Tableau Public Tutorial

Brief Introduction to Tableau Public

Tableau is a company of interactive data visualization for bloggers, journalists, researchers, advocates, proferssors and students to make their storytelling expressive and insightful. It offers five main products: Tableau Desktop, Tableau Server, Tableau Online, Tableau Reader and Tableau Public.

Tableau Public is a free service that lets anyone publish interactive data to the web. Once on the web, anyone can interact with the data, download it, or create their own visualizations of it. No programming skills are required. Be sure to look at the gallery to see some of the things people have been doing with it.

Tableau Public includes a free desktop product that you can download and use to publish interactive data visualizations to the web. The Tableau Public desktop saves work to the Tableau Public web servers – nothing is saved locally on your computer. All data saved to Tableau Public will be accessible by everyone on the internet, so be sure to work only with publically available (and appropriate) data.

It works with all Windows platforms and virtual environments for Windows. In this lab session, we are going to use its free release, Tableau Public, version 8.2, which essentially contains a set of interactive charts, advanced functions for data processing and JavaScript API.

Downloading and Installing Tableau Public and Dataset

Tableau Public:

1. You can download Tableau Public 8.2 from the following link:
   http://www.tableausoftware.com/public/
2. After installation, a shortcut can be found in the Start Menu.

Dataset:

Please download the data set "OlympicAthletes_0.xlsx" below:

https://public.tableau.com/s/sites/default/files/**OlympicAthletes_0**.xlsx

or

https://github.com/chrismilleruk/couch-stream/blob/master/OlympicAthletes_0.xlsx

Getting Started with Tableau Public

In general, it consists of three steps to visualize a dataset with Tableau Public.

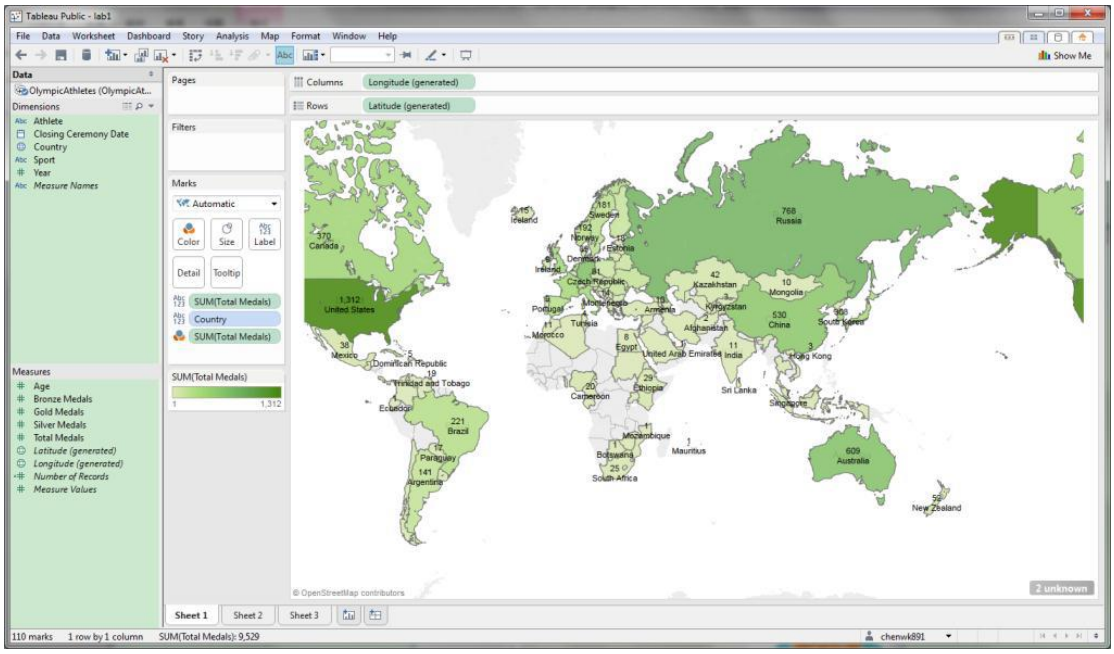
1. **Acquire and process the data.** Usually we can obtain the data from the web, such as government website, open data repositories, web services, etc. It could be accessed from the Excel document, database, or even web API. Let us take excel document for example. We usually need to preprocess the file to make it compatible with Tableau. For instance, an Excel file ready for import looks like below.

