

The video you just watched showed you how to make and visualize JOINS in Tableau. Now, you can use the datasets and instructions in this activity to perform the JOINS yourself. Feel free to refer back to the previous video if you get stuck.

In earlier activities, you worked in Tableau to create a data visualization. In this activity, you will review a scenario, link different data sources in Tableau, and create visualizations using multiple datasets.

By the time you complete this activity, you will be able to make visualizations out of data from multiple sources. This will enable you to visualize comparisons and combinations of data, which will allow you to share more complex projects in your career as a data analyst.


What you will need


Click the link to create a copy of the datasets and download them. Note that upon opening the CO2 dataset below, you will immediately see the About page load with a single URL code at top. Navigate to the bottom of the document, and you can access all the raw and cleaned CO2 data by clicking on the different sheets. All of these sheets will be accessed by the Tableau platform once loaded. If you don't have a Google account, download the datasets directly from the attachments below.

Link to datasets: [CO2](#), [energy](#), [total population](#), and [gdp total](#)

OR

Download datasets:

 [CO2](#)
XLSX File

 [Energy data](#)
XLSX File

 [total population](#)
XLSX File

 [gdp total](#)
XLSX File

The scenario

Imagine you are working as a data analyst at a policy research institute. For your current project, you need to create a visualization that shows the CO2 emissions per capita for each country from 2000-2011. You need to provide a visual presentation that not only allows someone to visually compare CO2 emissions between countries from year to year, but also provides information about each country's population, GDP, and energy use.

You already have a dataset that includes emissions for each country between the years 1960 - 2011. But, the information that you need on energy use, total population, and GDP you had to collect from a government website. Each dataset is in a separate file. Moreover, some of the information is missing for some countries.

Often you will work with datasets that are missing information. Whether or not you need to find this missing information will depend on your project. In this case, you will notice that the missing information is from the 1960s, 1970s, and 1980s.

Luckily, your project is only concerned with the data from 2000-2011. You need an efficient way to utilize some data from one source, and some data from other sources. Taking just the information that you need from each source and creating a new data source takes a lot of time.

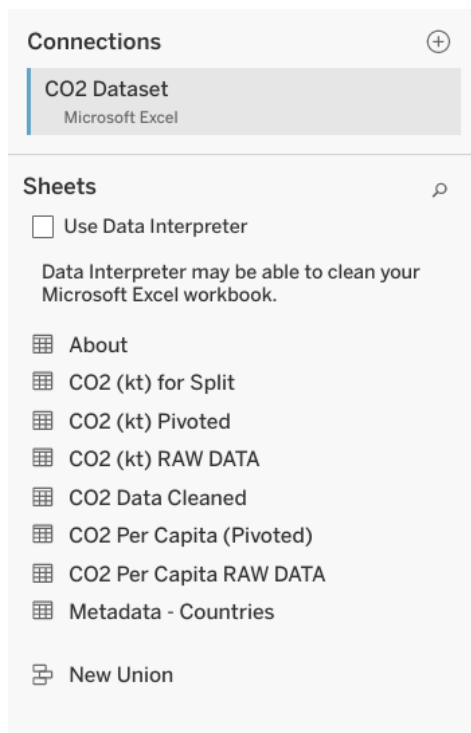
Tableau allows you to link data from different sources, as well as import data from different formats. While you won't be working with one in this assignment, Tableau allows you to use a Web Data Connector. This tool allows you to import the data you need directly from another site. Your visualizations will update when the data sources for your visualization are updated.

Load the data

1. Log in to [Tableau Public](#).

Note: Tableau frequently updates its user interface. The latest changes may not be reflected in the screenshots, but the principles in this activity remain the same. Adapting to changes in software updates is an essential skill for data analysts, and we encourage you to practice troubleshooting. You can also reach out to your community of learners on the discussion forum for help.

2. Go to your profile and click Create a Viz.
3. From the Connect to Data window, go to the Files tab and open the CO2 dataset you downloaded earlier.
4. From the Data Source tab on the bottom of the interface, go to the Connections header at the top of the left-side column.



