Lab 03 Tableau – Treemaps and Dashboard

COMP7507 Visualization and Visual Analytics

Oct 8ᵗʰ, 2024

Goal

In this lab session, we will go through an example on how to use Tableau for problem solving when you have some analysis ideas in mind. You will also learn how to publish your visualization results online and share with your partners.

Example 01: Analysis of 2012 Olympic Athletes

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

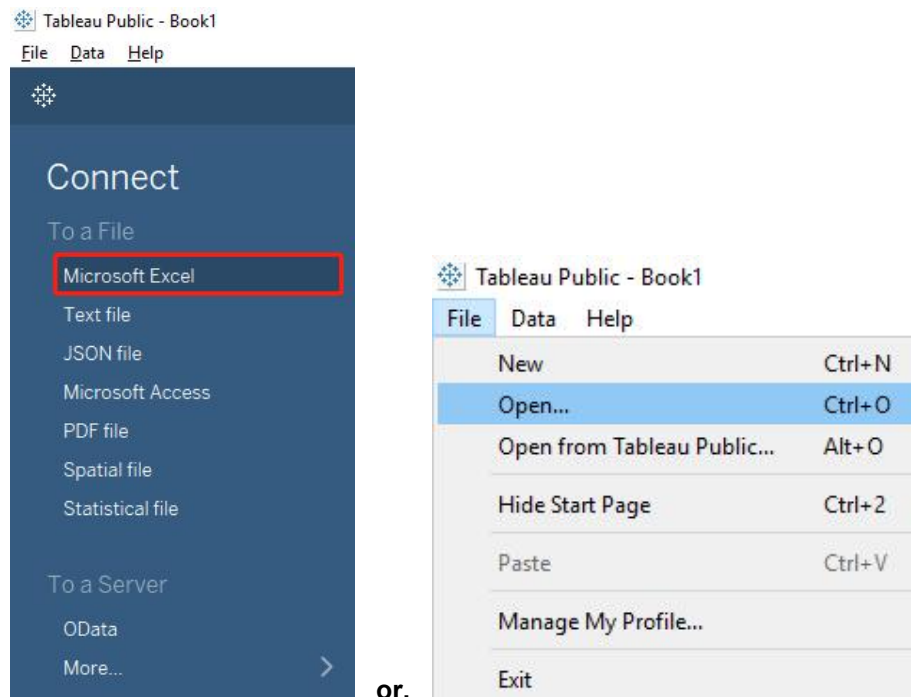Now, assume we have interest on these questions for our dataset:

- **How many medals have athletes won since the 2000 Games?**

- **How has the number of gold medals changed over time?**

- **Which country has won the greatest number of gold medals?**

To address these questions, we will go through the visualization processes step by step.

1. **Check the data**. Open OlympicAthletes_0.xlsx (download the file from Moodle page). We can see that there are 10 data fields for each record. It contains strings (e.g. name), numbers (e.g. age, number of medals), time (e.g. year, closing ceremony date), locations (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 Med	Silver Medals	Bronze Medal	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
9	Ryan Lochte	27	United States	2012	8/12/2012	Swimming	2	2	1	5
10	Allison Schmitt	22	United States	2012	8/12/2012	Swimming	3	1	1	5
11	Natalie Coughlin	21	United States	2004	8/29/2004	Swimming	2	2	1	5
12	Ian Thorpe	17	Australia	2000	10/1/2000	Swimming	3	2	0	5
13	Dara Torres	33	United States	2000	10/1/2000	Swimming	2	0	3	5
14	Cindy Klassen	26	Canada	2006	2/26/2006	Speed Skatin	1	2	2	5
15	Nastia Liukin	18	United States	2008	8/24/2008	Gymnastics	1	3	1	5
16	Marit Bjørgen	29	Norway	2010	2/28/2010	Cross Countr	3	1	1	5

2. **Import data**. Open Tableau Desktop or Tableau Public. We get a welcome page. Click **Microsoft Excel** on the quick start page to load an excel spreadsheet or Click **File->Open** to load the excel file.
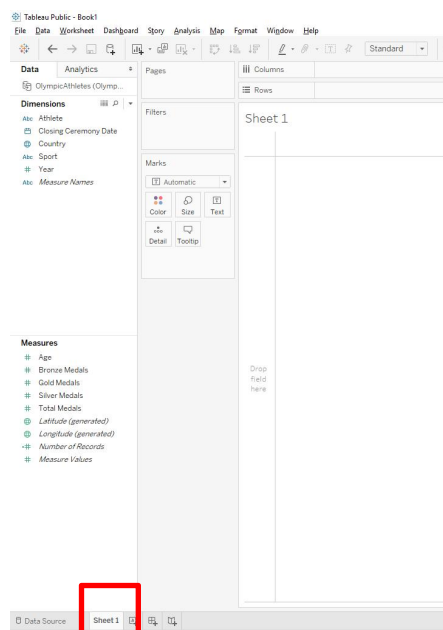

In this example, we load **OlympicAthletes_0.xlsx** and enter the data source page.

Then click **Sheet 1 tab** at the bottom to enter the workspace:

We now obtain a set of **Measure Names** and **Measure Values** ready for exploration. Note that the types of **Measure Names** are not the same: Athletes, Sport, Year are '**strings**'; Closing Ceremony Date is '**time**'; Country is of type '**locations**'. For **Measure Values**, there are more statistics calculated from the original data fields, such as Age, number of medals, etc.