	A	B	C	D	E	F	G	H	I	J
1	Athlete	Age	Country	Year	Closing Ceremony Date	Sport	Gold Medals	Silver Medals	Bronze Medals	Total Medals
2	Michael Phelps	23	United States	2008	8/24/2008	Swimming	8	0	0	8
3	Michael Phelps	19	United States	2004	8/29/2004	Swimming	6	0	2	8
4	Michael Phelps	27	United States	2012	8/12/2012	Swimming	4	2	0	6
5	Natalie Coughlin	25	United States	2008	8/24/2008	Swimming	1	2	3	6
6	Aleksey Nemov	24	Russia	2000	10/1/2000	Gymnastics	2	1	3	6
7	Alicia Coutts	24	Australia	2012	8/12/2012	Swimming	1	3	1	5
8	Missy Franklin	17	United States	2012	8/12/2012	Swimming	4	0	1	5

   Each column represents different filed of the variable, and the first row consists of the field names of the dimensions and the data start from the second row. Once it is properly formatted as above, we can use Connect to Data in Tableau to import it.

2. **Generate the charts out of the data and design the dashboard** (details will be introduced in the following sections). After the data get imported, the major work of visualization is to explore the relationship between different data fields (or

dimensions) and show them. In Tableau Public, we can achieve it by choosing the proper combination of data fields, generating the charts and layout the charts in the work space. The final results are shown as a dashboard of several components for interaction.



Example of using Tableau for visualizing the number of medals in different countries

3. **Share it to the web.** We can publish the final visualization results by sharing it as a webpage. This requires registration of Tableau account.

Example: 2012 Olympic Athletes

In this section, we walk through the whole process of visualization. We use an open data set of medals and athletes from each Olympic Games since the 2000 Games in Sydney, Australia.

Before start, we may have some analysis ideas:

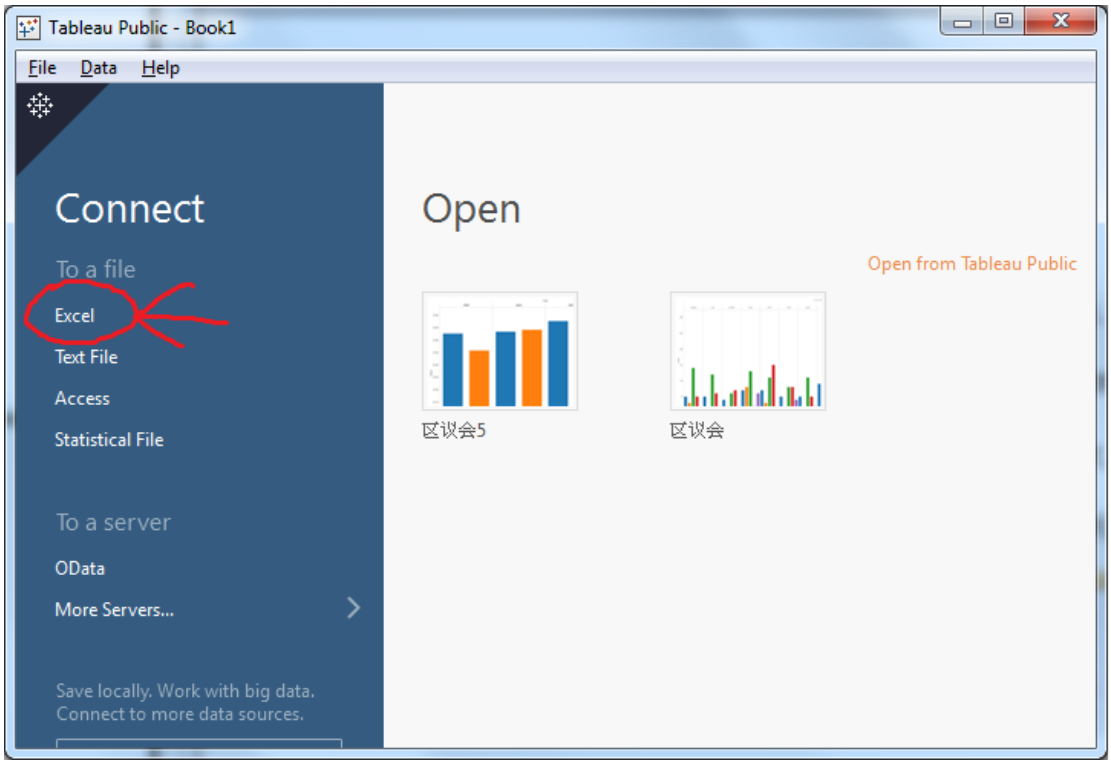
- *How many medals have athletes won since the 2000 Games?*
- *How have the number of gold medals changed over time?*
- *Which countries have won the most number of gold medals?*

Okay, let us start.

1. **Check the data.** Open "OlympicAthletes_0.xlsx". We can see that there are 10 data fields for each record. It consists of strings (e.g. name), numbers (e.g. age, number of medals), time (e.g. year, closing ceremony data), places (e.g. country), categories(e.g. sport type).