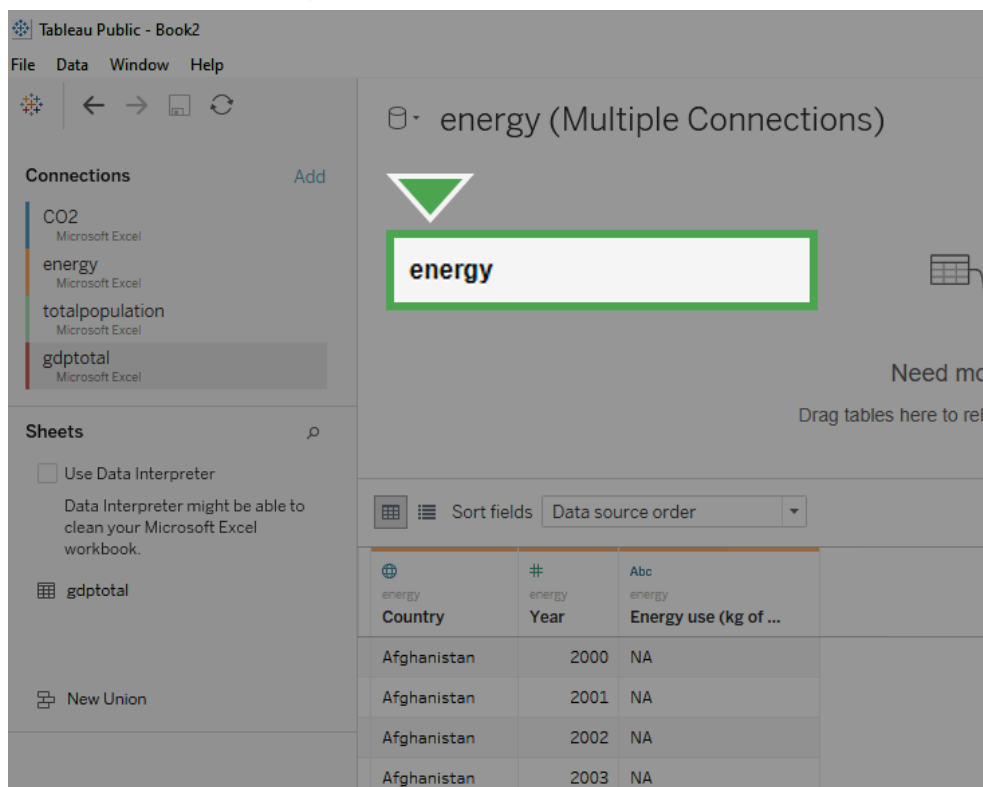
5. Click the + icon to add another data source. Start with the energy dataset.

