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?

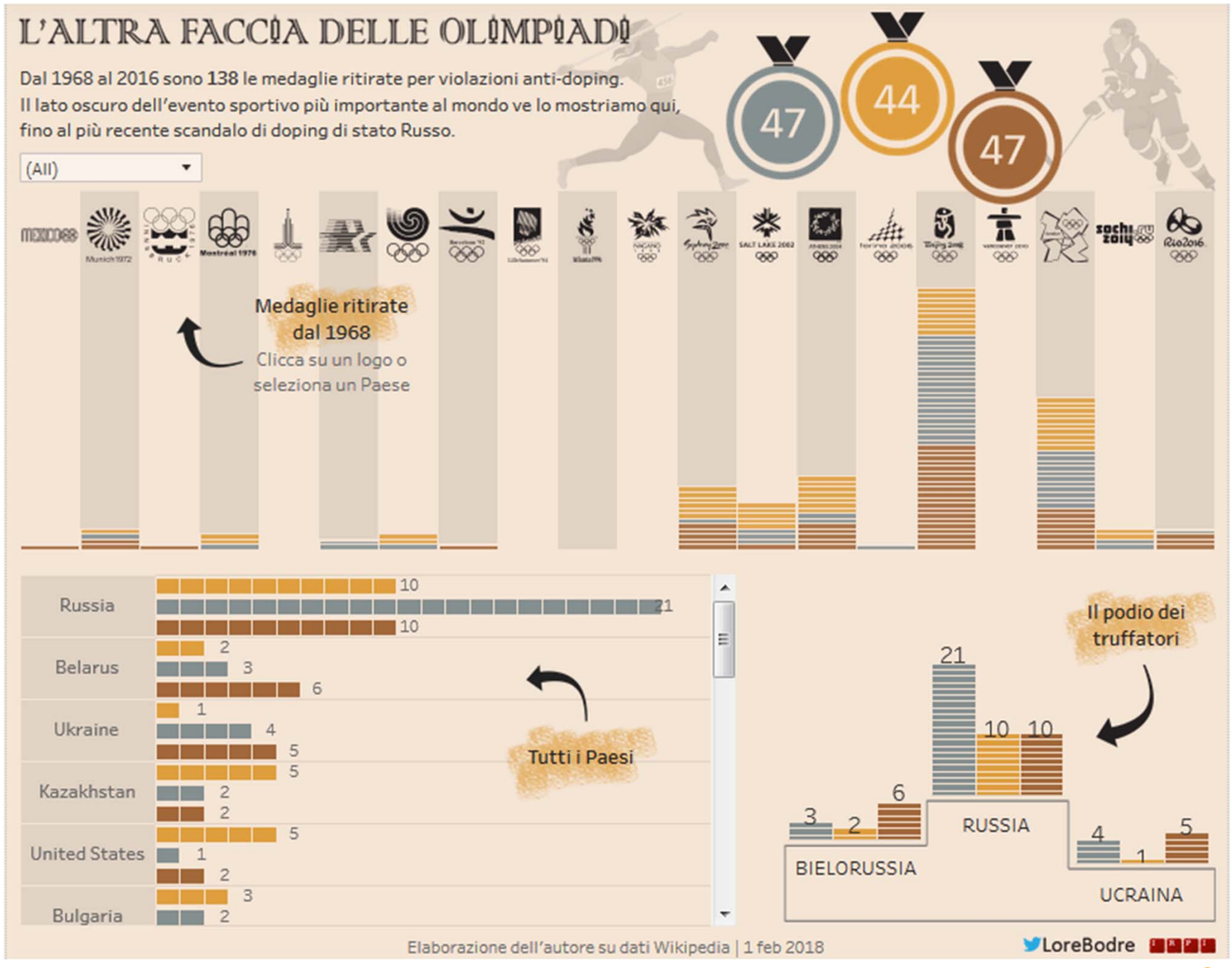
- Installed on all lab machines in John Honey 110
  - Start into Windows
  - Look for “Tableau Desktop”

- 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

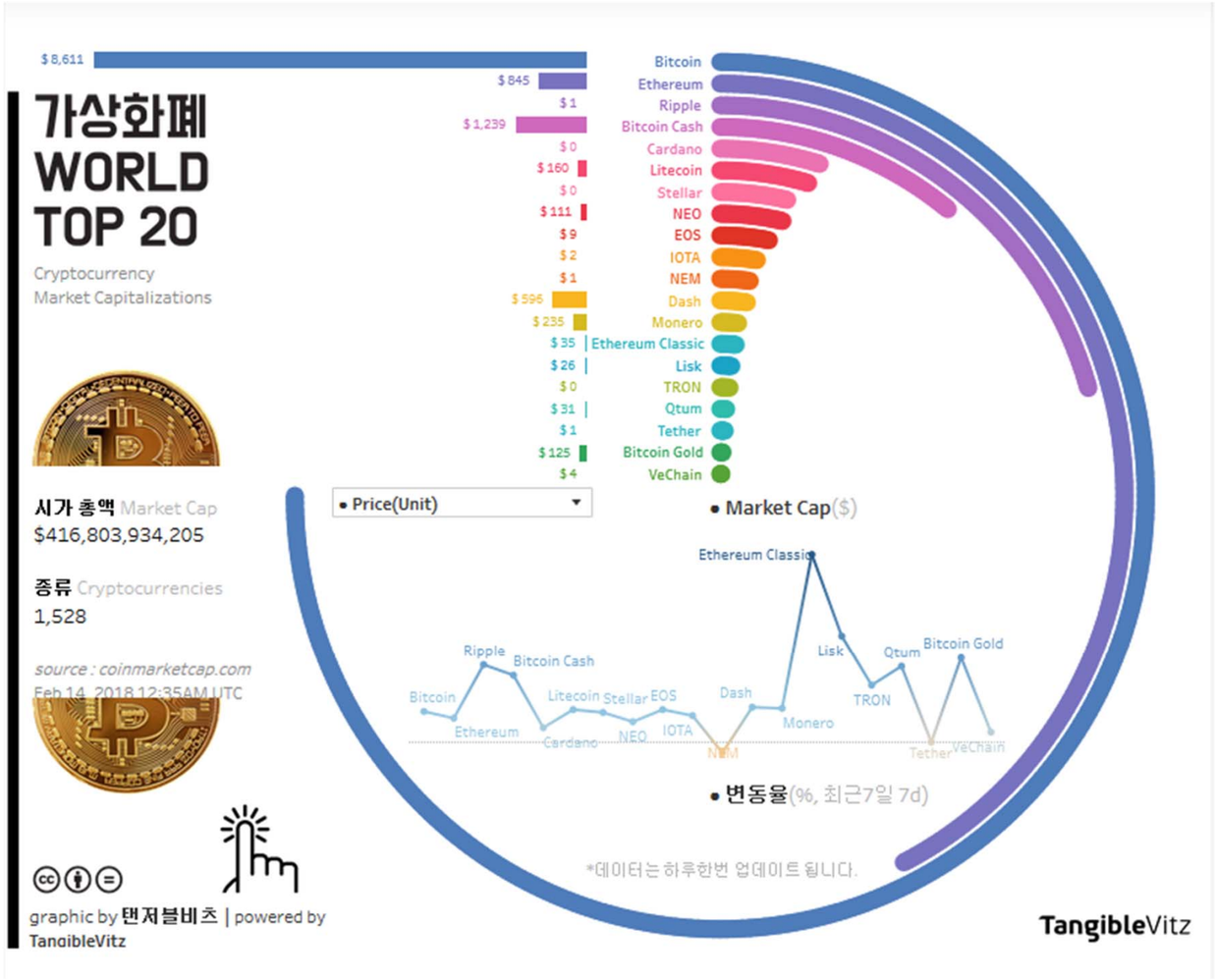
Lorenzo B.



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

comparison of cryptocurrencies by value/supply

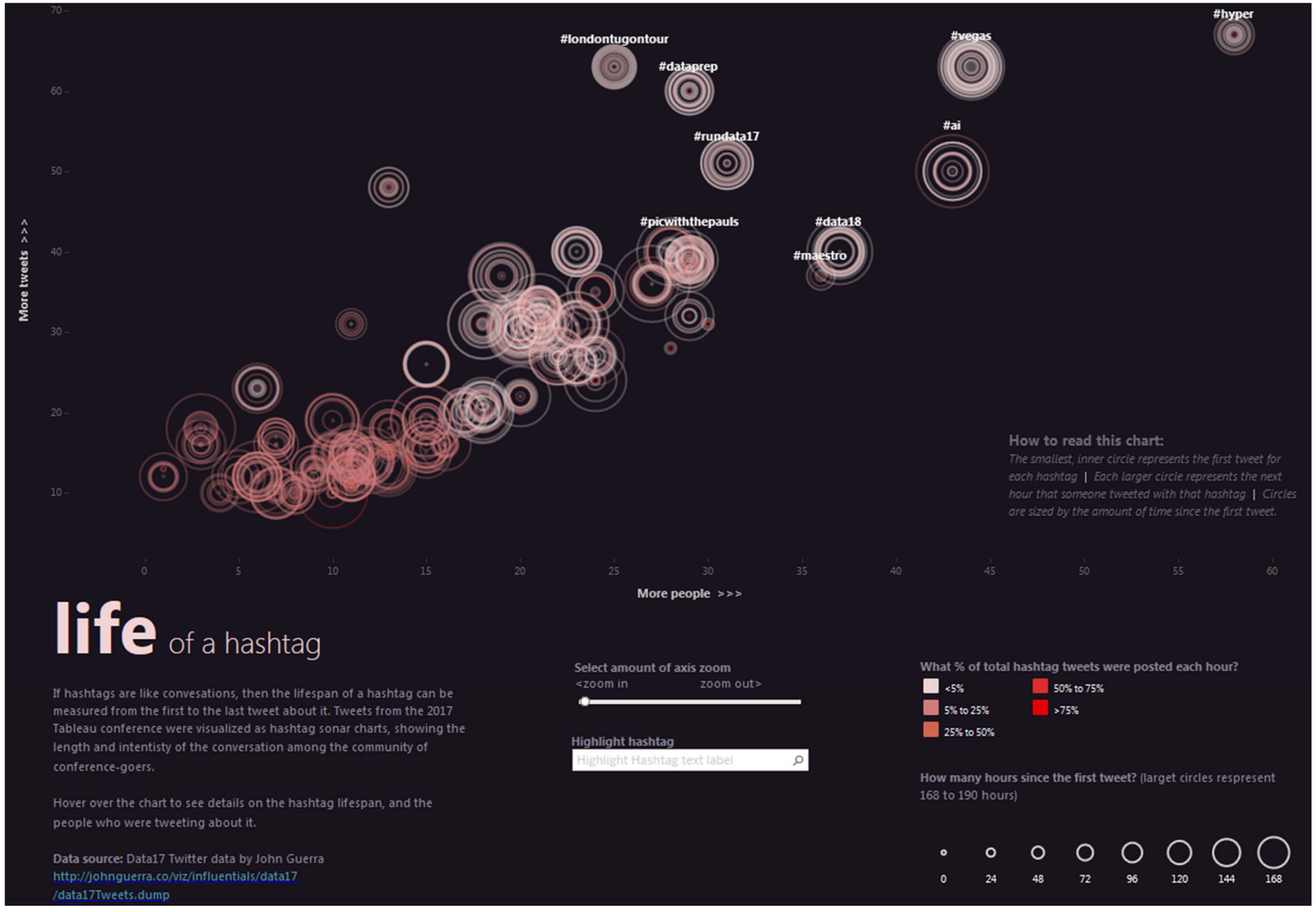
탠저블비츠 TangibleVitz



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

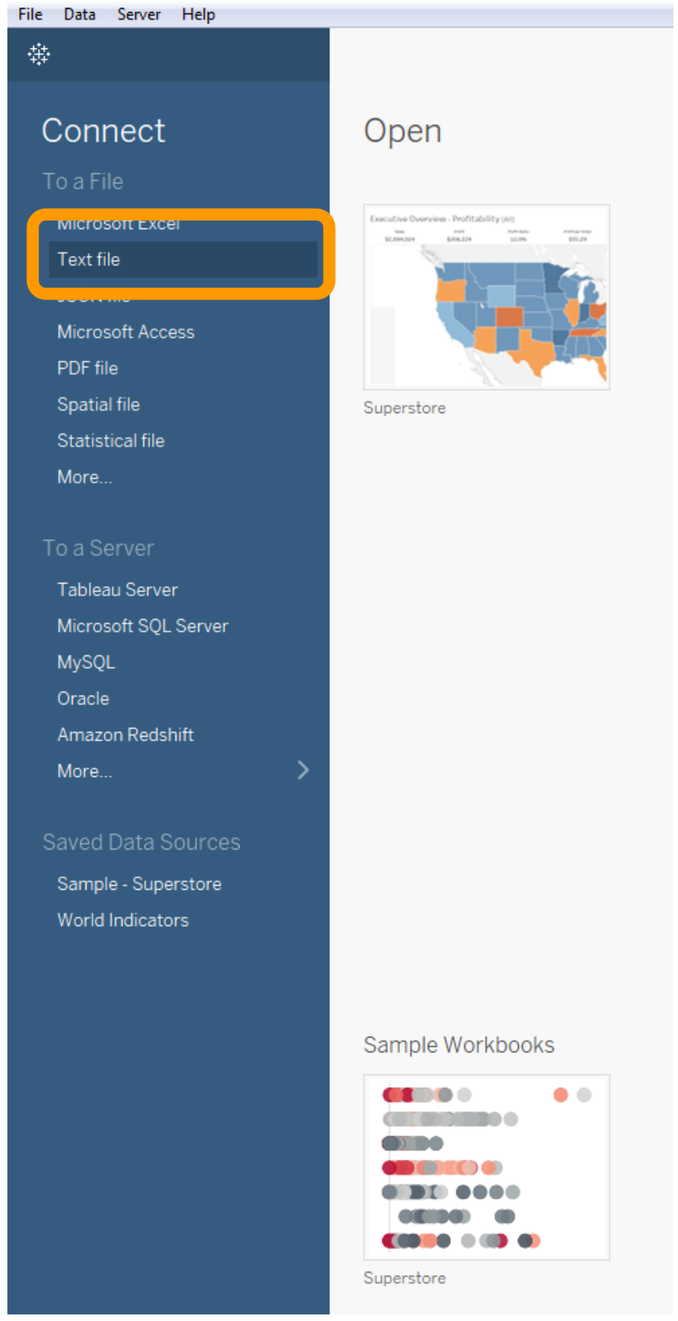
loading the data

data: studres/CS5044/Tutorials/Tableau/data

- Find ecofootprint.csv in the “data” folder
- Open Tableau
- Load this data as “Text File”

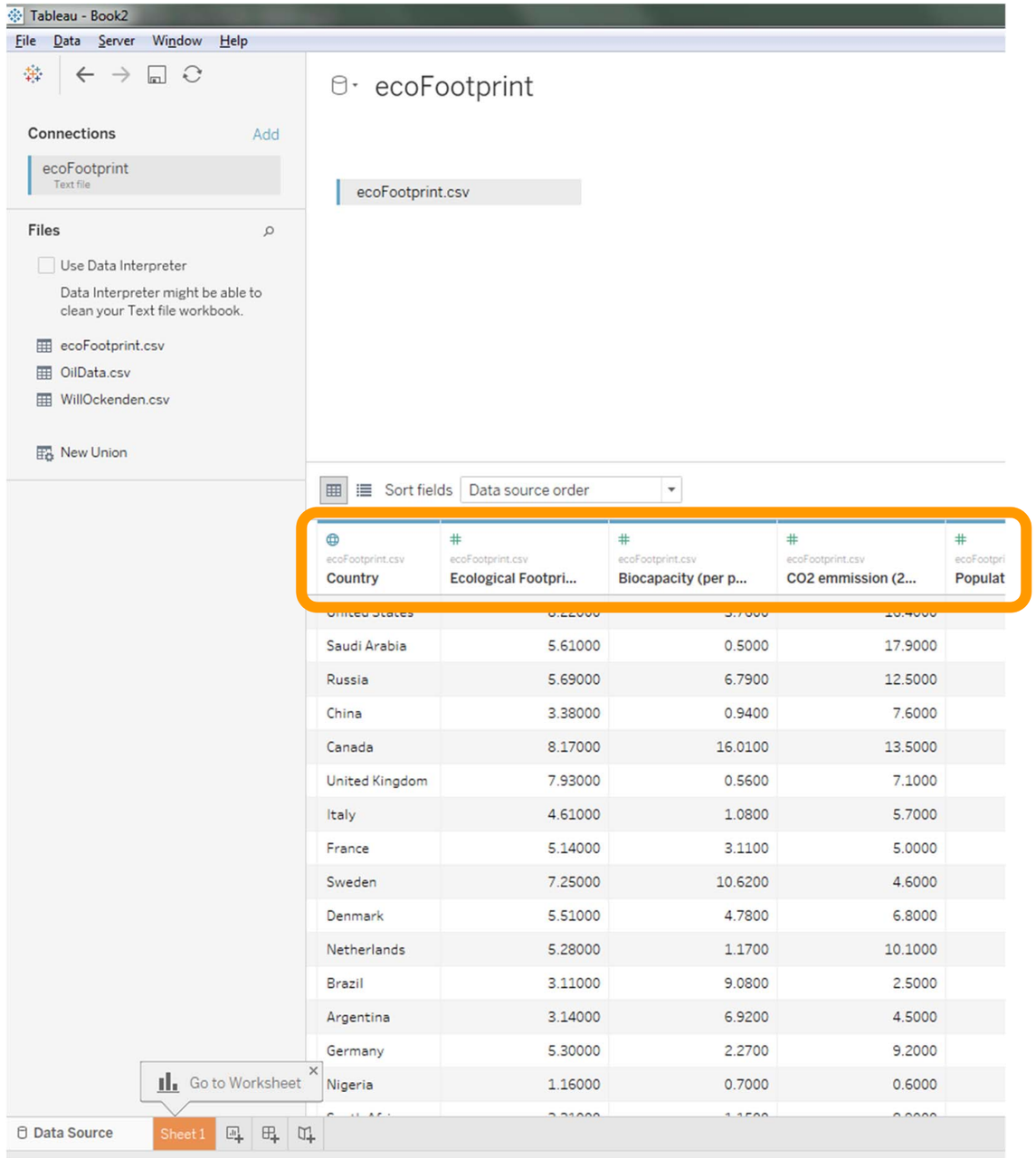
- More info about the data:

