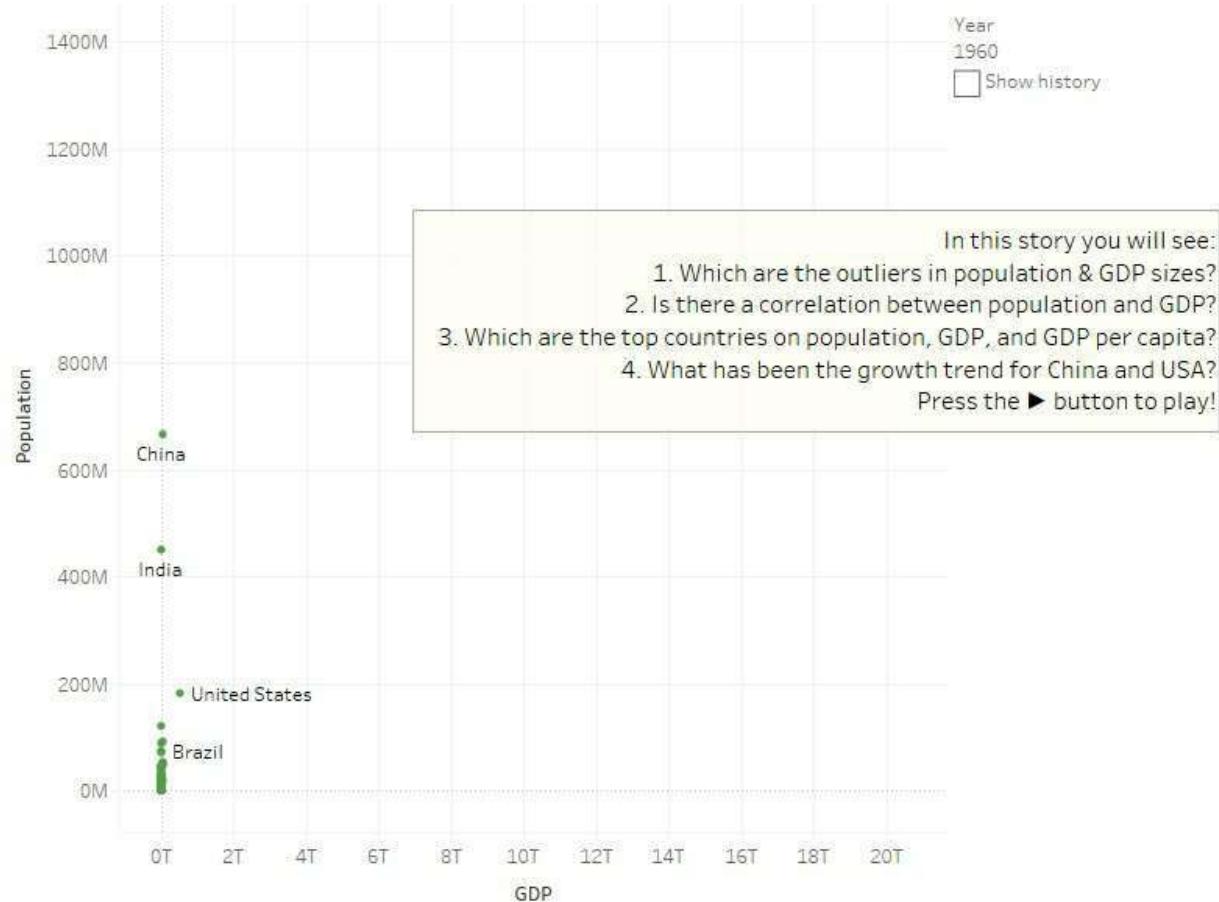
Workout 5. World population dashboard

On this workout you will make a very interesting story – like the Billboard Hot 100 workout –, about the history of world population, and analyze the correlation of the population with the gross domestic product (GDP) by country.

Data	■ ■ ■ □ □
Design	■ ■ □ □ □
Graphs	■ ■ ■ □ □
Interactivity	■ ■ ■ ■ ■

World GDP & Population Analysis

1 2 3 4 5 6



Workout Start

Remember step one before starting a dashboard: know your data. Open the three W5 csv files and observe the contents.

W5 Population by country - Excel																														
Ponce, Carlos (GE Renewable Energy)																														
File Home Insert Page Layout Formulas Data Review View Developer Smart View Help DV Desktop Search Share Comments																														
A1																														
Data Source																														
1	Data Source	World Development Indicators																												
2	Last Update	=====																												
3	Country	N	Country	C	Indicator	I	Indicator	I	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979		
4	Aruba	ABW	Population	SP.POP.TK	54211	55439	56225	56695	57032	57360	57715	58035	58386	58726	59063	59440	59840	60243	60528	60867	60586	60366	60103	59980						
5	Afghanistan	AFG	Population	SP.POP.TK	916973	9169410	931441	9345203	9744781	9583203	10174838	10389290	1063708	10893770	1117624	1147345	11791215	121083615	12689100	12943058	13171308	13411188	13411050	13						
6	Angola	AGO	Population	SP.POP.TK	5454933	5531472	5608359	5679454	5730044	5770570	5781214	5774243	5771652	5803254	5899369	6049777	6248552	6496952	6761380	7024000	7279509	7533735	7790707	8058067	E					
7	Albania	ALB	Population	SP.POP.TK	1608800	1655860	1711319	1762621	1814135	186473	1915473	1965593	2022272	2081695	2135479	2187835	2243126	2296752	2359124	2404831	2458525	2513546	2562626	2617832	Z					
8	Andorra	AND	Population	SP.POP.TK	13411	14175	15370	18412	17469	18549	19647	20758	21850	23058	24276	25559	28892	28232	29520	30705	31777	32771	33737	34818						
9	Arab World	ARB	Population	SP.POP.TK	9219733	9424510	9733442	1016408	1056108	1095108	1136108	1156108	1179000	1200900	1230000	12507000	12997000	13177000	13380000	13723000	13893000	14033000	14514000	14538000	14514000	14				
10	United Arab	ARE	Population	SP.POP.TK	92418	100796	112128	125130	138039	149857	159876	169771	182627	203106	2454124	277471	30974	3584624	467451	548301	67922	735344	885508		981749	1				
11	Argentina	ARG	Population	SP.POP.TK	2048179	2081266	21153625	21488912	2184425	22159625	22494053	2282809	23108267	23570111	23880101	24259301	2403175	2505478	2540232	2557077	26204681	26601398	27001047	27471630	Z					
12	Armenia	ARM	Population	SP.POP.TK	1874122	1941492	209526	207758	2346001	2211319	2276634	2391227	2401143	2462528	2525068	2587708	2659486	2717786	273748	2832759	2889579	2944383	2997419	3049109	E					
13	American	ASM	Population	SP.POP.TK	20123	20602	21233	22034	23672	25428	25989	26703	27363	27954	28567	29100	29596	30052	30456	31269	31845									
14	Antigua and	ATG	Population	SP.POP.TK	54131	55001	55841	56702	57641	58698	59915	61241	62521	63550	64177	64357	64133	65160	66275	672159	68239	69245	702545	712652	71952					
15	Australia	AUS	Population	SP.POP.TK	10278477	10481000	10742600	10950000	11167000	11388000	11651000	1179000	12009000	1230000	12507000	12997000	13177000	13380000	13723000	13893000	14033000	14514000	14538000	14514000	14					
16	Austria	AUT	Population	SP.POP.TK	7086239	7129804	7175813	7223801	7270884	7322006	7370894	7415048	7461055	7467058	7509482	7544201	7586115	7598080	7635535	7678908	7705535	7758430	7802305	7849425						
17	Azerbaijan	AZE	Population	SP.POP.TK	3853397	4030322	4171426	4315127	4455687	4592607	4721523	4843868	4960232	5071927	5180203	5284531	5382631	5483082	5579072	5647133	5768721	5883132	5957925	6035640	E					
18	Burundi	BDI	Population	SP.POP.TK	2970352	2964427	3026290	3094374	3170490	3253218	3336927	3413904	3479704	3579074	3670880	3770880	3777087	3854445	3942955	4051234	412124									
19	Belgium	BEL	Population	SP.POP.TK	9133499	9183949	9220578	9289779	9378113	9453667	9527807	9580991	9618756	9648032	9655540	9673162	9711115	9741720	9772419	9800700	9818227	9830358	9839534	984382	S					
20	Benin	BEN	Population	SP.POP.TK	2431622	2465867	2502806	2542859	2585965	2623628	2682159	2733037	2791500	2850681	291240	2978572	3043567	3113675	3187412	3265126	3326155	3347173	3433449	3523938	3618526					
21	Burkina Faso	BFU	Population	SP.POP.TK	4829288	4894580	4960326	5027821	5098890	5147807	5256363	5340309	5434041	5528174	5604600	5723381	5825173	5930483	6040041	6154545	6274037	6358935	6530819	6671656						
22	Bangladesh	BGD	Population	SP.POP.TK	4803350	4862848	5075215	5202067	5374170	53838112	57157054	58084289	60818458	6079705	6423342	6531638	6623974	6787452	6873923	6980235	70874238	70080810	7102381	74863584	75450012	7723945	75			
23	Bulgaria	BGR	Population	SP.POP.TK	7867314	7943118	8012946	8144340	8204168	8250657	8310226	836963	8454174	8489574	8536395	8576200	8620997	8678745	8720742	875859	8804183	8814032	8823940							
24	Bahrain	BHR	Population	SP.POP.TK	162427	167854	173144	178184	1826887	1871431	191790	195603	200653	206043	212605	219155	225957	231912	235957	241911	247074	253175	258175	263474	268752					
25	Bahamas	BHS	Population	SP.POP.TK	109534	115122	121191	127330	133709	140059	146182	152610	158648	164288	164288	169777	173849	177866	181517	185120	188998	192959	197101	201478						
26	Bosnia and	BH	Population	SP.POP.TK	3325668	3286603	335326	3417534	3478997	3535643	3566636	3632672	3675454	3717468	3760529	3803287	3851530	3897256	3942225	3985105	4025262	40631393	4103049	4138619	4					
27	Bolivia	BOL	Population	SP.POP.TK	9180000	8271108	8215182	8487238	8524242	8610000	88904980	8785648	8874552	89004004	9040004	9113576	9188908	927272	9411000	9467000	9481000	9483800	9520000	9584000						
28	Belize	BZL	Population	SP.POP.TK	106194	9734	103064	106199	110974	112947	116051	121692	124732	127150	130307	133260	135147	138899	141305	143064	145063	147062	150367	153607	161610	161812	164105			
29	Bolivia	BOL	Population	SP.POP.TK	44400	45500	46600	47700	48800	50100	51000	52000	53000	54000	55000	56000	57000	58000	59000	59300	59600	59800	59800	59800	59800	59800	59800	59800		
30	Bolivia	BOL	Population	SP.POP.TK	3225000	3261000	3301000	3341000	3381000	3421000	3461000	3501000	3541000	3581000	3621000	3661000	3701000	3741000	3781000	3821000	3861000	3901000	3941000	3981000	4021000	4061000	4101000			
31	Bolivia	BOL	Population	SP.POP.TK	3261000	3301000	3341000	3381000	3421000	3461000	3501000	3541000	3581000	3621000	3661000	3701000	3741000	3781000	3821000	3861000	3901000	3941000	3981000	4021000	4061000	4101000	4142100			
32	Bolivia	BOL	Population	SP.POP.TK	3.74E+08	4.02E+08	4.45E+08	4.72E+08	5.39E+08	6.04E+08	6.69E+08	7.56E+08	8.58E+08	9.33E+08	1.02E+09	1.1E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	
33	Bolivia	BOL	Population	SP.POP.TK	3.74E+08	4.02E+08	4.45E+08	4.72E+08	5.39E+08	6.04E+08	6.69E+08	7.56E+08	8.58E+08	9.33E+08	1.02E+09	1.1E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	
34	Bolivia	BOL	Population	SP.POP.TK	3.74E+08	4.02E+08	4.45E+08	4.72E+08	5.39E+08	6.04E+08	6.69E+08	7.56E+08	8.58E+08	9.33E+08	1.02E+09	1.1E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	1.26E+09	
35	Bolivia	BOL	Population	SP.POP.TK	3.74E+08	4.02E+08	4.45E+08	4.72E+08	5.39E+08	6.04E+08	6.69E+08	7.56E+08	8.58E+08	9.33E+08	1.02E+09	1.1E+09	1.26E+09	1.												

WS Country details - Excel

Country	Region	Income Group	Special Income Group	Table Name
ABW	Latin Amer High income	Aruba		
AFG	South Asia Low income	Afghanistan		
AGO	Sub-Saharan Lower middle income	Angola		
ALB	Europe & Upper middle income	Albania		
AND	Europe & High income	Andorra		
ARB		Arab World Arab World		
ARE	Middle East High income	United Arab Emirates		
ARG	Latin Amer Upper middle income	Argentina		
ARM	Europe & Upper middle income	Armenia		
ASM	East Asia & Upper middle income	American Samoa		
ATG	Latin Amer High income	Antigua and Barbuda		
AUS	East Asia & High income	Australia		
AUT	Europe & High income	Austria		
AZE	Europe & Upper middle income	Azerbaijan		
BDI	Sub-Saharan Low income	Burundi		
BEL	Europe & High income	Belgium		
BEN	Sub-Saharan Low income	Benin		
BFA	Sub-Saharan Low income	Burkina Faso		
BGD	South Asia Lower middle income	Bangladesh		
BGR	Europe & Upper middle income	Bulgaria		
BHR	Middle East High income	Bahrain		
BHS	Latin Amer High income	Bahamas, The		
BHK	Europe & Upper middle income	Bosnia and Herzegovina		
BLR	Europe & Upper mid Data belo Belarus			
BLZ	Latin Amer Upper middle income	Belize		
BMU	North Am High income	Bermuda		
BOL	Latin Amer Lower middle income	Bolivia		
BRA	Latin Amer Upper middle income	Brazil		
BRB	Latin Amer High income	Barbados		
BRN	East Asia & High income	Brunei Darussalam		
BTN	South Asia Lower middle income	Bhutan		
BWA	Sub-Saharan Upper middle income	Botswana		
CAF	Sub-Saharan Low income	Central African Republic		
CAN	North Am High income	Canada		
CEB		Central Eu Central Europe and the Baltics		

WS Country details

On granularity, it seems all tables are by the same level: country (264 countries to be exact, one missing on Country Details). On Country details we don't find the Country Name column, but we do find the **Country Code**.

You can see that the population and GDP tables have blank rows on the top, also that there are some null/empty cells on GDP for many countries. But there is something more important: notice how both the GDP and population tables show their numbers by year, but the year instead of being in a column, **it's the header for the metrics**.

You might think that regardless of this, Tableau should have no problems with the tables if you activate the interpreter, like below (connect to the Population table using the Data Interpreter):

Tableau - Book1

File Data Server Window Help

Connections Add

W5 Population by country

W5 Population by country.csv

Connection Live Extract Filters D | Add

W5 Population by country.csv

Sort fields Data source order ▾

Show aliases Show hidden fields 264 rows

Country Name	Country Code	Indicator Name	Indicator Code	1960	1961	1962	1963	1964	1965	1966	1967	1968
Aruba	ABW	Population, total	SP.POP.TOTL	54,211	55,438	56,225	56,695	57,022	57,380	57,715	58,065	
Afghanistan	AFG	Population, total	SP.POP.TOTL	8,986,973	9,169,410	9,351,441	9,543,209	9,744,781	9,956,320	10,174,836	10,399,326	
Angola	AGO	Population, total	SP.POP.TOTL	5,454,933	5,551,472	5,608,539	5,679,458	5,735,044	5,770,570	5,781,214	5,774,245	
Albania	KRZ	Population, total	SP.POP.TOTL	1,609,800	1,659,800	1,711,319	1,762,621	1,814,135	1,864,791	1,914,573	1,965,590	
Andorra	AND	Population, total	SP.POP.TOTL	18,411	14,375	18,370	18,412	17,489	18,549	19,647	20,758	
Arab World	ARB	Population, total	SP.POP.TOTL	92,197,753	94,724,510	97,334,442	100,034,179	102,832,760	105,736,431	108,758,610	111,893,364	
United Arab Emirates	ARE	Population, total	SP.POP.TOTL	92,410	100,796	112,110	125,130	138,039	149,857	159,976	169,771	
Argentina	ARG	Population, total	SP.POP.TOTL	20,481,779	20,817,266	21,152,062	21,488,912	21,824,425	22,159,650	22,494,035	22,828,869	
Armenia	ARM	Population, total	SP.POP.TOTL	1,874,121	1,941,492	2,009,526	2,077,578	2,145,001	2,211,819	2,276,084	2,339,127	
American Samoa	ASM	Population, total	SP.POP.TOTL	20,123	20,602	21,253	22,034	22,854	23,672	24,462	25,348	
Antigua and Barbuda	ATG	Population, total	SP.POP.TOTL	64,121	65,001	66,841	66,702	67,641	68,698	69,815	61,241	
Australia	AUS	Population, total	SP.POP.TOTL	10,276,477	10,483,000	10,742,000	10,950,000	11,167,000	11,388,000	11,651,000	11,799,000	

Go to Worksheet X Ponce, Carlos

Data Source Sheet 1

However, having your data like this is very limiting. Imagine you want to show a line graph of population throughout the years: **you can't**. You can't because you need one Measure for population and one Dimension for Years, and you can't make them as calculated variables. How would you create the Year dimension, if each row contains all the years?

Before fixing this, add your GDP table and do an **inner join** with the Population table, using Country Code and Country Name as keys.

The screenshot shows a Tableau interface with two data sources: 'W5 Population by country.csv' and 'W5 GDP by country.csv'. A 'Left' join is applied. The resulting table contains numerous columns, including 'Country Name' and 'Country Code' from the Population table, and 'Indicator Name' and 'Indicator Code' from the GDP table. The data includes various countries like Aruba, Afghanistan, Angola, Albania, Andorra, Arab World, United Arab Emirates, Argentina, Armenia, American Samoa, Antigua and Barbuda, and Australia.

We see many columns on the resulting table, if you click on the highlighted button you will see them listed vertically, which is helpful in this case. Look for **Country Code** and **Country Name** from the **GDP** table, select them, right-click and choose **Hide** (but don't hide the ones from the Population table). Look for **Indicator Name** and **Indicator Code** from both Population and GDP tables and hide them as well.

The screenshot shows the same Tableau interface as before, but with a context menu open over the 'Indicator Code' column. The 'Hide' option is highlighted in blue, indicating it is selected. Other options visible in the menu include 'Rename', 'Copy Values', 'Create Calculated Field...', 'Pivot', and 'Merge Mismatched Fields'.

Now do a second **inner join** with the Country Details table based on Country Code.

The screenshot shows the Tableau Data Source interface with two joined datasets: 'W5 Population by country.csv' and 'W5 Country details.csv'. A third dataset, 'W5 GDP by country.csv', is listed under 'Available Datasets' but is not currently selected. The 'Join' dialog is open, showing the configuration for the second join:

- Join Type:** Inner
 - Data Source:** W5 Country details.csv
 - Table:** W5 Country details
 - Key:** Country Code (W5 Country details.csv)
 - Key:** Country Code (W5 GDP by country.csv)
- Sort Fields:** Data source order
- Rows:** 263

The main data view shows a partial list of countries with their respective income groups, regions, and GDP values from 1960 to 1963.

Hide the Special Notes column.

The screenshot shows the Tableau Data Source interface with the same joined datasets. The 'Special Notes' column is highlighted in the 'Columns' pane, and a context menu is open with the 'Hide' option selected. The 'Columns' pane also lists other columns from the joined datasets.

Country Code (W5...)	Region	Income Group	Special Notes	1960 (W5 GDP by ...)	1961 (W5 GDP by ...)	1962 (W5 GDP by ...)	1963 (W5 GDP by ...)
ABW	Latin America & Carib...	High income	null	537,777,811.11	548,888,895.56	546,666,677.76	751,111.14
AFG	South Asia	Low income	null	Afghanistan	null	null	null
AGO	Sub-Saharan Africa	Lower middle income	null	Angola	null	null	null
ALB	Europe & Central Asia	Upper middle income	null	Albania	null	null	null
AND	Europe & Central Asia	High income	null	Andorra	null	null	null
ARB	null	null	Arab World aggregate	Arab World	null	null	null
ARE	Middle East & North A...	High income	null	United Arab Emirates	null	null	null
ARG	Latin America & Carib...	Upper middle income	null	Argentina	null	null	24,450,604,877.61
ARM	Europe & Central Asia	Upper middle income	null	Armenia	null	null	null
ASM	East Asia & Pacific	Upper middle income	null	American Samoa	null	null	null
ATG	Latin America & Carib...	High income	null	Antigua and Barbuda	null	null	null
AUS	East Asia & Pacific	High income	Fiscal year end: June	Australia	19,577,668,271.92	19,652,816,664.80	19,692,485,160.71

By now you should have a data source looking like this:

The screenshot shows the Tableau Data Source view with three joined CSV files:

- W5 Population by country.csv**: 264 rows
- W5 Country details.csv**: 263 rows
- W5 GDP by country.csv**: 263 rows

A tip at the bottom left states: "Only Country Details has 263 rows. Therefore, if we used inner join between Population and Country Details, it's expected that we remove that missing country."

Tip:

We talked before about the risks of not doing a join properly. A good validation for your data joins is to check the number of rows that show on the right textbox (263). When we looked on our tables, we said they had 264 rows each, only Country Details had 263. Therefore, if we used inner join between Population and Country Details, it's expected that we remove that missing country.

Now back to the issue of the years on the headers. To fix this, we need to do a procedure called **pivoting**.

In a spreadsheet, it would mean for you to create a new column called Year, a new column called GDP, and for all your 263 rows put Year = 1960 (the first year on your table) and on the GDP column the values that show by country. Then, you would have to duplicate your rows and again, put Year = 1961 and take the GDP values from 1961. You would have to repeat this process for each year you have, which considering that we have data up to 2018 would take you hours.

Fortunately, Tableau integrates a feature that does this for you on one click. Select all the columns with year as header, right click and select **Pivot**.

Tip:

For avoiding clicking one by one, click the first year 1960, hold Shift instead of Ctrl, and click the last year of the table. Then, still holding shift, now also hold Ctrl and click the **Country Name** and **Country Code** columns on the middle to unselect them, so you don't pivot those.

The screenshot shows the Tableau Data Source view for the 'W5 Population by country' dataset. The interface includes a navigation bar with File, Data, Server, Windows, Help, and a toolbar with various icons. On the left, there's a Connections pane with 'W5 Population by country' selected, and a Files pane listing various CSV files. The main area displays a table with several columns: 2015 (W5 GDP by ..., 2016 (W5 GDP by ..., 2017 (W5 GDP by ..., 2018 (W5 GDP by ..., Country Name, and Country Code. A context menu is open over the table, with the 'Pivot' option highlighted under the 'Transform' section. Other options in the menu include Rename, Copy Values, Hide, Create Calculated Field, and Merge Mismatched Fields. The bottom of the screen shows the Tableau ribbon with tabs like Data Source, Sheet 1, etc., and a status bar indicating 'Force, Carbs'.

It will take a moment; at the end, you will have this result:

The screenshot shows the Tableau Data Source interface. On the left, the 'Connections' pane lists several CSV files, with 'W5 Population by country' selected. The main workspace displays a joined dataset with the following structure:

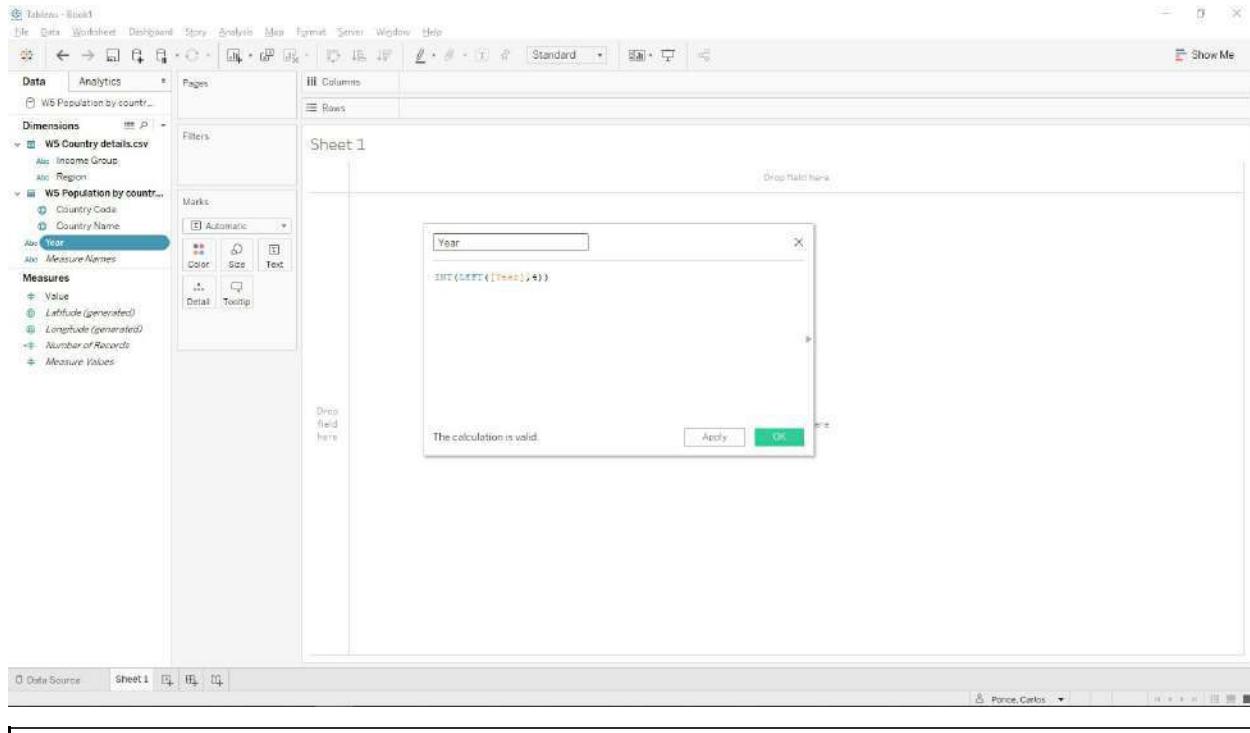
Pivot Field Names	Pivot Field Values	Region	Income Group	Country Name	Population
1960 (W5 GDP by cou...	null	Latin America & Carib...	High income	Aruba	ABW
1960 (W5 GDP by cou...	537,777,611.11	South Asia	Low income	Afghanistan	AFG
1960 (W5 GDP by cou...	null	Sub-Saharan Africa	Lower middle income	Angola	AGO
1960 (W5 GDP by cou...	null	Europe & Central Asia	Upper middle income	Albania	ALB
1960 (W5 GDP by cou...	null	Europe & Central Asia	High income	Andorra	AND
1960 (W5 GDP by cou...	null	null	null	Arab World	ARB
1960 (W5 GDP by cou...	null	Middle East & North A...	High income	United Arab Emirates	ARE
1960 (W5 GDP by cou...	null	Latin America & Carib...	Upper middle income	Argentina	ARG
1960 (W5 GDP by cou...	null	Europe & Central Asia	Upper middle income	Armenia	ARM
1960 (W5 GDP by cou...	null	East Asia & Pacific	Upper middle income	American Samoa	ASM
1960 (W5 GDP by cou...	null	Latin America & Carib...	High income	Anguilla and Barbuda	ATG
1960 (W5 GDP by cou...	18,577,668,271.92	East Asia & Pacific	High income	Australia	AUS

Now you have way more rows, because you now have 263 x the number of years you pivoted (60 years on each table so $120 = 31,560$ rows). But you also have way less columns, and you have a proper Year and GDP/Population column. Rename both by double clicking Pivot **Field Names** and **Pivot Field Values**.

For most things in Tableau, there is more than one way of doing it. An alternative way you could have set up your data source is by making the Population and GDP tables separately instead of joining them (like when you do a data blending, you put the tables separately). You pivot each and that way you have a separate table with GDP numbers and another with Population.

Anyways, our approach also works the same. It only requires this couple of extra steps:

1. Create a new “**Year Clean**” calculated value which is equal to the first four characters of the existing **Year** dimension. You do this with `INT(LEFT([Year],4))` (we add the `INT()` function so it converts the string result “2019” to numeric 2019).



Explained:

INT(), FLOAT(), DATE(), and STR() are some of the typical type conversion functions used in calculated values. You will frequently use these when concatenating different dimensions and they are of different types. You can find more on the function reference, accessible on the arrow on the right when making a calculated value.

Move it to discrete. Create another calculated value called GDP, that checks if the **Year** value contains “.csv”. This because if you go back to your resulting table, the lines that come from the GDP table have **Year** values like “1960 (W5 GDP by country.csv)”. This is also the reason why we created Year Clean, to remove that extra labelling.

The formula would be as follows: IF CONTAINS([Year], “.csv”) THEN [Value] ELSE 0 END

If the Year value does not contain “.csv”, then it’s a Population value, because those have the years clean.

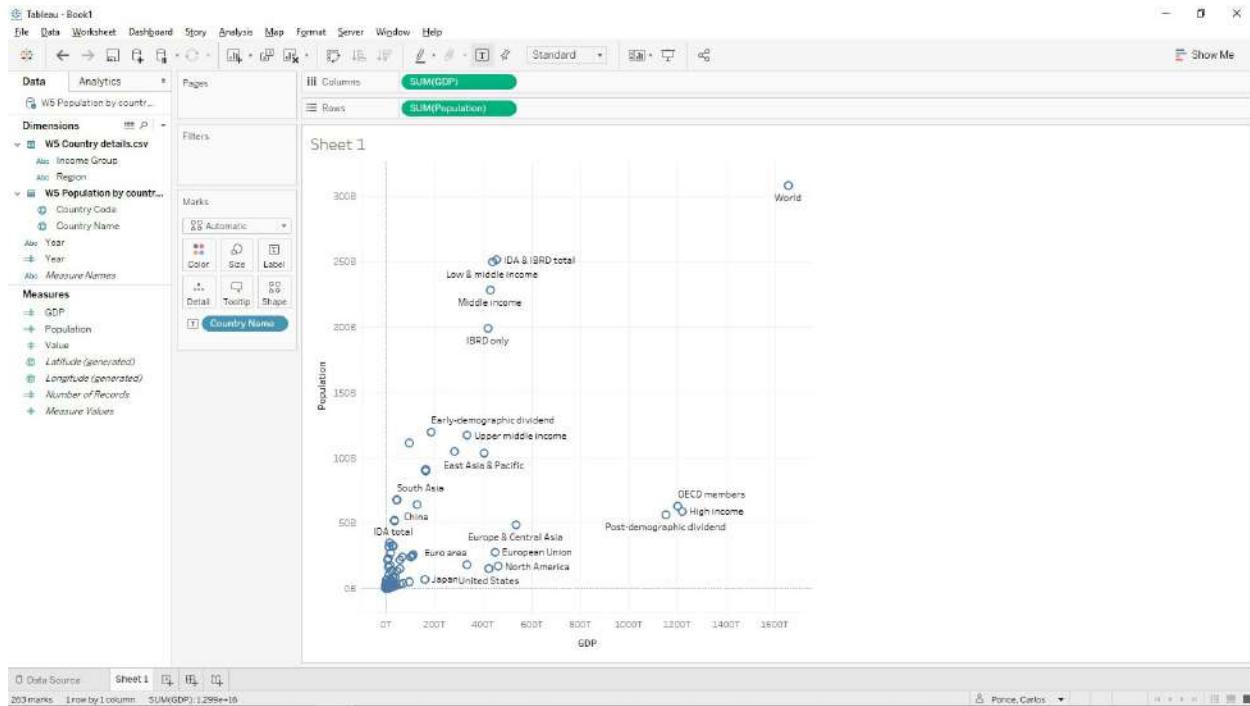
The screenshot shows the Tableau Data Source interface. On the left, the 'Dimensions' and 'Measures' panes are visible. In the 'Measures' pane, the 'GDP' measure is selected. A tooltip window for 'GDP' displays the formula: `IF CONTAINS([Text],"GDP") THEN [Value] ELSE 0 END`. The status bar at the bottom right indicates 'Ponce, Carlos'.

Finally, create a calculated value called **Population**, which has the same formula as GDP, but add a NOT before the CONTAINS so it's the inverse.

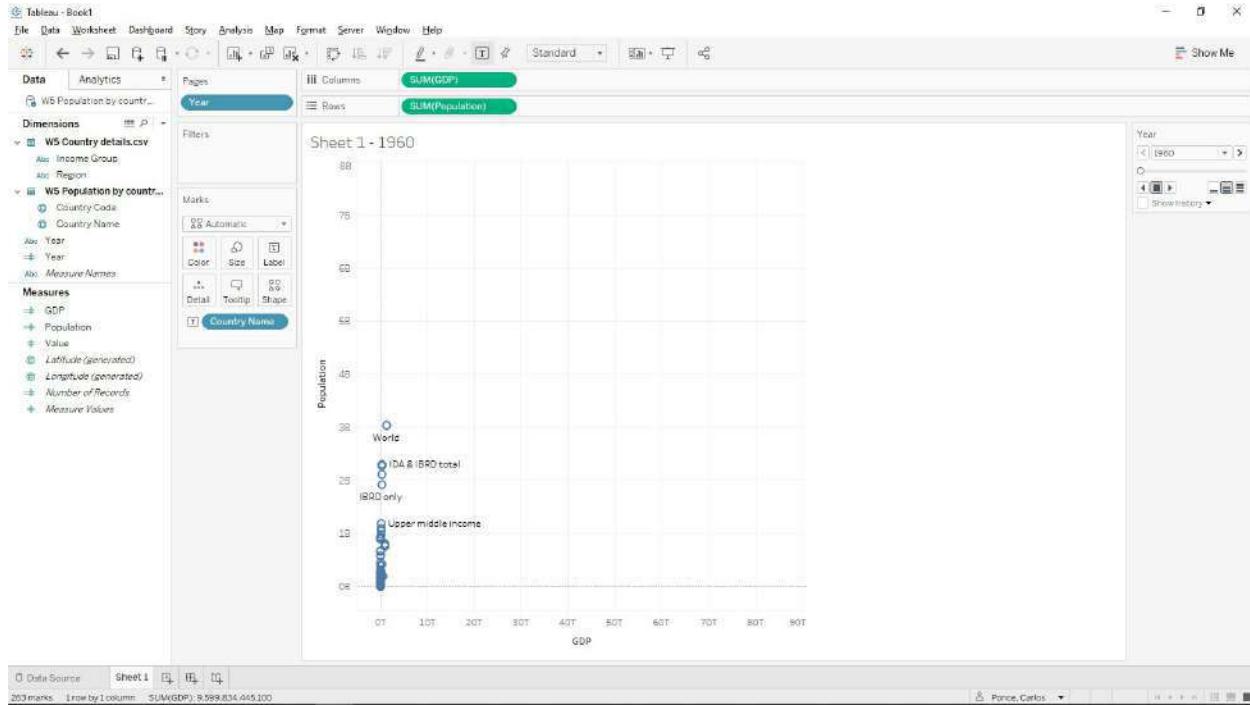
The screenshot shows the Tableau Data Source interface. The 'GDP' measure is still selected in the 'Measures' pane. A tooltip window for 'Population' displays the formula: `IF NOT CONTAINS([Text],"GDP") THEN [Value] ELSE 0 END`. To the right of the formula, a tooltip for the 'NOT' operator is shown: 'IF NOT <expr> THEN <then> END'. The status bar at the bottom right indicates 'Ponce, Carlos'.

Now you have your dataset ready! Let's start with our first sheet.

Add **SUM(GDP)** to Columns, **SUM(Population)** to Rows, and **Country Name** to Label.



Next step: add Year to the Pages shelf, the one above the Filters shelf.



Explained:

Notice this new shelf: it's the control that appears when you drag a dimension to the **Pages** shelf. It's a navigator between the different values (in this case

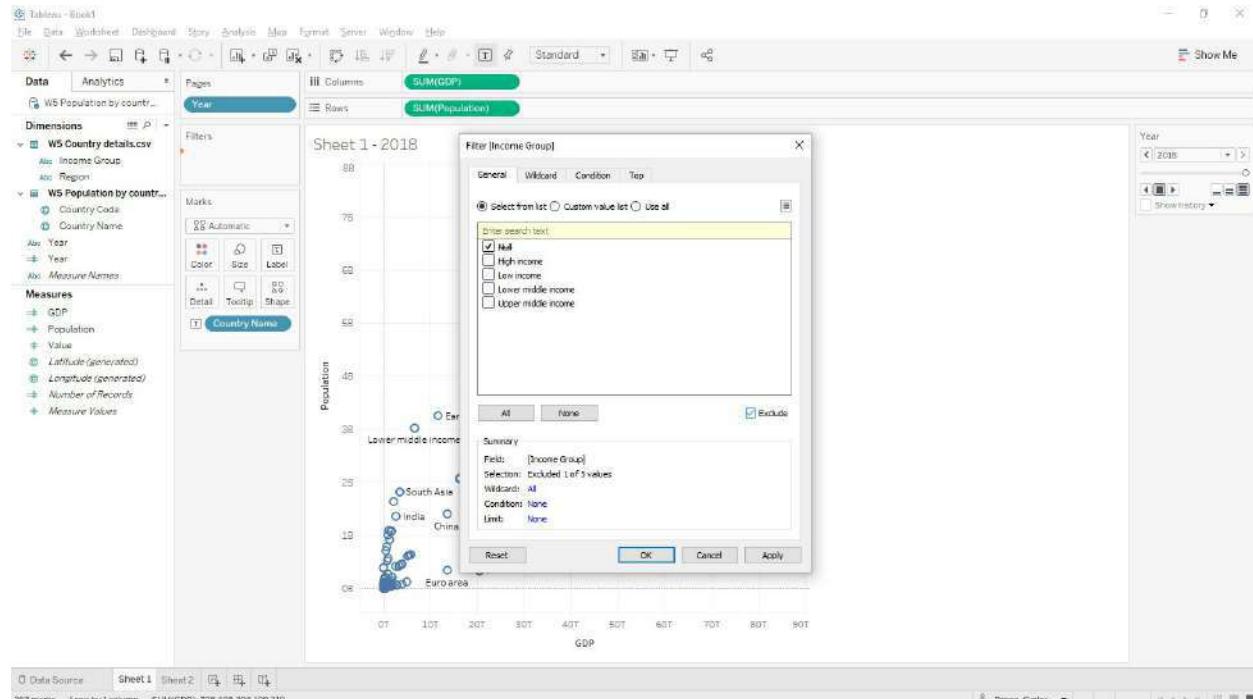
years), on the top you have a dropdown to change from them. Below you can see a slider that works for the same thing. At the bottom you see new buttons; they are for playing an animation that moves from values. Use the left buttons to play right-to-left, pause, or play left-to-right, and the other three ones to change the speed of the animation. Finally, the **Show history** option lets you see the line of movement for your points, and you can additional configuration by clicking the name and test it.



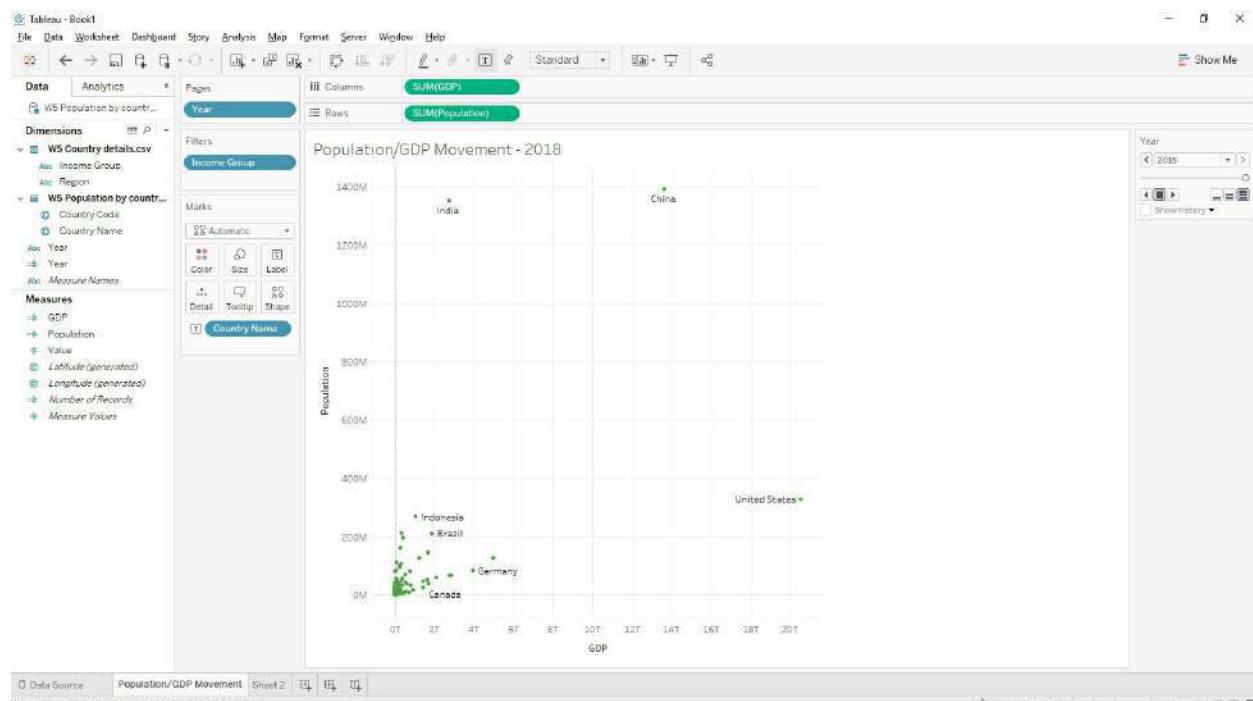
Looking on the dots, we see some strange labels for “countries”, like “World”, “IBRD”, and “Upper middle income” to name a few. We must filter them out to have an actual scatter plot of countries. We could go one by one to remove which ones are not real countries, but instead we’ll do a faster test. If on a separate sheet you list all the countries by name and to the side their income group (which comes from the Country Details table), you will see that those fake countries have **null** values.

Country Name	Income Group
Sudan	Lower middle income
Suriname	Upper middle income
Sweden	High income
Switzerland	High income
Syrian Arab Republic	Low income
Tajikistan	Low income
Tanzania	Low income
Thailand	Upper middle income
Timor-Leste	Lower middle income
Togo	Low income
Tonga	Upper middle income
Trinidad and Tobago	High income
Tunisia	Lower middle income
Turkey	Upper middle income
Turkmenistan	Upper middle income
Turks and Caicos Islands	High income
Tuvalu	Upper middle income
Uganda	Low income
Ukraine	Lower middle income
United Arab Emirates	High income
United Kingdom	High income
United States	High income
Upper middle income	Null
Uruguay	High income
Uzbekistan	Lower middle income
Venezuela	Lower middle income
Venezuela, RB	Upper middle income
Vietnam	Lower middle income
Virgin Islands (U.S.)	High income
West Bank and Gaza	Lower middle income
World	Null

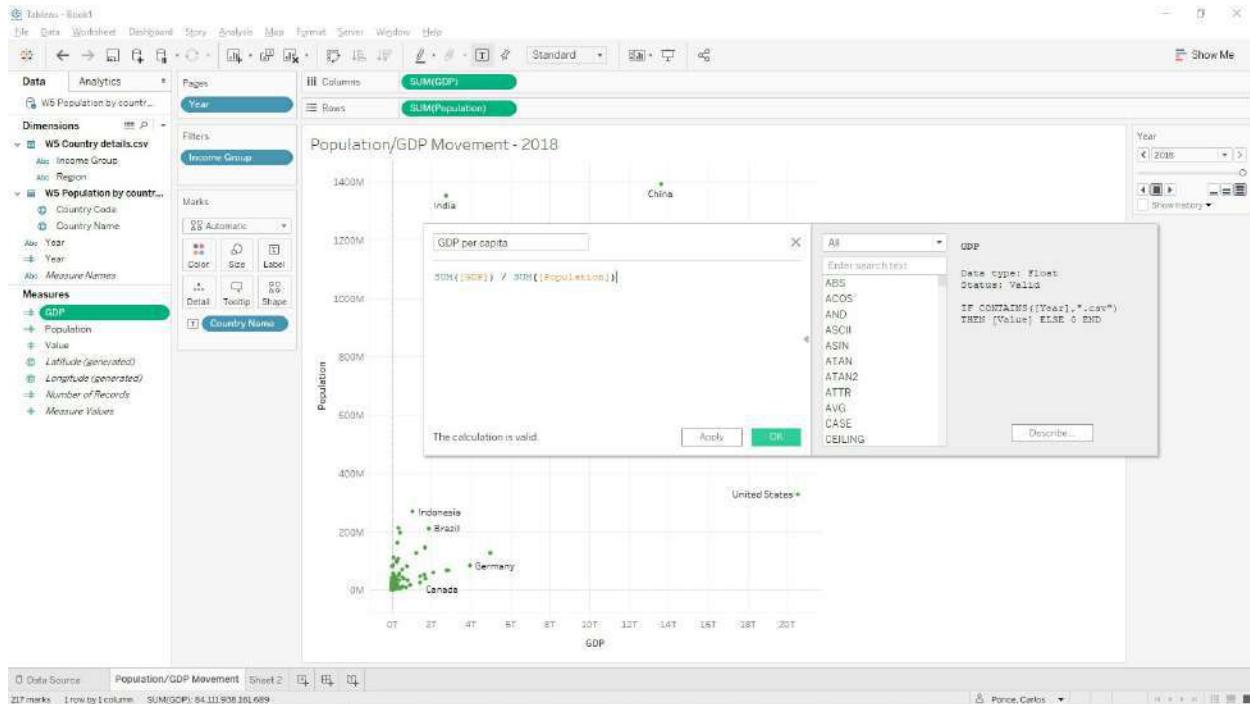
Therefore, add a filter on our scatter to exclude all entries with **Income Group** = null.



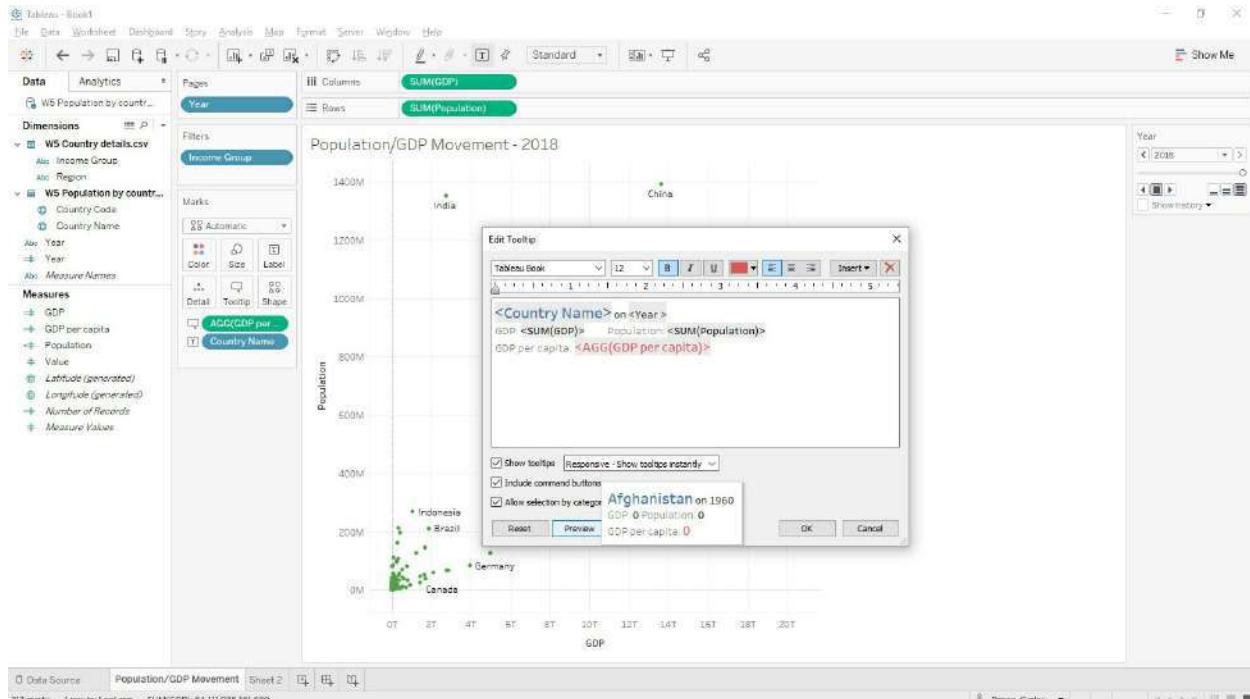
Rename to *Population/GDP Movement* and apply some styling to the dots.



Create a new calculated value called “GDP per capita”, equal to the sum of GDP divided by the sum of population.



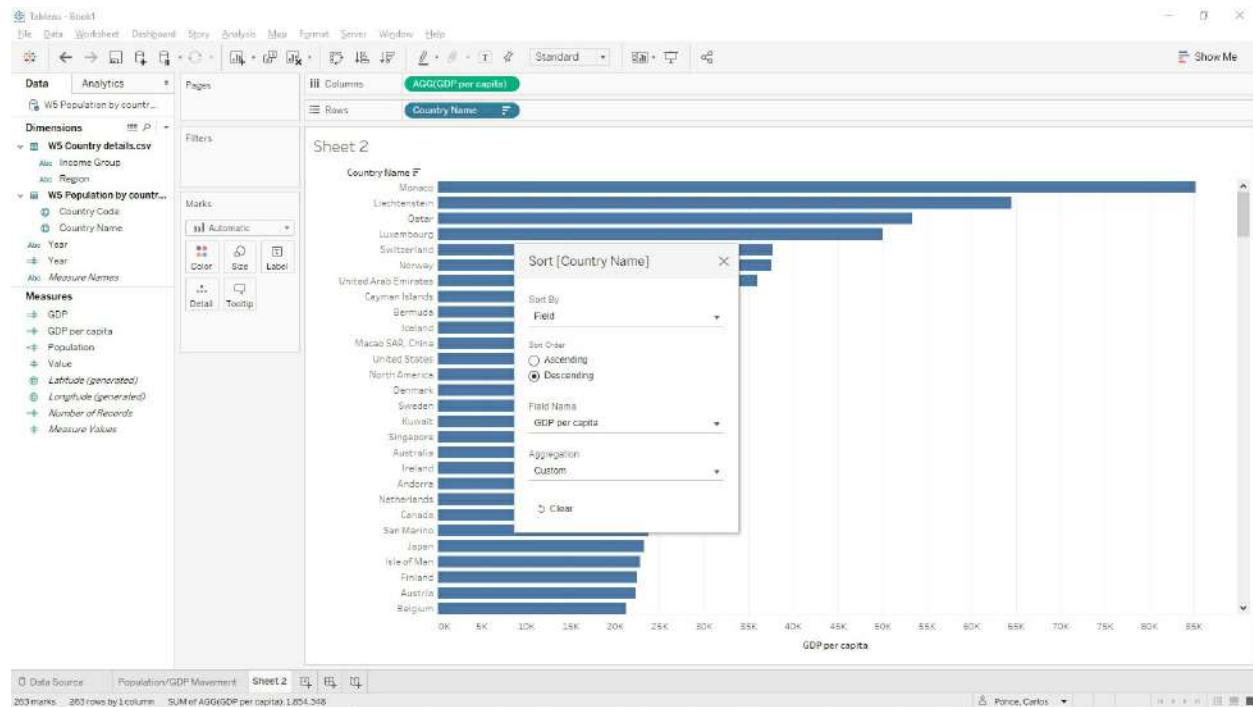
Now enhance the tooltip on our sheet as below:



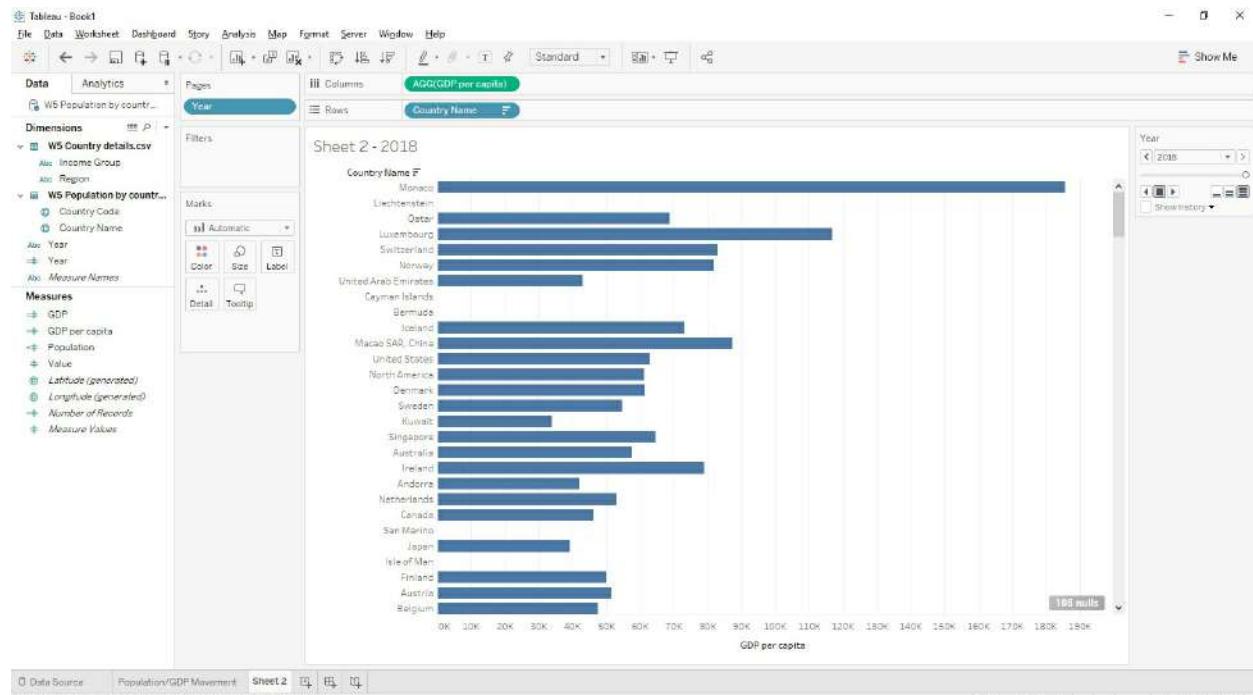
The fonts used are 14 on <Country Name>, 12 on <AGG(GDP per capita)>. You can reference the bolding and coloring from the image.

Empty Sheet 2 with the button on the top. Add **GDP per capita** on columns and **Country Name** on rows and apply sort on the last by GDP per

capita descending.



