

Vis Exercise 02

Introduction to Tableau Desktop

Uta Hinrichs

CS5044 – Information Visualization



University of
St Andrews

What is Tableau?

- Software tool to visualize and analyze your data
- Versatile data sources
 - CSV, JSON, MS Excel
 - Databases: MySQL, PostGreSQL, MongoDB...
 - Data warehouses/cloud: Amazon Redshift...
- Some data processing
 - Transformation of data types: geographic region → latitude/longitude
 - Some statistics
 - Possibilities to derive new data types
- Based on the principles of information visualization

Tableau Software

- Tableau Desktop <https://www.tableau.com/en-gb/products/desktop>
 - Authoring tool for interactive visualizations
- Tableau Public <https://public.tableau.com/en-us/s/>
 - Free platform for creating and publishing interactive visualizations on the web
 - Visualization is public
 - Data is public
- Tableau Prep <https://www.tableau.com/en-gb/products/prep>
 - Data wrangling and transformation in Tableau
 - Clean, re-shape and combine data for analysis
- Tableau Server
 - Managing data-driven projects

→ We will focus on Tableau Desktop

where can I get Tableau Desktop?

- Installed on all lab machines in John Honey 110
 - Start into Windows
 - Look for “Tableau Desktop”
- You can install it on your own computer
 - Option 1 – get a free student license (valid for 1 year)
 - <https://www.tableau.com/academic/students>
 - Option 2 – get a license for this module (valid until May or so)
 - Download Tableau Desktop from here: <https://www.tableau.com/tft/activation>
 - Install Tableau Desktop. When prompted enter your school email address for “business email” and “School of Computer Science” for “organization”
 - Activate with your product key: TC87-BA48-9590-AC99-A88C

some example Tableau visualizations

athletes convicted of doping

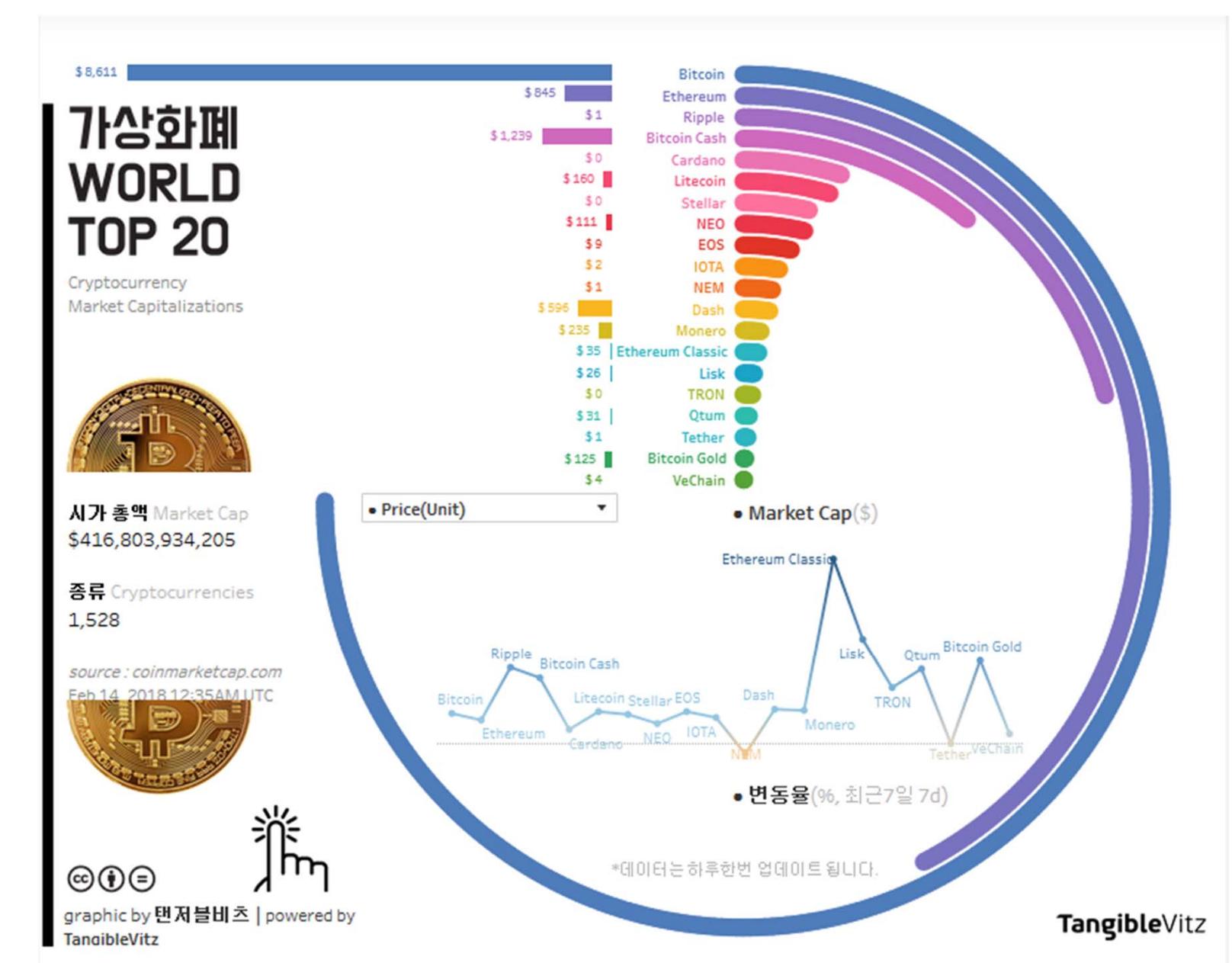
Lorenzo B.



<https://public.tableau.com/en-us/s/gallery/stripped-medals?gallery=votd>

comparison of cryptocurrencies by value/supply

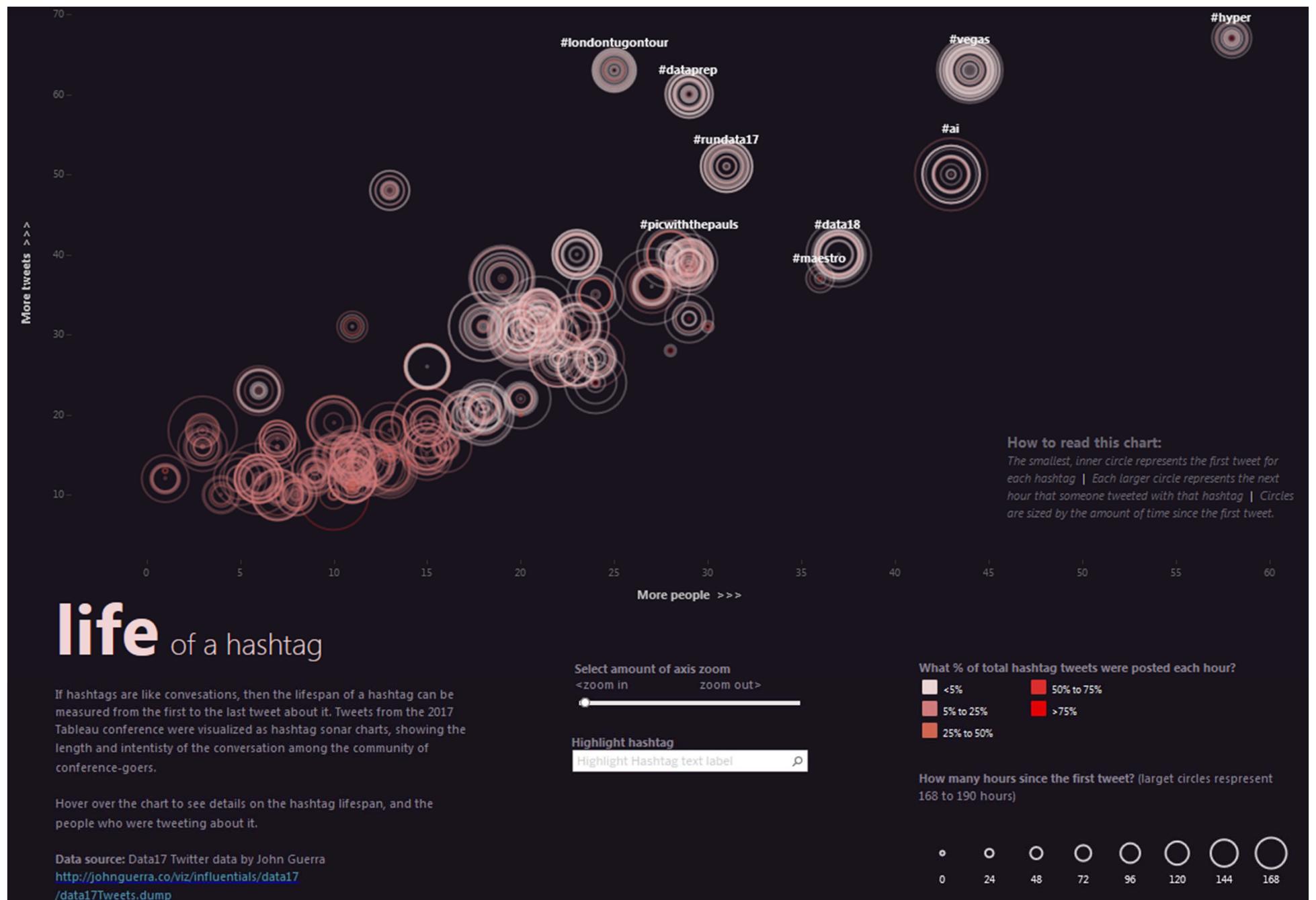
탠저블비츠 TangibleVitz



<https://public.tableau.com/en-us/s/gallery/cryptomarket?gallery=votd>

life of a hashtag

Lilach Manheim



<https://public.tableau.com/en-us/s/gallery/life-hashtag?gallery=votd>

notes on Tableau learning material

creating visualizations in Tableau

- Tableau folder on StudRes
 - CS5044/Tutorials/Tableau
- How-to instructions on:
 - Manipulating data
 - Interactive Elements
 - Visualising geospatial data

→ Ask me if you want to know specific things
- Additional resources
 - Official free tutorials by Tableau: <https://www.tableau.com/learn/training>
 - See also ResourceList

designing visualizations in Tableau – some advice

- Don't let the tool drive your visualization design
 - Come up with visualization ideas THEN try to implement them in Tableau
- Think critically about the design solutions that Tableau is offering
 - Colour schemes
 - Spatial layout
 - Labelling
 - Legends
 - ...
- Explore different possible design solutions (beyond the “Show me” panel)
- If you have an idea, and you don't know how to do it
 - Try to find solutions online
 - Ask in class

creating visualizations in Tableau

- Today: Intro to Tableau
 - Getting started
 - Creating basic interactive visualizations in Tableau
 - Follow-along demo + tutorial
- Some preparation
 - Got to [studres/CS5044/Tutorials/Tableau](#)
 - Open [Tableau_tutorial_01.pdf](#) [go to Slide 12]
 - Go to your home directory
 - Create a folder “Tableau”
 - Go to [studres/CS5044/Tutorials/Tableau/](#) and download the [data folder](#) into the Tableau folder in your home directory

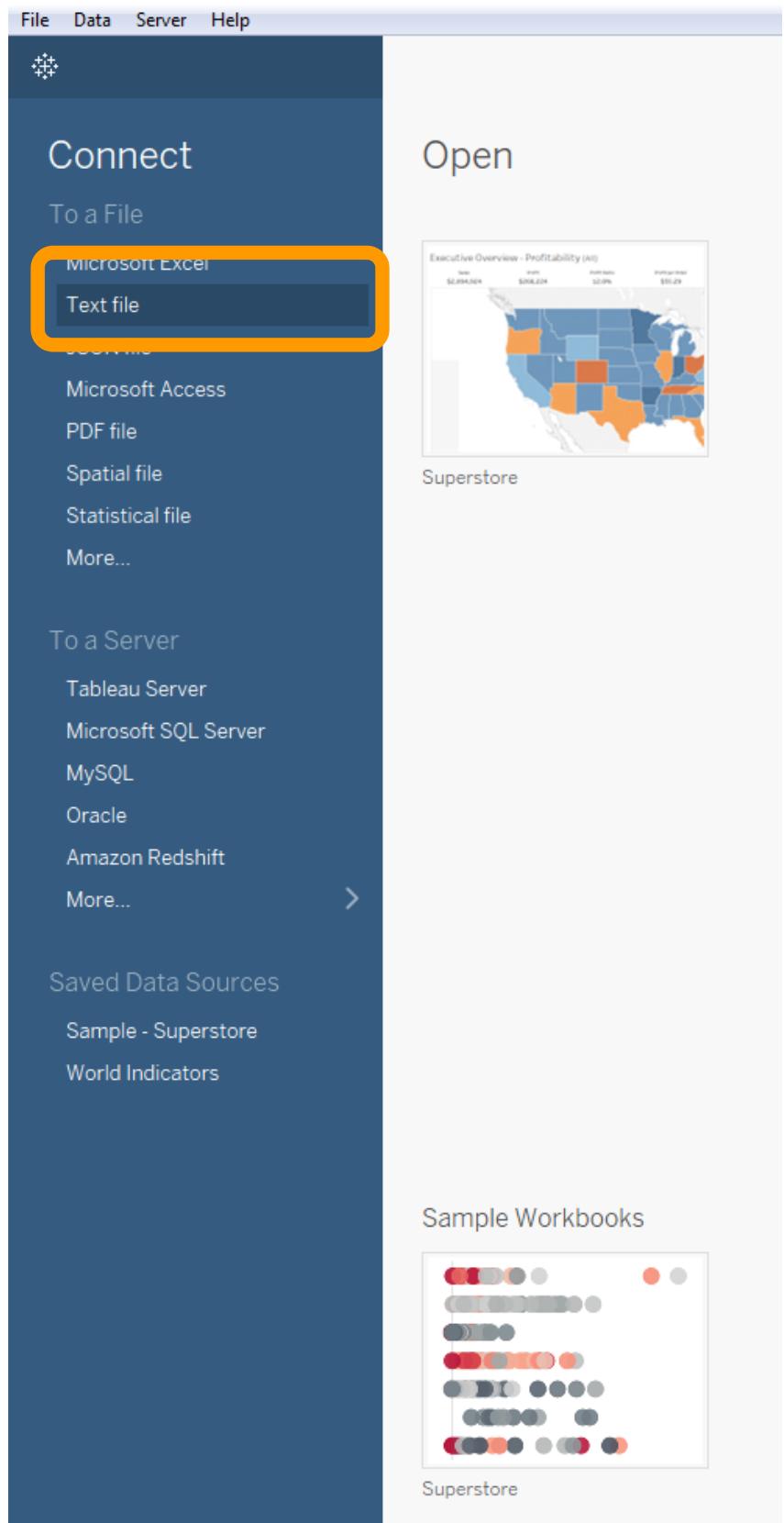
loading the data

data: [studres/CS5044/Tutorials/Tableau/data](#)

- Find [ecofootprint.csv](#) in the “data” folder
- Open Tableau
- Load this data as “Text File”

- More info about the data:

https://en.wikipedia.org/wiki/List_of_countries_by_ecological_footprint



loading the data

- You will be taken to the “Data Source” view
- The table shows all attributes in your data
- Note the icons above each column – Tableau recognizes the types of data that is being loaded
- Countries are automatically recognized as geographic data

The screenshot shows the Tableau Data Source view. In the top navigation bar, 'File', 'Data', 'Server', 'Window', and 'Help' are visible. Below the navigation is a toolbar with icons for refresh, back, forward, and search. The main area is titled 'Connections' with an 'Add' button. A connection named 'ecoFootprint' (Text file) is selected. To the right, under 'Files', there is a list of files: 'ecoFootprint.csv', 'OilData.csv', and 'WillOckenden.csv'. Below this is a 'New Union' section. The central part of the screen displays a table with the following columns and data:

Country	Ecological Footprint	Biocapacity (per person)	CO2 emission (2010)	Population
United States	5.22000	0.7000	10.4000	
Saudi Arabia	5.61000	0.5000	17.9000	
Russia	5.69000	6.7900	12.5000	
China	3.38000	0.9400	7.6000	
Canada	8.17000	16.0100	13.5000	
United Kingdom	7.93000	0.5600	7.1000	
Italy	4.61000	1.0800	5.7000	
France	5.14000	3.1100	5.0000	
Sweden	7.25000	10.6200	4.6000	
Denmark	5.51000	4.7800	6.8000	
Netherlands	5.28000	1.1700	10.1000	
Brazil	3.11000	9.0800	2.5000	
Argentina	3.14000	6.9200	4.5000	
Germany	5.30000	2.2700	9.2000	
Nigeria	1.16000	0.7000	0.6000	

loading the data

- You can also directly modify the attribute type of your data

The screenshot shows the Tableau Data Source editor interface. On the left, the 'Connections' section lists 'ecoFootprint' (Text file). Below it, the 'Files' section lists 'ecoFootprint.csv', 'OilData.csv', and 'WillOckenden.csv'. A 'New Union' button is also present. The main area displays a data preview for 'ecoFootprint.csv' with columns: 'Country/Region', 'Actual Footprint...', 'Biocapacity (per person)', 'CO2 emission (2005)', and 'Population'. A context menu is open over the 'Country/Region' column header, specifically over the first row ('United Kingdom'). This menu is titled 'Geographic Role' and includes options: 'None', 'Airport', 'Area Code (U.S.)', 'CBSA/MSA (U.S.)', 'City', 'Congressional District (U.S.)', 'County', 'NUTS Europe', 'State/Province', and 'ZIP Code/Postcode'. The 'Country/Region' option is selected. A large orange circle highlights this entire context menu. A smaller orange box highlights the 'Country/Region' option in the list. At the bottom of the preview, there's a 'Go to Worksheet' button.