6. Repeat step 5 for the other datasets, gdp total and total population.

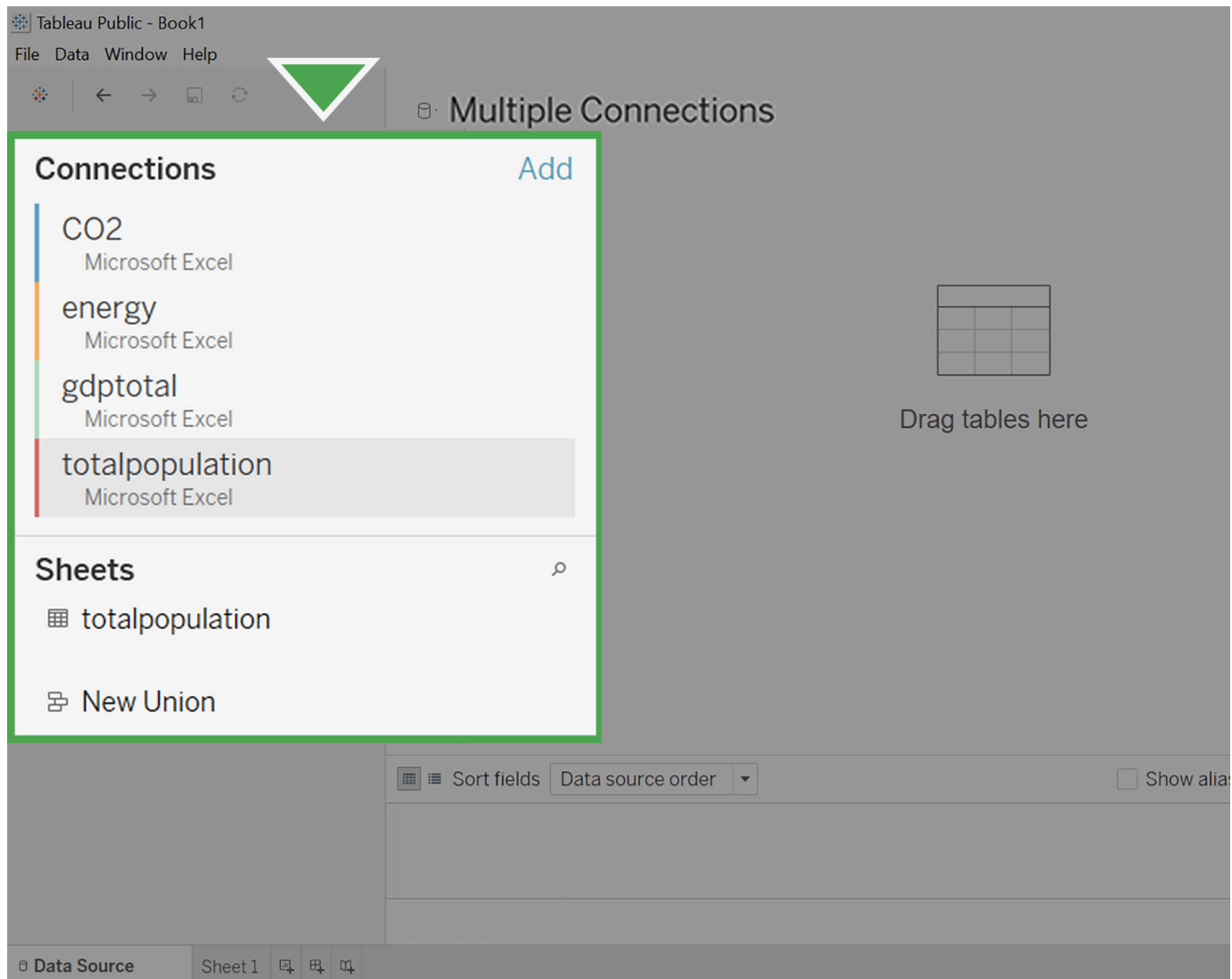
Now, you should have all four datasets loaded into Tableau. The datasets will be on the left-hand side of your screen under Connections.

Note: As you progress through the activity, make sure to save your progress by clicking File, then Publish or the Publish button in the top right corner of the screen. If you are asked to "Create an Extract," do so. It may take some time to create an extract of the data you are using in this activity, but it is essentially the same as saving your progress.

You'll notice that Tableau has already added one of the data sources into the area Multiple Connections. In the screenshot below, Energy is already loaded.



If one of the datasets has already been loaded in, you can remove it by dragging the box to the left-hand side (the grey area) of the screen. Click on the box labeled Energy in the center-top of the screen and drag it off to the left to delete it. Once you remove the data from Multiple Connections, it should appear similar to the image below.



Make connections with JOINS

Now, you'll set up the connections between the different datasets by creating JOINS between them. You learned about JOINS in the previous course.