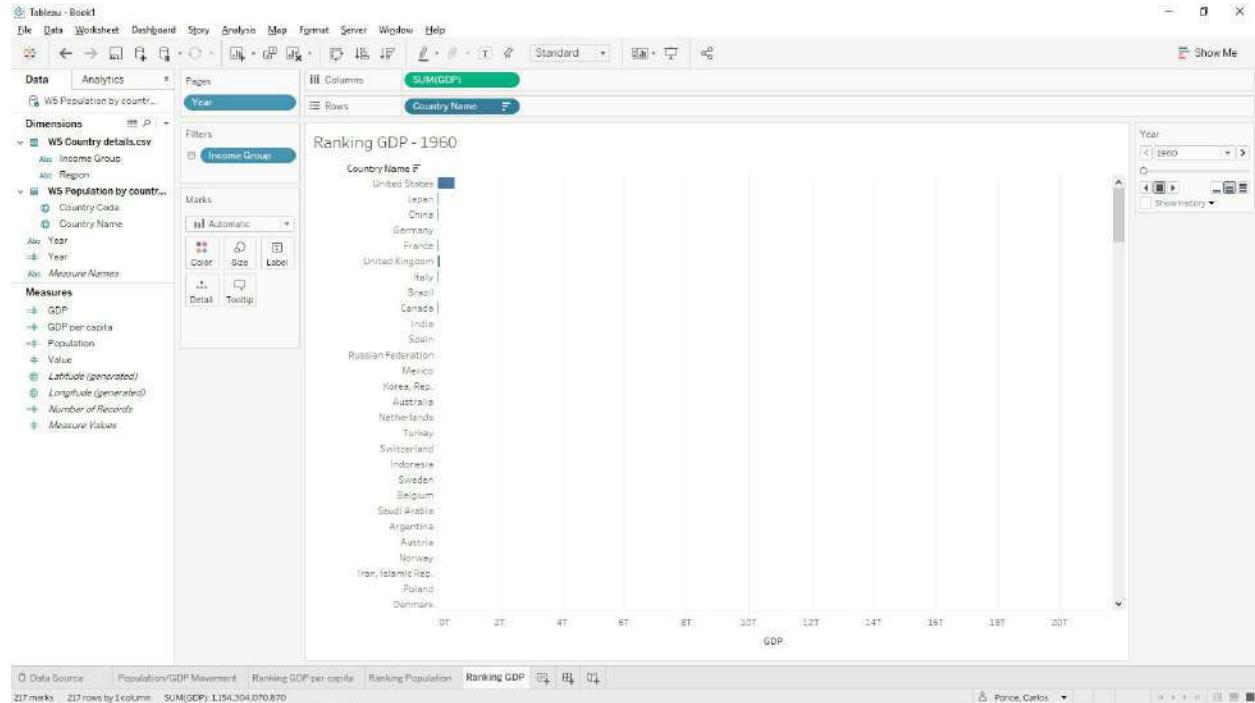
Add **Year** to **Pages** again, play with it, and notice how on 2018 we got many empties. This is normal; if we go back to the source, we saw GDP had a lot of nulls.



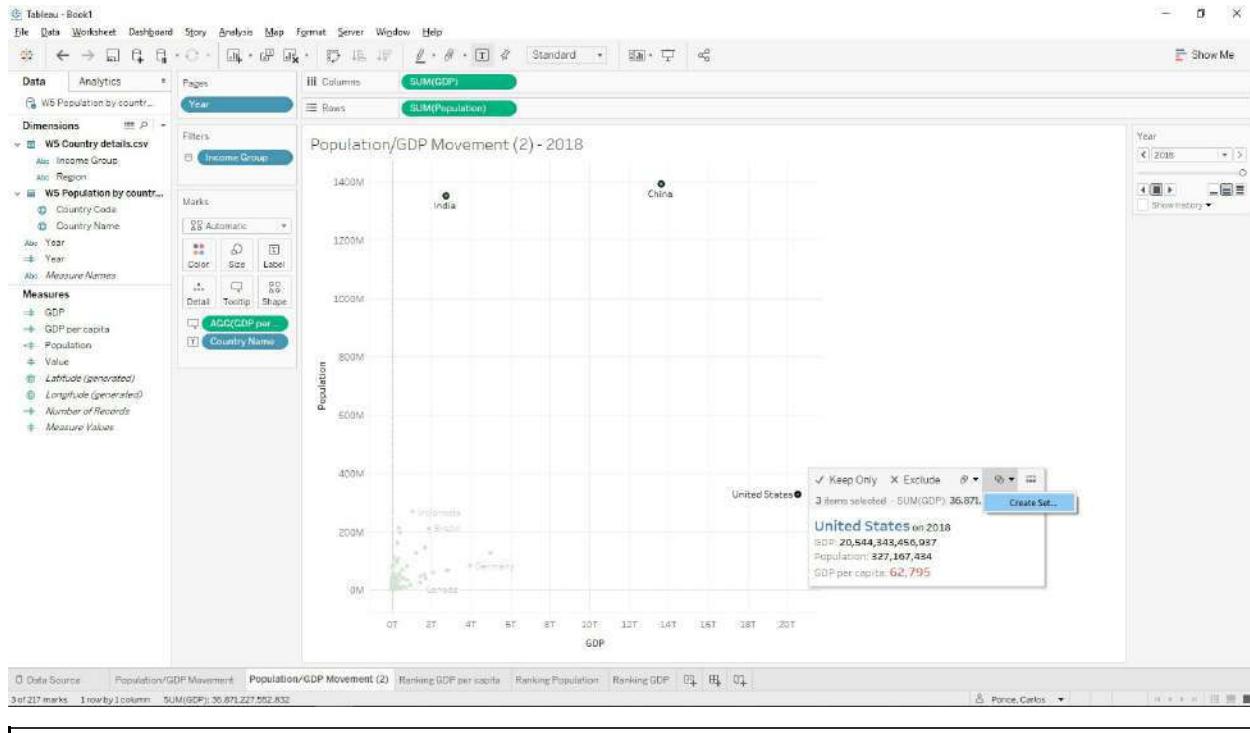
Call the sheet *Ranking GDP per capita*. Duplicate it and make the same

graph but for *Population* and *GDP* (just change the measure on columns and reapply sort on it).

Make the **Income Group** filter apply to all sheets (from the first sheet).



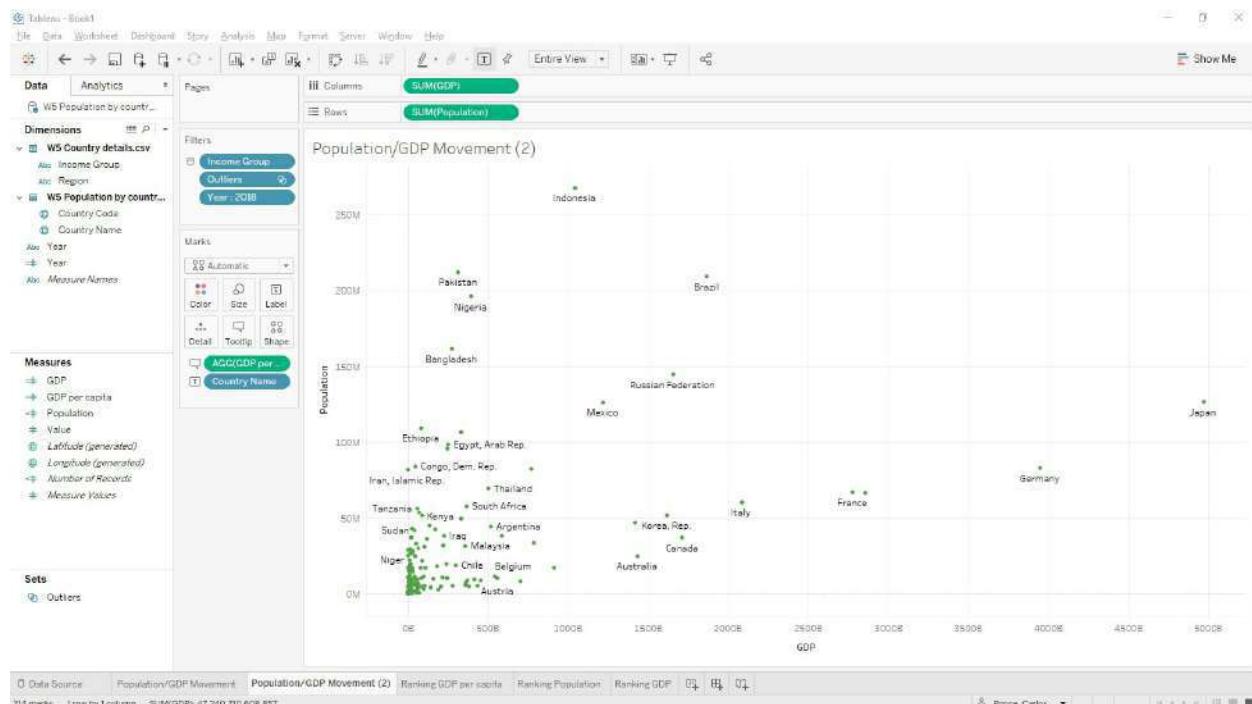
Duplicate your first sheet and create a set with the three outliers by selecting them (ctrl + click each or draw a rectangle around them) and clicking **Create Set**. Call it “Outliers” and use it as a filter.



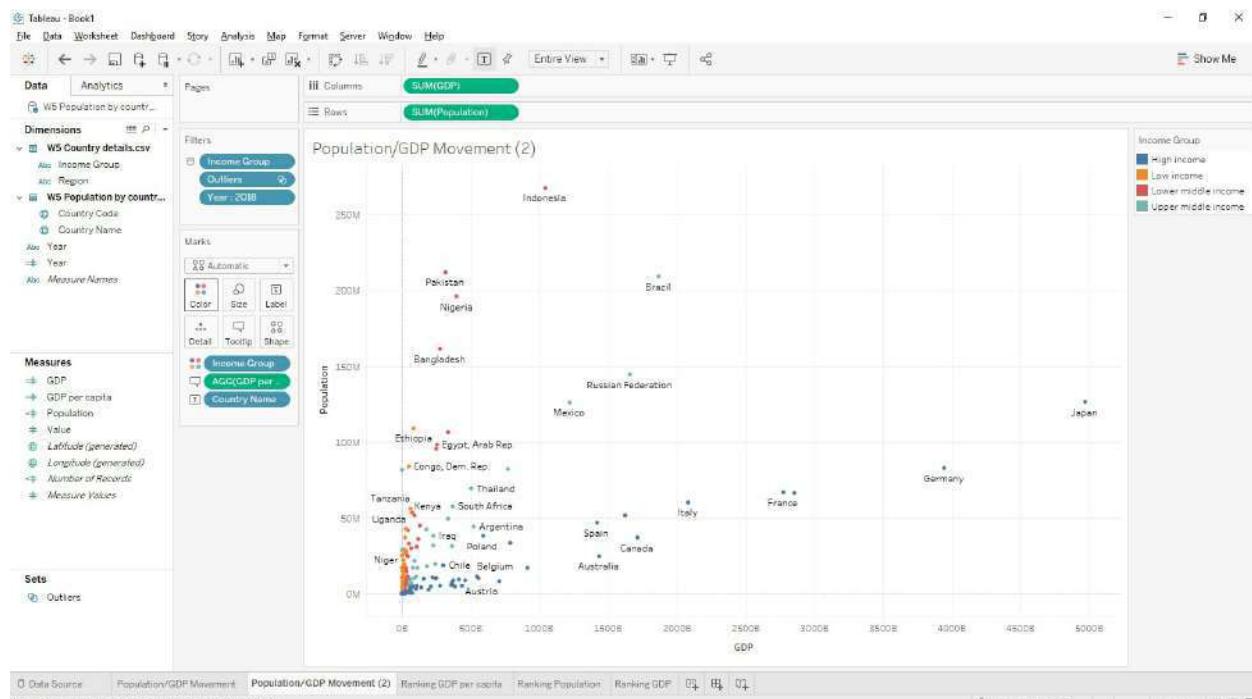
Explained:

Sets represent data that comes from a group of filters. They are rarely used but work as a faster alternative when you want to analyze a portion of your data.

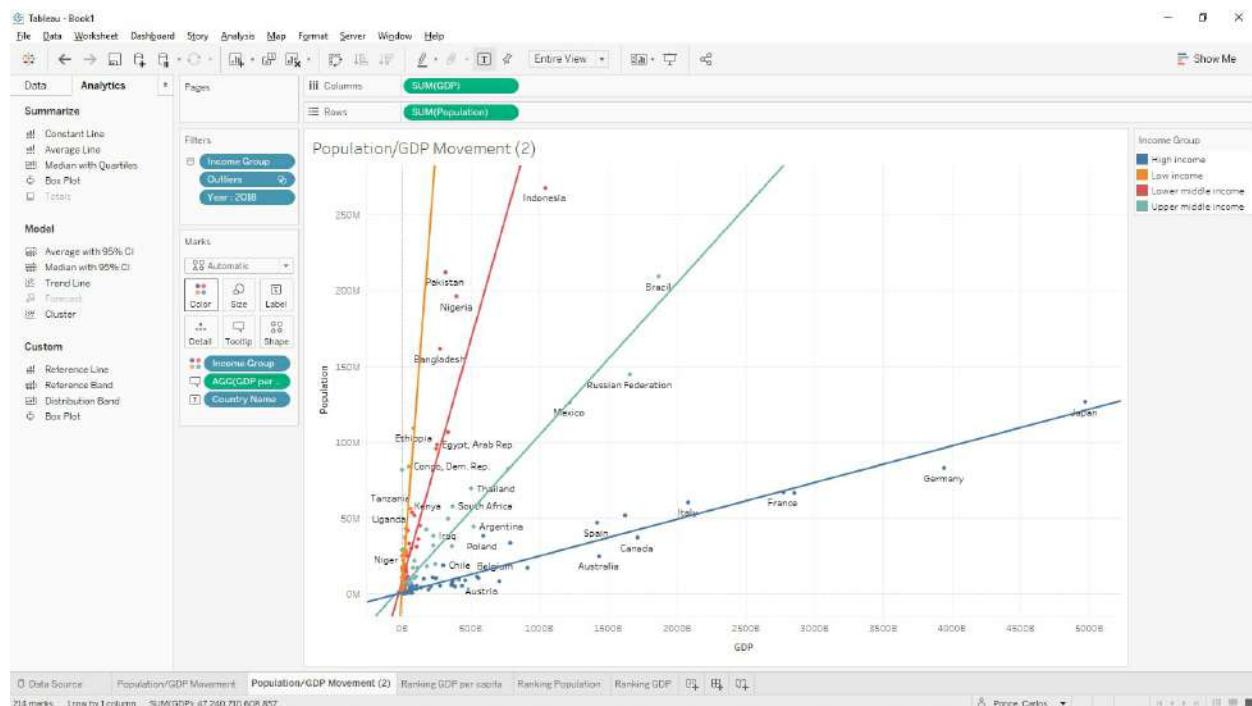
Move **Year** from Page to Filters, select only 2018. Set the graph to Entire View.



Drag Income Group to Color.

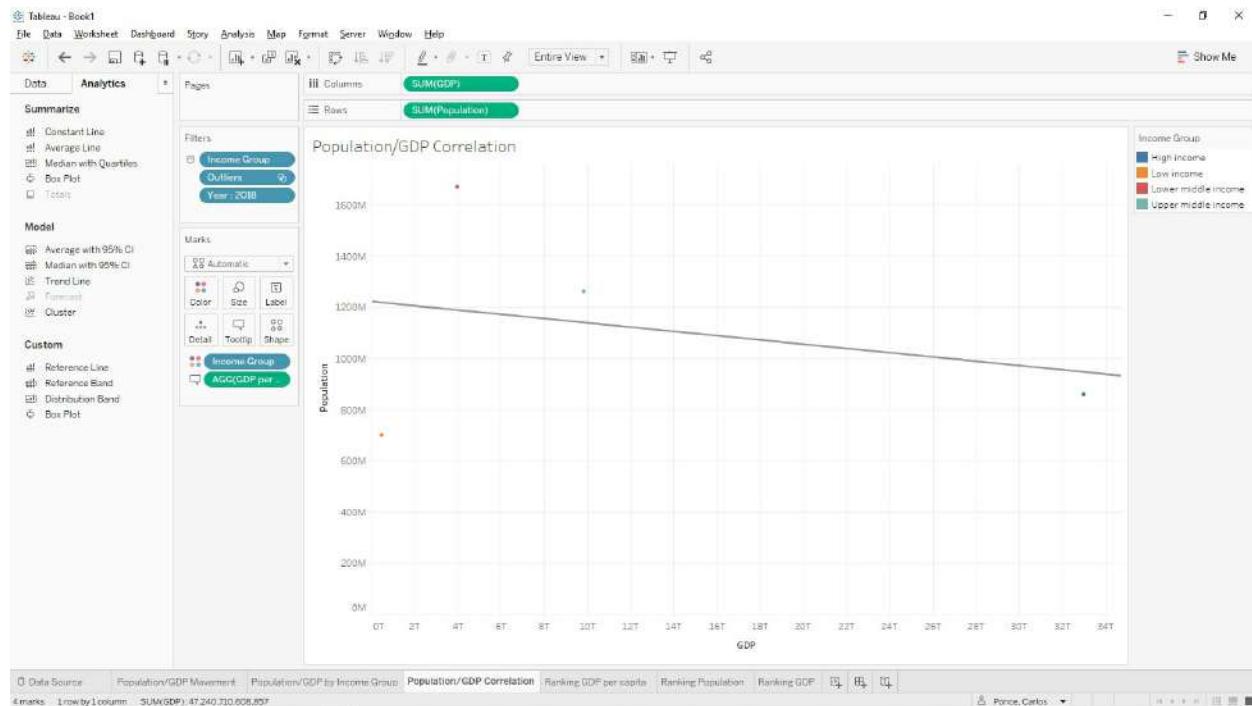


Add a **Linear Trend Line**, it will be made by **Income Group**. Observe how there is a difference in trend by income group.

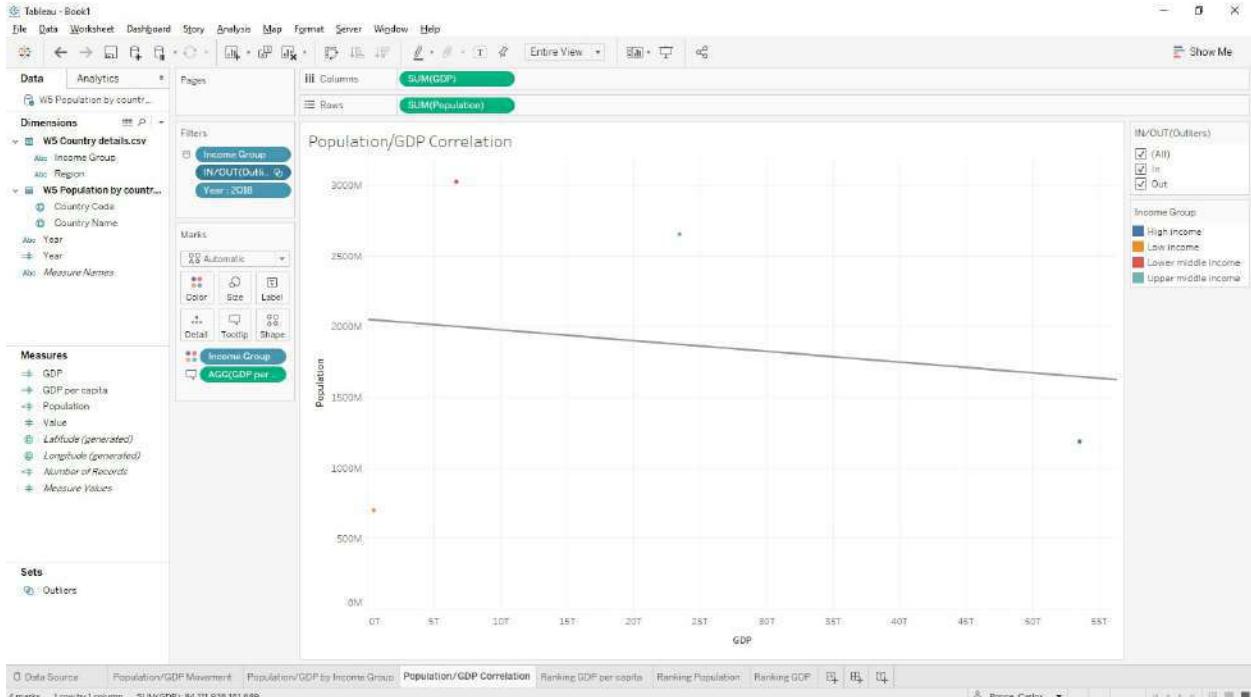


Rename the sheet to *Population/GDP by Income Group*.

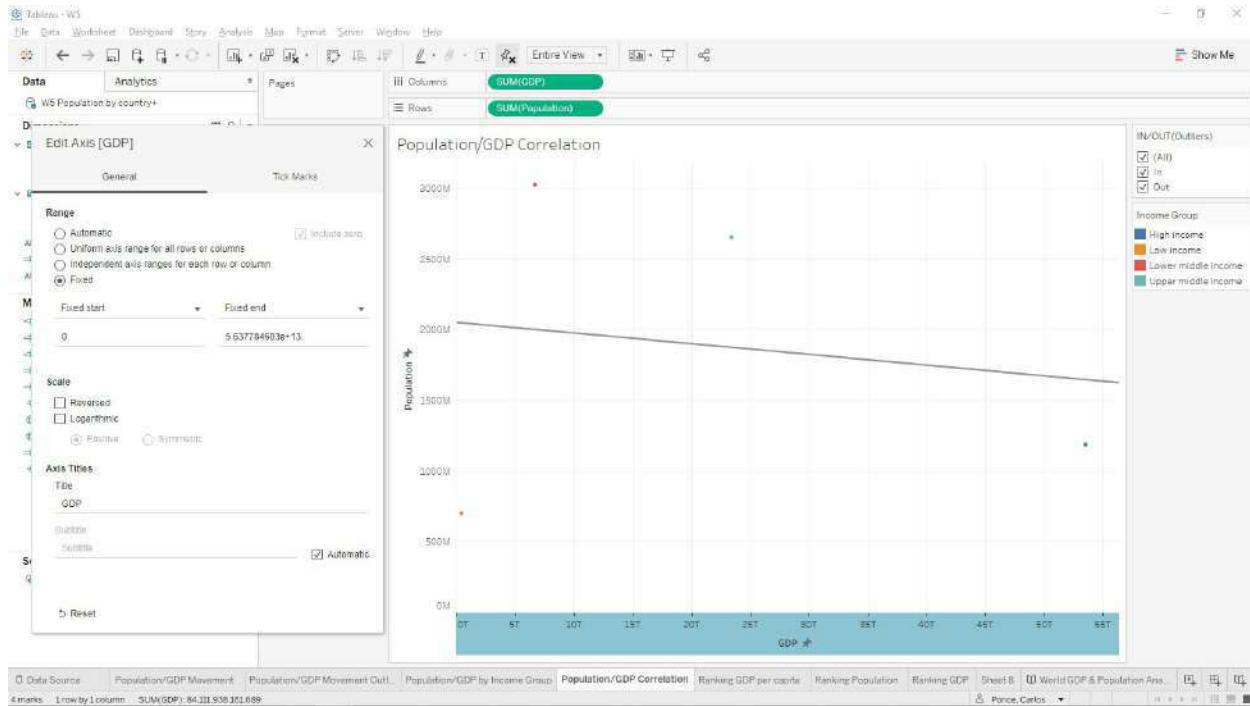
Duplicate it and remove **Country Name** from Text. Watch how there is a low correlation between population and GDP when you disregard **Income Group**. According to the linear trend, with less population there is more GDP, but it's not a strong slope, so we cannot determine this is true.



On the other hand, what happens if we consider the outliers? Right click the filter, select **Show in/out**, select **In**, and **Show filter**. By comparing to the previous look, you can see that **Out** represents the outliers. The tendency looks a bit the same.

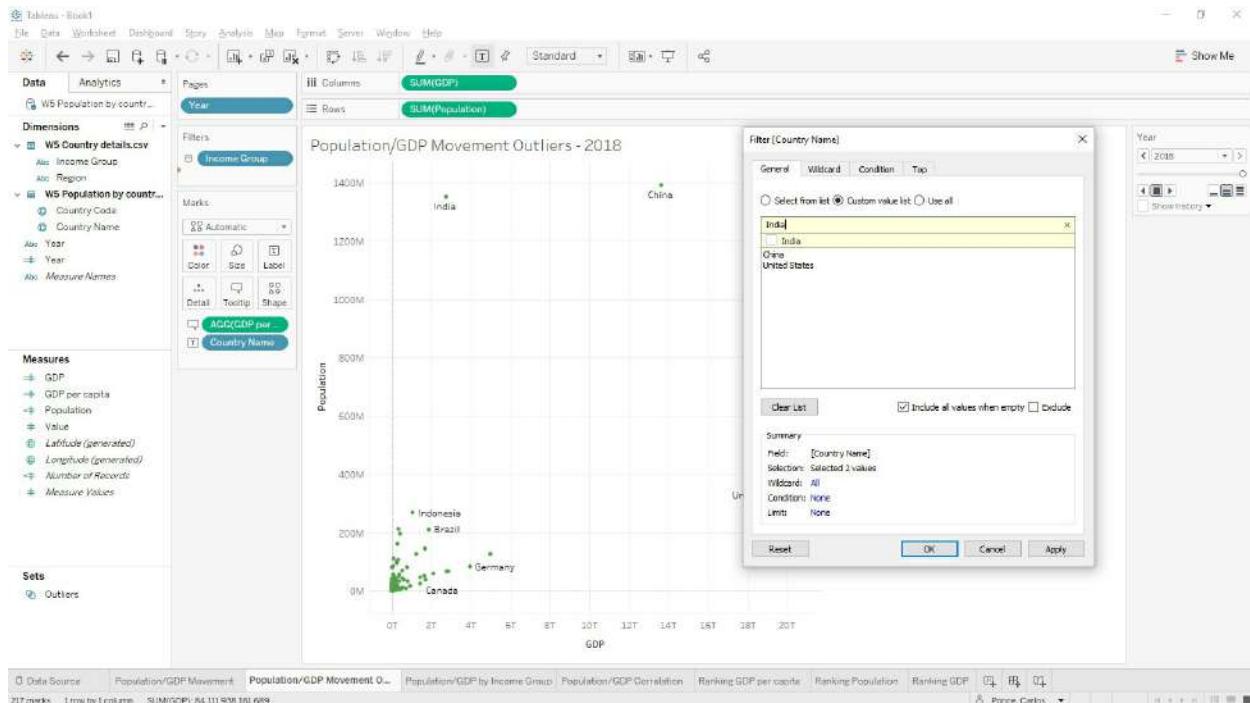


However, there is a problem; the axes are moving. So, having both In and Out selected, right click the **Population** axis, click **Edit axis**, set the range to **fixed**, and keep the values that show there. Do the same for the **GDP** axis.

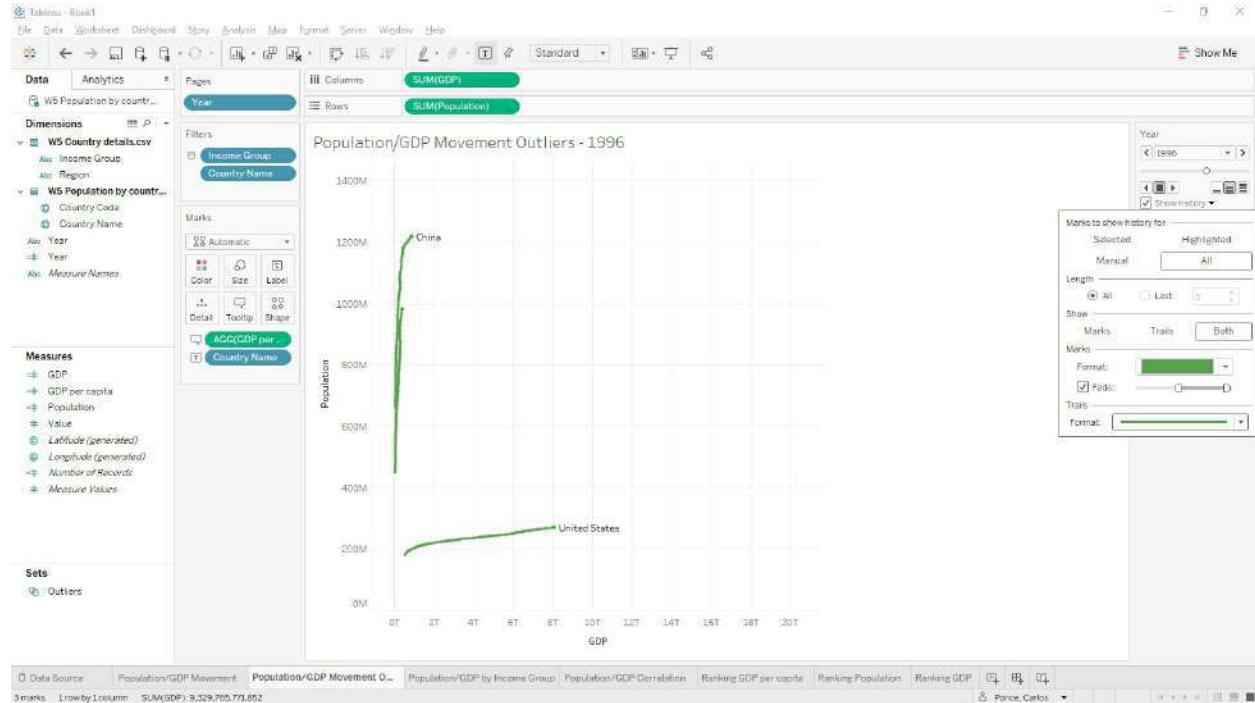


Now we can see the slope gets even less pronounced when considering outliers. Rename the sheet to *Population/GDP Correlation*.

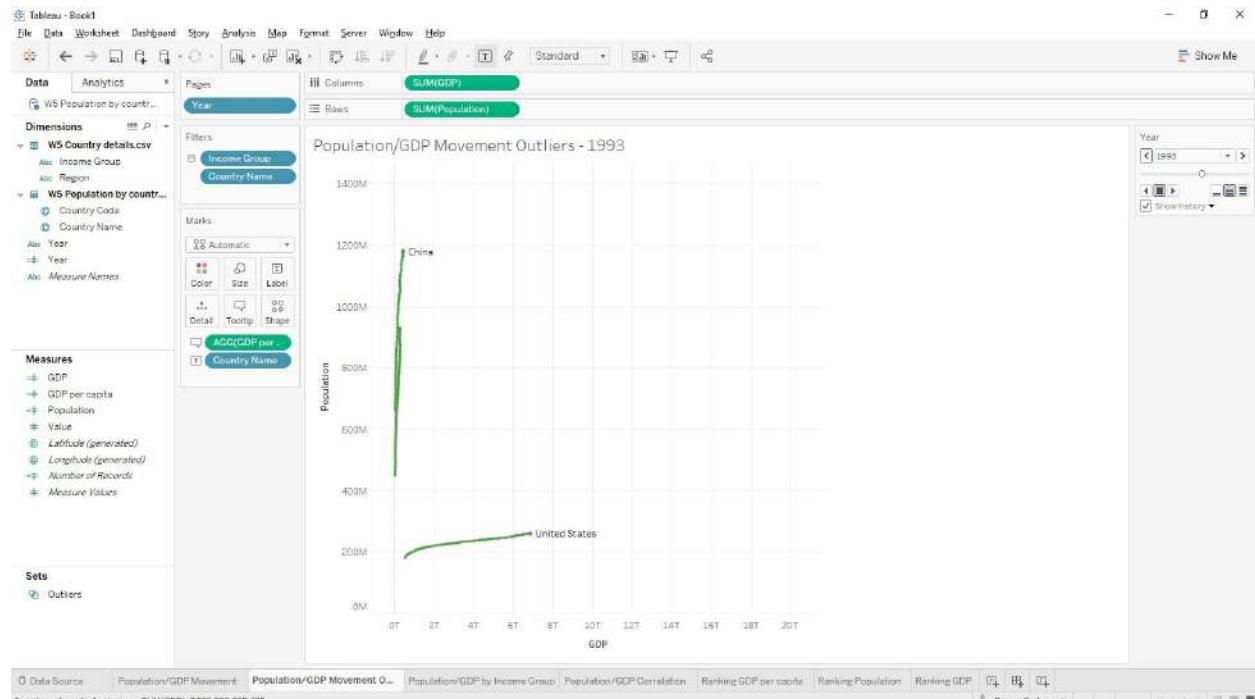
Duplicate sheet one, now filter only the three countries using the **Country Name** filter. Don't use the set because that is also limited to year 2018. You can type the names on the **Custom value list** like below or select them from the list.



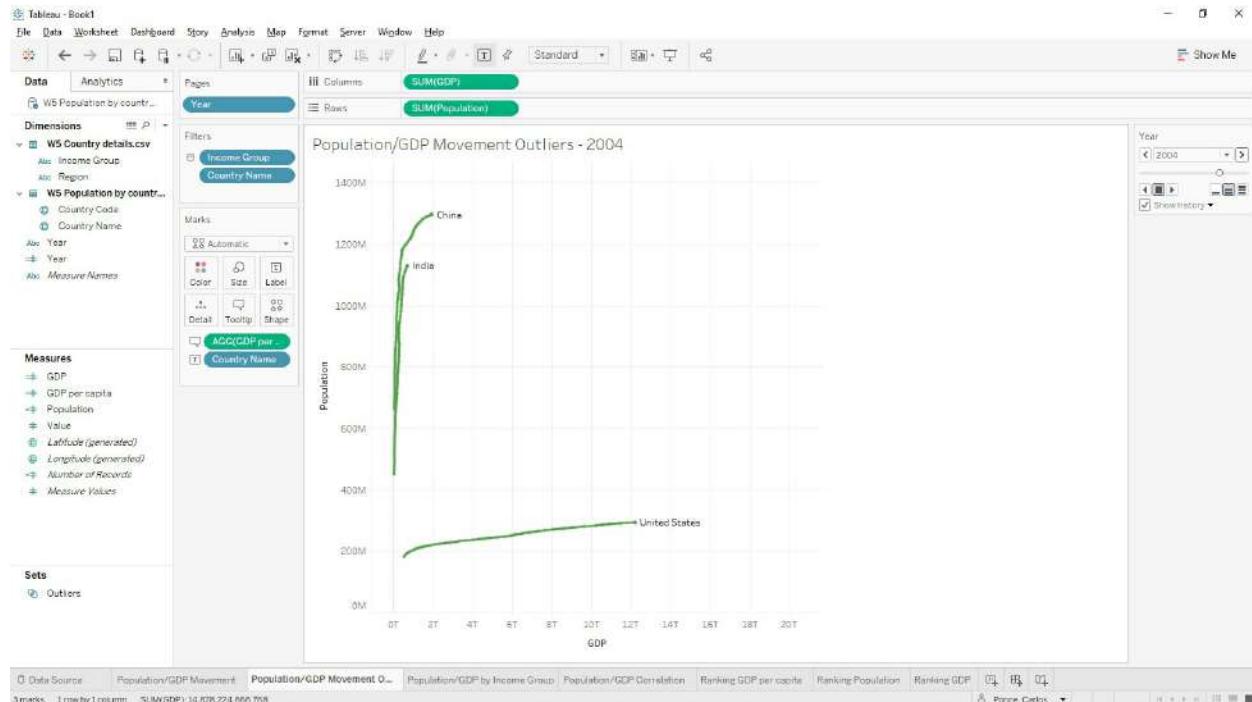
Now show the movement of the dots with **Show history**:



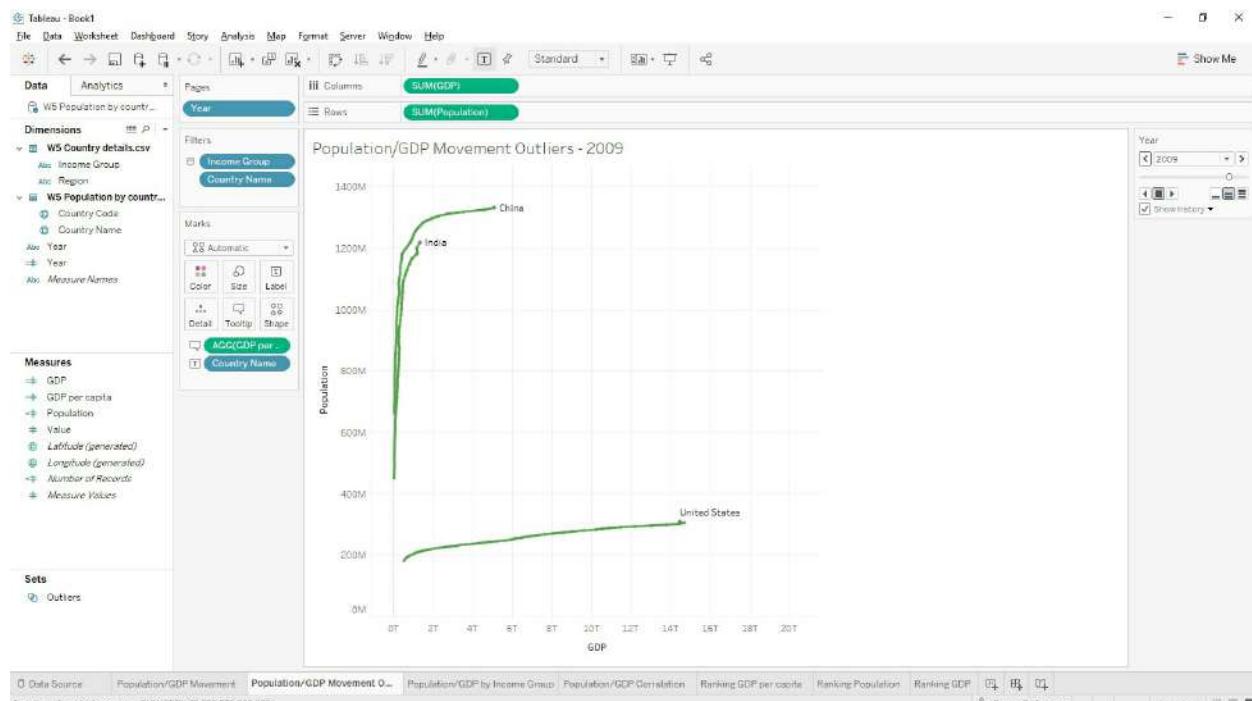
Watch the changes through years and make conclusions.



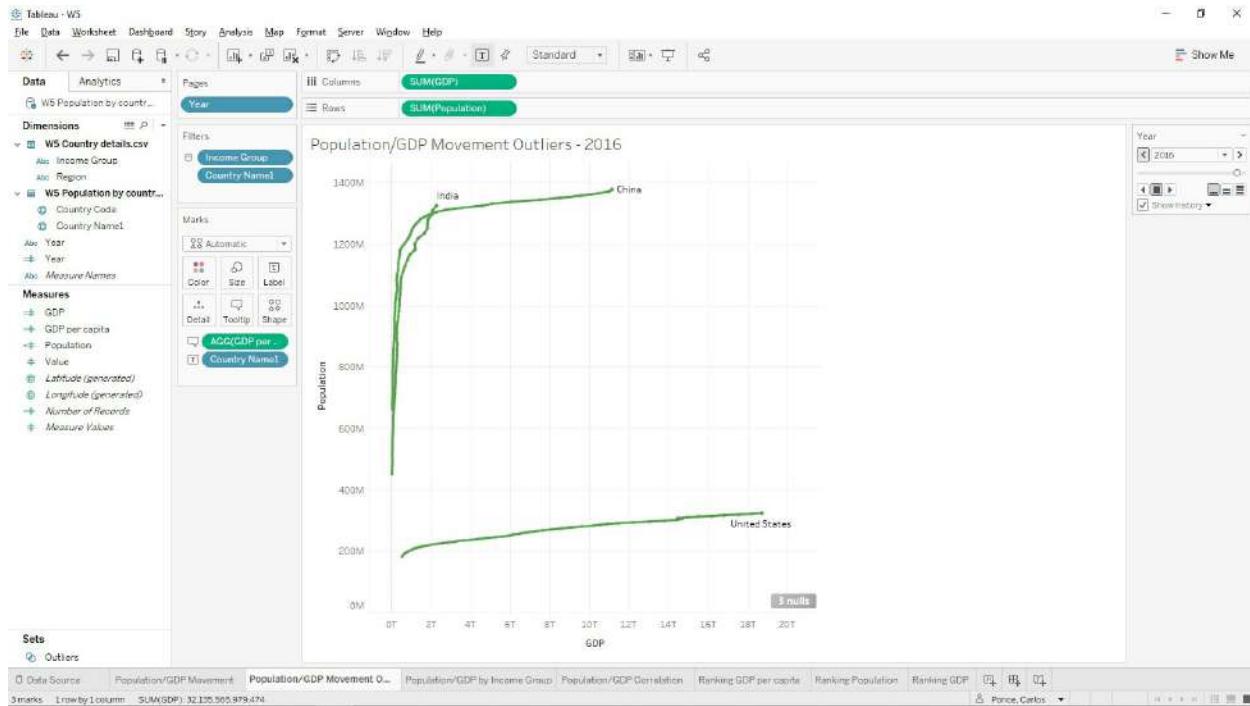
Up until 1993 china and India were not making major advances on GDP, only on population.



On 2004, we see China starts to move the trend.



2009 is the only year USA's GDP went down; the year of the housing crisis.



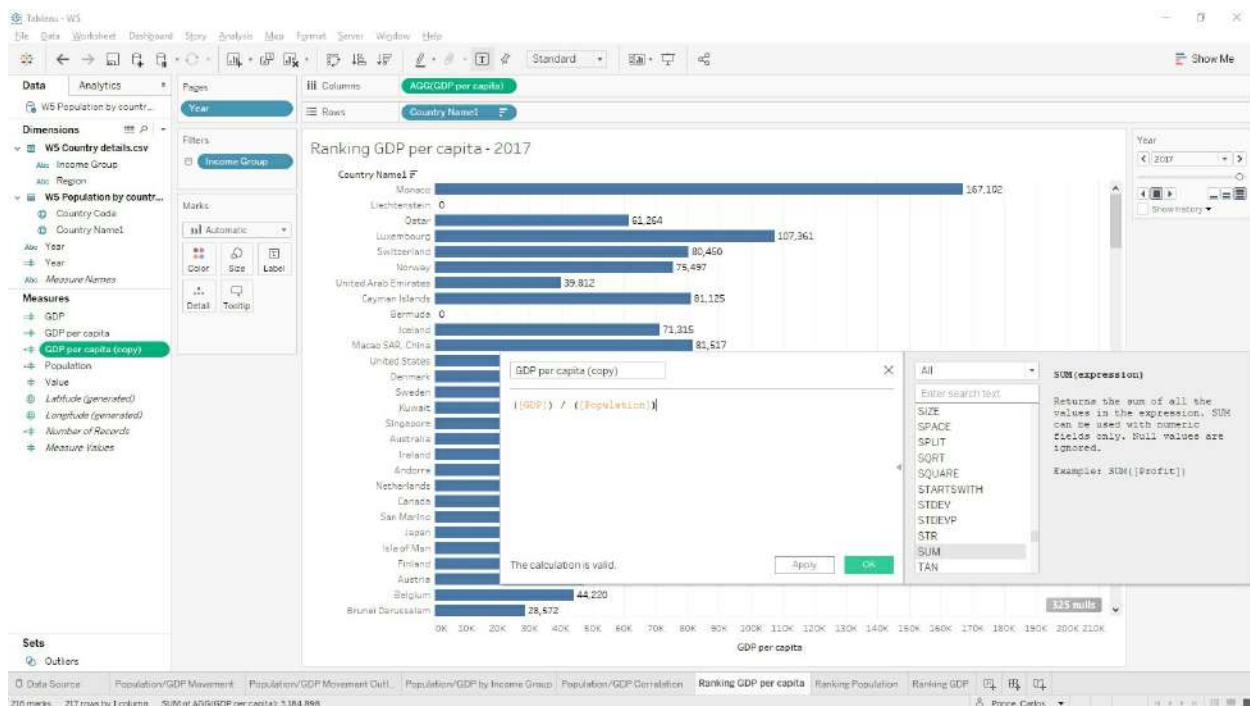
2016 is the year China's GDP growth decelerated. According to web sources, it was due to structural reforms.

This is the value of the Pages shelf; scroll quickly through a dimension to find trends.

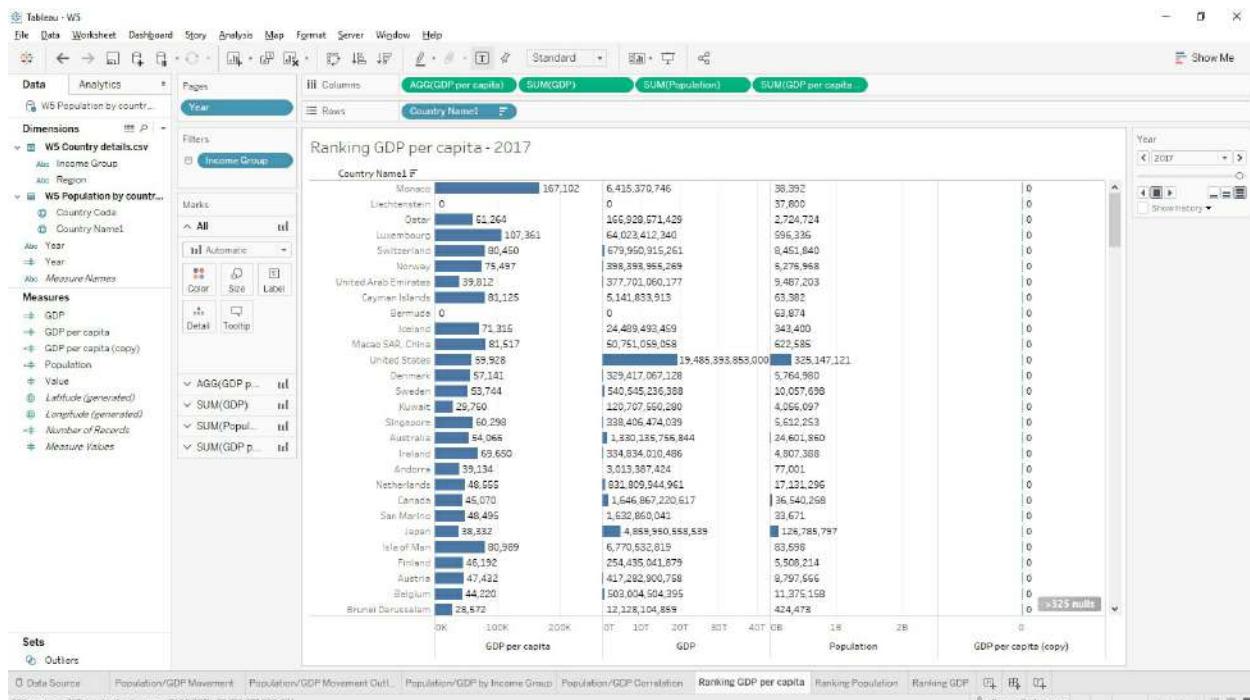
With all the above graphs, we're now able to make a story. But if you wish, you can remake the analysis of some of the sheets to make your own, using what you have learned so far. Maybe look at the Population/GDP movement of your country, or a subset based on a criterion of your preference.

Let's do one final graph. As you saw, *Ranking GDP per capita* has a lot of blank bars because of the missing data; let's fix that by showing instead the average GDP per capita **for the last 18 years** (2000-2018). There's a problem, though; you cannot simply make **AVG(GDP per capita)** because **GDP per capita is already an aggregated calculation** (you can try and see the error for yourself).

One option would be to remove the **SUMs** in the formula so it's a non-aggregated calculation that now happens on the row level, so that you can now aggregate as average, like this:



However, even though our data is at year level as we're displaying, remember the pivots we did on the beginning; our GDP and population numbers are separated in different rows, therefore our calculation is all equal to zero (go do data source and see on each row how if GDP has a number, population is blank, and vice versa):



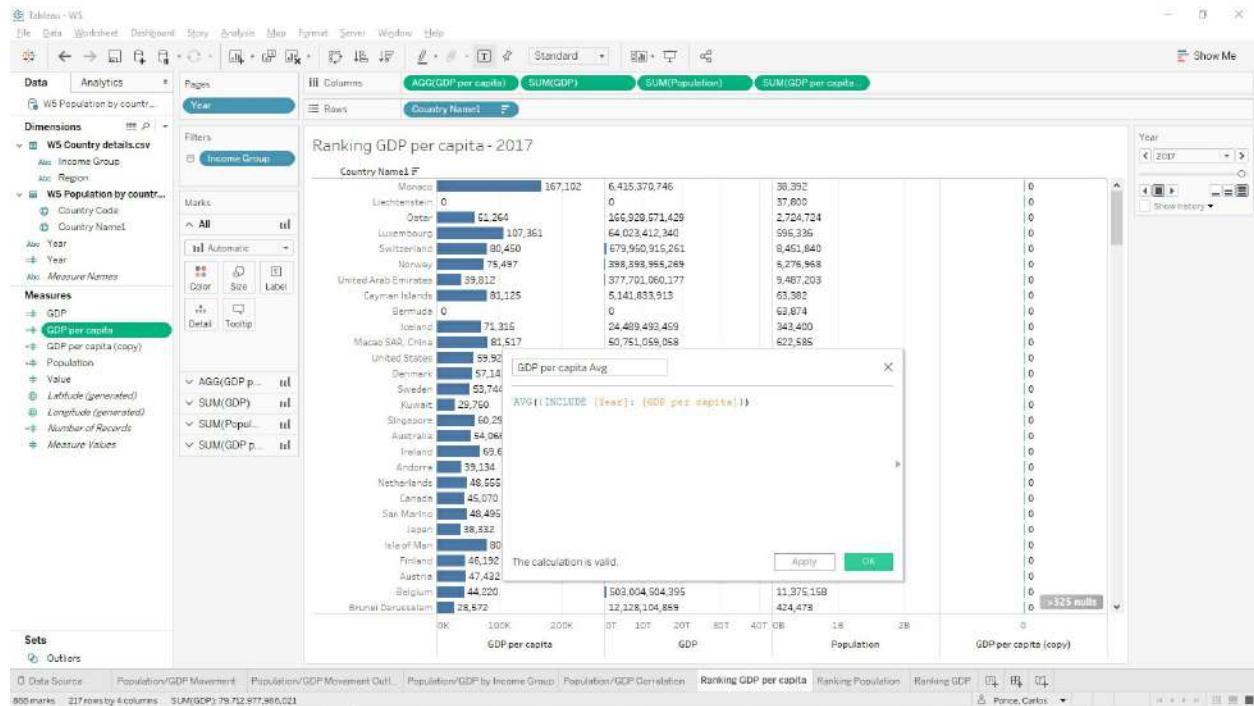
Example:

Country	GDP	Population	GDP per capita
Qatar	166.9M	0	166.9M/0=0
Qatar	0	2.7M	0/2.7M=0

That's why we can't de-aggregate our **GDP per capita** function. So then, is it possible to average an already aggregated calculation?

The answer is yes, through **Level-of-detail calculations**. Create a new calculated field named **GDP per capita Avg**, and paste this formula:

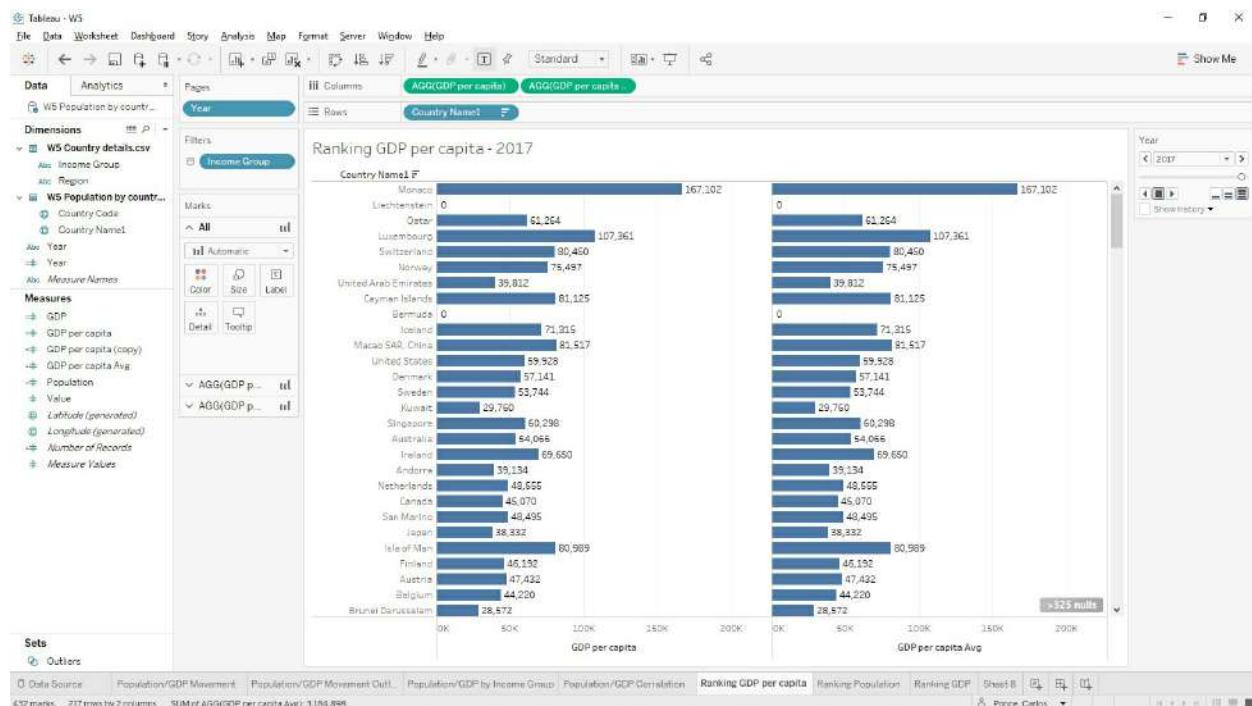
AVG({INCLUDE [Year]: [GDP per capita]}) //This is our Year Clean dimension.