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



loading the data

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

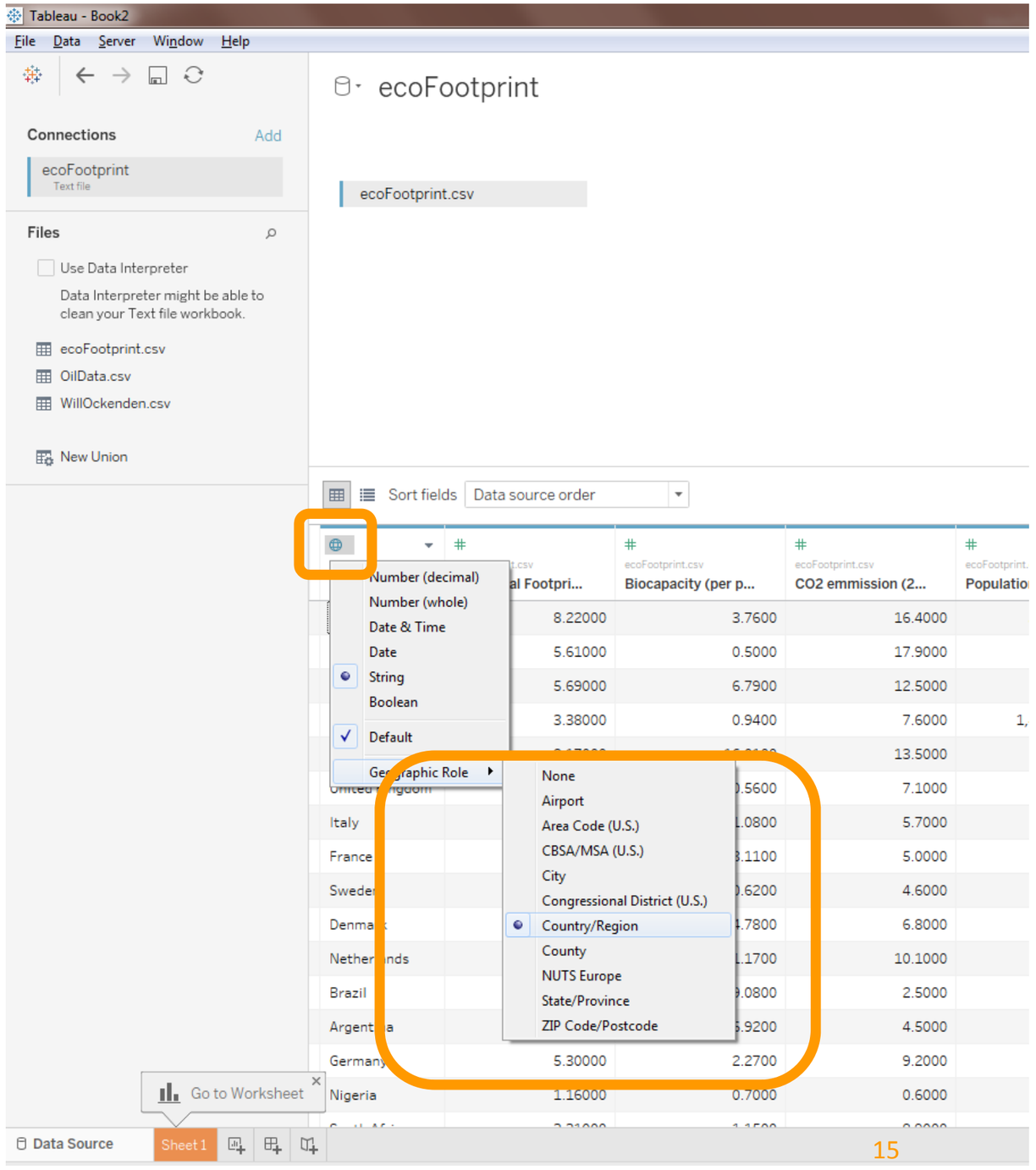


Tableau worksheets

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

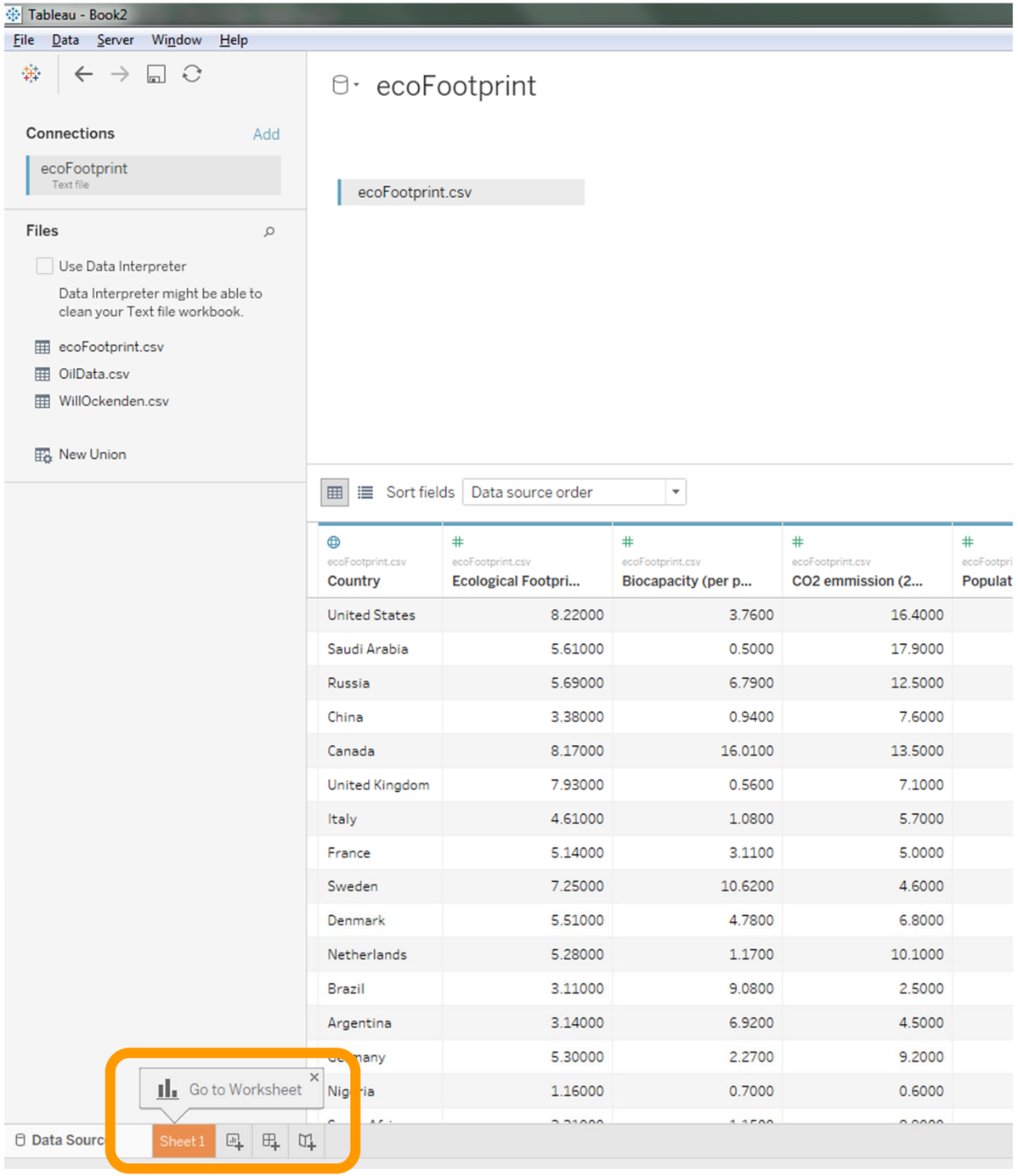
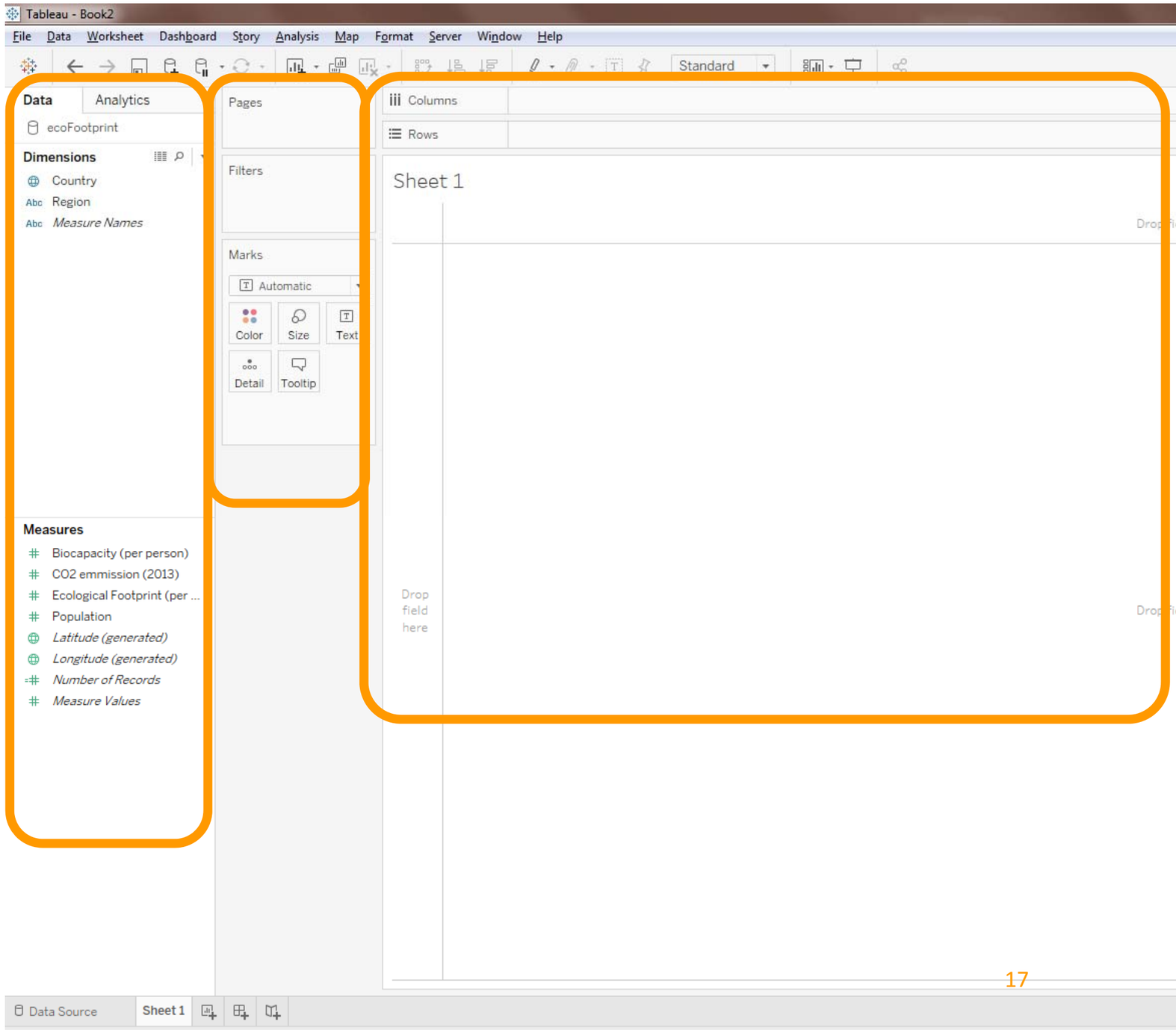


Tableau worksheets

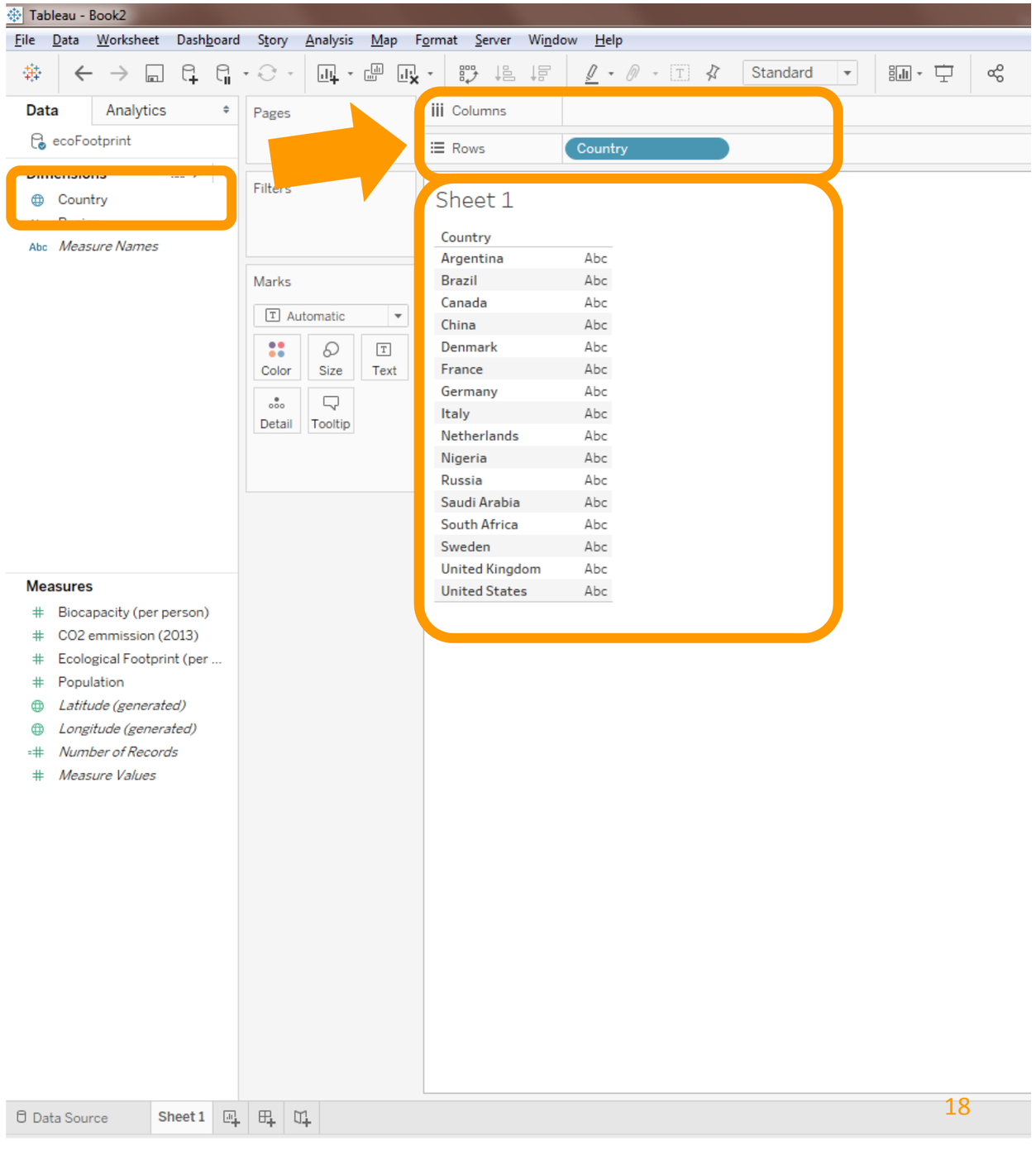
- Left: list of data attributes
  - Dimensions → categorial data
  - Measures → quantitative data
- Middle: choices for marks and visual variables
- Right: visualization space that is structured into rows and columns



visualizing countries

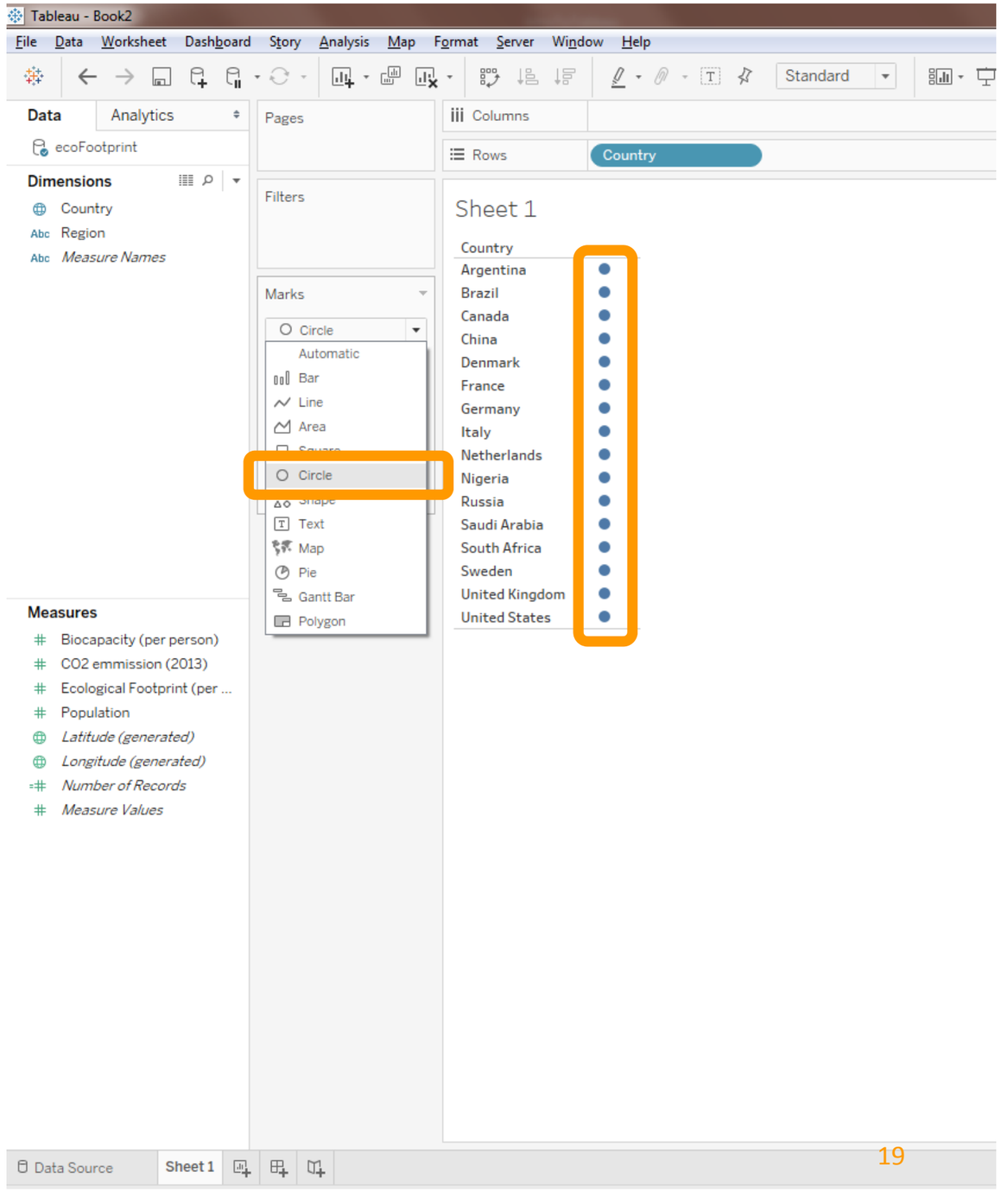
- Drag the “country” attribute into the “Rows” pane