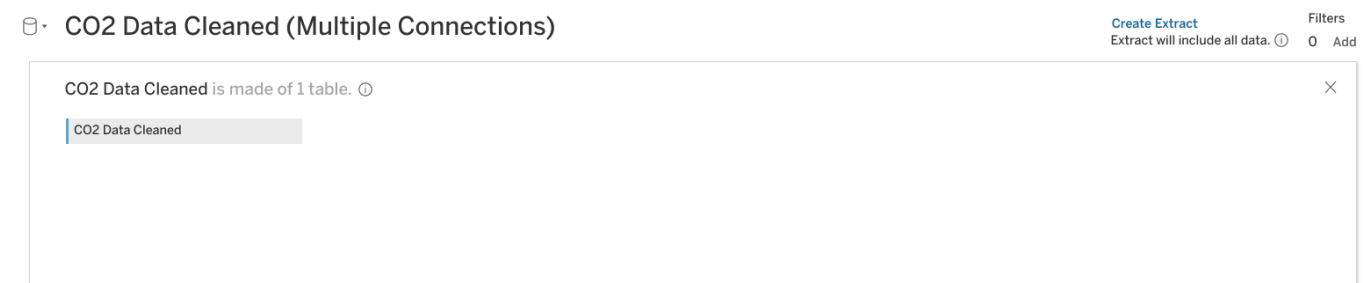
As a refresher: INNER JOINS and OUTER JOINS are types of relationships that can be used to combine data based on common columns of information.

In Tableau, you'll notice that the former JOINS window has now become a multi-purpose Relationships window. By double-clicking your data tables, you can edit the JOINS instead of the relationships.

Follow these instructions to create JOINS in Tableau:

1. Click on CO2 under Connections.
2. Under Sheets, you will notice all the different sheets in the CO2 dataset. Find CO2 Data Cleaned and double-click on it to load it.
3. Hover your cursor over the right side of the CO2 Data Cleaned box and click on the arrow.
4. Select Open to open the CO2 Data Cleaned dataset. *Make sure you complete this step.* This allows you to change the physical table, which will allow you to create JOINS. Otherwise, you will only be able to edit Relationships. Usually, you could use either option to accomplish the same goal. But for the purposes of this activity, we specifically want to use JOINS.


Your screen should appear similar to the screenshot below.







5. Click on the energy dataset under Connections.
6. Drag the energy sheet across to the CO2 Data Cleaned box under Multiple Connections. A pop-up window for a Join will appear.
7. The popup window may automatically populate with Year from CO2 Data Cleaned and Year1 from Energy. If not, put Year on the left side of the chart and Year1 on the right side.
8. Click on Add new join clause under Year. A dropdown menu will appear.
9. Select Country Name on the left side and Country on the right side.
10. Click the X to close the dropdown menu.

CO2 Data Cleaned+ (Multiple Connections)

CO2 Data Cleaned is made of 2 tables. ⓘ

CO2 Data Cleaned —  Energy

Join

 Inner
  Left
  Right
  Full Outer

Data Source		Energy
Year	=	Year1
Country Name	=	Country
Add new join clause		

Sort fields

CO2 Data Cleaned	CO2 Data Cleaned	CO2 Data Cleaned	CO2 Data Cleaned	CO2 Data Cleaned
Country Code	Country Name	Region	Year	CO2 (kt)

11. Click Update now to examine the dataset. You will notice that Year and Year1 have a number sign above them. Change the data type to date for each of these columns.

12. In the column, Year click on the # (not the arrow next to it) and select Date from the available options.

After completing the first field, you will notice a red exclamation mark between CO2 Data Cleaned and Energy. This indicates that the columns you have joined are no longer of the same data type. One is formatted as date, and the other numeric.


You will also notice that after changing Year (CO2 Data Cleaned) to a Date type, the data preview pane will no longer display properly.

13. To fix this, go to the column list in the lower left of the screen.

CO2 Data Cleaned+ (Multiple Connections)

Create Extract Filters
Extract will include all data. ⓘ 0 Add

CO2 Data Cleaned is made of 2 tables. ⓘ










CO2 Data Cleaned —  Energy

CO2 Data Cleaned

Name

CO2 Data Cleaned

Fields

Type	Field Name	Physic...	Rem...
	Country Code	CO2 Dat...	Count...
	Country Name	CO2 Dat...	Count...
	Region	CO2 Dat...	Region
	Year	CO2 Dat...	Year
	CO2 (kt)	CO2 Dat...	CO2 (...)
	CO2 Per Capita (metric tons)	CO2 Dat...	CO2 P...
	Country	Energy	Country
	Year1	Energy	year
	Energy use (kg of oil equiva...	Energy	Energy...