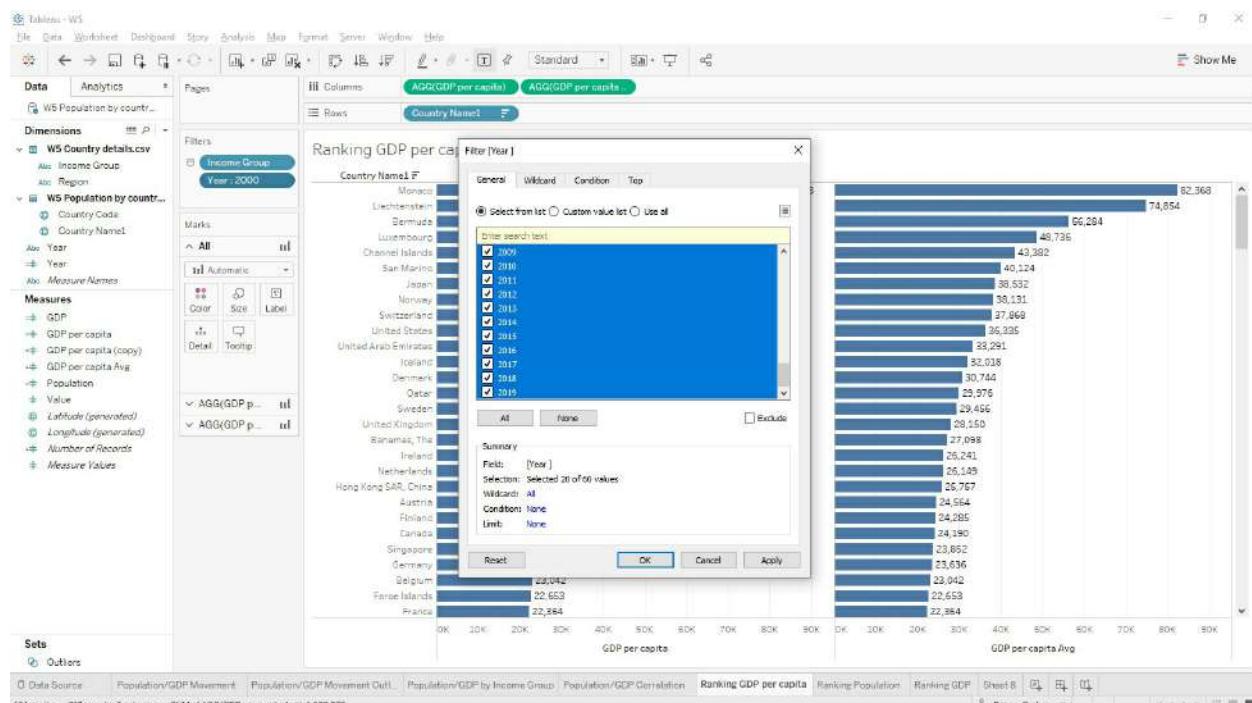
What this formula is doing is:

- 1) Getting the GDP per capita.
- 2) Making an internal array of GDP per capita values per Year.
- 3) Getting the average of this array of values.

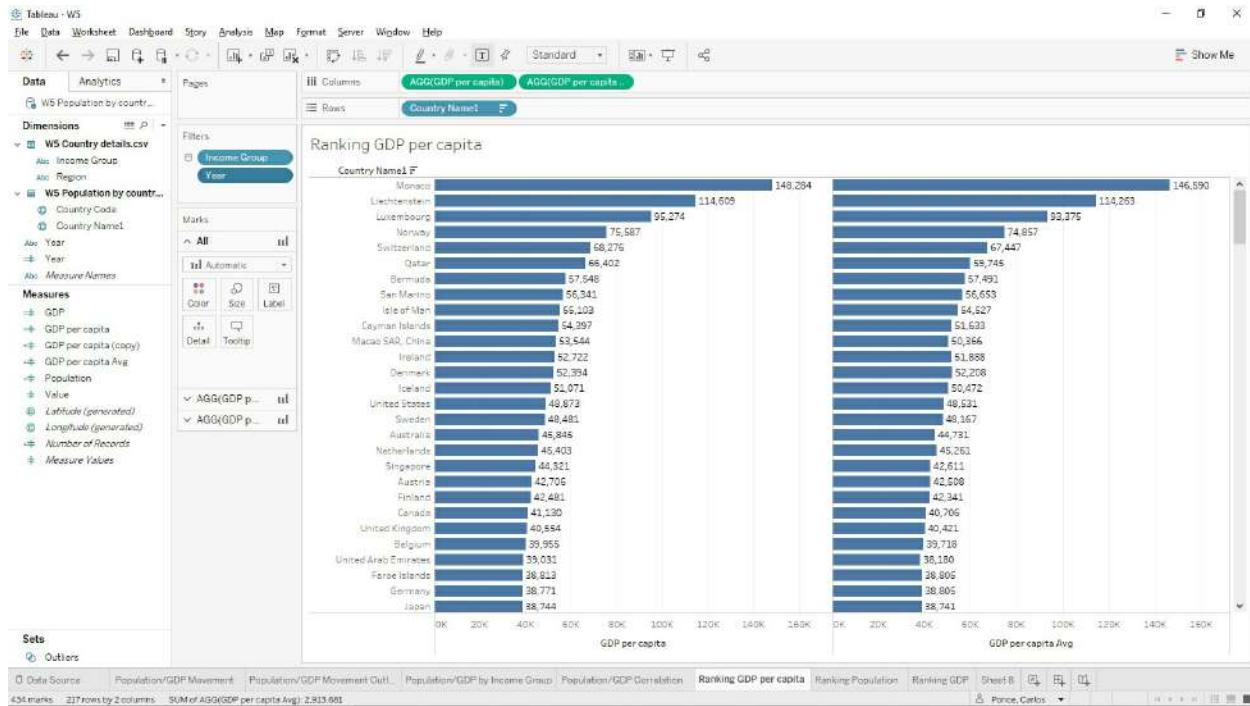
When you drag this to the Columns shelf, you'll notice it gives the same result as the original formula:



This is because we have a **Year** filter on Pages. Move it to Filters and include from 2000 to 2018. You can select years faster if you select the first year (after toggling the checkbox) so it gets a blue highlight, hold shift, and select the last year:



Notice how the values now changed a little:



This is because now, for the years that did not have GDP value for a country, we're not counting the population of that year either.

Visually, it would be like this:

	2001	2002	2003
Qatar GDP	0	5,000	20,000
Qatar Pop	20	30	40

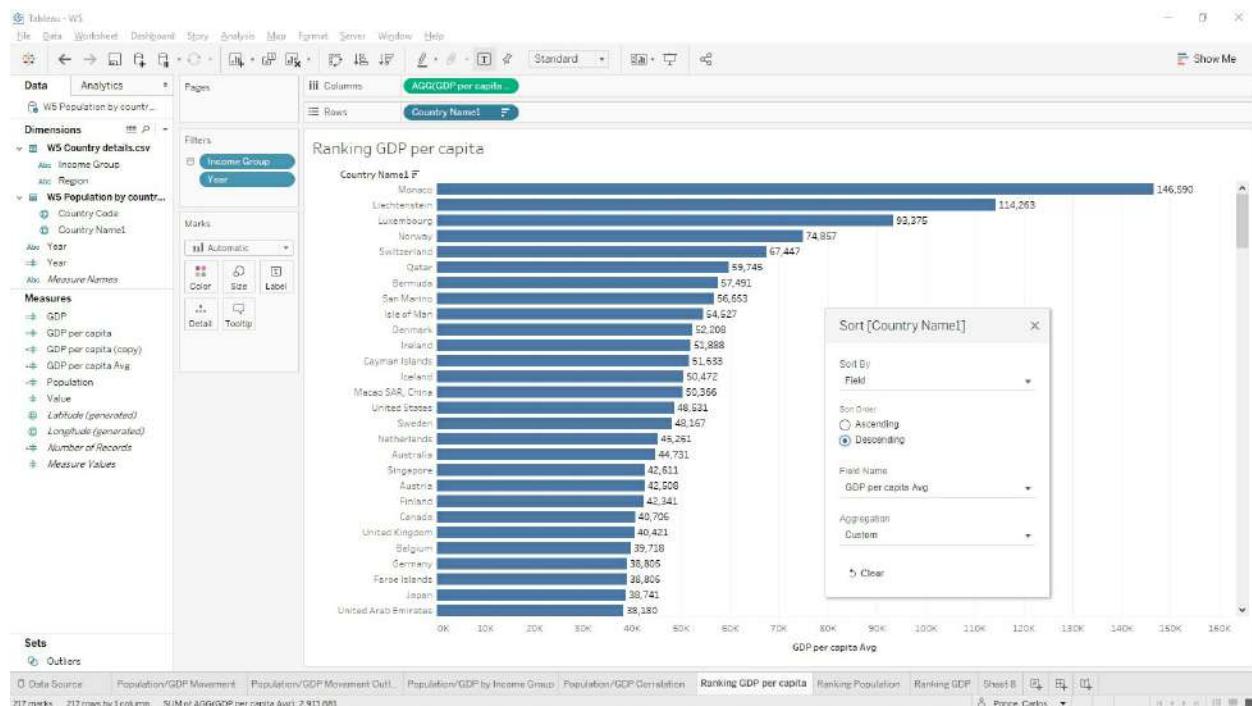
Original formula: $(20+30+40)/(0+5,000+20,000)= 277.7$

LOD formula: $0/20+5,000/30+20,000/40= 666.6$

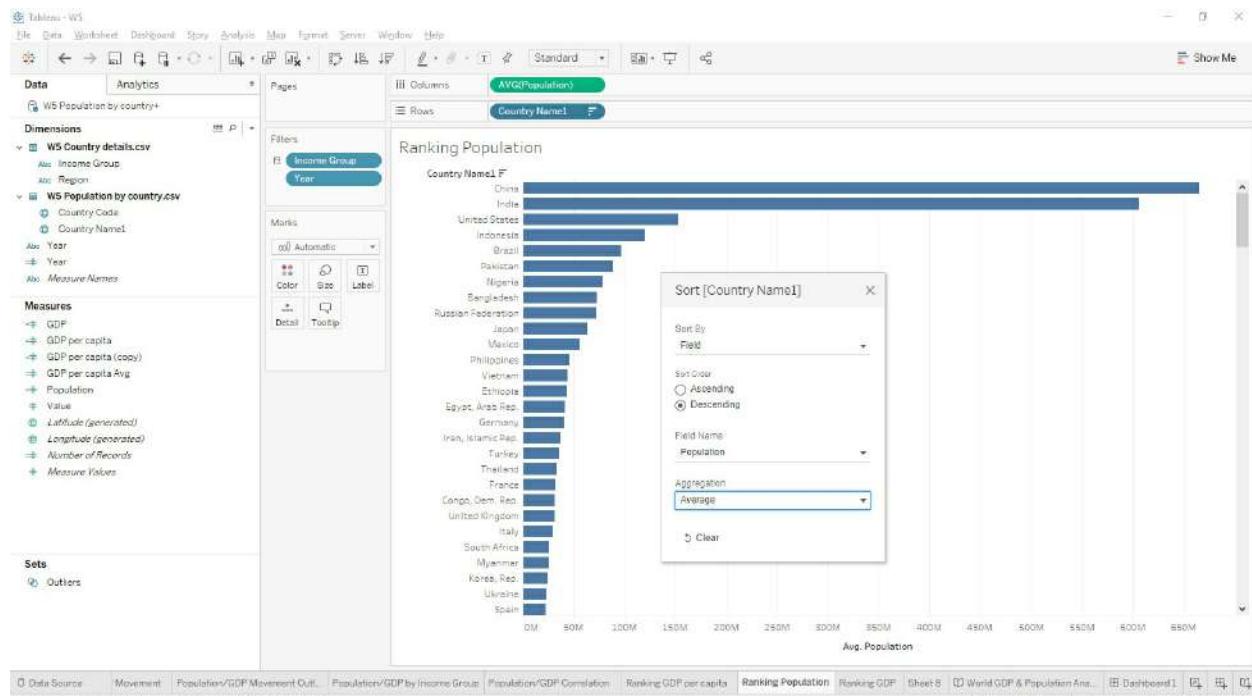
See how originally the aggregated calculation of GDP per capita just sums all the GDP, then divides it by the sum of all the populations, while the LOD formula gets the GDP per capita value per year and makes the appropriate average.

This kind of concepts are important to understand in order to present accurate and reliable data.

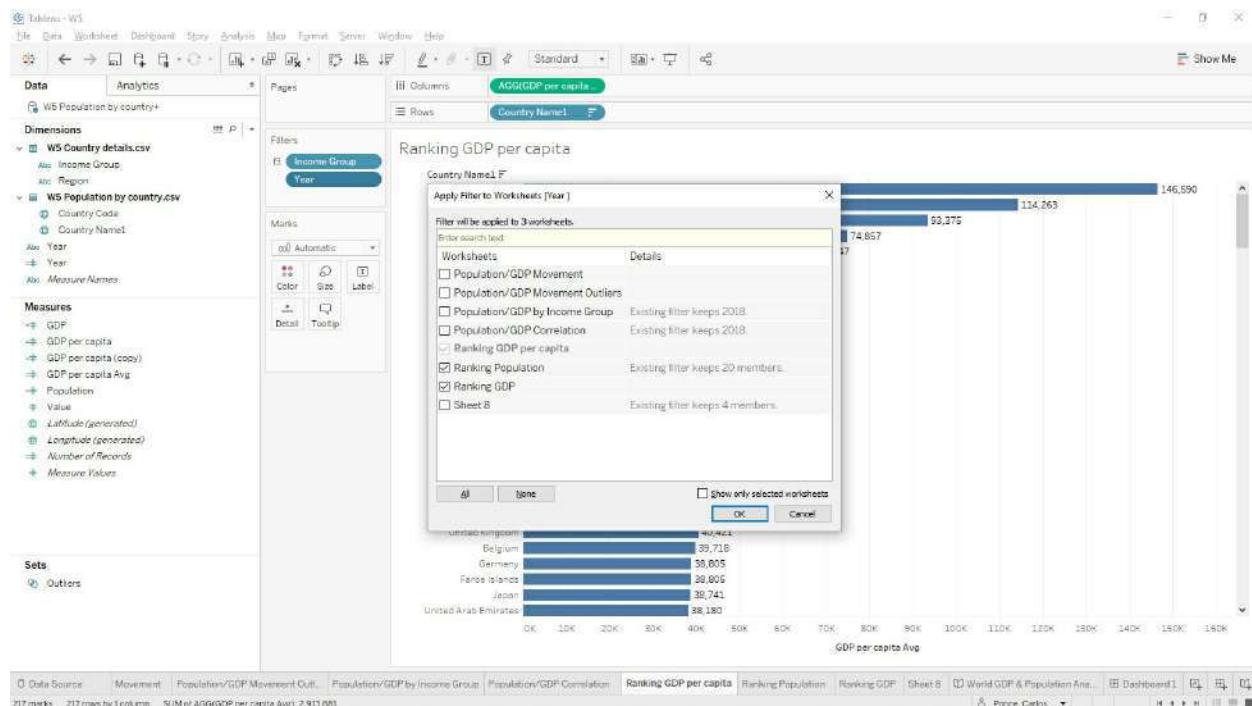
Now on *Ranking GDP per capita* put your new **GDP per capita Avg** on the columns and reapply the sort on **Country Name** by this new measure.



Repeat for *Ranking Population* and *Ranking GDP* as well, but you don't need to make an LOD calculation for these two, just change the aggregation and sort to average:



For cleanliness, you can apply the Year filter from *Ranking GDP per capita* on the other two ranking sheets, so they're aligned. Right click filter > Apply to worksheets > selected worksheets:

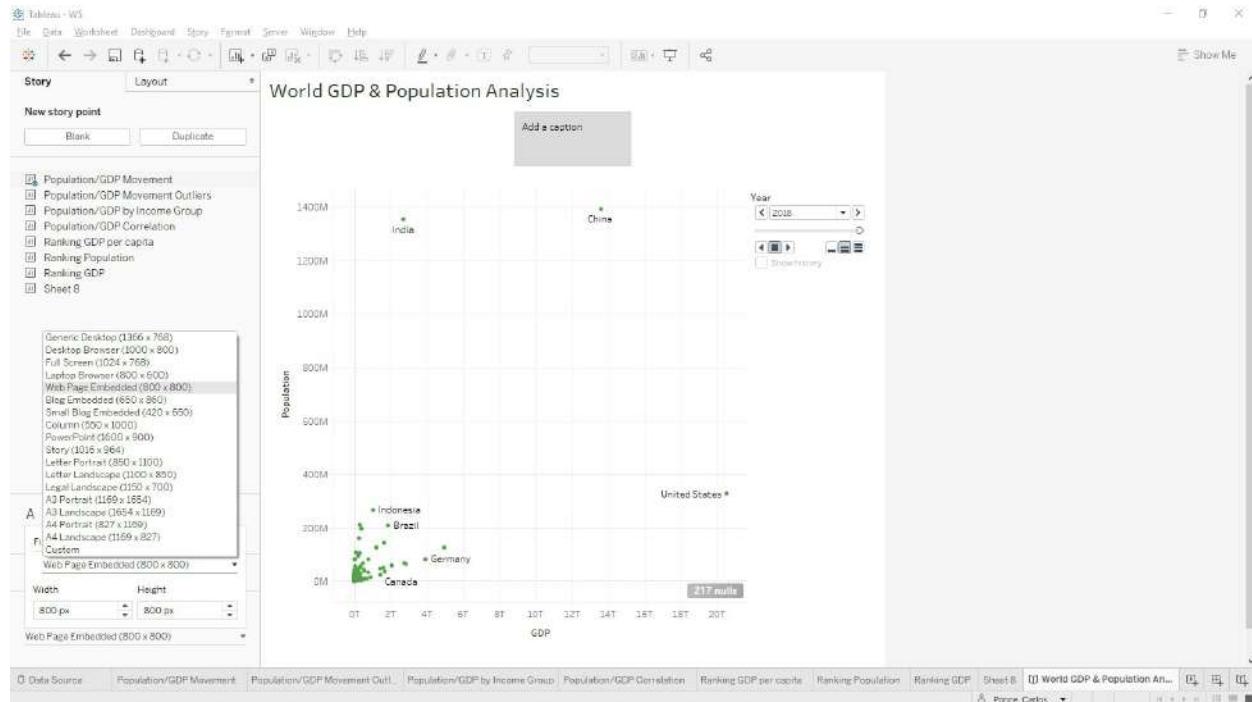


Time to make a story! Press the **New Story** button.

The screenshot shows the Tableau desktop interface with the "Story" feature selected. A new story point named "Story 1" is created. The main canvas area is empty with a placeholder text "Drag a sheet here". On the left side, there is a sidebar listing various sheets: "Population/GDP Movement", "Population/GDP Movement Outliers", "Population/GDP by Income Group", "Population/GDP Correlation", "Ranking GDP per capita", "Ranking Population", "Ranking GDP", and "Sheet 8". There are also checkboxes for "Show title" and "Size" (set to "Story (1020 x 904)").

Name it *World GDP & Population Analysis*.

Drag *Population/GDP movement* to the canvas. Resize the story to **Web Page Embedded** (or any of your choice) by the option on the bottom left.

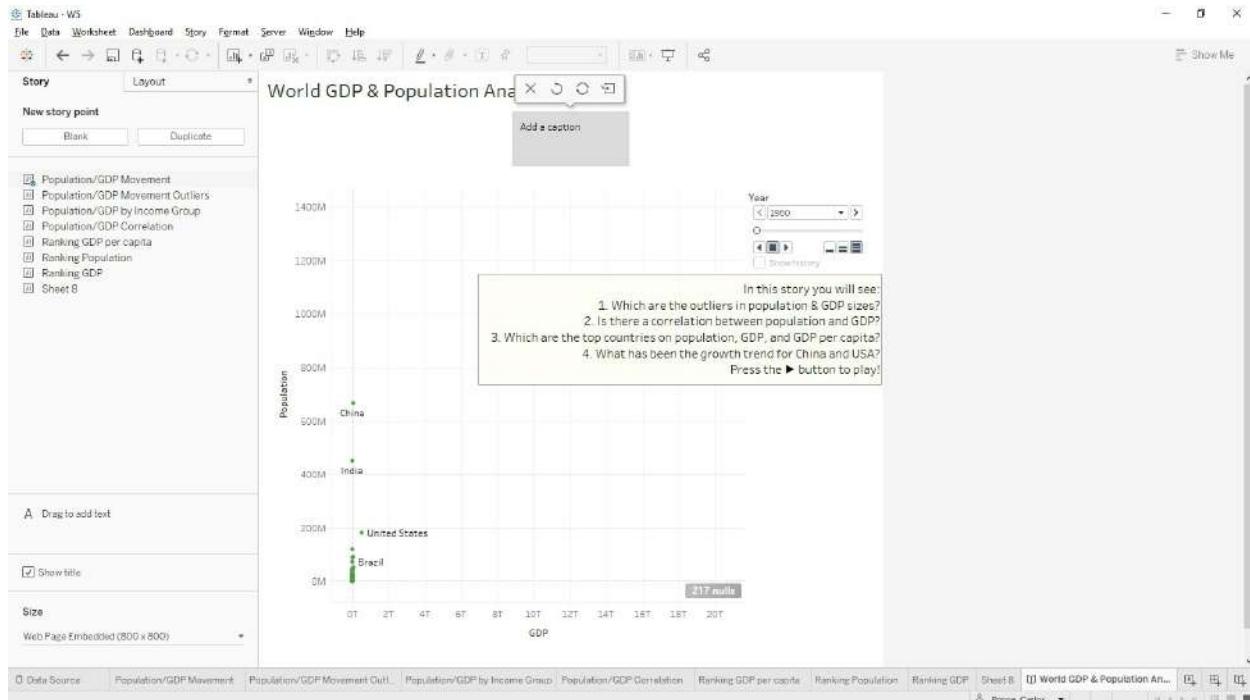


Add the following text:

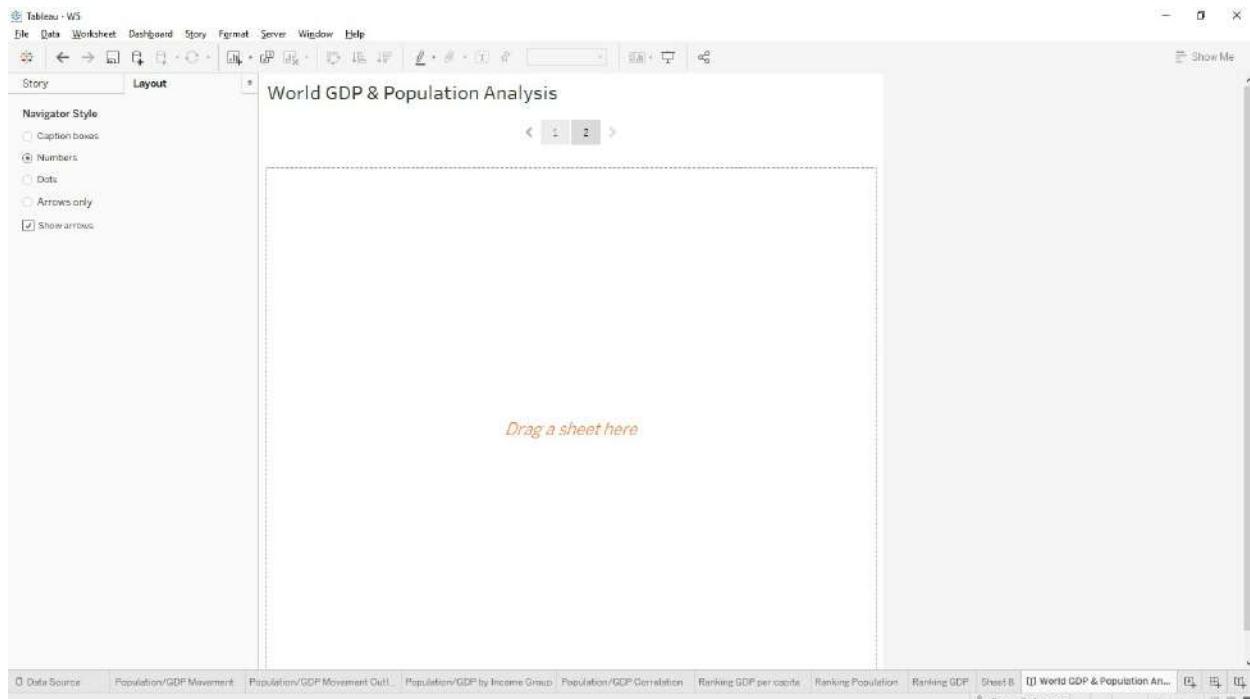
In this story you will see:

1. Which are the outliers in population & GDP sizes?
2. Is there a correlation between population and GDP?
3. Which are the top countries on population, GDP, and GDP per capita?
4. What has been the growth trend for China and USA?

Press the ► button to play!



Make a new **blank** story point, now go to the **Layout** tab and choose **Numbers**.

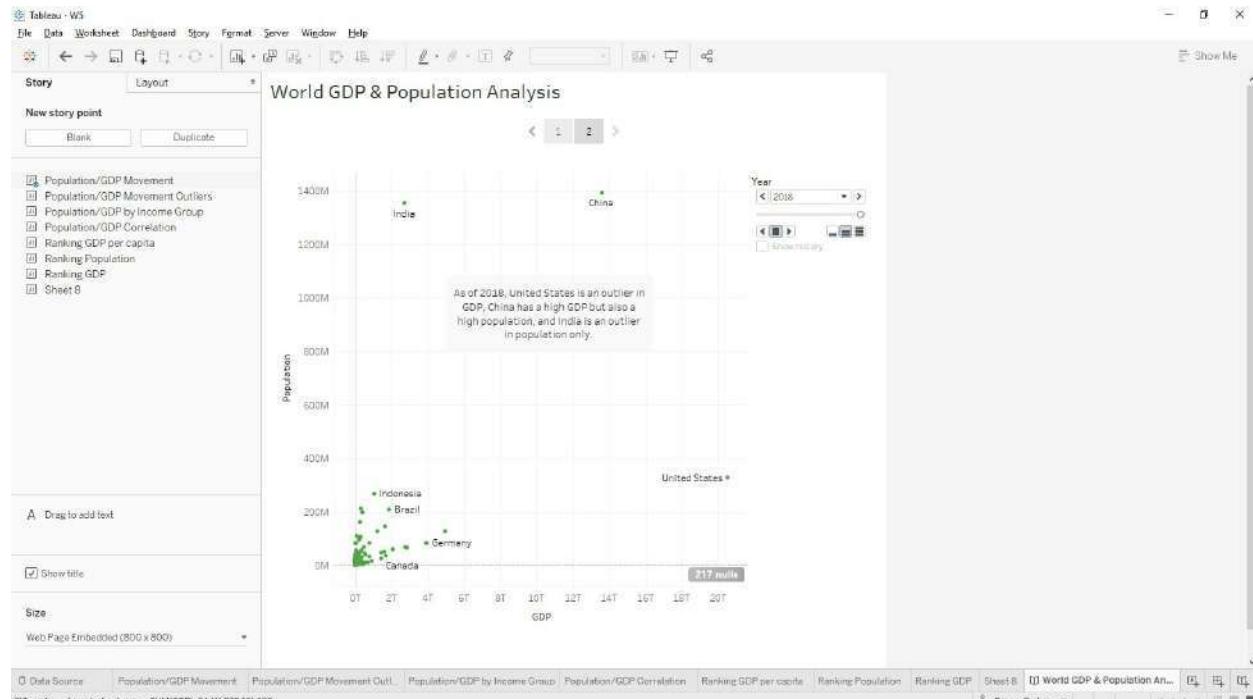
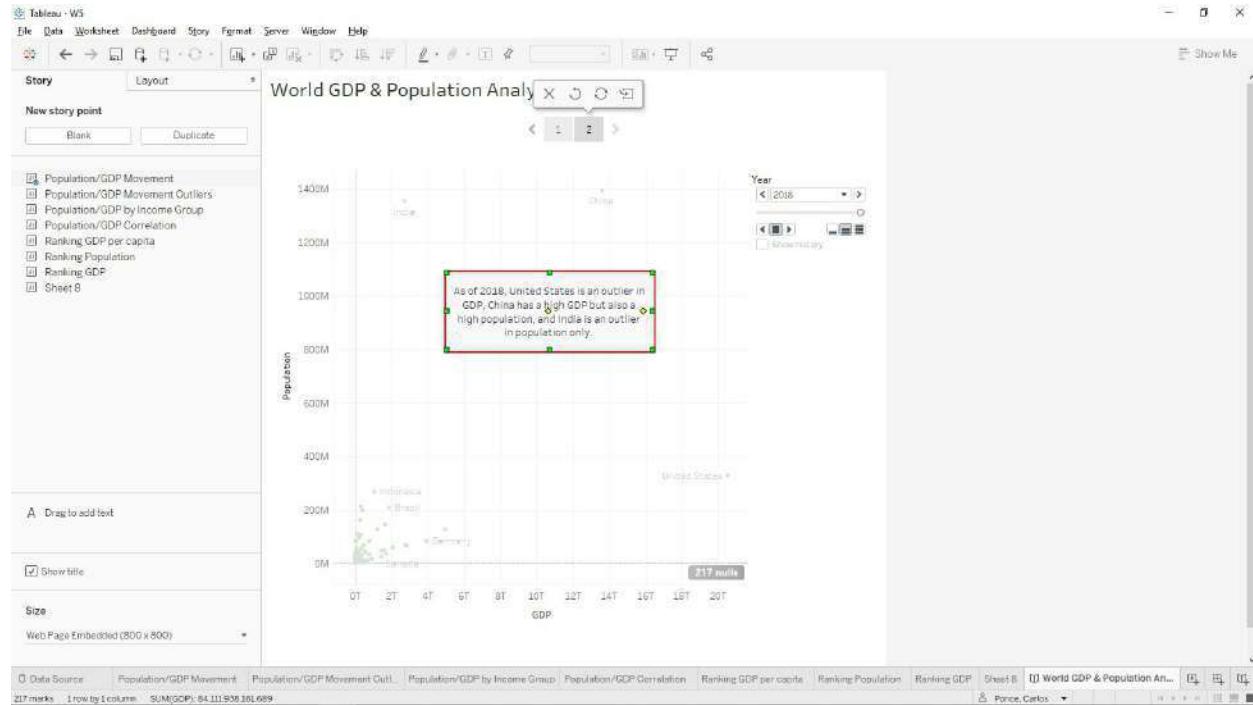


Drag the same sheet again, right click on the United States point, Annotate > **Mark**, and paste this:

As of 2018, United States is an outlier in GDP, China has a high GDP but

also a high population, and India is an outlier in population only.

Click the new textbox to modify size and position.

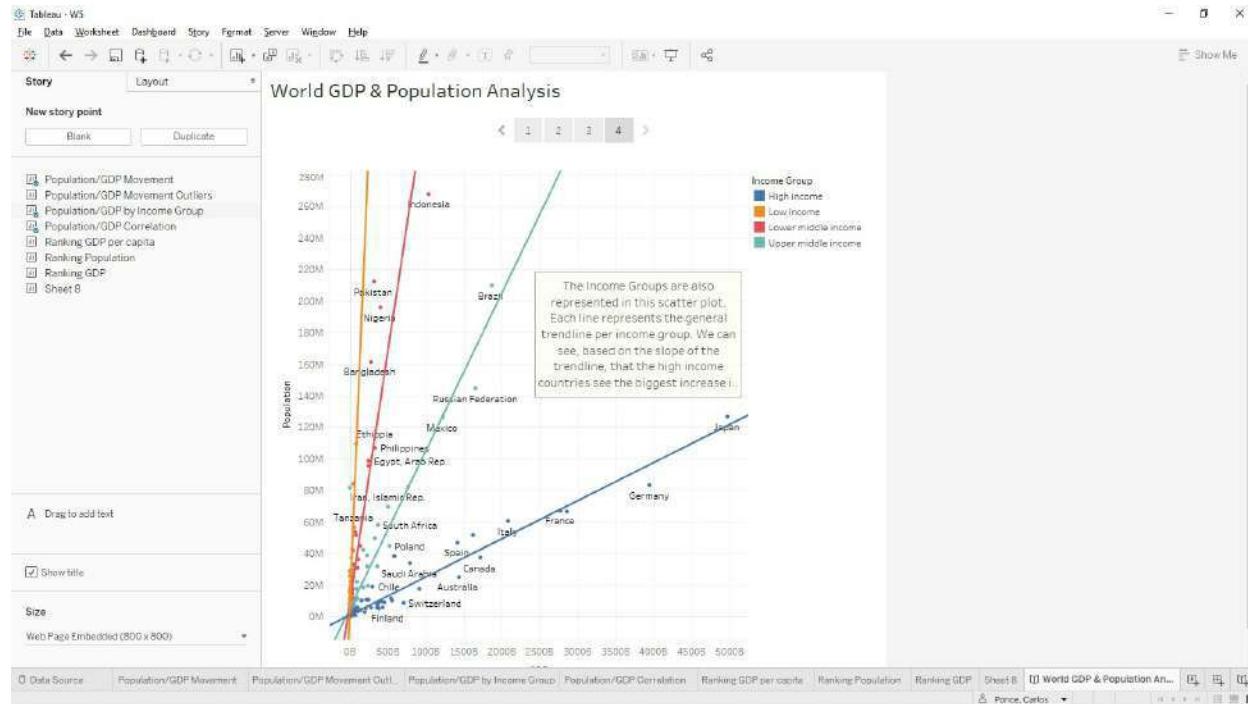


New story point, drag second sheet, drag a new text that says:

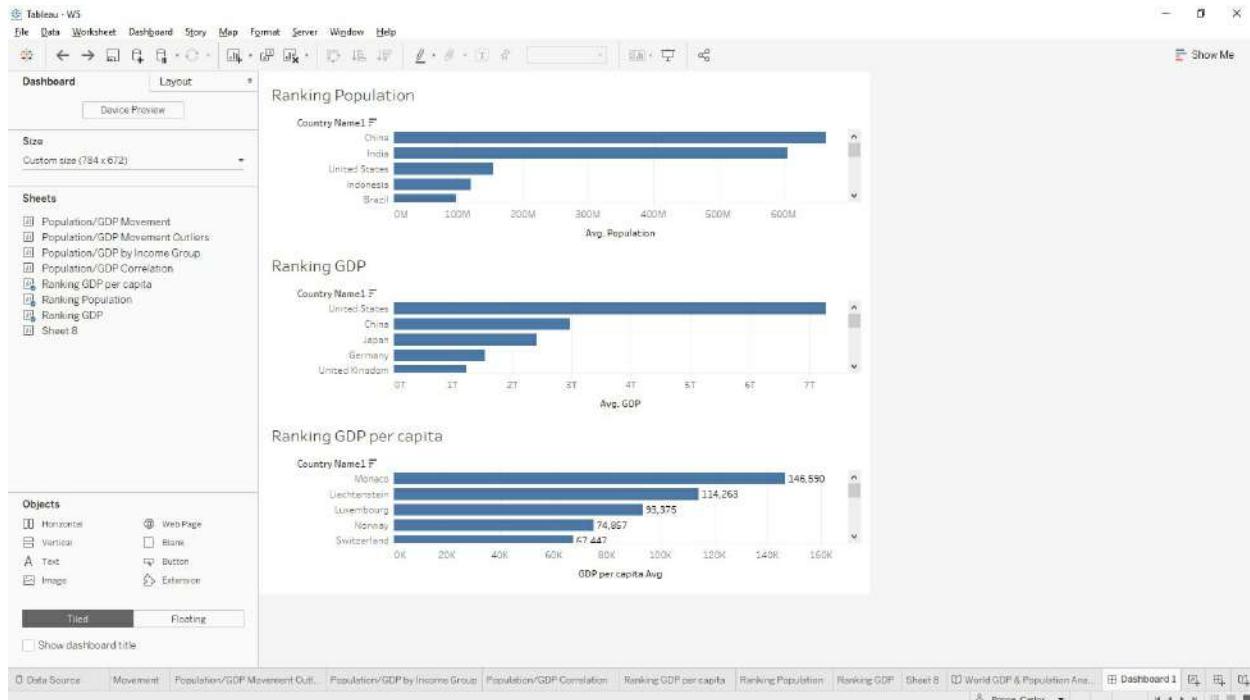
“There was found a low correlation between population and GDP in general. Removing the three outliers (uncheck “out”) made no effect on it.”

New story, drag third sheet, drag new text:

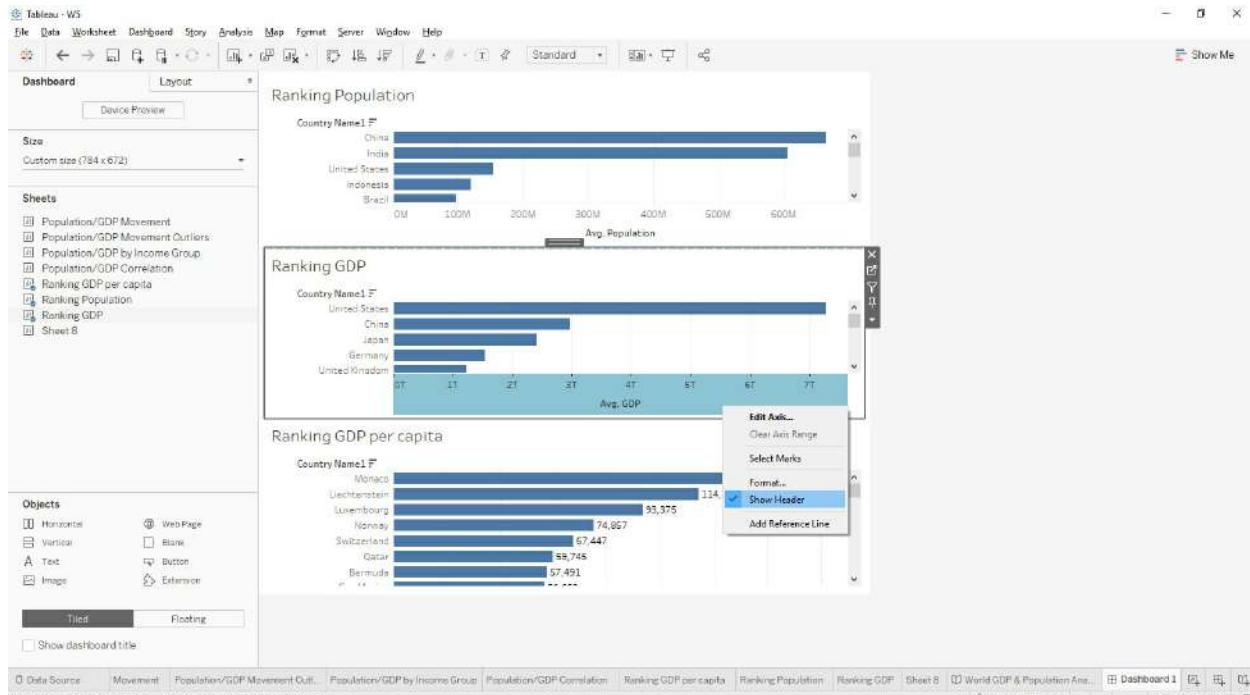
“The Income Groups are also represented in this scatter plot. Each line represents the general trendline per income group. We can see, based on the slope of the trendline, that the high-income countries see the biggest increase in GDP per capita with the growth of population.”



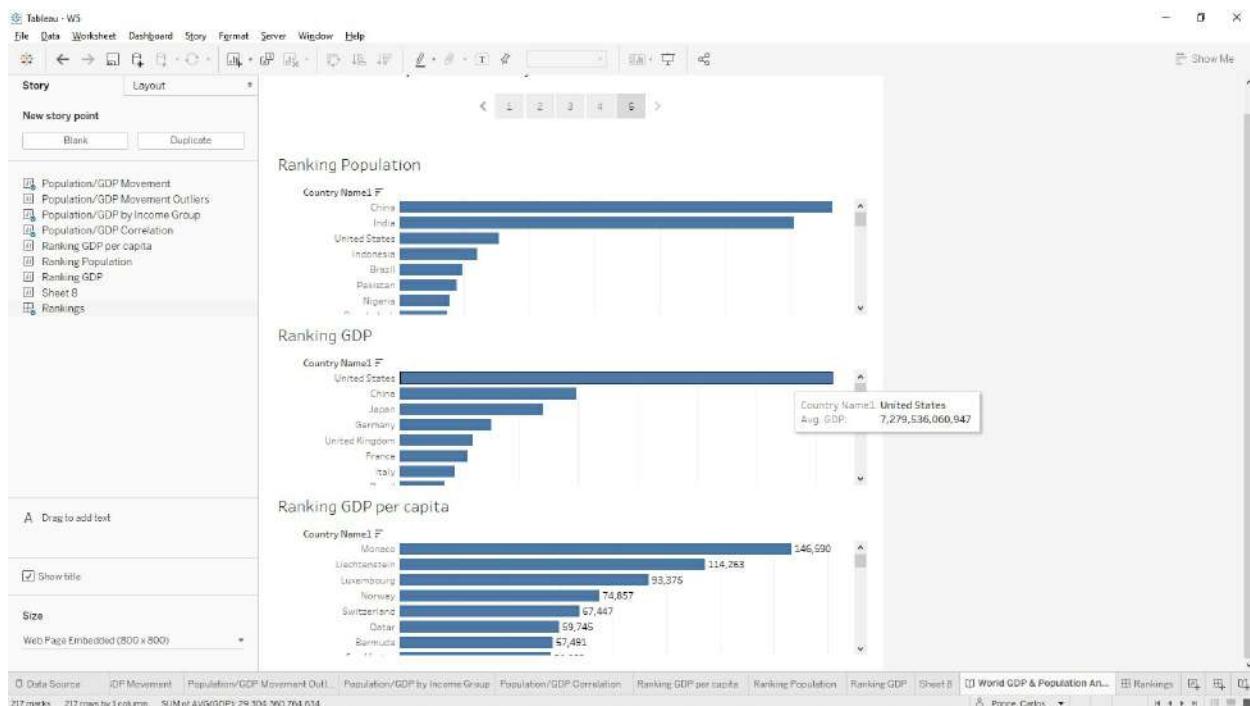
For the next story point, let's create a new dashboard. Add just a vertical container and your three ranking sheets in the order: population, GDP, and GDP per capita. Change the dashboard size to fit the story we're making (custom size).



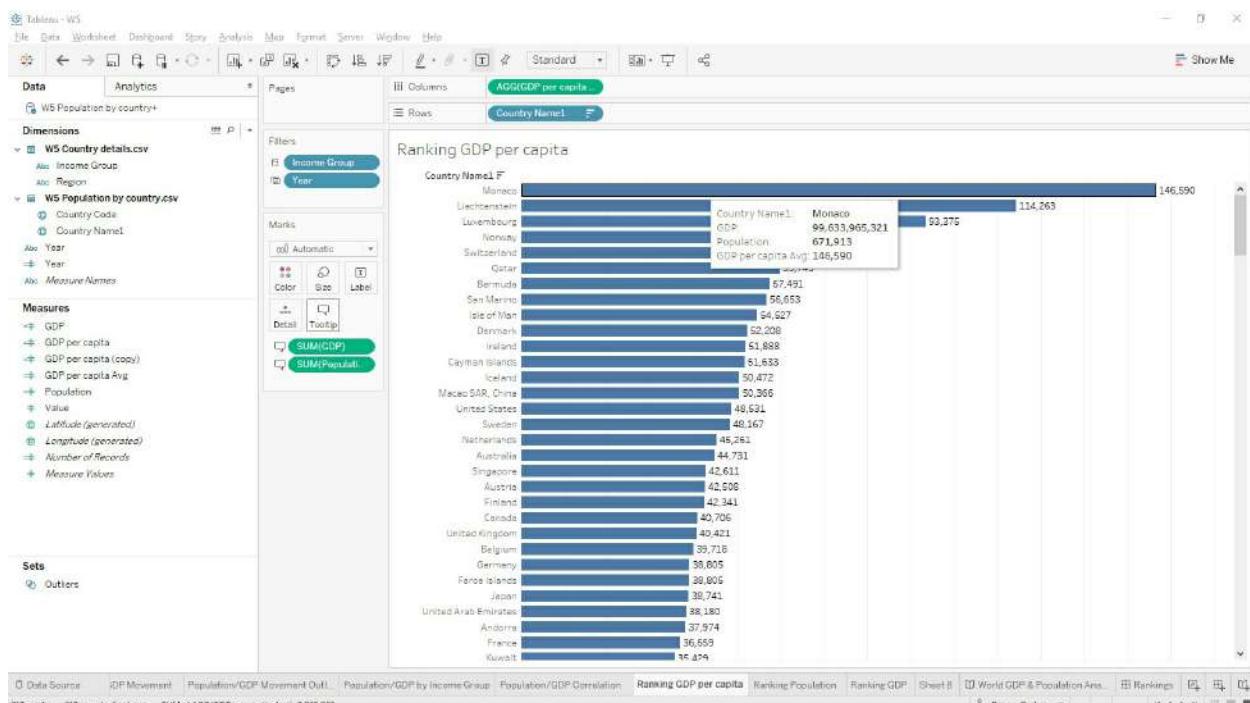
For the three graphs, right click the x axis and click **Show Header** so we hide it.



Distribute contents evenly on your Vertical just in case it's not good, rename this dashboard to *Rankings*, and show it as the next story point.

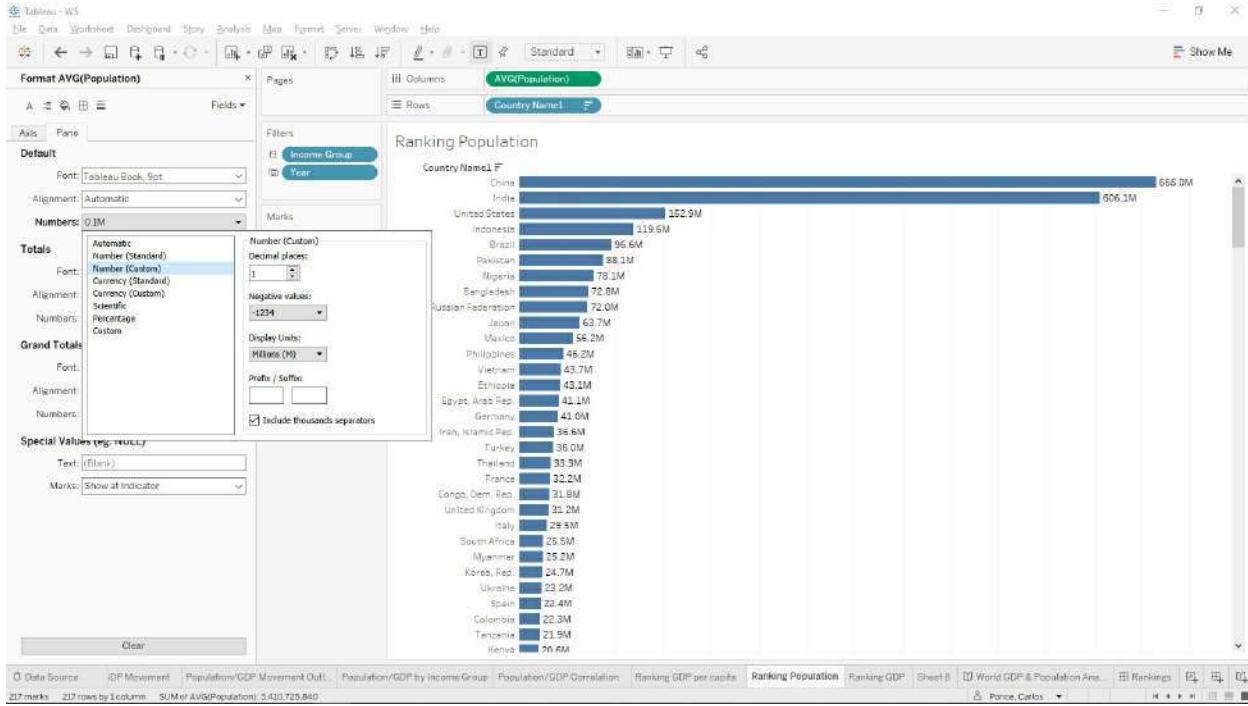


Add on the tooltip of each graph the three measures, so we can compare (remember not to use GDP per capita but **GDP per capita Avg**). Just head to each ranking and drag the two missing measures to the Tooltip property of Marks.



While you're at it, also show the numbers for all graphs clicking on Label > Show mark labels.

The numbers for population and GDP look too big, let's put them on millions. On *Ranking Population*, right click the measure on the columns shelf, click **Format**, go to the **Pane** tab, and under **Default** set the format to be **Number (Custom)** in millions and with one decimal place.

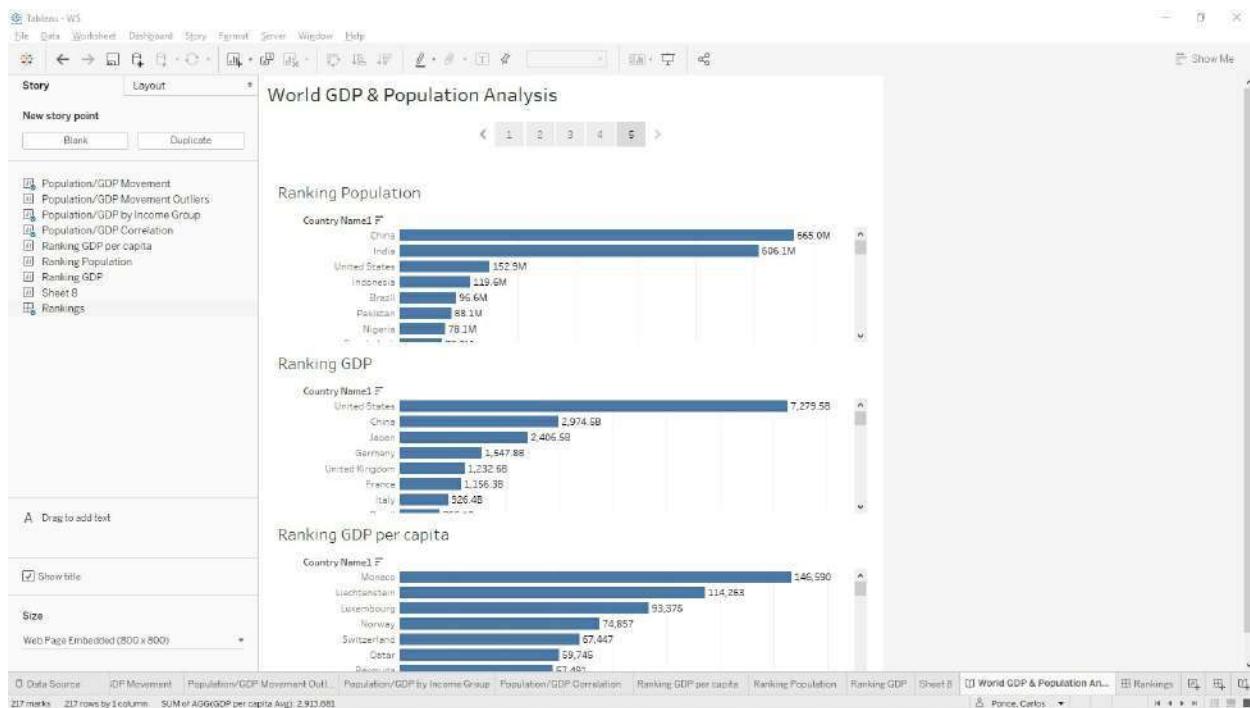


Repeat for *Ranking GDP* but put that on billions instead.

Tip:

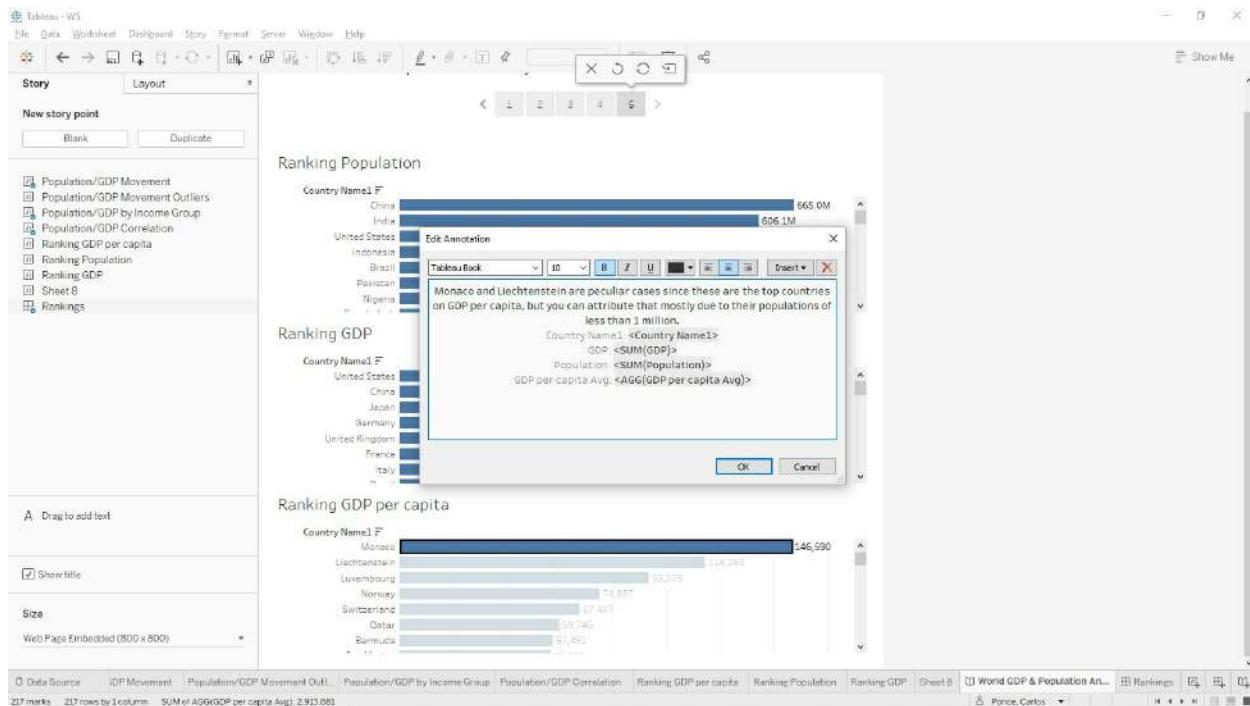
If you need to put in millions but without the “M”, or do any special format, you can go to **Custom** where you’ll be able to fully edit the format as in Excel. You can learn more about this [here](#).