Country/Region	Actual Footprint...	Biocapacity (per person)	CO2 emission (2005)	Population
United Kingdom	8.22000	3.7600	16.4000	1,400,000
Italy	5.61000	0.5000	17.9000	55,000
France	5.69000	6.7900	12.5000	65,000
Sweden	3.38000	0.9400	7.6000	9,000
Denmark	5.47000	1.0100	13.5000	5,000
Netherlands	5.56000	0.5600	7.1000	15,000
Brazil	1.08000	1.0800	5.7000	1,000
Argentina	3.11000	0.6200	4.6000	1,000
Germany	4.78000	1.1700	10.1000	6,000
Spain	4.78000	1.1700	10.1000	6,000
Portugal	3.11000	0.0800	2.5000	1,000
Chile	5.92000	0.9200	4.5000	1,000
China	2.27000	5.30000	9.2000	1,000
Nigeria	0.7000	1.16000	0.6000	1,000
Malta	1.15000	2.21000	0.2000	1,000

Tableau worksheets

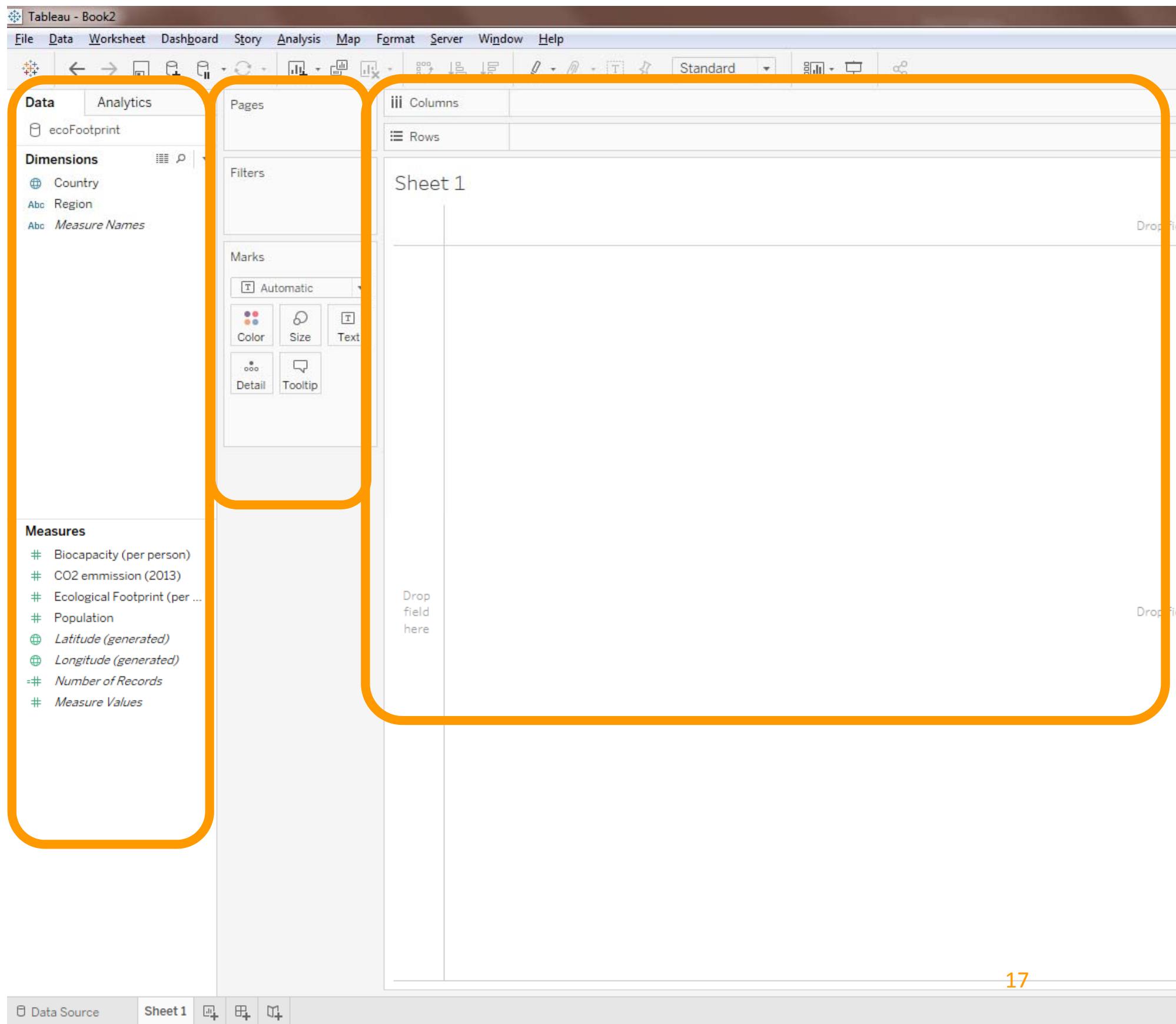
- Navigate to “Sheet1” to open your first worksheet where all the visualization actions will take place.

The screenshot shows the Tableau software interface. The top menu bar includes File, Data, Server, Window, and Help. The left sidebar displays 'Connections' with 'ecoFootprint' selected, and 'Files' containing 'ecoFootprint.csv', 'OilData.csv', and 'WillOckenden.csv'. A checkbox for 'Use Data Interpreter' is checked, with a note explaining it might clean the Text file workbook. Below the files is a 'New Union' option. The main area shows a data grid with columns: Country, Ecological Footpri..., Biocapacity (per p...), CO2 emission (2...), and Populat. The first few rows of data are: United States, 8.22000, 3.7600, 16.4000; Saudi Arabia, 5.61000, 0.5000, 17.9000; Russia, 5.69000, 6.7900, 12.5000. At the bottom, there's a 'Data Source' tab and several icons. A red box highlights a 'Go to Worksheet' button with a bar chart icon, which is part of a tooltip. The 'Sheet1' tab is also highlighted in red.

Country	Ecological Footpri...	Biocapacity (per p...)	CO2 emission (2...)	Populat
United States	8.22000	3.7600	16.4000	
Saudi Arabia	5.61000	0.5000	17.9000	
Russia	5.69000	6.7900	12.5000	
China	3.38000	0.9400	7.6000	
Canada	8.17000	16.0100	13.5000	
United Kingdom	7.93000	0.5600	7.1000	
Italy	4.61000	1.0800	5.7000	
France	5.14000	3.1100	5.0000	
Sweden	7.25000	10.6200	4.6000	
Denmark	5.51000	4.7800	6.8000	
Netherlands	5.28000	1.1700	10.1000	
Brazil	3.11000	9.0800	2.5000	
Argentina	3.14000	6.9200	4.5000	
Germany	5.30000	2.2700	9.2000	
Nigeria	1.16000	0.7000	0.6000	
Costa Rica	2.21000	1.1500	0.2000	

Tableau worksheets

- Left: list of data attributes
 - Dimensions → categorial data
 - Measures → quantitative data
- Middle: choices for marks and visual variables
- Right: visualization space that is structured into rows and columns



visualizing countries

- Drag the “country” attribute into the “Rows” pane
- All values for “country” will be distributed by row in the resulting view

The screenshot shows the Tableau interface with the 'Rows' pane highlighted by a large orange rectangle. An orange arrow points from the 'Rows' pane to the resulting data view on the right. The 'Dimensions' pane also has a small orange box around the 'Country' item.

Sheet 1

Country	
Argentina	Abc
Brazil	Abc
Canada	Abc
China	Abc
Denmark	Abc
France	Abc
Germany	Abc
Italy	Abc
Netherlands	Abc
Nigeria	Abc
Russia	Abc
Saudi Arabia	Abc
South Africa	Abc
Sweden	Abc
United Kingdom	Abc
United States	Abc

Measures

- # Biocapacity (per person)
- # CO2 emission (2013)
- # Ecological Footprint (per ...)
- # Population
- # Latitude (generated)
- # Longitude (generated)
- # Number of Records
- # Measure Values

choosing marks

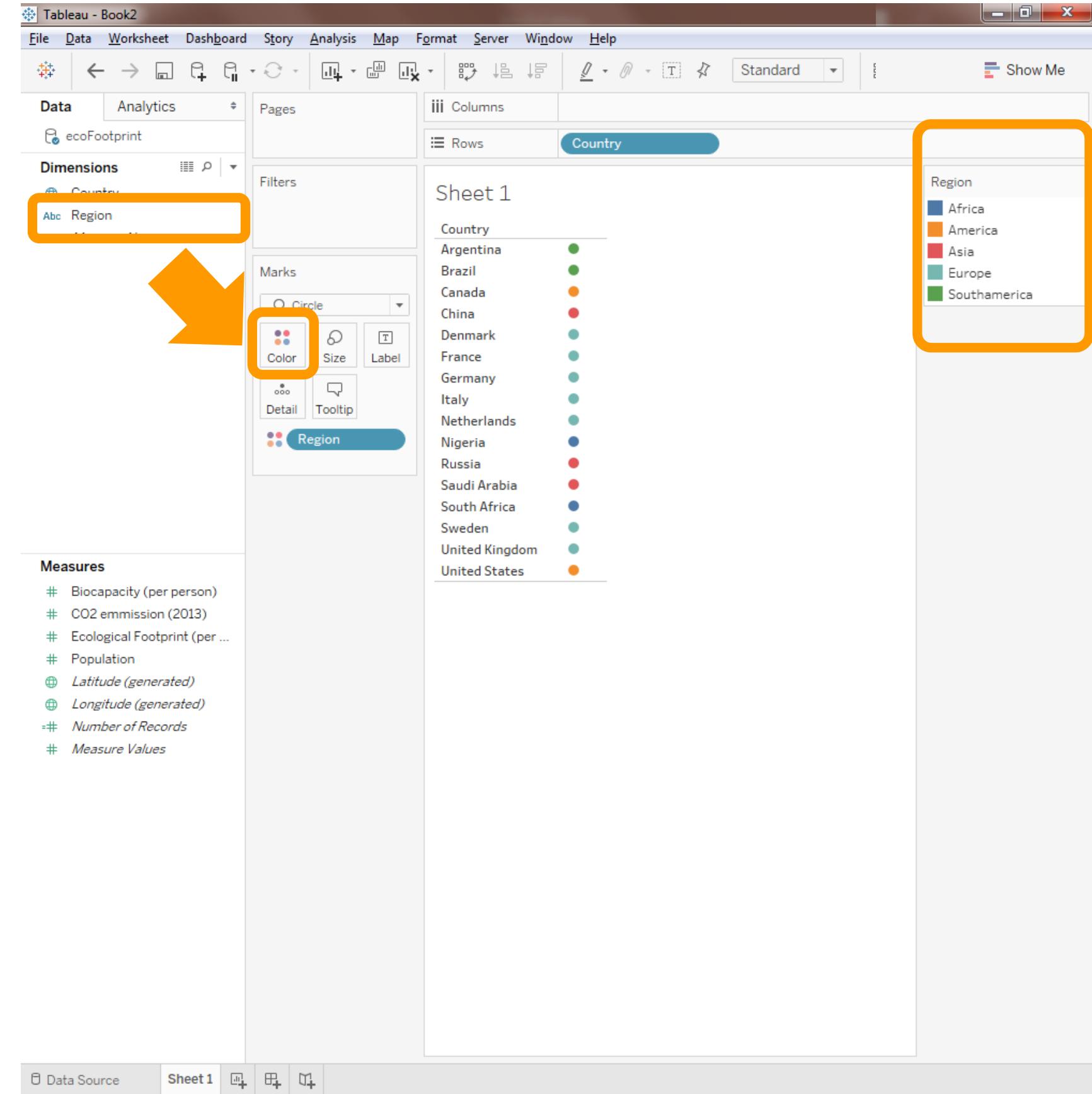
- You can choose the type of marks to represent countries

The screenshot shows the Tableau interface with the following details:

- Top Bar:** Tableau - Book2, File, Data, Worksheet, Dashboard, Story, Analysis, Map, Format, Server, Window, Help.
- Data Source:** ecoFootprint
- Dimensions:** Country, Region, Measure Names
- Marks Shelf (highlighted with an orange box):** Circle (selected), Bar, Line, Area, Square, Shape, Text, Map, Pie, Gantt Bar, Polygon.
- Sheet 1 (Country dimension):** A list of countries with corresponding blue dots representing the chosen mark type.
- Bottom Bar:** Data Source, Sheet 1, and other standard Tableau navigation icons.

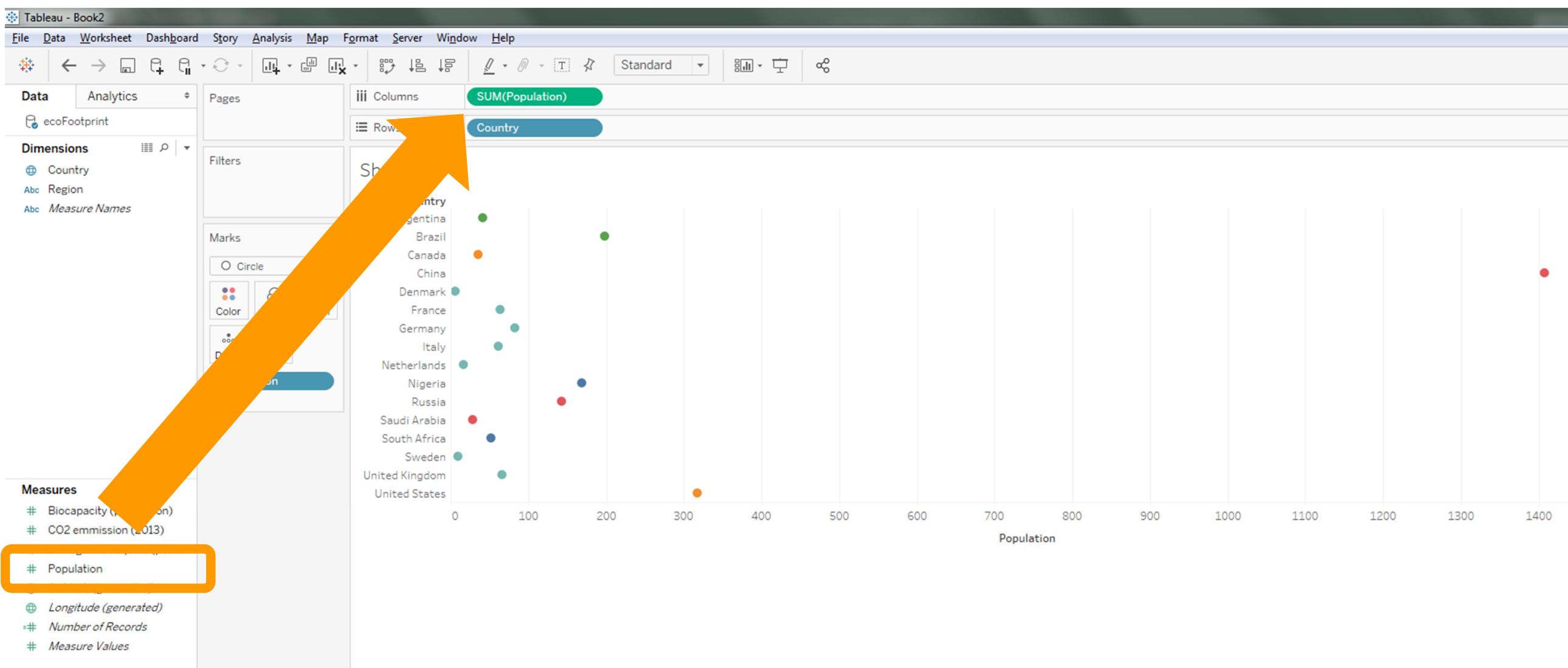
choosing visual variables

- You can also choose visual variables to change the characteristics of marks according to the data.
- Here we choose colour (hue) to represent the countries by region.



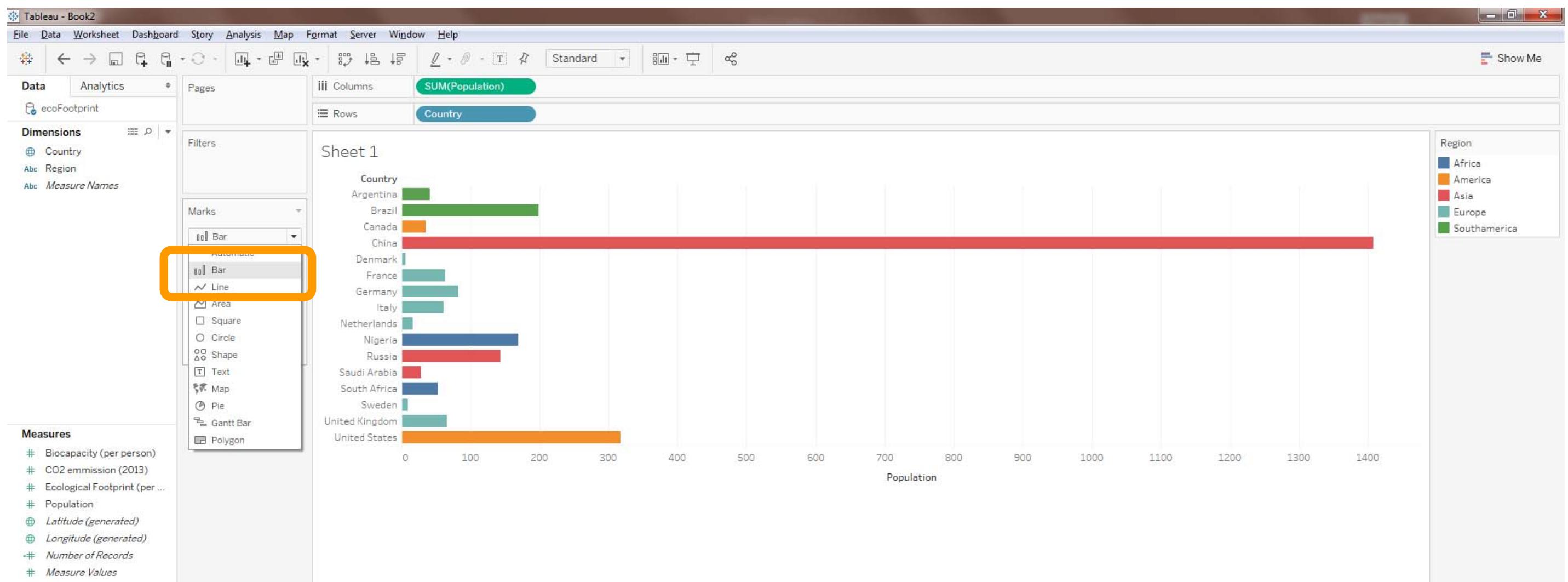
country by population

- We can now bring in an additional attribute to show the population per country. Drag the population attribute from the “measures” into the “columns” pane



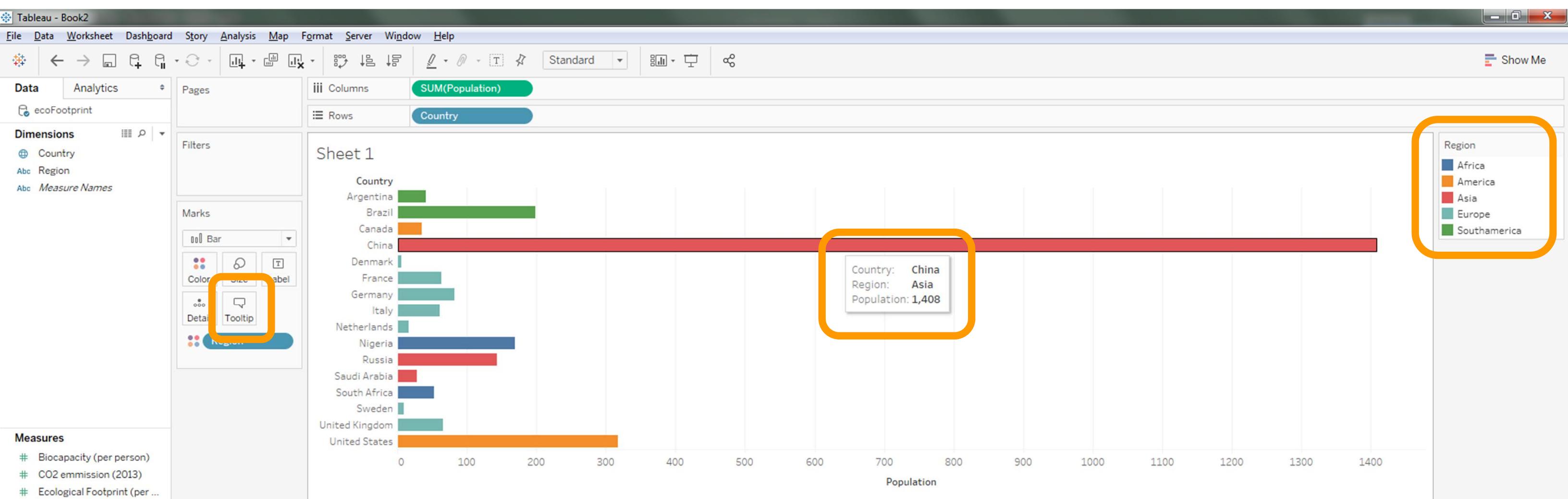
country by population

- Not happy with the visualization? Try a different type of marks.