Data preview unavailable

14. Select the # icon next to the Year1 (Energy) column. Then, change the data type to Date.



Name

CO2 Data Cleaned

Fields

Type	Field Name	Physic...	Rem...
	Country Code	CO2 Dat...	Count...
	Country Name	CO2 Dat...	Count...
	Number (decimal)	CO2 Dat...	Region
	Number (whole)	CO2 Dat...	Year
	Date & Time	CO2 Dat...	CO2 (...)
	✓ Date	CO2 Dat...	CO2 P...
	String	Energy	Country
	Boolean	Energy	year
	Default	Energy	Energy...
	Energy use (kg of oil equiva...	Energy	Energy...

You may need to click Update Now in the preview pane to display the data properly. Make sure to repeat this step when you change more data types later on in this exercise.

The red exclamation mark will disappear. You might notice that all the years have been put into a month/day/year form with the default month and day as January 1st. This will not create any problems when creating a visualization, as you will filter the data by year.

Connect additional datasets

Before adding any additional joins, the data type for Year(Gdptotal) needs to change.

A pop-up window will appear for the join. It might already be populated with Year1 under Datasource and Year(Gdptotal) under gdptotal.

1. Click on gdptotal under Connections.
2. Under Sheets, drag the gdptotal sheet into the white space underneath the energy box.
3. Go to the column list in the lower left of the screen, scroll until you find the column Year(Gdptotal). Click on # above it. A drop-down menu will appear.
4. Select Date from the drop-down menu.

If the data preview does not display properly, fix the date type in the lower left pane.

5. Click on the Venn diagram between energy and gdptotal. Click on Add new join clause under year. A drop-down menu will appear.
6. Under CO2 Data Cleaned click on Country Name.

The screenshot shows the Tableau interface with two datasets, 'Energy' and 'gdptotal', connected by a Venn diagram. A 'Join' dialog box is open, displaying the join type as 'Inner'. The 'Data Source' is 'Year1' and the 'gdptotal' field is 'Year (Gdptotal)'. Below the dialog, a dropdown menu is open for the 'Year (Gdptotal)' field, showing various field types including 'Date', which is selected. The background shows a preview of the data with columns for Country Code, Country Name, Region, and Energy.

7. Click on the empty field under gdptotal across from Country Name. A dropdown menu will appear.
8. Set the right side of the join statement to Country1.
9. Close the Join pop-up by clicking on its exit button.

Now you are going to join totalpopulation, the last of the four datasets that you downloaded.

10. Click on totalpopulation under Connections.
11. Under Sheets, drag the totalpopulation sheet into the white space to the right of the energy and gdptotal boxes.

A pop-up window will appear for the join. It should already be populated with Year under Datasource and Year(totalpopulation) under totalpopulation.

12. Go to the column list in the lower left of the screen, scroll until you find the Year(totalpopulation) column. Click on # above it. A drop-down menu will appear.
13. Select Date from the drop-down menu.

If the data preview does not display properly, fix the date type in the lower left pane.

- Congratulations! You have successfully joined four different sources of data.

If the dataset had gone beyond those dates, you would have filtered out the unneeded years in your visualization.