3. **Design single sheet.** Let us start to answer the previous questions.

1) *How many medals have athletes won since the 2000 Games?*

To answer this question, we need to consider the relationship between the number of Total Medals and the name of Athletes. We first drag **Total Medals** to **Columns** and drag **Athletes** to **Rows**, and we have a bar chart as below. Click the **sort** icon at the tool bar line. Finally, we find that Michael Phelps win the most medals since 2000.
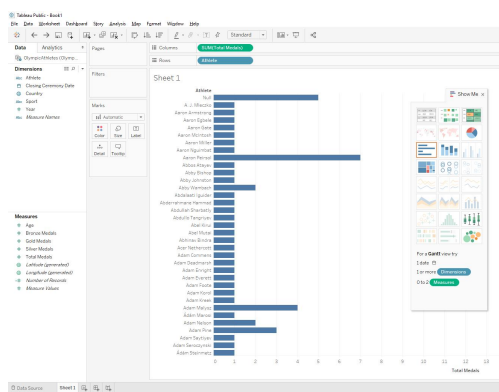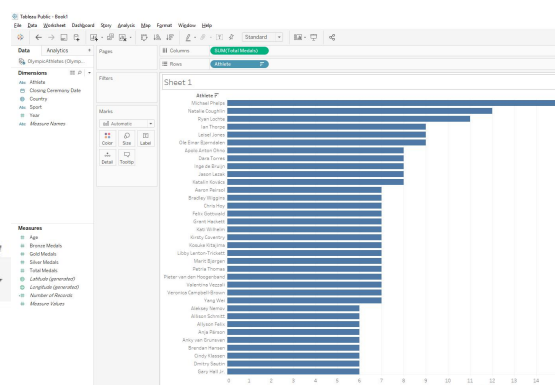
Furthermore, drag **Gold Medals**, **Silver Medals** and **Bronze Medals** to the **Columns**, and we will obtain a new chart showing the details of medals obtained by each athlete. This enables us to sort the table by different type of medals as well.
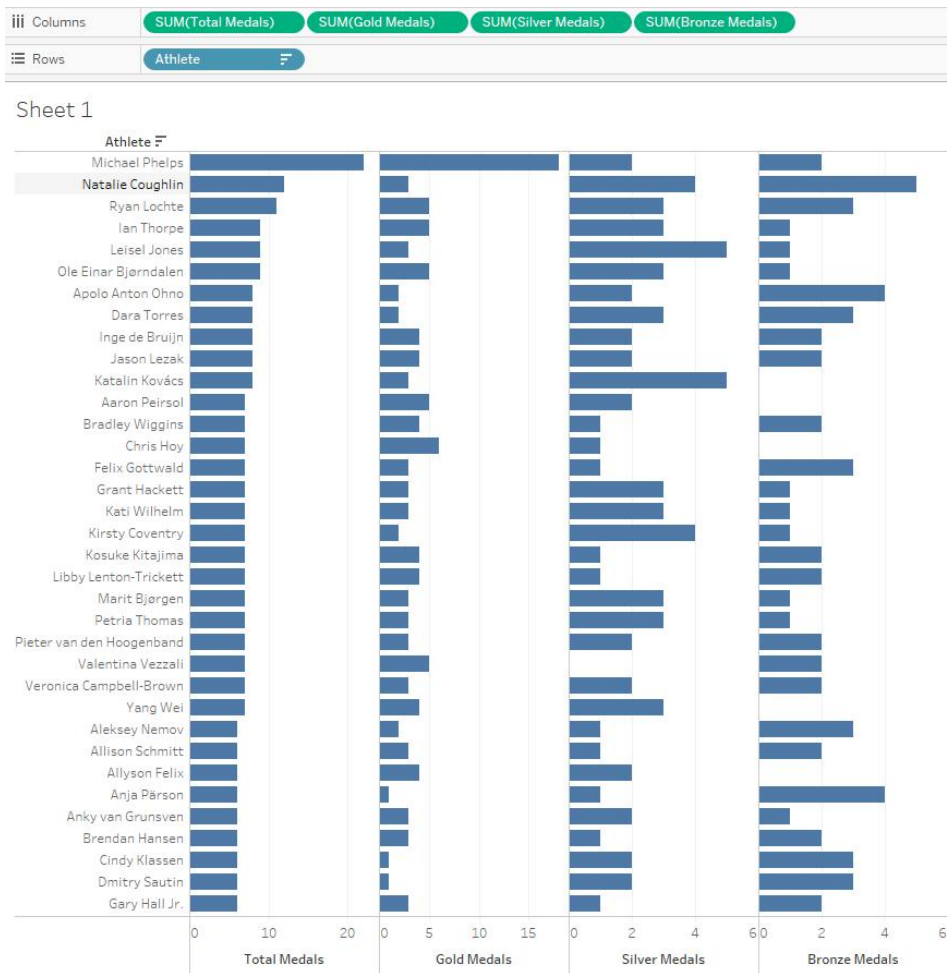
Then click **side-by-side bars** in **Show Me** panel, and we will obtain a novel view that every measurement is in the same scale. Click **Athlete** label and click the **sort** button, you will have the below graph.
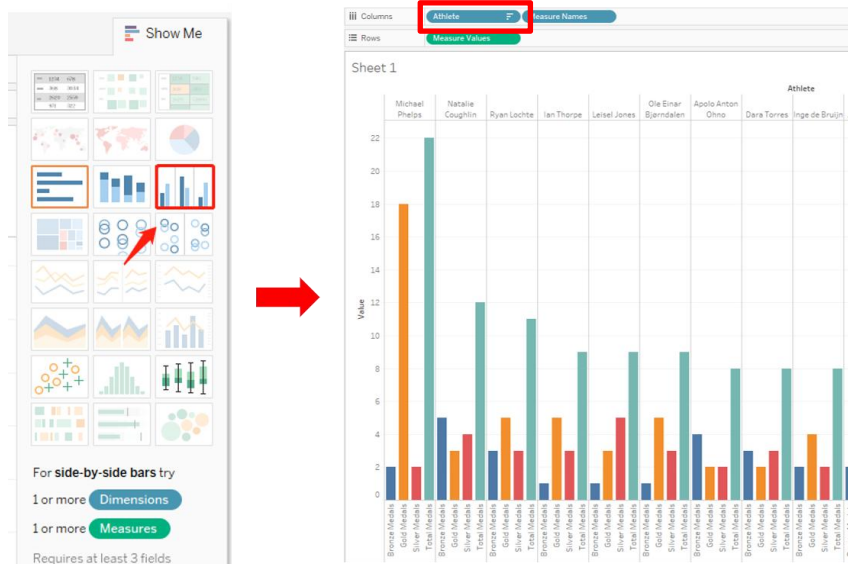
## 2) How has the number of gold medals changed over time?