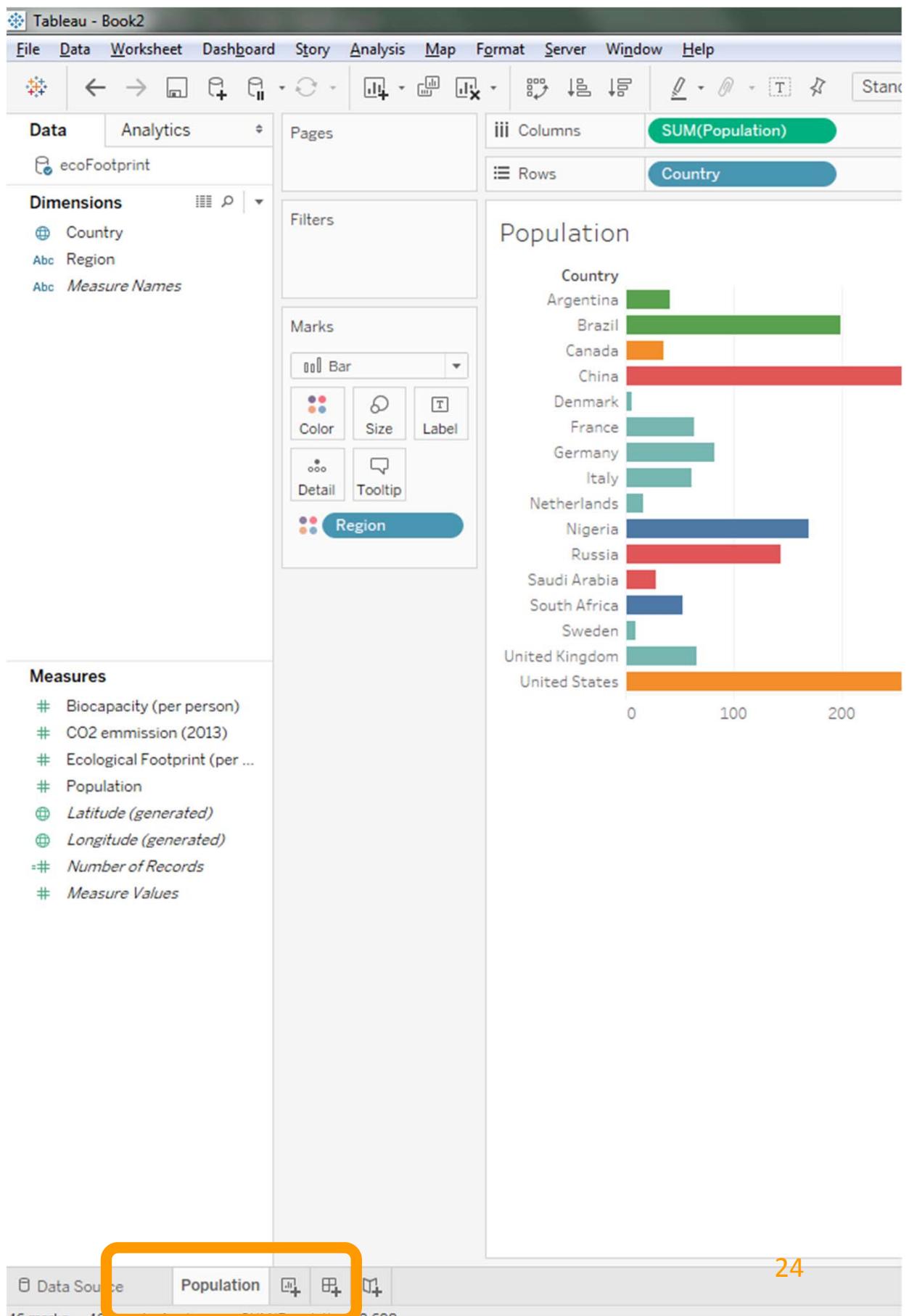
tooltips

- Interactive tooltips and legends are provided automatically, but you can customize them via the tooltip button in the “marks” pane or by double-clicking the legend.



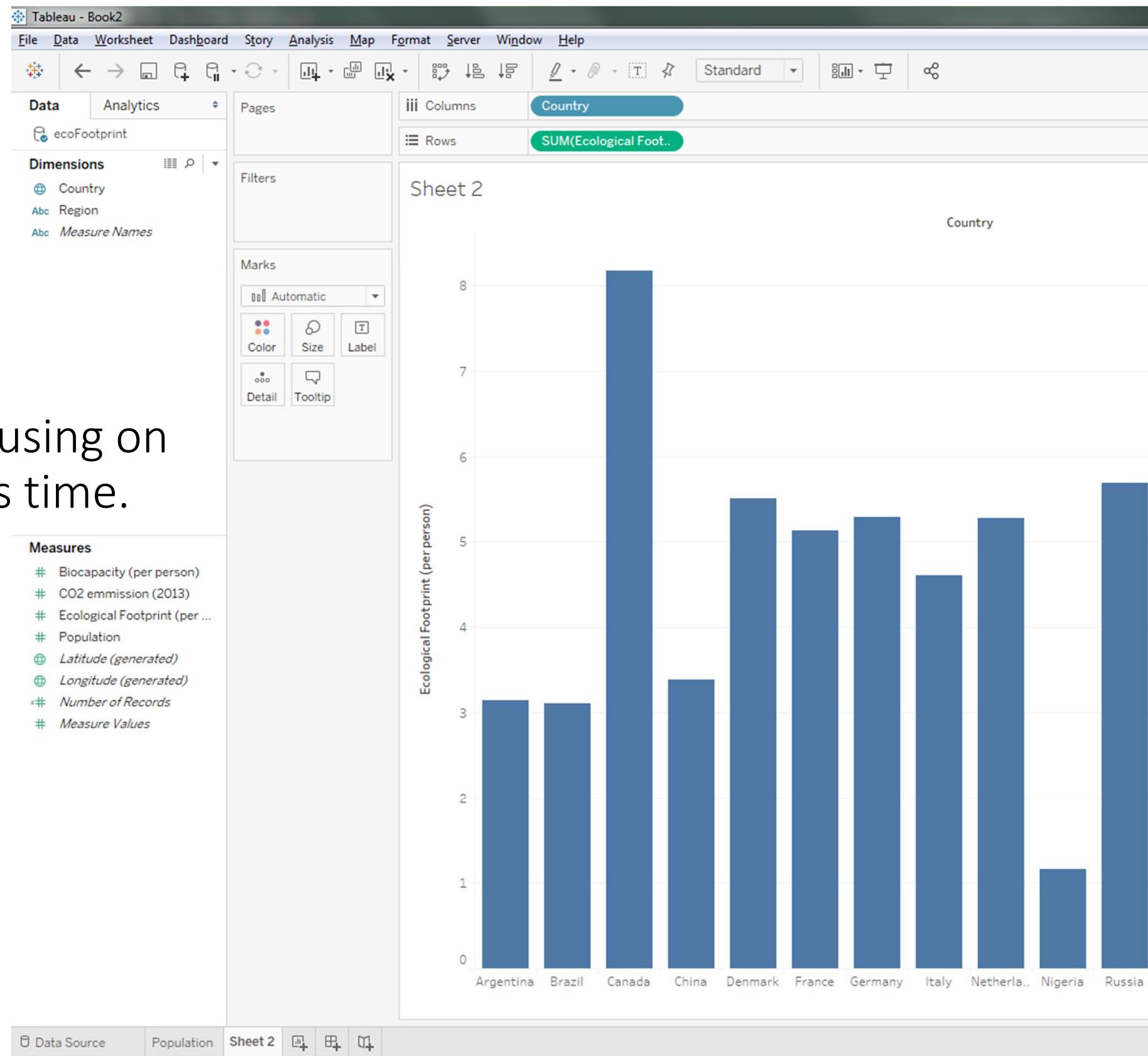
sheet labeling

- Always label your sheets, so you can easily distinguish them.
- Double-click on the tab of your sheet to rename it.



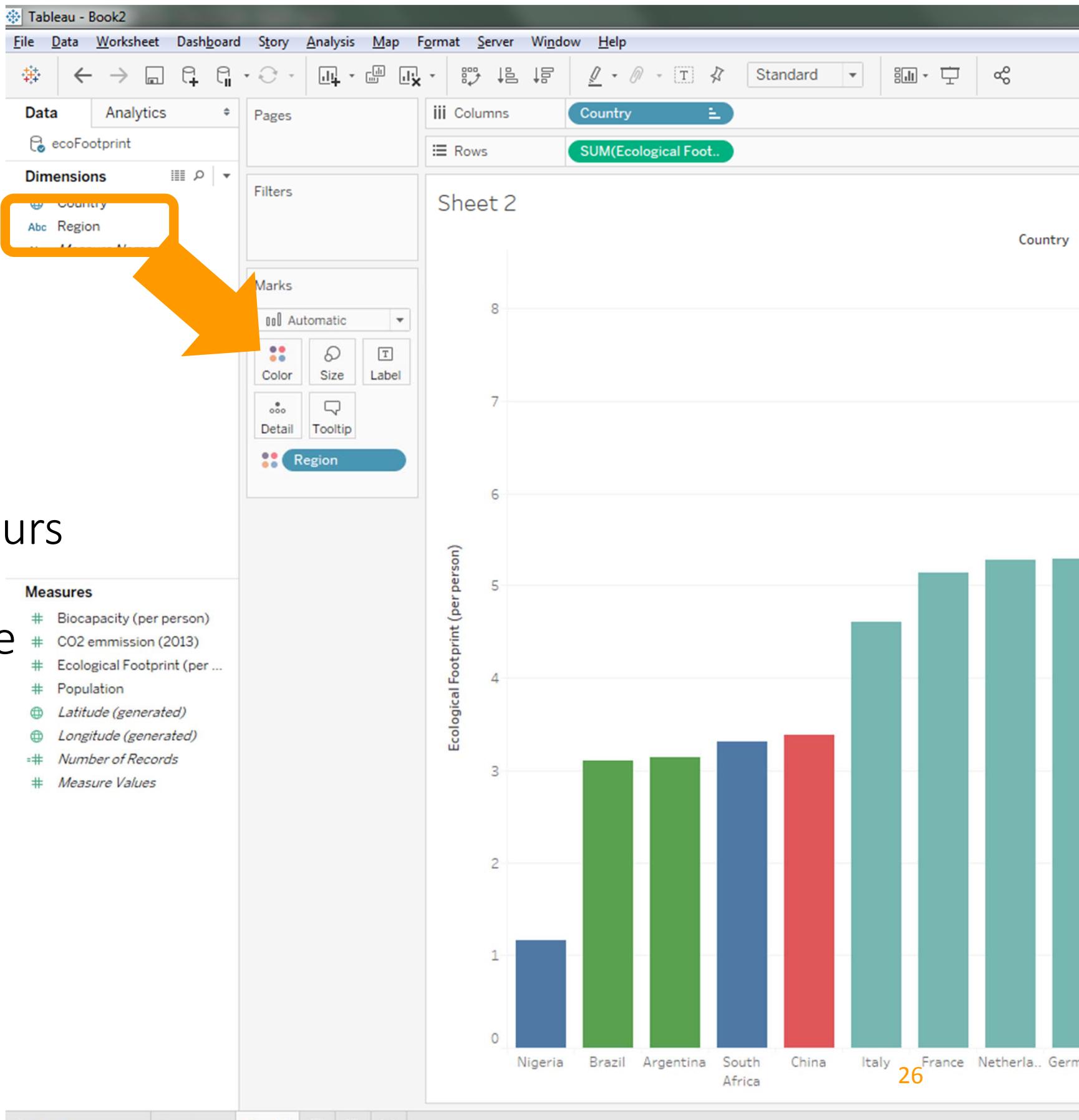
country by eco footprint

- Let's create a new sheet for a new visualization
- Let's create a similar bar chart, focusing on country vs. ecological footprint this time.



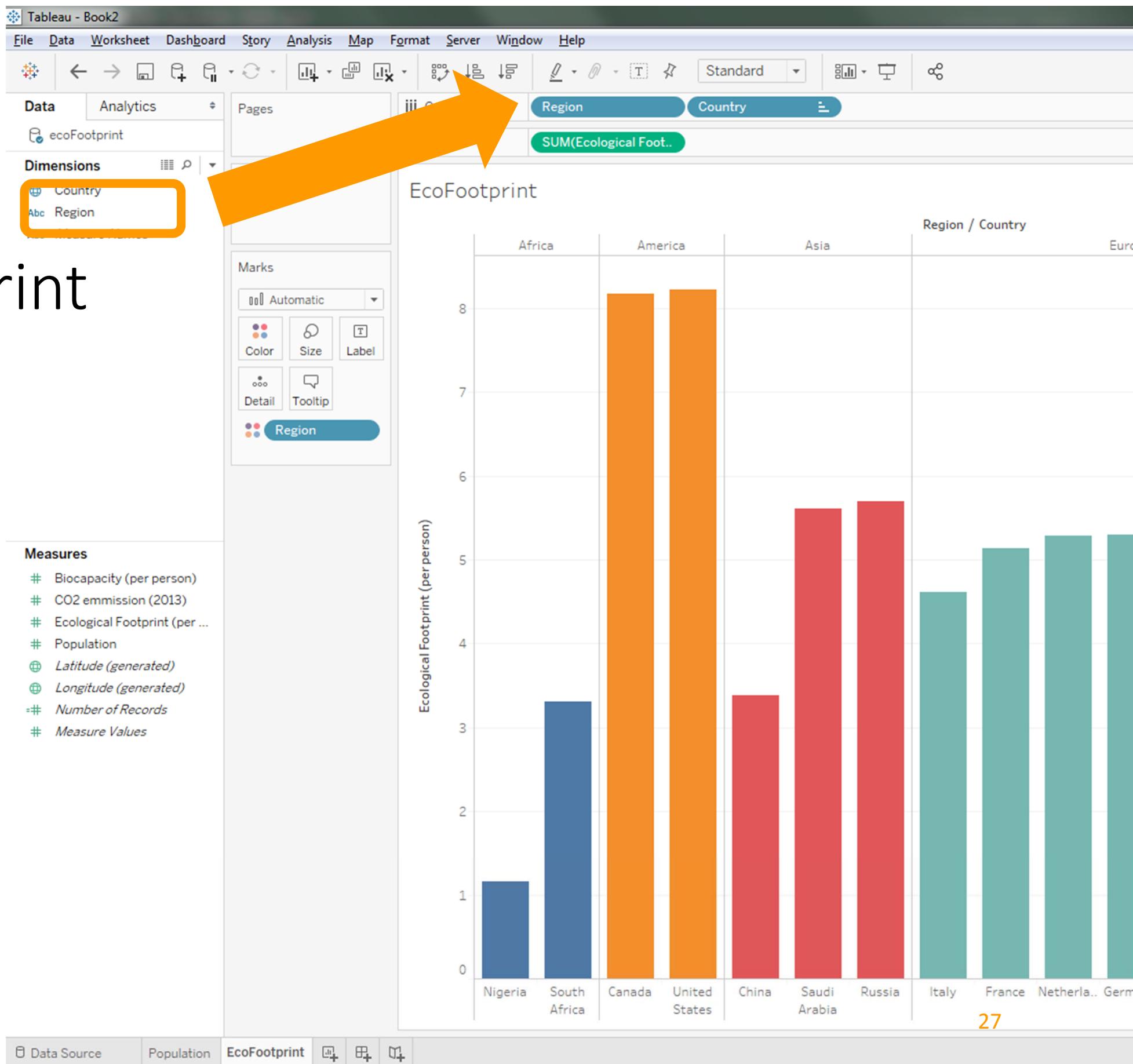
country by eco footprint

- Again, we choose colour (hue) to distinguish between regions.
- Note that even if you change the colours of individual regions (double-click the legend), the colour scheme will still be consistent across the different sheets
→ Currently population and eco footprint