- All values for “country” will be distributed by row in the resulting view



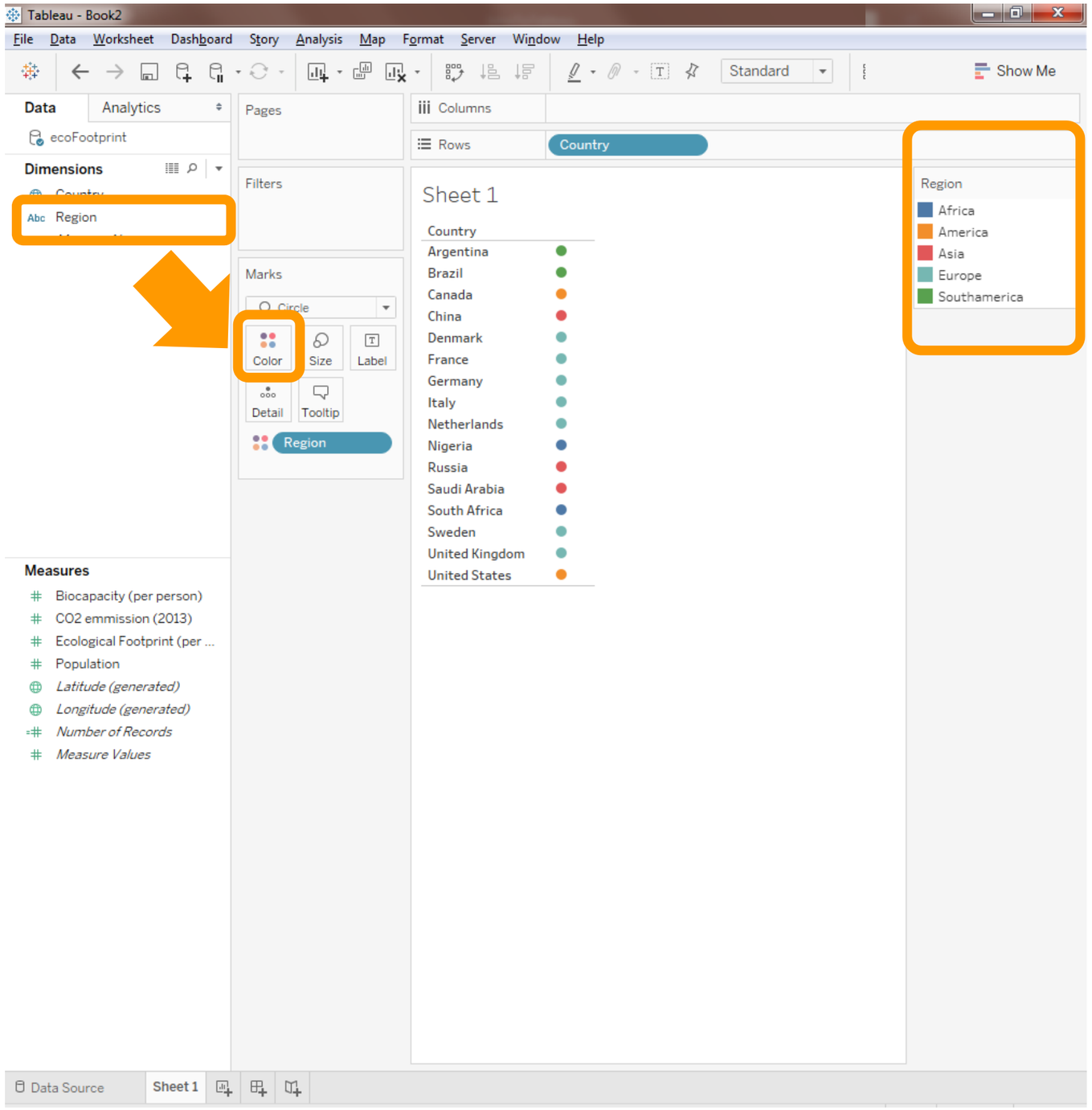
choosing marks

- You can choose the type of marks to represent countries



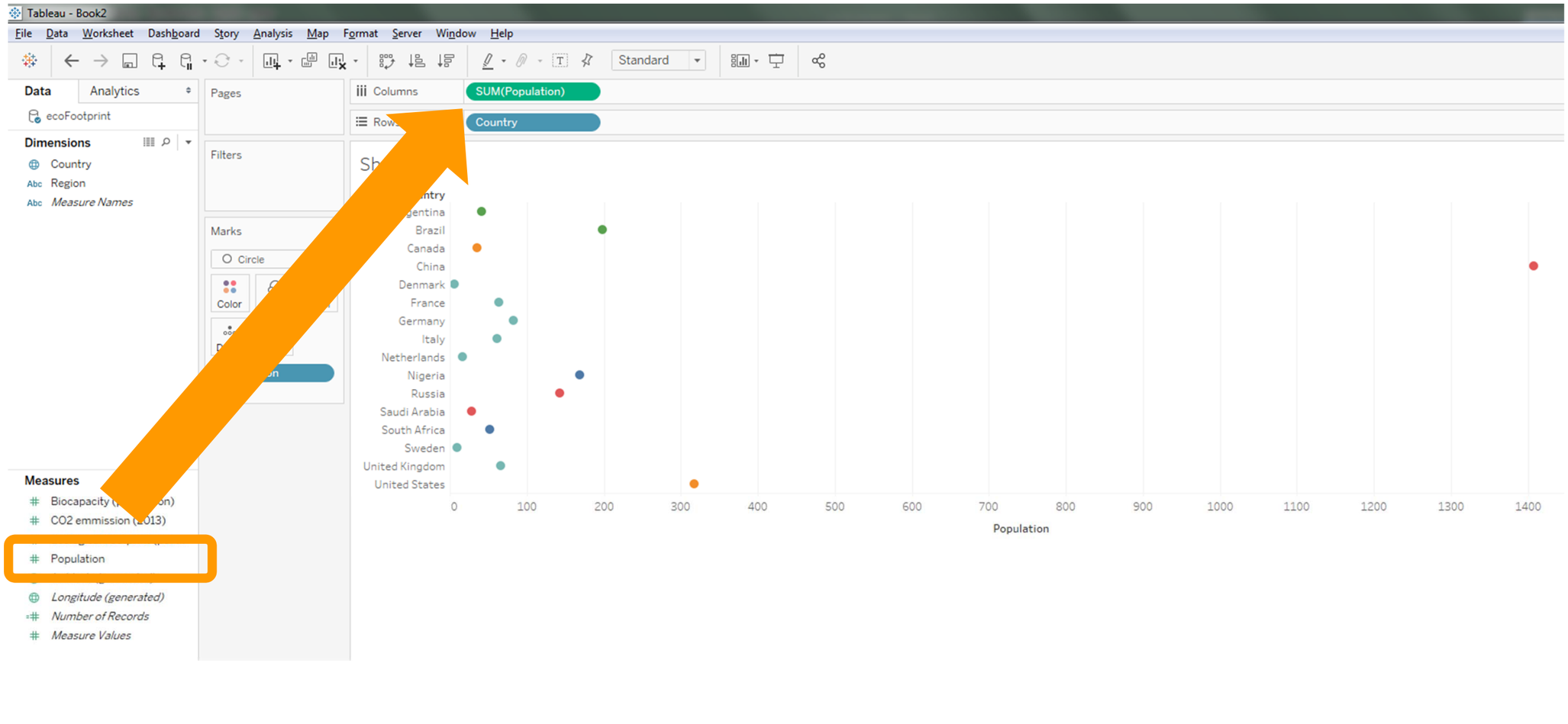
choosing visual variables

- 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.



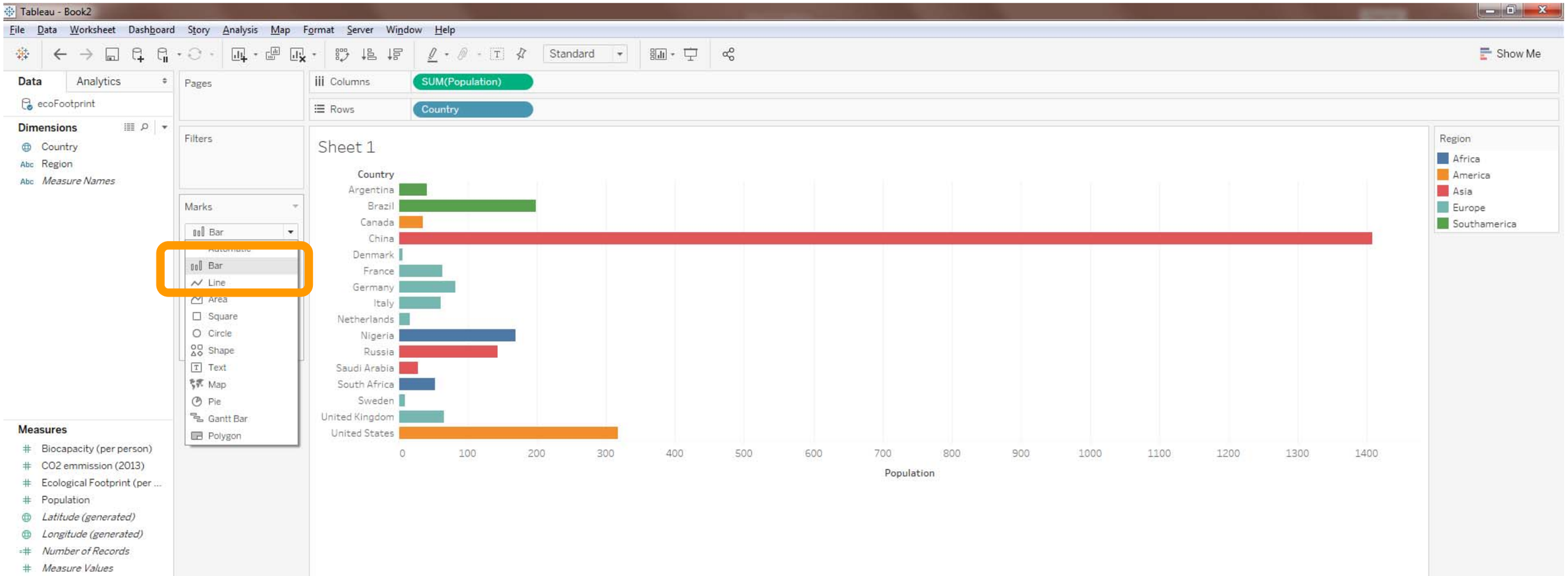
country by population

- 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



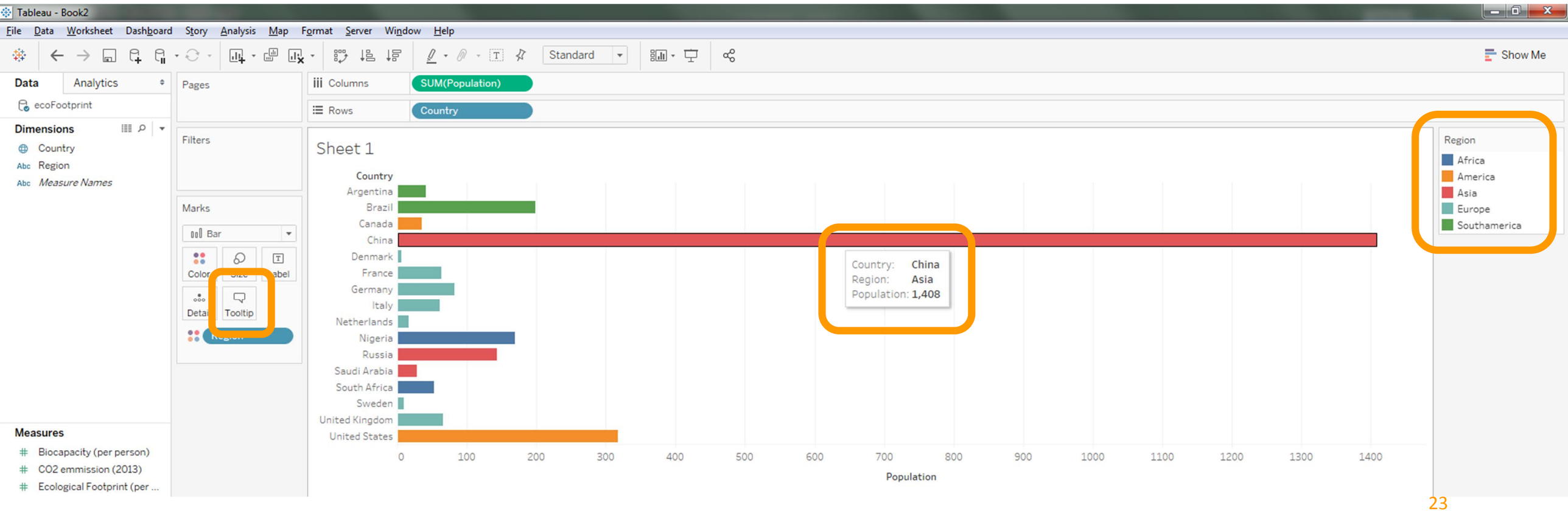
country by population

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



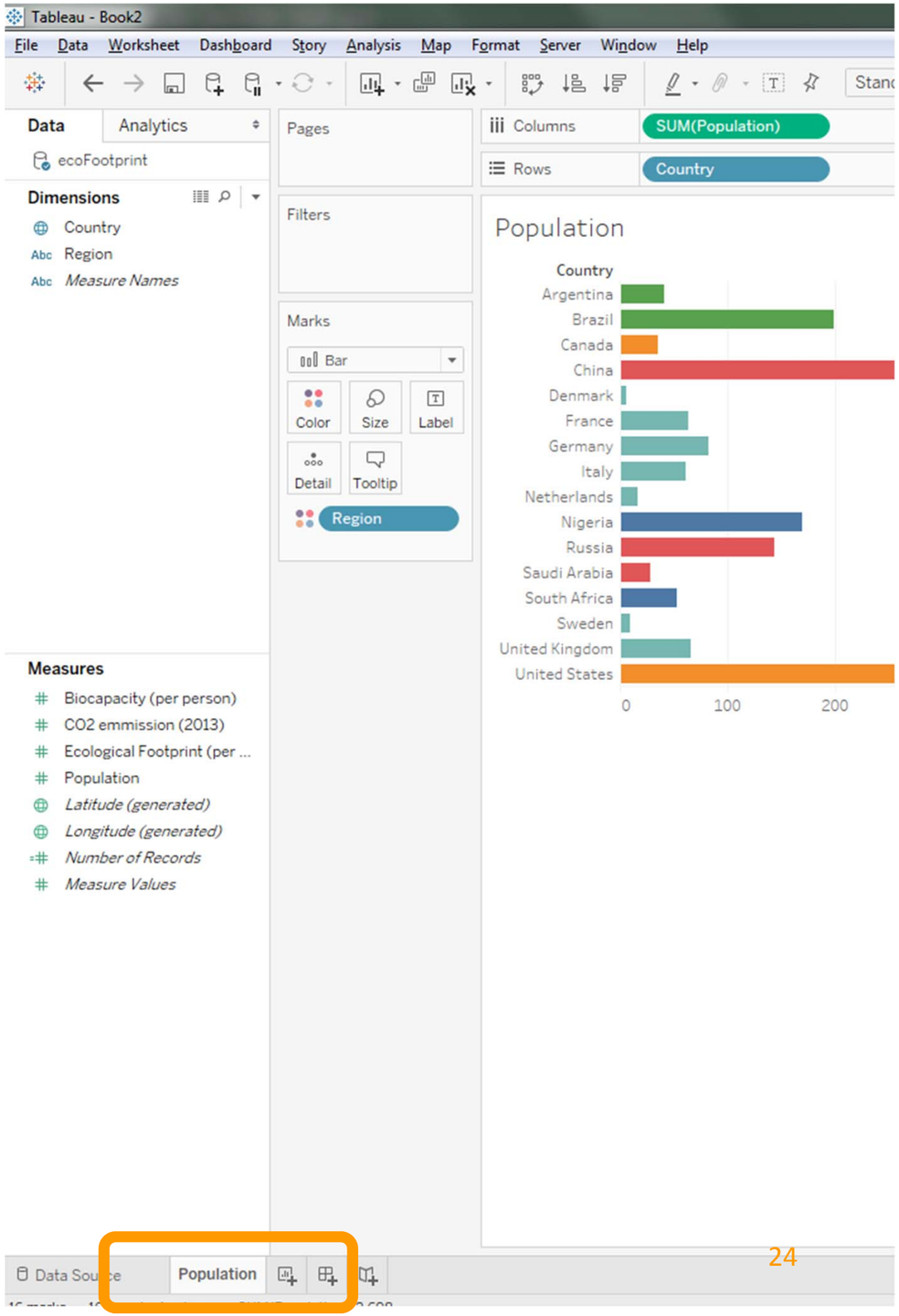
tooltips

- 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.