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).

Then, multiple select (press control and click) these three items (**Closing Ceremony date, Country** and **Gold Metals**), and then click **lines**(continuous) in **Show Me**. We will obtain a chart showing the number of gold medals changing over years.
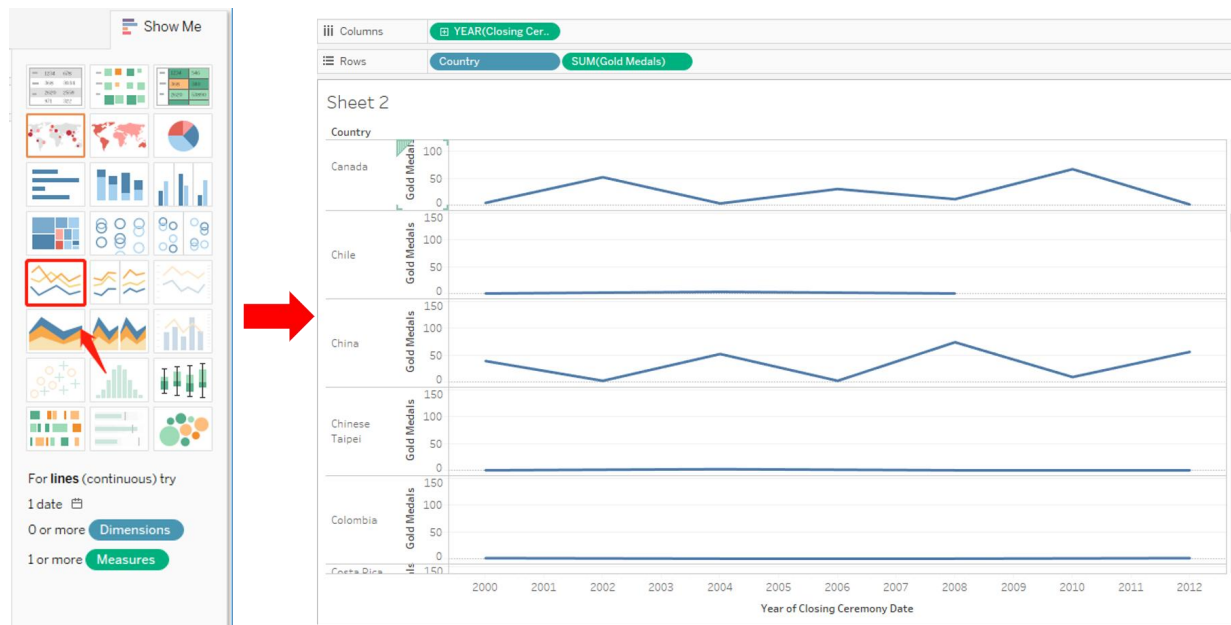