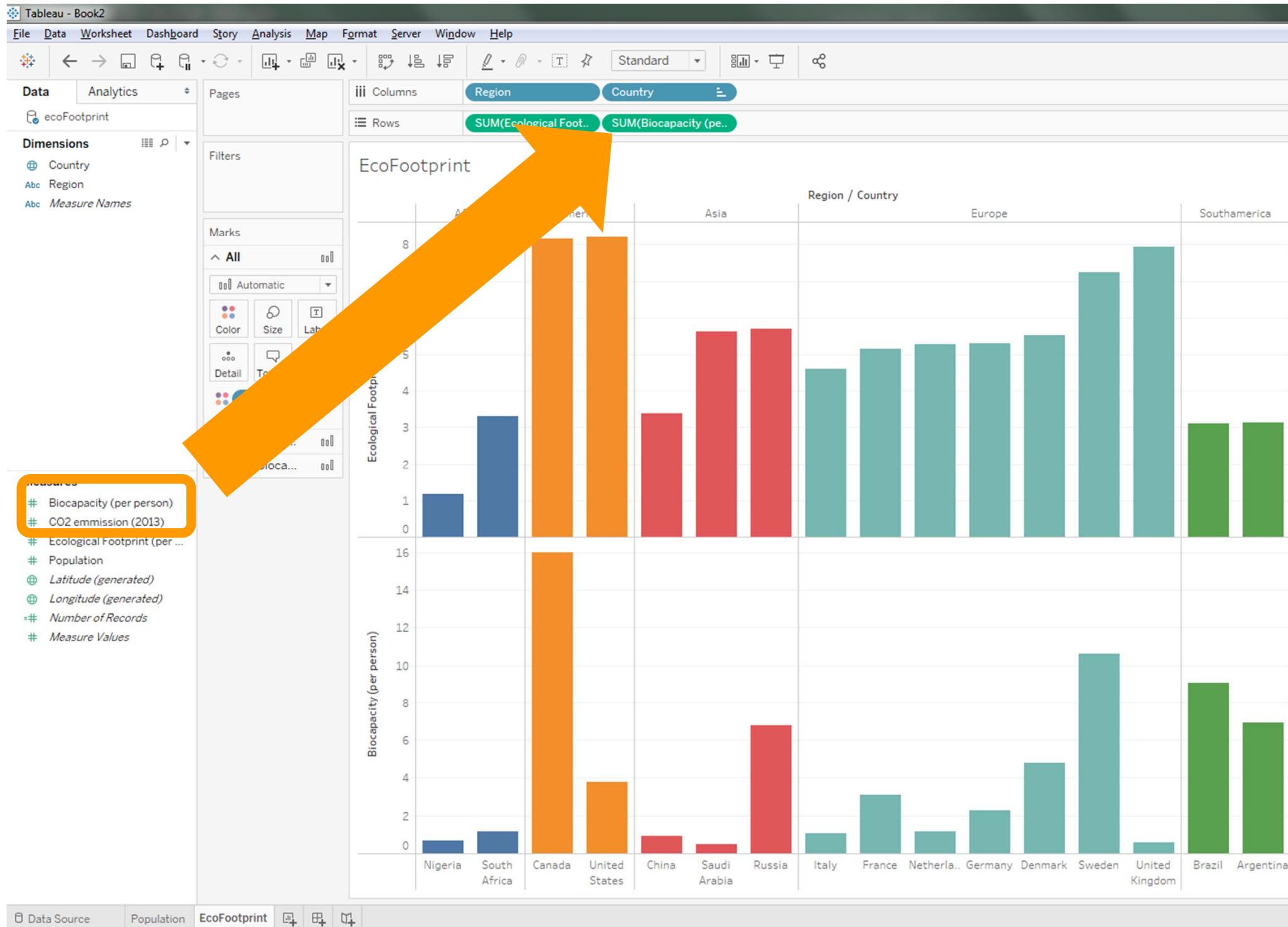
country by eco footprint

- Ordering by region



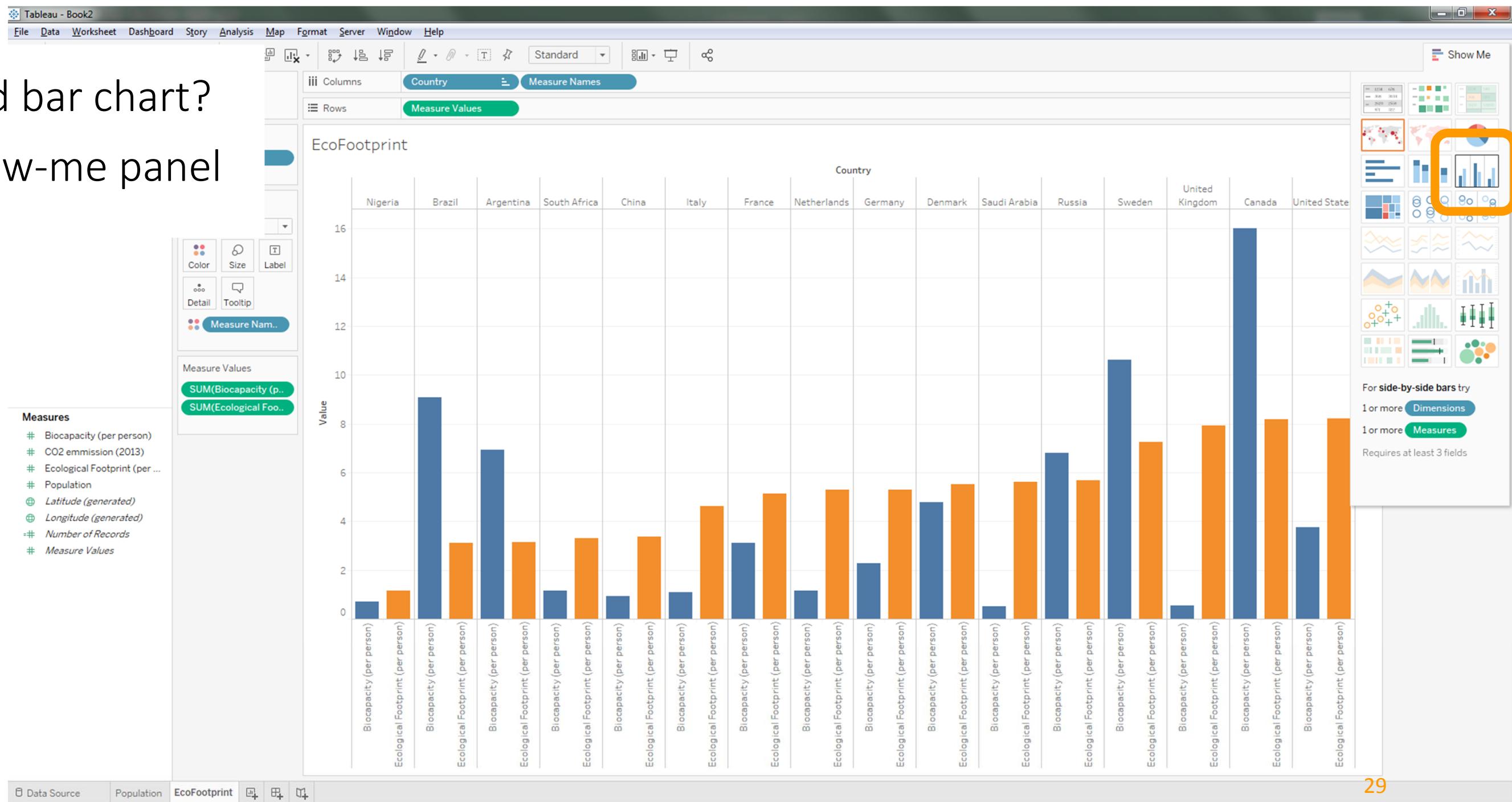
eco footprint vs. biocapacity

- Two bar graphs?



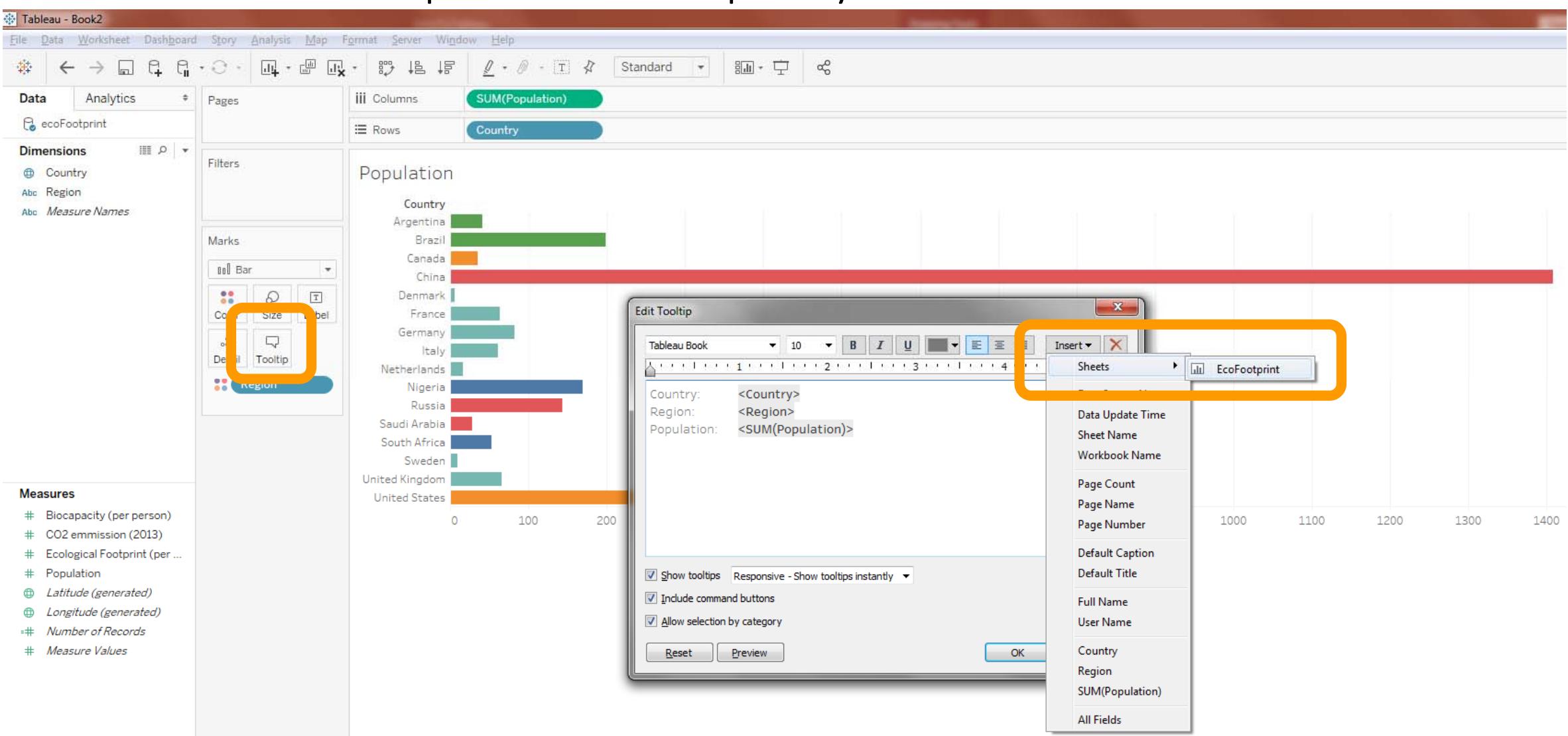
eco footprint vs. biocapacity

- Grouped bar chart?
- Via show-me panel



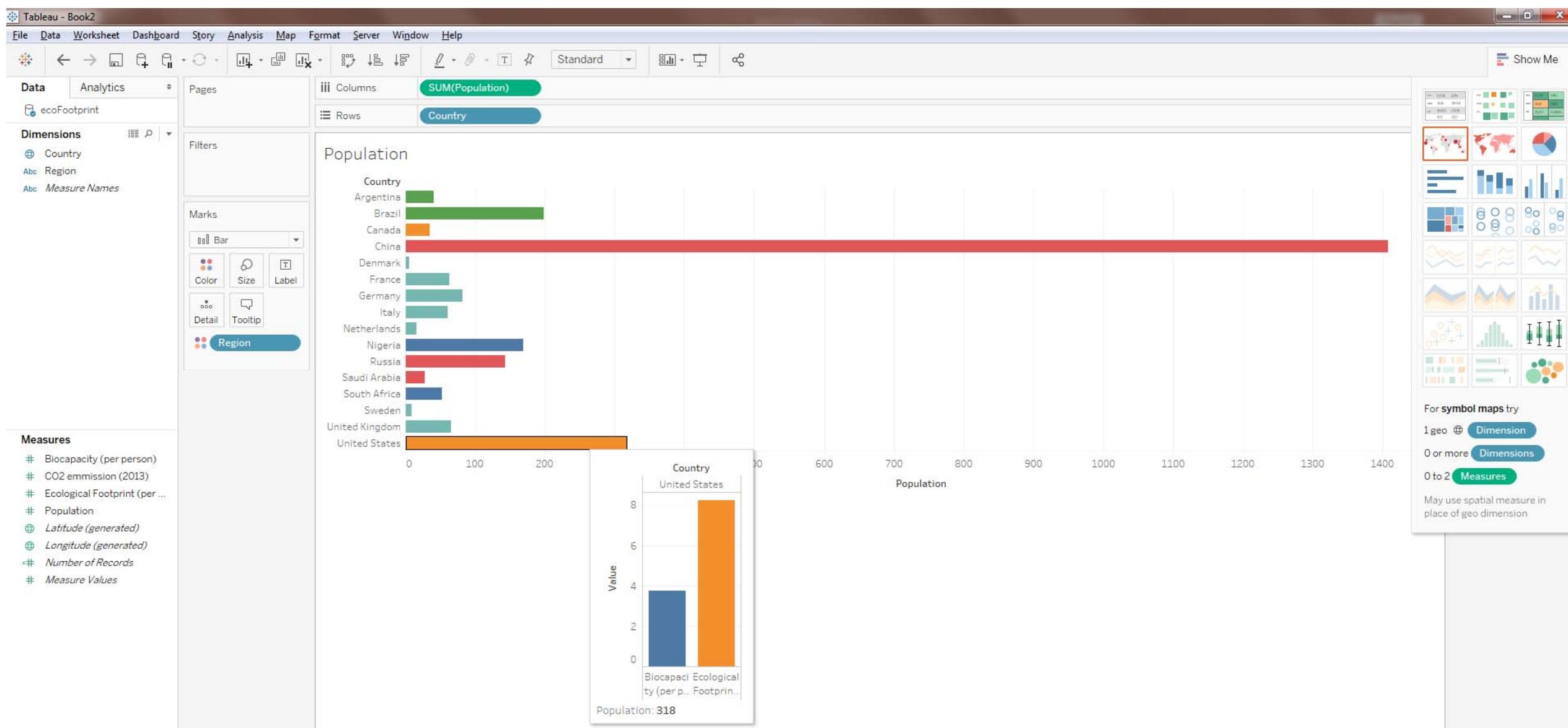
updating tooltips in “Population” visualization

- Go back to the “Population” worksheet and update the tooltip to integrate a mini tooltip visualization of eco footprint vs. biocapacity.



updating tooltips in “Population” visualization

- Make sure to avoid inconsistencies in colouring (here, orange and blue are used for different attributes, which is not good).



biocapacity “deficit/surplus”

- In the previous bar chart we can see that some countries clearly have less biocapacity (by person) compared to their eco footprint (by person).
- Let's highlight this more by calculating a country's biocapacity deficit/surplus
- Go back to the “Data Source” view and let's create a calculated field that shows the biocapacity deficit
 - biocapacity – ecological footprint

The screenshot shows the Tableau Data Source view for the 'ecoFootprint' data source. The 'Data Source' tab is selected at the bottom. A context menu is open over the fourth column header 'Biocapacity (per person)'. The 'Create Calculated Field...' option is highlighted with an orange arrow and a yellow border.

Country	Ecological Footprint (per person)	Biocapacity (per person)	
United States	8.22000	3.76	
Saudi Arabia	5.61000	0.50	
Russia	5.69000	6.79	
China	3.38000	0.94	
Canada	8.17000	16.01	
United Kingdom	7.93000	0.56	
Italy	4.61000	1.0800	5.7000
France	5.14000	3.1100	5.0000
Sweden	7.25000	10.6200	4.6000
Denmark	5.51000	4.7800	6.8000
Netherlands	5.28000	1.1700	10.1000
Brazil	3.11000	9.0800	2.5000
Argentina	3.14000	6.9200	4.5000
Germany	5.30000	2.2700	9.2000
Nigeria	1.16000	0.7000	0.6000

biocapacity “deficit/surplus”

- Creating a calculated field
 - [Biocapacity (per person)] - [Ecological Footprint (per person)]

The screenshot shows a Tableau interface with a calculated field dialog box highlighted by an orange circle. The dialog box is titled "Calculation1" and contains the formula: "[Biocapacity (per person)] - [Ecological Footprint (per person)]". The calculated field is being applied to the "Country" dimension. A tooltip for the calculated field shows the formula and the source fields: "EcoFootprint" and "Number of Records". A message at the bottom of the dialog box states "The calculation contains errors".

Country	Ecological Footprint (per person)	Biocapacity (per person)
United States	8.22000	3.7600
Saudi Arabia	5.61000	0.5000
Russia	5.69000	6.7900
China	3.38000	0.9400
Canada	8.17000	16.0100
United Kingdom	7.93000	0.5600
Italy	4.61000	1.0800
France	5.14000	3.1100
Sweden	7.25000	10.6200
Denmark	5.51000	4.7800
Netherlands	5.28000	1.1700
Brazil	3.11000	9.0800
Argentina	3.14000	6.9200
Germany	5.30000	2.2700
Nigeria	1.16000	0.7000
China	7.78000	9.4700
United States	8.22000	3.7600
America	13.5000	34.84
Europe	7.1000	65.65
Southamerica	2.5000	198.66
Africa	0.6000	168.83

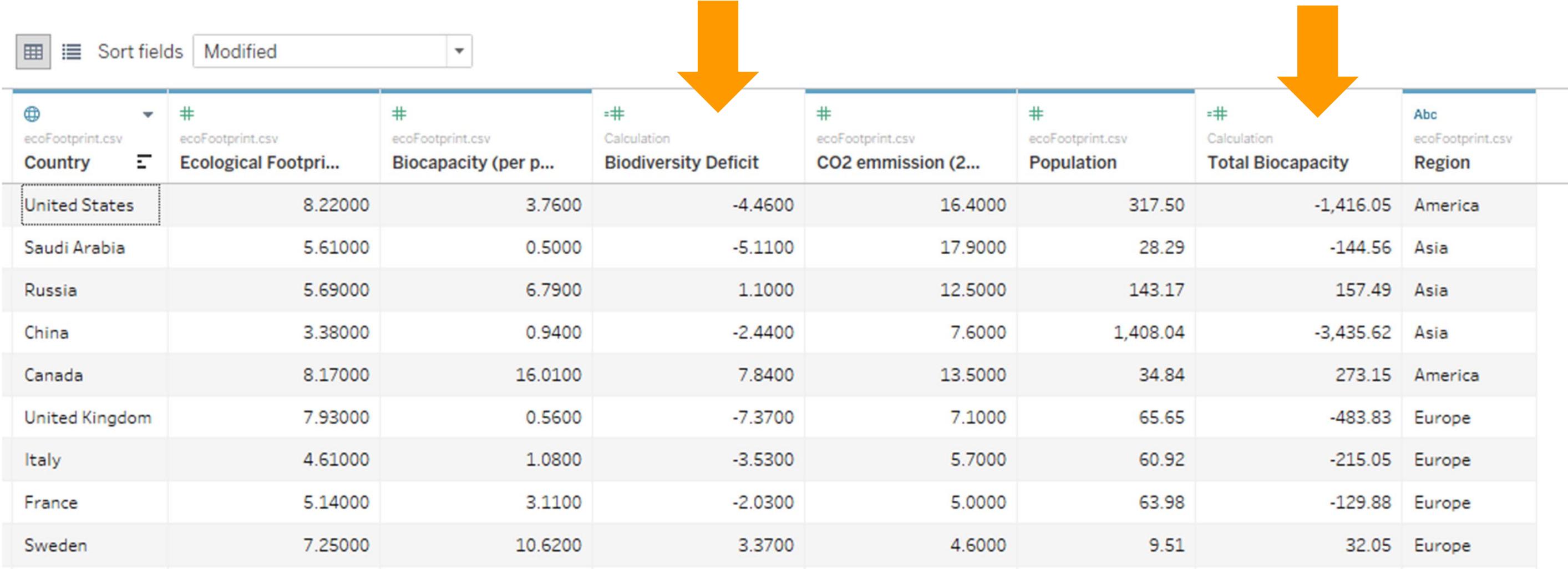
considering population size

- Create another calculated field in the same way for “total biocapacity deficit/surplus”
 - $[Biocapacity\ deficit/surplus] * [Population]$

The screenshot shows a Tableau interface with a data source named "WillOckenden.csv" containing a table "New Union". The table has columns: Country, Ecological Footprint (per person), and Biocapacity (per person). A calculated field dialog box is open, titled "Total Biocapacity", with the formula $[Population] * [Biodiversity Deficit]$. The dialog includes an "OK" button and a message "The calculation is valid." A yellow oval highlights the formula input field and the "OK" button.