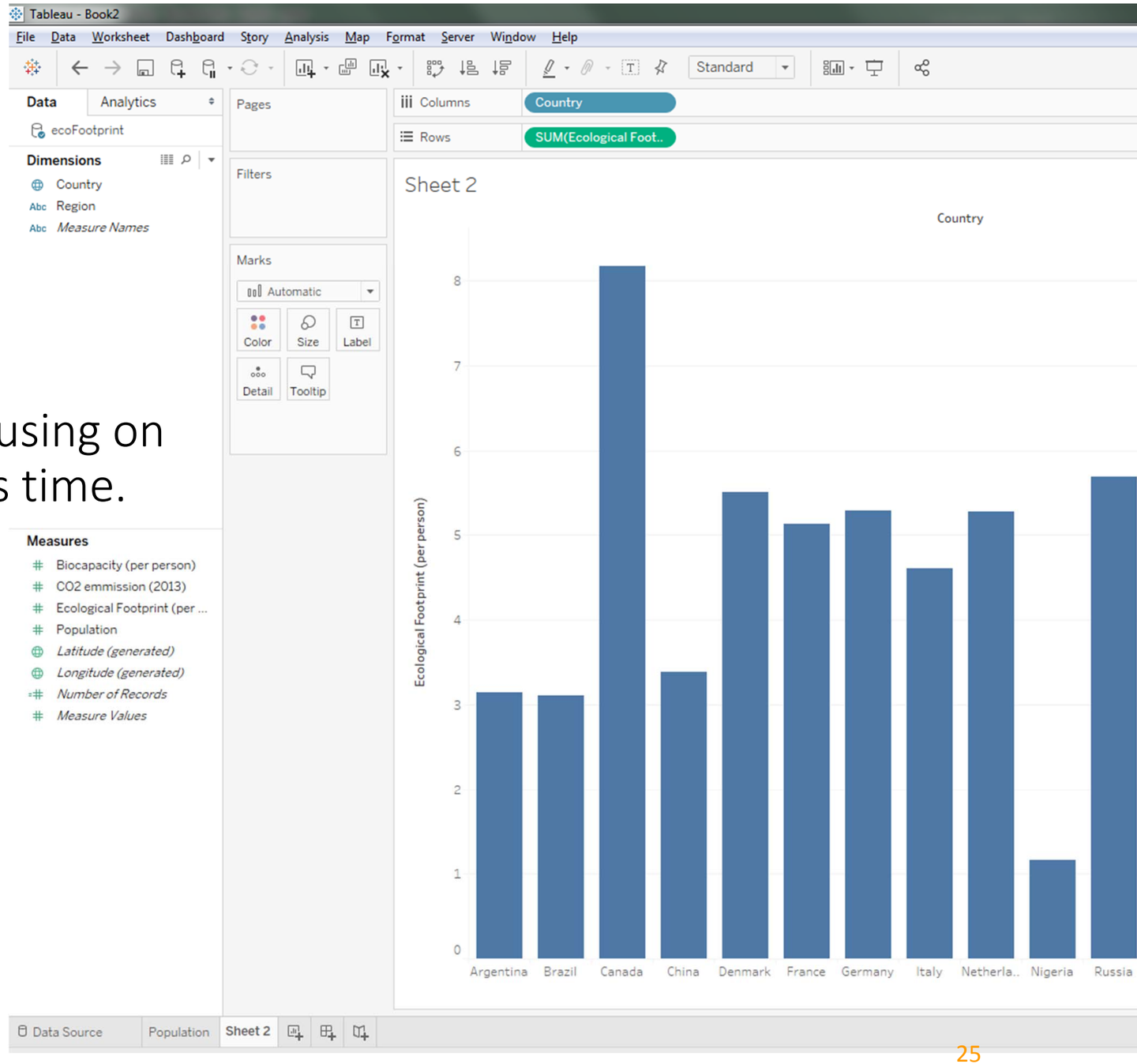
sheet labeling

- 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

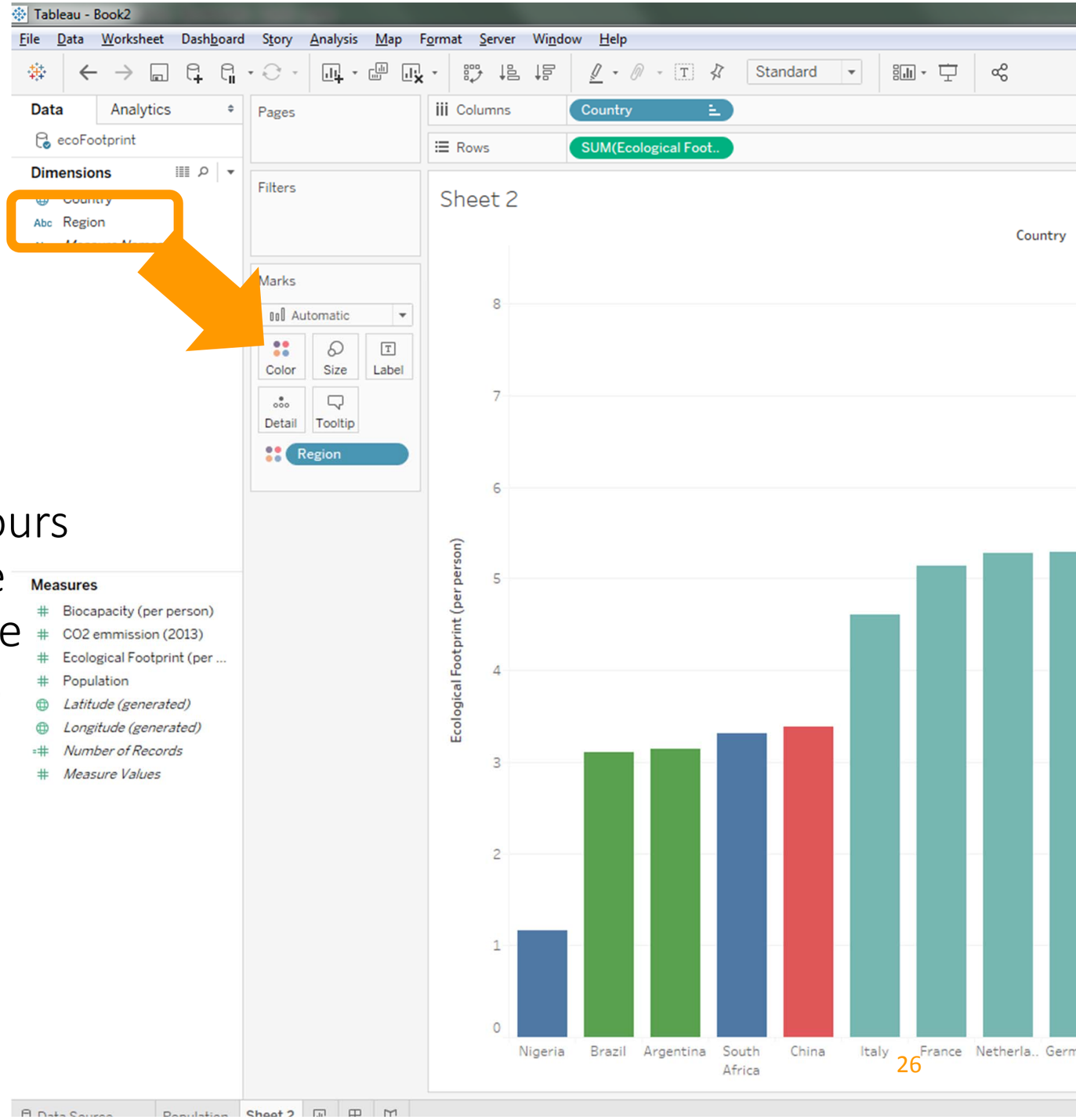
- Let’s create a new sheet for a new visualization
- Let’s create a similar bar chart, focusing on country vs. ecological footprint this time.



country by eco footprint

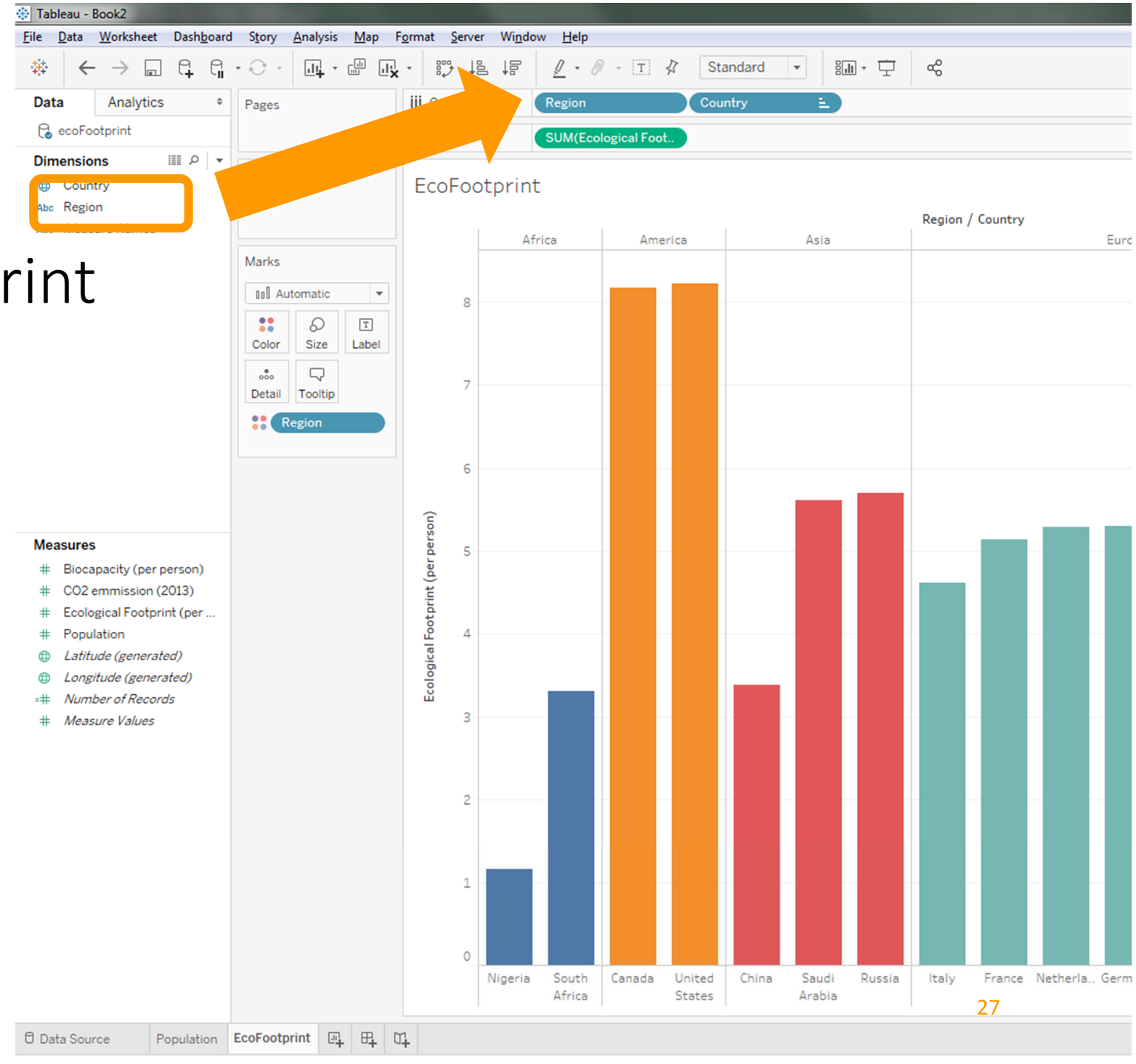
- Again, we choose colour (hue) to distinguish between regions.