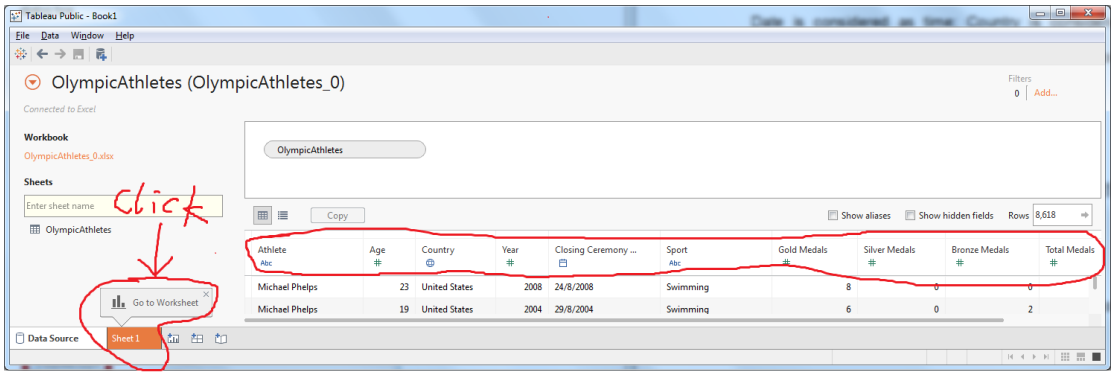
	A	B	C	D	E	F	G	H	I	J
1	Athlete	Age	Country	Year	Closing Ceremony Date	Sport	Gold Medals	Silver Medals	Bronze Medals	Total Medals
2	Michael Phelps	23	United States	2008	8/24/2008	Swimming	8	0	0	8
3	Michael Phelps	19	United States	2004	8/29/2004	Swimming	6	0	2	8
4	Michael Phelps	27	United States	2012	8/12/2012	Swimming	4	2	0	6
5	Natalie Coughlin	25	United States	2008	8/24/2008	Swimming	1	2	3	6
6	Aleksey Nemov	24	Russia	2000	10/1/2000	Gymnastics	2	1	3	6
7	Alicia Coutts	24	Australia	2012	8/12/2012	Swimming	1	3	1	5
8	Missy Franklin	17	United States	2012	8/12/2012	Swimming	4	0	1	5

2. **Import data.** Open Tableau Public. We get welcome page. Click Excel to connect to data. Find the "OlympicAthletes_0" dataset, and import it.



Wait for the program to import and we will obtain a set of dimensions and measures ready for exploration. Note that the types of the data fields are not the same: Athletes, Sport, Year are considered as strings; Closing Ceremony Date is considered as time; Country is considered as locations. As for measurement, there are more statistics calculated from the original data fields, such as Age, number of medals, etc. These dimensions and measures can be used for proper chart generation. Then click "Sheet

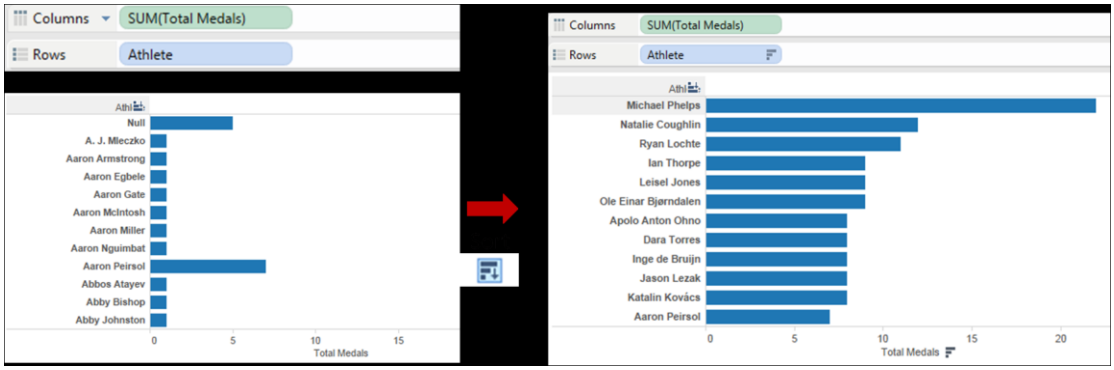
1".