Country	Ecological Footprint (per person)	Biocapacity (per person)
United States	8.22000	-0.7600
Saudi Arabia	5.61000	-0.5000
Russia	5.69000	-0.7900
China	3.38000	-0.3400
Canada	8.17000	16.0100
United Kingdom	7.93000	0.5600
Italy	4.61000	1.0800
France	5.14000	3.1100
Sweden	7.25000	10.6200
Denmark	5.51000	4.7800
Netherlands	5.28000	1.1700
Brazil	3.11000	9.0800
Argentina	3.14000	6.9200
Germany	5.30000	2.2700
Nigeria	1.16000	0.7000
South Africa	2.21000	-1.1500

the new data table



The screenshot shows a Microsoft Excel spreadsheet with a data table. The table has the following columns:

- Country
- Ecological Footprint (per person)
- Biocapacity (per person)
- Biodiversity Deficit
- CO2 emission (2010)
- Population
- Total Biocapacity
- Region

The data rows include:

Country	Ecological Footprint (per person)	Biocapacity (per person)	Biodiversity Deficit	CO2 emission (2010)	Population	Total Biocapacity	Region
United States	8.22000	3.7600	-4.4600	16.4000	317.50	-1,416.05	America
Saudi Arabia	5.61000	0.5000	-5.1100	17.9000	28.29	-144.56	Asia
Russia	5.69000	6.7900	1.1000	12.5000	143.17	157.49	Asia
China	3.38000	0.9400	-2.4400	7.6000	1,408.04	-3,435.62	Asia
Canada	8.17000	16.0100	7.8400	13.5000	34.84	273.15	America
United Kingdom	7.93000	0.5600	-7.3700	7.1000	65.65	-483.83	Europe
Italy	4.61000	1.0800	-3.5300	5.7000	60.92	-215.05	Europe
France	5.14000	3.1100	-2.0300	5.0000	63.98	-129.88	Europe
Sweden	7.25000	10.6200	3.3700	4.6000	9.51	32.05	Europe

showing total biocapacity surplus/deficit

- Create a new worksheet
- Now we can, again, create a simple bar chart by mapping “total biocapacity surplus/deficit” by country.



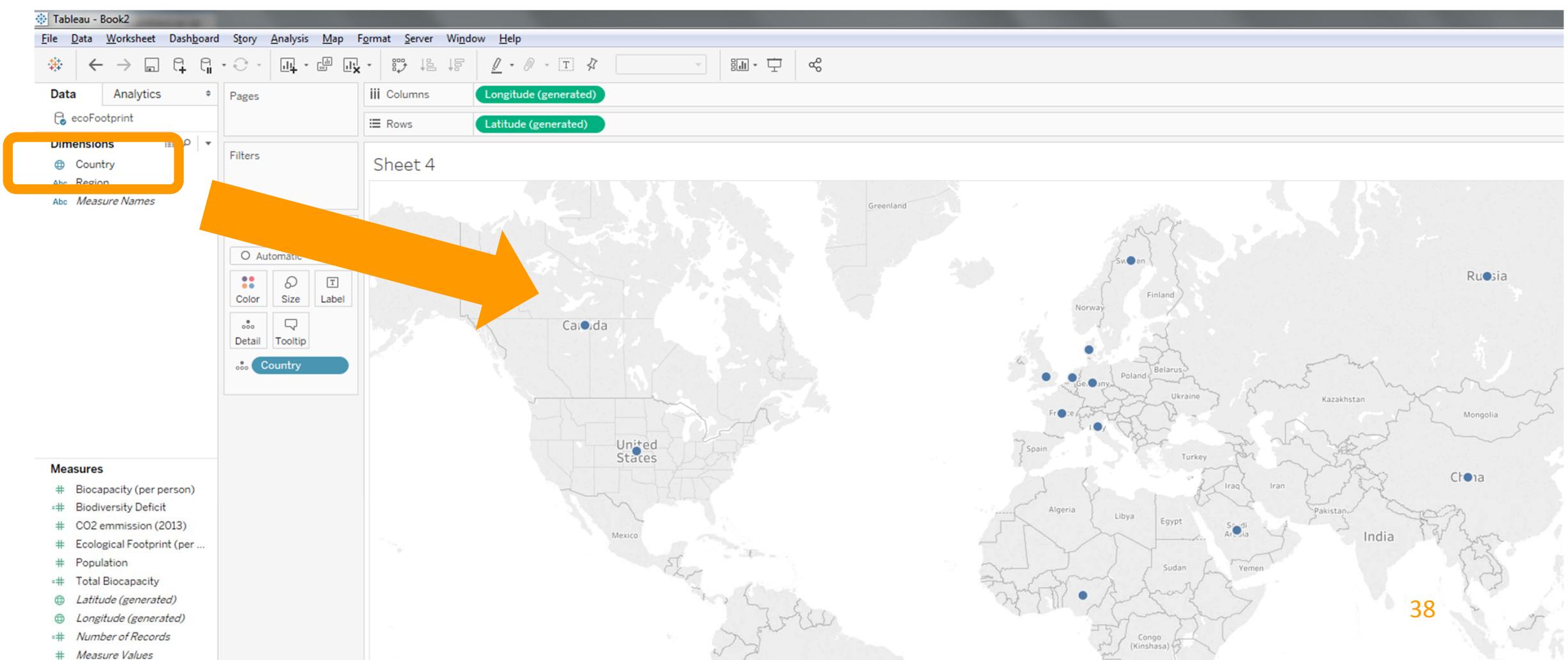
showing total biocapacity surplus/deficit



- Now map “biocapacity surplus/deficit” to colour
- A diverging colour scale further highlights surplus and deficit.

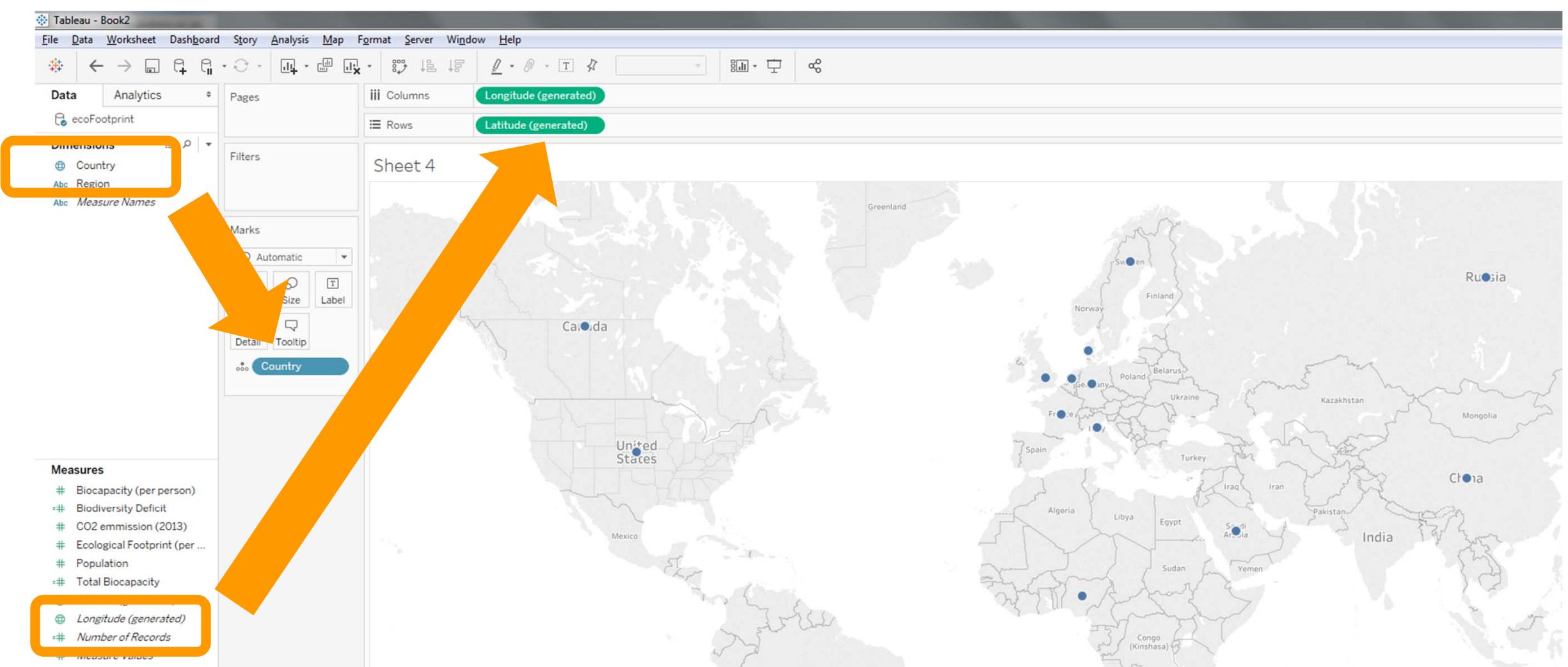
visualizing the data on a geographic map

- We create a geographic map, showing countries and their population
- Create a new worksheet
- Simply drag “Country” – geographic dimension – into the canvas.



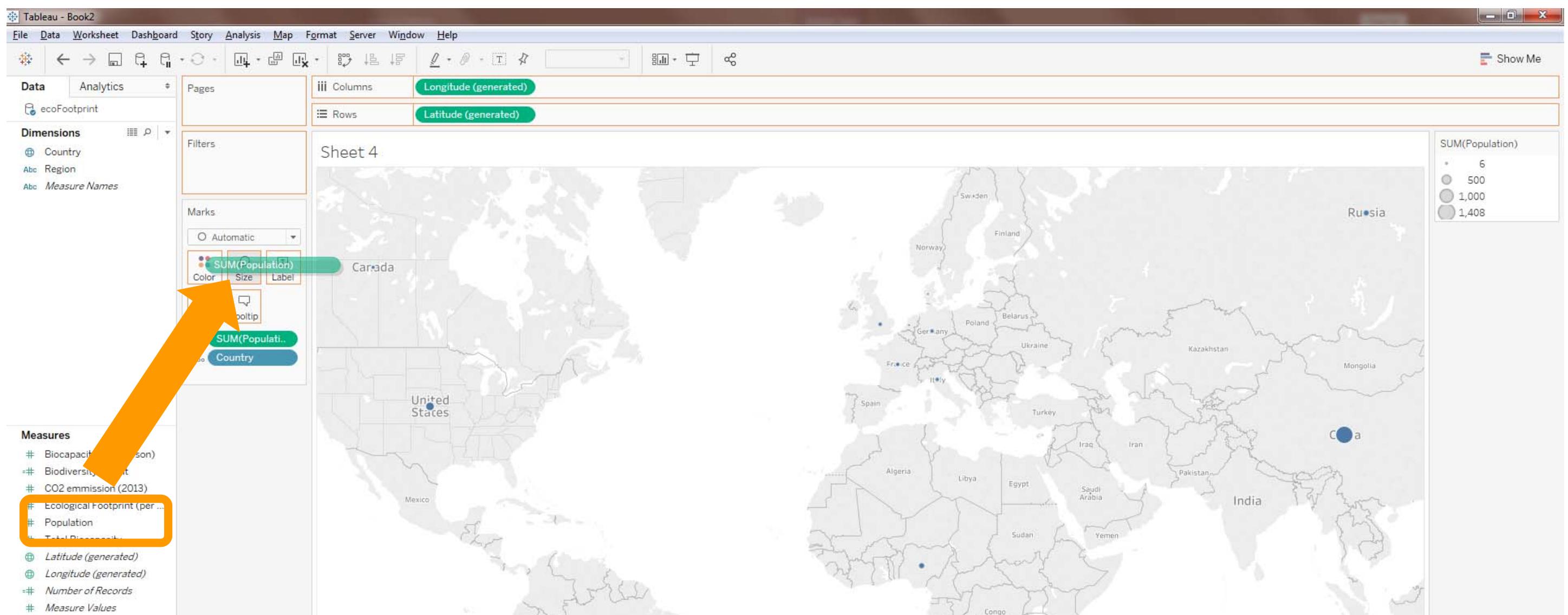
geographic map

- You can also work directly with the latitude and longitude measures, which are automatically generated by Tableau.
- Drag “Country” into the “Marks” pane to show countries as circles.



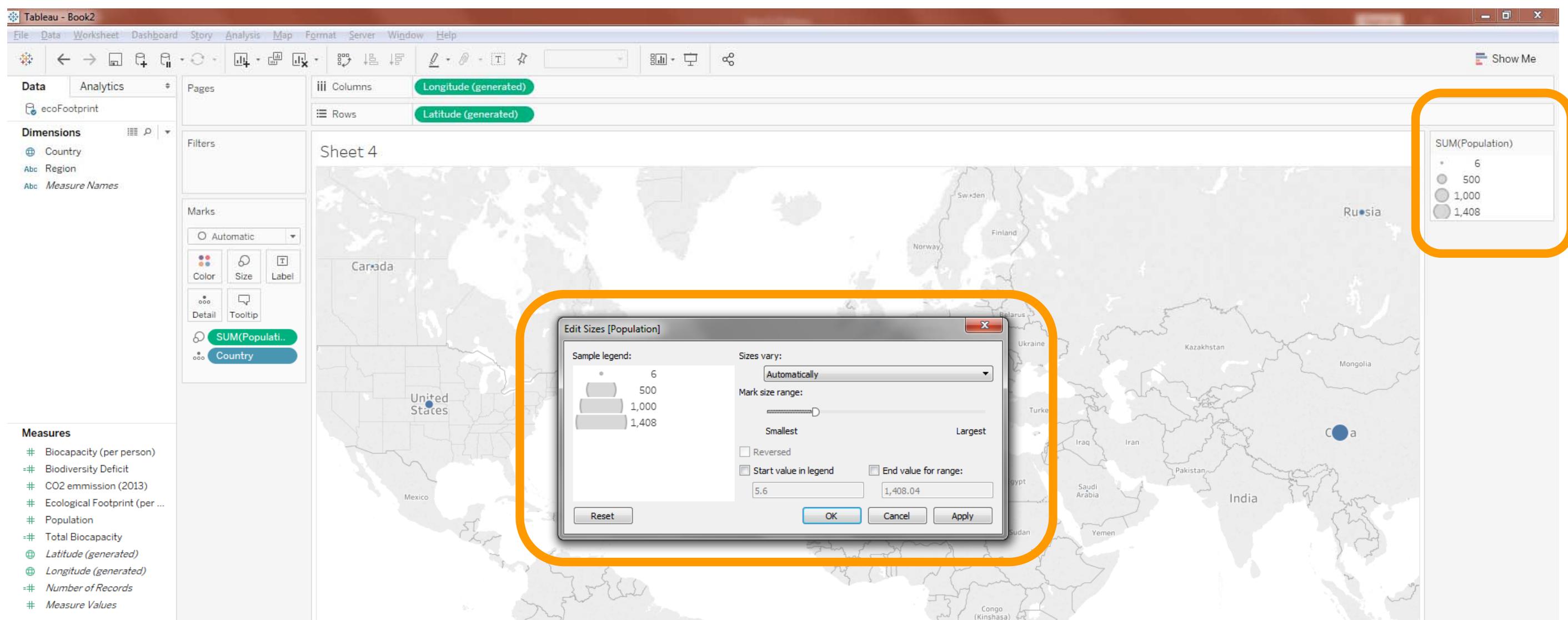
geographic map

- Now show population as circle size.



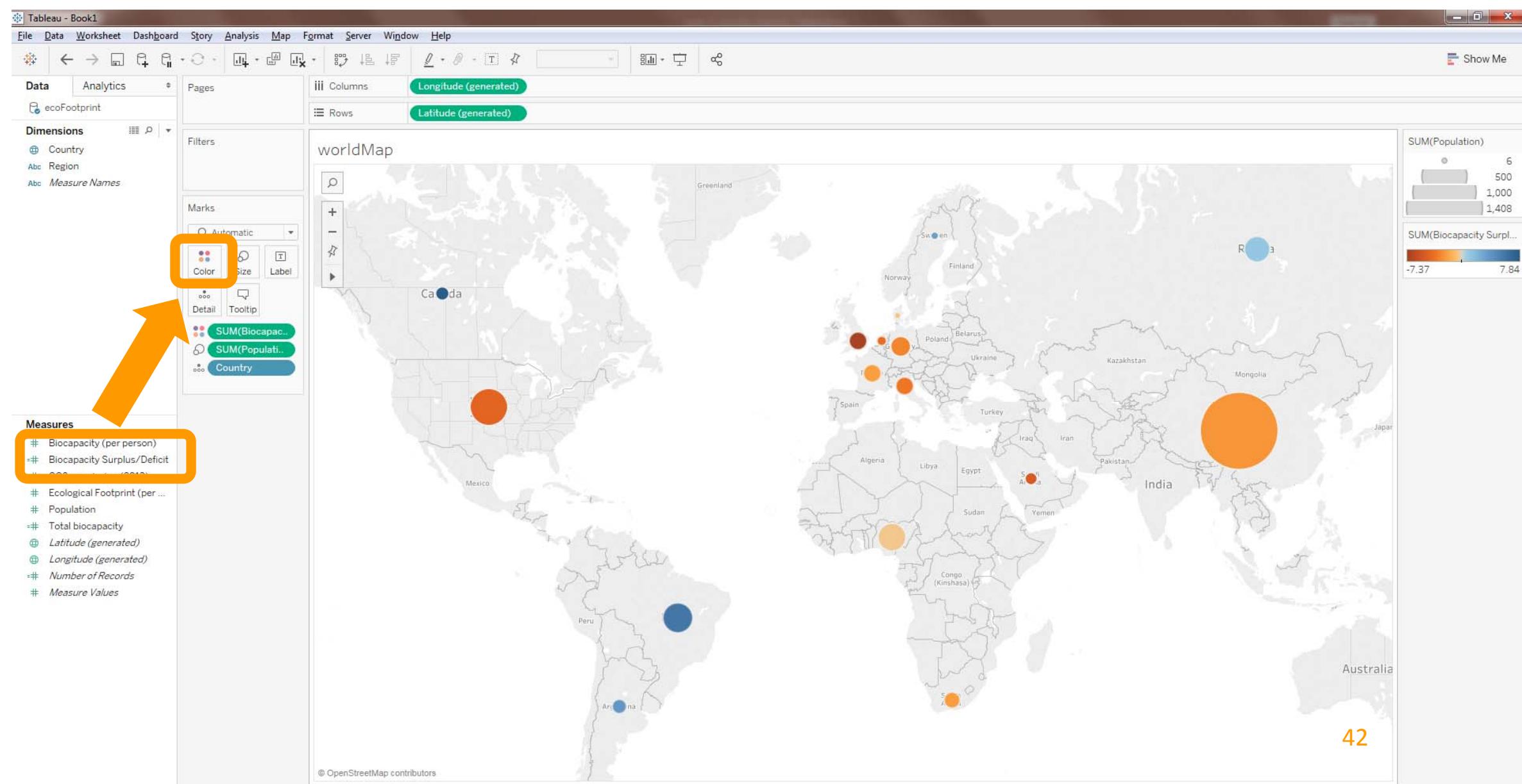
geographic map

- Double-clicking the legend allows you to adjust the circle sizes relative to the data.