However, it is still not intuitive as we are more interested in the comparison among countries. Click **lines** (**discrete**) in **Show Me panel**, the dimension '**Country**' will be moved to **Marks** panel. To differentiate the countries, we need to drag **Country** to **Color icon** in the **Marks** panel. We can further filter out some countries and keep only three countries: China, Russia and United States by right clicking the **Country** label and choosing **Filter**. As a result, a chart with three lines is shown for detailed comparison.
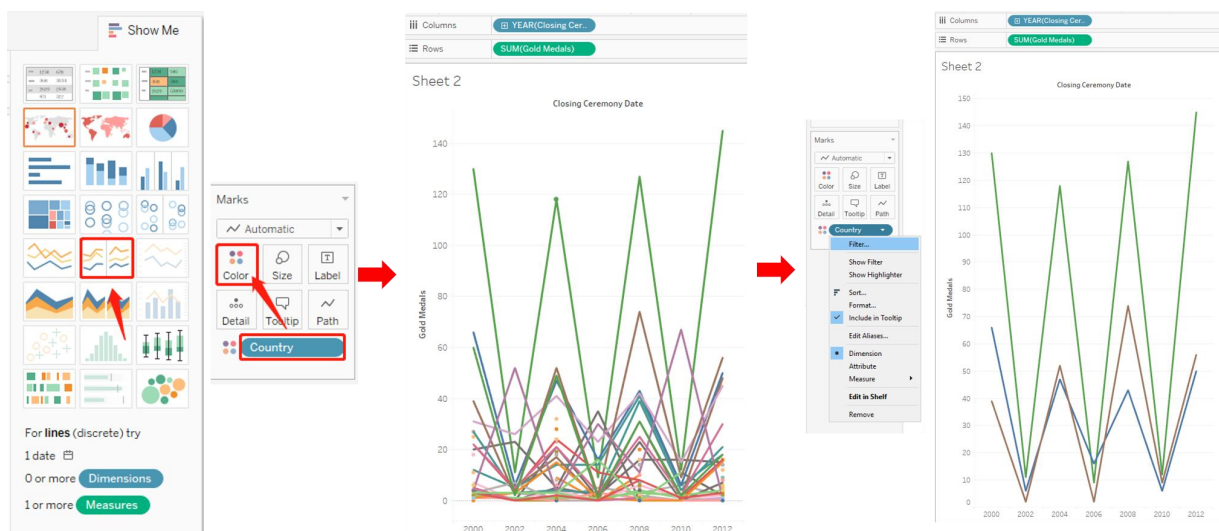
3) Which country has won the greatest number of gold medals?

Another way to compare multiple countries is to visualize them with a map. Create a **new sheet**, select **Country, Latitude** and **Longitude**. Then choose **symbol maps** in **Show Me** panel (use middle button on mouse to zoom in/out the map). We will see a map of dots showing position of different countries. Moreover, we can 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 further adjusted by dragging **Country** and **Gold Medal** to **Label** icon in **Marks** panel.
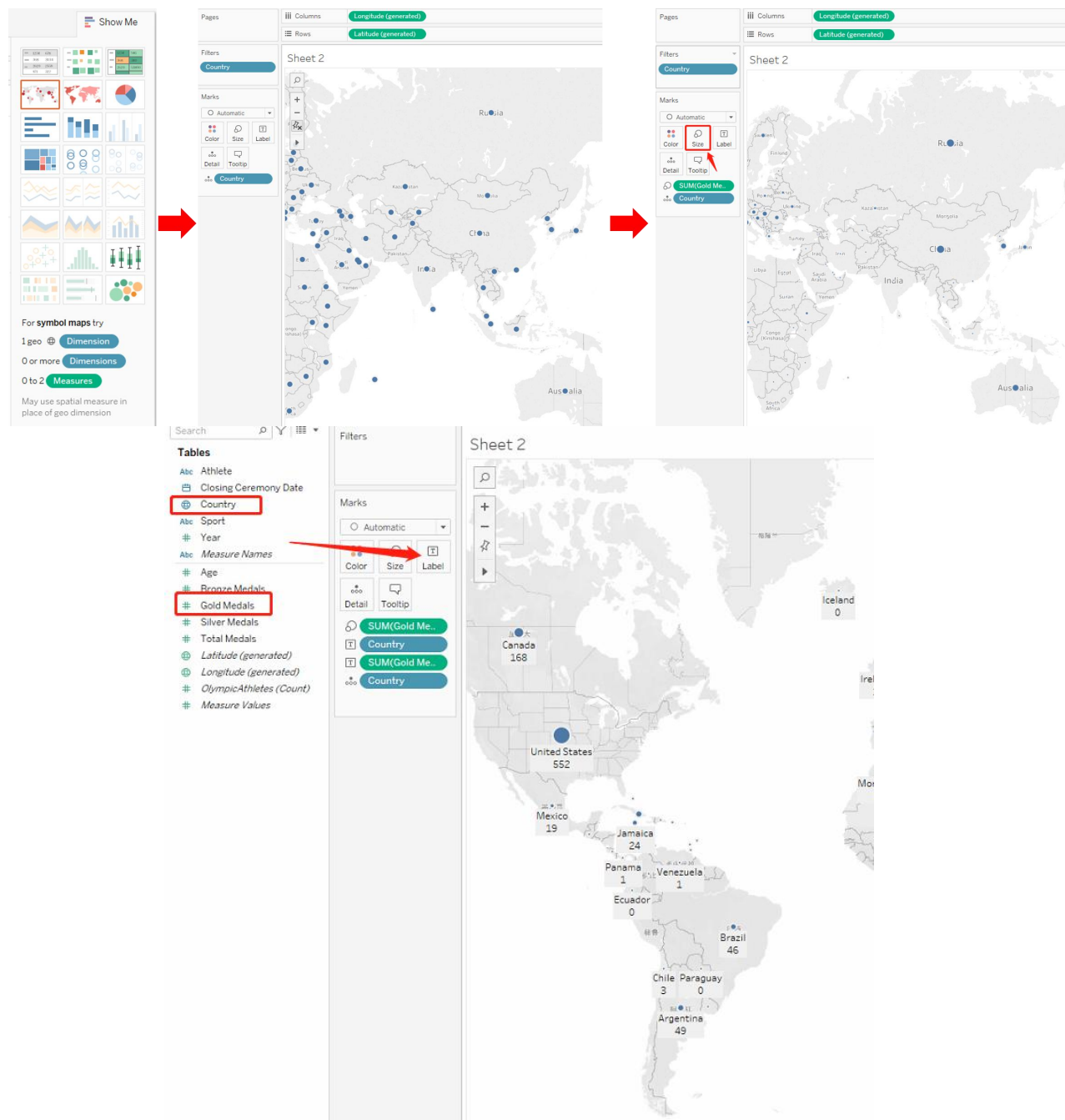