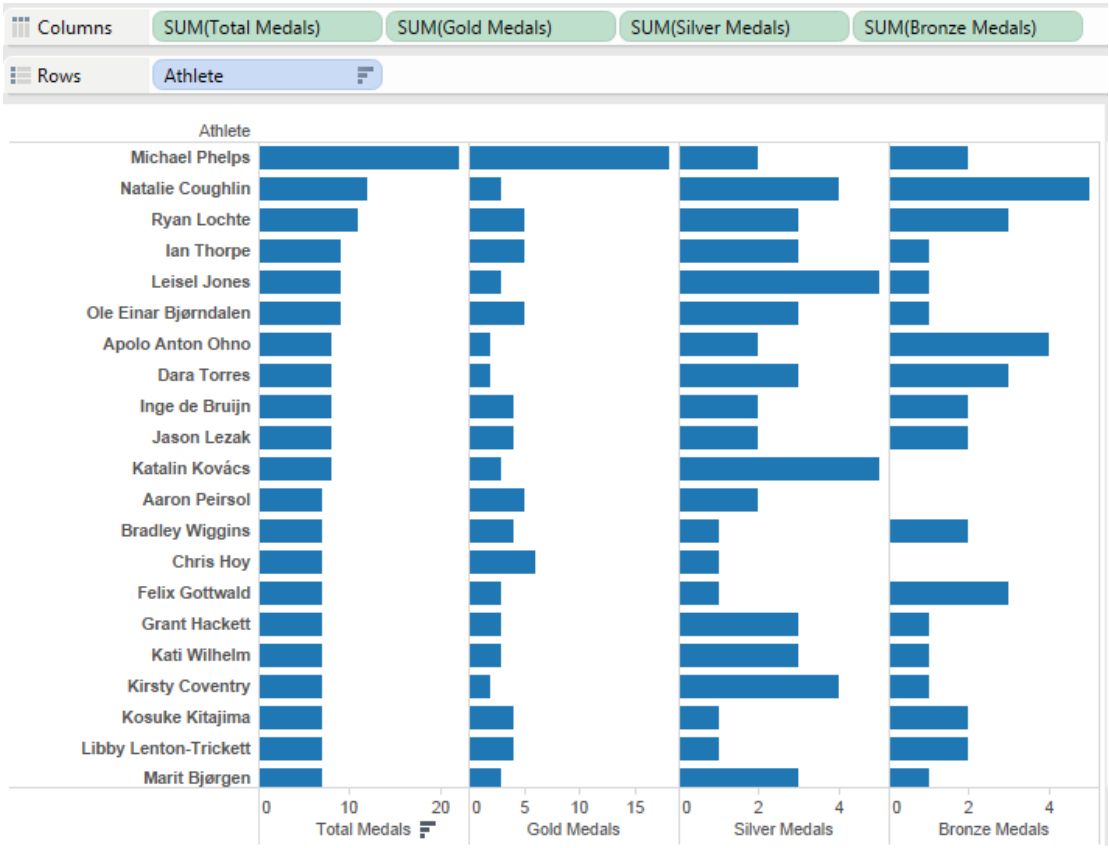
3. **Design single sheet.** Let us start by answering the previous questions.

How may medals have athletes won since the 2000 Games?

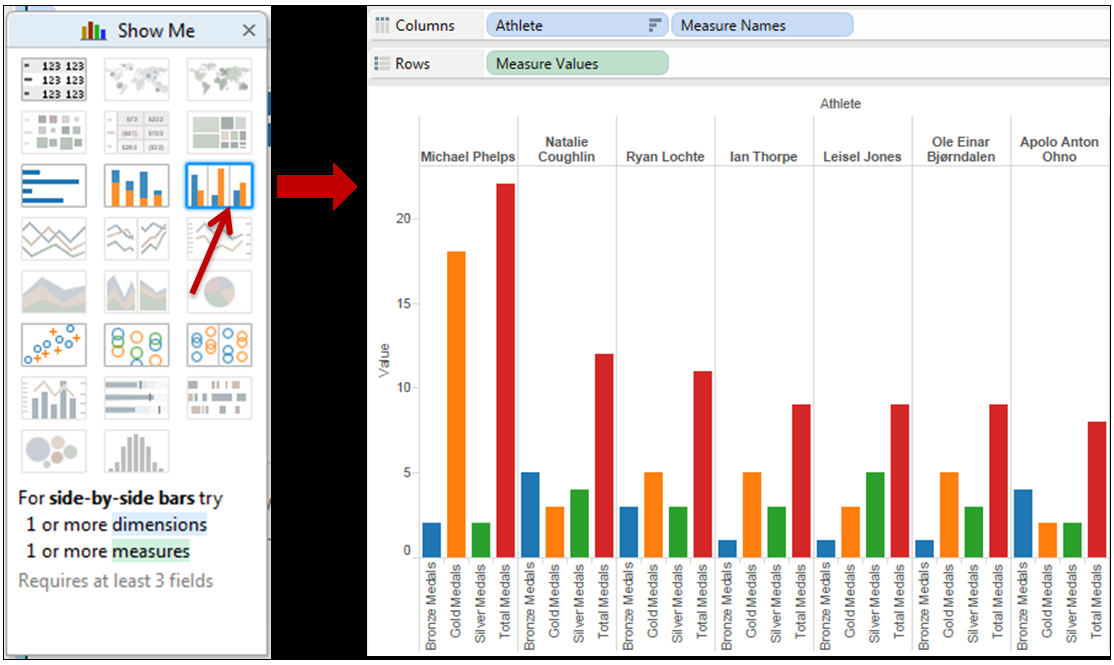
To answer this, the number of Total Medals, and the name of Athletes are related. We first drag Total Medals to Columns and drag Athletes to Rows. Ignore the warning of too many Athletes, and we have a bar chat as below. Right click sort icon when your mouse hovers on the header region and then click sort. Check Descending and Field, then click OK. Finally we find that Micheal Phelps win the most medals since 2000.