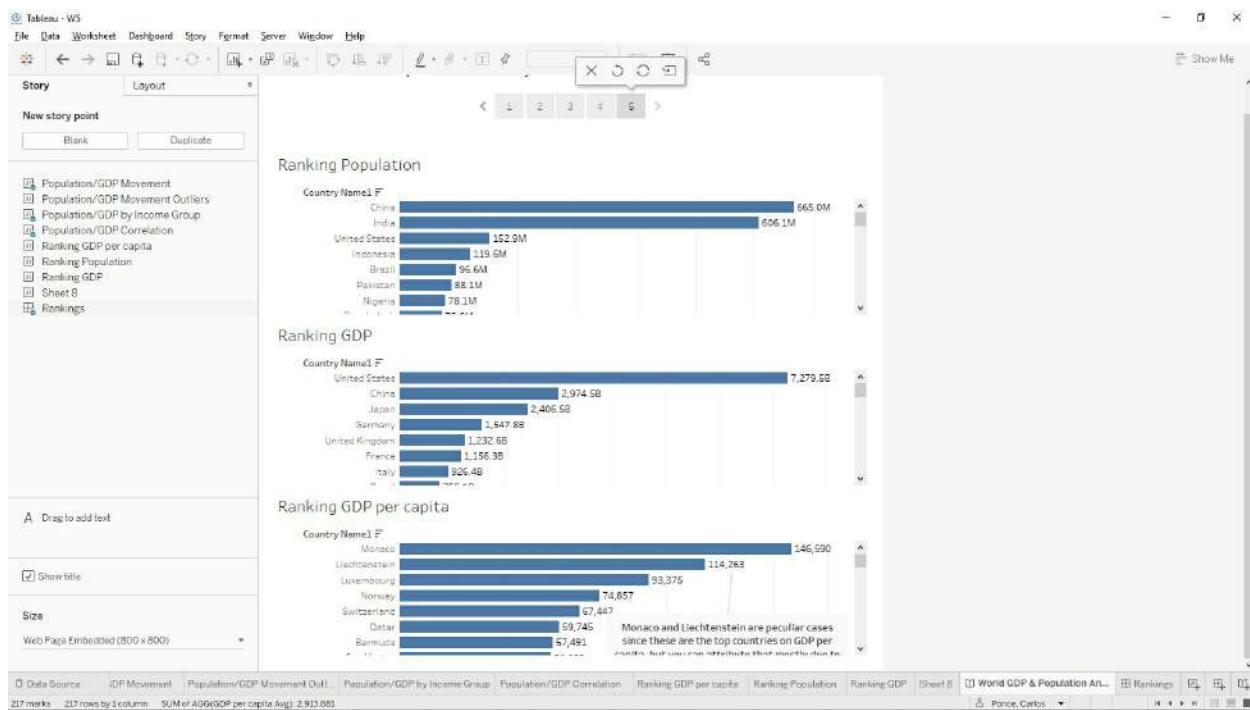
Below the result:



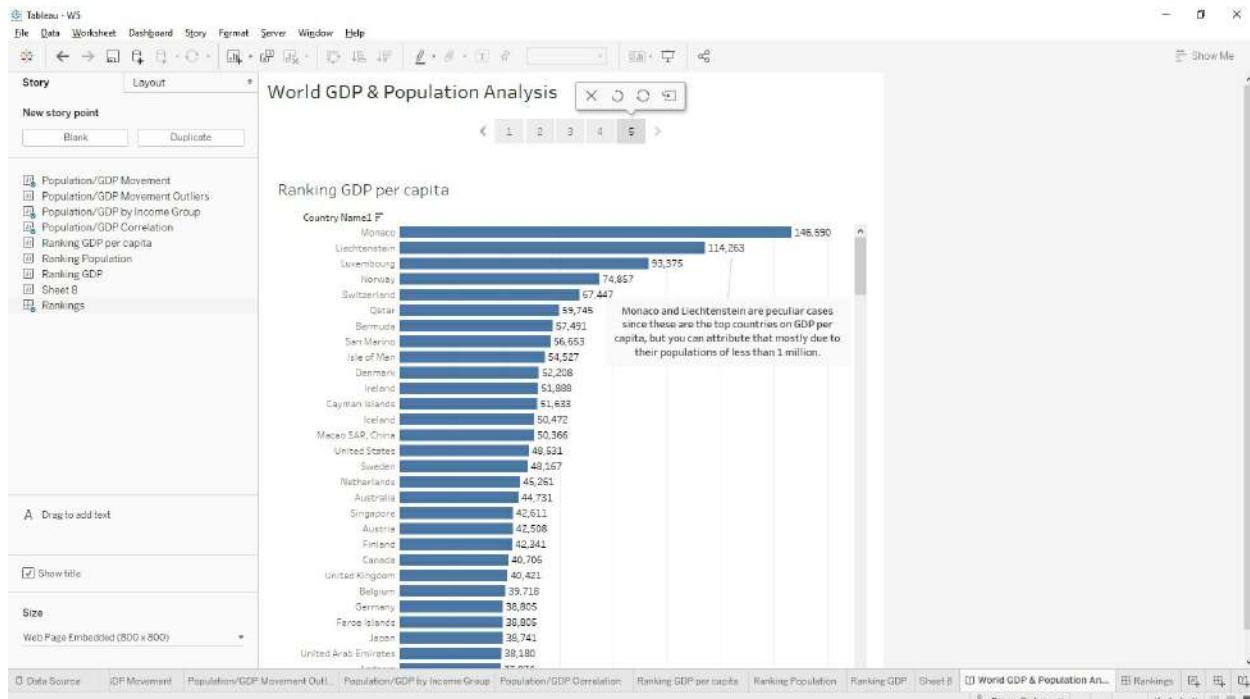
Right click Monaco on *Ranking GDP per capita*, Annotate > **Mark**, remove the automatic text and paste this:

“Monaco and Liechtenstein are peculiar cases since these are the top countries on GDP per capita, but it is due to their populations of less than 1 million.”





The annotation is not very visible. Given that we are not really looking at anything on the population and GDP graphs, remove them from the dashboard so you can leave more space for *GDP per capita* and the annotation.

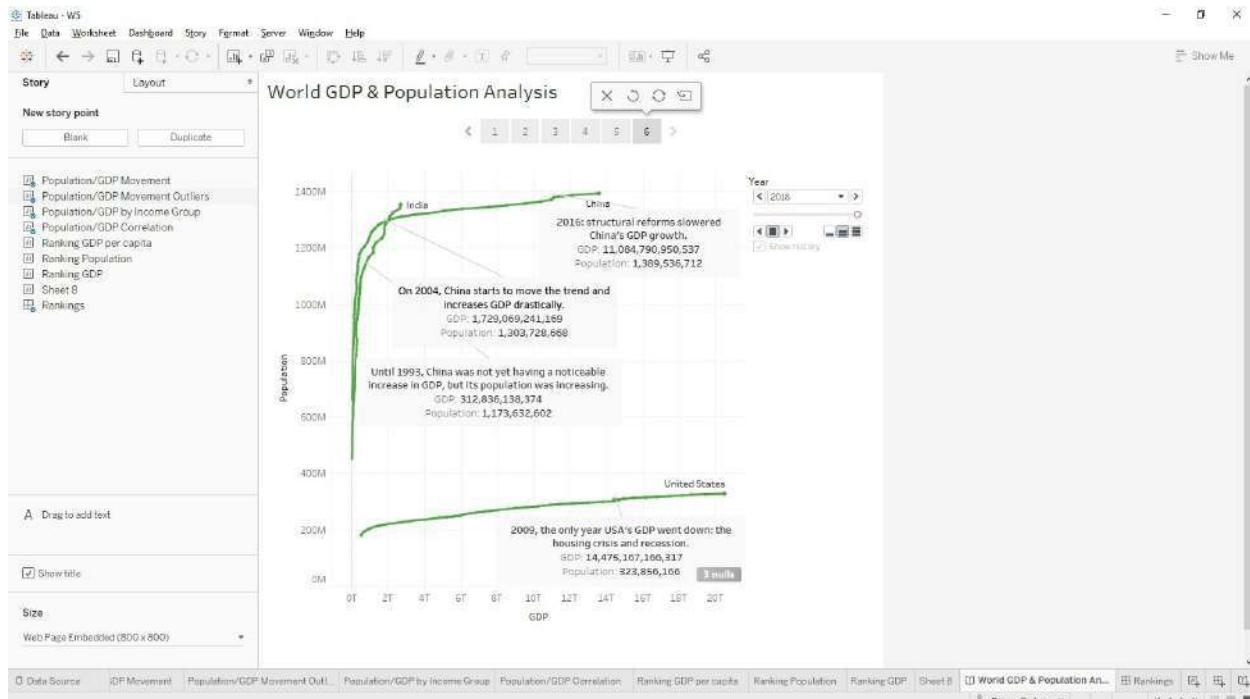


For the last story point, add *Population/GDP Movement Outliers*. Let's add the four insights we commented previously from that graph. Move the Pages

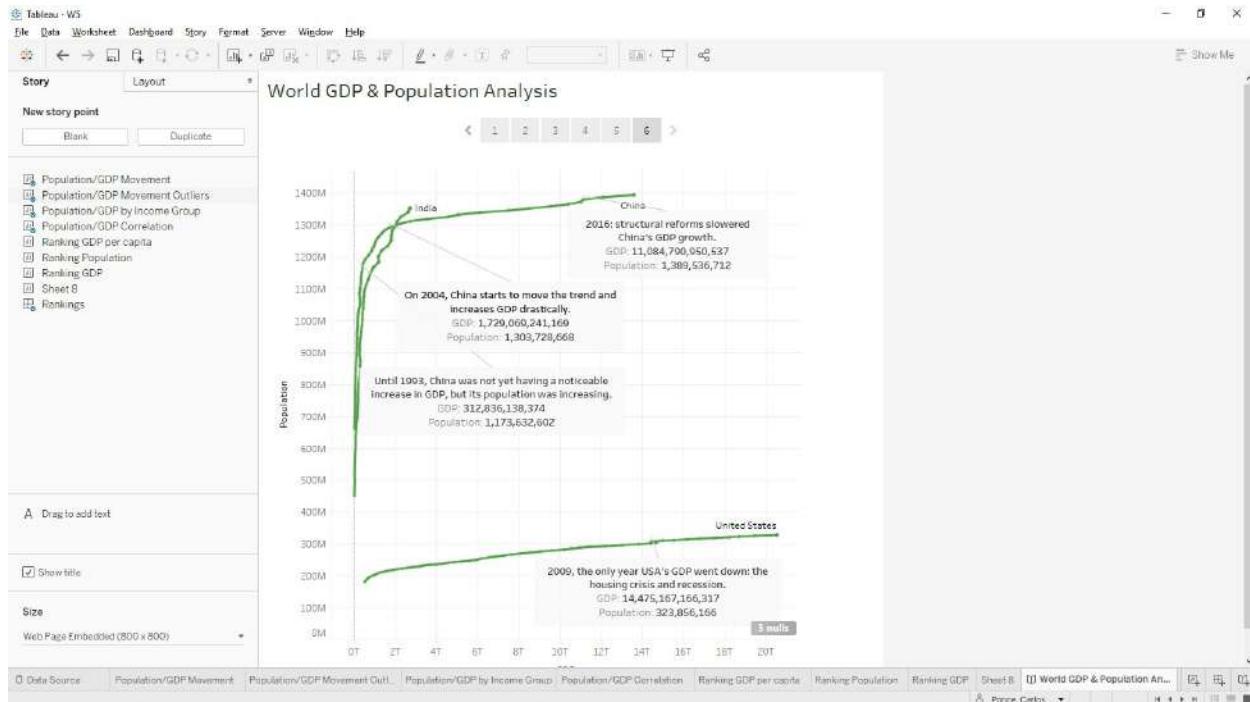
shelf value to 2018, right-click on each of the below points, Annotate > **Point**, and paste the comment included:

Point:	Annotation:
China, 1993	Until 1993, China was not yet having a noticeable increase in GDP, but its population was increasing.
China, 2004	On 2004, China starts to move the trend and increases GDP drastically.
USA, 2009	2009, the only year USA's GDP went down: the housing crisis and recession.
China, 2016	2016: structural reforms lowered China's GDP growth.

Find below the end result:



Hide the **Year** Page card from the sheet itself, so it disappears on the story as well.



Congratulations, fifth workout done! With this, you have mastered all you need to know on Stories. Also, you have learned:

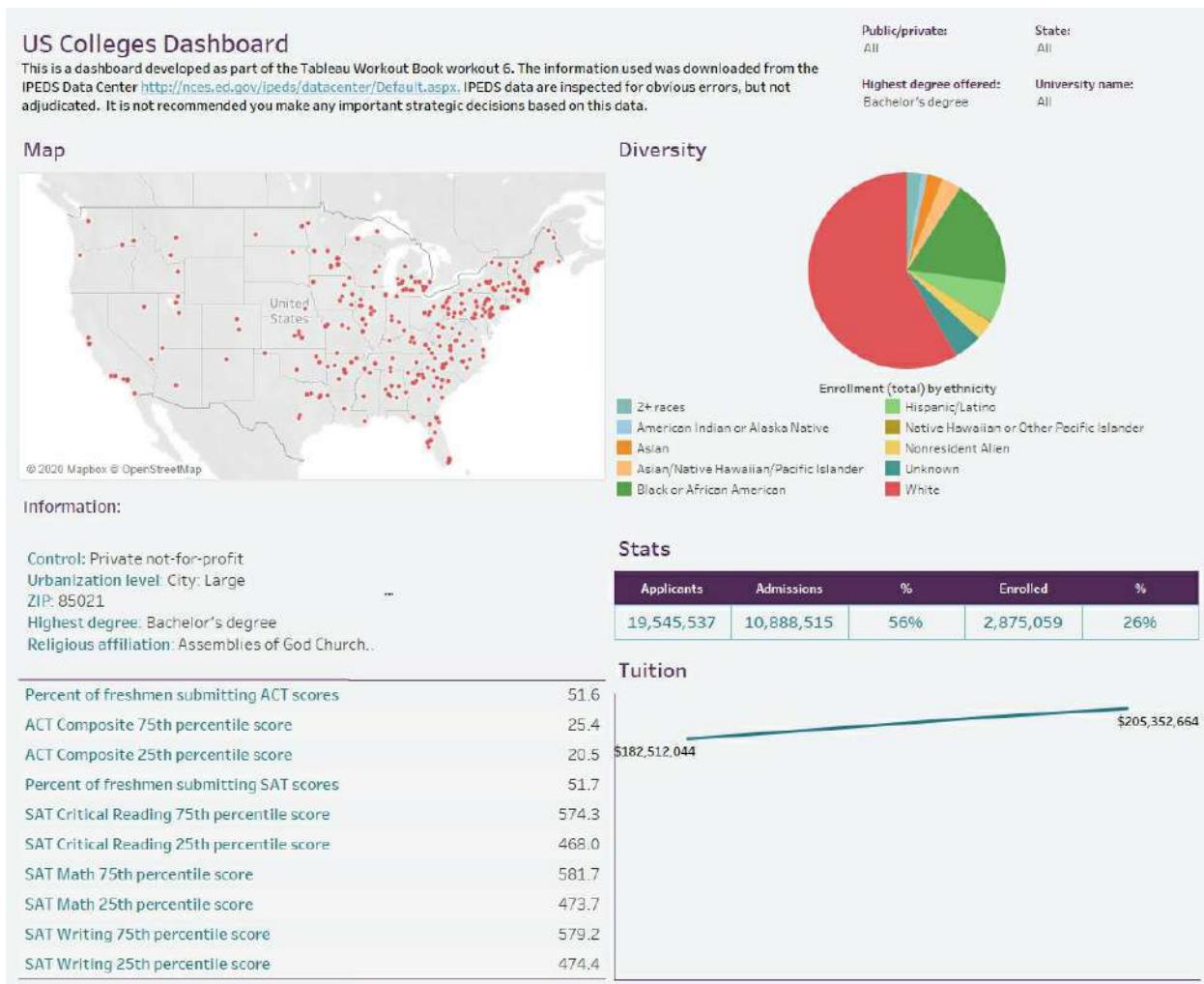
- Pivoting data sources
- Validating a proper data join
- Type conversion functions
- The Pages shelf
- Sets
- Linear Trendline
- Level-of-detail calculations: INCLUDE
- More on customizing and annotating stories
- Custom number formats

You're now halfway done through Tableau Workout Book! Let's go to the second half:

Workout 6. USA Colleges dashboard

On this workout you will make an informational sheet about the colleges in United States, showing interesting data such as % of enrolled/applicants and the tuition costs over time.

Data	■ ■ ■ □ □
Design	■ ■ ■ □ □
Graphs	■ ■ ■ ■ □
Interactivity	■ ■ ■ ■ ■ □



Workout Start

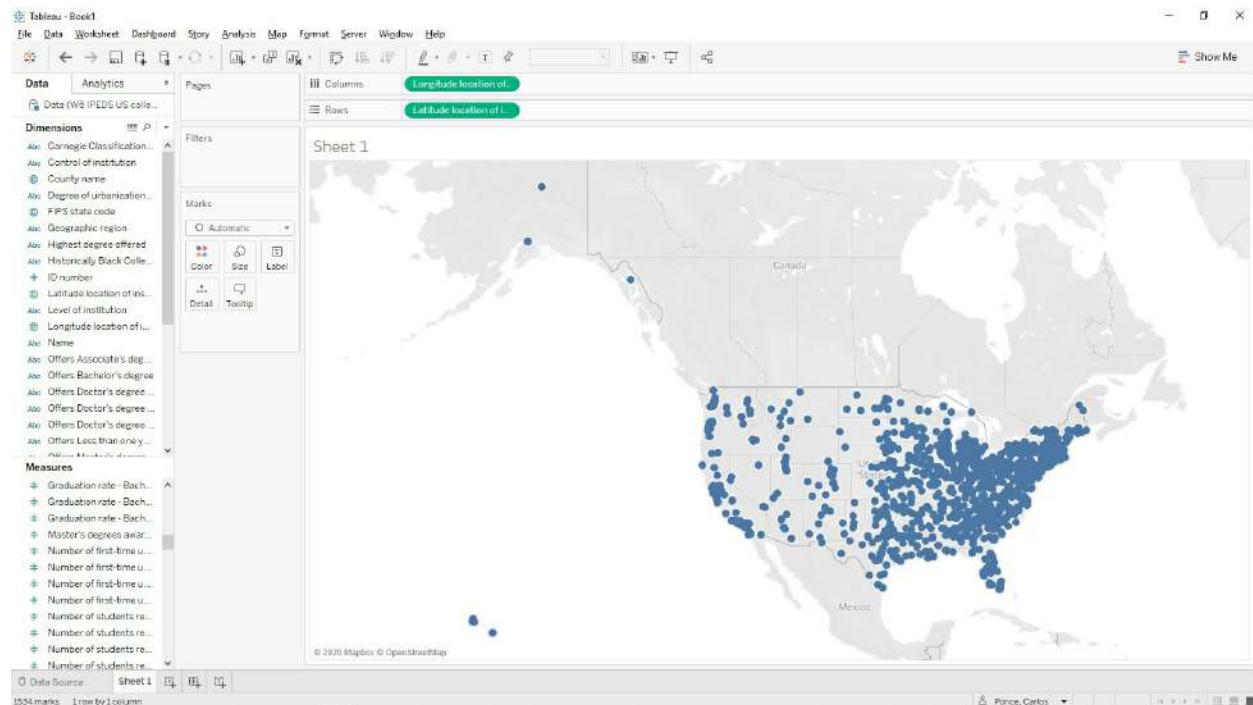
Analyze W6 IPEDS US Colleges.xlsx. It has four tabs: a data tab with the main table, a variables mapping table that we don't need, some notes, and a

comprehensive metadata tab that we also don't need to read.

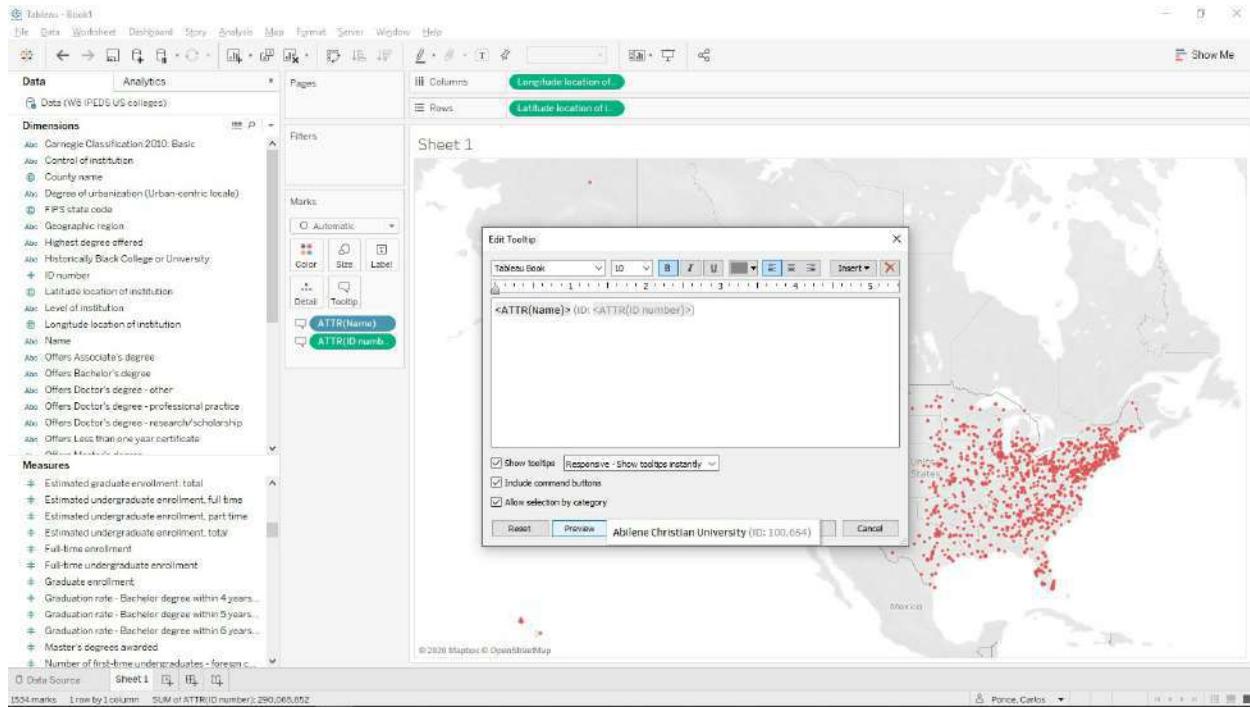
Our Data tab contains 1,534 rows, each of which represent a university in USA. We will make an informational dashboard for people interested in learning about US universities based on this data.

Start connecting to the tab and open *Sheet 1* (remember it's an Excel file this time, so you must connect to an Excel not a Text file on the startup screen).

Convert the **Latitude location of the institution** and **Longitude location of the institution** measures to dimensions and drag each to the rows and columns shelves, respectively.



Modify the size and color of the marks and drag **Name** and **ID Number** to the tooltip. Make the tooltip look like below:



Call the sheet *Map* and create a second one called *Diversity*. On it, we want to show the percentage of enrollments that belong to each ethnicity that our table is providing (American, Asian, African-American, etc.). However, we have these percentages in separate columns, reason for which we must pivot them on the Data Source tab. Do it as we did on workout 5: select the 33 variables related to enrollment percentage by ethnical group, right-click and choose **Pivot**.

Look for the two new columns at the end and rename to **Enrollment by ethnicity Name** and **Enrollment by ethnicity value**.

Go back to *Diversity* sheet, drag **Enrollment by ethnicity Name** to rows. We must make three new dimensions, so we can take only the total enrollment rows, the undergraduate enrollment rows, or the graduate enrollment rows without mixing each other. Create Enrollment (total) by ethnicity with this formula:

IF STARTSWITH([Enrollment by ethnicity name],"Percent of total") THEN
[Enrollment by ethnicity name] END

The screenshot shows the Tableau Data Editor interface. On the left, the 'Dimensions' and 'Measures' panes are visible. In the center, a calculated field is being edited under the 'Enrollment by ethnicity' column. The formula is:

```
IF STARTSWITH([Enrollment by ethnicity name],"Percent of total")
THEN [Enrollment by ethnicity name]
END
```

The calculation is valid, and the OK button is highlighted.

Do other two variables for Enrollment (graduate) and Enrollment (undergraduate), just change the second parameter of STARTSWITH() to be “Percent of graduate” and “Percent of undergraduate”, respectively.

To validate, drag the three new dimensions to rows, and confirm it looks like this:

The screenshot shows the Tableau Data Editor interface with four new calculated dimensions added to the 'Dimensions' pane:

- Enrollment (graduate) by ethnicity
- Enrollment (total) by ethnicity
- Enrollment (undergraduate) by ethnicity
- Enrollment by ethnicity name

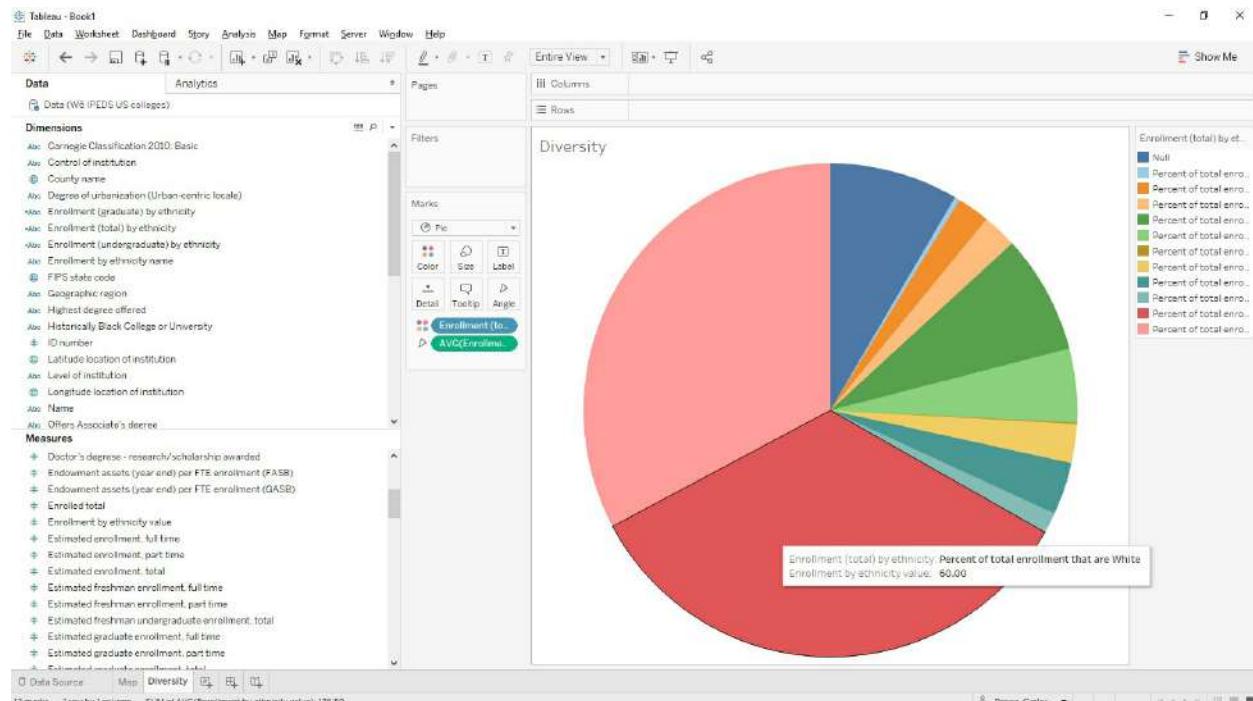
These dimensions are now listed in the 'Dimensions' pane along with the original dimensions like Carnegie Classification and County name.

Tip:

Using calculated values as filters as we just did is a common thing in practice. Sometimes, this is the only way to perform complex calculations, such as getting the variation of sales vs plan (one calculation to get only sales, a second to get only the plan, and a third to get the variation using the previous two).

When you face difficult operations, always think of this approach. You could call this process as unpivoting, it will be very helpful on those cases. However, if you see you're making way too many calculated values, consider using an external ETL tool that transforms your data in a way that is more useful, because the more calculations you have, the less performance your dashboard will have (meaning that it will take minutes to load and to change filters or parameters).

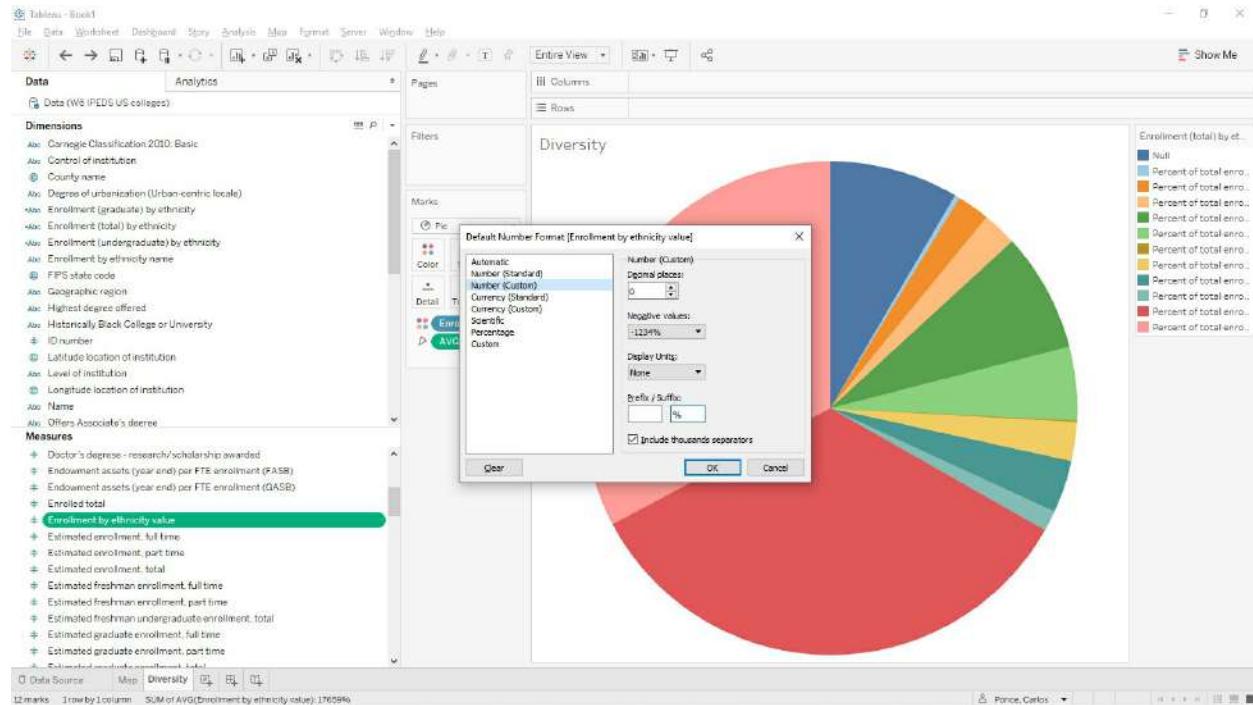
Delete all dimensions from rows and drag **Enrollment (total) by ethnicity** to the Color property. Before dragging **Enrollment by ethnicity value**, change its default aggregation property to AVG. Now change the Marks type from Automatic to Pie and drag the measure to the Angle property. Increase its Size property and set the graph to fill the entire view.



Notice how if you select all parts of the pie it gives more than 100%. Also,

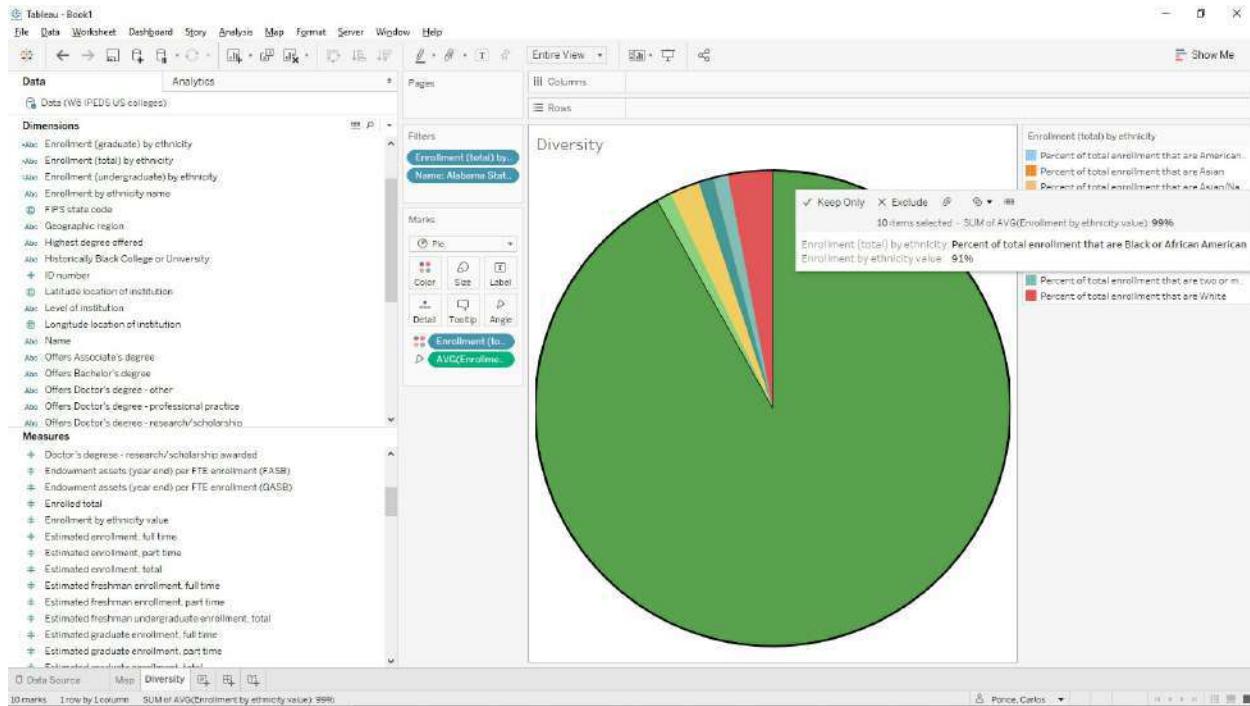
that the percentages are not being shown in a proper format.

First, fix the default format property of our measure so it's in percentage with no decimals. Do this by setting the format to **Number (custom)**, instead of setting it to **Percentage**, because that will multiply your numbers by 100.

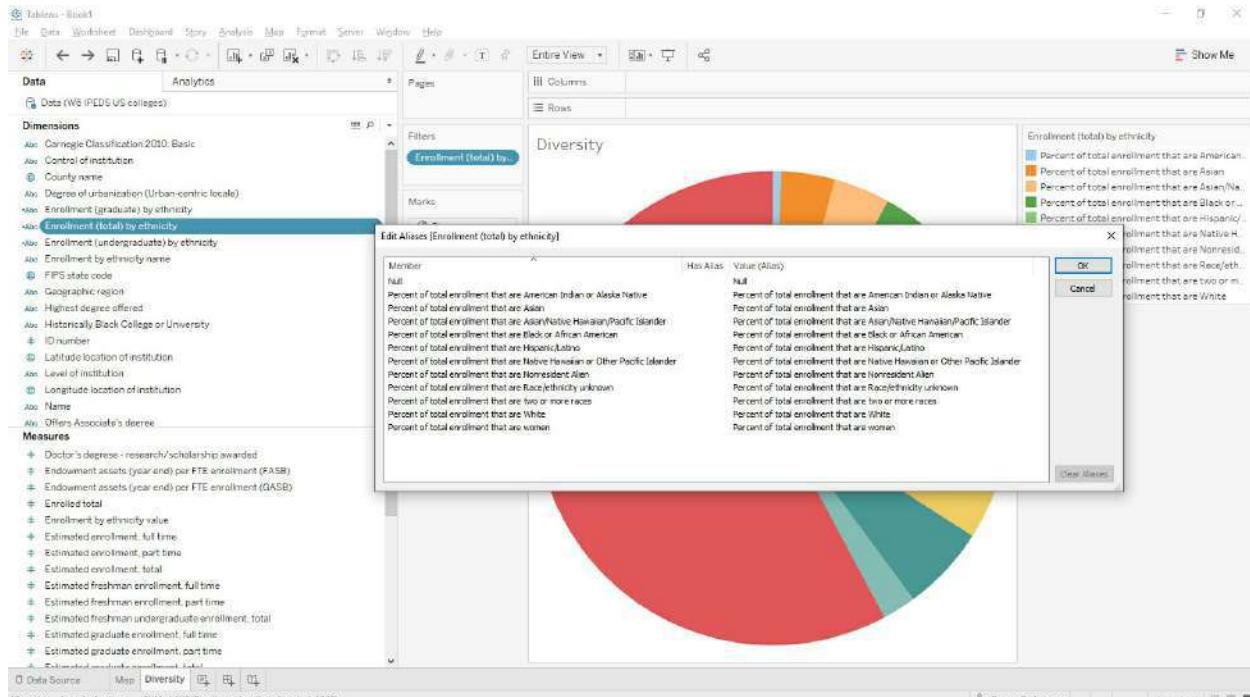


Now that the format is good, let's exclude the parts that should not be on this pie chart. We are talking about the values that have nothing to do with ethnicity. Start removing Null and percentage of enrollment that are women. Now our total is 104%. However, there is another thing; the percentages were calculated on a college level, right now we are seeing on a total level, so we are not seeing real percentages here but average percentages, so it's improbable we get to 100%.

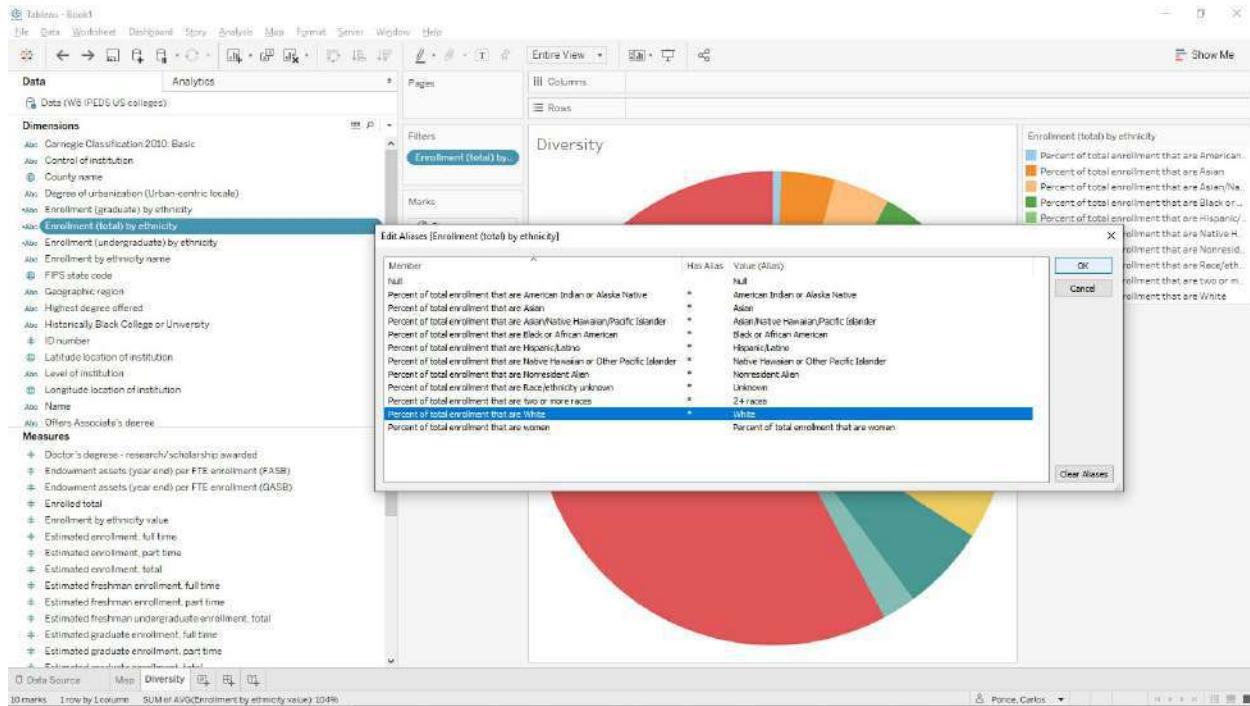
What we can do is filter by any university using **Name**, and confirm that the total there is 100%, like this:



The labels in the color legend are too long... we can clean them by using aliases; right-click the dimension from the Data pane and select Aliases.



Change the Value (alias) column to show shorter names as below:



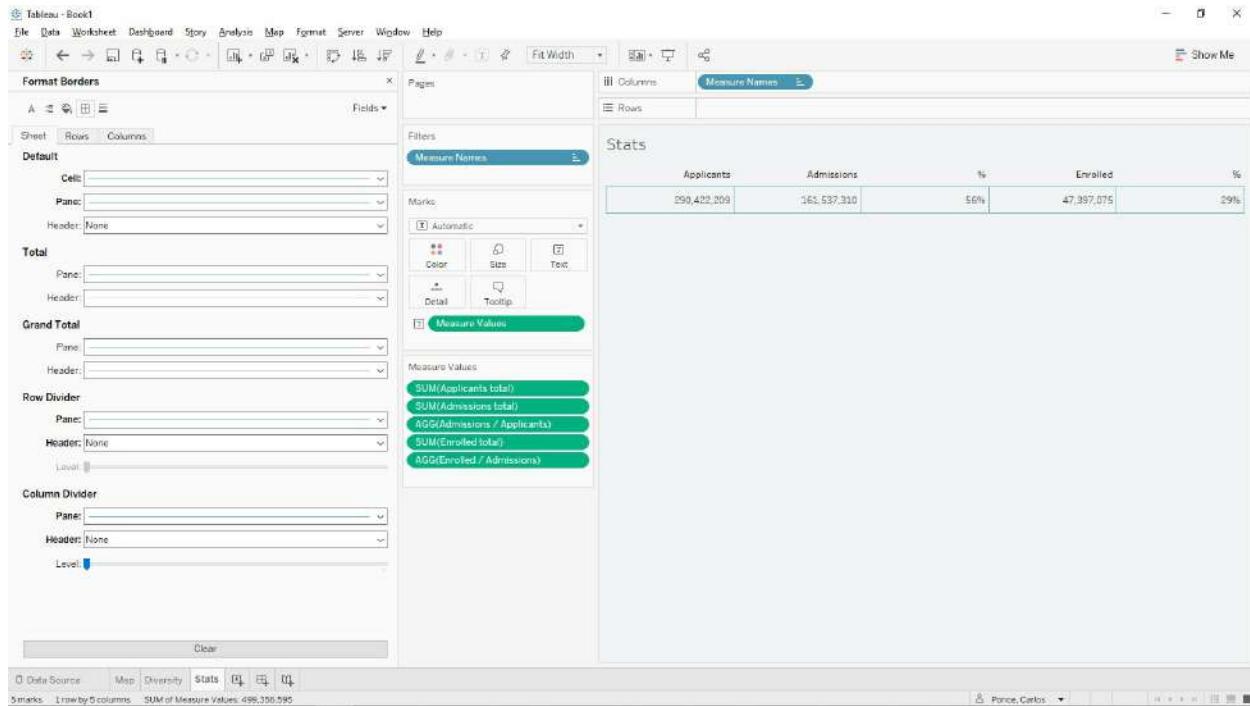
Let's move to the third sheet, call it *Stats*. We want to see how the number of applicants rolls up to the number of enrollments in a neat table. Remember that if you have any questions regarding the columns, you can open the Data Definitions tab on our source table.

Add **Applicants total**, **Admissions total**, and **Enrolled total** to the “Abc” cell. Make two calculations, one that takes the **SUM** of admissions and divides it by the sum of applicants, and one that takes the sum of enrolled and divides it by the sum of admissions. Call them **Admissions / Applicants** and **Enrolled / Admissions**, respectively.

The screenshot shows the Tableau Data Editor interface. On the left, the Dimensions and Measures shelves are visible, listing various fields like 'Carnegie Classification 2010: Basic', 'Control of institution', 'Degree of urbanization (Urban-centric locale)', etc. In the center, the 'Measure Names' shelf is selected. A floating window displays a calculated field named 'Admissions / Applicants' with the formula `SUM({\$Admissions total}) / SUM({\$Applicants total})`. Below this window, a message says 'The calculation is valid.' At the bottom of the screen, there are tabs for 'Data Source', 'Map', 'Diversity', 'Stats', and 'Marks'. The 'Marks' tab is currently active.

Set their default number format to percentage with no decimals and add them to the **Measure Values** shelf. Click the **Swap Rows and Columns** button, set table to fit width, and apply the next format:

1. Change the **headers aliases** by right clicking them, to remove the “total” part, and name the two calculated values “%” and “%”. You must add the space on the second because it does not allow repeated aliases.
2. Drag the lines so the height of the headers’ row and the data row are bigger.
3. On Format > Shading, apply the custom color #f1f3f4 on Default – Worksheet on the Sheet tab. This is part of a color palette found on [colorhunt.co](#).
4. On Format > Borders, apply the custom color #79bac1 on the Sheet tab for all lines except for the Headers, like below:



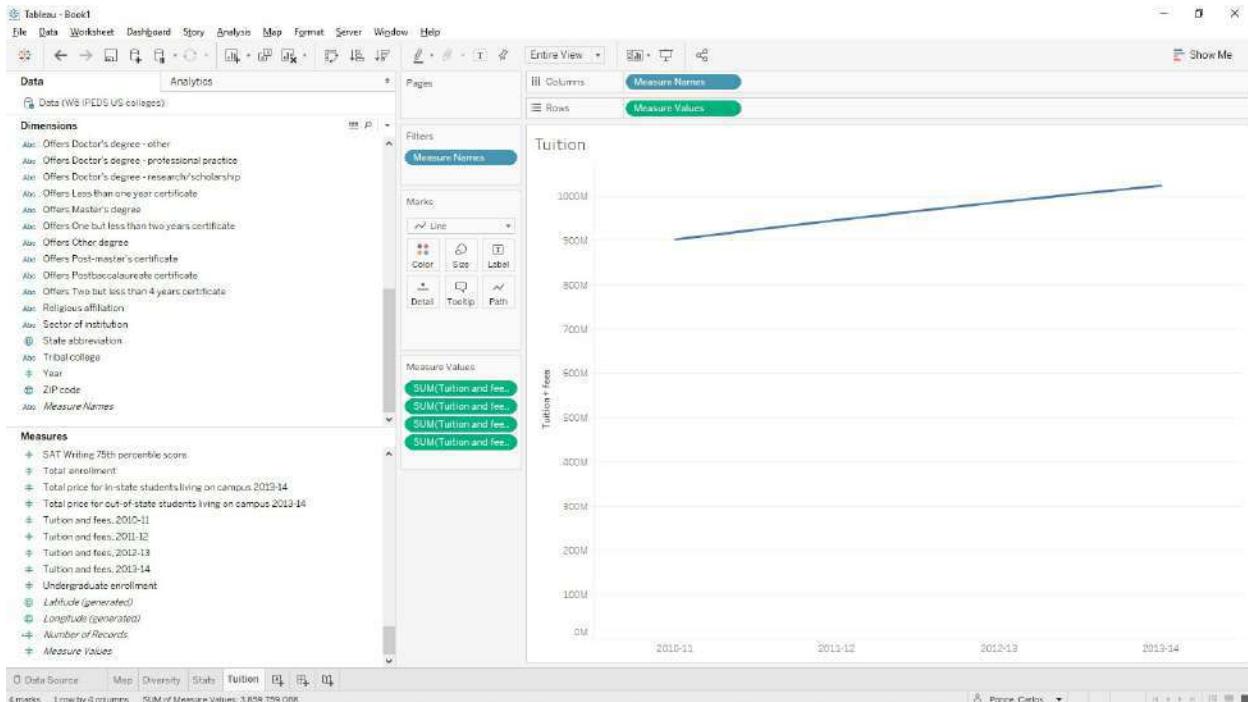
5. Go to the Text property and change the **Alignment** to Middle Center.
6. On Text again, increase the font size of your measure values to 12, change the text color to #2a7886.
7. Back to Format > Shading, select on the Fields dropdown **Measure Names**, apply background color #512b58 and Font color white.

Tip:

Remember: you format from the highest level (workbook) to the most detailed (fields), since the formatting done on the lower level overrides that of the higher level. That is why in this example we started applying format to the worksheet and then for Measure Names. See how our purple background applied over the light gray one.

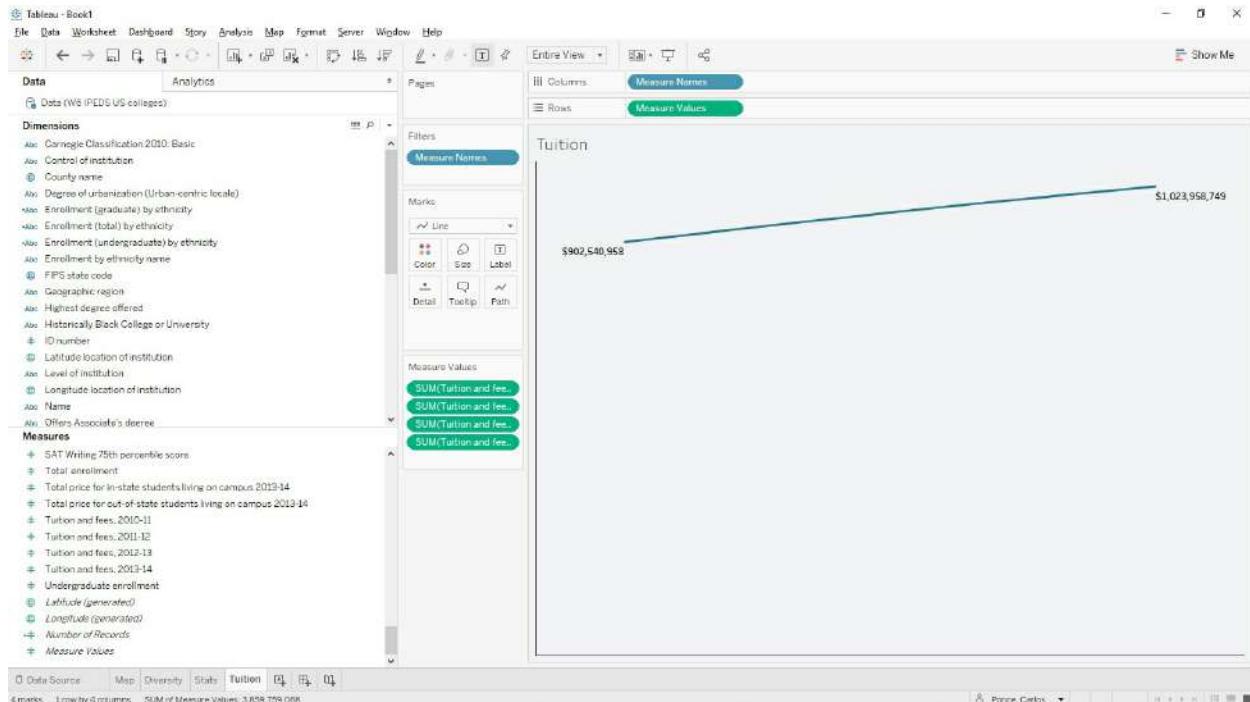
Make a new sheet called *Tuition*. We want to show the tuition numbers by academic year. But we face the same issue as with the percentages by ethnical group; the numbers are in different measures. However, for this case and to avoid having to mess again with the Pivot, let's use them as they are.

Drag the Measure Names to the Columns and the Filters shelves. On Filters, only enable the four tuition measures. Now drag Measure Values to Rows. Change the marks to be line, the graph to be entire view, the y-axis name to “Tuition + fees”, and the header alias for each measure so it looks like this:



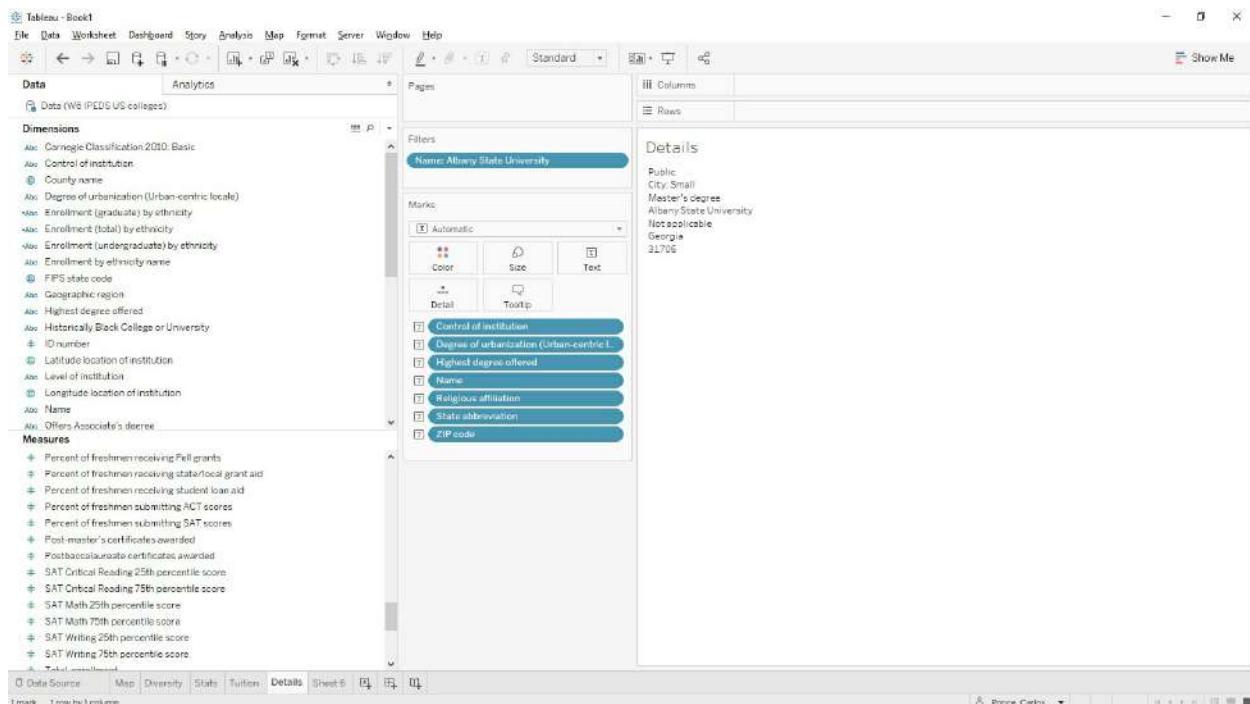
Let's show the numbers on the graph. For this, change the default number format of our four measures to be in **Currency (Custom)** without decimals nor display units, then **Show mark labels**.

Now for the formatting, let's give it a minimalististic aspect by hiding all axes. On Format > **Shading** > Sheet > Default, apply the same light gray from *Stats*. On Format > **Lines**, on the **Sheet** tab, look for **Axis Rulers** and change the line color to the purple we are using. Then on the **Rows** tab set **Grid Lines** to None. On Marks Colors, set it to be the darker blue we are using and set Markers to none. Finally, on Marks Text, make it bold, align it to bottom, and set the marks to label only on Min/Max.



Now, on to the fifth and final section: *Details*.

On a new sheet, add a filter on Name, select any university. Drag the dimensions you see on below image on the Marks' Text property.



Edit text so that the cell contents have their header next to them.

Call this sheet *DetailsA*. Create Details B and add the measures you see on below image. But before you do, change their default aggregation to average:

Measure Names	Measure Values
Percent of freshmen submitting ACT scores	\$3.1
ACT Composite 75th percentile score	28.6
ACT Composite 25th percentile score	20.6
Percent of freshmen submitting SAT scores	\$3.9
SAT Critical Reading 75th percentile score	579.3
SAT Critical Reading 25th percentile score	470.3
SAT Math 75th percentile score	588.3
SAT Math 25th percentile score	480.5
SAT Writing 75th percentile score	575.6
SAT Writing 25th percentile score	470.9

Be sure that, when you add them, you see the **Measure Names** and **Measure Values** pills on your table. Also, that you're putting them on the same order.