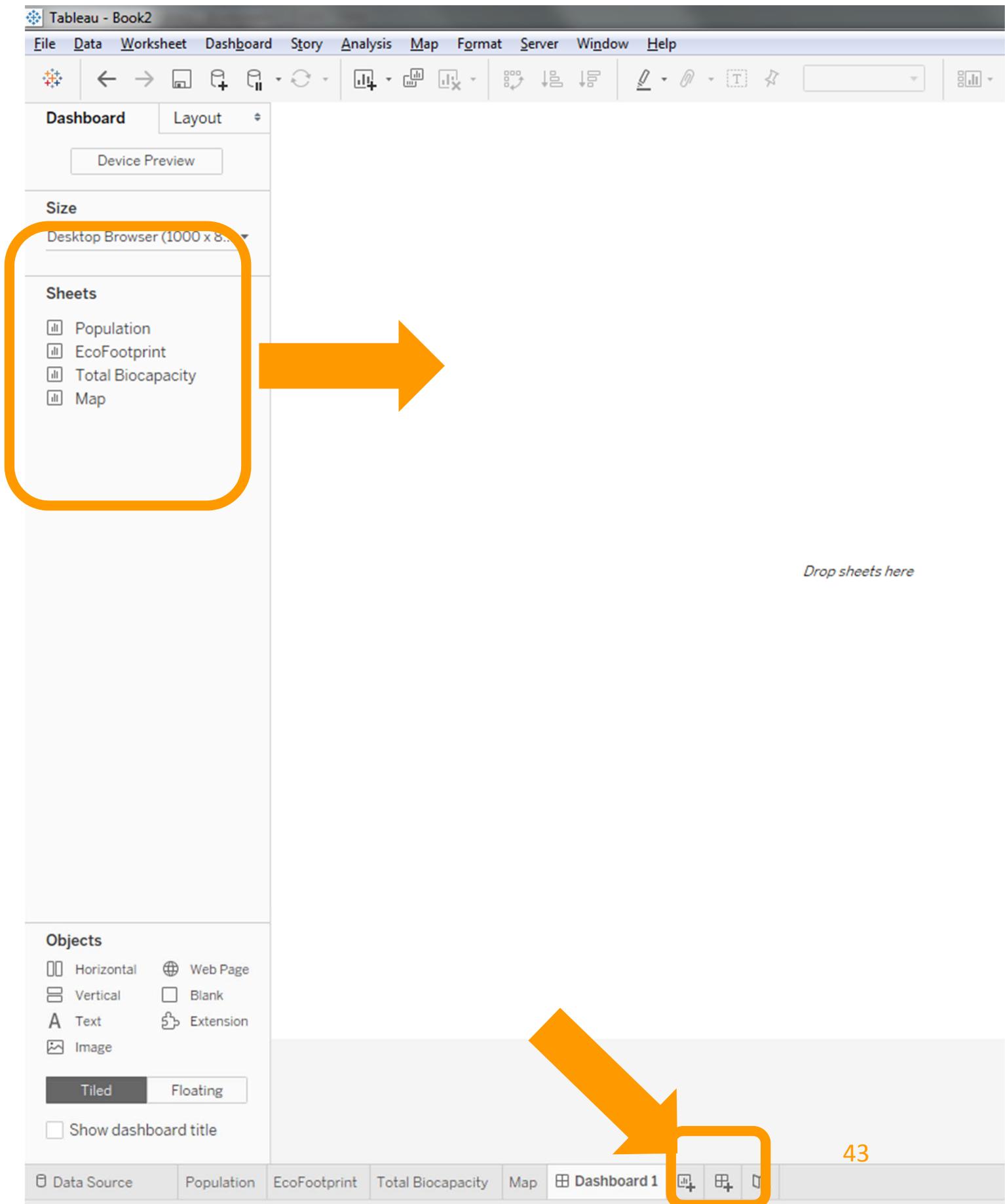
geographic map

- We apply colour to also show the “biocapacity surplus/deficit (per person)”.



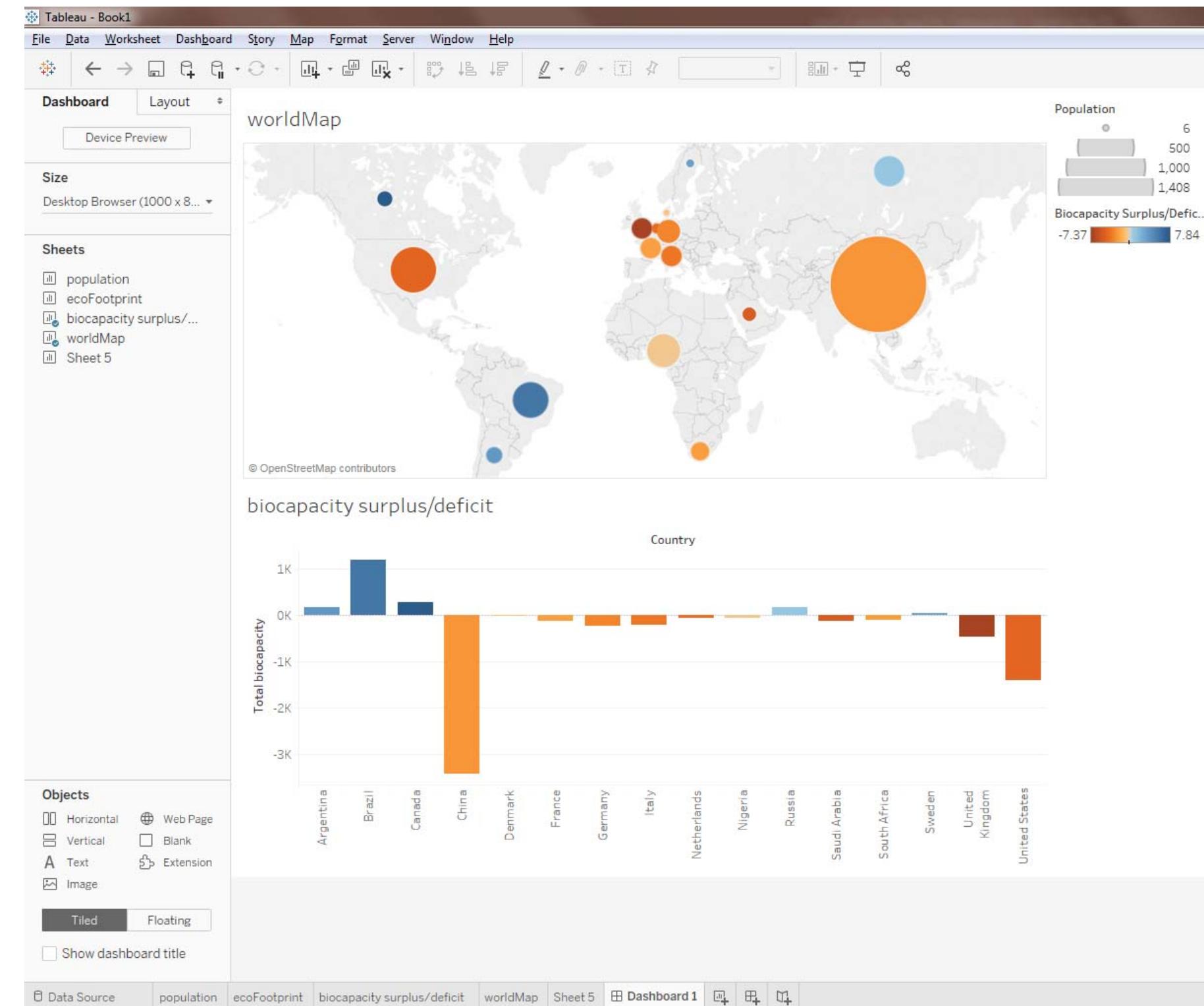
creating a dashboard

- Now let's put multiple visualizations together into a dashboard.
- To the left we can see the different visualization worksheets we have created.
- We can drag them into the dashboard canvas.



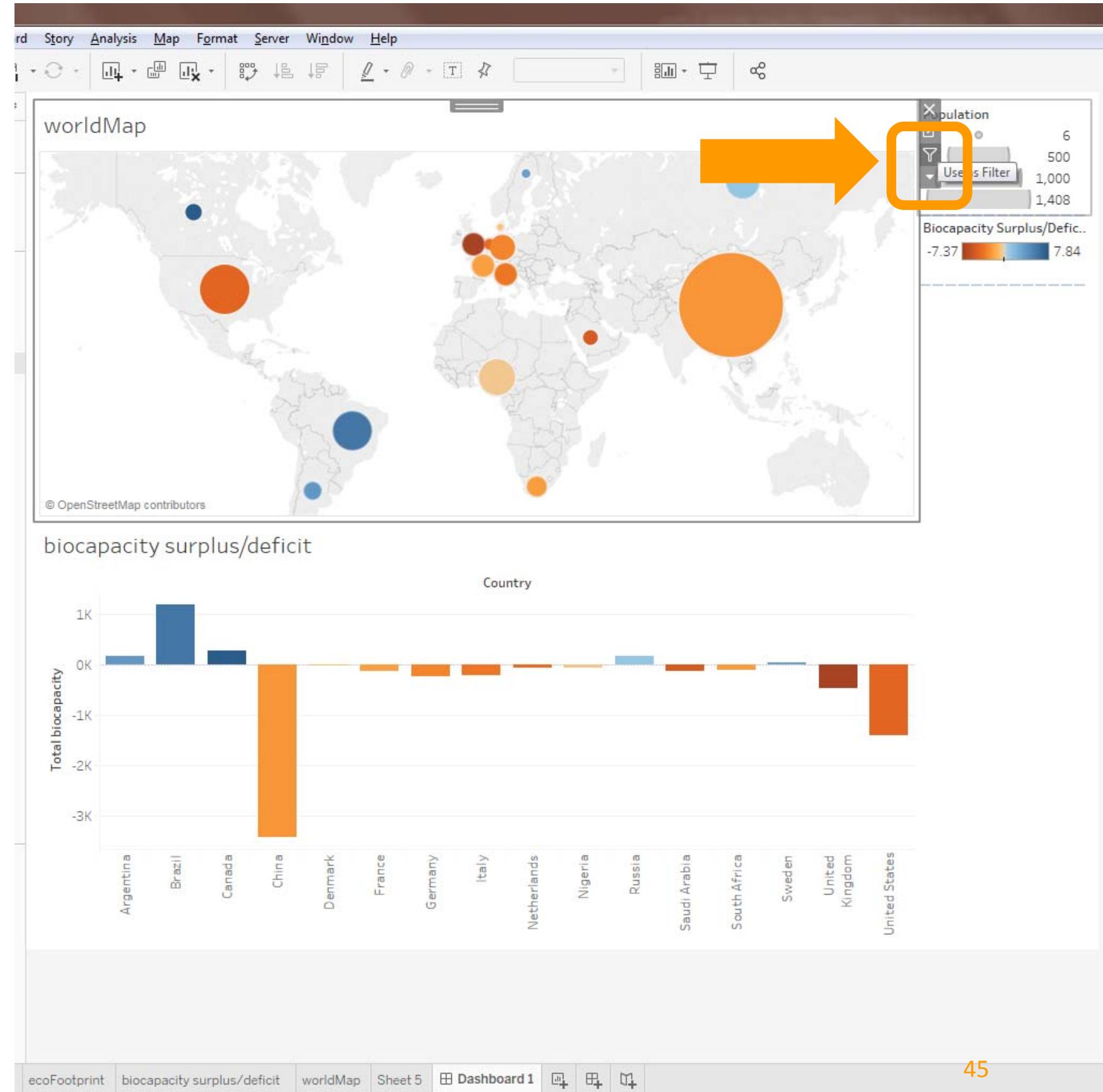
creating a dashboard

- We are dragging the map and the “biocapacity surplus/deficit” bar chart into the dashboard canvas.



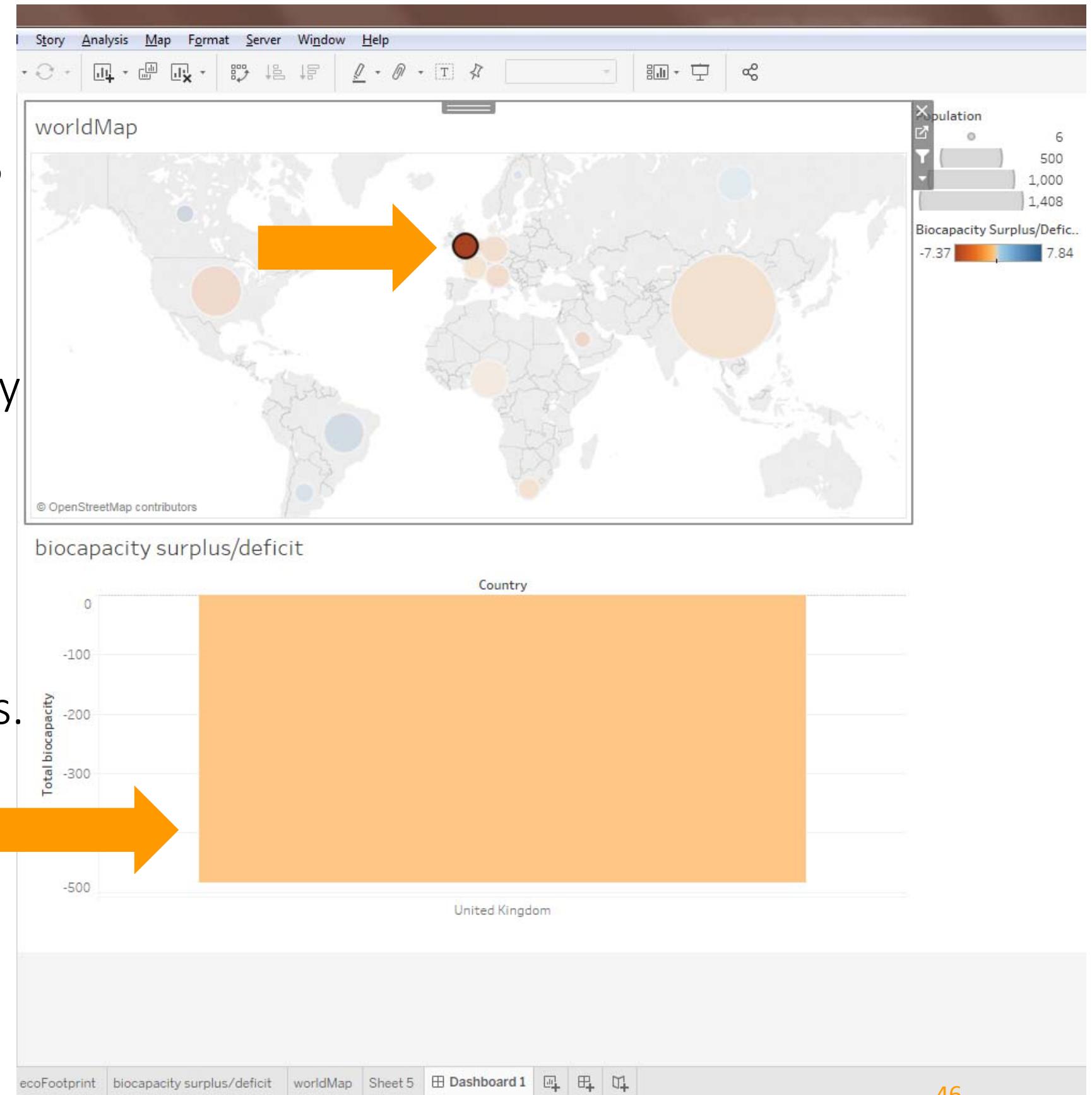
interactive filters

- Let's make the map an interactive filter for the bar chart.
- If we select a country in the map, the "Biocapacity" chart will be automatically filtered to only show this country.



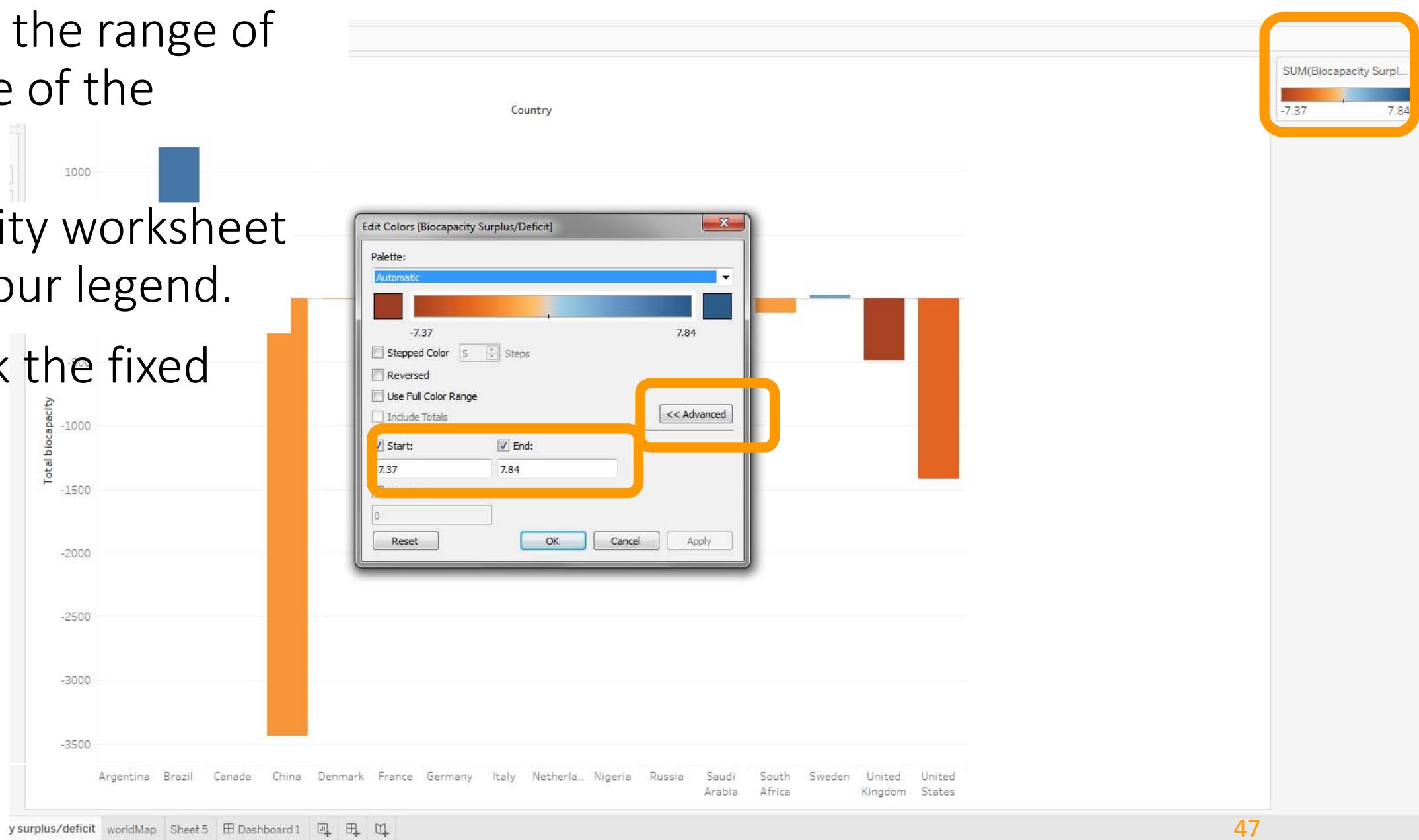
inconsistent colour scales

- However, this changes the colour scale in the Biocapacity chart, as only the selected value is considered.
- The value ranges in both visualizations are inconsistent, although both show the same values.