In addition, we can also visualize the average of age of athletes for each country by dragging **Age** to **Color** icon in **Marks** panel.
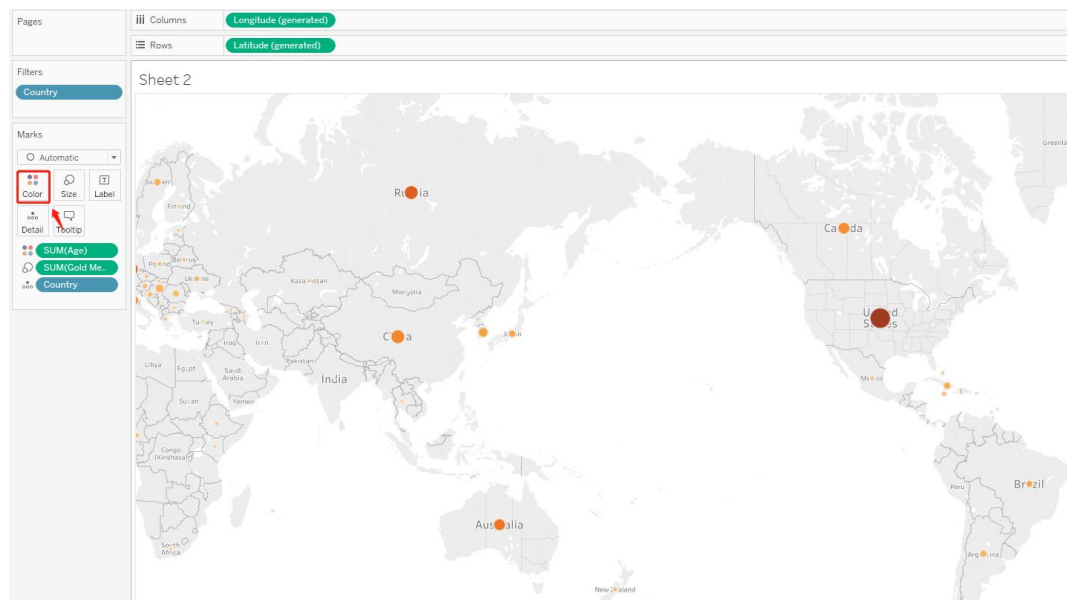
Now we can answer our three questions by three separate infographics.

4. Design the dashboard.

Suppose we focus on the analysis of countries this time, and we are interested in **how the medals are distributed over these countries,** but we want to use multiple infographics to show the distribution. Now, let's create three separate sheets by clicking the **new sheet** icon at the bottom of the window


New Sheet   New Dashboard
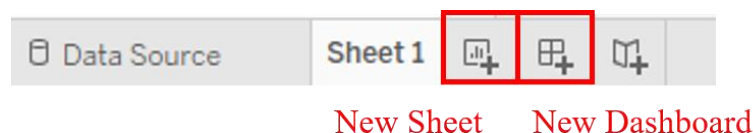
To visualize the portion of medals won by different countries. We first select **Country** and **Total Medals** and choose **treemaps** in Show Me panel. Make sure in the **Marks** panel, **Total Medals** is represented by **Size** and **Color**, and **Country** is shown by **Label**.

In the second sheet, visualize the distribution of the medals by map. Select **Country, Latitude, Longitude**, and choose **maps** in **Show Me** panel. Drag **Total Medals** to **Color icon** and drag **Country** and **Total Medals** to **Label icon** in **Marks** panel.
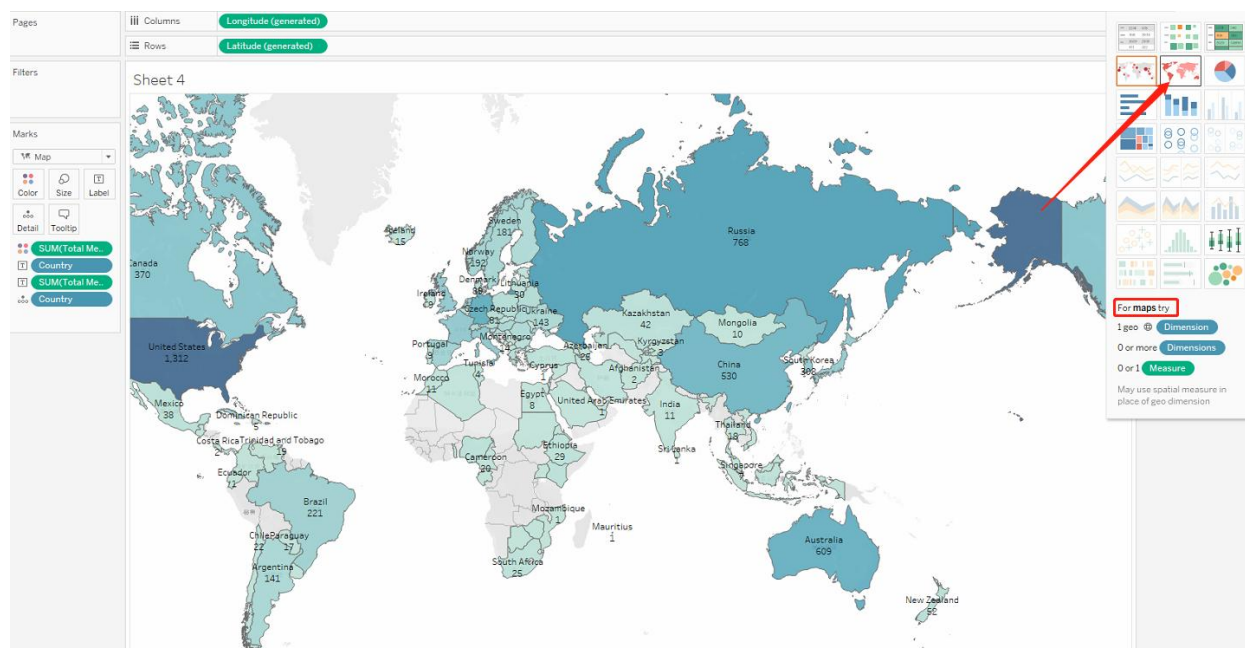