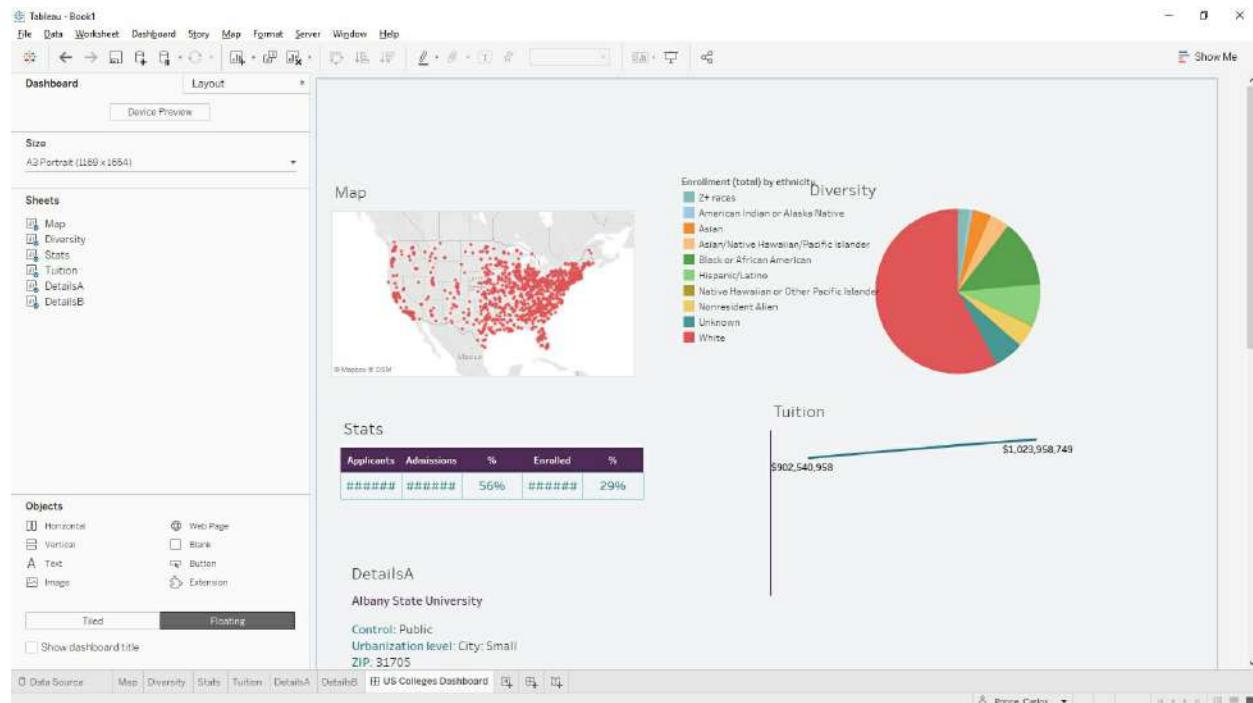
We now have our five main sections of the dashboard ready, but the

formatting is not equally applied for all of them. First, go back to each of the tabs and make sure the workbook background color (light gray) is applied on all. Then, on *DetailsA* and *DetailsB*, apply the color palette, set to fit width, and increase the font to 11.

The screenshot shows the Tableau desktop interface with two tabs open: *DetailsA* and *DetailsB*. Both tabs have a light gray background. The *DetailsA* tab displays a list of university details, including Name: Albany State University, Control: Public, Urbanization level: City, Small, ZIP: 31705, Highest degree: Master's degree, Religious affiliation: Not applicable, and State: Georgia. The *DetailsB* tab displays a list of measure values, including Percent of freshmen submitting ACT scores, ACT Composite 75th percentile score, ACT Composite 25th percentile score, Percent of freshmen submitting SAT scores, SAT Critical Reading 75th percentile score, SAT Critical Reading 25th percentile score, SAT Math 75th percentile score, SAT Math 25th percentile score, SAT Writing 75th percentile score, and SAT Writing 25th percentile score. A floating 'Edit Label' dialog is open, showing the text 'Tableau Book' with a font size of 11pt.

Now we are ready to do our dashboard. On a blank one, set the background to be our lighter blue, the size to be A3 Portrait, and set the dashboard to be

made using the **Floating** approach. drag the sheets onto it like below:



Add a dashboard title by clicking Dashboard > Show title. Call the sheet “US Colleges Dashboard”, and then edit the dashboard title you just added to add below this additional text:

“This is a dashboard developed as part of the Tableau Workout Book workout 6. The information used was downloaded from the IPEDS Data Center <http://nces.ed.gov/ipeds/datacenter/Default.aspx>. IPEDS data are inspected for obvious errors, but not adjudicated. It is not recommended you make any important strategic decisions based on this data.”

The screenshot shows the Tableau Desktop interface with the following elements:

- Top Bar:** File, Data, Worksheet, Dashboard, Story, Map, Format, Server, Widgets, Help.
- Dashboard Header:** US Colleges Dashboard. A note states: "This is a dashboard developed as part of the Tableau Workout Book workout 6. The information used was downloaded from the IPEDS Data Center <http://nces.ed.gov/ipeds/datasetcenter/default.aspx>. IPEDS data are inspected for obvious errors, but not adjudicated. It is not recommended you make any important strategic decisions based on this data."
- Left Sidebar:**
 - Sheets:** Map, Diversity, Stats, Tuition, DetailsA, DetailsB.
 - Objects:** Horizontal, Vertical, Text, Image, Web Page, Blank, Button, Extension.
 - Layout:** Tiled (selected), Floating.
 - Show dashboard title:** Checked.
- Content Area:**
 - Map:** A choropleth map showing college enrollment by state.
 - Diversity:** A pie chart titled "Enrollment (total) by ethnicity" showing proportions for various racial and ethnic groups.
 - Tuition:** A line chart titled "Tuition" showing tuition fees from \$902,540.95 to \$1,023,958.74.
 - DetailsA:** A card for Albany State University with controls: Public, Urbanization level: City: Small, ZIP: 31705.
- Bottom:** Data Source (selected), Map, Diversity, Stats, Tuition, DetailsA, DetailsB, US Colleges Dashboard, and a floating title editor window.

Now start doing the floating approach: resize and drag all the elements of your dashboard until you get to a layout that you like. You can base on the below design (remove the extra height that you don't need):

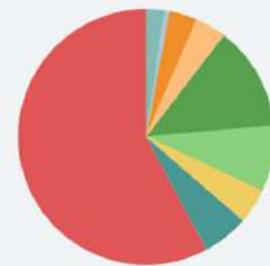
US Colleges Dashboard

This is a dashboard developed as part of the Tableau Workout Book workout 6. The information used was downloaded from the IPEDS Data Center <http://nces.ed.gov/ipeds/datacenter/Default.aspx>. IPEDS data are inspected for obvious errors, but not adjudicated. It is not recommended you make any important strategic decisions based on this data.

Map



Diversity



Enrollment (total) by ethnicity	
2+ races	Green
American Indian or Alaska Native	Light Blue
Asian	Orange
Asian/Native Hawaiian/Pacific Islander	Light Orange
Black or African American	Dark Green
Hispanic/Latino	Light Green
Native Hawaiian or Other Pacific Islander	Yellow-Gold
Nonresident Alien	Gold
Unknown	Dark Blue
White	Red

Information:

Albany State University

Control: Public
Urbanization level: City: Small
ZIP: 31705
Highest degree: Master's degree
Religious affiliation: Not applicable
State: Georgia

Percent of freshmen submitting ACT scores	53.1
ACT Composite 75th percentile score	25.6
ACT Composite 25th percentile score	20.6
Percent of freshmen submitting SAT scores	53.9
SAT Critical Reading 75th percentile score	578.3
SAT Critical Reading 25th percentile score	470.3
SAT Math 75th percentile score	588.3
SAT Math 25th percentile score	480.5
SAT Writing 75th percentile score	576.6
SAT Writing 25th percentile score	470.9

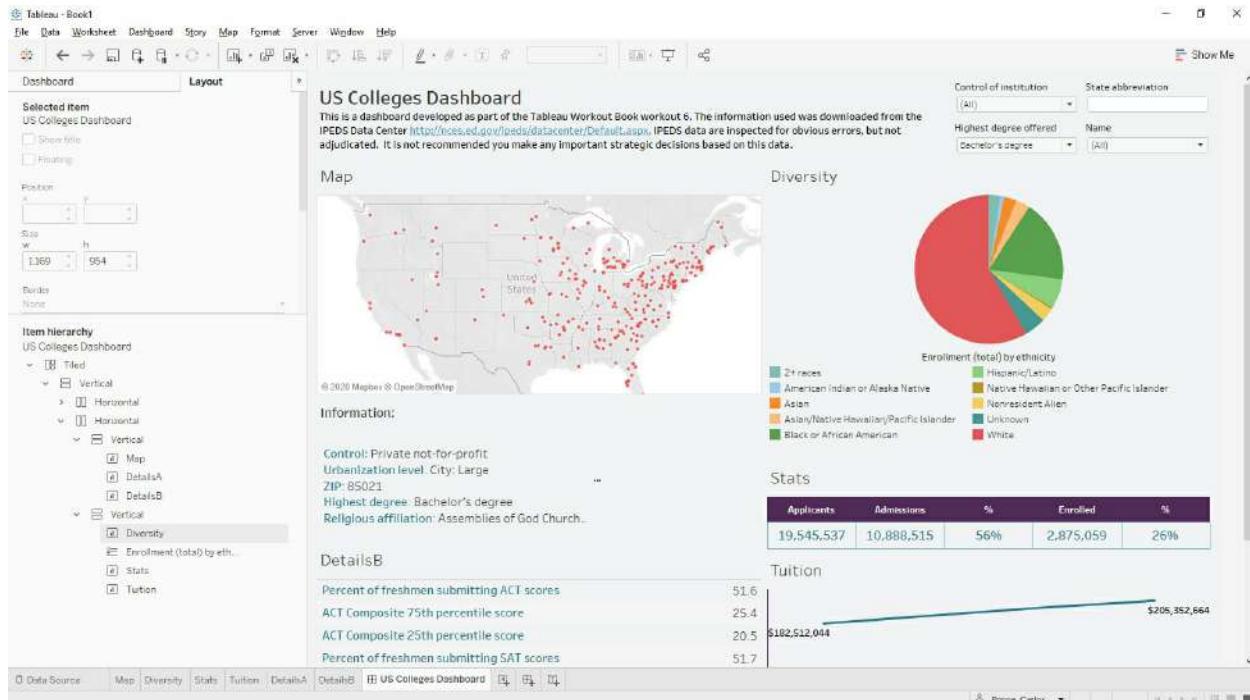
Stats

Applicants	Admissions	%	Enrolled	%
290,422,209	161,537,310	56%	47,397,075	29%

Tuition



It's looking good so far, but we're missing a key element: filters. Apply global filters for **State Abbreviation**, **Control of institution**, **Highest degree**, and **Name**. Show them on the top like this:

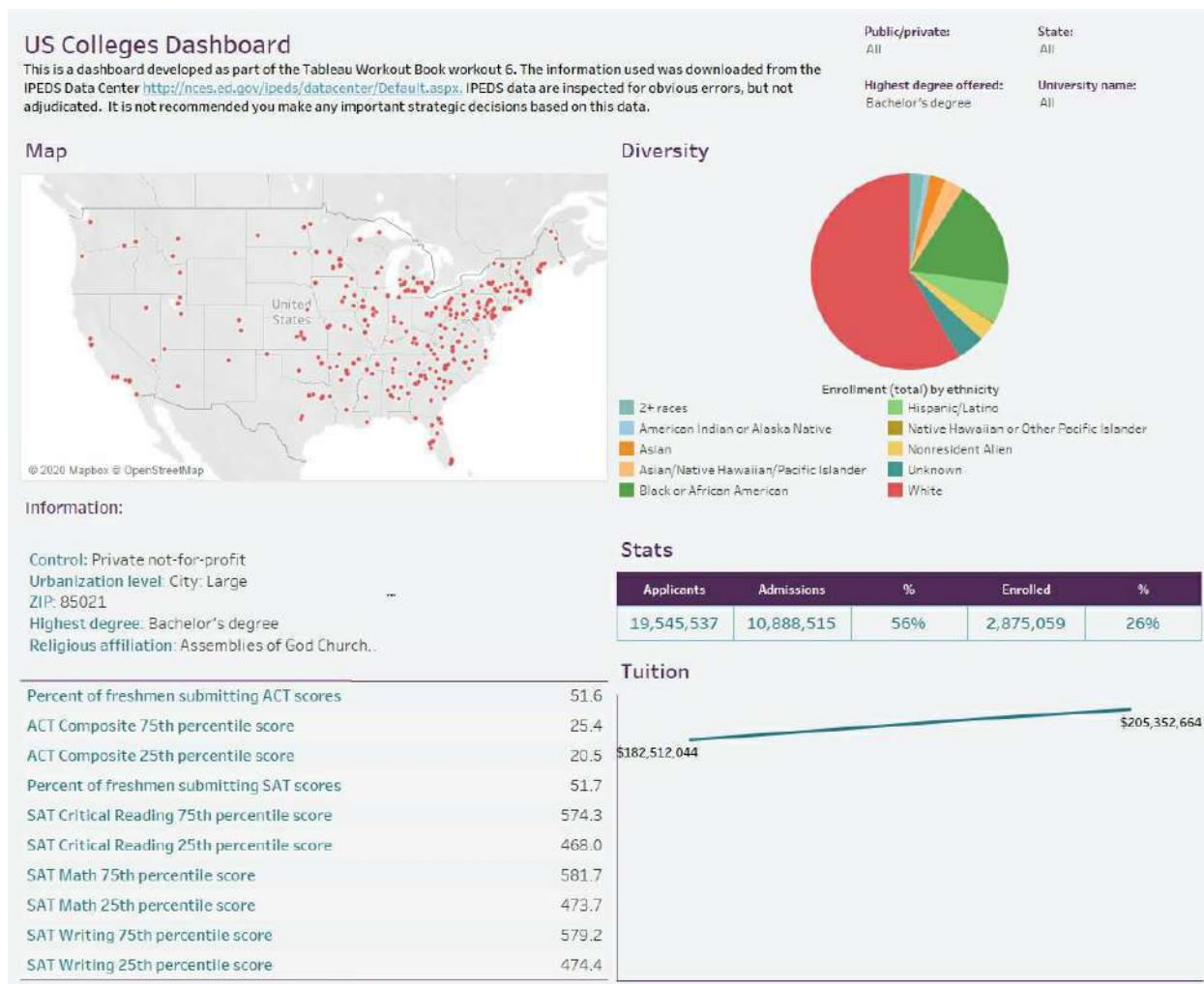


Have you had issues so far with the **floating** approach? Maybe when you reduced the dashboard's height your floating elements readjusted, adding scrolling bars on all of them. Or adding the global filters became a more complicated task. I wanted to show you on this workout how the floating approach might look easy at a glance when doing a dashboard from the ground up, but how painful it is when you make changes to it afterwards.

If you didn't have trouble, feel free to continue doing it with floating. If you did, you can start over by typing Ctrl + Shift + Backspace, or the **Clear Sheet** button on the top, then redoing it with the tiled approach as we did on last image.

Now, apply formatting to the filter titles, so it's bold and purple. Also do it for the sheet titles. Set the Map to **Use as Filter**.

As a final touch, go back to *DetailsB* and paint those gray horizontal lines to purple on Format > Borders > Rows > Row Divider > Pane.



Play with your dashboard and make sure it's working fine. **Congratulations, you have now finished the sixth workout!** As you could tell, this time I went very high-level on the explanations on how to do every step. This will help you start to get more confident with Tableau and to force you to recap previous workouts or even search on the Internet how to do a certain thing (but hopefully you did not had to recur to that because all has been covered on previous workouts).

Here are the new topics we went through:

- STARTSWITH() function
- Using calculated values as filters
- Values and headers' aliases
- Recap on formatting layers (worksheet vs field)
- Dashboards: Floating approach

Go to next page to start the seventh workout.

Workout 7. FIFA 19 player rank dashboard

On this workout you will build a comparative dashboard of the players of Electronic Arts's famous game FIFA 19.

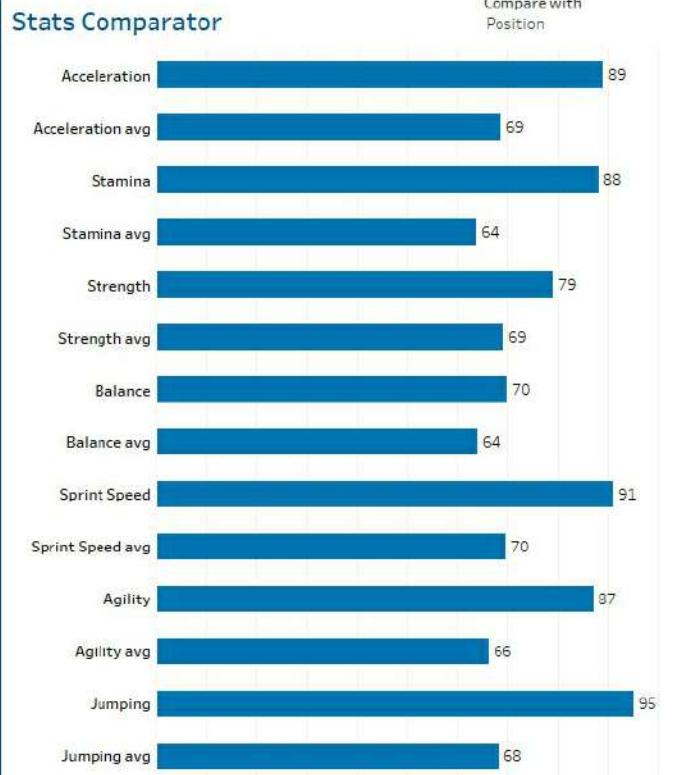


FIFA 19 Players Comparison Cristiano Ronaldo

33 years old from Portugal



Nationality	Club	Position	
All	Juventus	All	
Name	Club	Nationality	Jersey Num..
A. Barzagli	Juventus	Italy	15
Alex Sandro	Juventus	Brazil	12
B. Matuidi	Juventus	France	14
C. Pinsoglio	Juventus	Italy	21
Cristiano R...	Juventus	Portugal	7
D. Rugani	Juventus	Italy	15
Douglas Co...	Juventus	Brazil	11
E. Can	Juventus	Germany	23
F. Bernard...	Juventus	Italy	20
G. Chiellini	Juventus	Italy	3
J. Cuadrado	Juventus	Colombia	11
João Cancelo	Juventus	Portugal	2
L. Bonucci	Juventus	Italy	19
L. Spinazzola	Juventus	Italy	37
M. Benatia	Juventus	Morocco	4
M. De Sciglio	Juventus	Italy	2
M. Kean	Juventus	Italy	18
M. Mandžu...	Juventus	Croatia	17
M. Perin	Juventus	Italy	22
M. DianicA	Juventus	Senegal	5



Workout Start

Connect to W7 FIFA19 dataset.csv and observe the end table. What we have is the list of all the players in the FIFA 19 game, with metrics ranging from age, salary, to ability per football position (goalkeepers will have nulls in all positions except the ones from GK).

Name your first sheet *Player list*. Drag **Name**, **Club**, **Nationality**, and **Jersey Number** to the Rows shelf. Then drag **Position** to the Text property in Marks. Add **Position**, **Club**, and **Nationality** to the Filters shelf as well, and show them as Single Value (dropdown). Finally, set the table to Fit Width.

The screenshot shows the Tableau Data Editor interface with the following details:

- Top Bar:** File, Data, Worksheet, Dashboard, Story, Analysis, Map, Format, Server, Window, Help.
- Toolbar:** Includes icons for zoom, refresh, search, and various data analysis tools.
- Data Pane (Left):**
 - Dimensions:** Body Type, CM, CS, CDM, CF, Club, Club Logo, CM, CM, Flag, Height, ID, Jersey Number, Joined, LAM, LB, LCB, LCM.
 - Measures:** Acceleration, Age, Aggression, Agility, Balance, Ball Control, Composure, Contract Valid Until, Crossing, Curvo, Dribbling, FL, Finishing, FX Accuracy.
- Filters (Bottom Left):** Position (selected), Club, Nationality.
- Marks (Bottom Left):** Automatic, Color, Size, Text, Detail, Tooltip, Position.
- Table View (Main Area):**
 - Sheet 1:** A table with columns: Name, Club, Nationality, Jersey Num., and Position.
 - Data:** A list of player records, such as A. Abdeslam (Club Desportos Temuco, Argentina, 27, LM), A. Abou (New York Red Bulls, Cameroon, 20, ST), etc.
 - Row Labels:** Jersey numbers (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100).
 - Right Panel:** Shows filters for Position (All), Club (All), and Nationality (All).
- Bottom Navigation:** Data Source, Sheet 1, Fit Width, and other dashboard controls.

The easy table is ready. Now create a new sheet named *Stats Comparator*. For context, what we want to build is a dashboard that not only lists the players' stats, but that allows the users to compare a player's stats versus the average for the position.

On the Data pane under Measures, right-click any measure, then **Group by**, and select **Folders**. This will enable you to make folders for your dimensions and measures.

Tip:

It's a good practice to keep your measures grouped in folders with explanatory names. In the workplace, dashboards may grow and have many calculated values – I have had once a dashboard with more than 300, which I don't recommend since that makes your dashboard very slow. Because of this, it is better that you start grouping from the beginning.

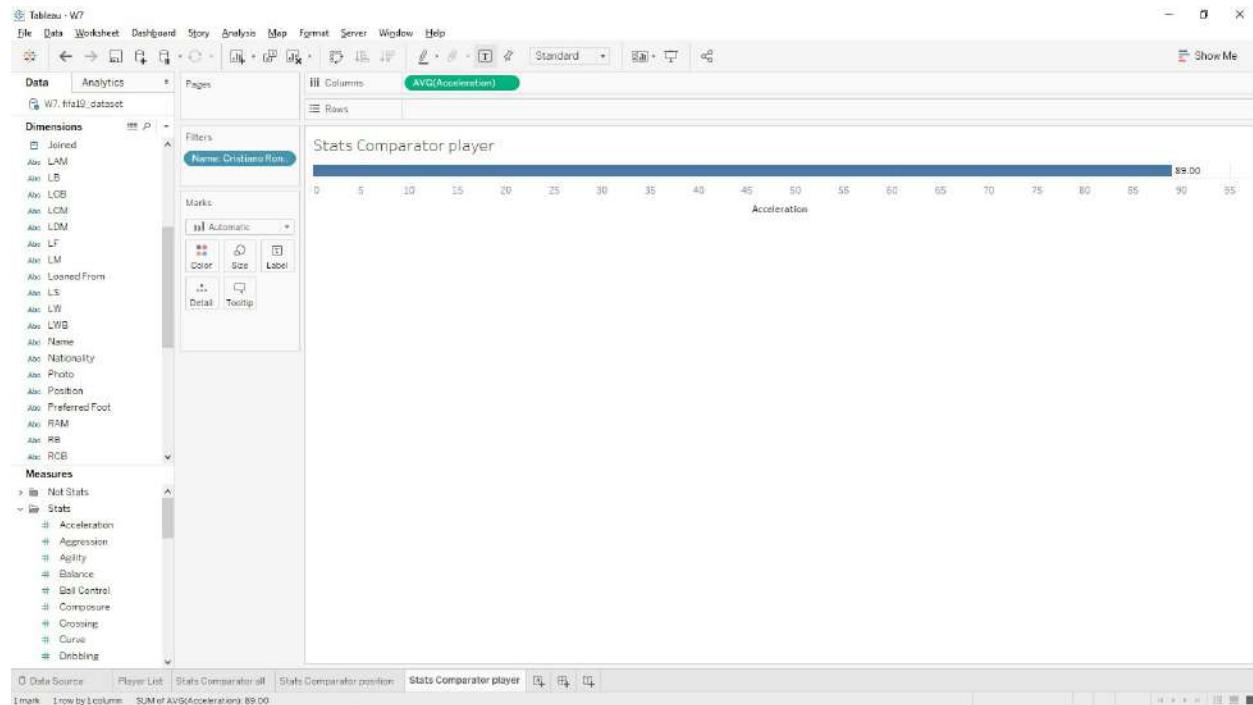
Now select **Age**, **Contract Valid Until**, **Weak Foot**, **Skill Moves**, and **F1**,

and group them on a new folder called *Not Stats* (create a new folder with right click > Folder > New Folder). Then group the rest of the measures and put them on a folder called *Stats*. You can collapse/hide the *Not Stats* folder for now.

Also, since we are comparing averages, change the default aggregation property for all the measures under the *Stats* folder to average instead of sum (don't do one by one, remember you can press the first on the list, then holding Shift click the last measure, then you can hit the right click to change it).

Drag **Acceleration** to the Columns shelf on your new sheet. This is the average acceleration for all players in the database.

But we want to filter by player and by position, so to do it slow and clear, duplicate the sheet two times. In one, apply the **Position** filter equal to ST and in the second apply the **Name** filter equal to Cristiano Ronaldo (this is for demonstrative purposes). Add “position” and “player” to the name of those two sheets, respectively:

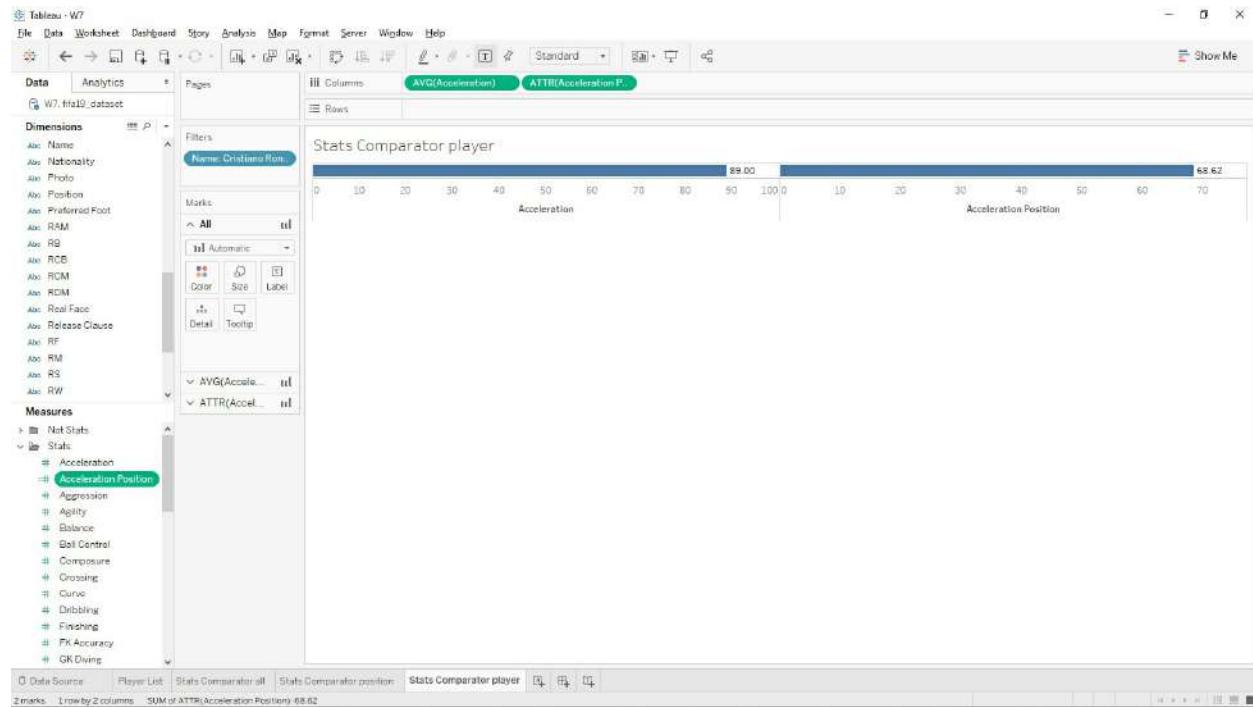


How can we show in one same graph one bar which only considers the average for the player and another bar with the average for the position? The answer is: Level of Detail calculations.

Create a calculated field named **Acceleration Position** and add it on the

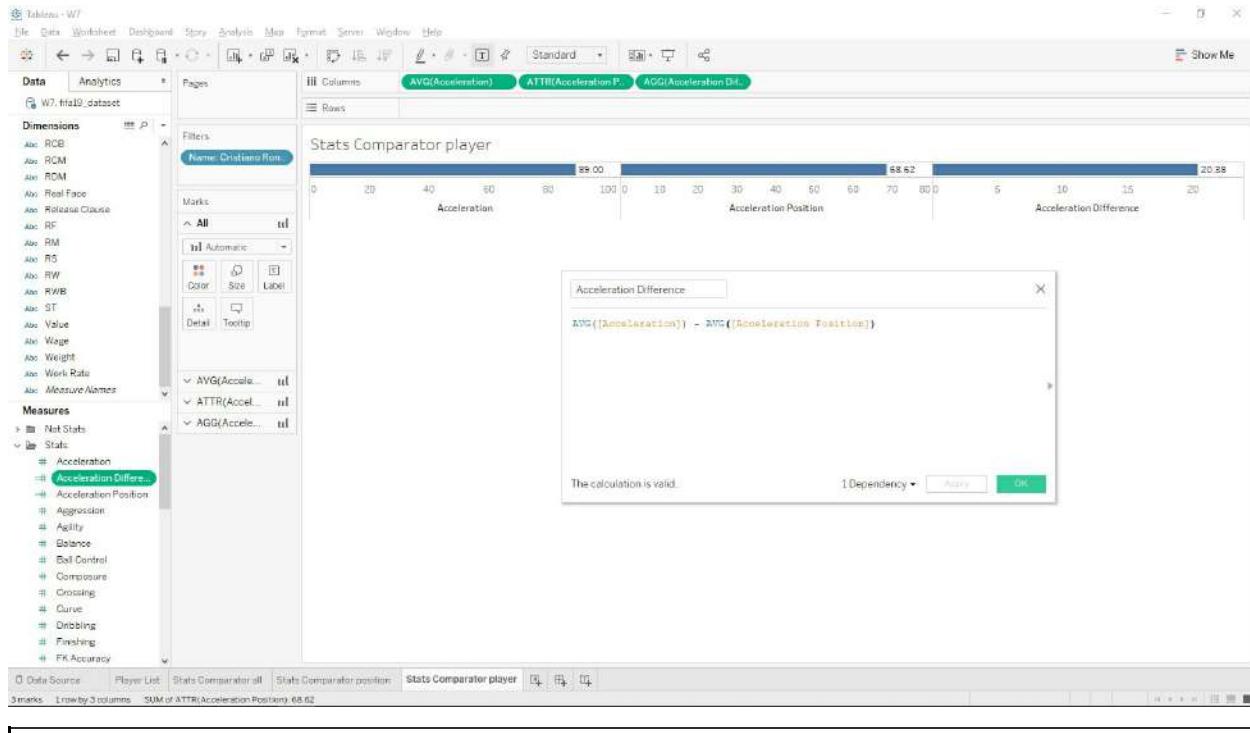
Columns shelf of the *Stats Comparator Player* sheet:

{FIXED [Position]: AVG([Acceleration])}



Go back to the *Stats Comparator position* sheet to confirm that the amount that your new calculated field is showing is the right one: **68.62**. Remember: we have three Level-of-detail calculation functions available: FIXED, INCLUDE, and EXCLUDE, but FIXED is the one that you will use the most.

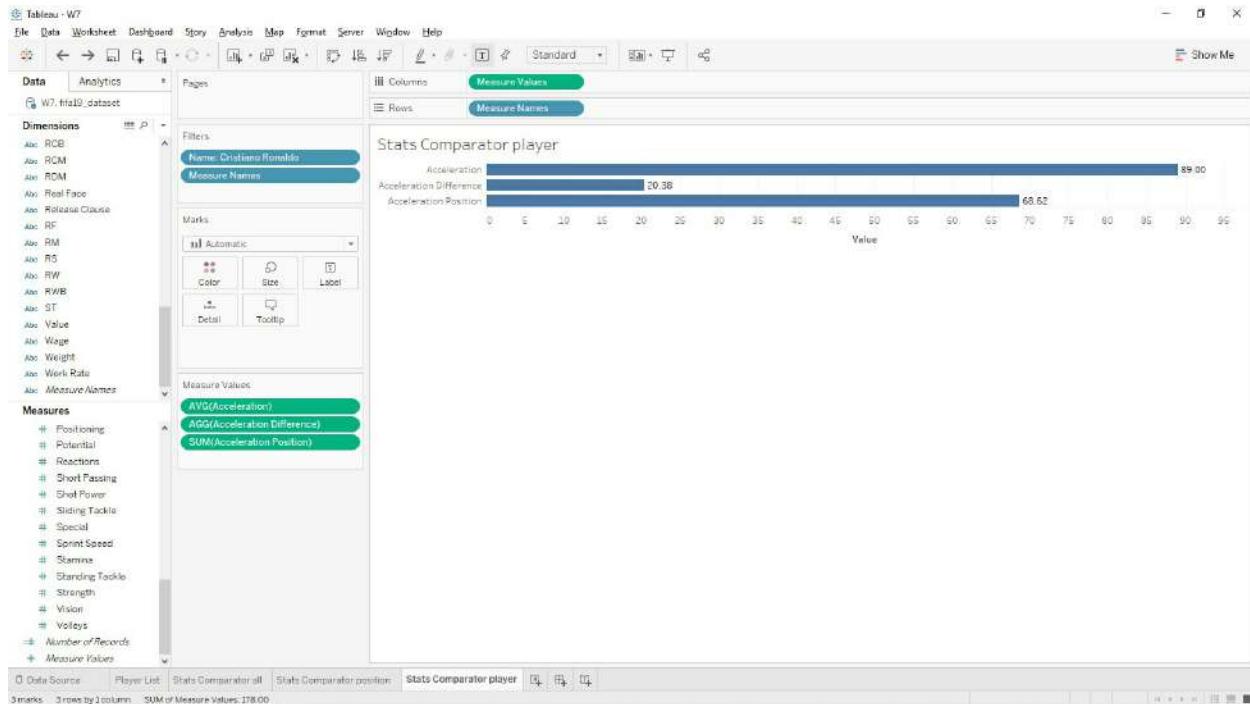
Now, we have our two elements to compare; let's create the calculated variable that will state by how many points is the difference. Create Acceleration Difference = AVG([Acceleration]) - AVG([Acceleration Position]), and also add it to the graph to validate it's giving the right number: 20.38.



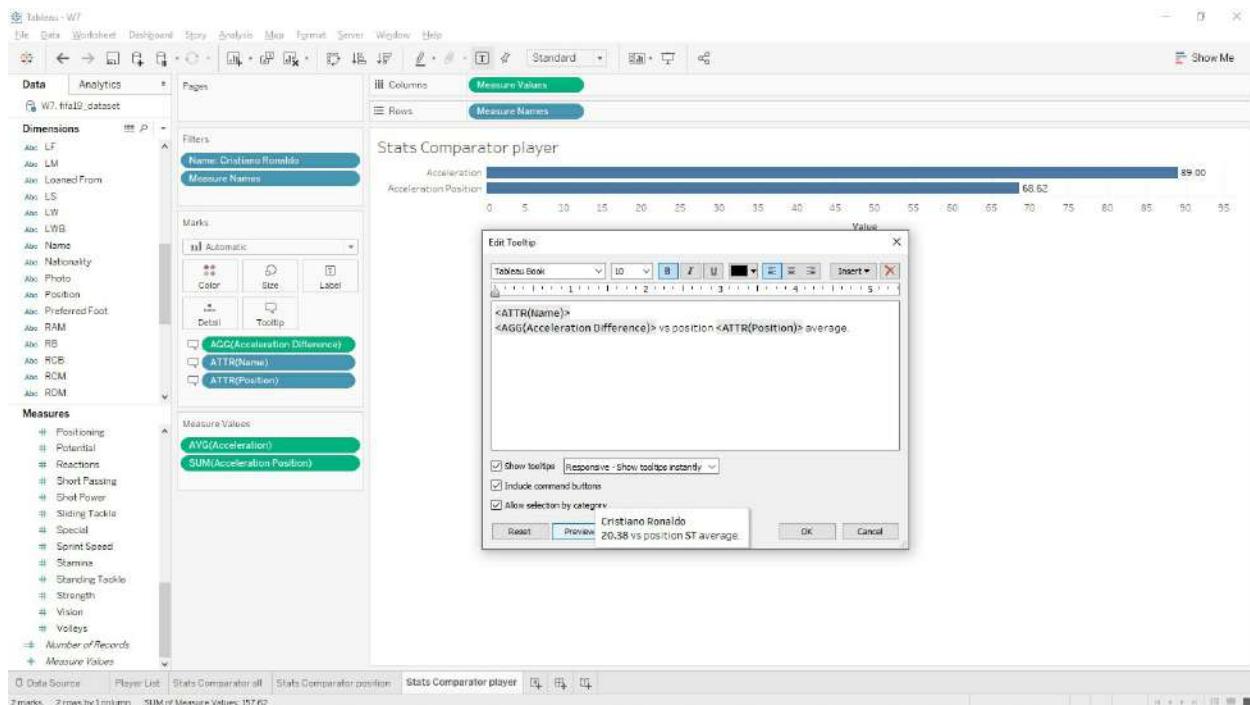
Tip:

It's highly recommended to always do proper testing on your calculated variables, even if you are very sure that what you are doing will be correct. It is common that we omit small but important details such as aggregation, order of operations, and other aspects of a calculation. It is also possible that your calculated value works well when showing a certain dimension in your table, but that fails when you modify it. Remember: it's best to go slow and assure the quality of your workbooks.

Now that we have our calculated values set, let's set up the graph. Remove all three measures from the Columns shelf, and instead put **Measure Names** onto Rows. Add a filter on it so you only show our three main measures, then drag **Measure Values** onto the Columns shelf.

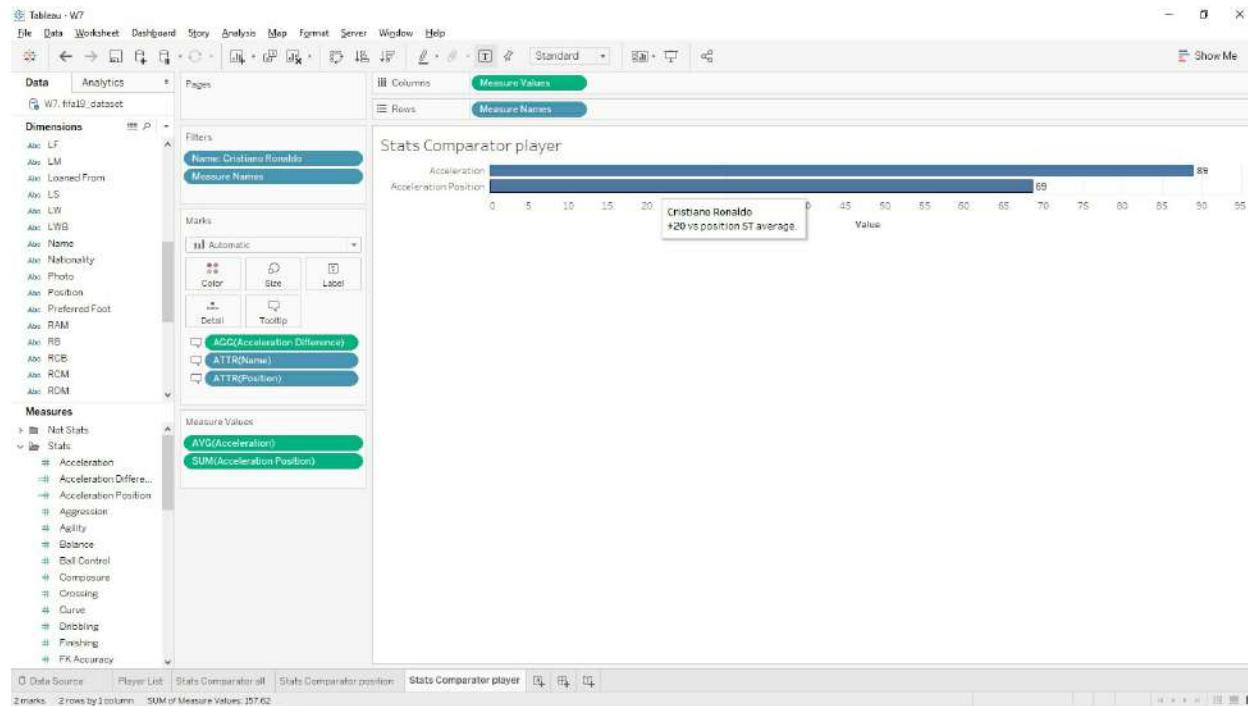


Now we have a vertical display of our measures which looks better. Move **Acceleration Difference** from the Measure Values shelf and put it on the Tooltip property of marks. Also add from the Dimensions pane **Name** and **Position** to the Tooltip, and edit it so it looks like below:



We will deal with the color and formatting after we finish our graphs. But let's deal with number format: for **Acceleration** and **Acceleration Position**,

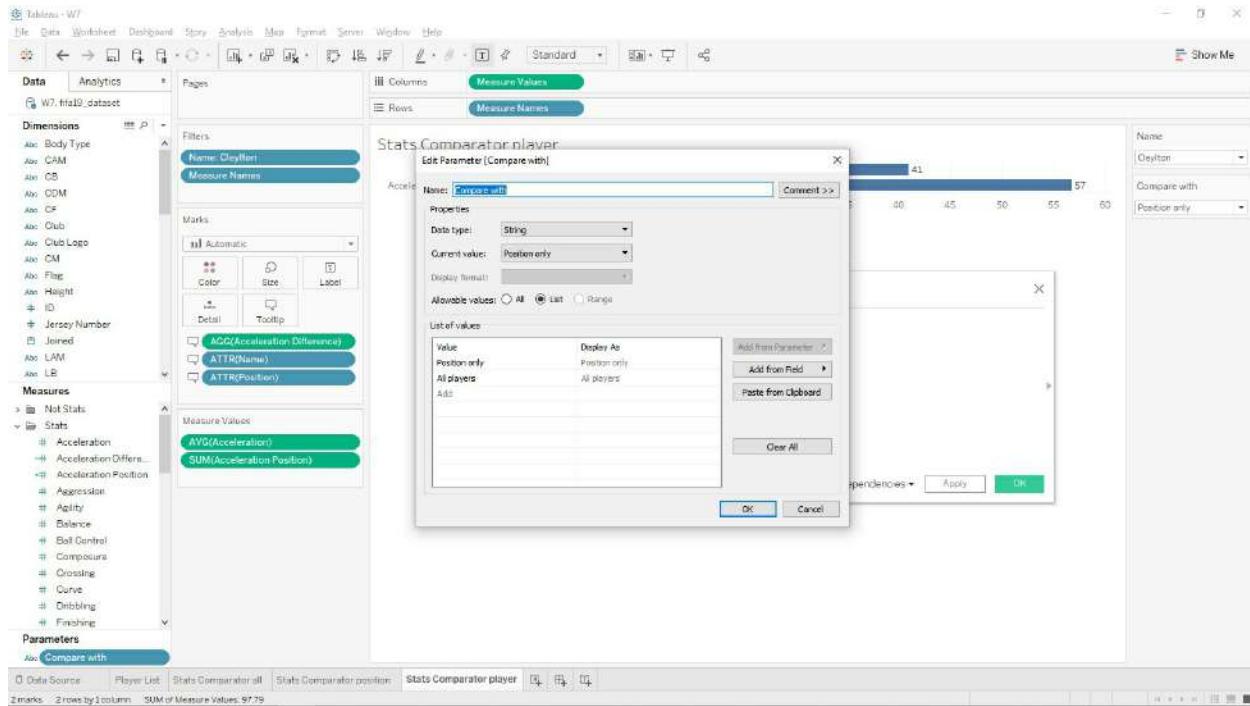
make it default to number with no decimals, and make Acceleration Difference default to custom, using this format: +#,##0;-#,##0. It's the same format as Number (Custom) with 0 decimals, just that we are adding the “+” sign on the positives, so it's more readable on the tooltip (it will show “+20” instead of “20”).



Show the **Name** filter as Single Value (dropdown) and start testing your graph. Use different player names, read that the tooltip number is updated correctly on every change, that the position is changing as well depending on the player, that the bars are reflecting the correct numbers (you can go back to *Stats Comparator position* to check the position amount).

Now that we have confirmed it is working, let's add a second step of complexity: how about we also enable the user to compare stats versus all players?

Create a parameter called **Compare with**, set data type to string, and for values set a list of only two options: “Position only” and “All players”.

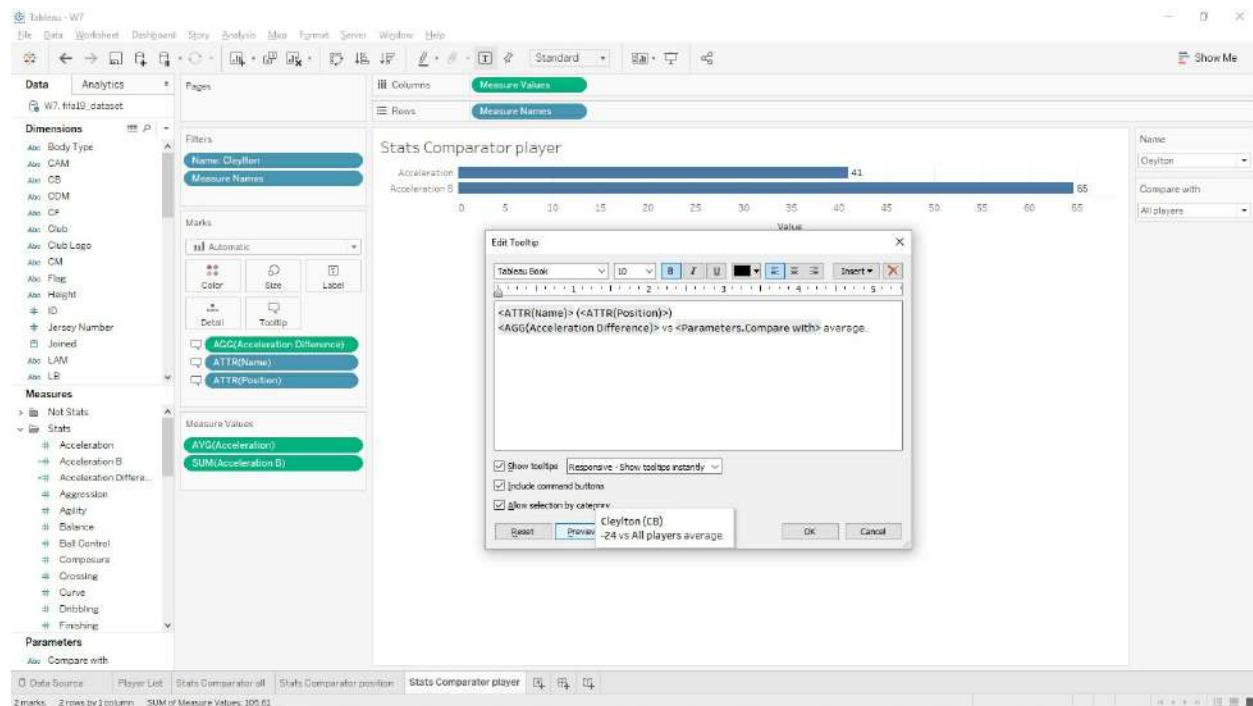


Show it on your *Stats Comparator player* sheet, and modify your **Acceleration Position** calculated value; change the name to **Acceleration B** (to make it clear it's the second acceleration metric but not exclusively for position average) and modify the formula so that if Compare with is “Position only” uses the current calculation, else it should display the average for all players. Try to do it first by yourself, and when done compare with the correct formula on the next page:

Acceleration B

```
IF [Compare with] = "Position only" THEN  
    {FIXED [Position]: AVG([Acceleration])}  
ELSE  
    {AVG([Acceleration])}  
END
```

Now change the parameter between its two values and compare with the other two sheets (*Stats Comparator all* and *Stats Comparator position*). You will see that it's working properly. There's a thing, though; our tooltip is fixed to label the comparison as if it is always by position. Fix it by inserting the Compare with parameter in the tooltip and making changes needed to make it look properly:

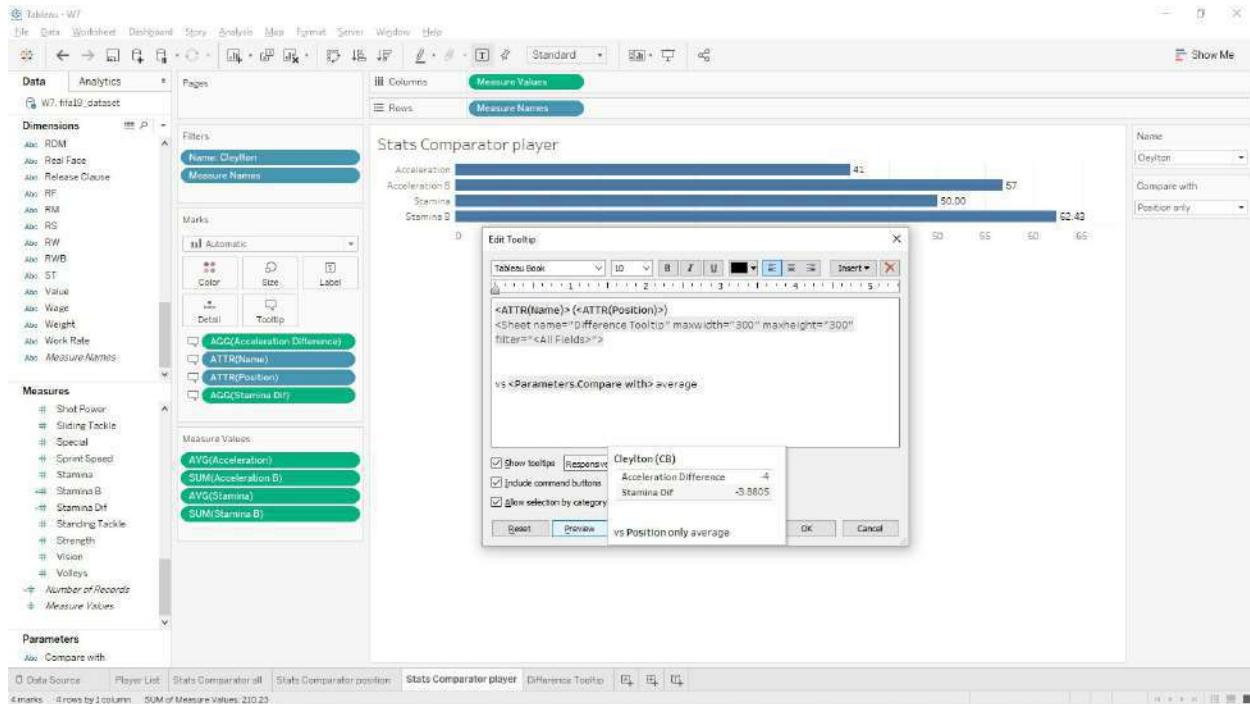


Now create **Stamina B** and **Stamina Difference** based on the same logic. Put **Stamina** and **Stamina B** on the Measure Values shelf, while putting **Stamina Difference** on the tooltip. Notice the problem we have with the tooltip; Tableau is showing **Acceleration Difference** even when hovering the mouse over the stamina bars. Therefore, there is a problem to display number by number individually.

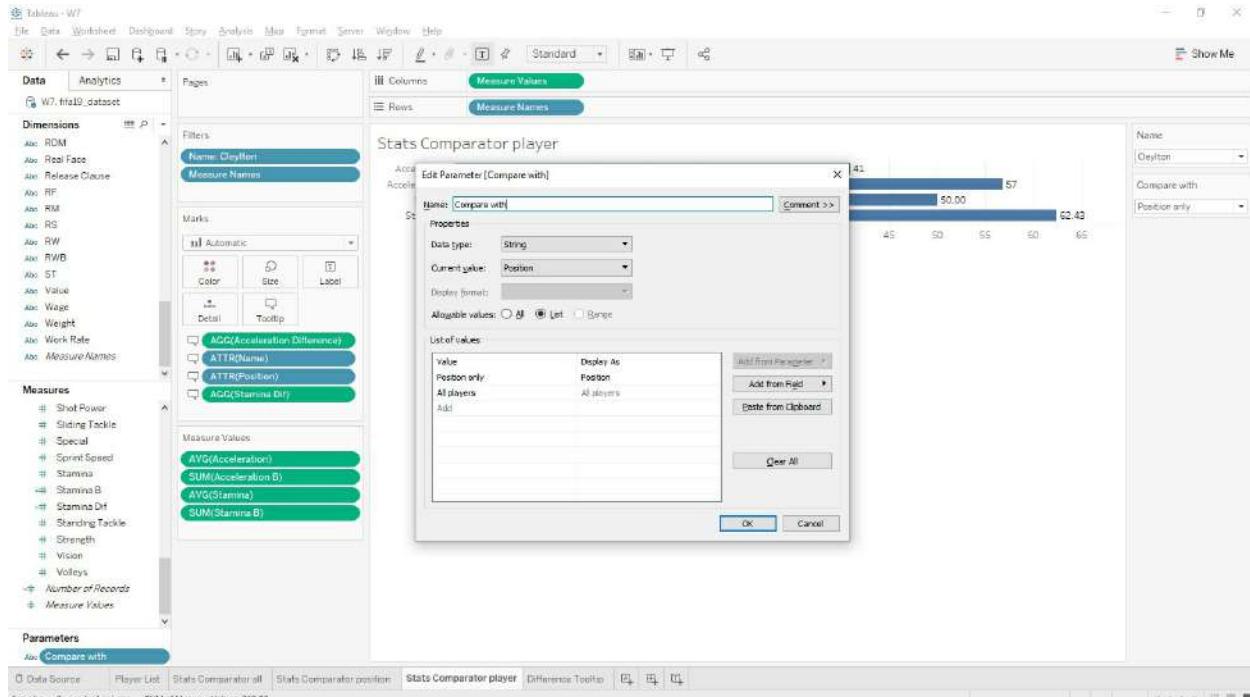
To give this a solution, we will add as a tooltip a table with all the

differences, which for now are **Acceleration Difference** and **Stamina Difference**. Create a new sheet called *Difference Tooltip*, drag **Measure Names** to rows, apply a filter to show our two difference measures, drag **Measure Values** to Text inside Marks. Go back to *Stats Comparator player*, open the tooltip, remove the part of the second row where it's referencing **Acceleration Difference**, and instead insert the sheet we just made on that row.