- 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



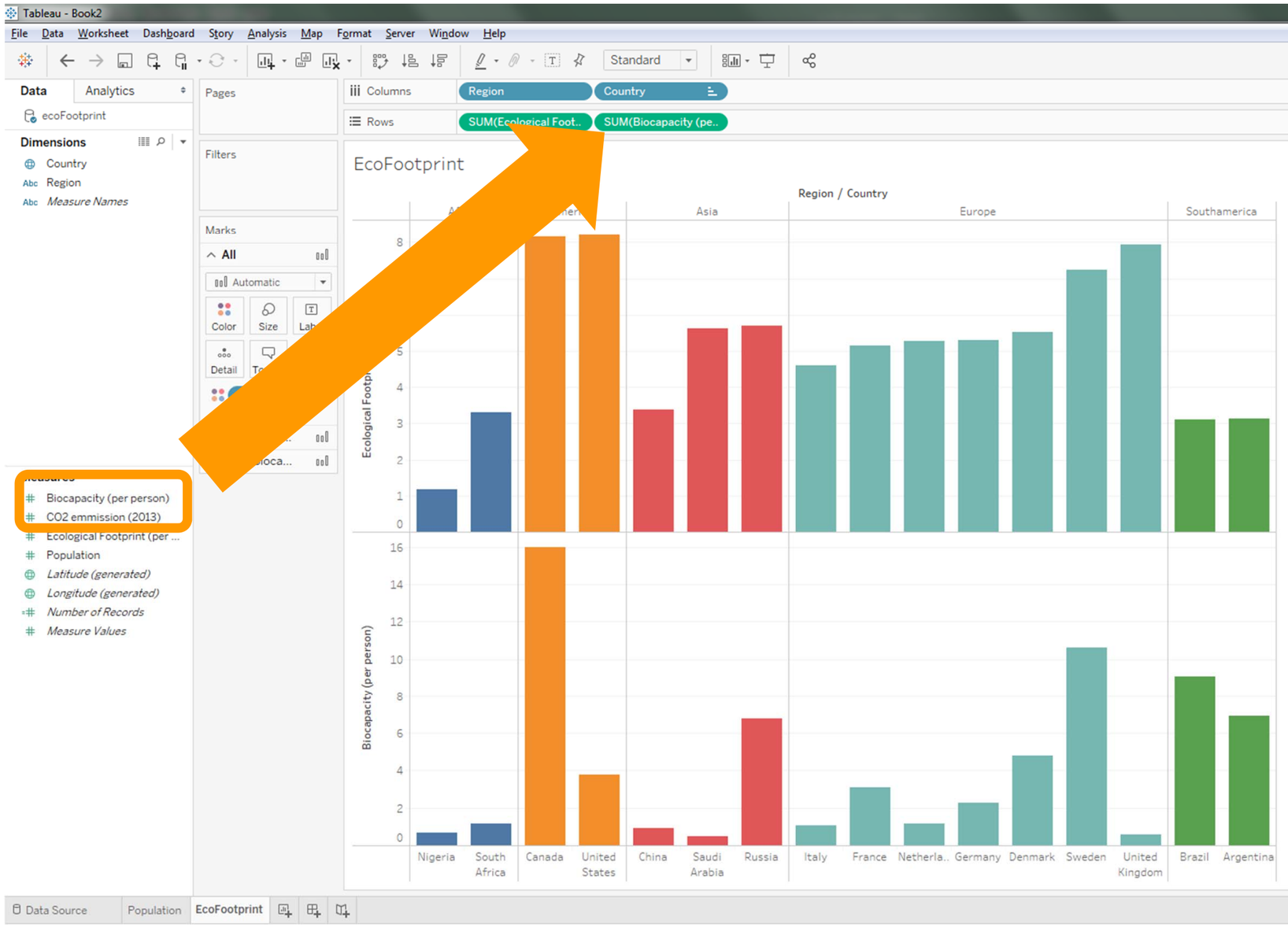
country by eco footprint

- Ordering by region



eco footprint vs. biocapacity

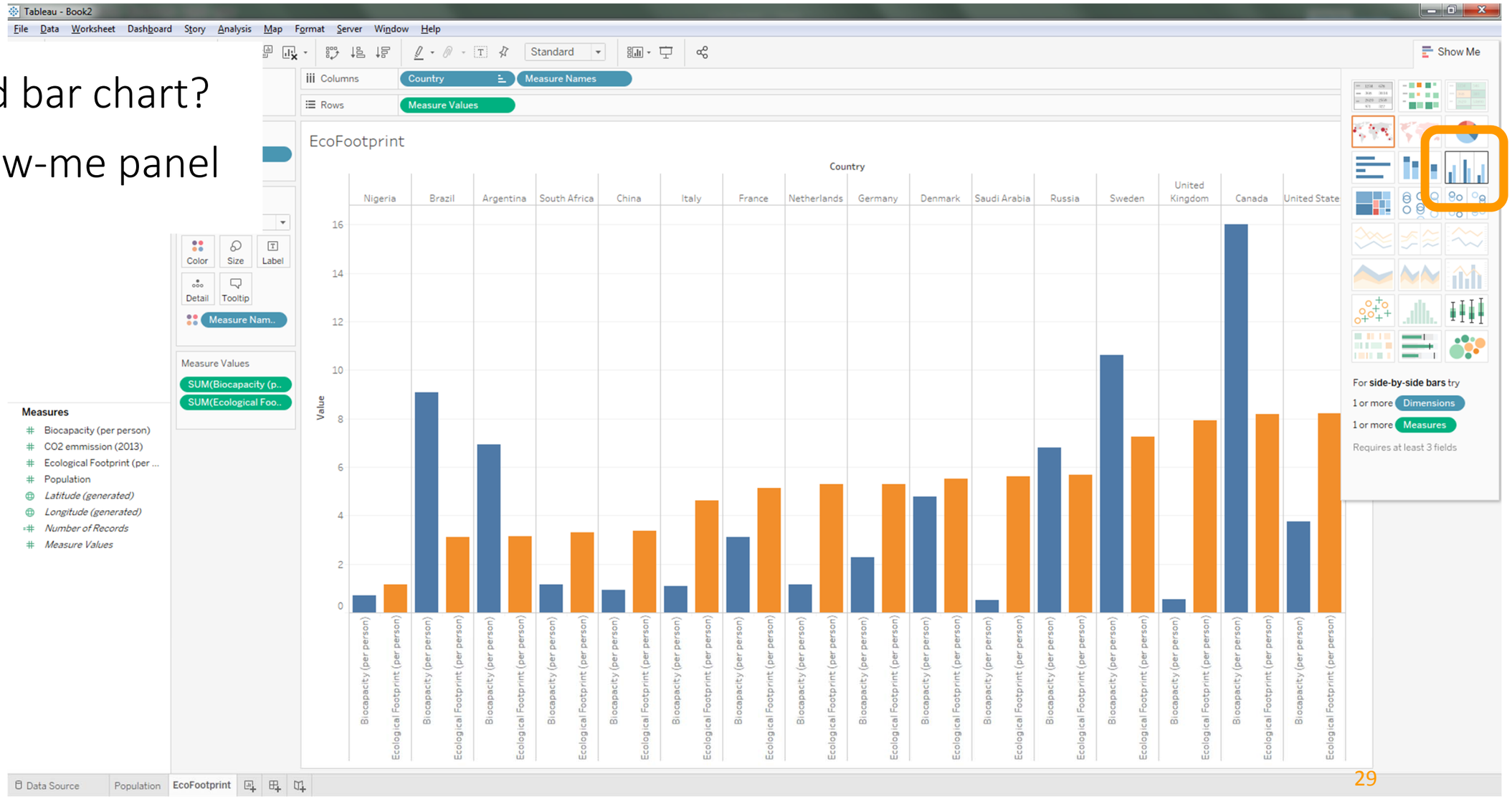
- Two bar graphs?



eco footprint vs. biocapacity

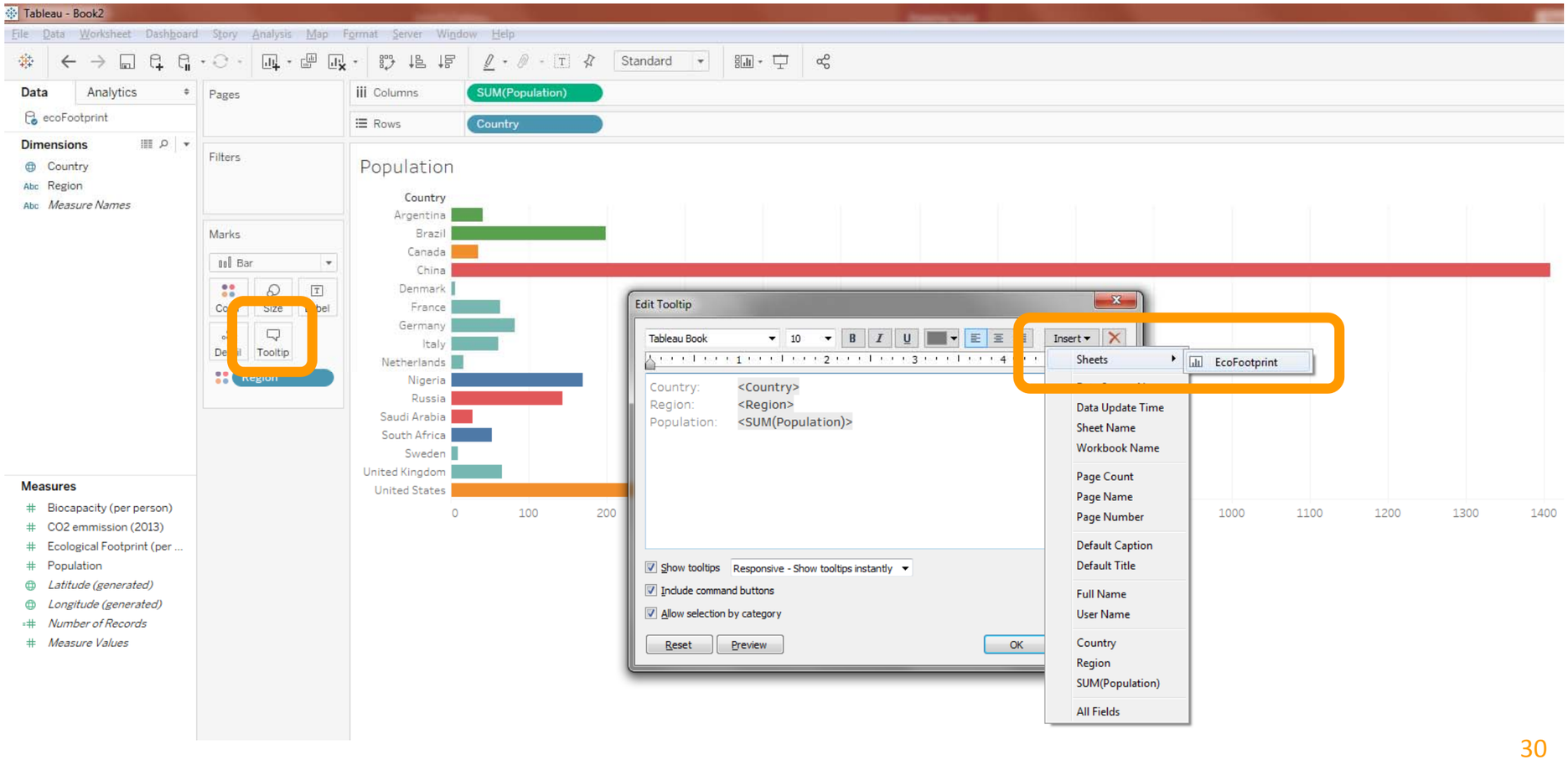
- Grouped bar chart?

→ Via show-me panel



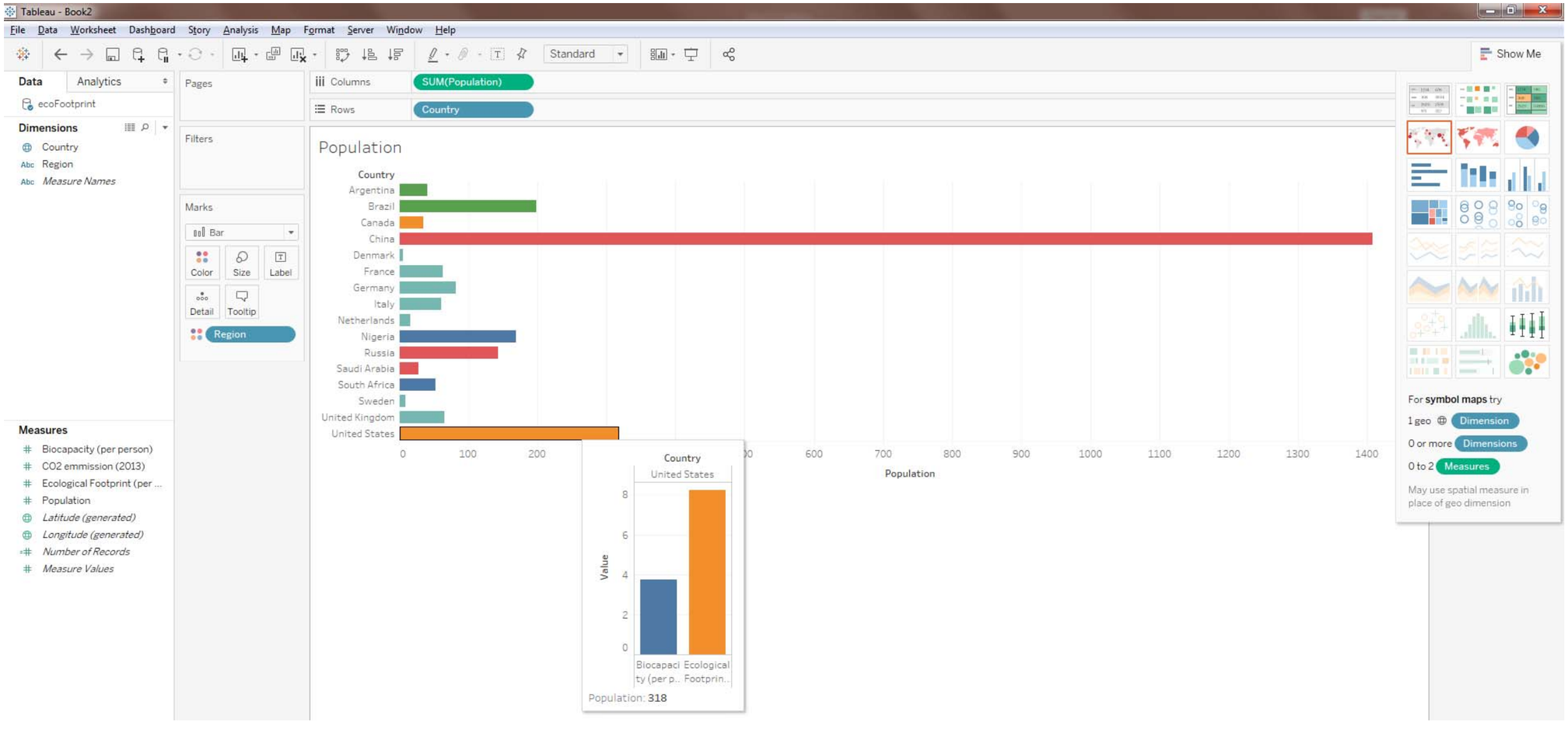
updating tooltips in “Population” visualization

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



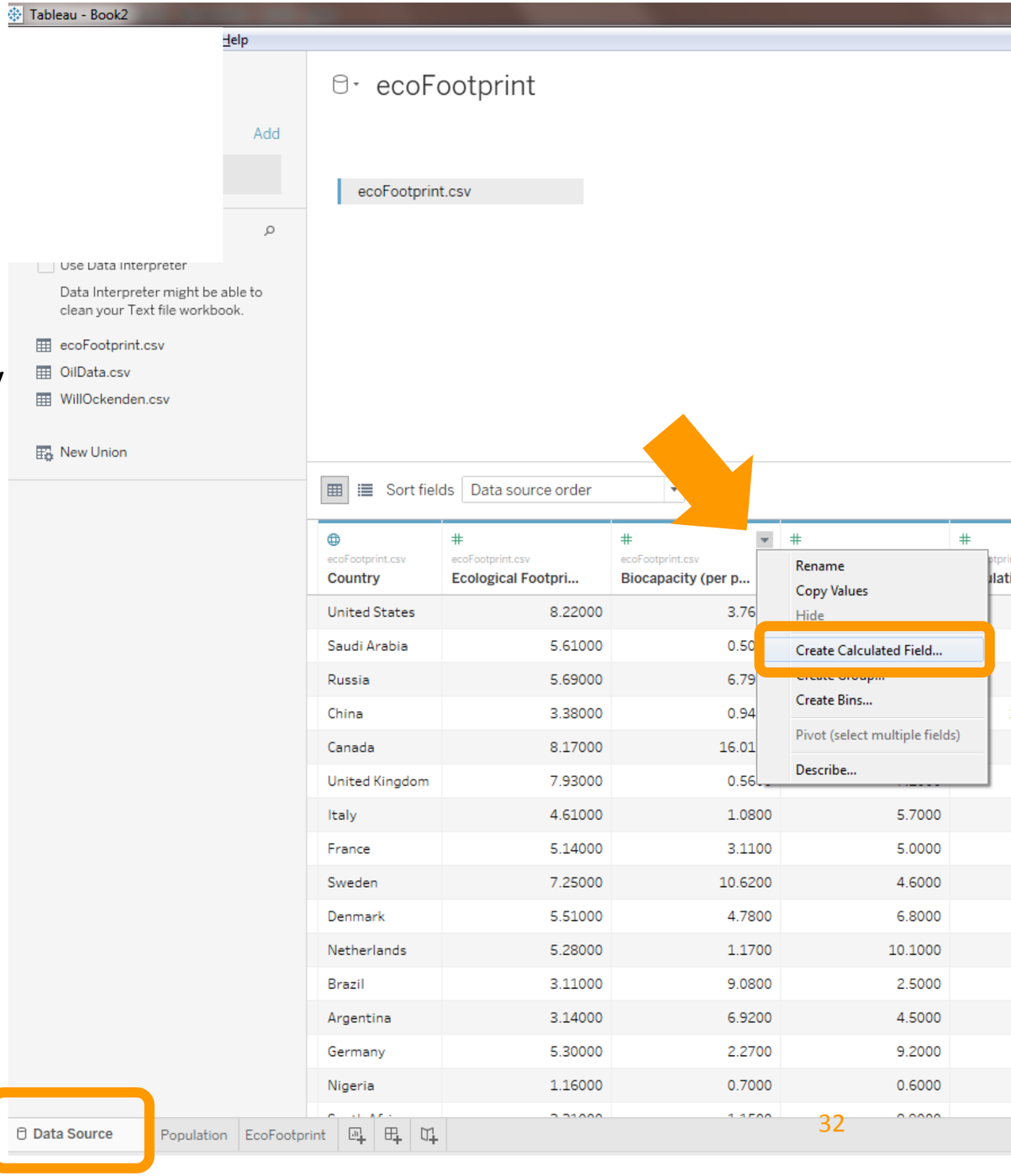
updating tooltips in “Population” visualization

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



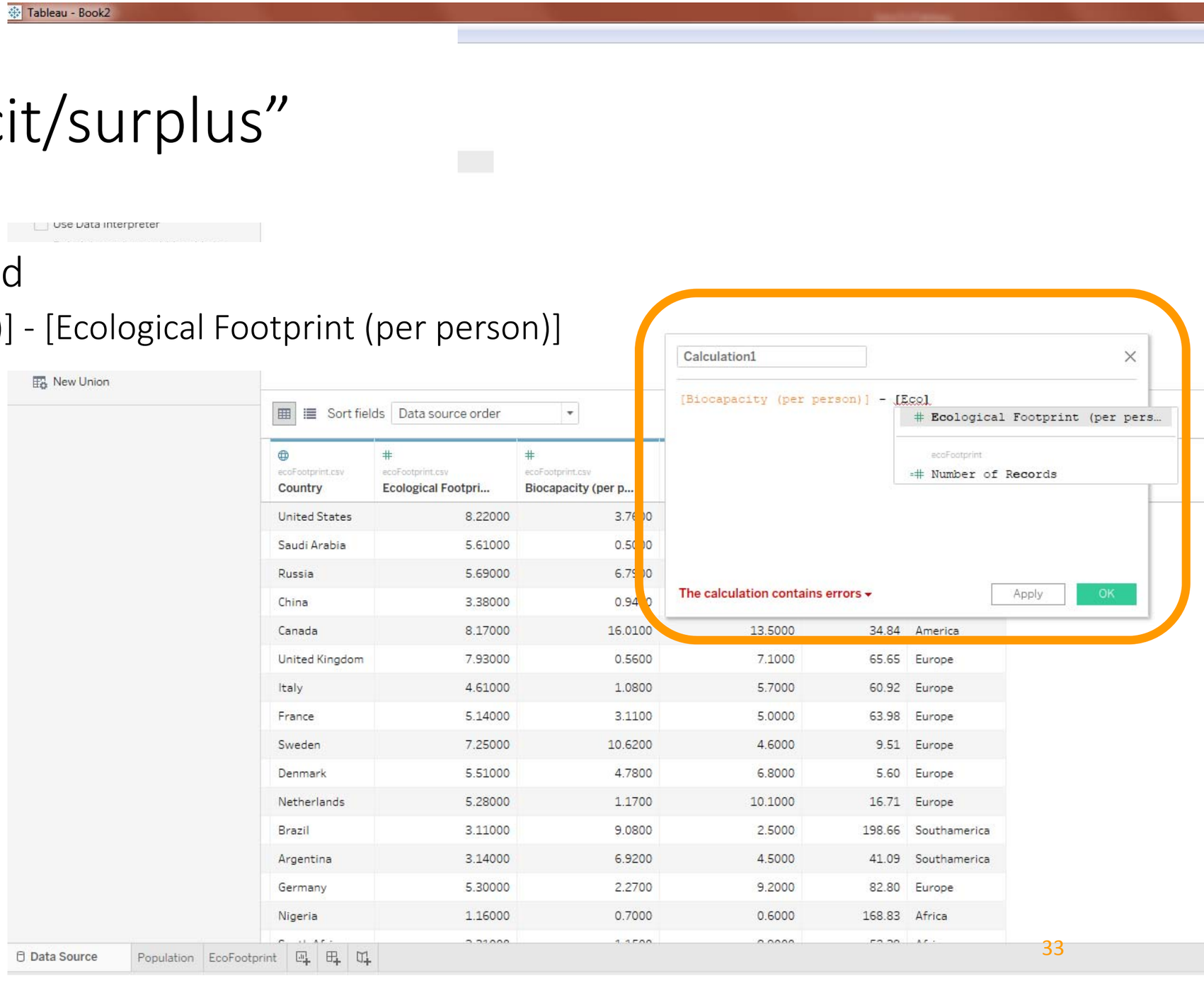
biocapacity “deficit/surplus”