Furthermore, drag Gold Medals, Silver Medals and Bronze Medals to the Columns, and we obtain a new chart showing the details of medals obtained by each athlete. This enables us to sort the table by different type of medals as well.

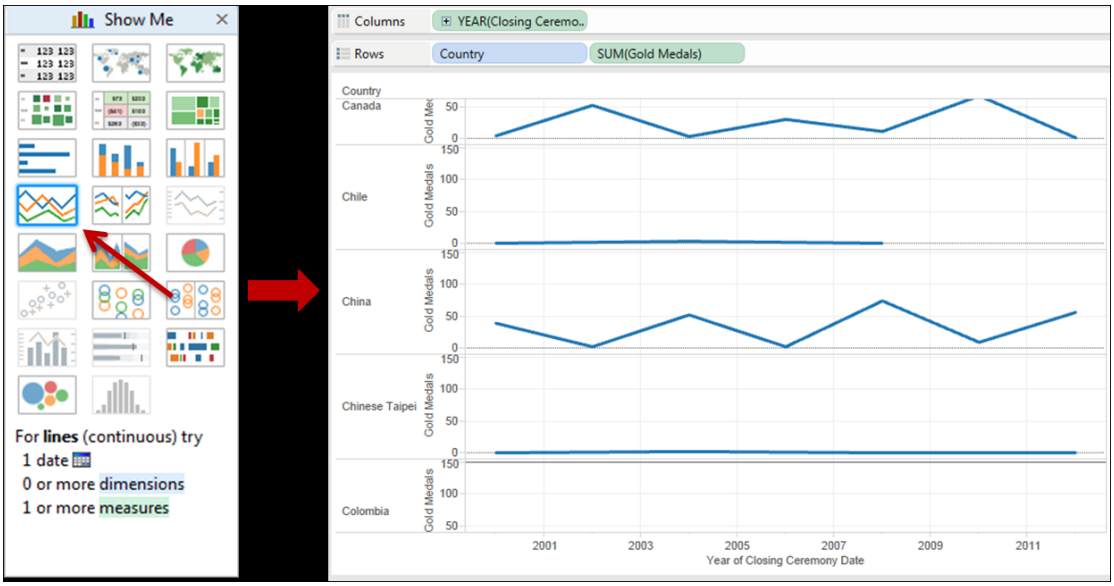


However, you may find that the scales of each bar chart are not the same. Because the sum of Gold, Silver and Bronze Medals are Total Medals, we need a view that follows this principle as well. This can be done by merging the four bar charts into one. Click side-by-side bars in Show Me panel, and we will obtain a novel view that every measurement is in the same scale.

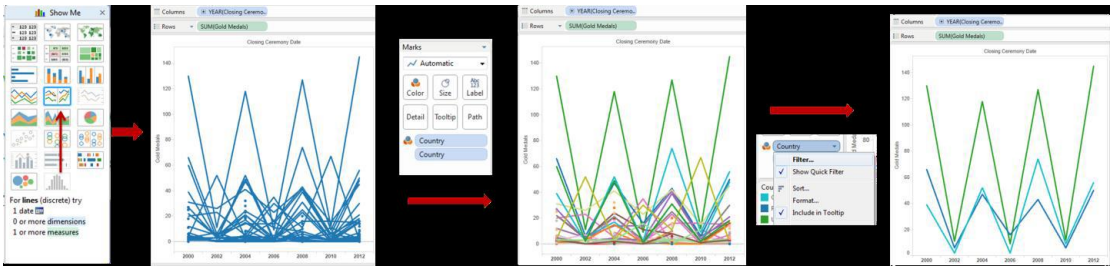


How has the number of gold medals changed over time?

To answer this, we consider Closing Ceremony Date, Country, Gold Medals. Create a new sheet (clicking the new sheet icon at the bottom of the window). Multiple select (press control and click) these three items (Closing Ceremony data, Country and Gold Metals), and then click lines in Show Me. We will obtain a chart showing the number of gold medals changing over years.



There is still not intuitive as we are more interested in the comparison between countries. Click lines (discrete) in Show Me panel and it will be layout again and the dimension Country is omitted but put into one chart. To differentiate them, we need to drag Country to Color icon in the Marks panel. Still, there are too many countries to see. We can further filter countries and leave three countries, China, Russia and The United States. As a consequence, a chart with three lines is shown for detailed comparison.