20. Find the Abc icon above the Energy use column. Change it to Number (decimal).

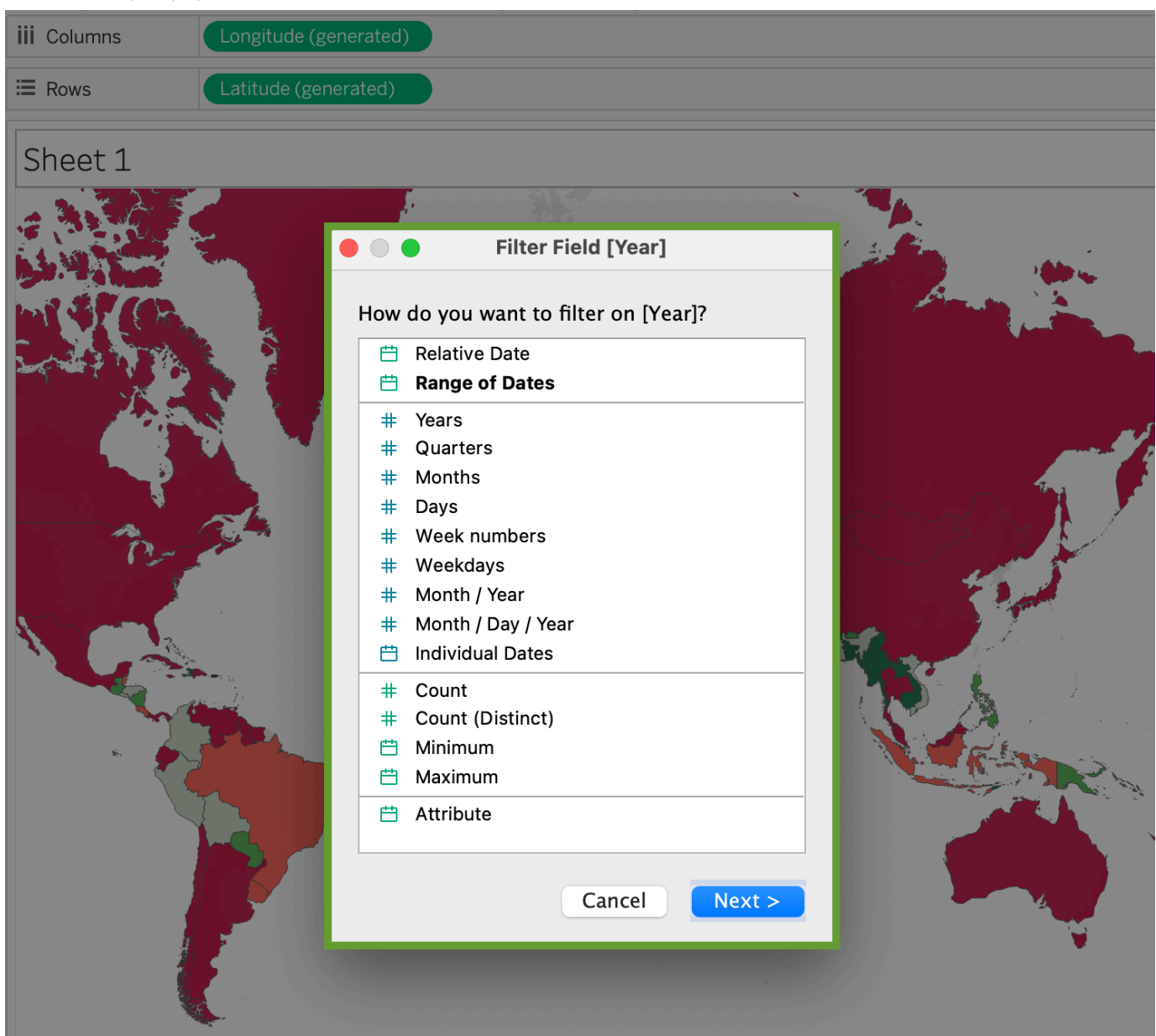