You should have this result on the Preview:



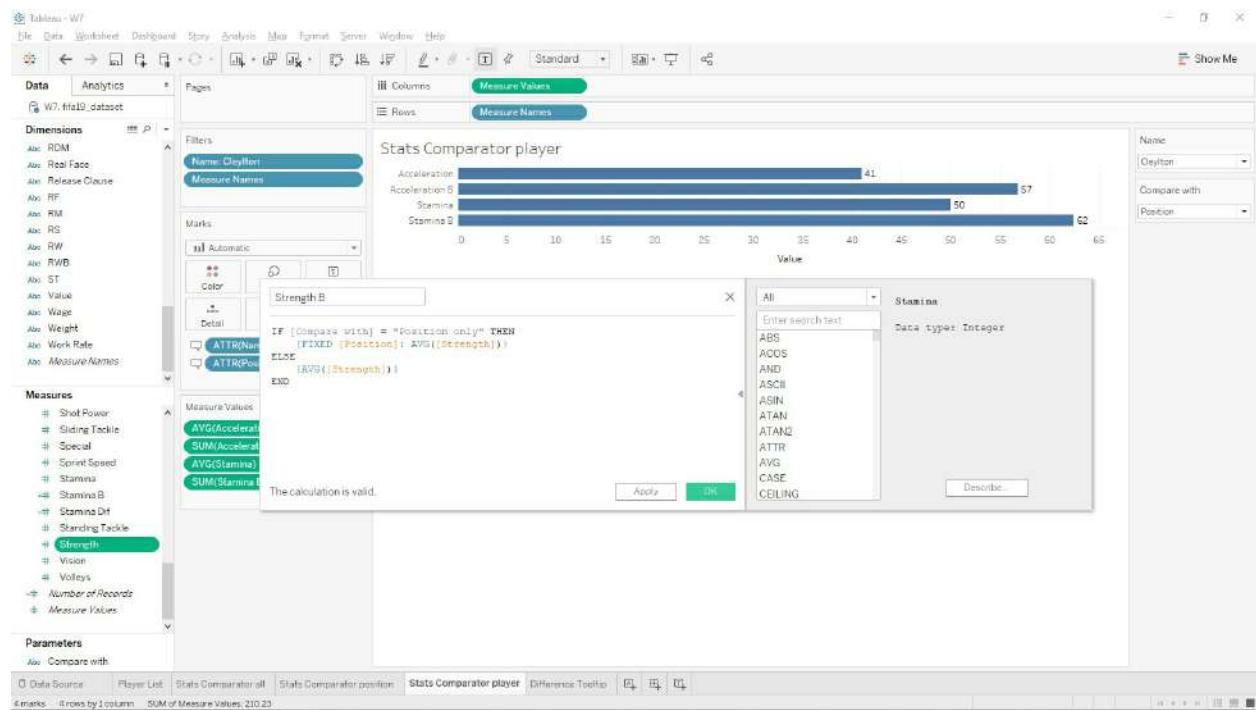
It would be better if instead of saying “vs Position only average” it said, “vs Position average”; edit your parameter, and on the Position only line, inside the display as column, put “Position”.



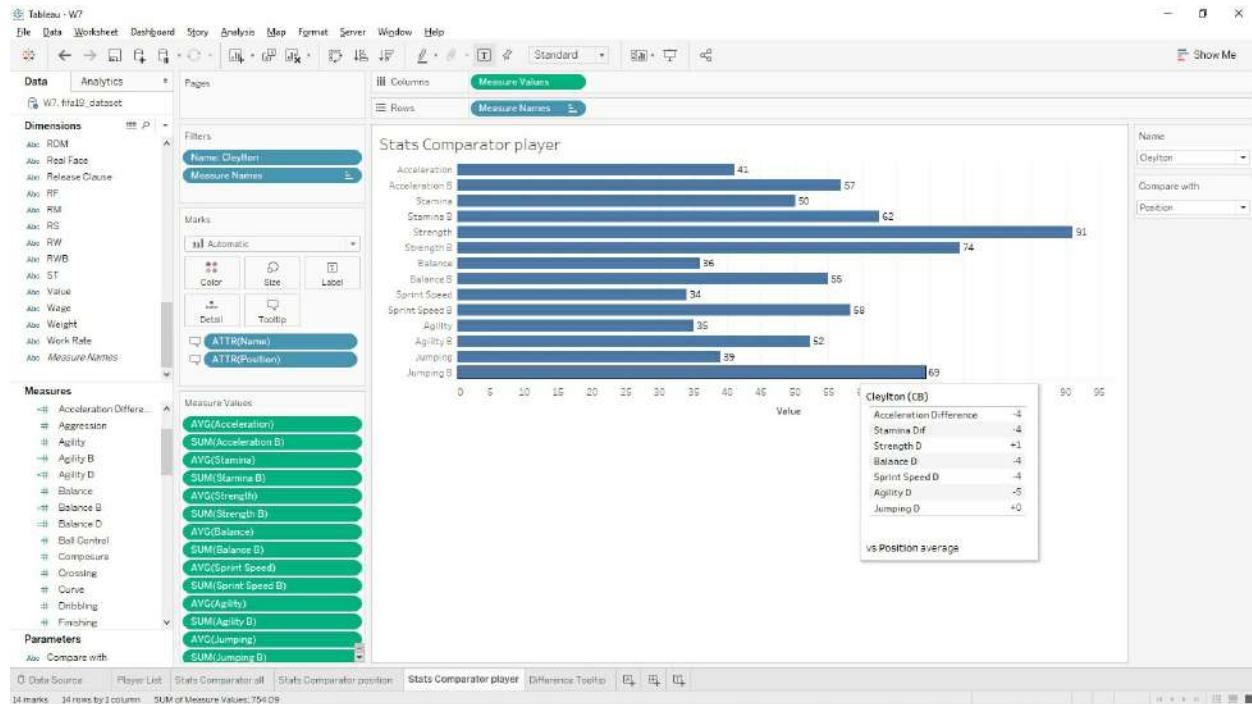
You can remove the Acceleration Difference and Stamina Difference measures from the Marks pane, since you’re not using them.

Now for the tedious part: using the Shift key, apply the same default number formatting for all your measures inside the Stats folder, remembering to exclude the difference measures because those have a custom “+” for the positives. Then also apply that custom format for Stamina Difference.

Finally, repeat the creation of the **B** and **Difference** measures for at least the next measures: **Strength**, **Balance**, **Sprint Speed**, **Agility**, and **Jumping**. Then add them to the *Stats Comparator player* and *Difference Tooltip* sheets. Copy and paste the B formula and then just replace the two AVGs for the next measure to go faster:

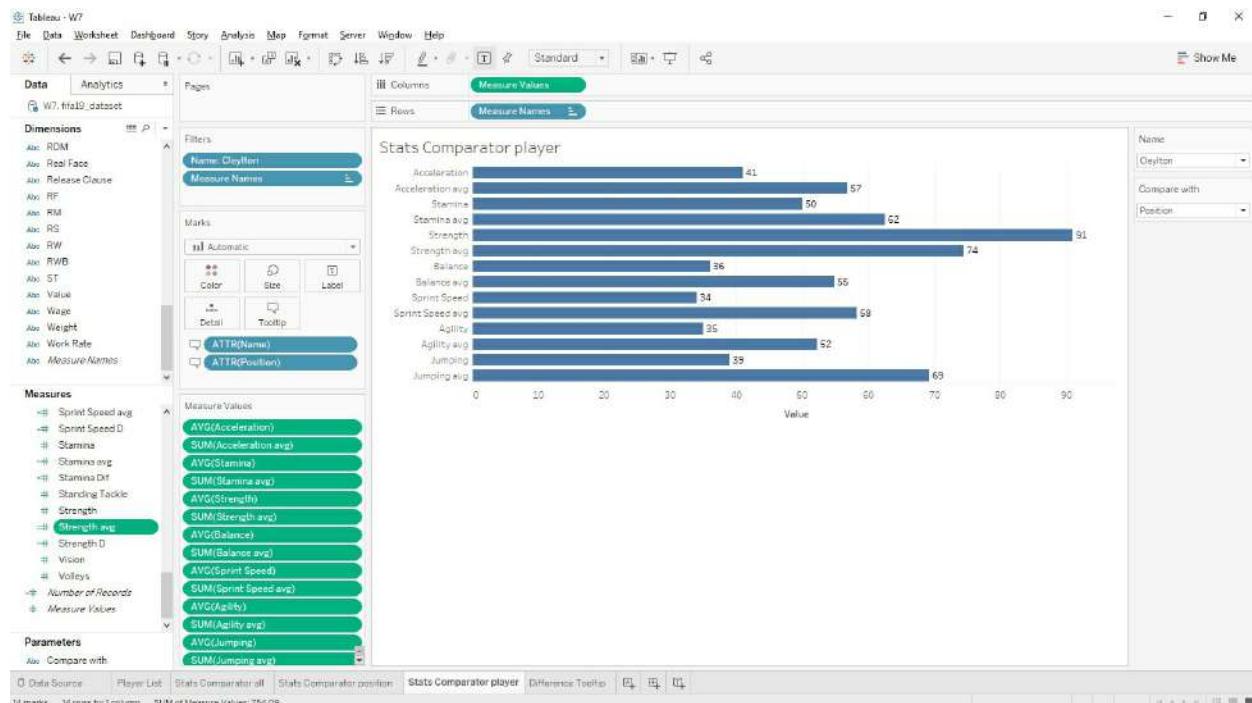


You can also wait until the end to apply the default number formatting to all in two clicks. You should end up with this:

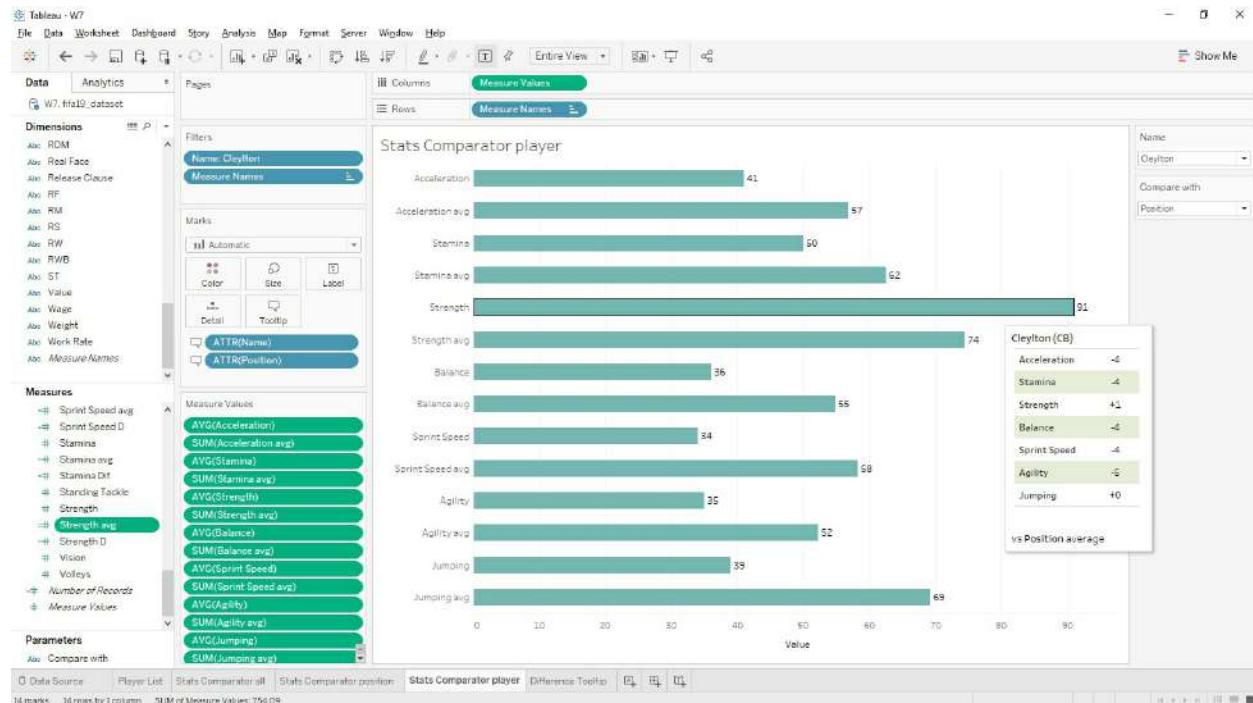


Go to *Difference Tooltip* and edit the aliases of your measures so instead of saying “Strength D” it says “Strength”.

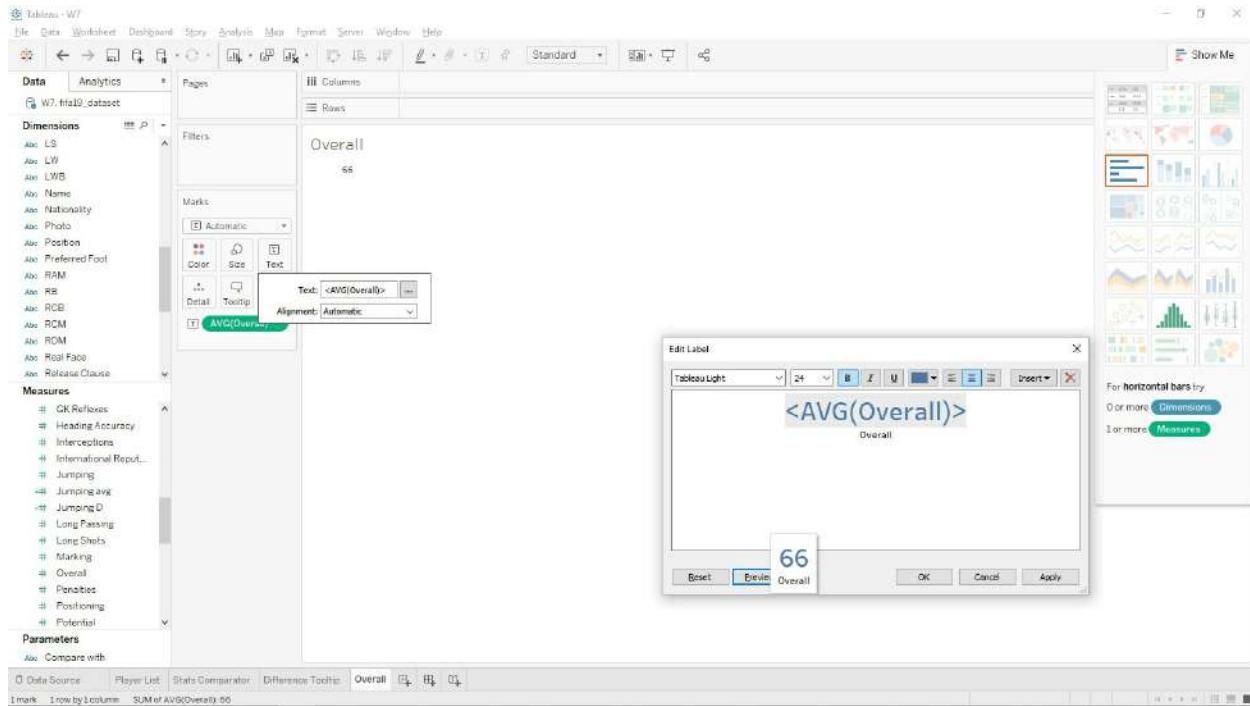
Also, in *Stats Comparator player*, change the aliases for the B measures so it says “Acceleration avg.” instead of “Acceleration B”. You can either change the aliases by right-clicking the axis on the sheet or by changing the original name of the calculated value (I will do the second):



Set the graph to fit Entire View, reduce the bars' size to half, and change the color if you like. On *Difference Tooltip* increase the rows' size and width so it looks less cluttered and add some formatting to it.



You can now delete the other two stats comparator sheets and remove the “player” part of our final sheet’s name. Now create a new sheet called **Overall**. Drag the **Overall** measure from the Stats folder to the sheet, then apply text formatting:



Now do another sheet called *Nationality*, convert the **Nationality** dimension to Geographic > Country/Region, drag it to the Rows shelf, and convert the graph to filled maps from the Show Me! button. Go back to *Stats Comparator* and apply the Name filter to all sheets but *Player Selector*. You will notice that the difference tooltip numbers are now showing correctly (if you didn't notice, then this is a good reminder to always review your results).

Go to *Nationality* and fix the map position using the controls on the top-left of the sheet so you can see it completely and it doesn't zoom.

Now create the *Title* sheet, add the Name filter from the previous sheets, then add **Name**, **Age**, and **Nationality** to the Text property and apply formatting to it.

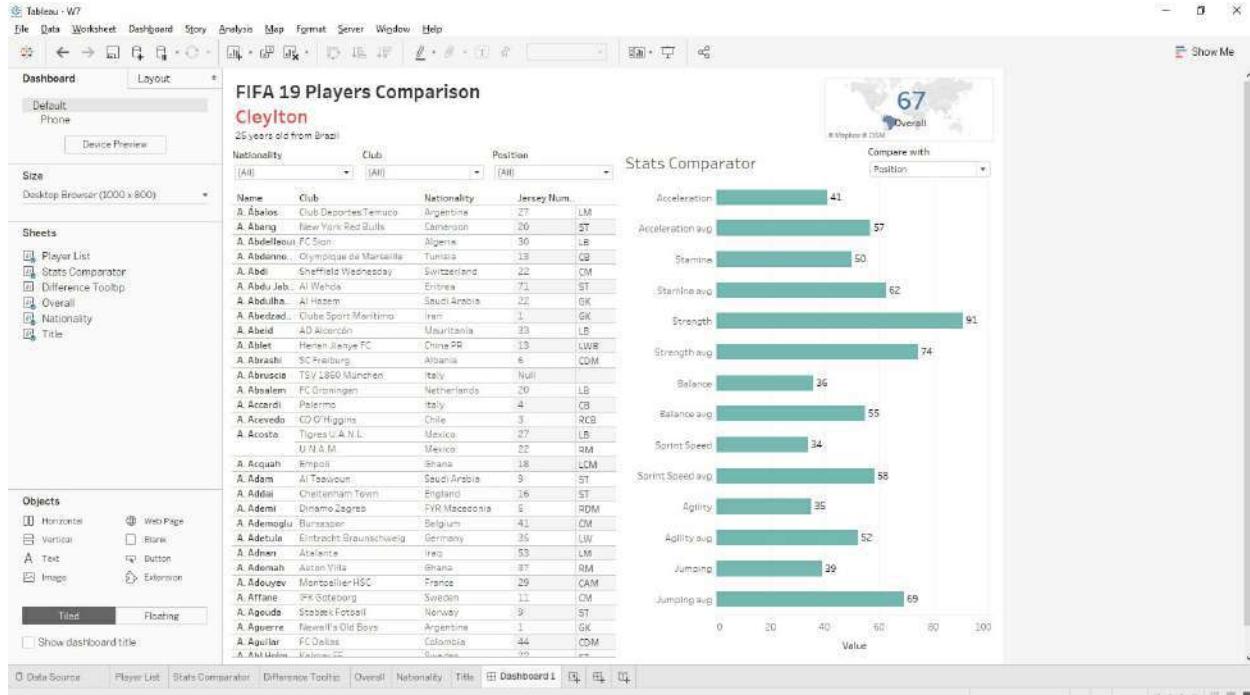
The screenshot shows a Tableau desktop interface with a worksheet titled "FIFA 19 Players Comparison". The title card displays "Cleyton" and "25 years old from Brazil". The stats comparator graph on the right shows various player statistics like Acceleration, Stamina, Strength, Balance, Sprint Speed, Agility, and Jumping, each with a corresponding green bar chart.

Time for the dashboard. Create a new one, and do this layout based on what you know about tiled dashboard design:

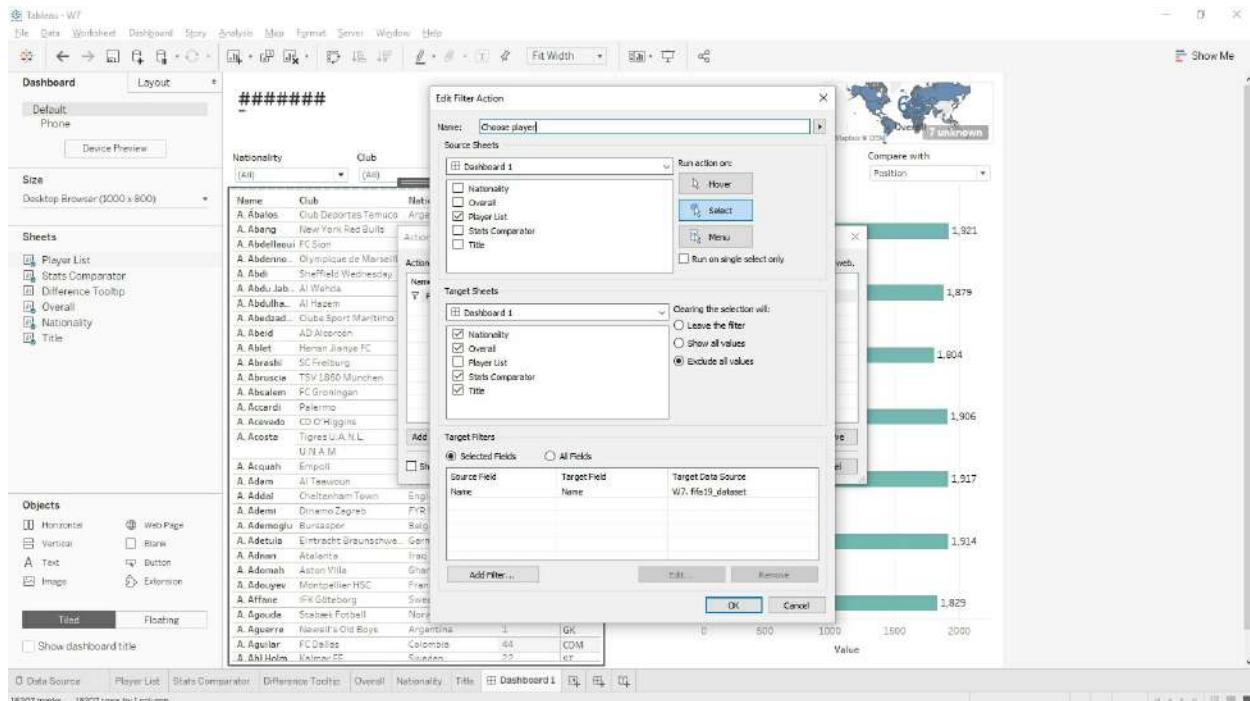
The screenshot shows a Tableau dashboard titled "FIFA 19 Players Comparison" for player "Cleyton". The dashboard includes a title card with "Cleyton" and "25 years old from Brazil", a "Nationality" filter, a "Stats Comparator" section with horizontal bar charts, and a "Player List" table showing player statistics like Name, Club, Nationality, Jersey Num., Position, and various performance metrics.

Click the arrow on Compare with and select Floating. Put it inside the Stats comparator graph, to the top-right corner, and remove the Name filter. Hide the sheet titles for all sheets except Stats Comparator. Adjust the width of the Nationality sheet so it's smaller.

Drag the *Overall* sheet between the *Title* and *Nationality* sheets, make it floating as you did with your parameter, and put it above the *Nationality* sheet. Format the sheet so it does not have that white background in Shading > Sheet tab > Default > Worksheet.



Toggle the Use As Filter option for *Player List*, and go to Dashboard > Actions to edit it so it is configured like below:



What we did is:

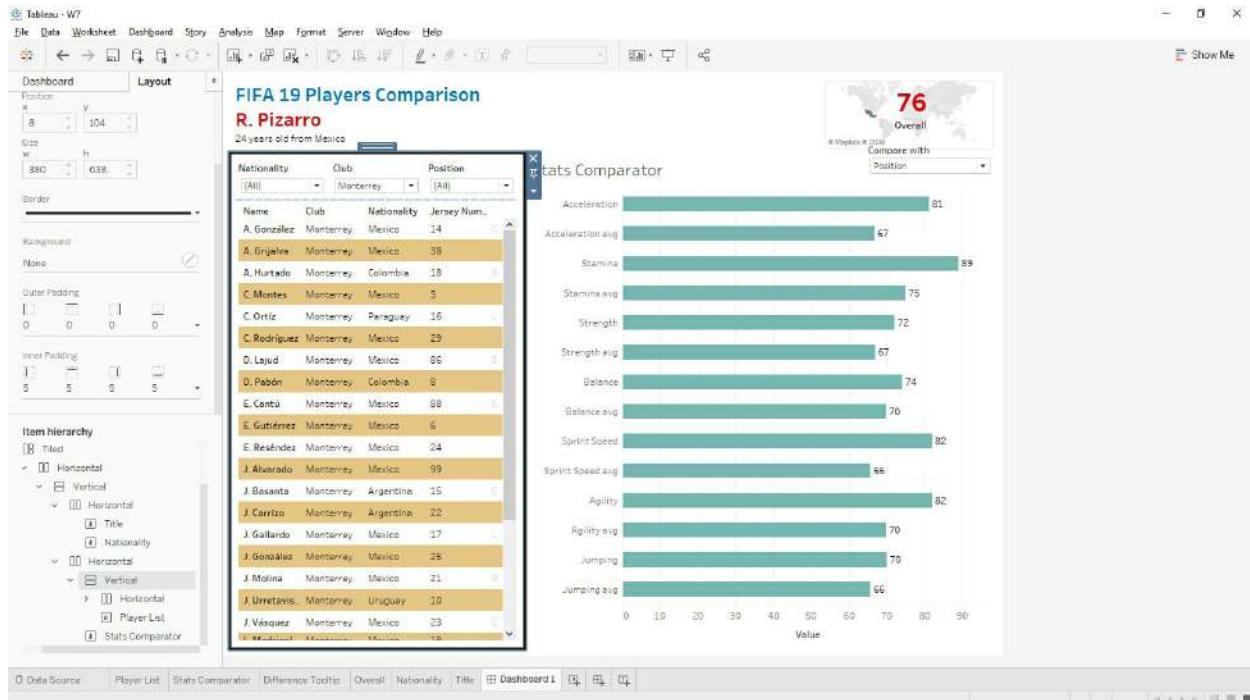
- 1) Make that it applies to all sheets except *Player List*, because if it applies to it, then we can't choose another player.
- 2) Adjust so that when the user deselects a player it doesn't remove the filter, but rather exclude all values so the sheets disappear until the user selects another player.
- 3) Make that this sheet can only apply a **Name** filter on the other sheets, and nothing on Club, Nationality, or the other dimensions in the table.

One last thing: for each of the three filters above the players list, right click and toggle Only Relevant Values. What this will do is not show the user any teams that do not apply when he/she is already filtering by **Nationality** or **Position**. Example: if I am looking on Afghan players, it doesn't make sense that on the Club dropdown I get clubs where no Afghan player is.

Now select any player on the list and see how it works. Also toggle the **Compare with** parameter and make sure everything works. If you don't get the scroll bar for *Player List*, go back to the sheet and set the size to Standard and not Fit Width. If you see your horizontal container of the top increasing in size by itself when changing filters, fix its height so it stops moving.

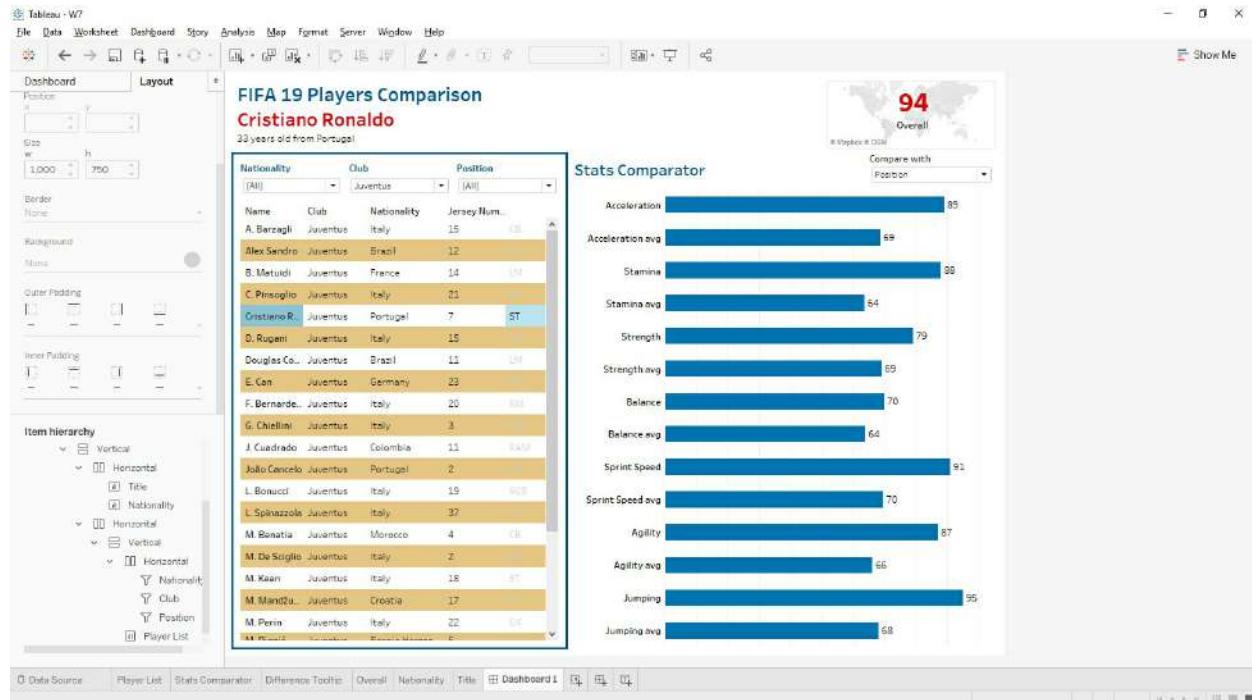
Now on to the formatting last steps:

1. Reduce the height of your dashboard to 750 or to a size that makes it fit your screen without a scroll bar.
2. Using the next color palette, format your header: #d30208 (red), #e5c685 (light brown), #171714 (black), #015386 (dark blue), #0074b1 (blue) so it looks like the image below. Edit *Title* and *Overall*.
3. On *Player List*, modify the Row banding so it's to the full row, and use the light brown for it instead of gray. Also remove the row and column divider borders and increase a bit the row height.
4. On the Vertical container that holds the player list, apply a thick border with the black color of our palette, and an inner padding of 5, making sure the width is still enough to show the **Position** column, like in below image:



5. For the Horizontal container that holds the three filters for players list, change the title of the three filters to be of our blue color.
6. On Stats Comparator, change the title to be of dark blue color and bold. Hide the Value axis and for the Measures axis make the labels bold. Finally, change the bars color to be the blue of our palette.
7. For the Nationality map, change the color to our light brown and eliminate the automatic border.

You should have this as result:



Congratulations, you have now finished your seventh dashboard! The topics covered were:

- Deep calculated values development and test
- Parameter-based calculations
- Graphs in tooltip
- Folders
- Floating items in dashboards

Workout 8. Olympics dashboard

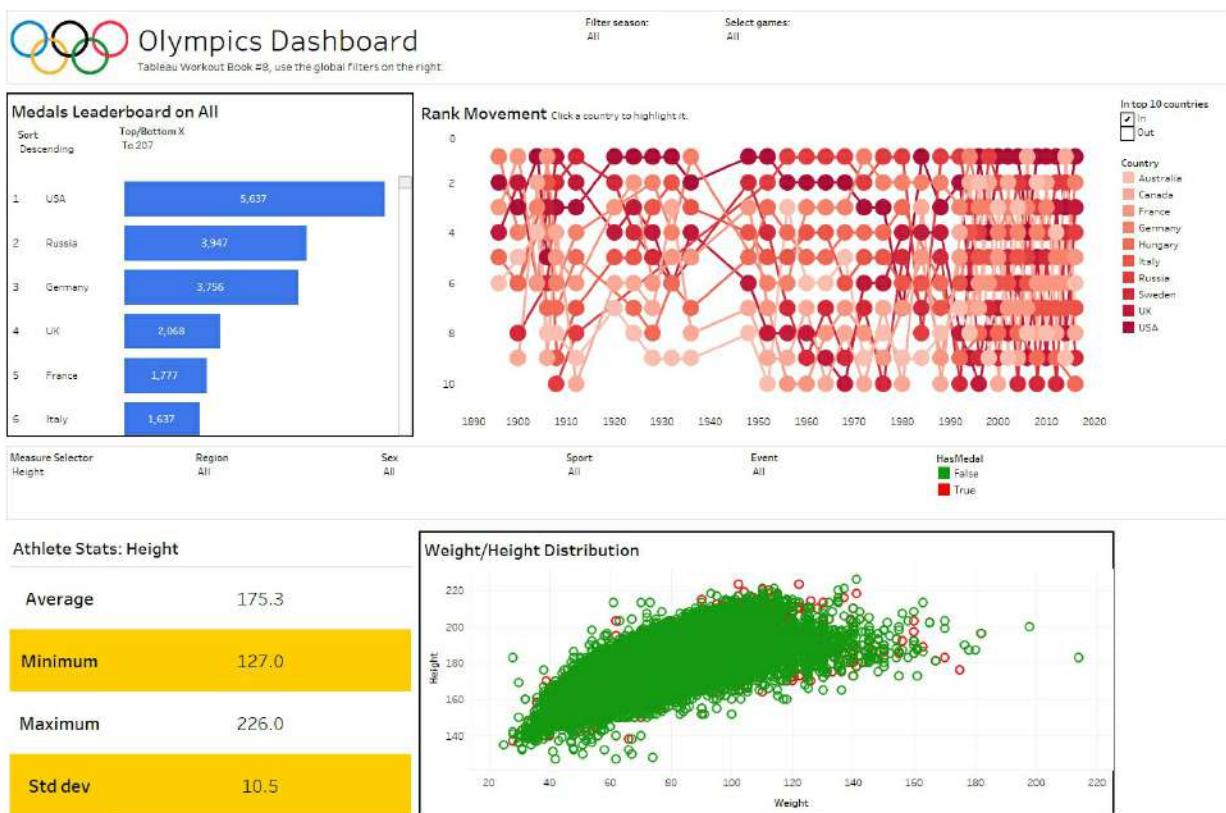
On this workout you will develop a dashboard containing all relevant information about the history of Olympics; which countries have made the most medals, which countries are the best in which events, the average athlete's stats by event, and the correlation of physical attributes and sport.

Data

Design

Graphs

Interactivity

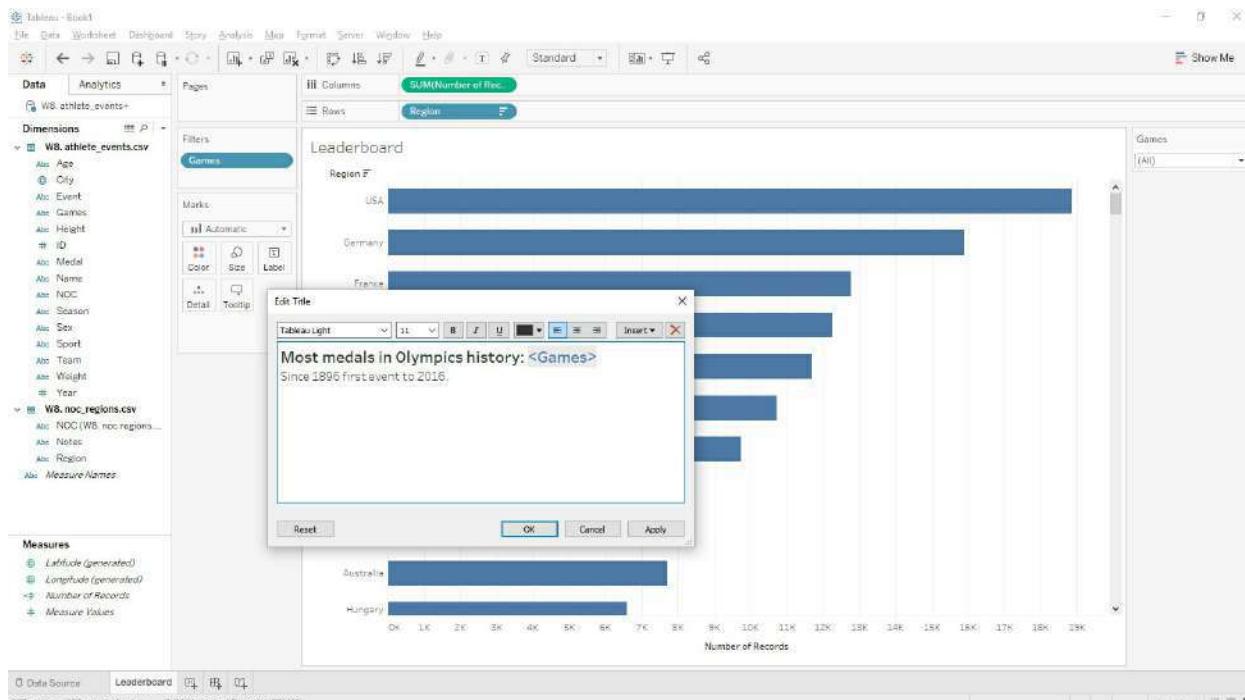


Workout Start

Connect to W8. Athlete events.csv and join to it W8. Noc_regions.csv by NOC. We do join because the Noc regions table does not have duplicates.

The screenshot shows the Tableau Data Source interface. On the left, there's a sidebar with 'Connections' and 'Files'. Under 'Files', several CSV files are listed, including 'W8.athlete_events.csv' which is currently selected. A note says 'Data Interpreter might be able to clean your Text file workbook.' Below this are other files like 'W3.Arribb.listings.csv', 'W3.Arribb.reviews.csv', etc. In the main area, two tables are shown: 'W8.athlete_events.csv' and 'W8.noc_regions.csv'. A 'Join' button is active, and the 'Left' option is selected. The 'Data Source' dropdown shows 'W8.athlete_events.csv' and 'NOC (W8.noc_regions...)'. A note at the bottom says 'Add views from closure...'. At the bottom, there are buttons for 'Go to Worksheet' and 'Sheet 1'.

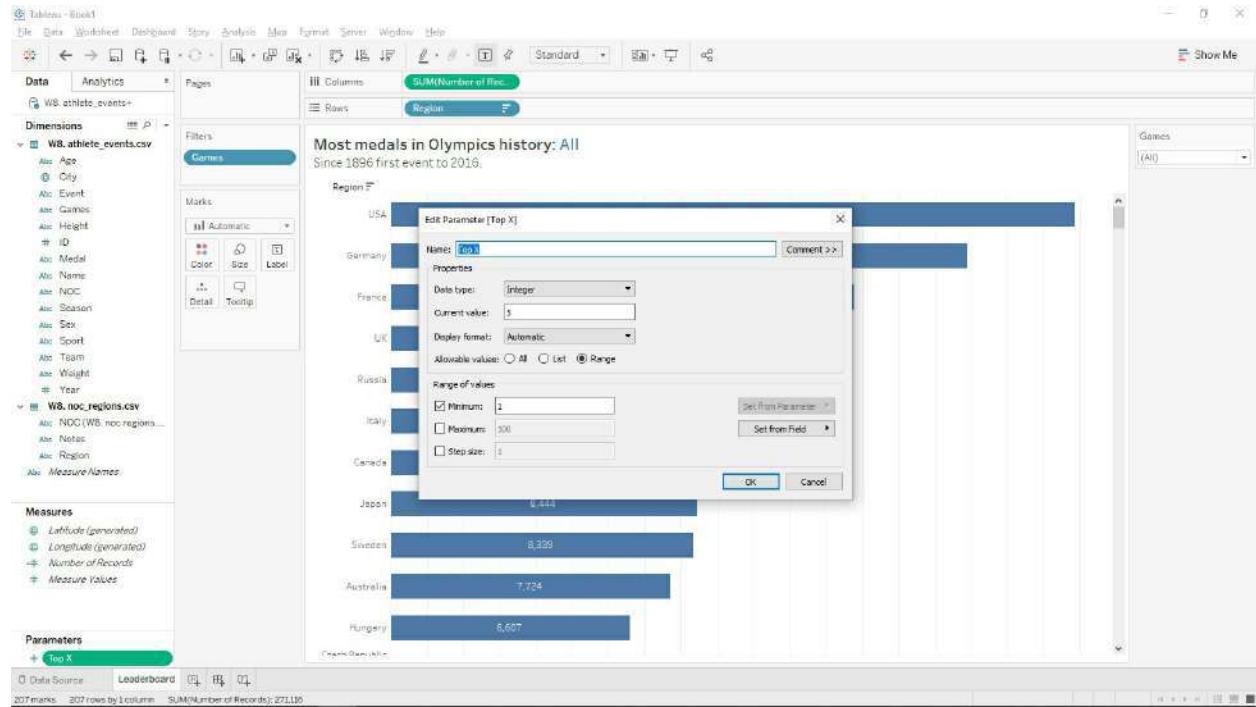
Start by making a sheet called *Leaderboard*. Move **Region** to Rows, **Number of Records** to Columns. Order **Region** by **Number of Records** descending. Add a **Games** filter and show as Single Value (dropdown). Increase the width of the bars and reduce the bars' size to a bit above half. Finally, edit the sheet title to look like below:



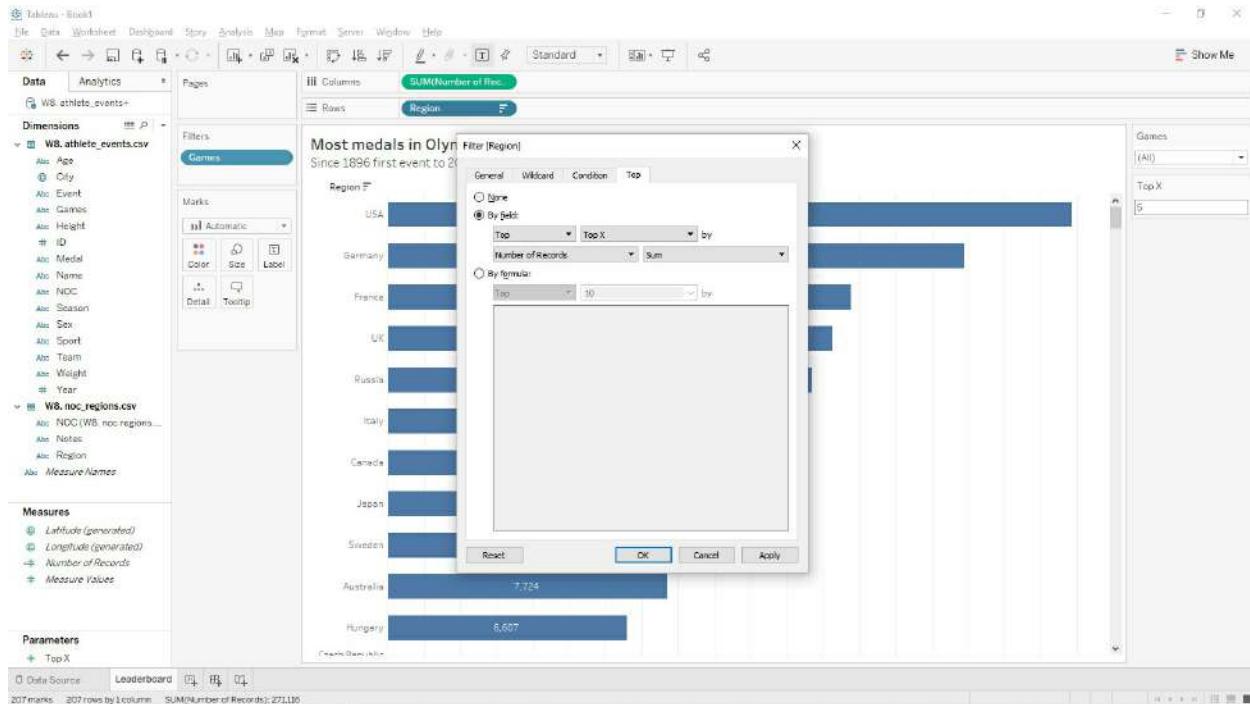
Show the mark labels from the Label property and align them horizontally to

be in the middle (inside the bar). Finally, hide the x-axis.

Now let's add some parameters and filters to make it more interactive. Start with making a parameter called **Top X** of value int, set 5 as current value and for allowable values define a range that only states minimum as 1.



Now add a filter on **Region** and set it as top filter by our **Top X** parameter, based on the sum of **Number of Records**.



With this, the users can now apply a top filter on our leaderboard to just see top 5 countries. In case they don't want to filter, they can type a big number like 1,000 so they see all of them.

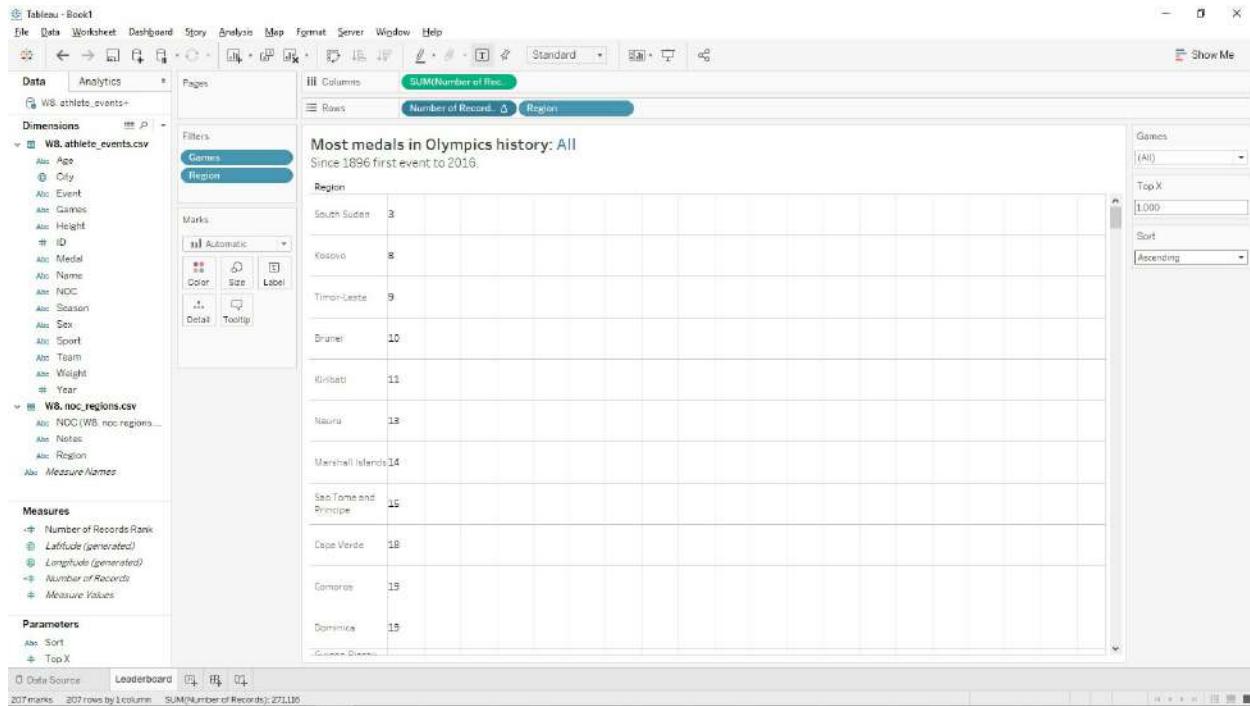
Next feature: create a parameter called **Sort** of type string that only allows values “Descending” and “Ascending”. Then create a calculated value called **Number of Records Rank** that is based on the next formula:

```

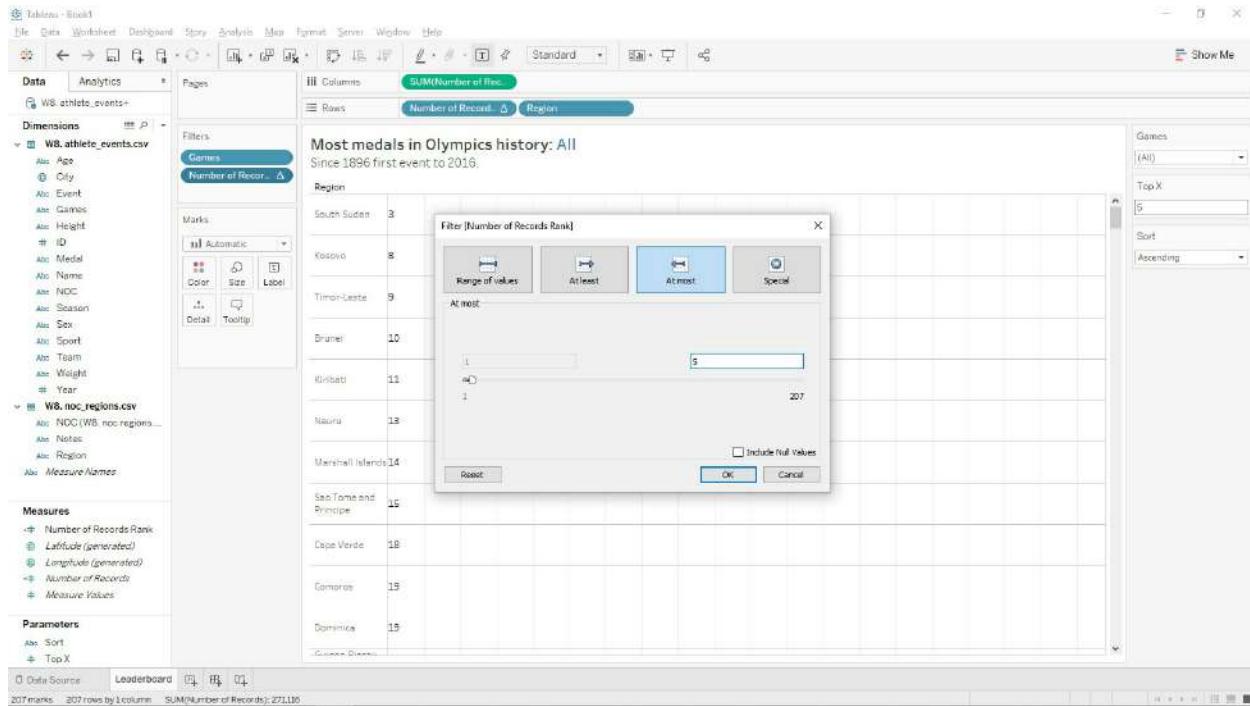
IF [Sort] = "Descending" THEN
    RANK(SUM([Number of Records]),'desc')
ELSE
    RANK(SUM([Number of Records]),'asc')
END

```

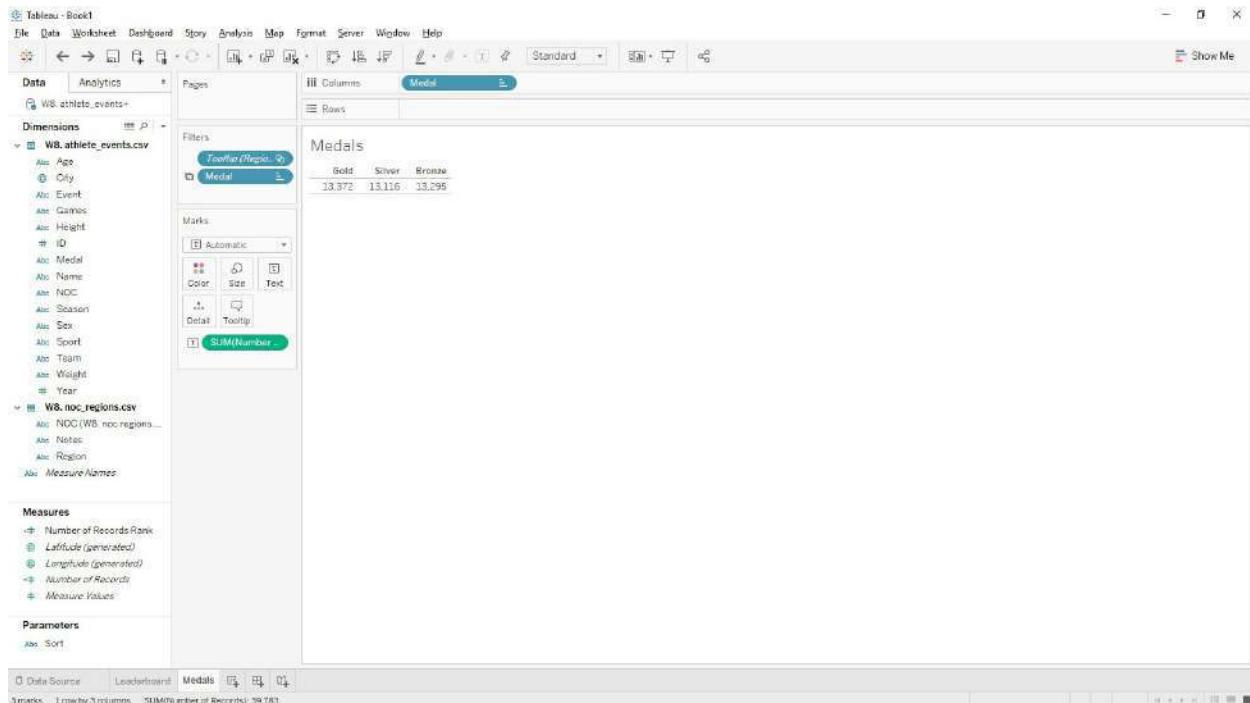
What this does is allow the user to change the sorting through a parameter. Make **Number of Records Rank** discrete and drag it on Rows right before **Region**, then clear the sort on **Region**. Finally, show the **Sort** parameter and test it (also change the **Top X** parameter to make sure both work together).



If you did the testing, you will see that when you filter **Top X** to 5 or a small number and have **Sort** set as Ascending, you are seeing the top countries instead of the bottom, sorted ascending. We want to have the bottom countries when **Sort** = Ascending, so to change this, remove the **Region** filter and instead add a **Number of Records Rank** filter. Convert the filter to continuous and set it to be of type At Most, like below. With this, you can delete the **Top X** parameter.

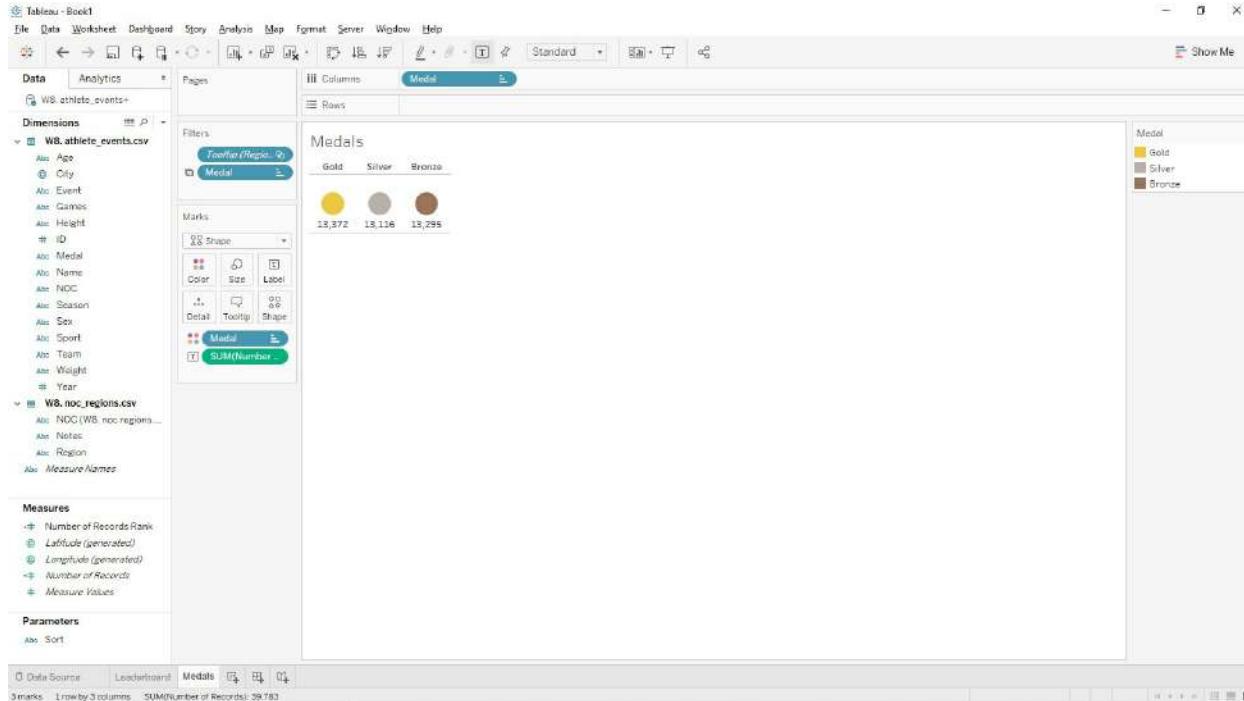


Create a second sheet called *Medals*, drag **Medal** to Columns, **Number of Records** to Text in Marks, hide the Medals column name from the sheet, exclude Medal = N/A (put this filter also on *Leaderboard*), and sort them in gold-silver-bronze.