- 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



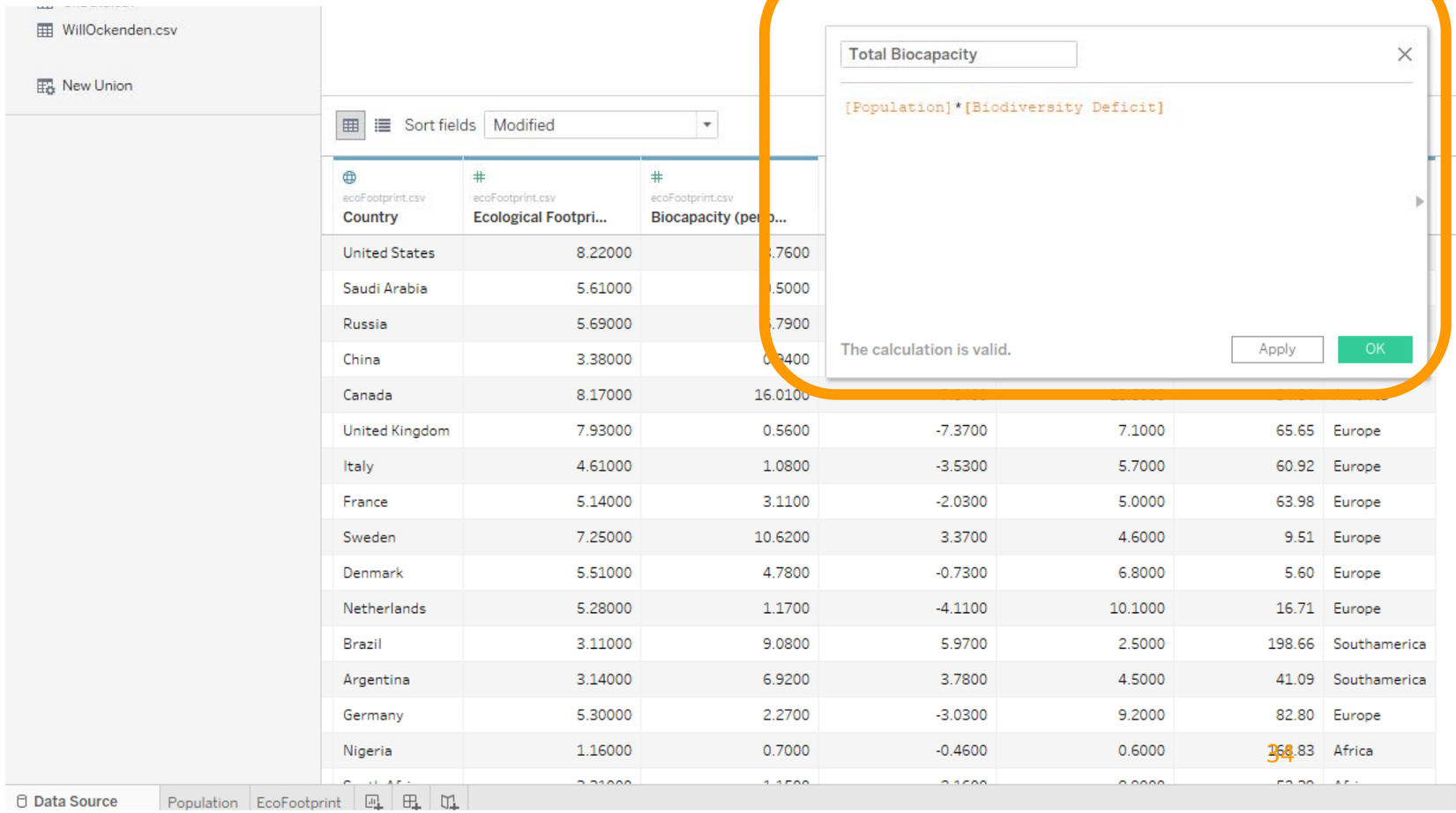
biocapacity “deficit/surplus”

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



considering population size

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



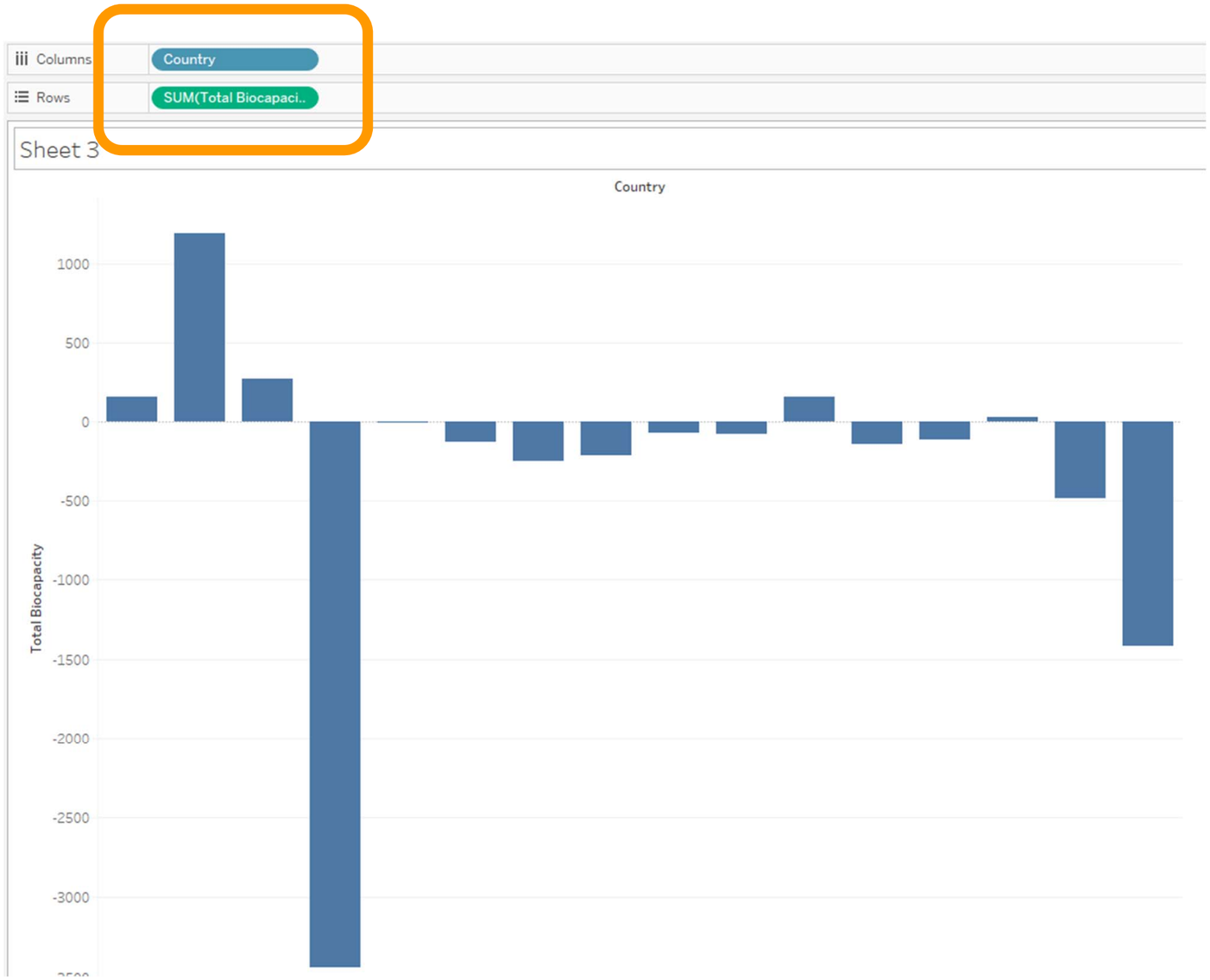
the new data table

Sort fields Modified

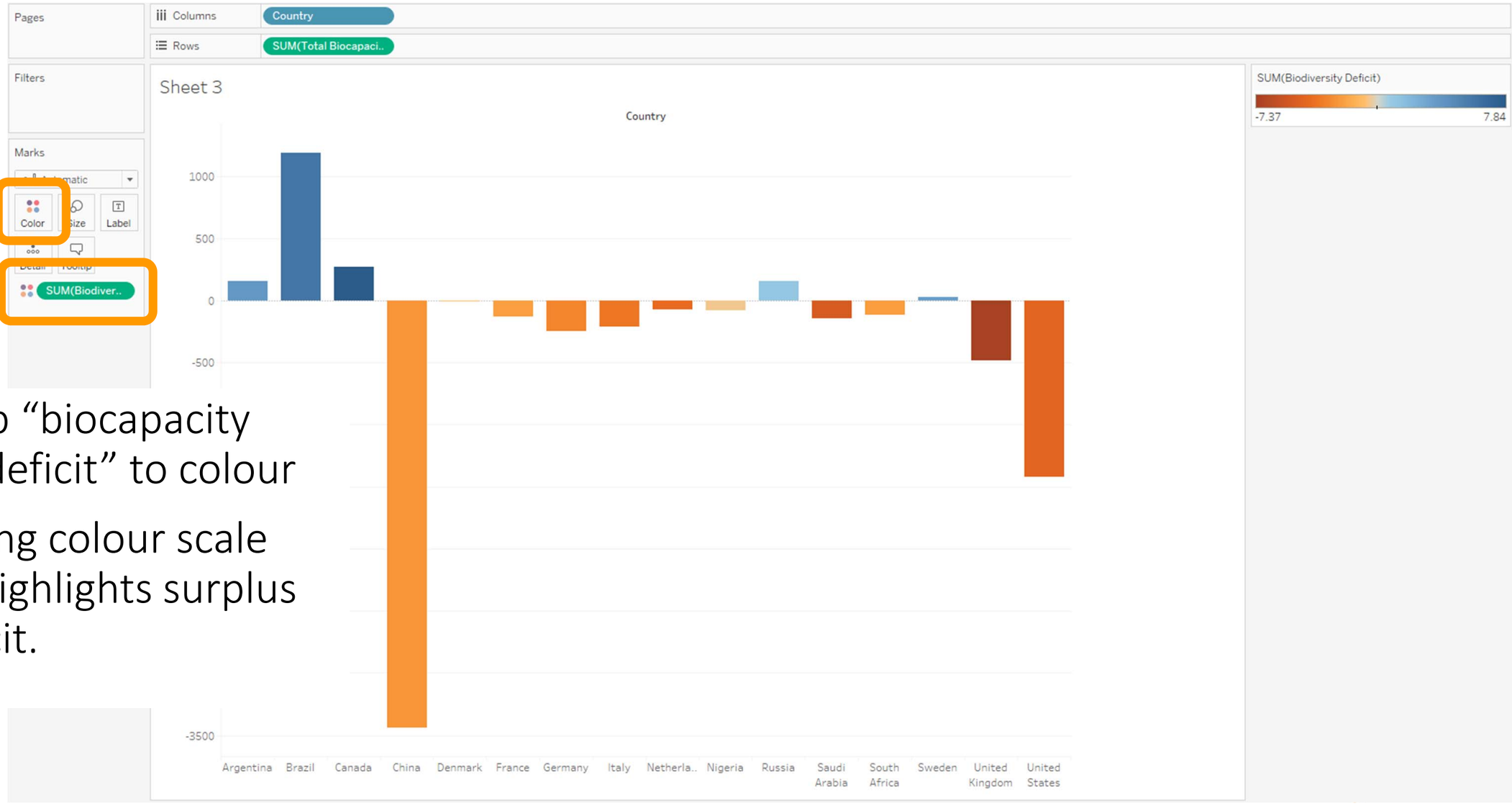
ecoFootprint.csv Country	ecoFootprint.csv Ecological Footpri...	ecoFootprint.csv Biocapacity (per p...	Calculation Biodiversity Deficit	ecoFootprint.csv CO2 emmission (2...	ecoFootprint.csv Population	Calculation Total Biocapacity	ecoFootprint.csv Region
United States	8.22000	3.7600	-4.4600	16.4000	317.50	-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.



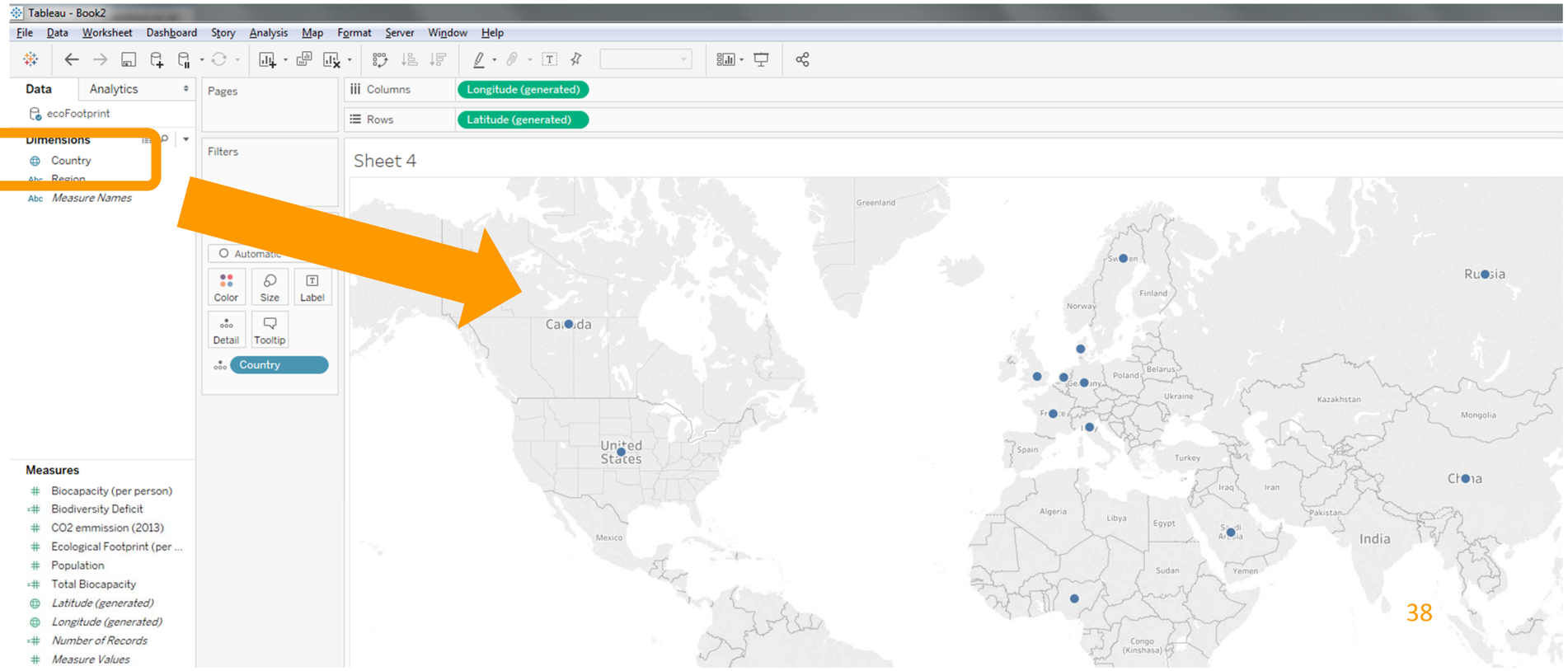
showing total biocapacity surplus/deficit