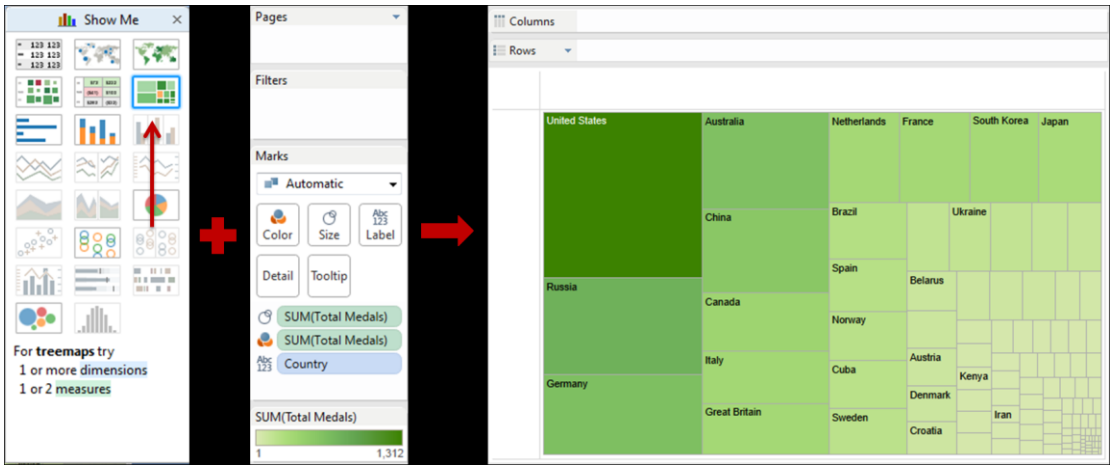
Which countries have won the most number of gold medals?

Another way to compare many countries is to visualize them with a map. Select Country, Latitude, Longtitude, and choose symbol map in Show Me panel. We will see a map of dots showing position of different countries. What is more, we can further specify the size of dots to visualize the total number of gold medals for each

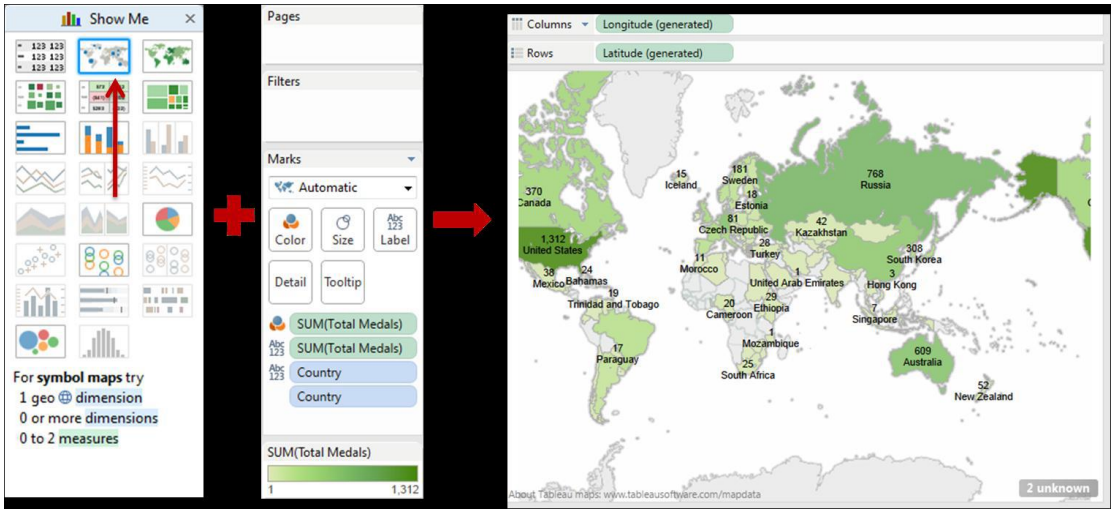
country. This can be done by dragging the Gold Medals to Size icon in Marks panel. The chart can be fine tuned by dragging Country and Gold Medal to Label icon in Marks panel. In addition, we can also visualize the average of age of athletes for each country by dragging Age to Color icon in Marks panel.

4. **Design the dashboard.** After we visualize the data from different perspectives, our next job is to put all these together. Suppose we focus on the analysis of countries, and we are interested in how the medals are distributed over these countries. First of all, we can create three separated sheets by clicking the new sheet icon at the bottom of the window.