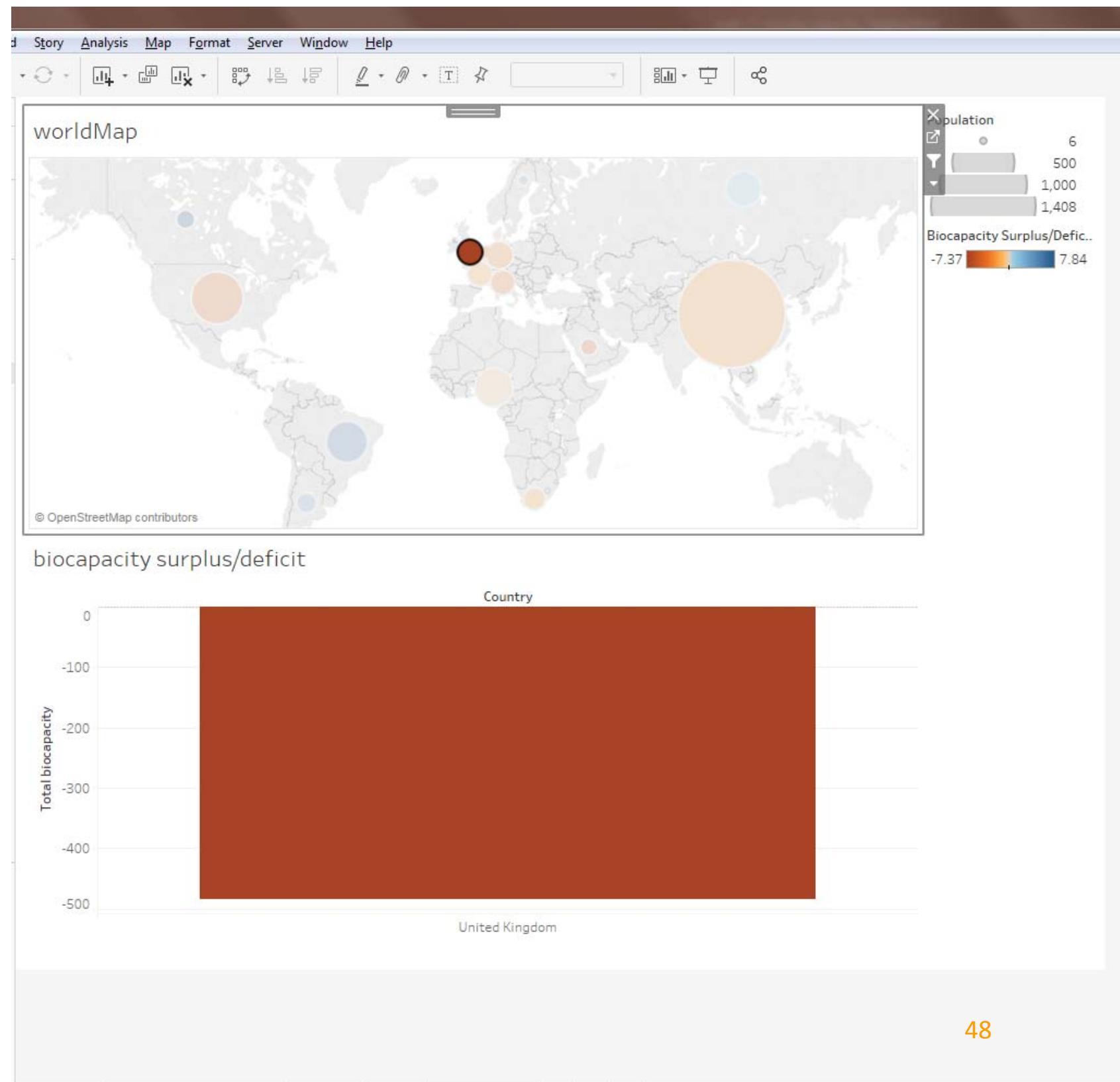
inconsistent colour scales

- To fix this, we have to fix the range of values in the colour scale of the Biocapacity chart.
- Go back to the Biocapacity worksheet and double-click the colour legend.
- Under “Advanced” check the fixed ranges.
- Click “Apply”

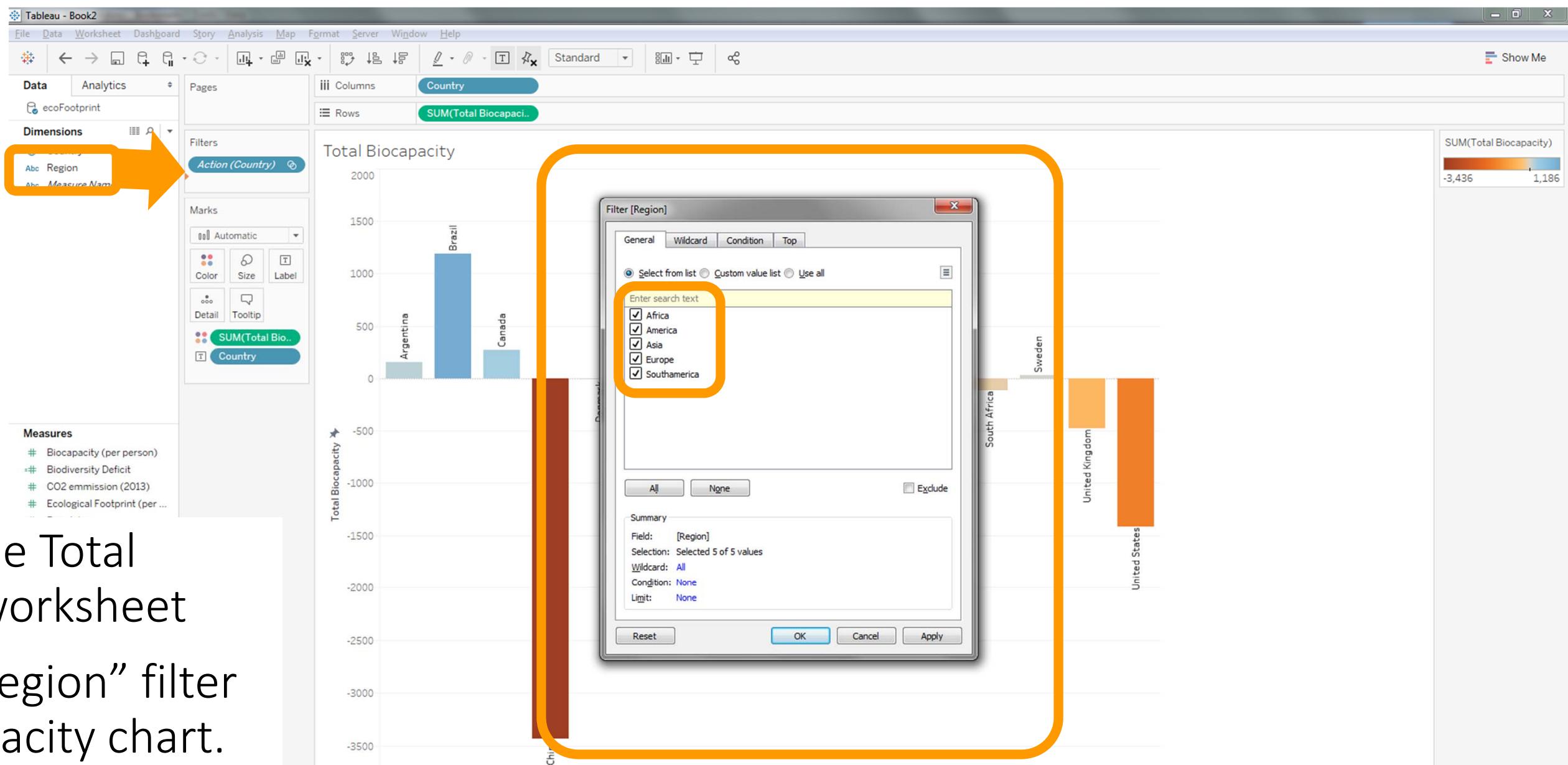


inconsistent colour scales

- Back in the dashboard, the colours between map and Biocapacity chart should now be consistent when filtering for particular countries
- Apply the same “fixed” colour scale to the map and make the Biocapacity chart a filter of the map.



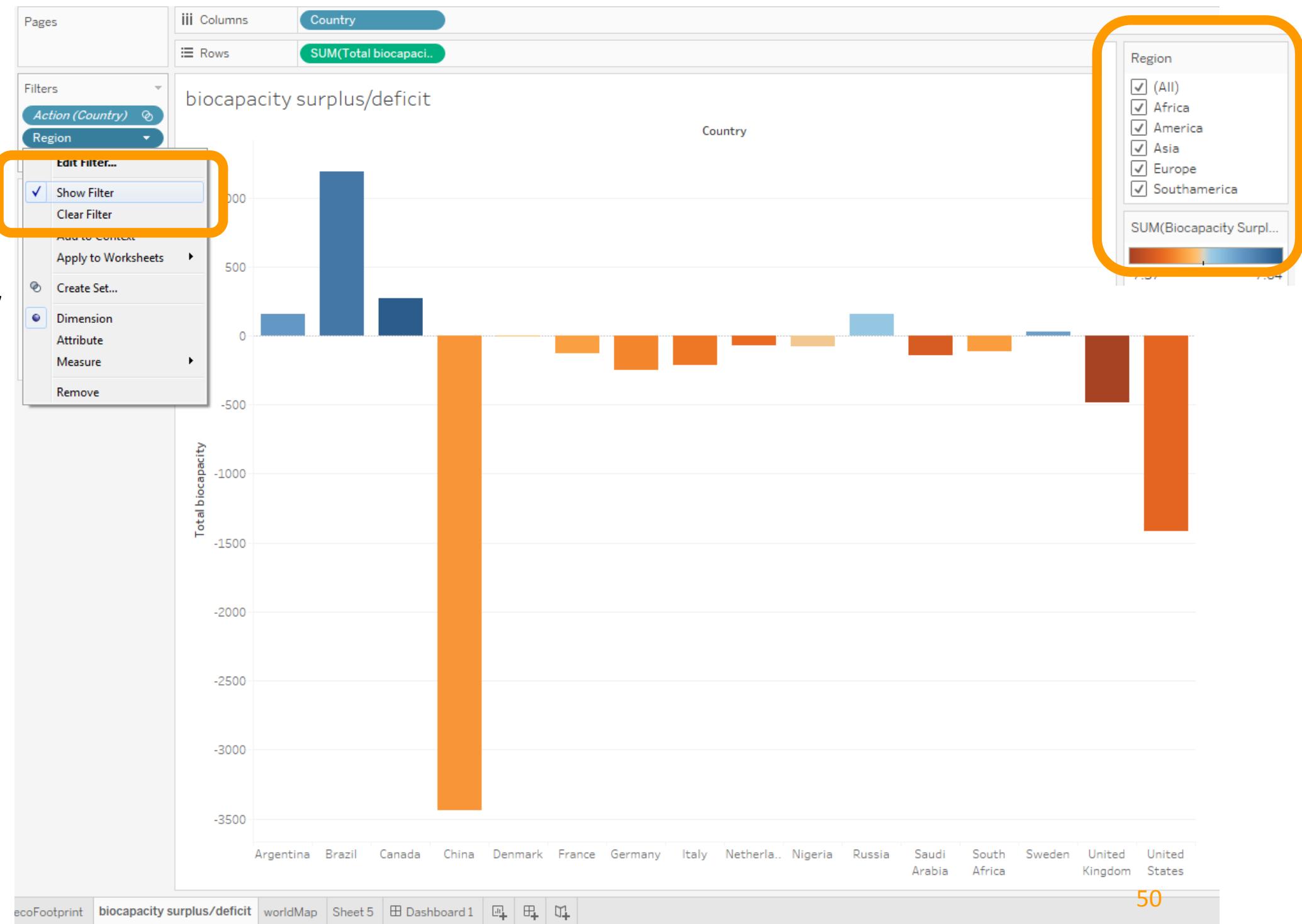
applying additional filters



- Go back to the Total Biocapacity worksheet
- Let's add a “region” filter to the Biocapacity chart.

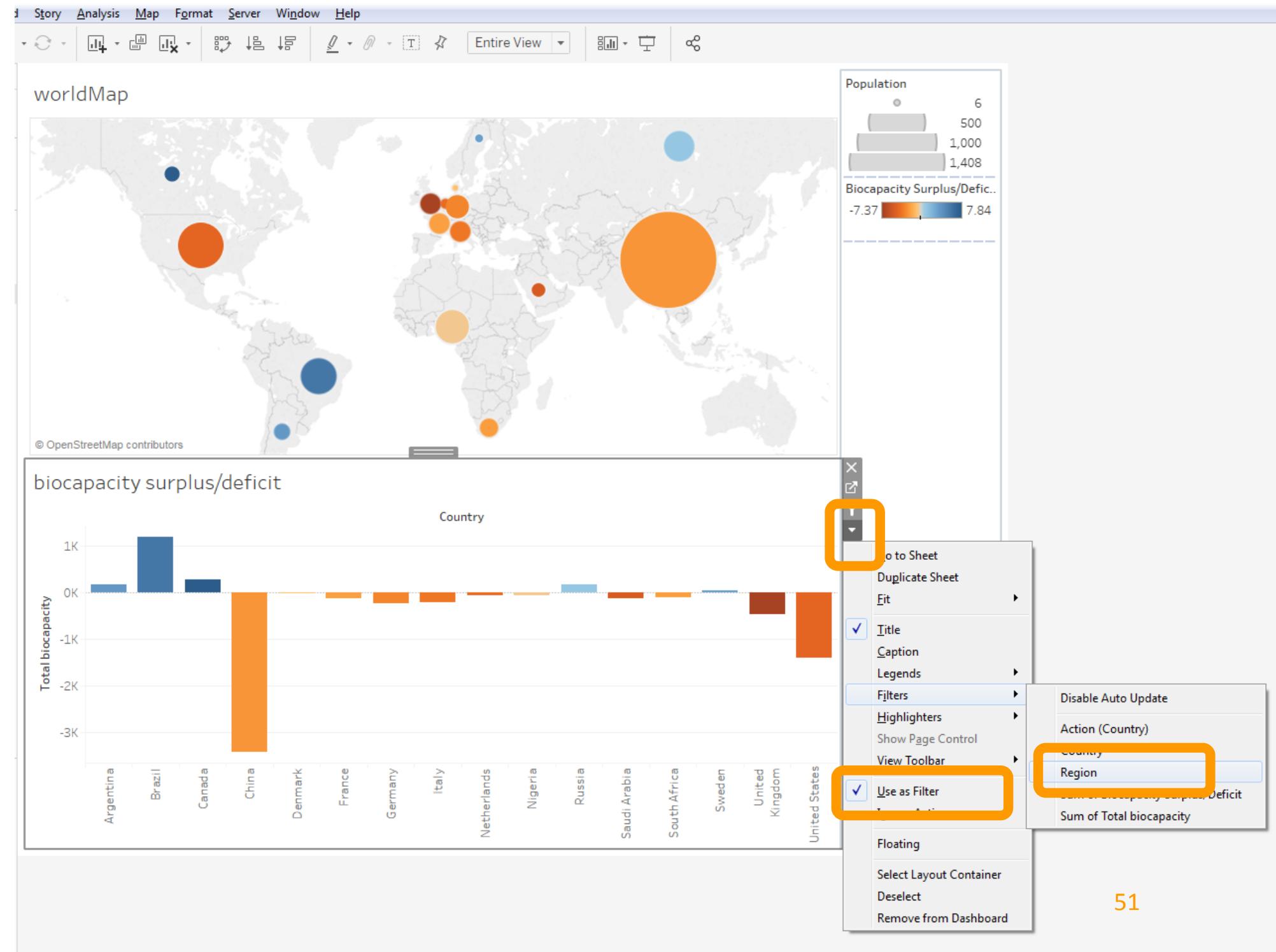
applying additional filters

- To make the filter visible as an interactive element, press the small arrow to the right and select “Show Filter”



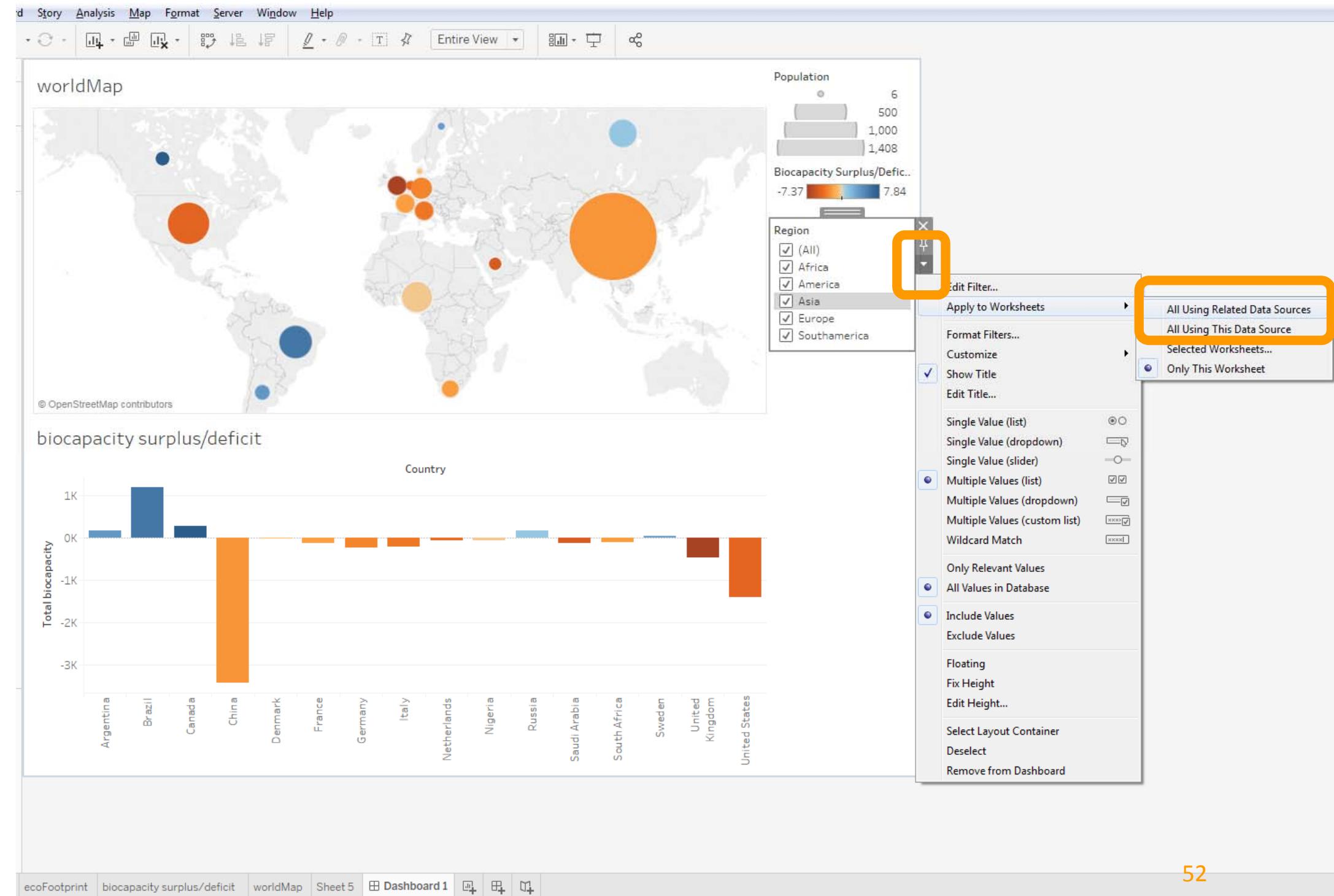
applying additional filters

- The filter will not be visible in your dashboard yet.
- You have to select it from the “Filters” options of the Biocapacity worksheet.