In the third sheet, we use **packed bubbles** to help us quickly spot the dominating countries in certain fields, such as swimming, athletics and table tennis. Drag **Sport** to **Color icon**, **Total Medals** to **Size icon** and **Country** to **Label icon**, choose **packed bubbles** in the **Show Me** panel. Filter the **Sport** and limit the scope of our observation on three fields (**Athletics, Swimming** and **Table Tennis**).
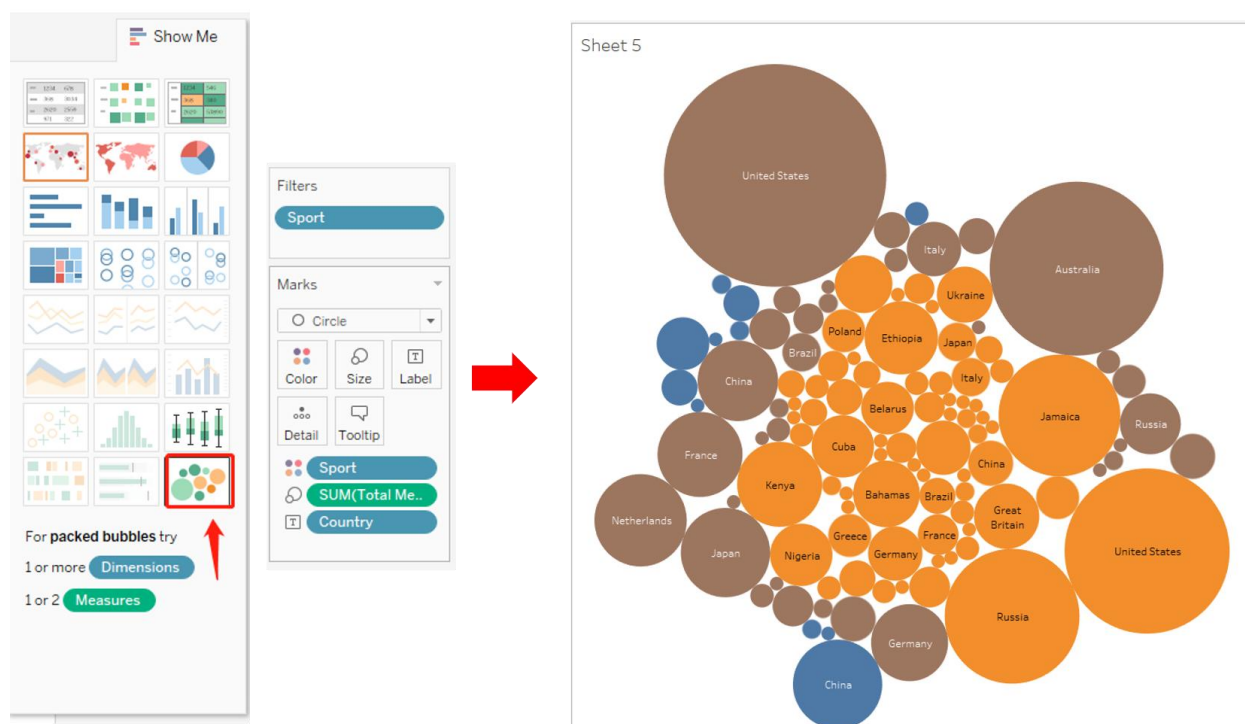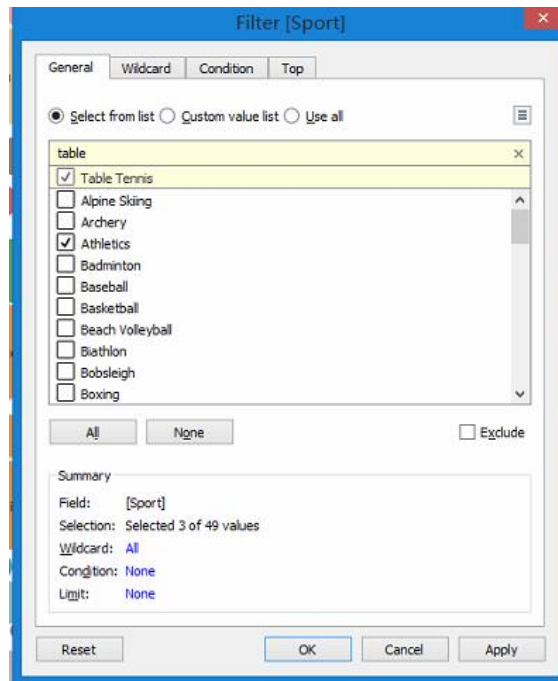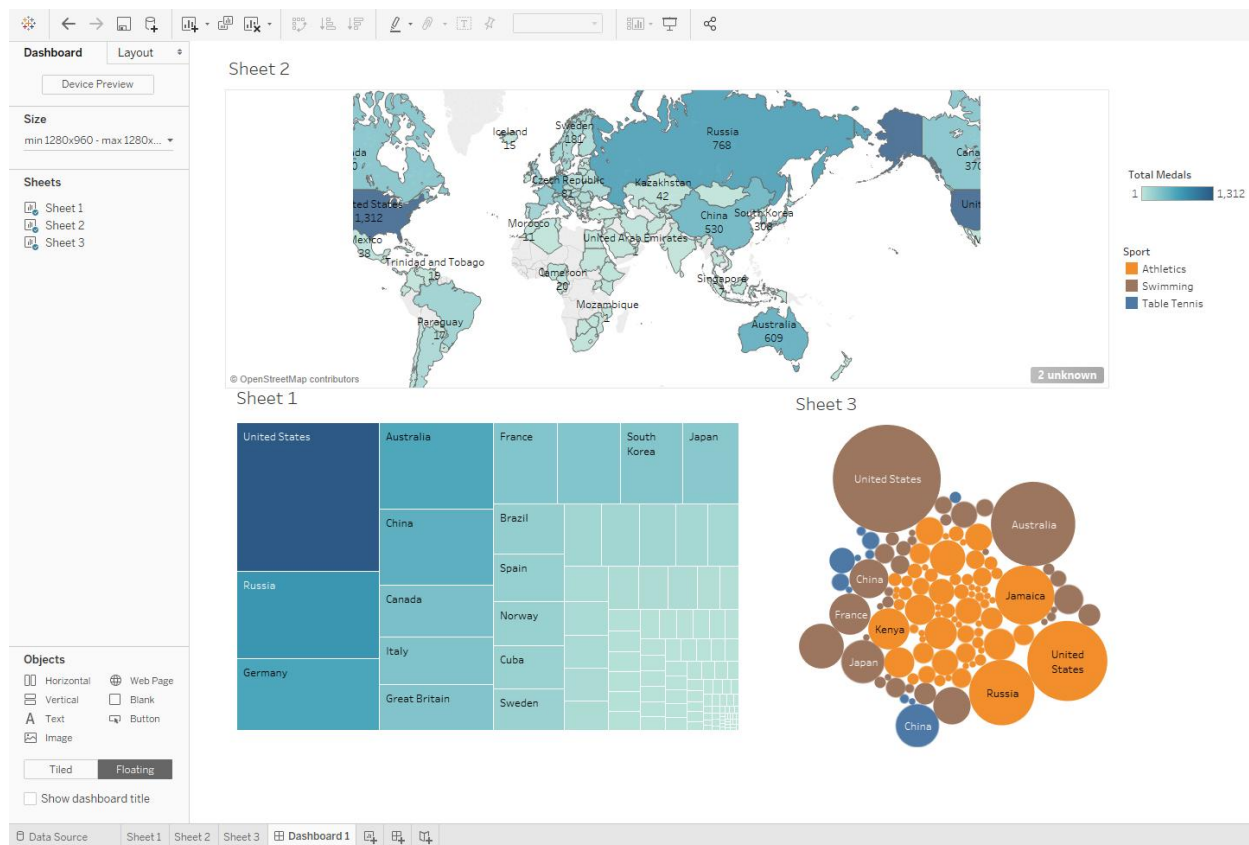
Finally, we can put these three sheets together by pressing the **Dashboard icon** below.
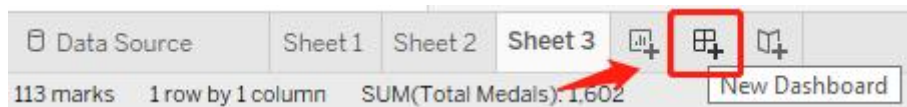