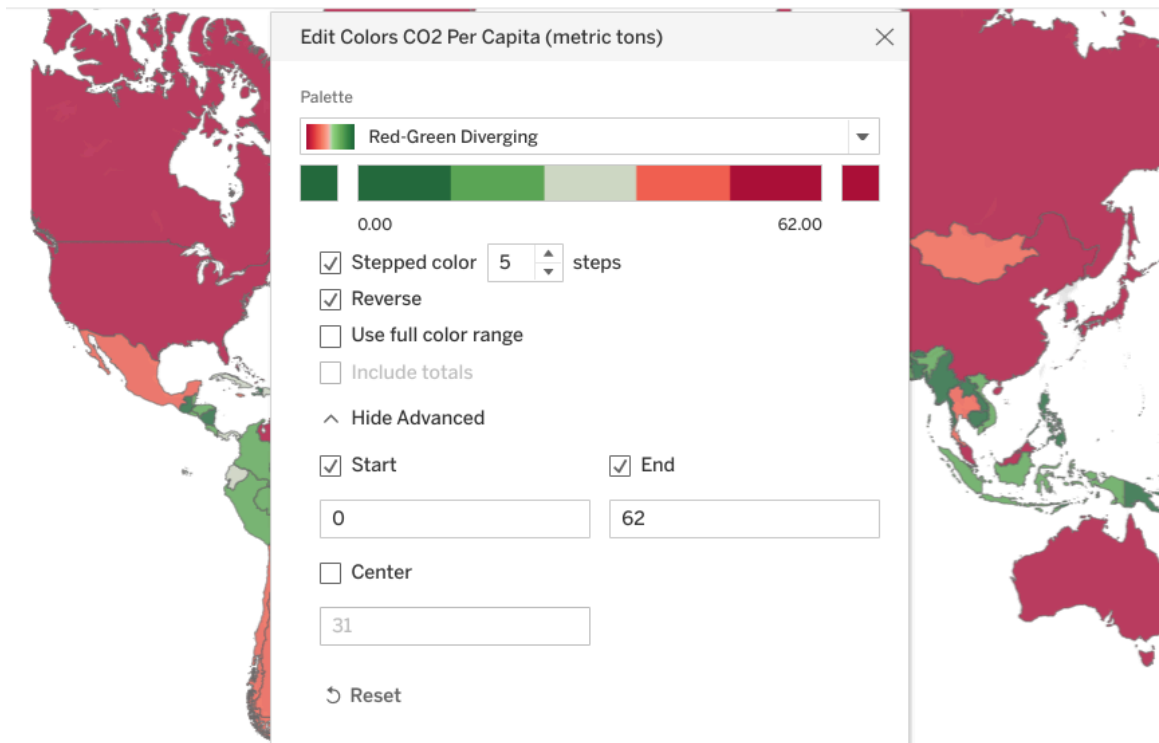
- If the data preview does not display properly, fix the date type in the lower left pane.