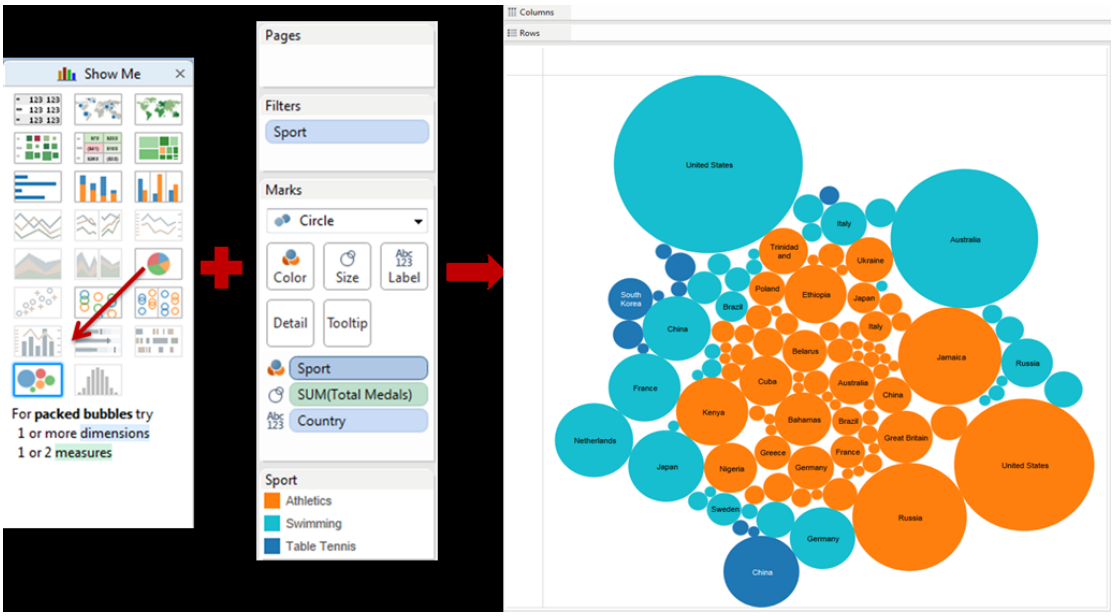
The first figure we focus on the portion of medals won by different countries. Therefore, first select Country and Total Medals and choose tree-map in Show Me panel. Drag Total Medals to Size icon and Color icon and drag Country to Label icon in Marks panel.



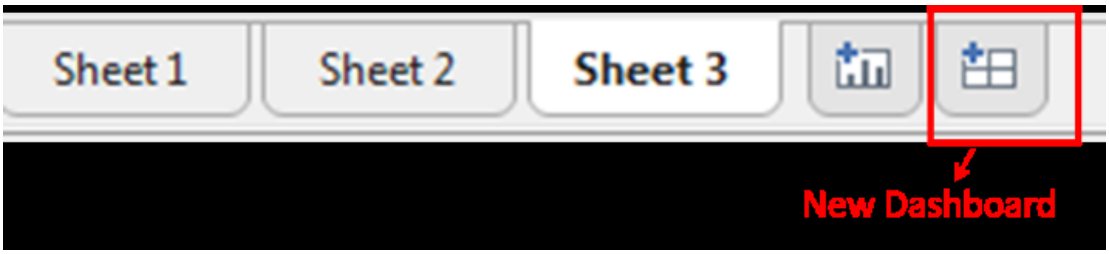
The second figure we focus on the distinction of the medals. Select Country, Latitude, and choose symbol map in Show Me panel. Drag Total Medals to Color icon and drag Country and Total Medals to Label icon in Marks panel.



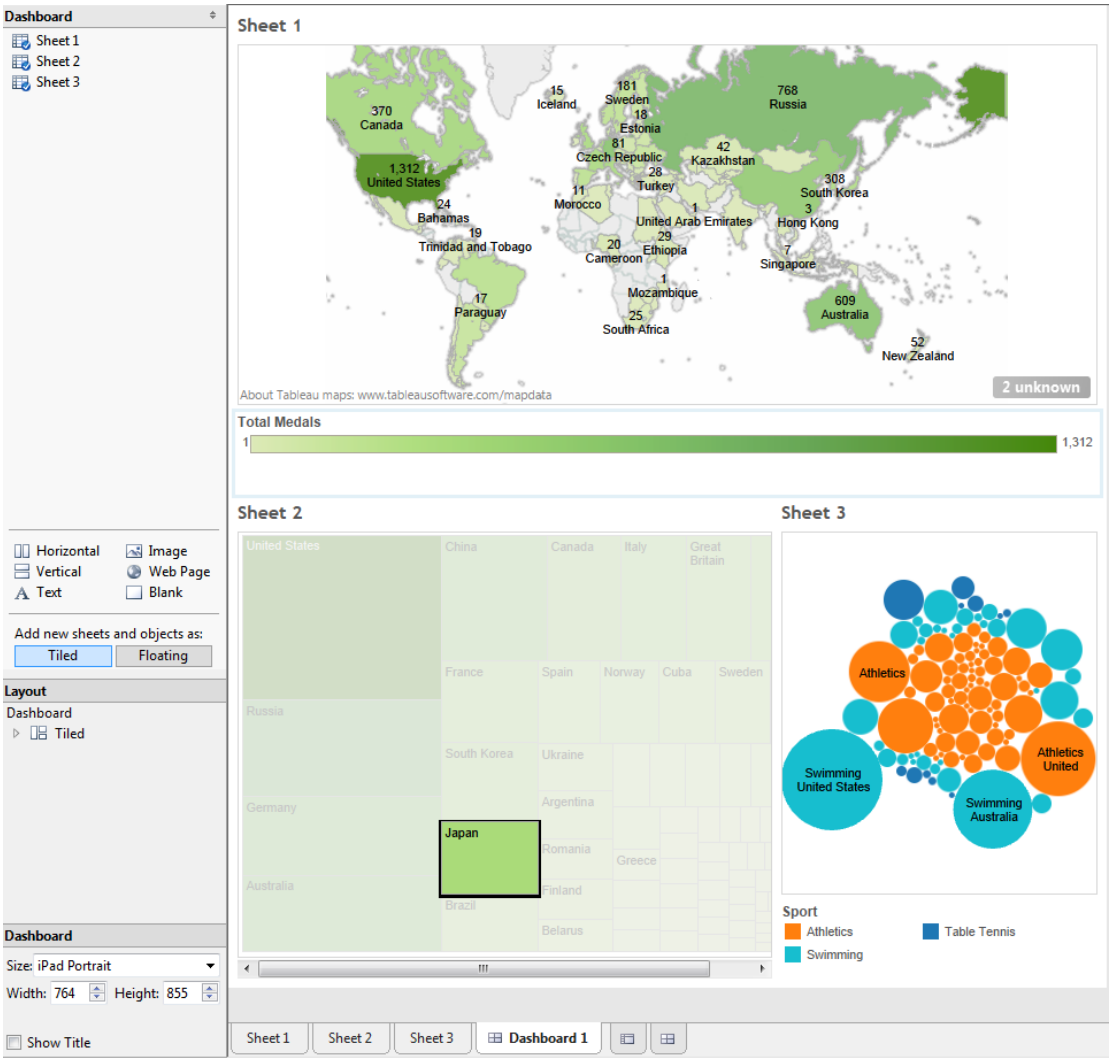
The third figure boils down by sport type. This time we use packed bubbles to help us quickly spot the dominating countries in certain fields, such as swimming, athletics, and table tennis. Drag Sport to Color icon, Sum of Total Medals to Size icon and Country to Label icon. Filter the Sport and limit the scope of our observation on three fields.