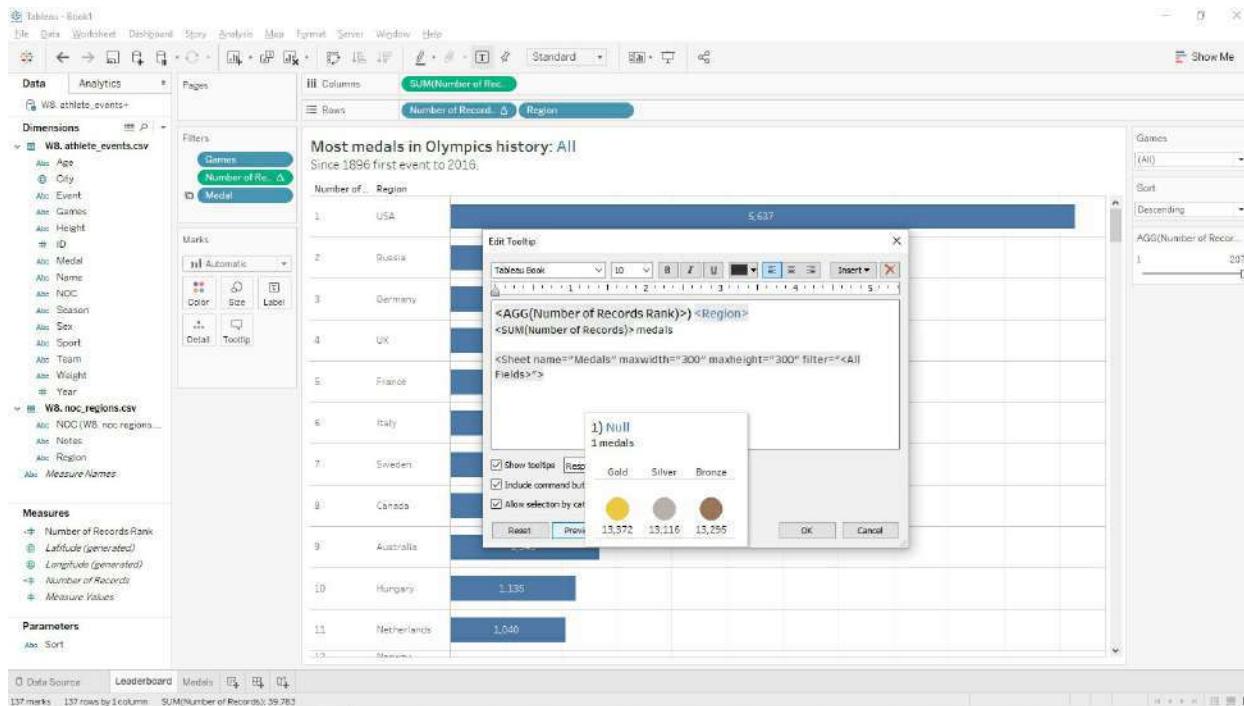


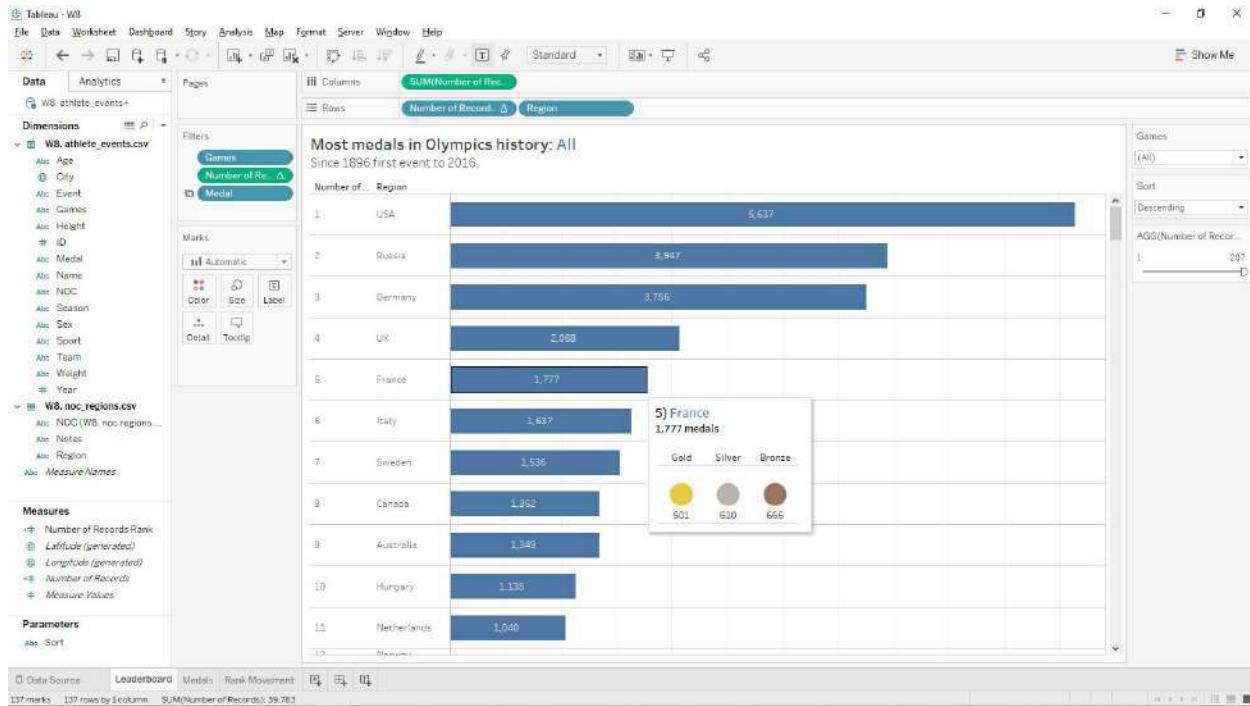
On the Marks dropdown, select Shape, and while holding Ctrl, drag Medal from Columns to Color in Marks. Change the colors so that gold is yellow,

silver is gray, and bronze is brown. Change the Shape property to be a filled circle instead of an empty one and expand the row height so it's bigger. Finally, change the vertical alignment on the Label property to bottom and horizontal alignment to center. It should be looking like this:

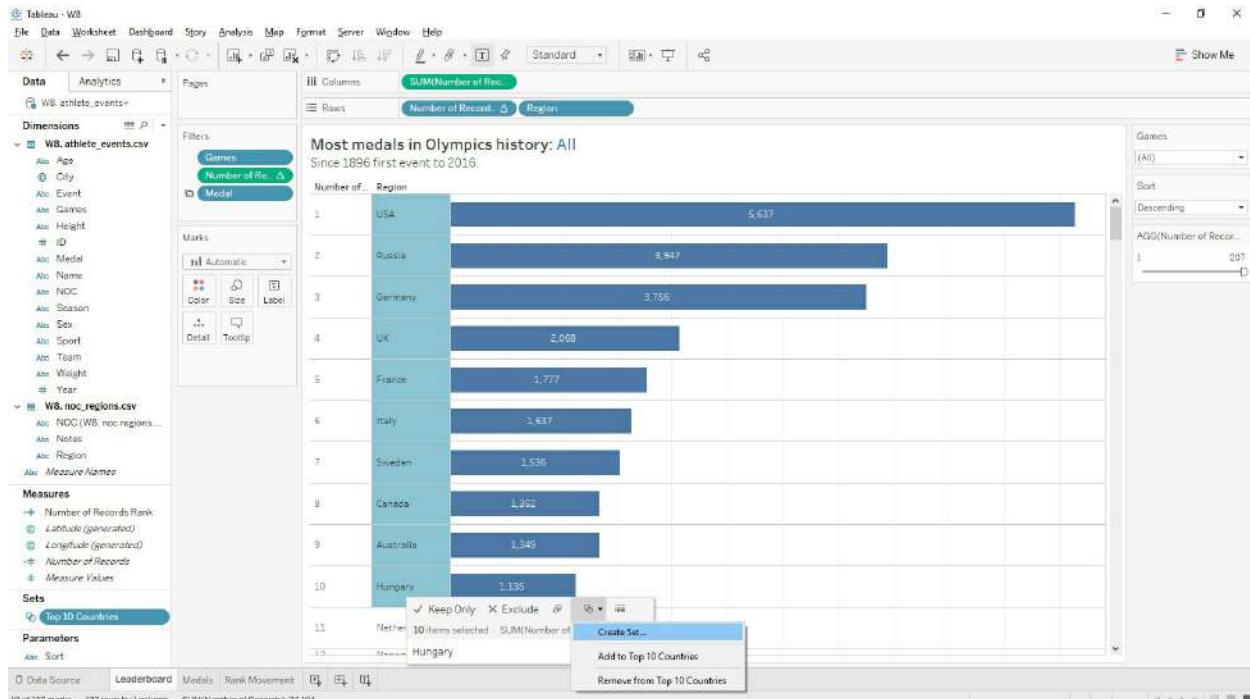


Now, add it as a tooltip on *Leaderboard*:



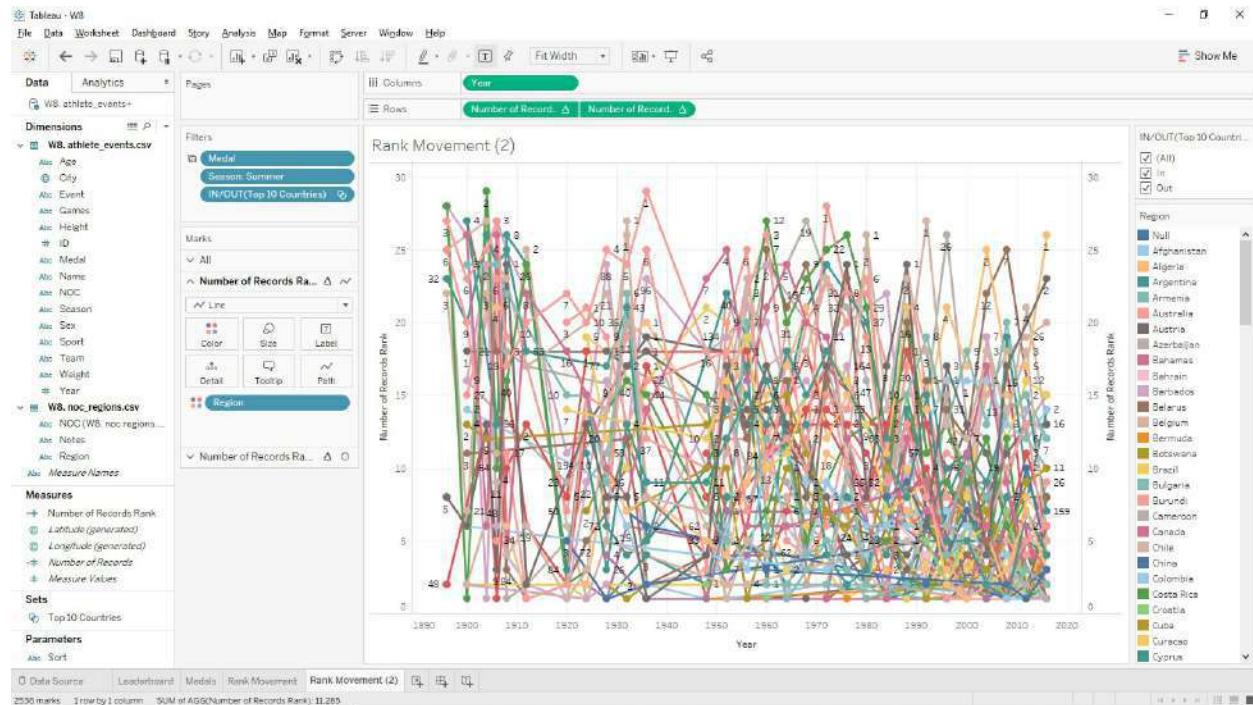


Select the top 10 regions, right click and create a new set called **Top Countries**.

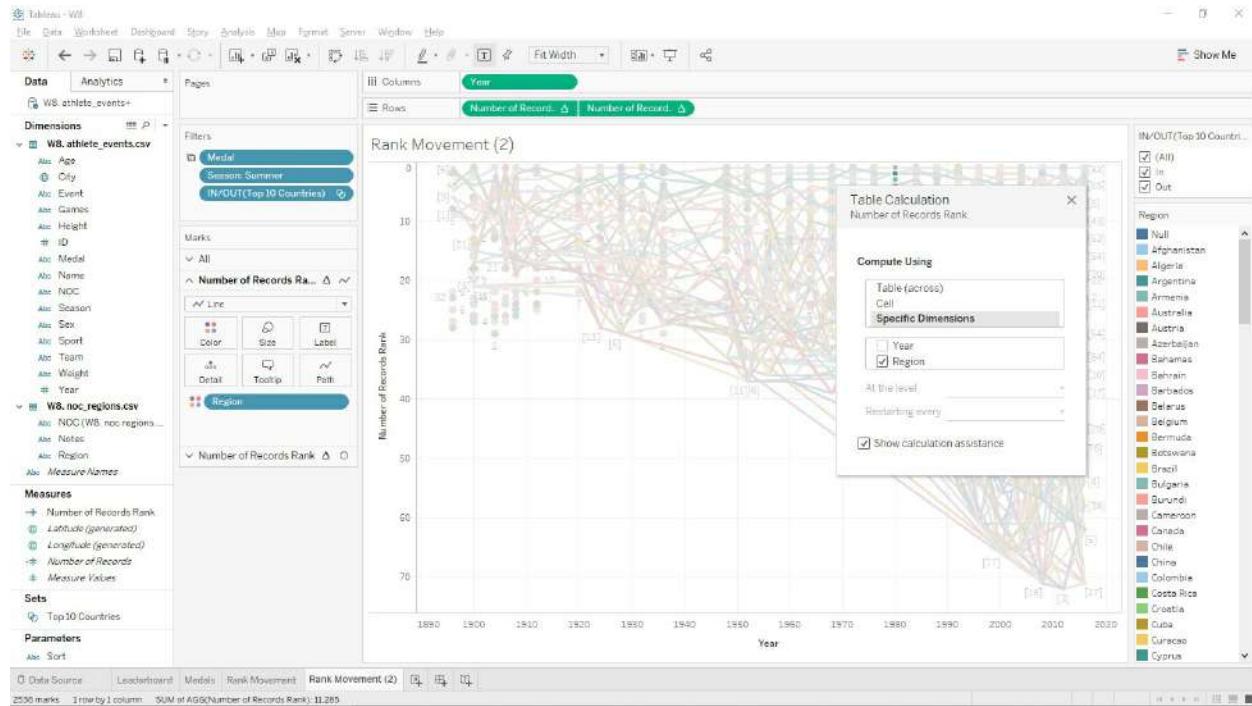


Now let's do the third sheet: *Rank Movement*. Apply the same Medal filter you have on the other sheets, add the set we just created also as a filter, and set Season = Summer.

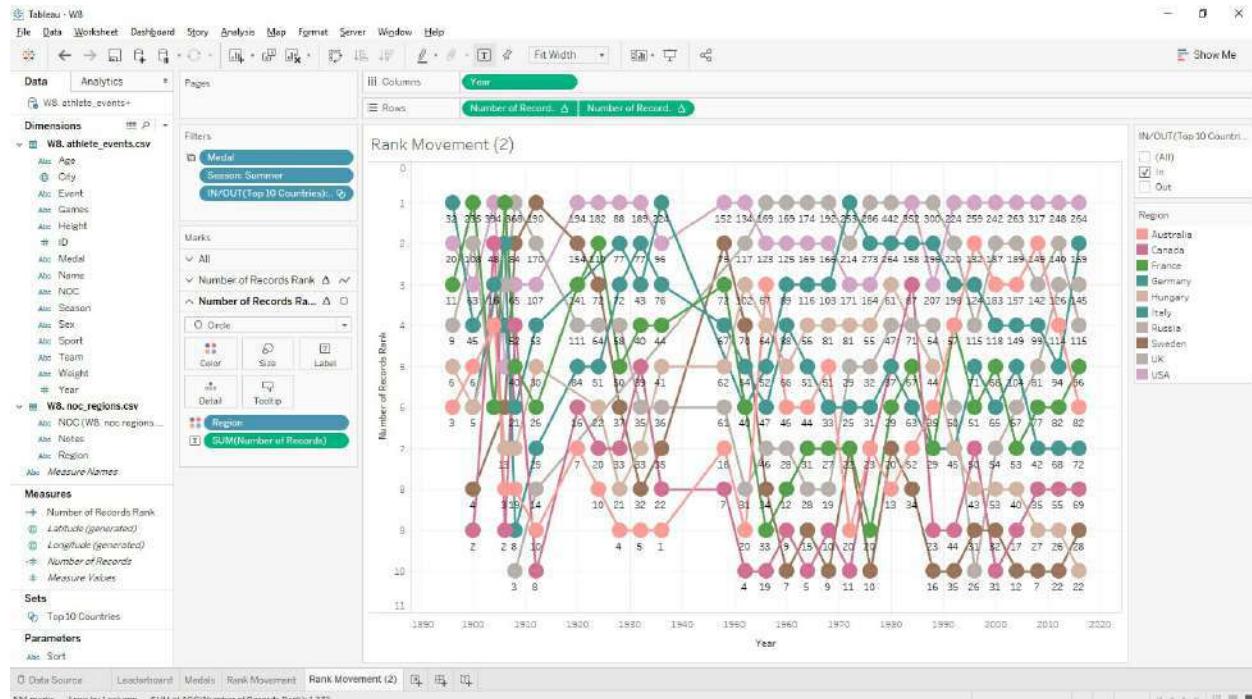
Drag **Year** to Columns and **Number of Records Rank** to rows. Convert the last one to continuous right from the Rows shelf, then holding Ctrl drag it to the right so you have it two times. Set that second pill as Dual Axis, also toggle Synchronize Axis so both y-axes are on the same scale. On the Marks card, notice you have three tabs: All, Number of Records Rank, and Number of Records Rank (2). Open All and drag **Region** to Color. Open the second tab and from the dropdown select Line, then open the third marks tab to select Circle and drag **Number of Records** to Text.



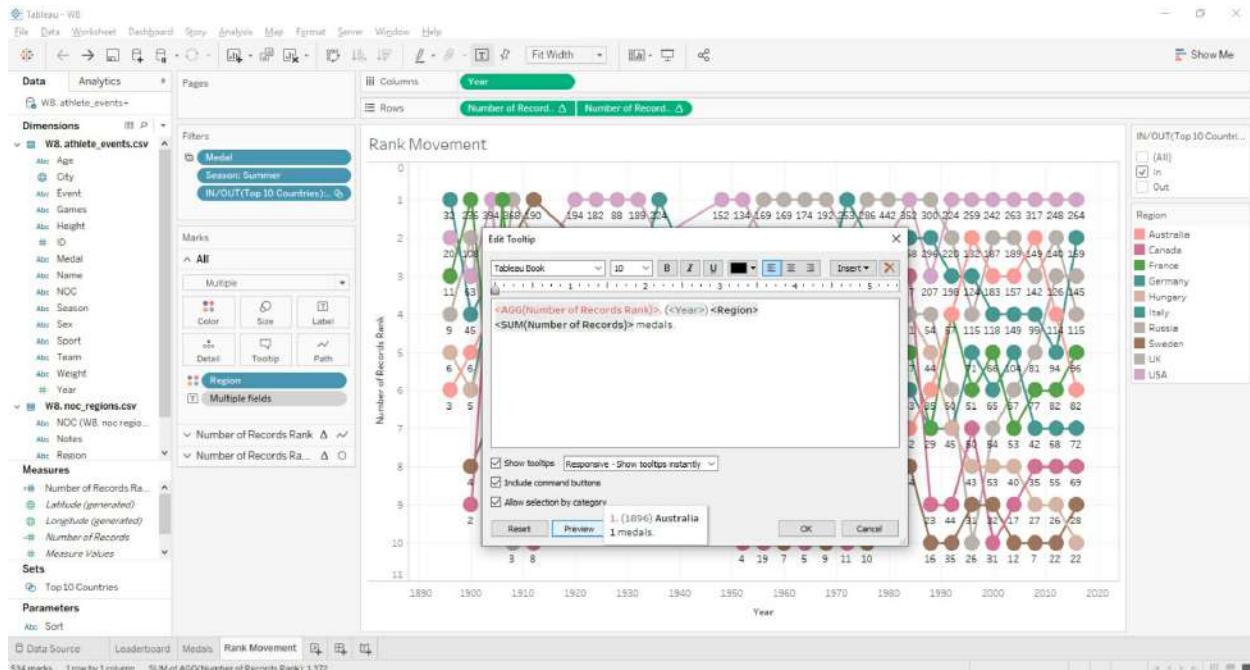
Hide the second y-axis on the right, reverse the scale of the first y-axis on the left, and edit the table calculation of your two pills on Rows so instead of being calculated on Table (across) level it's on Specific Dimensions > Region (not Year).



Deselect “Out” from your set filter, notice how the graph becomes clearer. Adjust the circle size to half.



Modify the tooltip so it looks like below:



We basically have the graph ready, but it would be better if we could do something with the set filter, since it's very unintuitive. Change the filter title to “In top 10 countries” and remove the “All” option. Finally, change the y-axis name to “Place”.

Make the fourth sheet called “**Athlete Stats**”. Convert **Age** and **Height** from string to number (whole) and convert them to measures. Convert **Weight** to number (decimal) and convert it to measure as well.

Just like in the Airbnb workout #3, make a **Measure Selector** parameter and a **Measure Selected** measure with **Age**, **Height**, and **Weight**.

The screenshot shows the Tableau Data Editor interface. In the top navigation bar, 'File', 'Data', 'Worksheet', 'Dashboard', 'Story', 'Analysis', 'Map', 'Format', 'Server', 'Window', and 'Help' are visible. On the right, there's a 'Show Me' button. The main area displays a worksheet titled 'Athlete Stats'. The data pane on the left lists dimensions like 'WB_athlete_events.csv' (City, Event, Games, ID, Medal, Name, NOC, Sex, Sport, Team, Year) and measures like Age, Height, Number of Records, Weight, Latitude, Longitude, and Measure Values. A 'Dimensions' section is expanded. A 'Measures' section is also present. A 'Sets' section contains 'Top10Countries' and 'Top10Countries (copy)'. A 'Parameters' section includes 'Measure Selector' and 'Sort'. The 'Filters' pane has 'Measure Names' selected. The 'Marks' pane shows 'Automatic' with 'Color', 'Size', and 'Text' options. A tooltip window is open, showing the formula: `IF [Measure Selector] = "Age" THEN [Age] ELSEIF [Measure Selector] = "Height" THEN [Height] ELSEIF [Measure Selector] = "Weight" THEN [Weight] END`. To the right, a dropdown menu for 'Age' shows various mathematical functions: ABS, ACOS, AND, ASCII, ASIN, ATAN, ATANZ, ATTR, AVG, CASE, CEILING, CHAR, etc. Below the formula, it says 'The calculation is valid.' with 'Apply' and 'OK' buttons. A 'Describe...' button is also present.

Show the parameter, add **Measure Names** to Rows, filter it so it only shows **Measure Selected**, and drag **Measure Values** to Rows.

This screenshot shows the same Tableau Data Editor interface after the calculated field 'Age' has been created. The 'Measures' shelf now contains 'Age' and 'Measure Selected'. The 'Rows' shelf contains 'Measure Names'. The 'Measure Selected' pill is highlighted in blue. The 'Value' shelf shows a single value of 0.668575. The 'Measure Selector' parameter dropdown on the right is set to 'Age'. The rest of the interface is identical to the first screenshot, showing the 'Athlete Stats' worksheet and its data and mark definitions.

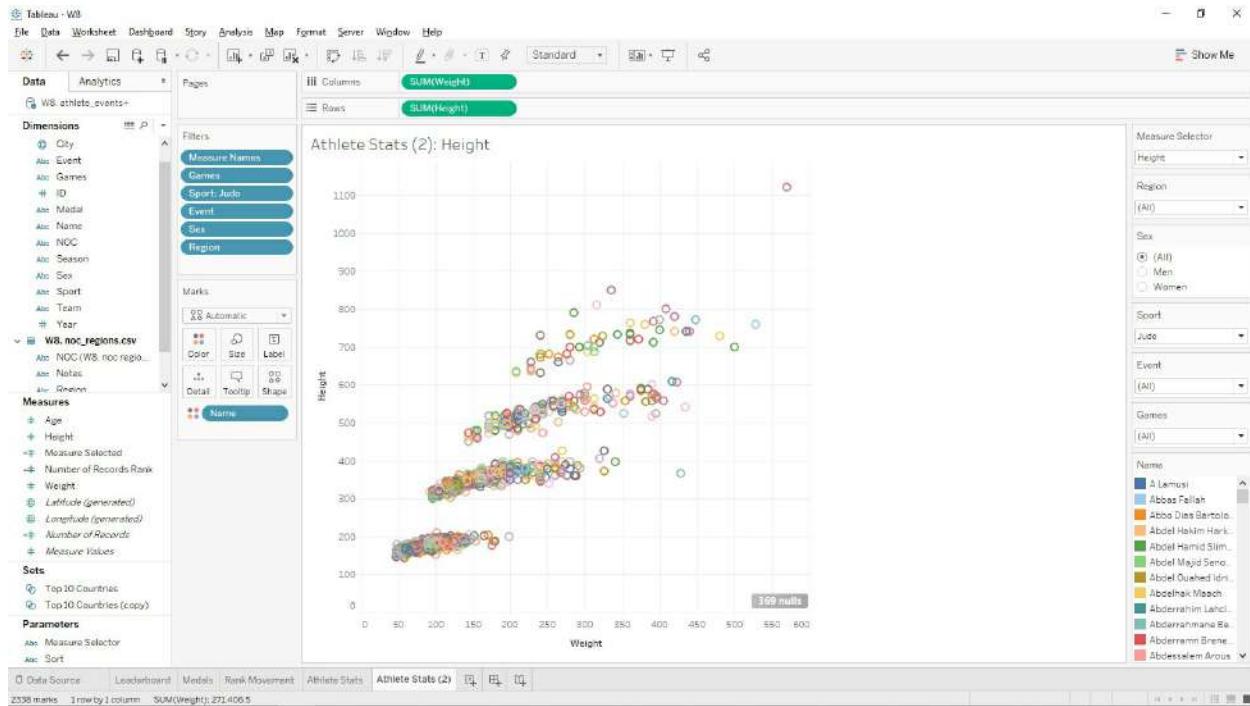
Drag **Measure Selected** with right click from the data pane to the Measure Values shelf and aggregate it as minimum. Repeat to also add maximum and standard deviation. Move **Measure Values** from Columns to Text. Finally, change the aggregation of the first pill on the Measure Values shelf from sum

to average and change the aliases for all of them to just be “Average”, “Minimum”, “Maximum”, and “Std dev”. Expand the row height as well and add the **Measure Selector** parameter in the sheet title.

Now add the next filters: **Games**, **Sport**, **Event**, **Sex**, and **Region**. Show them all as Single Value (dropdown) except **Sex**, do that as list and rename its value aliases F = “Women” and M = “Men”. Also, set all your filters to show Only Relevant Values.

Play with **Measure Selector** and your filters and see how you find interesting trivia, like that the oldest athlete to be in the Olympics was 97 years old, or that the average height of basketball athletes is 190cm while that of gymnastics is 162cm.

Finally, duplicate this sheet so you keep the filters, but drag out Measure Names and Measure Values. Then drag **Weight** to columns and **Height** to rows, changing the default aggregation from SUM to AVG.



Here you will be able to see the clusters of height/weight numbers among athletes for a certain sport. It's interesting to watch especially those sports that don't have any weight divisions like diving. Add the **Region** to the tooltip so you can also see the nationality of each athlete.

Hide **Measure Selector** from this view, fix both axes by setting filters to "All" then right clicking each (excluding the zeros), it is important because we don't want the scale to be moved when playing with the filters. Rename the sheet to *Weight/Height distribution*.

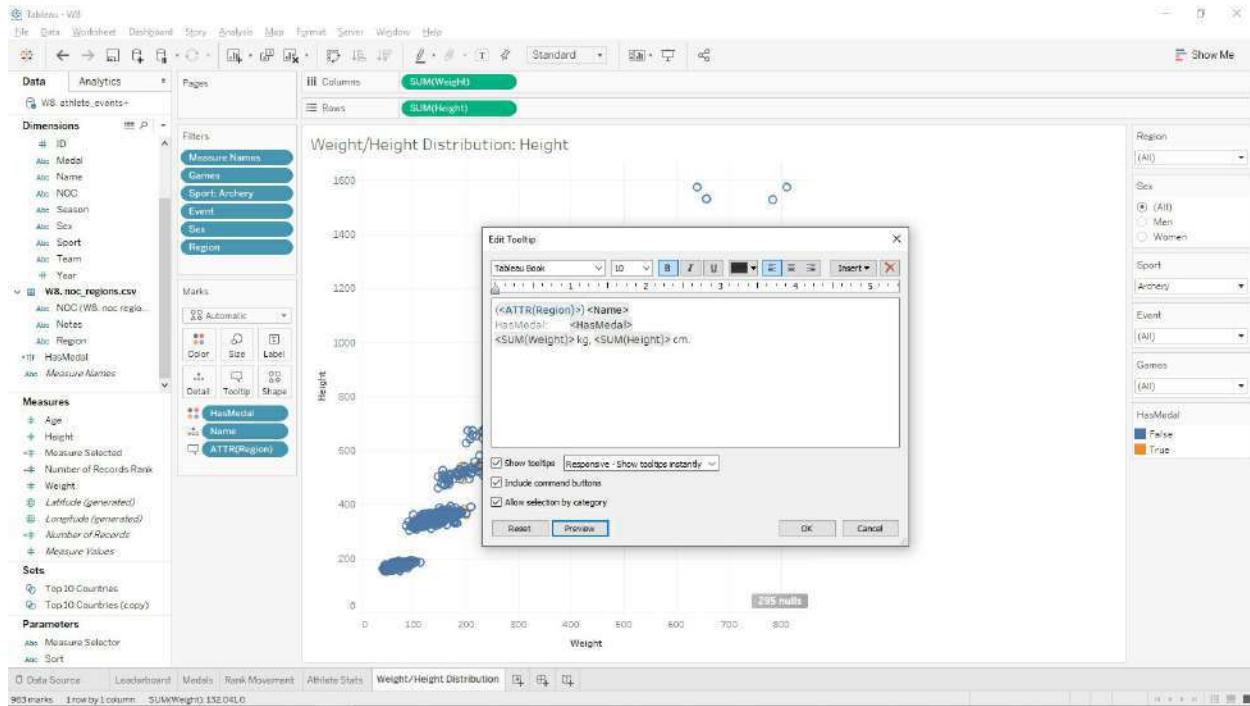
Having Names in Color is not very helpful... how about instead we have a Boolean that distinguishes the athletes with medals from those without?

Create a calculated field called **HasMedal** equal to:

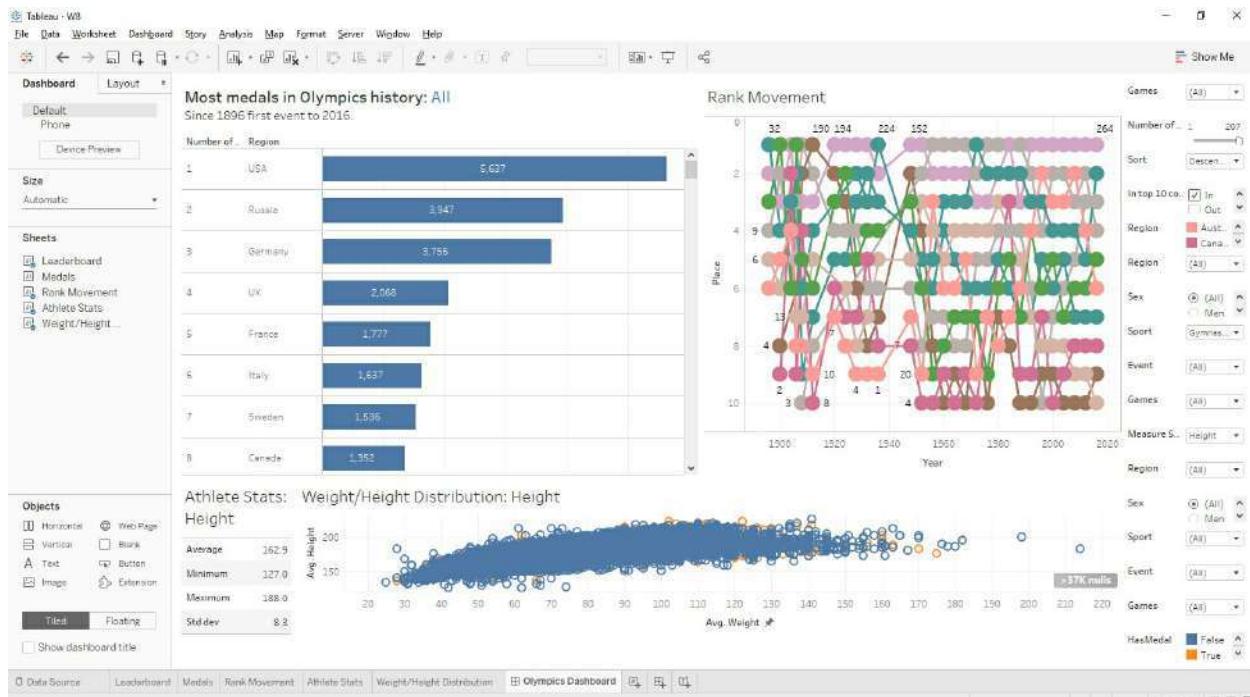
`IIF([Medal] = "NA", FALSE, TRUE)`

Then drag **Name** to the Detail property and put **HasMedal** on the Color.

Also, improve the tooltip so it looks more like below:



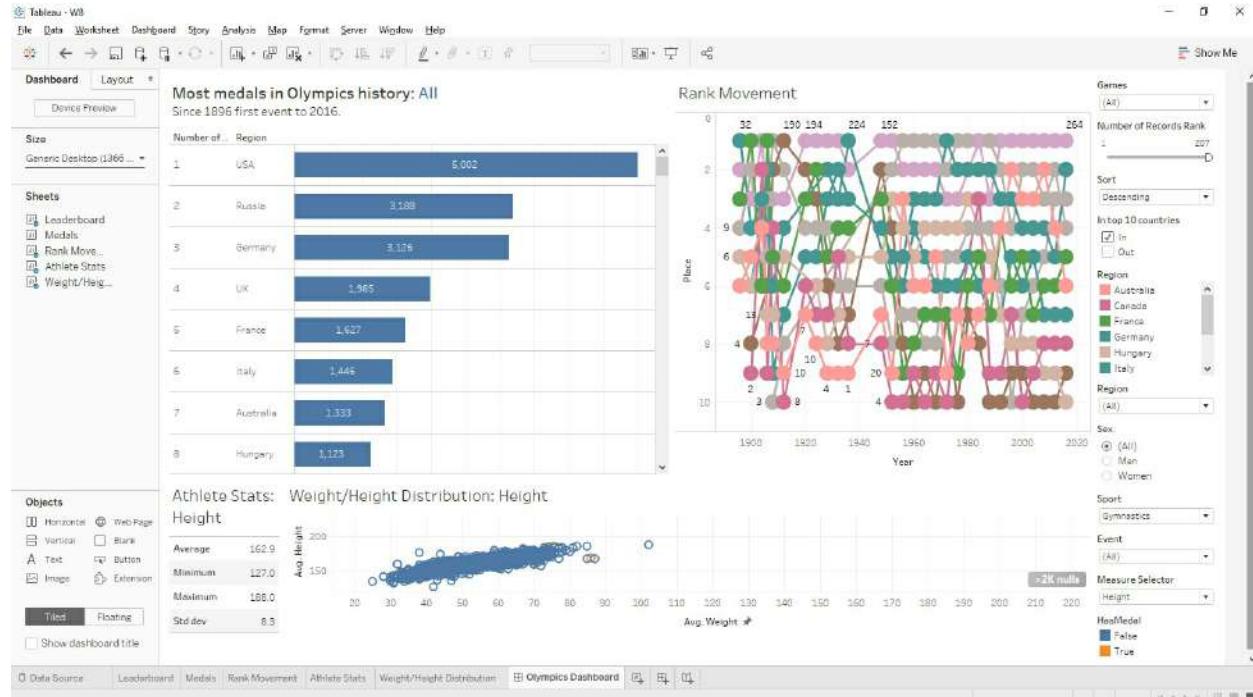
There goes our last graph, time to create the dashboard (this time without the guidance of the item hierarchy):



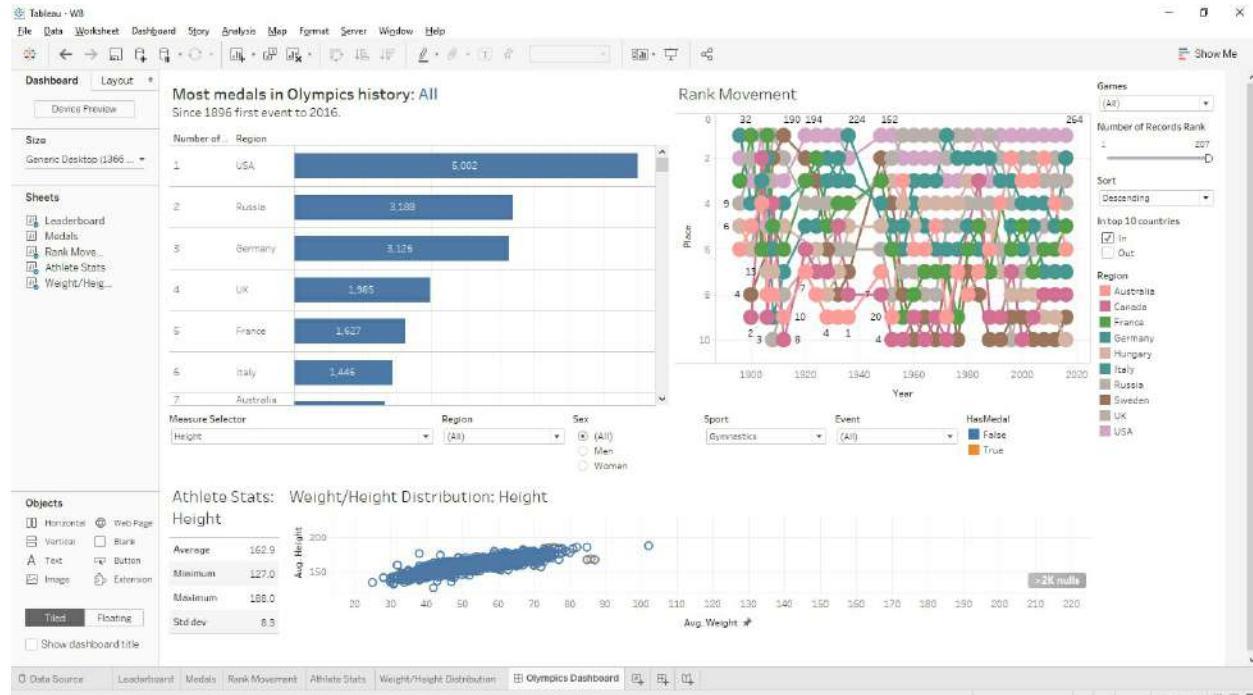
We have too many repeated filters; compare them all per sheet and link the repeated ones by applying one on all the sheets it appears. Example: apply the **Games** filter on *Leaderboard*, *Medals*, *Athlete Stats*, and *Weight/Height Distribution*. **Season** should be applied on all sheets. Then **Sport**, **Event**,

Sex, and **Region** will only be applied to *Athlete Stats* and *Weight/Height Distribution*.

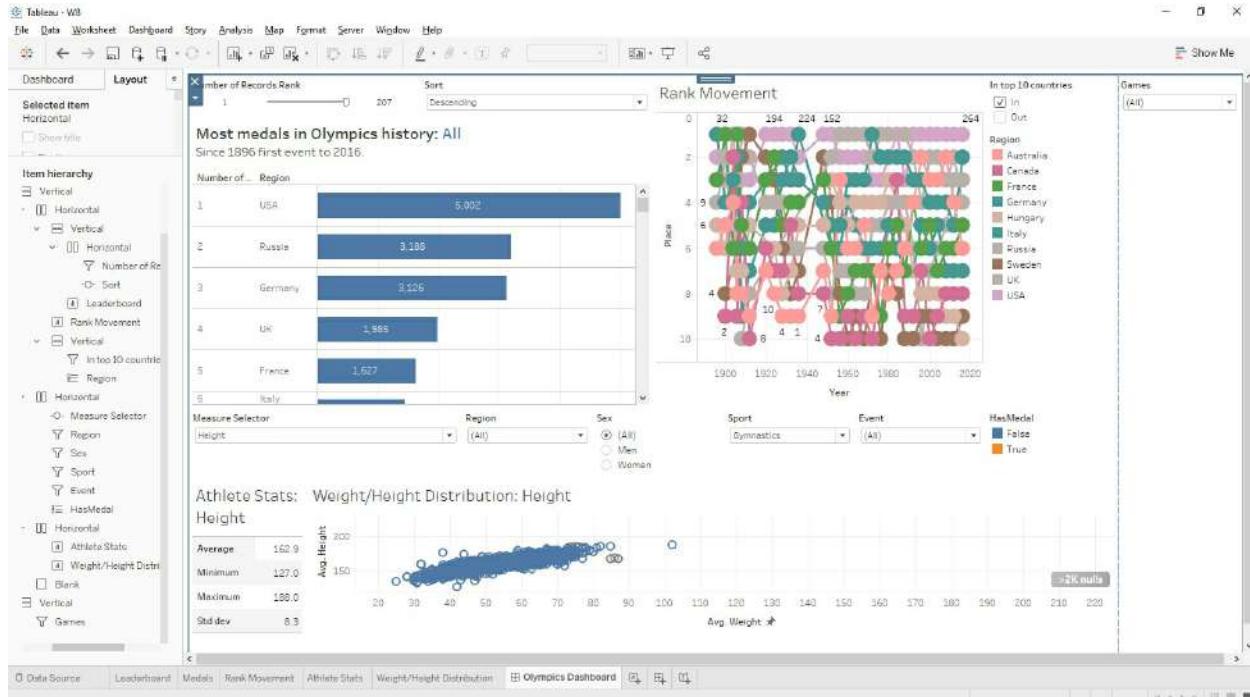
Redo the dashboard above and see it has less filters now.



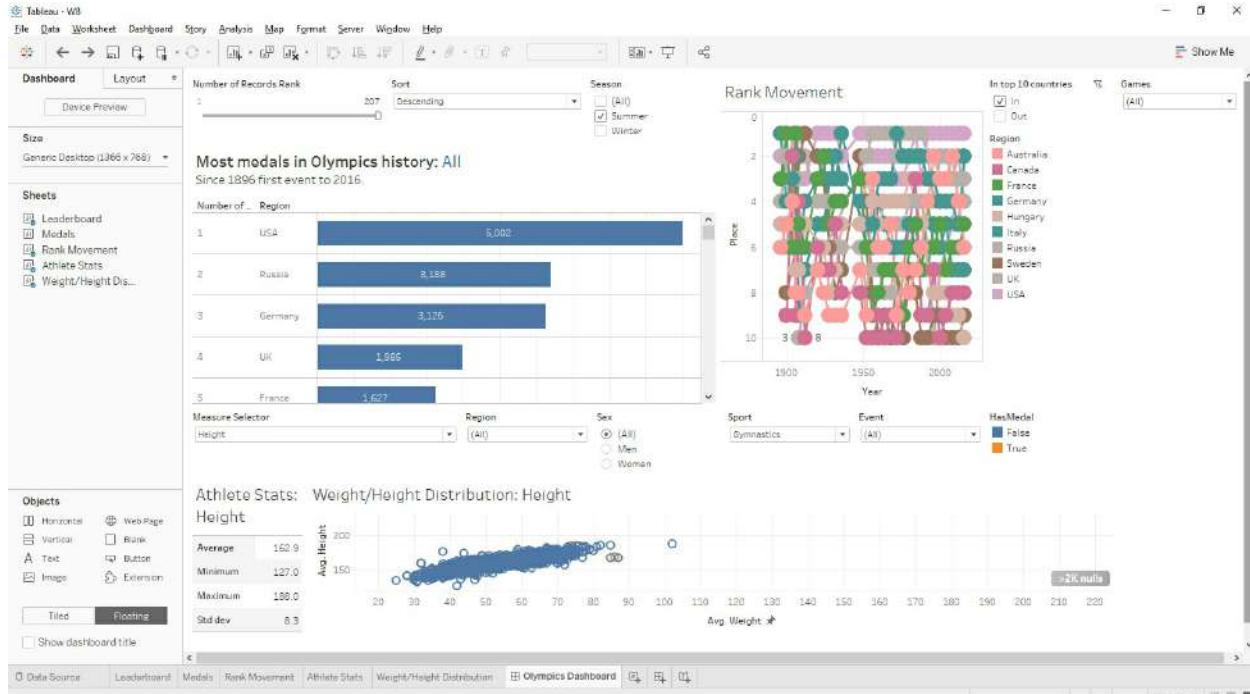
Start organizing the filters for the two sheets on the bottom side. Put them all above the two graphs in a horizontal manner.



Now move the filters belonging to *Leaderboard* above it and those of *Rank Movement* to its right, as in below image:



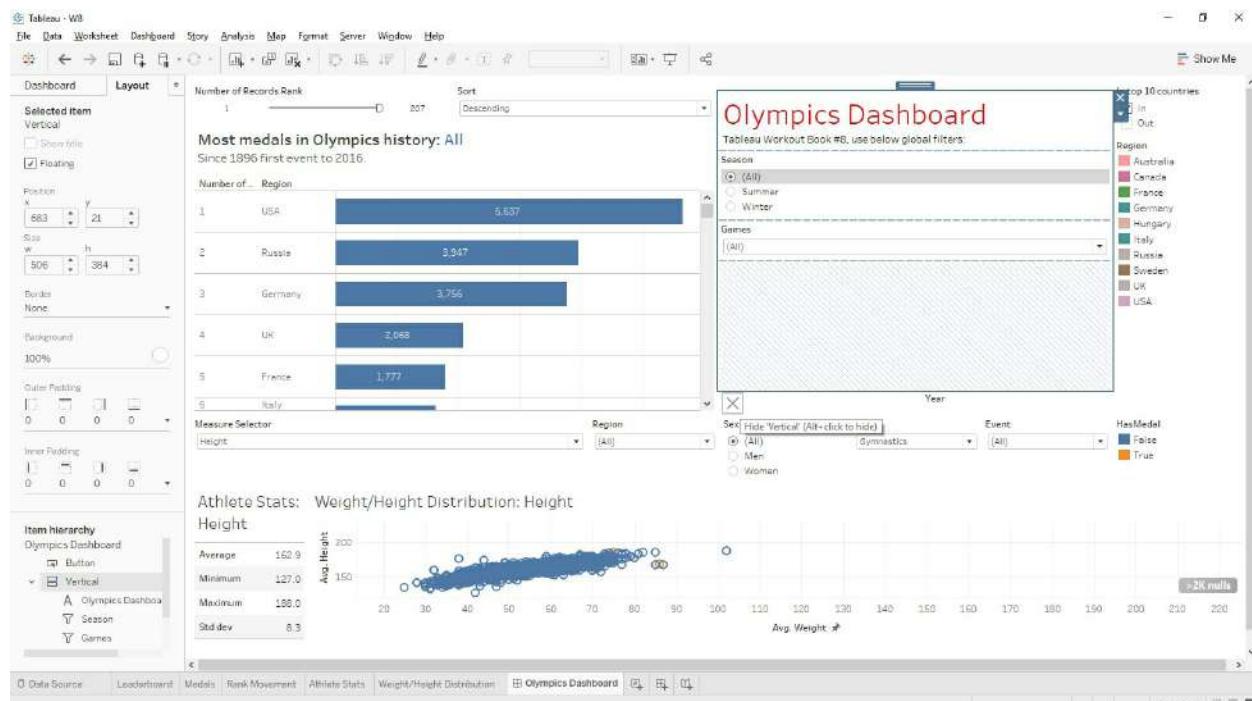
We are only left with the **Games** filter, but remember we also applied **Season** as filter for all sheets but *Medals*. To show Season, click on the arrow of any graph in the dashboard, go to Filters, and toggle it.



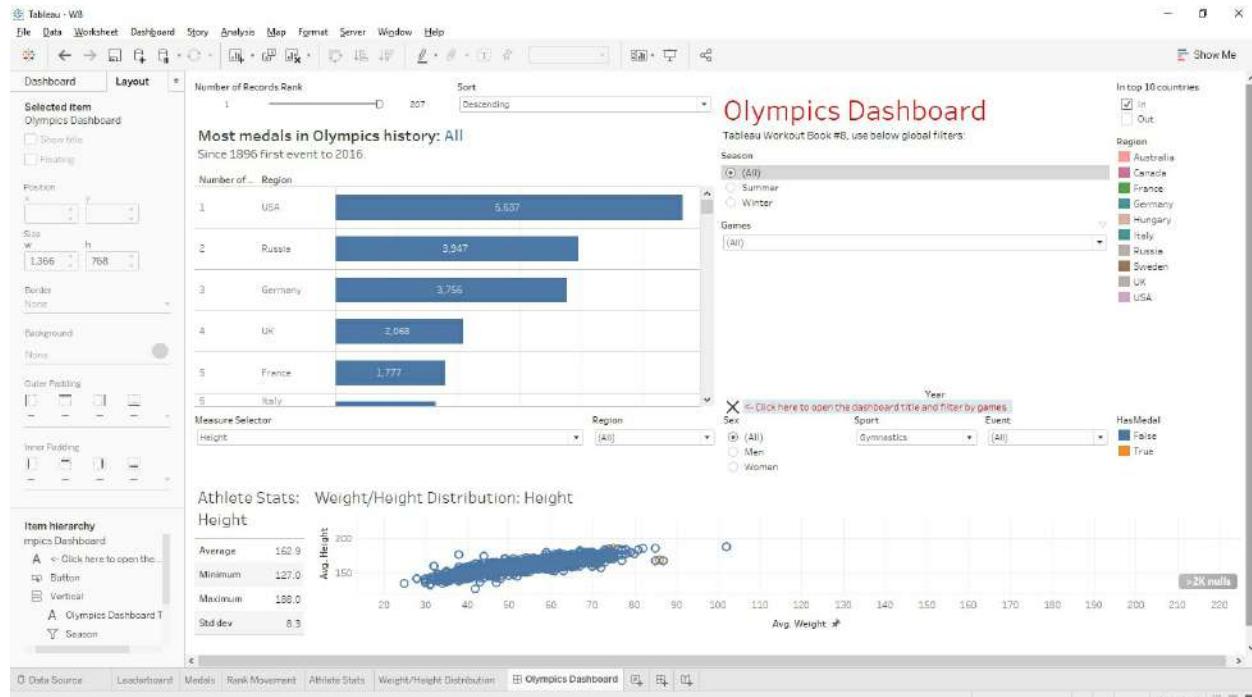
Now if it is not already there, move the **Season** filter to the vertical container that holds **Games**. Add a Text object inside the container as well, that looks like below:

The screenshot shows the Tableau desktop interface with the 'Olympics Dashboard' worksheet active. A floating text box titled 'Olympics Dashboard' is centered over the 'Rank Movement' visualization. The text box contains the text 'Tableau Workout Book #8, usa below global filters'. The dashboard also features a bar chart titled 'Most medals in Olympics history: All' showing the number of medals for top countries, and a scatter plot titled 'Athlete Stats: Weight/Height Distribution: Height' showing the relationship between average height and average weight.

Now select that container, then click the arrow and set it as floating. Click the arrow again and toggle the **Add Show/Hide button**. Add a white background to your container from the Layout pane at the left, so it's readable, adjust its size and put it over *Rank Movement*.

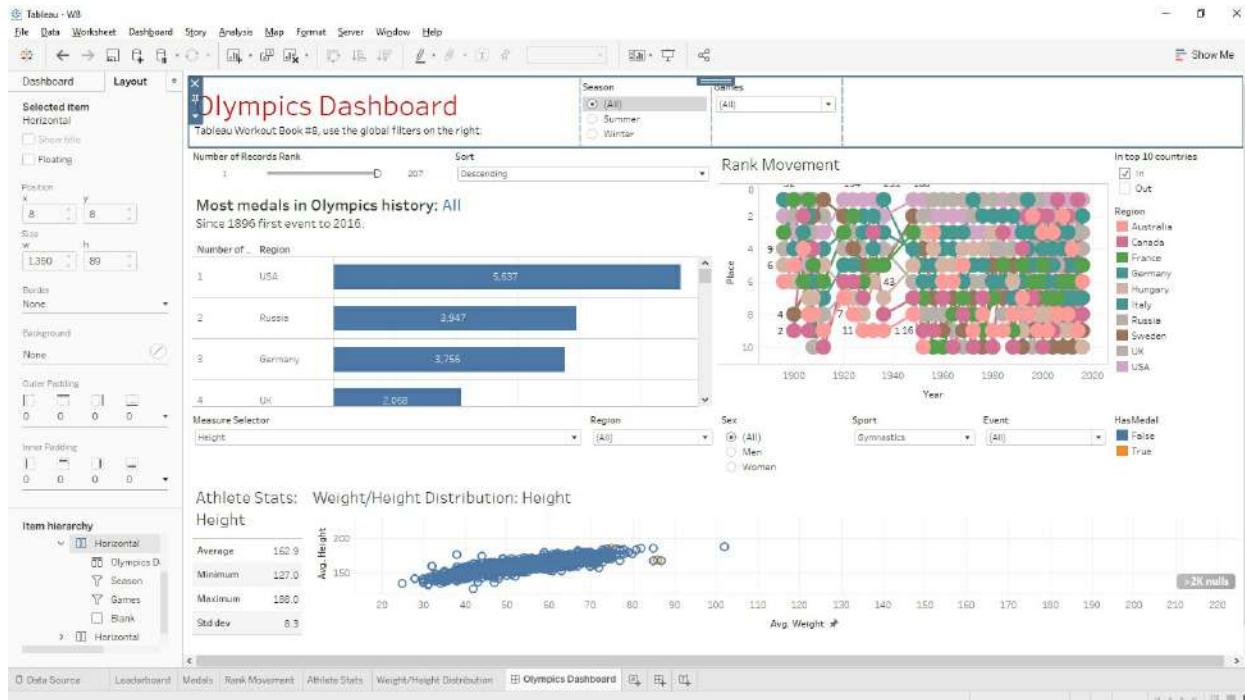


Set the **Games** filter to show only relevant values. If your show/hide button disappeared (the small cross), toggle it again, and put it below the container as shown in the above image. It's a bit unclear what that small cross is – even to us –, but it functions similarly to a hamburger button on an app; it lets you show and collapse a container. Let's add a text instruction to its side so it's clearer for a new user.