Then drag the three sheets we just created from the left **Sheets** panel into the dashboard (we need to adjust the size of the dashboard to make it fit well). Now, we get our visualization results like below.

**5. Publish the dashboard.** (Note: a tableau account is required)

a. If you are using Tableau Desktop, choose **Server** -> **Tableau Public** -> **Save to Tableau Public** to save your work on Tableau Public Server, or choose **File** -> **Save** to save your work locally.
b. If you are using Tableau Public, Choose **File** -> **Save to Tableau Public**, and log in with the registered Tableau account.
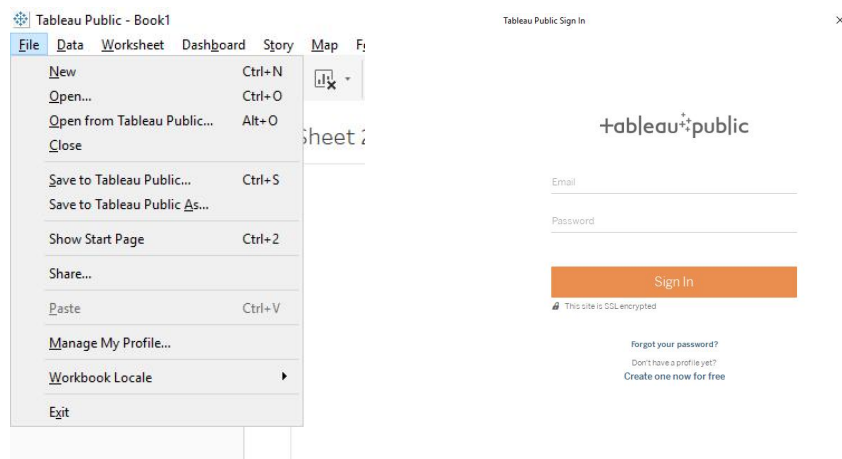
Save the workbook to Tableau Public:
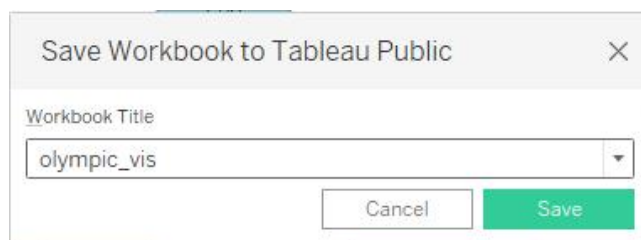
Now, your workbook will be uploaded to Tableau server and can be accessed by the links provided by Tableau Public server.
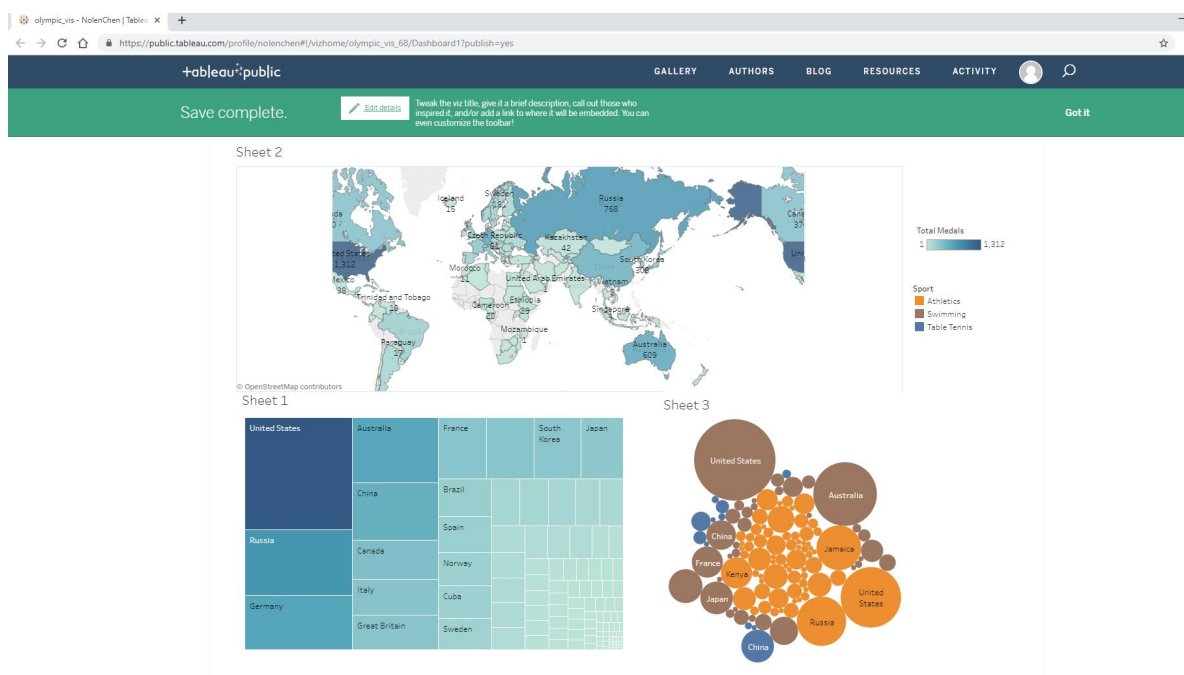
Some tasks for you

1) Connect to World Bank Indicators.xlsx (download from Moodle page) and try to build **a dashboard with 4 figures** to answer the following questions. You can summary **your answer** to each question in one sentence as **title of each figure**. (i.e. For figure 1 of the dashboard, you can set the title as 'A,B and C have the highest urban population')

   - Which countries have the highest urban population?
   - Where has Internet usage grown the most in the past decade?
   - What's the relationship between GDP and Health?
   - What's the relationship between Transit and Mobile Phone Usage?

2) Publish the final dashboard of task (1) to Tableau Public and provide a screenshot of your published dashboard to Moodle by **Oct. 22, 2024**.

Reference

[1] https://www.tableau.com/support

[2] http://www.tableausoftware.com/public/training

[3] http://www.tableausoftware.com/public/community/sample-data-sets

[4]http://www.tableausoftware.com/public/faq