At the bottom of your screen, you will notice a tab labeled Sheet 1.

-
- The screenshot displays the Tableau interface for a world map visualization. The left sidebar lists data sources: 'CO2 Data Cleaned', 'Energy', 'gdptotal', and 'totalpopulation'. The 'CO2 Data Cleaned' source is expanded, showing fields like 'Country Code', 'Country Name', 'Year1', 'Year (Gdptotal)', 'Country (totalpopulation)', and 'Year (totalpopulation)'. The 'Country Name' field is highlighted with a green box. The top shelf shows 'Columns' with 'Longitude (generated)' and 'Rows' with 'Latitude (generated)'. The bottom shelf shows 'Filters' with 'Country Name'. A red box highlights the 'Detail' and 'Tooltip' marks in the 'Country Name' field. The map shows numerous blue dots representing data points across the world.

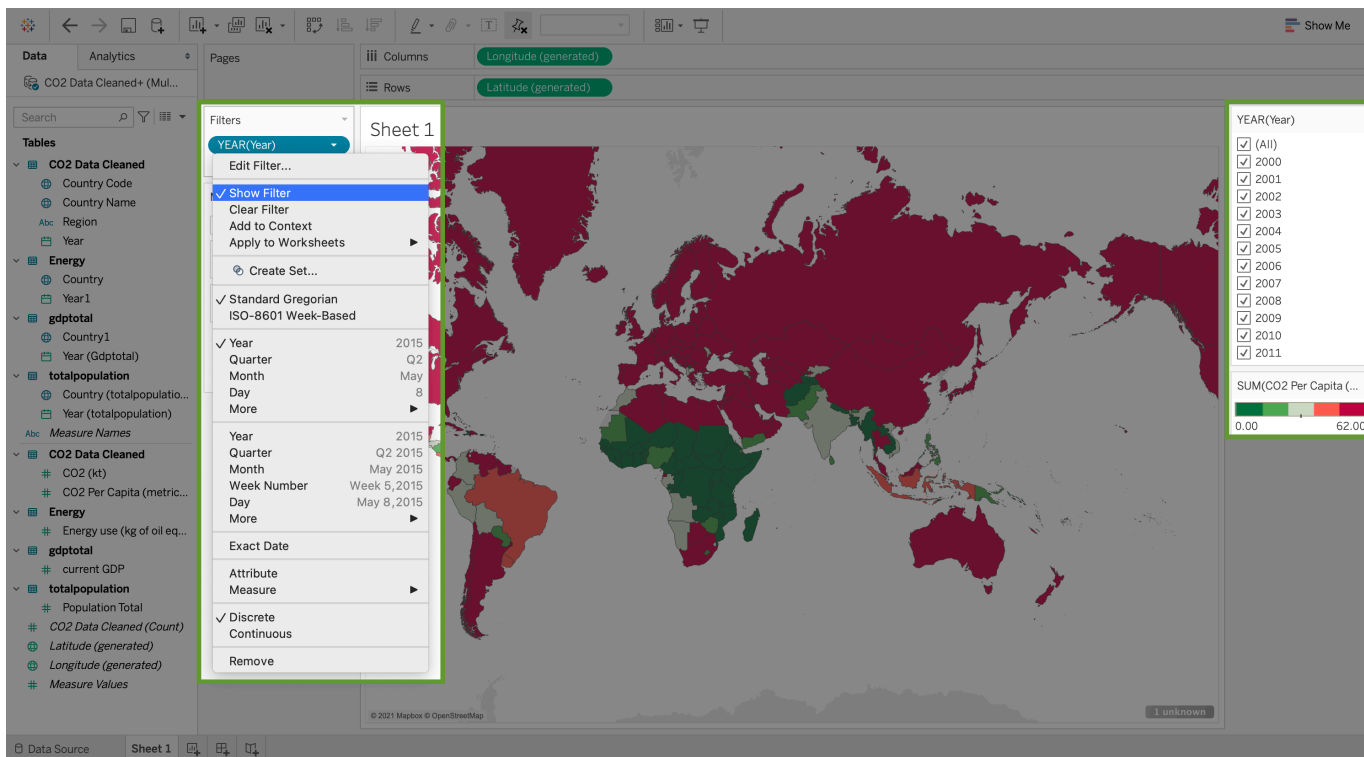
- The lowest CO2 Per Capita emission for any year is 0.0396 and the largest is 61.9898.

- Note: These values are the highest and lowest emissions between 2000-2011. Your screen should now look like this:



12. In the Filters box, right-click on YEAR(Year)

13. Select Show Filter. The filter will appear on the right side of the screen.



14. Click on the arrow to the right of YEAR(Year) on the far-right side of the screen.

15. Select Single Value (dropdown). Now the areas are colored only for the values of each year. Use the checkboxes in the list to choose which years you want to include in the visualization. You can select only the years between 2000-2011 to view the emissions relevant to the scenario. Congratulations! You've linked your data and made a comprehensive data visualization in Tableau.

Final Note: It is important to understand that you are not required to submit any exported dashboards or visualizations from the Tableau Public interface to move forward in this course. Your reflection answer is her for personal guidance and growth. Moreover, it is highly suggested that you export any visualizations from your gallery as a PDF or embedding the interactive dashboard for your portfolio.

For more information on exporting and embedding visualizations, please refer to the section titled, [Tableau Public: Export](#).

Reflection

In this activity, you used Tableau to link data together and visualize it. In the text box below, write 2-3 sentences (40-60 words) in response to each of the following questions:

What did linking data from multiple sources allow you to do with your visualization in Tableau?

What other kinds of datasets could you link to the four you used in this activity? What kinds of comparisons or insights could you make?

If you couldn't link data in this way, how would you make complex comparative datasets and visualizations like this?

What do you think?

Your answer cannot be more than 10000 characters.