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

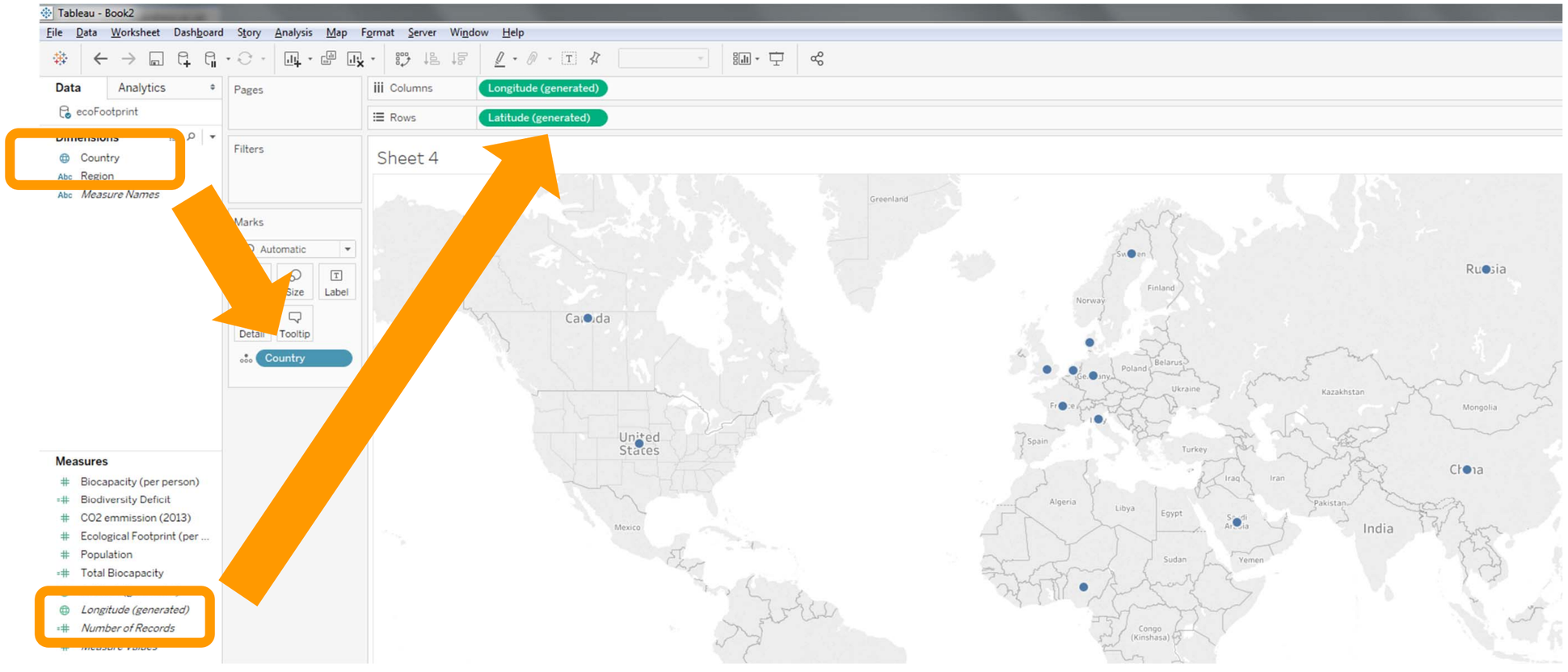
visualizing the data on a geographic map

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



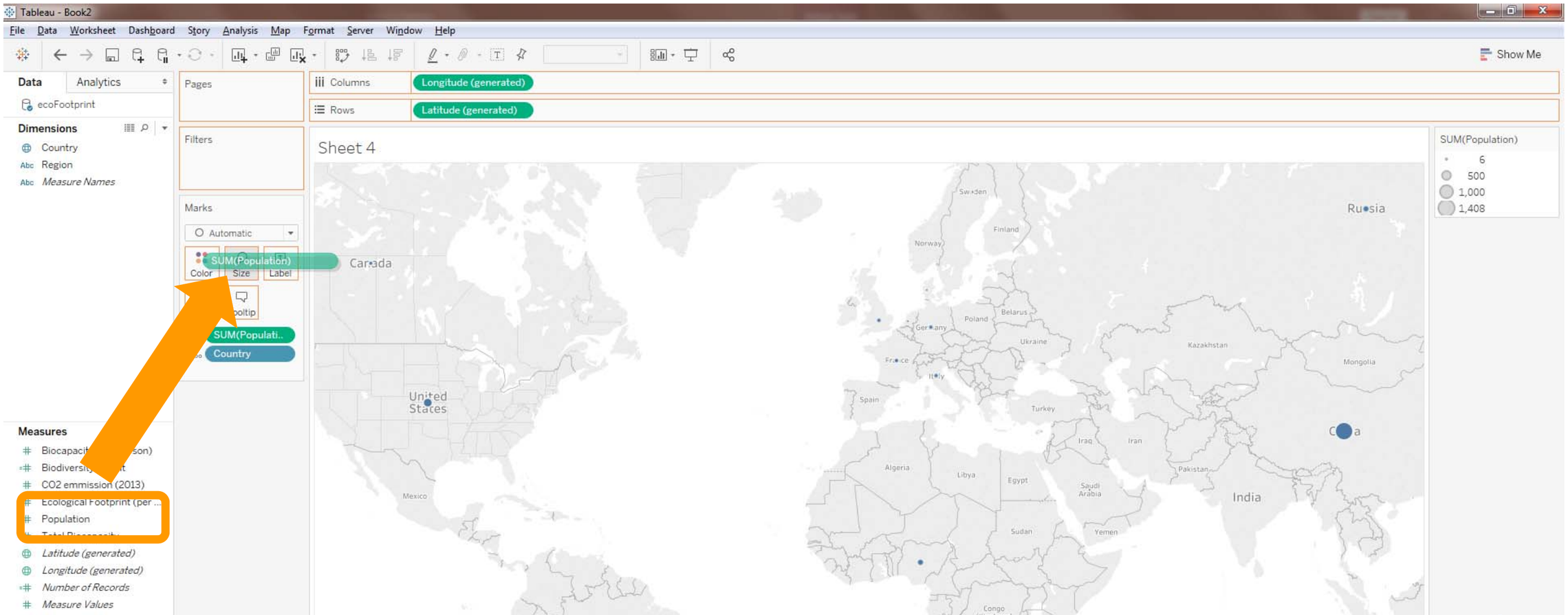
geographic map

- 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.



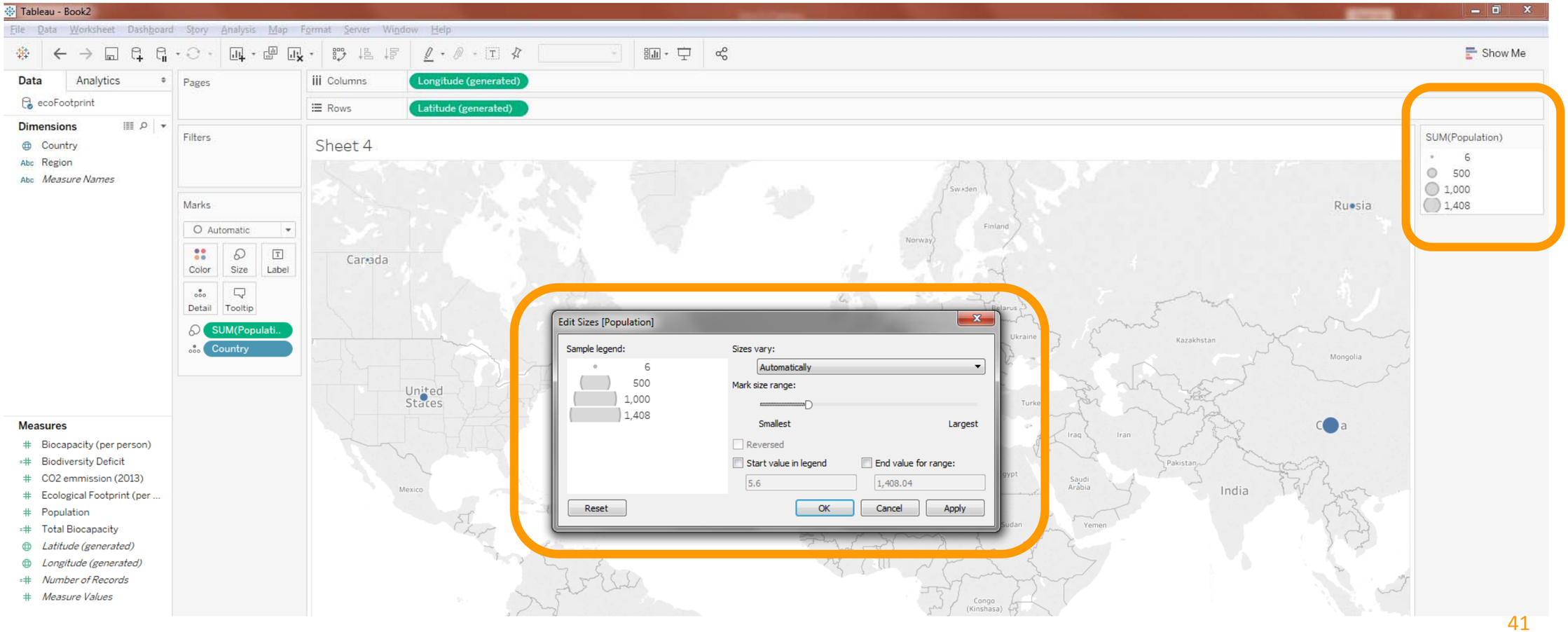
geographic map

- Now show population as circle size.



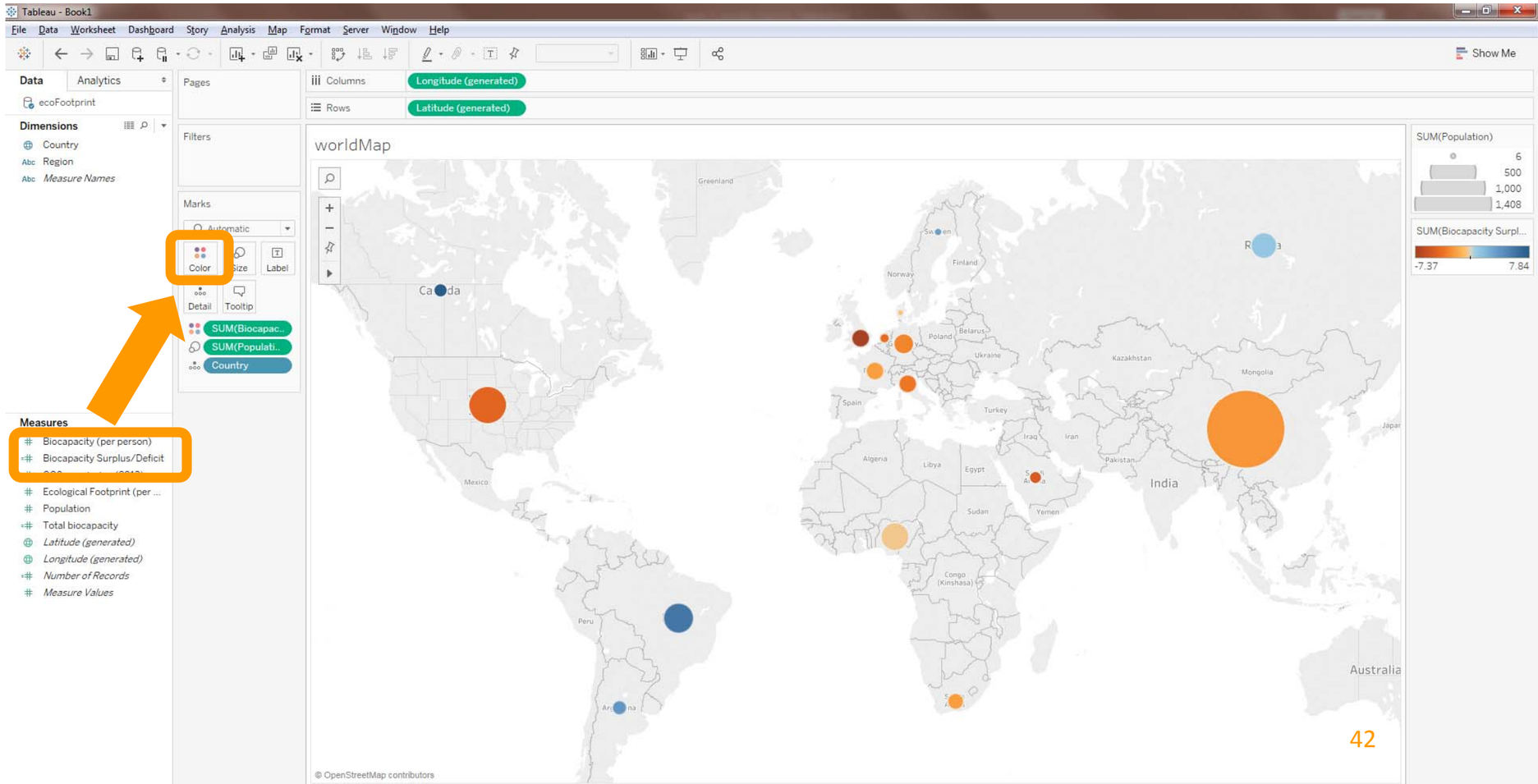
geographic map

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



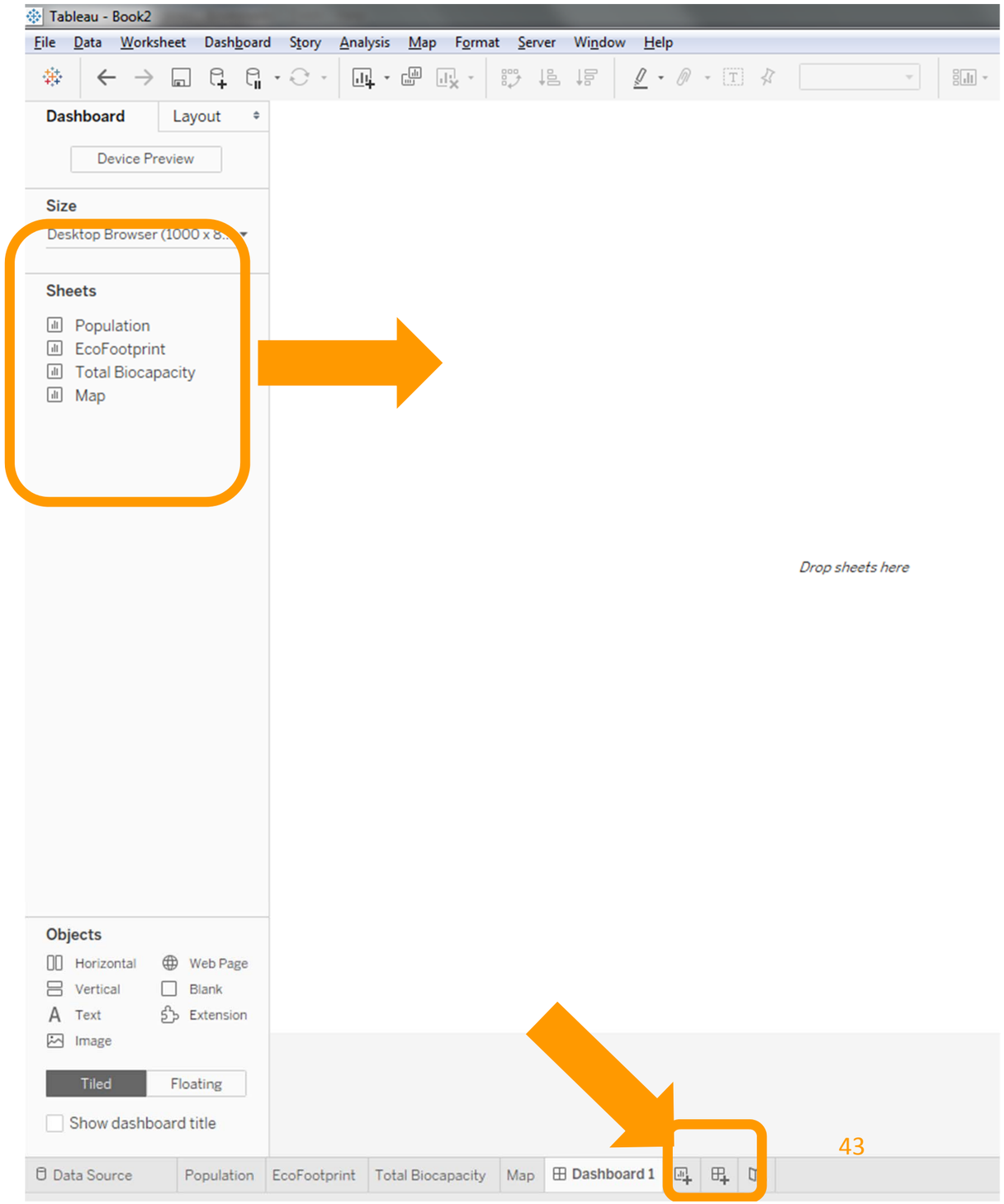
geographic map

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



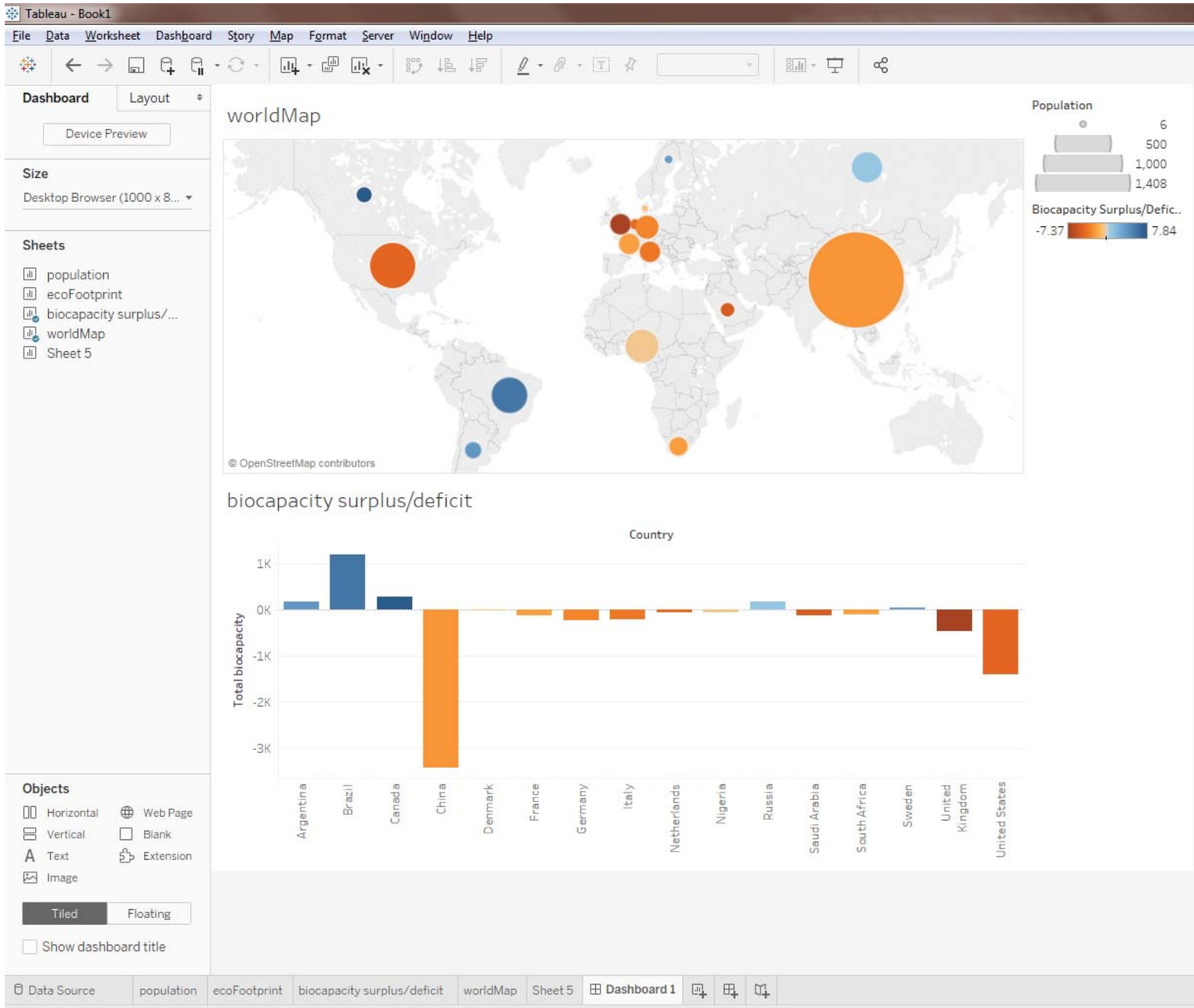
creating a dashboard

- 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.