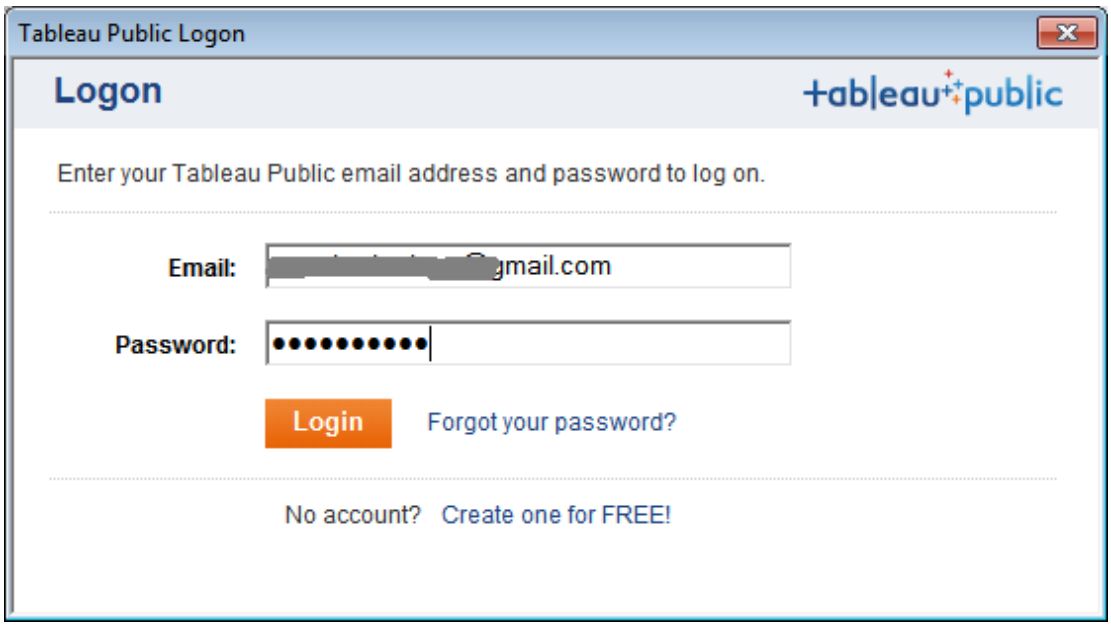
Finally, we can put these three sheets together by choosing the Dashboard icon below.



First, we need to consider the size of dashboard. Suppose we are going to show it in an iPad, so we select iPad Landscape in the dropdown menu. Next, drag Sheet 1, 2 & 3 into the dashboard. Finally we get our visualization results like below.



5. **Finally, we can save it to web**. (Note: you would need to register a account in order to complete this step) Choose File -> Save , and log in with the registered Tableau account.



Our visualization workbook will be uploaded to the server. And you can see it automatically in the browser.