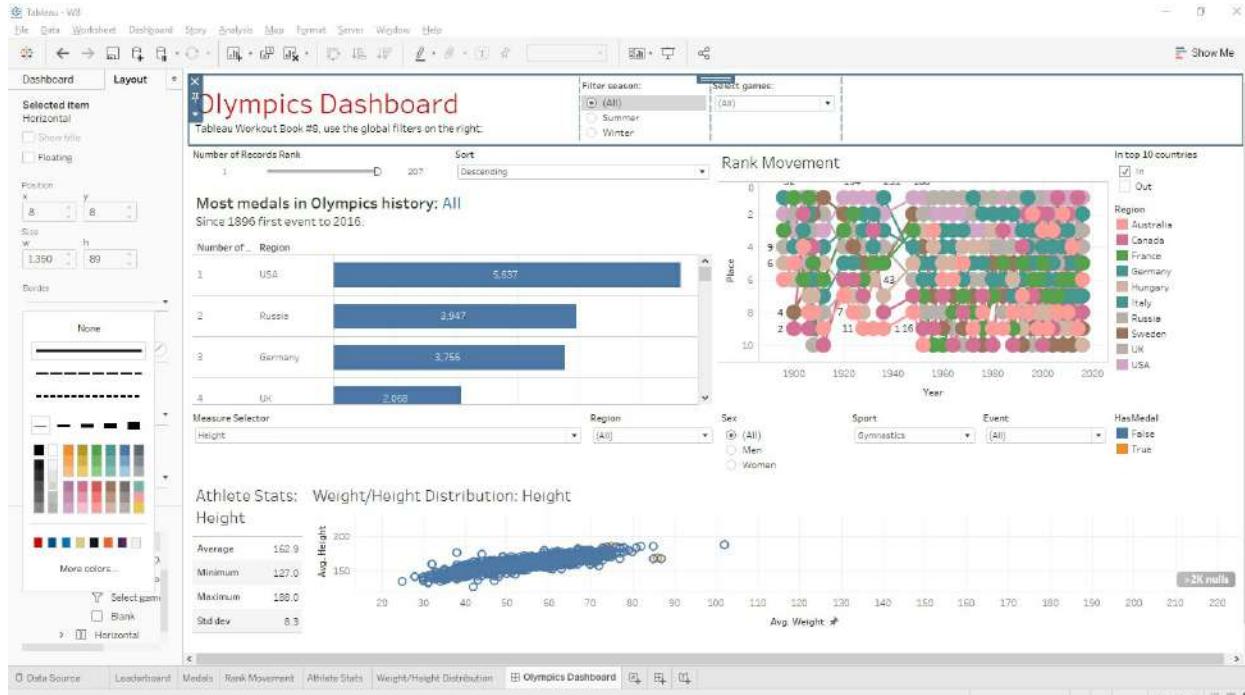


So now you have learned one of Tableau's newest features; how to be able to show and hide a group of filters. The cross is not a very self-explanatory symbol, but fortunately you can customize this clicking on the arrow and Edit Button, so you show an icon of your choice.

However, we will remove this feature from this dashboard for better design. Enable the dashboard title from the Action bar at the top. Copy and paste the title we put on the vertical container, but change the part where it says “below filters” to say “filters on the right”, since we’ll put them on the right of the title. Create a horizontal container, put the dashboard title inside, also move **Season** and **Games** to this container, then add a blank to the right so they’re more centered. Fix the width of this blank, and remove the vertical floating container along with the cross and the helping text.

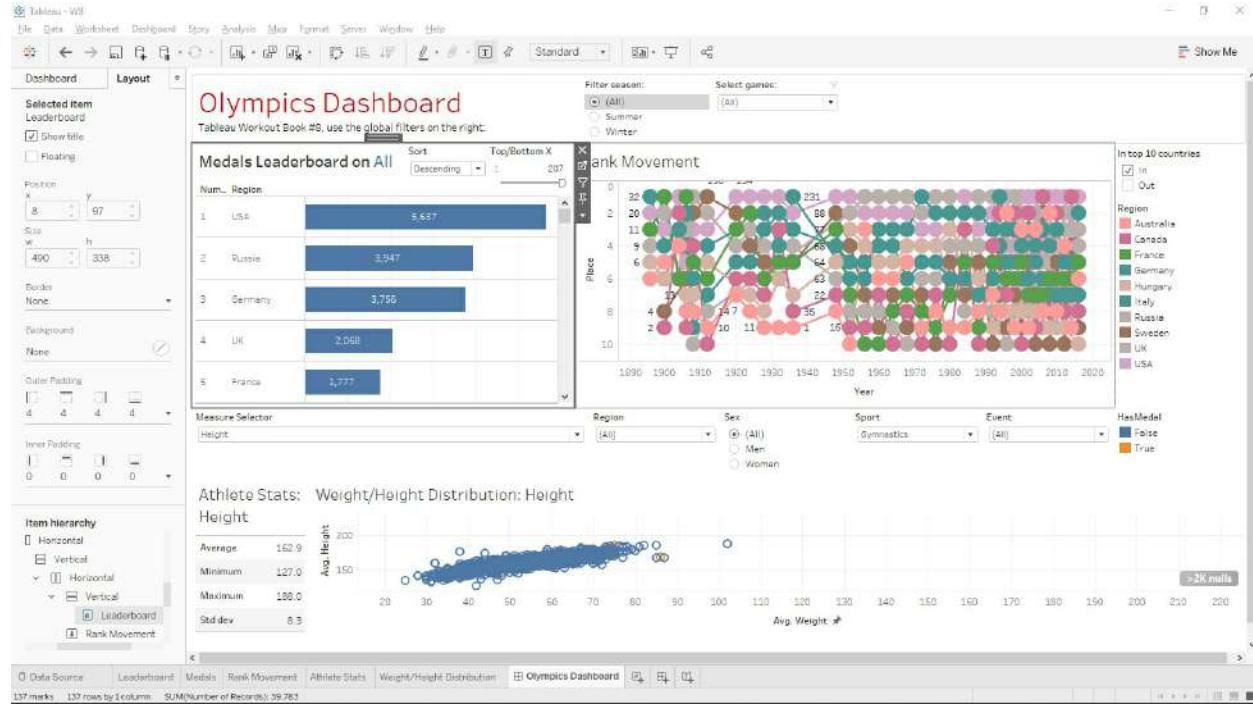


Now add a light gray border to the container:



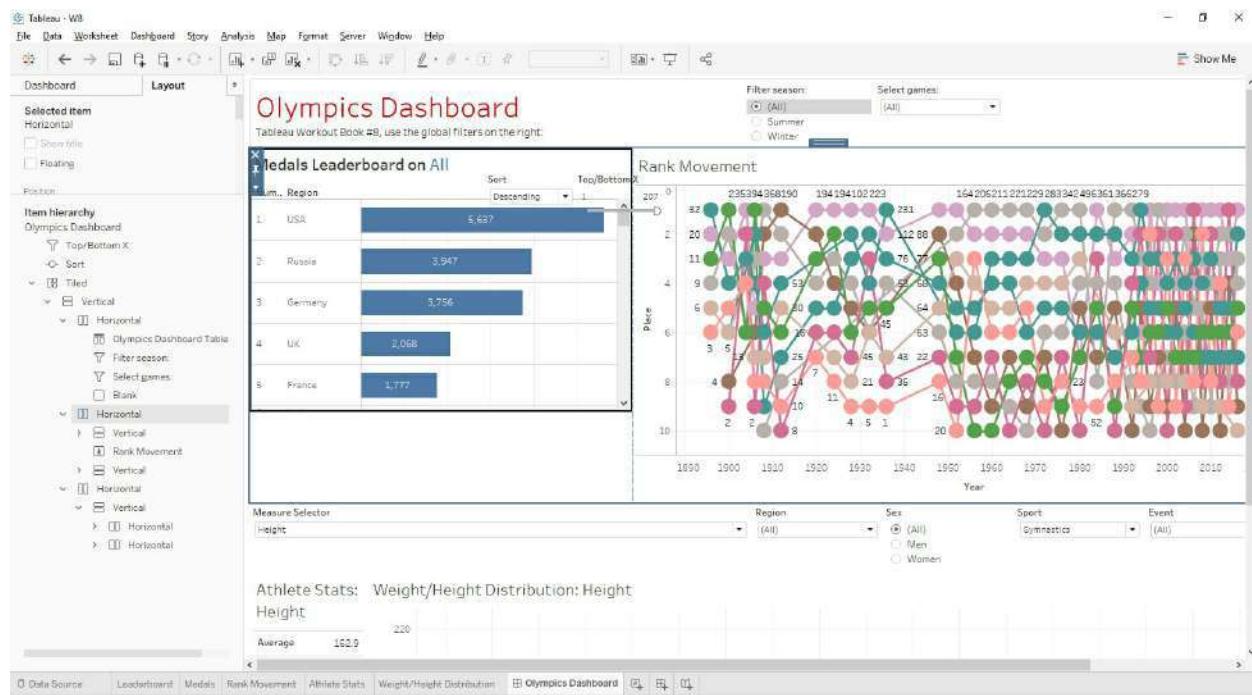
Change the Leaderboard title so it only says “Medals Leaderboard on “, and then the Games reference, removing the second line. Then reduce the width of the rank number column on the table itself so we can make this graph smaller in width.

Change the title of the “Number of Records Rank” filter to “Top/Bottom X”. Set that and the **Sort** parameter as floating and put them to the side of the *Leaderboard* title. You will need to squeeze them in size a bit so they don’t show above the blue bars of the leaderboard, like below:



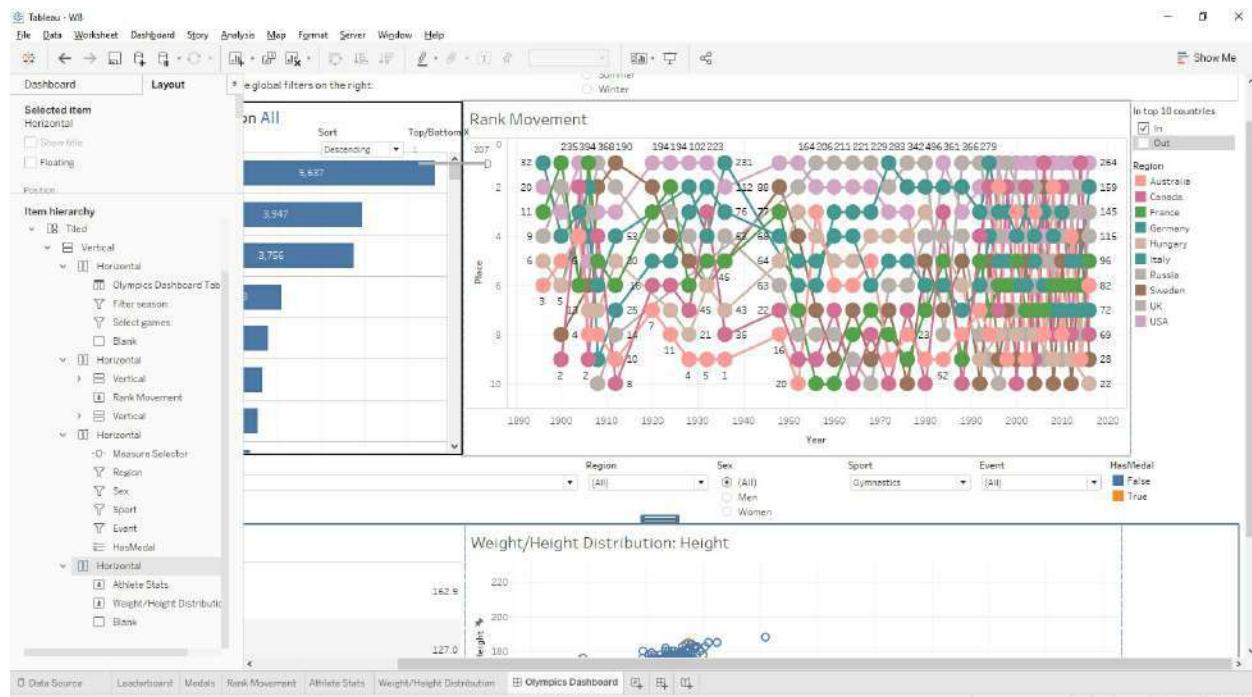
Fix *Leaderboard*’s width so it doesn’t move, then add a medium-small black border to it, and set Outer Padding to 0 on all sides.

Judging by the other sheets, it seems our dashboard is getting too crammed. Change its size to 1,500 x 1,000. Fix the height of Horizontal (1) (the one with our title) and start checking your item hierarchy for duplicate/unnecessary containers on your dashboard. Example, if you have a Horizontal container that only has one child element inside, and that child is a Vertical which you have in a higher level. Every time you are about to move a parent container, fix its height/width and then move it. Finally, remove any blanks we left.

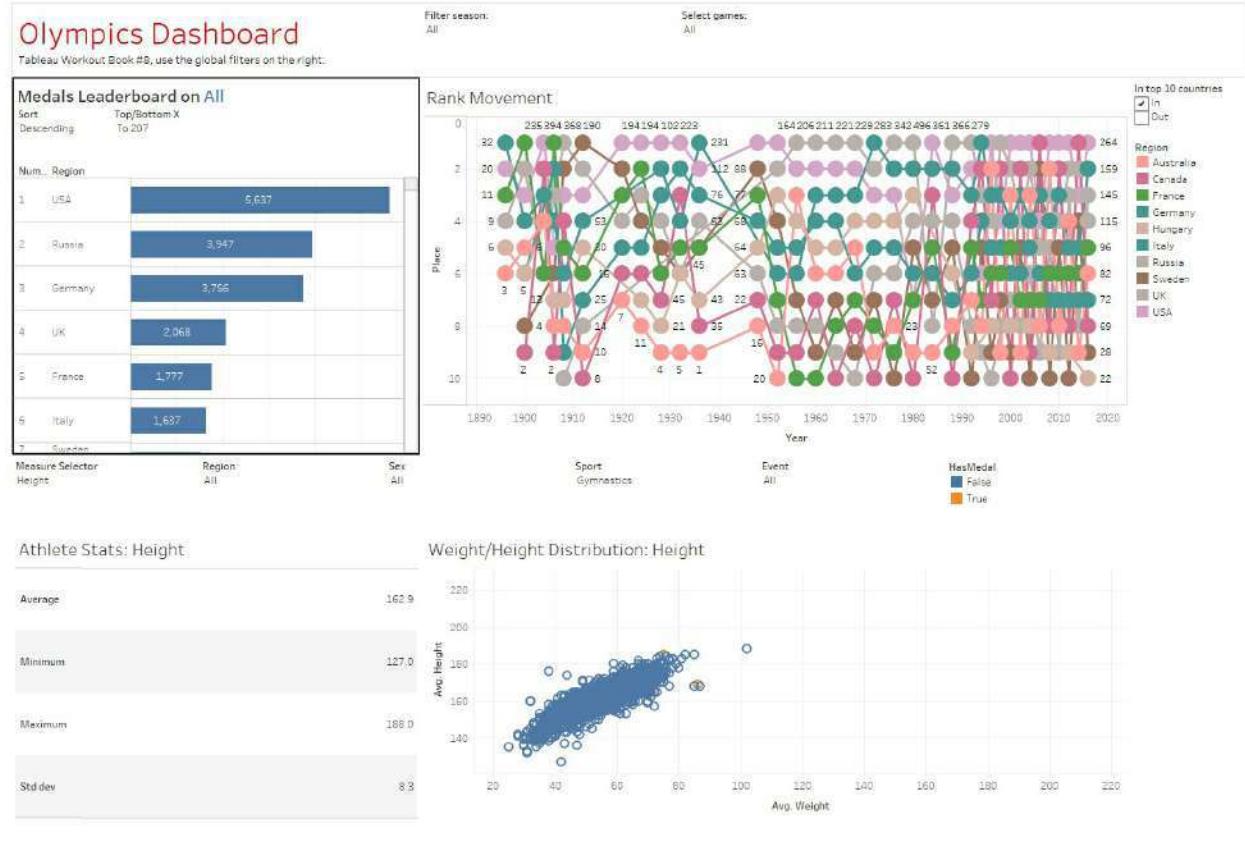


The point of this process is to start fixing the size of the containers so it does not behave crazy when moving items, or readjusting the dashboard size.

Add a blank on the last horizontal container so it's to the right of **Weight/Height Distribution**. Fix it to be the same width as the vertical container above that holds **Region** and **In top 10 countries**. We want our first and second containers with graphs to be aligned in width.



Set *Athlete Stats* to Entire View, distribute items evenly for the two containers with only filters and titles, and adjust the size and position of all elements so it resembles the image below:



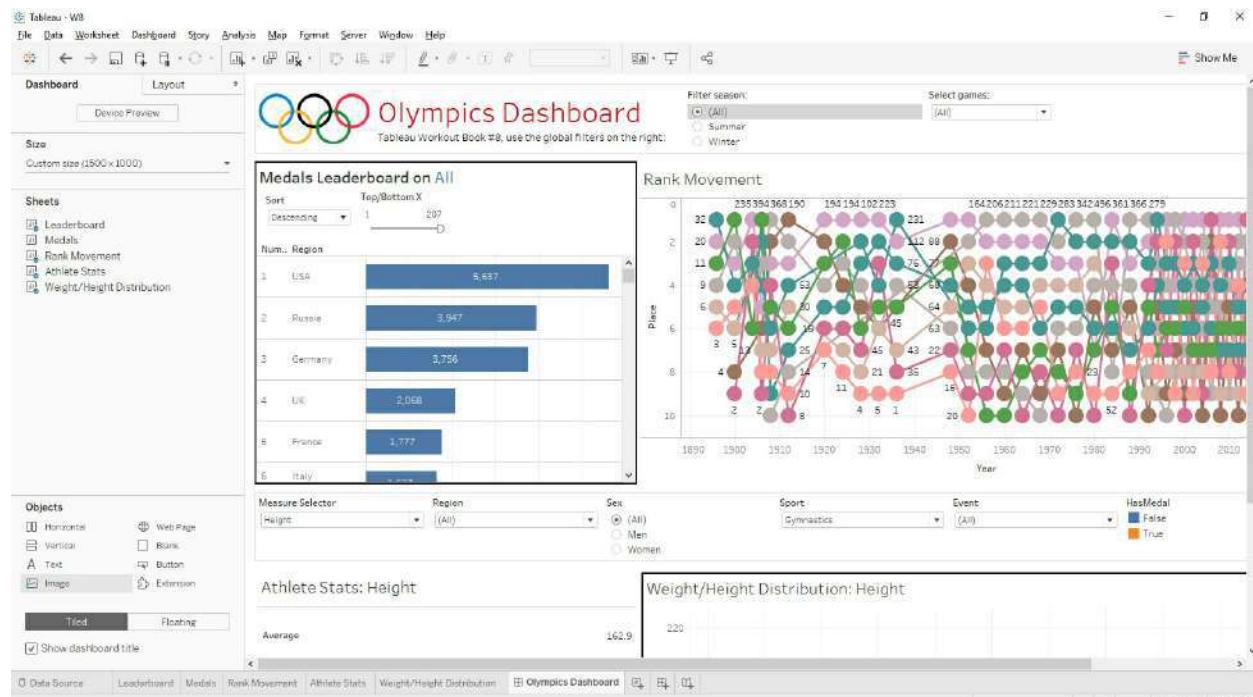
Tip:

You might see that when resizing items you get new Tiled containers in your Item hierarchy. This happens because you are aligning items that belong to different containers. What Tableau does for this is separate those items to exclusive containers that are set up with the same width/height. This might not be liked by some developers, so if you don't want these to show, avoid resizing an element to one of those small arrows that appear when you resize.

Now on to the formatting. For all four horizontal/tiled containers, add an Outer Padding of 5 to all sides. Readjust the two filters for *Leaderboard*. Now apply the same gray border for the third horizontal container and apply the same border from *Leaderboard* to *Weight/Height Distribution*.



The sizing and distribution of elements is now finished, time to finish with the last formatting details. Download an Olympics logo and put it on the first Horizontal by dragging an **Image** element.



Edit the dashboard title color to this red (#ff0000) and the second text line to black. Format the filters to have a black title and make the options black (by default it's a dark gray, slight difference but it's for consistency).

Go to *Leaderboard*, change the bar color to this blue (#3e76ec), and remove all borders and lines. Again we need to format the font for all worksheet's text elements to be full black, not dark gray, but instead of going sheet by sheet, format this from the Format Workbook > All option.

Tip:

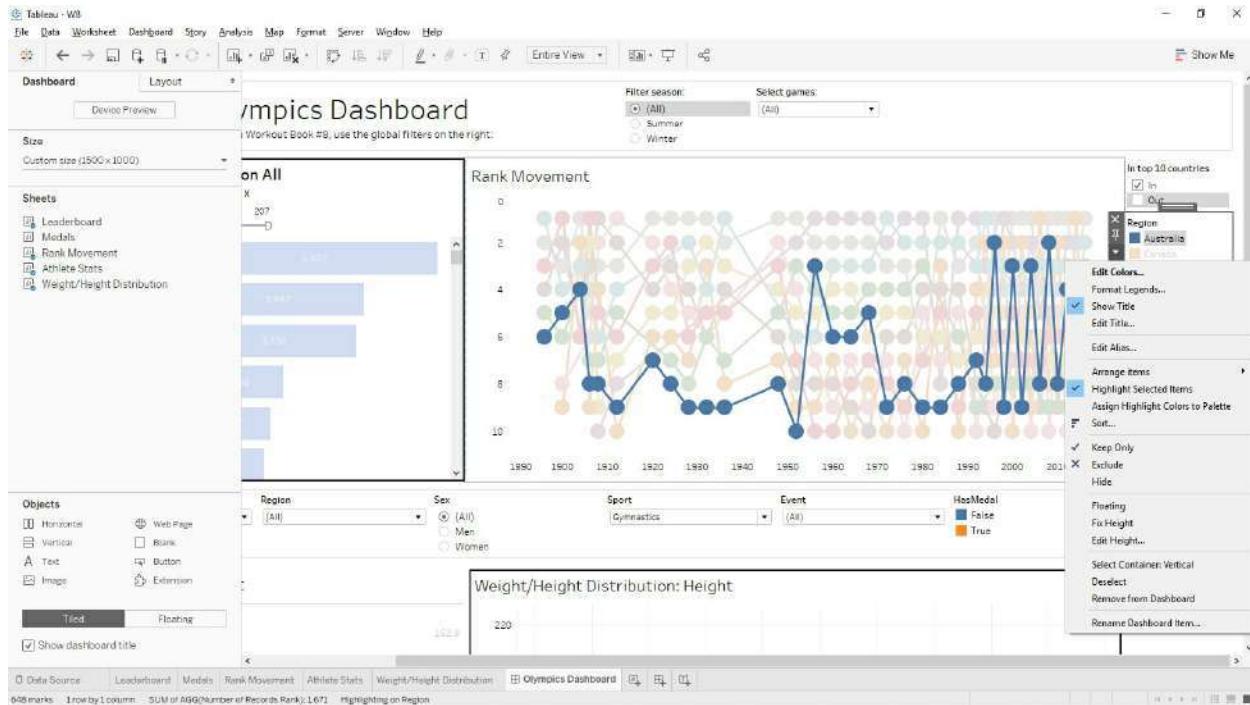
As said before, when formatting it is best to start from the top to the bottom level. Meaning we should have started with the Worksheet formatting, then after that we should have gone to the sheet by sheet format. You will see that after applying the last adjustment, the colors we applied on our labels before, like the dashboard title and Leaderboard's title, will go to black. But for this exercise, we will keep that change. Just bear this in mind on your future projects, since this might cause you rework.

Then set the color for the bars' labels back to white and hide field labels for rows so it looks like below:

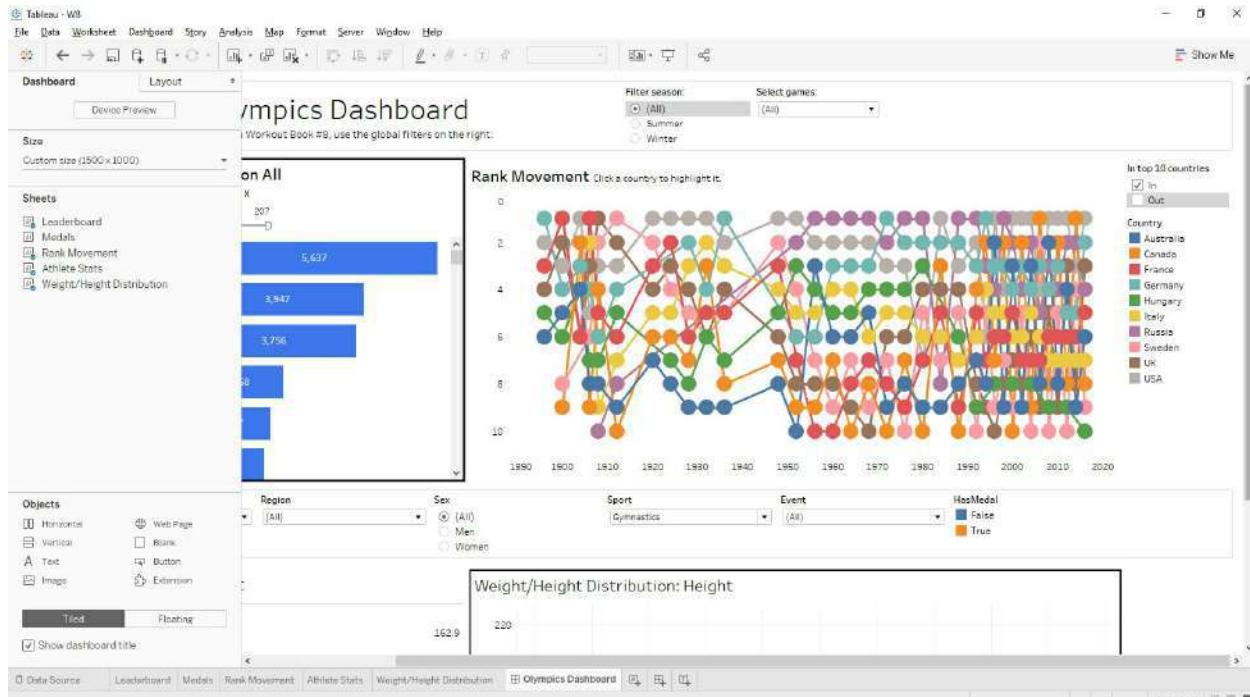


Now on *Rank Movement*, make the title bold, remove the axes titles ("Year")

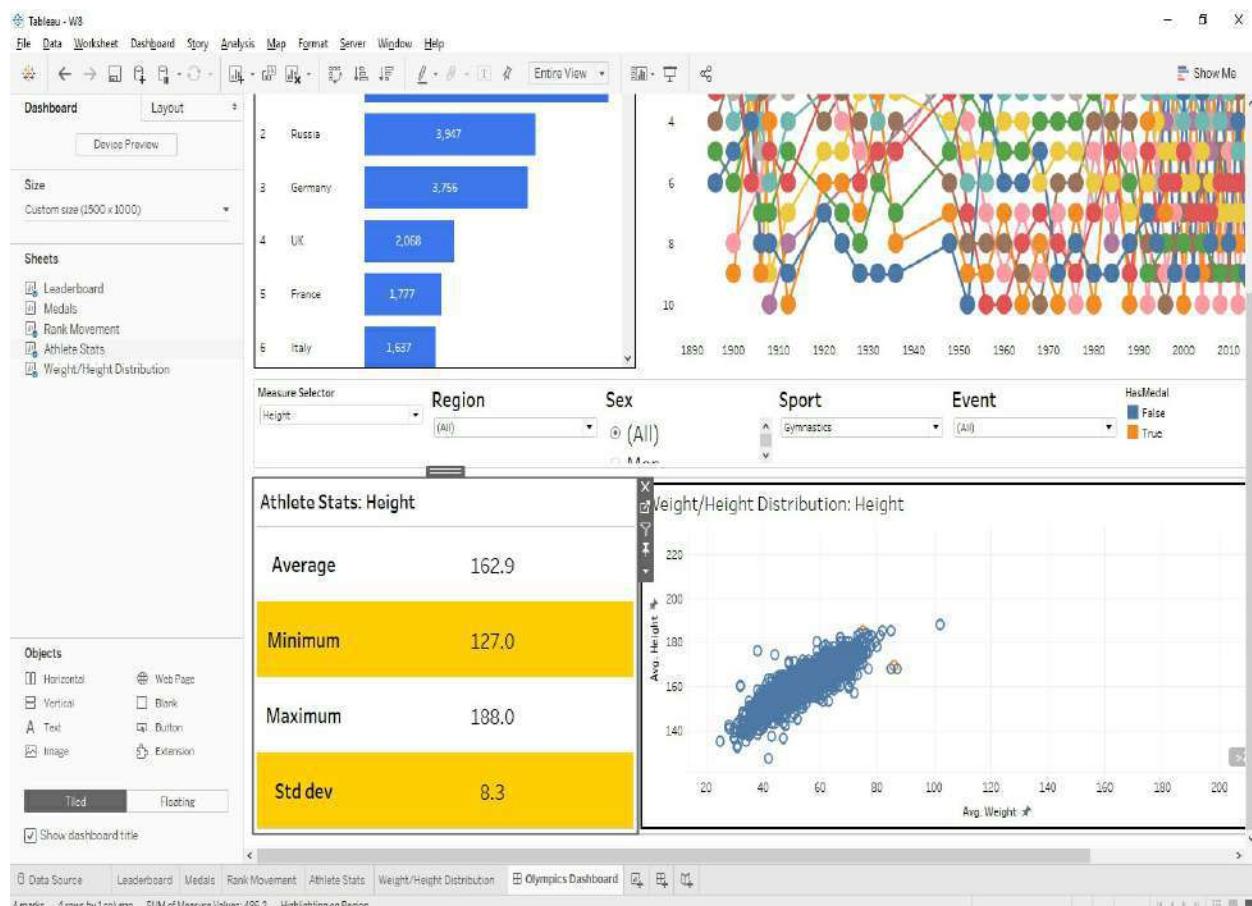
and “Place”), and remove all borders and lines. Reset the colors assignment for the top 10 Regions and remove the mark labels since they’re not visible (the numbers on the circles). Back on the dashboard, click on the **Region** legend so it shows its arrow, click it and toggle **Highlight Selected Items** so that when a user clicks a country it gets highlighted in the rank graph like below:



Now change the title of the color legend from “Region” to “Country” and to the *Rank Movement* sheet title add “Click a country to highlight it.” as a small help.



Now on to *Athlete Stats*. Make the title bold, format the font size of the whole worksheet to be 16 (not the workbook, only this sheet), and format the alignment to center for pane and header. Finally, change the row banding to be of this yellow (#ffce01).

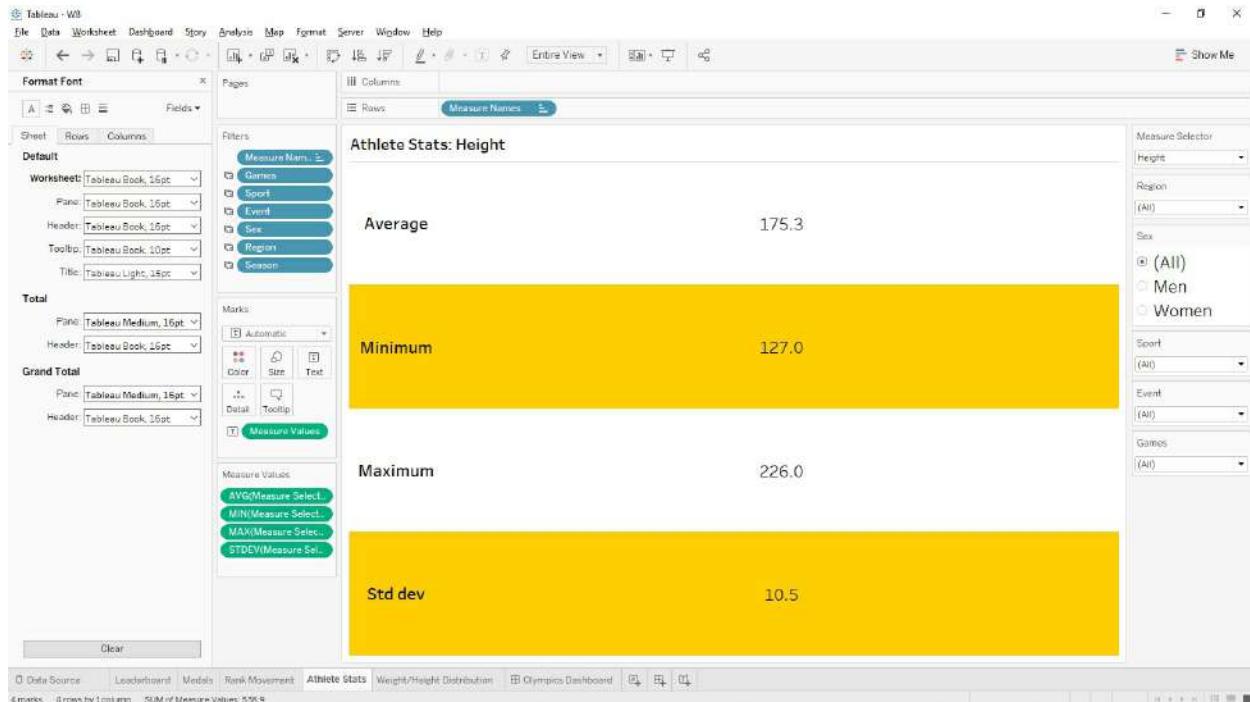


Now on the last sheet; make the title bold, remove the **Measure Selector** portion, and create a new color palette called “Olympics discrete” like we did on workout 3 using the following colors: blue (#3e76ec), black (#000000), red (#ff0000), yellow (#ffce01), and green (#179a13). You don’t need to delete any other custom palette (if you don’t remember [go back to it](#)). Remember to restart Tableau after you modify the preferences file.

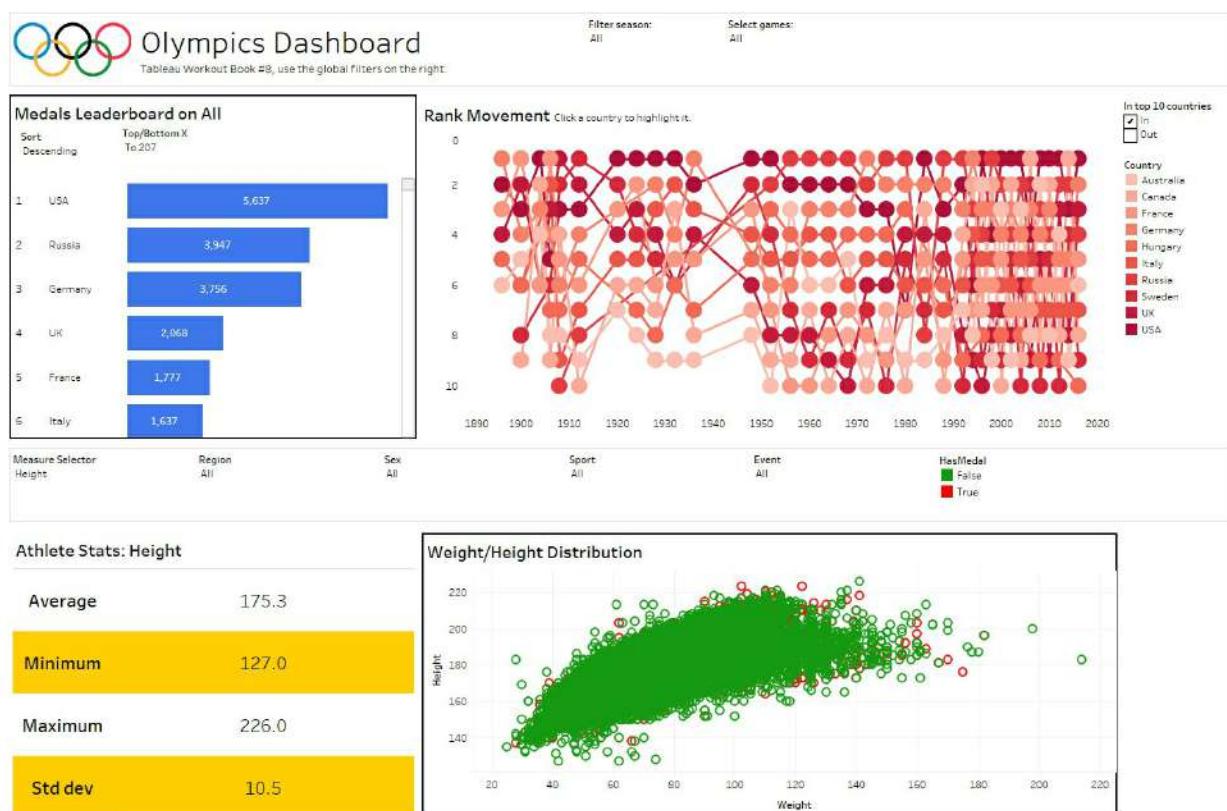
Using that palette, assign the no medals athletes the green color and red for the ones with medals. Finally, rename the axes so it doesn’t say “Avg.” for height or weight.



Seems like when we applied the Worksheet font size format on *Athlete Stats* we also affected the filters, if you look at some of them on the dashboard they look bigger. Go back to it, revert that format, and put it instead for the Pane and Header options. An easy way to revert format is pressing the Clear button at the bottom of the pane, it will revert to default all settings from the current pane you're at.



Go back to *Rank Movement* and change the color palette to be Red Sequential.



If you notice, what we did was assign a color from the Olympics logo to each

sheet, leaving the black for all the elements. This is a simpler design but it looks very clean. **Congratulations, you have finished your eight workout!**

The new topics reviewed are:

- Functionality to edit sorting by ascending/descending through a parameter
- Top X items
- Rank graph
- Functionality to show/hide filters on a dashboard

You have gone through all the topics there are for a Tableau developer. You have mastered from the basic concepts of creating graphs to the more specific functionalities that users ask for. You have learned how to use basic design concepts such as color palettes, formatting, sizing, and positioning. You have learned the technical calculation concepts from the basic IIF formulas to the level-of-detail calculations, which are unknown to many! Finally, you have learned the best practices and tips to work efficiently and fast.

It's now time to prove your capacity, the next workout will not include any instructions for you. What you'll receive is just the data source and the expected result. Are you ready? Turn the page for the final workout.

Workout 9. YouTube trending dashboard

On this workout you will make an analysis on the videos that appear on YouTube trending, allowing users to try to find a pattern on video's statistics and trending appearance, and to see which channels and categories appear the most.



Your data sources are:

W9 US Videos.csv

Its granularity is trending date – video. It is the main data source where you will find details on the videos and the dates they were trending.

Column	Description
video_id	Unique identifier of video.
trending_date	The date the video was displayed on the trending section.
title	The video title.
channel_title	The owner of the video.
category_id	The category ID of the video.
publish_time	The time the video was published.
tags	The tags the video was tagged.
views	The views the video had when it went trending.
likes	The likes the video had the moment it went trending.
dislikes	The dislikes the video had the moment it went trending.
comment_count	The comments the video had the moment it went trending.

thumbnail_link	The video thumbnail image.
comments_disabled	If comments were disabled for the video.
ratings_disabled	If ratings were disabled for the video.
video_error_or_removed	If the video was banned/removed.
description	The video description.

W9 US Category ID.json

Its granularity is category. Contains the mapping table for the Category ID on the main table. The only columns you need are:

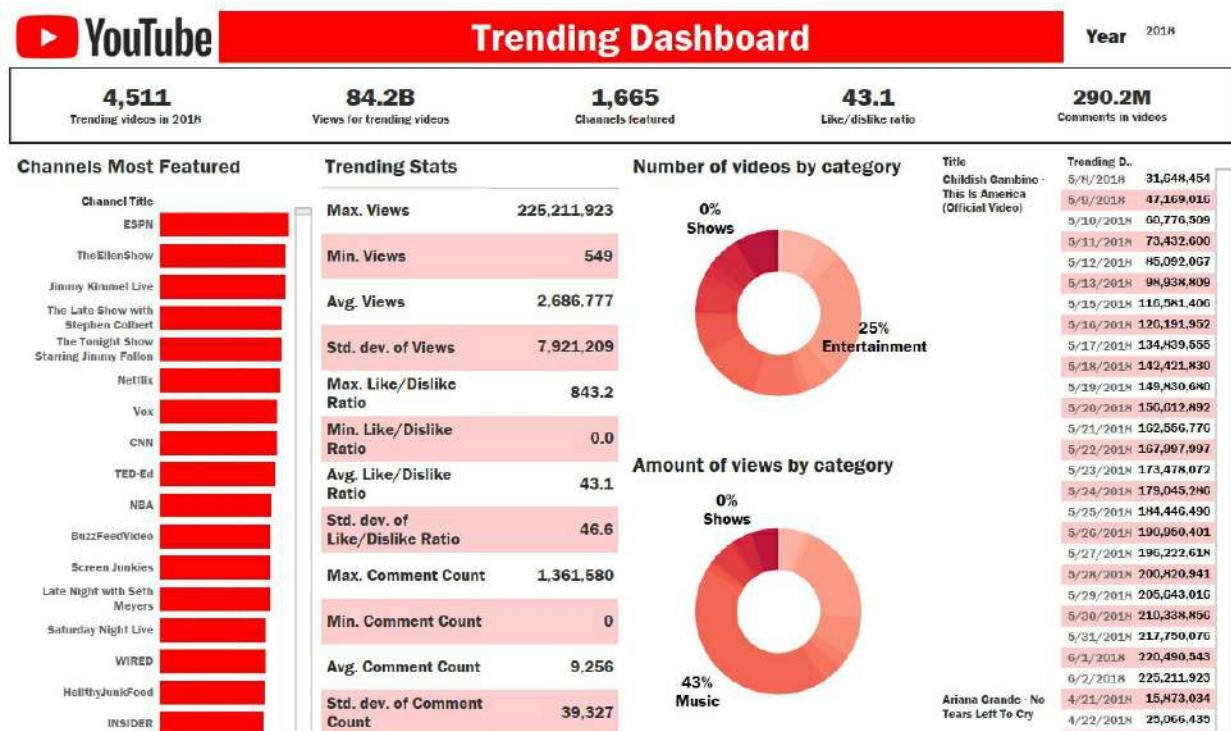
Column Description

Id	The category ID.
Title	The category name.

You can hide the rest.

Hint: add both tables as separate data sources and join them through data blending with **Category ID**.

You should get to this dashboard:



This is a dashboard made up of 10 sheets; the first five are just a number display of the main metrics with no tooltips; COUNTD>Title, SUM(Views), COUNTD(Channel name), SUM(Likes) / Sum(Dislikes), and SUM(Comments).

Regarding format:

Workbook Font Style: Franklin Gothic Demi

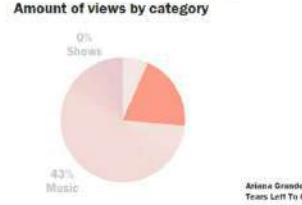
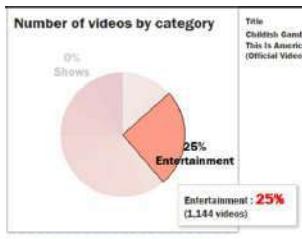
Colors: (red) #FF0000 and (black) #282828 as mains, also (light red) #FFCDCC for table shading, used the red color palette for the donut graphs

The next five sheets display:

- 1) **Channels most featured** display the channel names sorted by who has the most appearances in trending. Shows the number on the tooltip as below:



- 2) **Trending stats** shows the views, like/dislike ratio, and comment stats, with aggregation by maximum, minimum, average, and standard deviation. It has no tooltip.
- 3) **Number of videos by category** pie chart that shows by category how many videos have appeared in trending, calculated as a percentage. It has the tooltip shown below:



- 4) **Amount of views by category** pie chart like the previous one, only that this shows instead the sum of views for trending videos, calculated as a percentage as well. It has the same tooltip as the previous pie chart, but with views instead of videos.

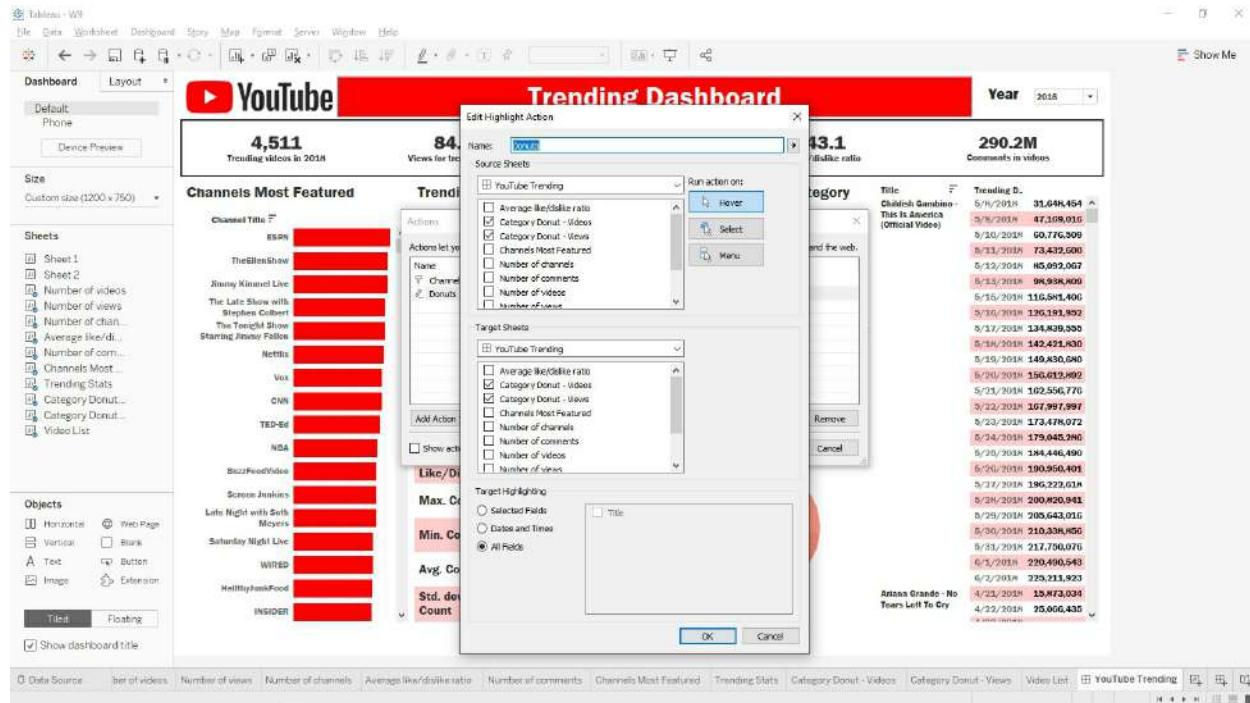
- 5) **Video List** displays the list of videos by title and trending date, showing the sum of views. Includes more details in the tooltip, like below:

The dashboard also has two actions:

- 1) When selecting a channel from the Channels Most Featured leaderboard, the rest of the dashboard elements get filtered by the clicked channel. When unclicked, the filter reverts.

- 2) When hovering over a category on either of the two pie charts, the category is highlighted on the other chart as well, so it's more visible

the proportion of a category in number of videos and amount of views.



You can compare your final dashboard with the one stored on the “other” folder from the downloaded materials. There you will also find the finished versions of all the workouts we did.

Congratulations, that was the final workout for this book.

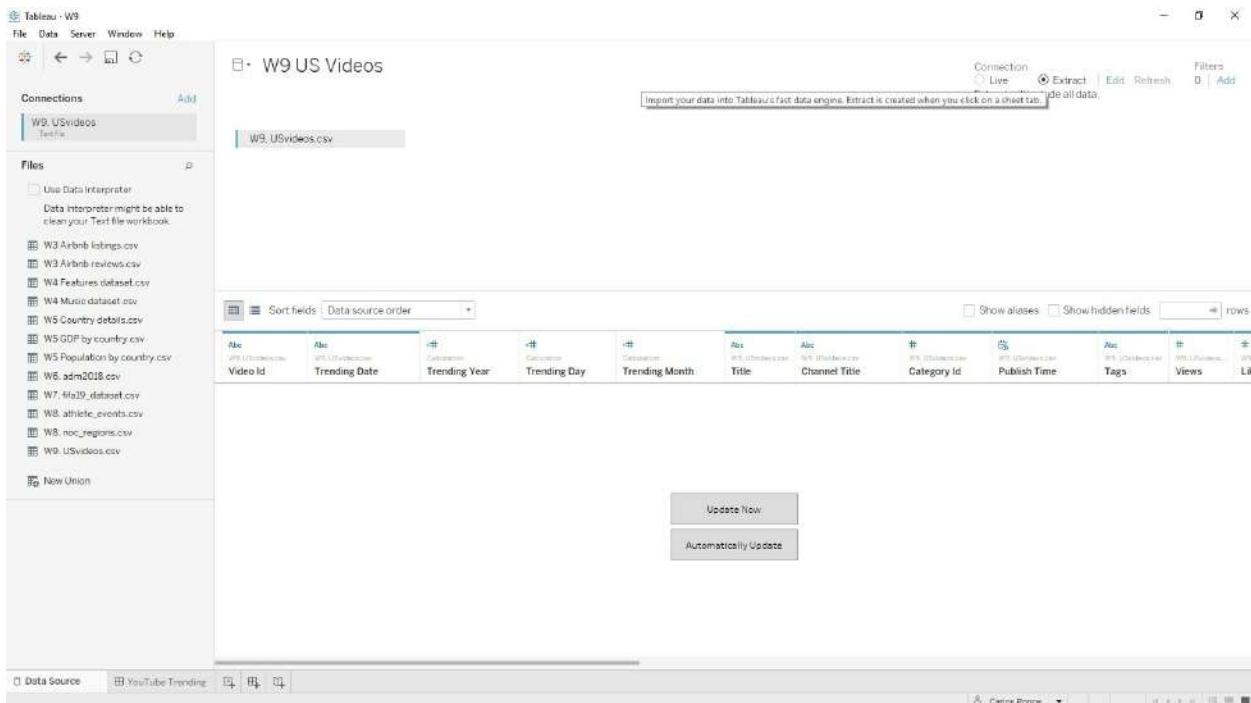
End

You made it to the end, congratulations! **If you purchase a Tableau license**, you can publish your dashboards on **Tableau Public** and share the links through LinkedIn, so people can see your level of expertise – **including recruiters**.

If you are not planning to have a license soon, you can instead export each of your workbooks as a **Packaged Workbook**, by going to the File option on the action bar. This will store your workbook and data sources together in one file which can be opened using [Tableau Reader](#), which is free.

To publish your workbooks on **Tableau Public** and share with your contacts:

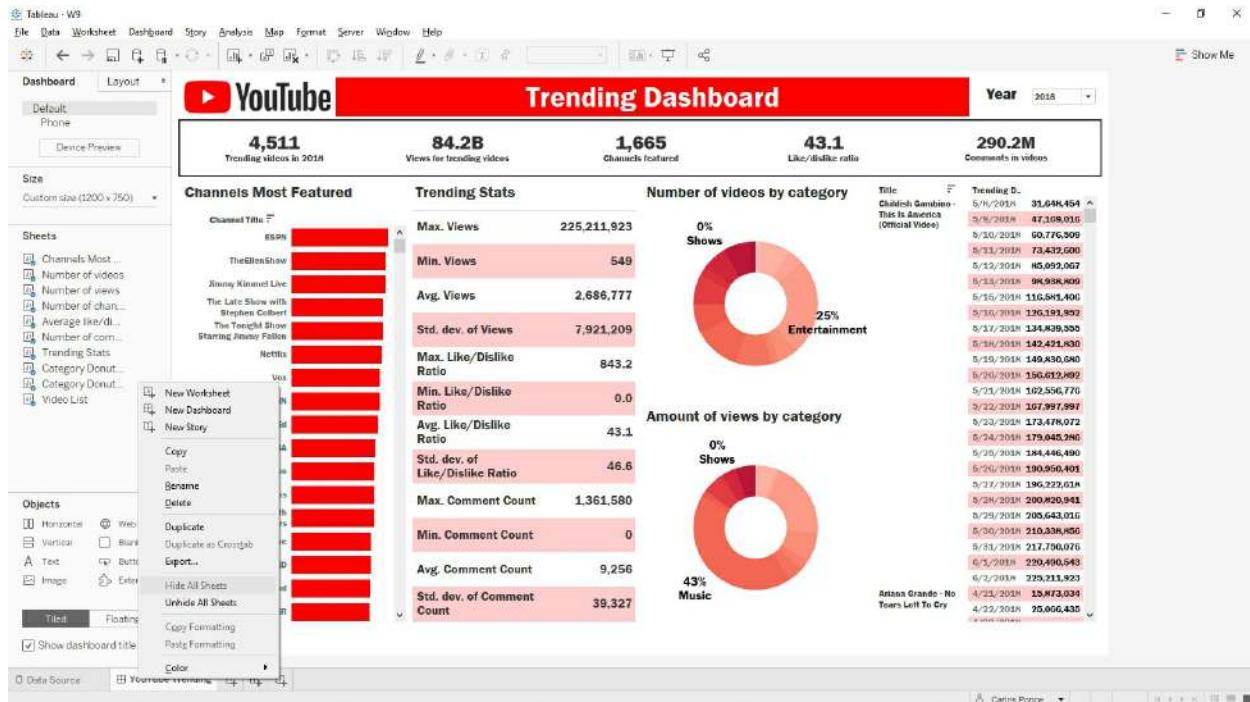
1. Create an account on public.tableau.com
2. Convert all the data sources on your workbooks to data extracts by selecting the option from the Data Source tab. You can save the hyper file anywhere.



The screenshot shows the Tableau interface with the following details:

- File Bar:** File, Data, Server, Window, Help.
- Connections:** W9_USvideos (selected).
- Files:** Use Data Interpreter (disabled), W9_Airbnb_listings.csv, W9_Airbnb_reviews.csv, W9_Features_dataset.csv, W9_Music_dataset.csv, NS_Country_details.csv, WS_GDP_by_country.csv, WS_Population_by_country.csv, W6_admin2018.csv, W7_ifta9_dofmort.csv, W8_athlete_events.csv, W8_noc_regions.csv, W9_USVideos.csv, New Union.
- Dashboard Tab:** YouTube Trending.
- Data Source Tab:** Preview of W9_USVideos.csv with columns: Video Id, Trending Date, Trending Year, Trending Day, Trending Month, Title, Channel Title, Category Id, Publish Time, Tags, Views, Lik.
- Buttons at Bottom:** Update Now, Automatically Update.

3. Hide all the sheets on your workbook so you only see the dashboard tab. You do this by right clicking a dashboard and selecting Hide All Sheets.



4. Go to Server > Tableau Public > Save to Tableau Public As..., log in, and give your dashboard a name.

If you liked this, please leave a review on Amazon so other people can know (I hope you did!). Finally, if you want more news and free Tableau tips follow my [Facebook page](#).