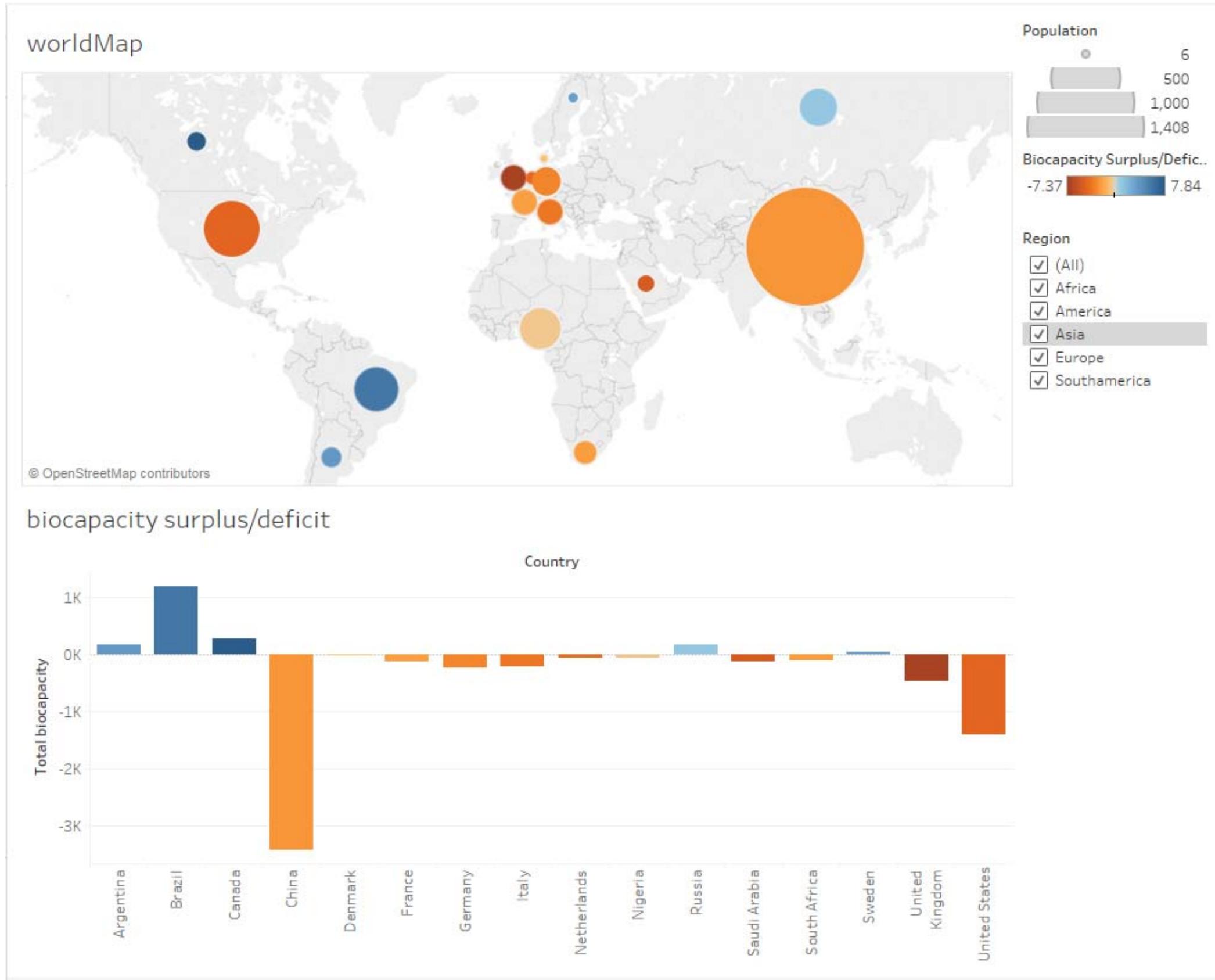


applying additional filters

- The filter is now visible but only applies to the surplus worksheet.
- In the filter options select
 - “Apply to Worksheets”
 - “All Using Related Data Sources”



all done



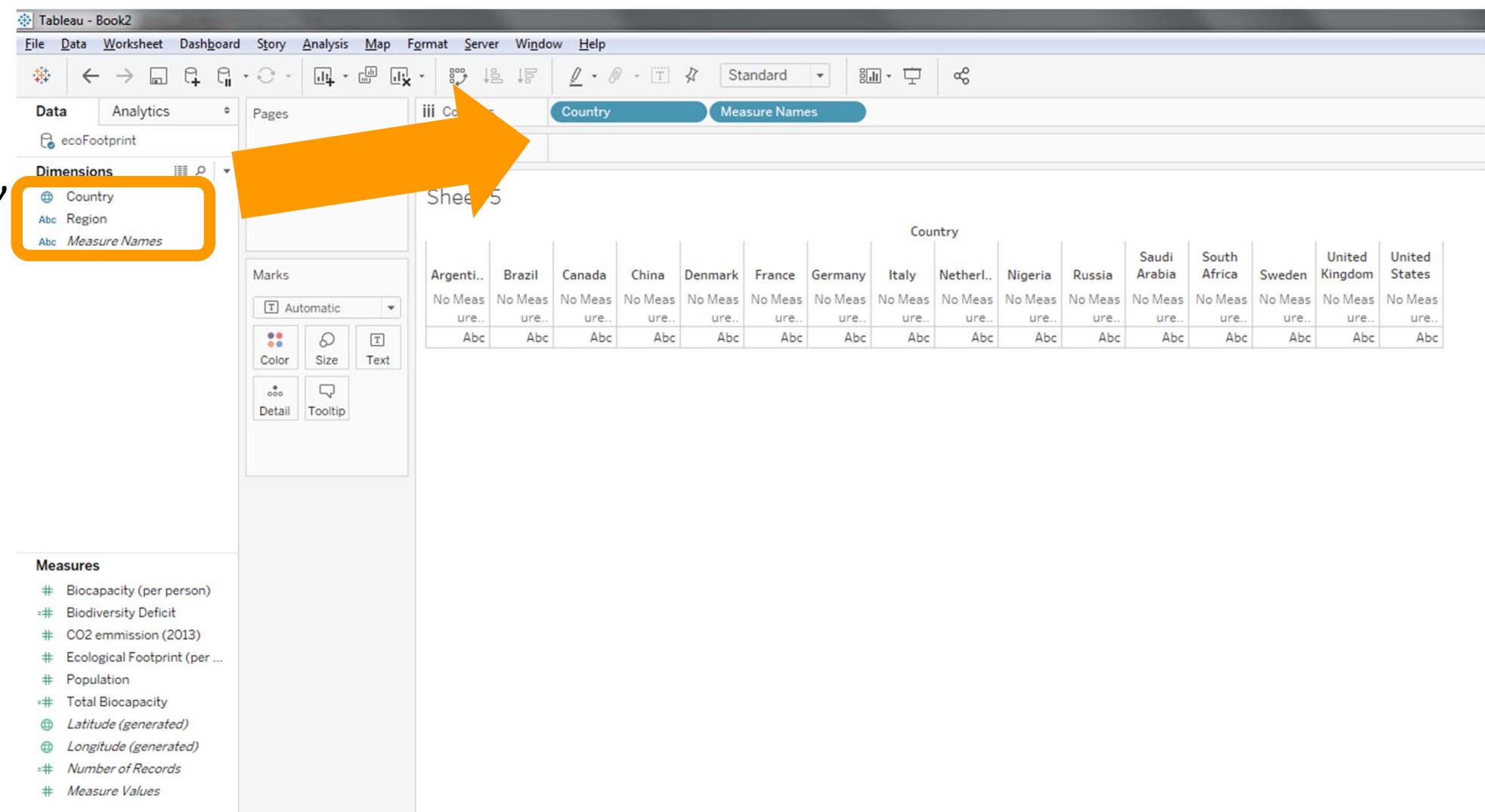
additional exercises, tutorials, data, resources

- Tableau_visExercise_01
 - Try to visualise this small dataset in your own time
- How-to instructions in on studies: [CS5044/Tutorials/Tableau/tableau_howTo](#)
 - Manipulating data
 - Interactive elements
 - Visualising geospatial data
- More general Tableau tutorials on resource list
- Lots of resources on the web!
- Next week Monday (Week 4) – Tableau tutorial II

Additional material

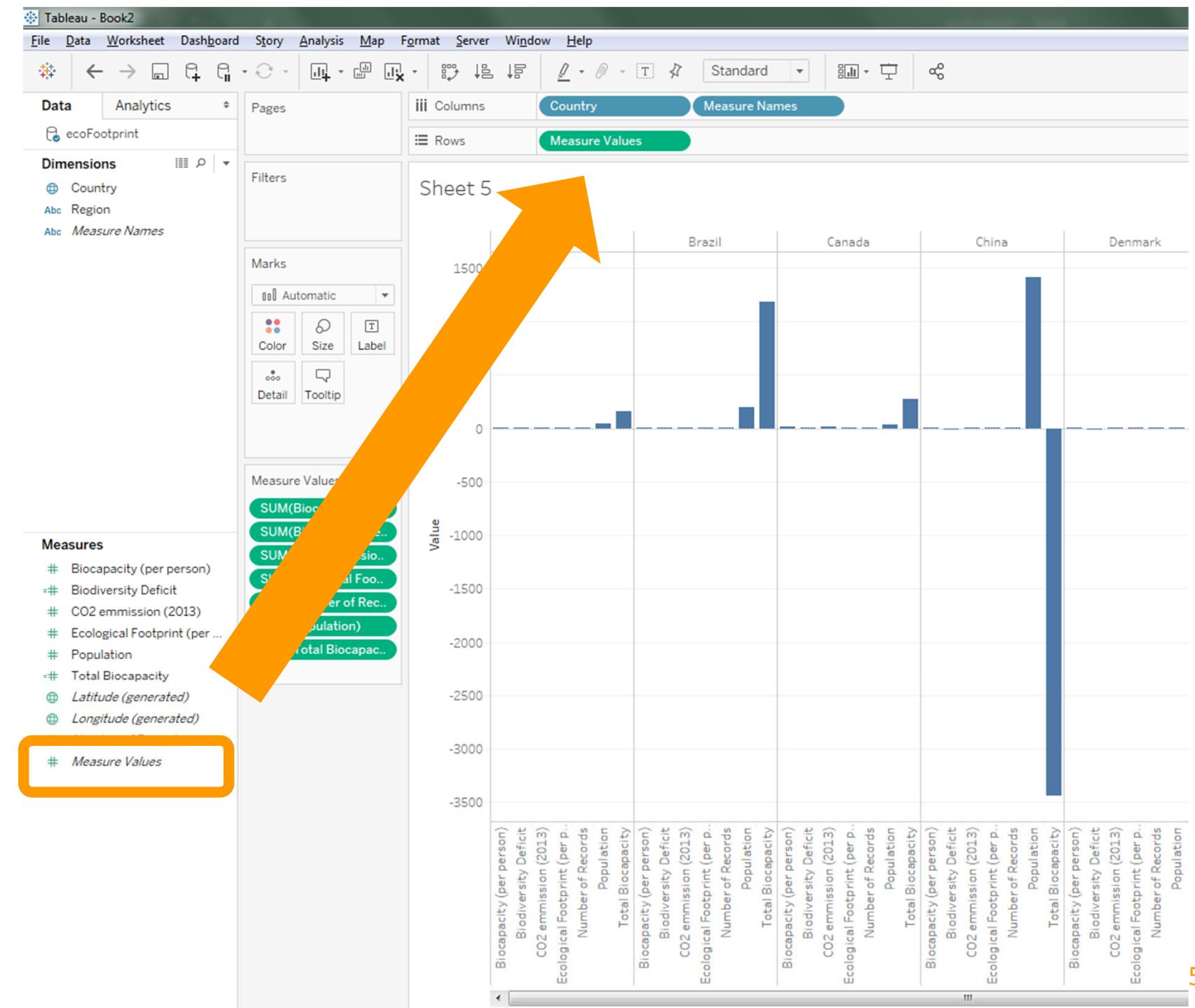
Creating a grouped bar chart manually → eco footprint vs. biocapacity

1. Drag the “Country” dimension to columns
2. Drag the “Measure Names” dimension to columns



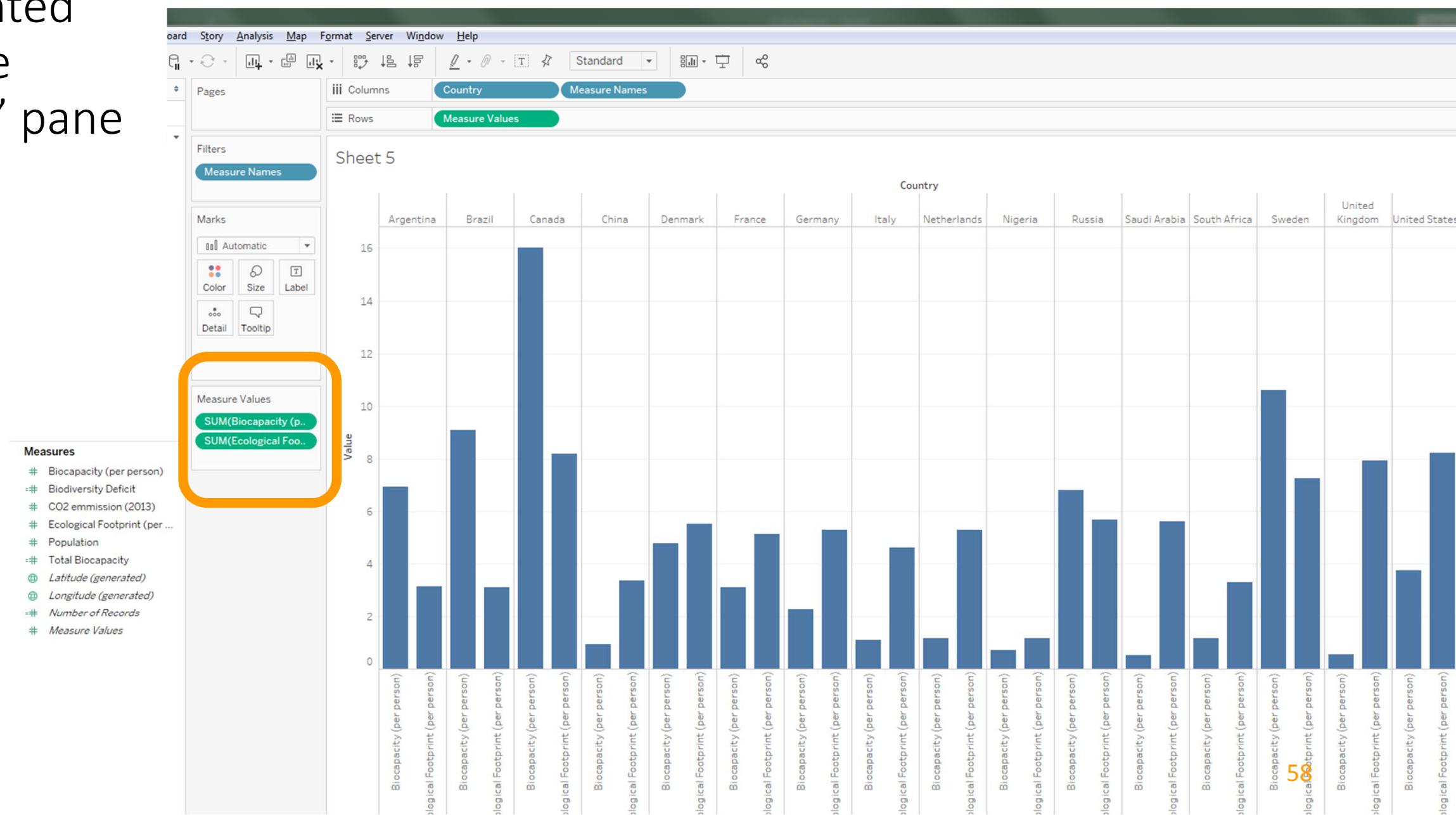
eco footprint vs. biocapacity

3. Drag the “Measure Values” to rows



eco footprint vs. biocapacity

4. Delete the unwanted “measures” in the “Measure Values” pane



eco footprint vs. biocapacity

5. Colour marks according to “Measure Names”