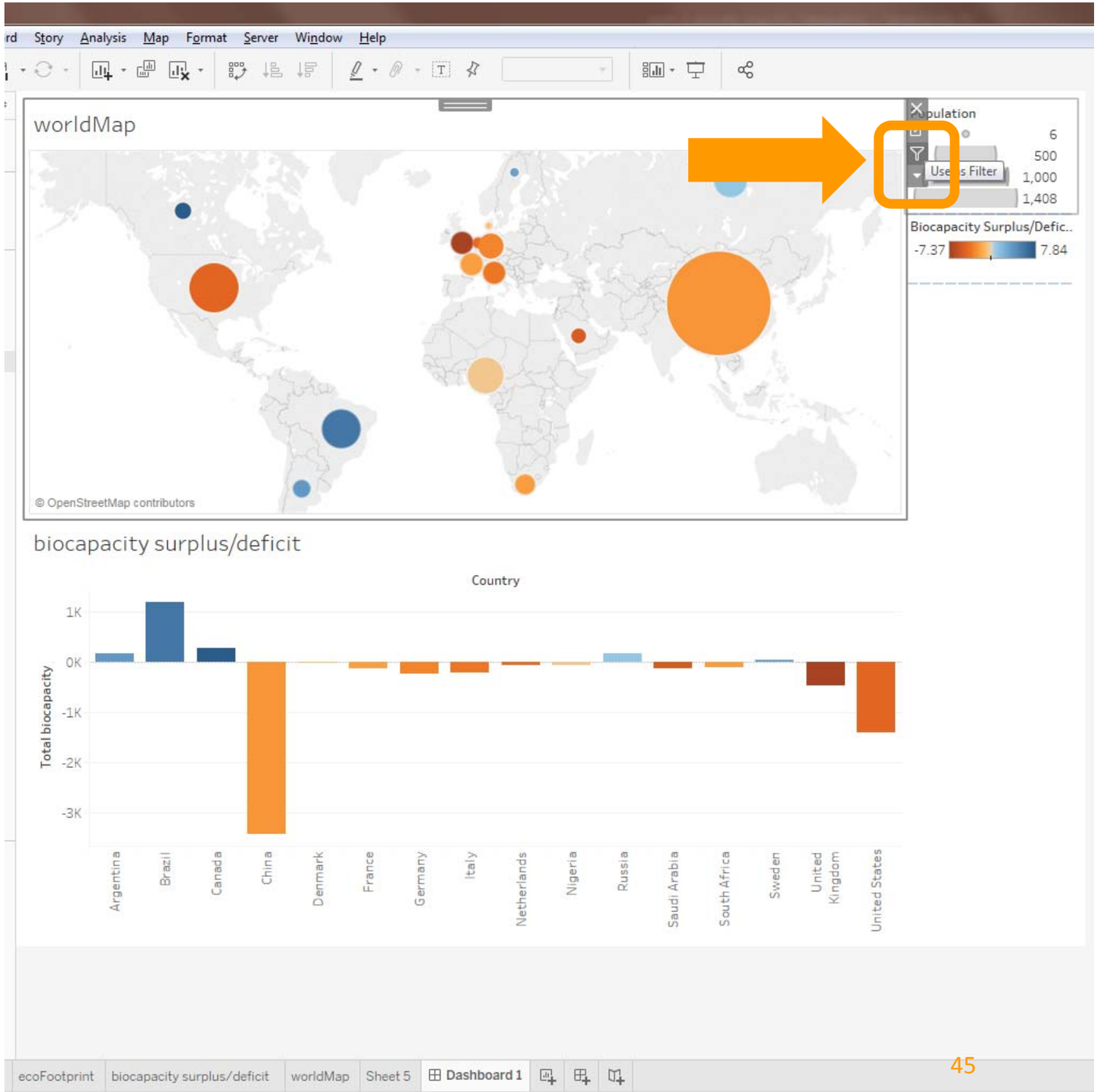
creating a dashboard

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



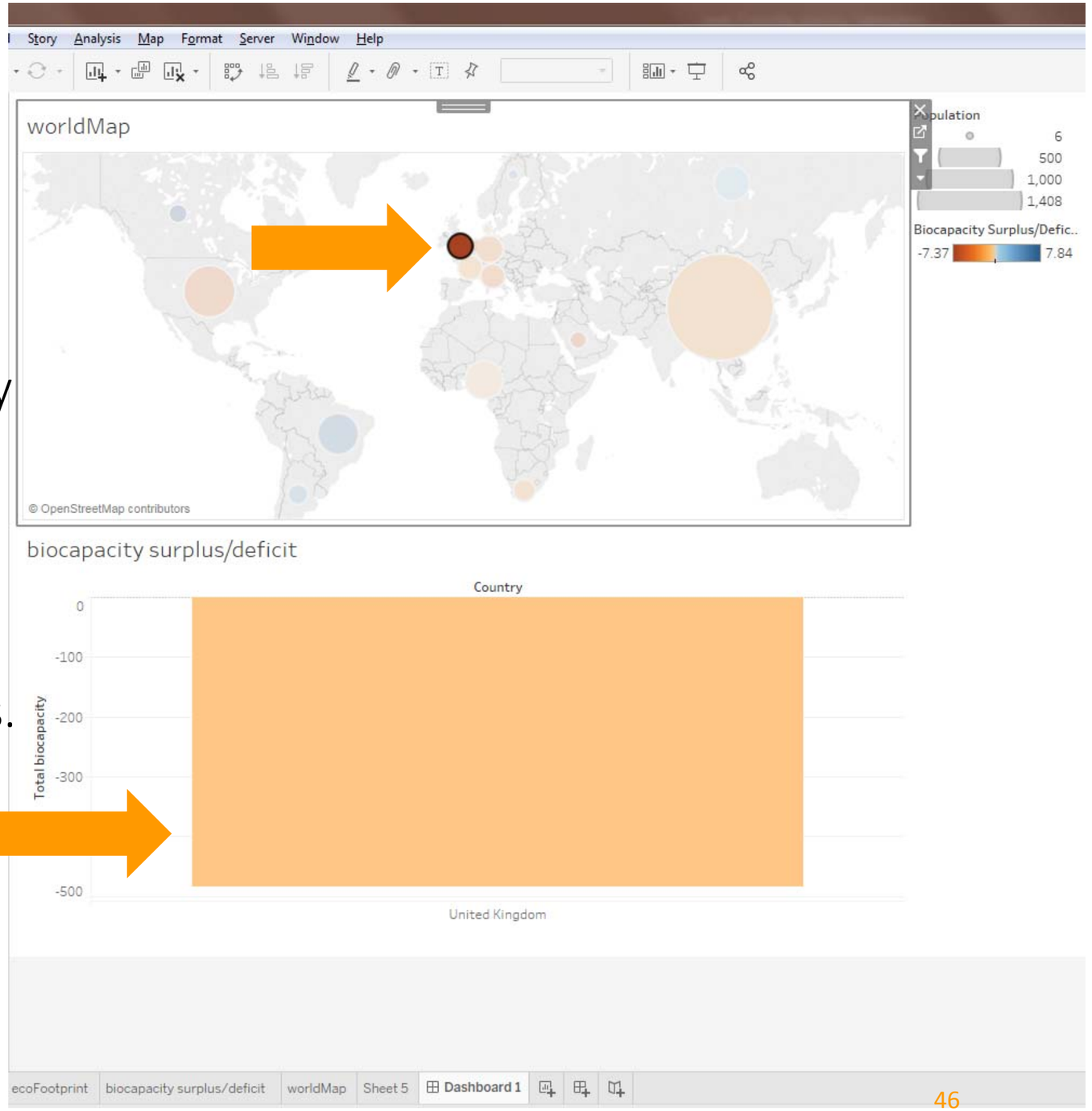
interactive filters

- 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.



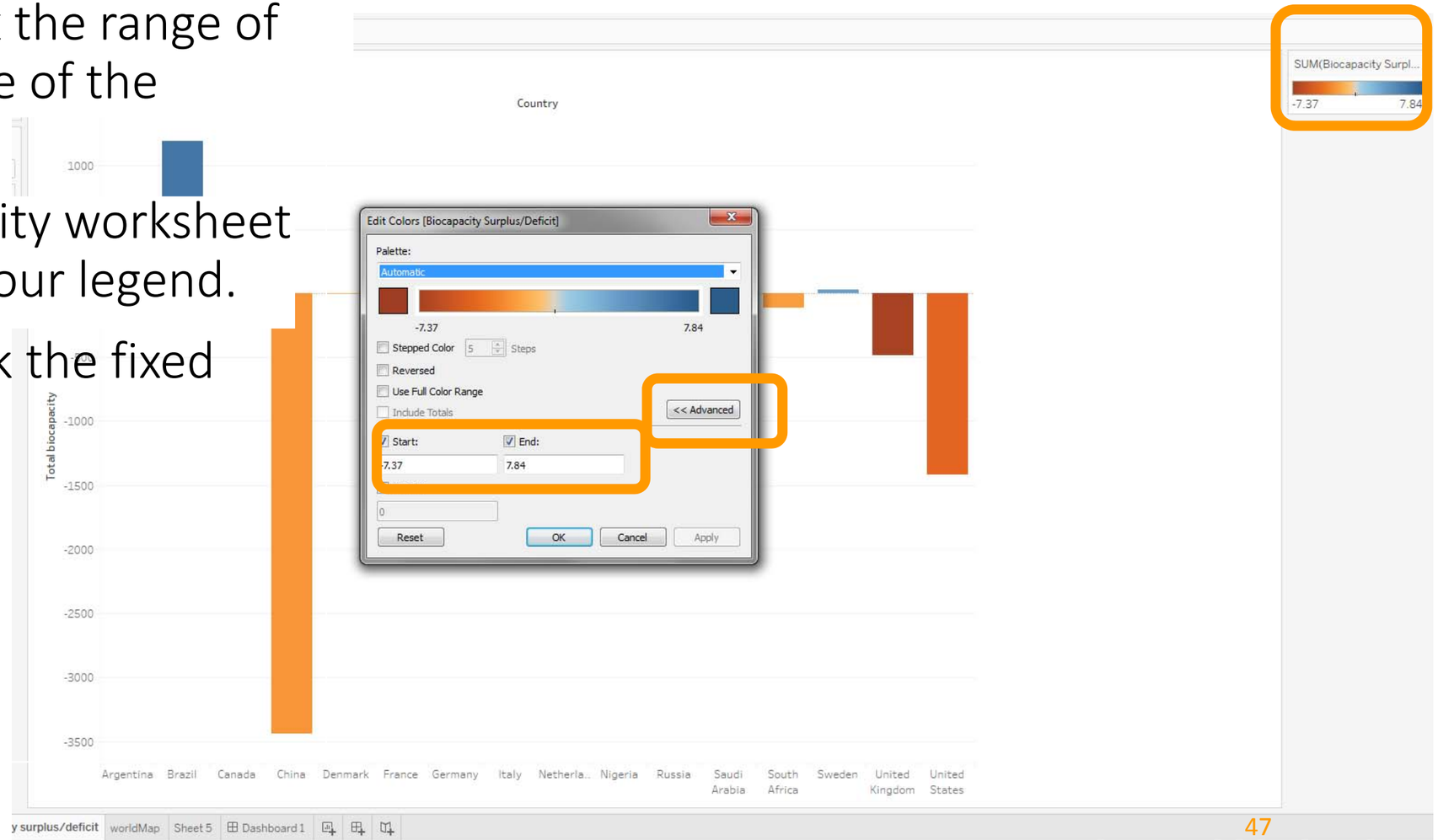
inconsistent colour scales

- 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

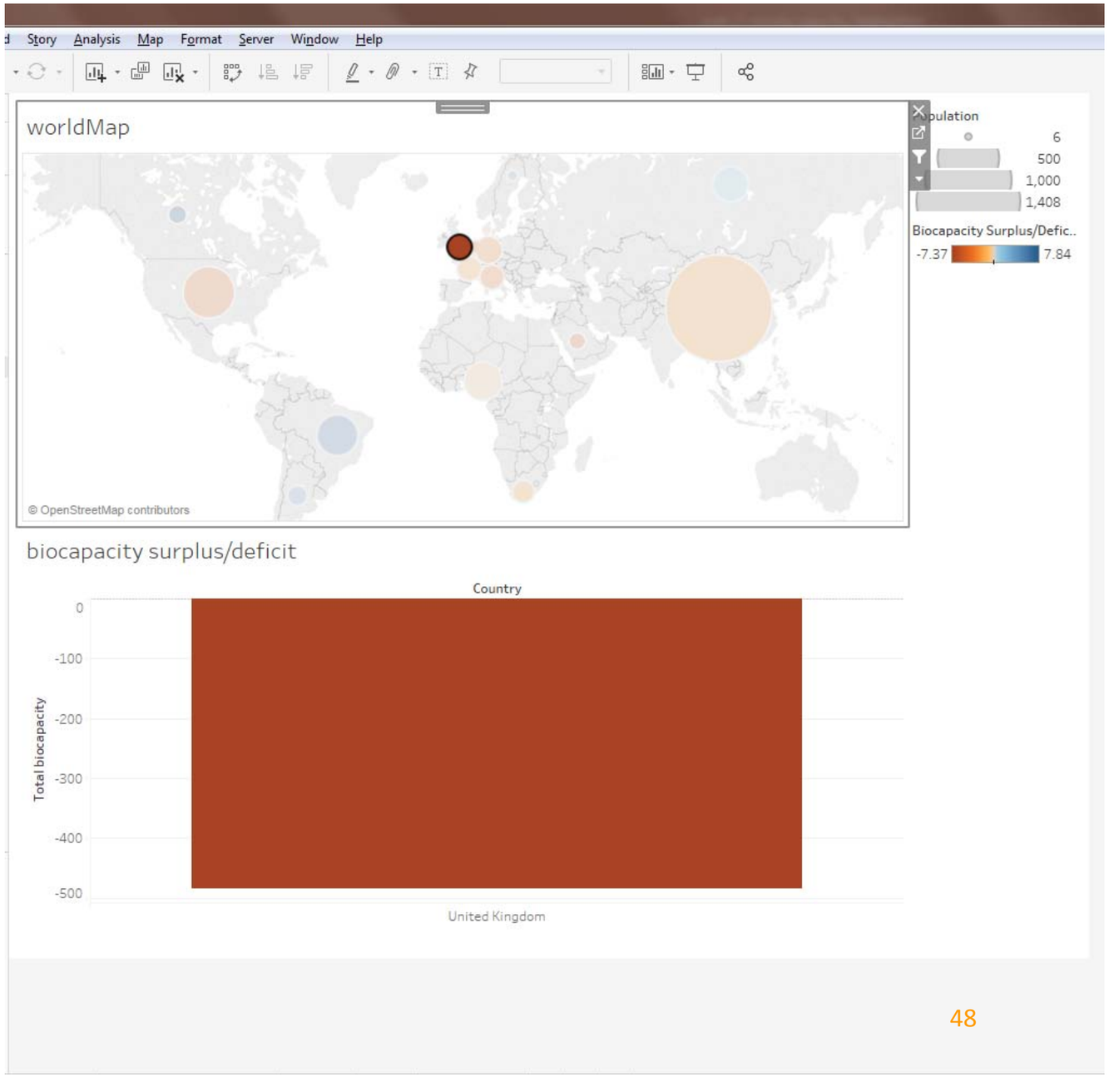
- To fix this, we have to fix the range of values in the colour scale of the Biocapacity chart.
- Go back to the Biocapacity worksheet and double-click the colour legend.
- Under “Advanced” check the fixed ranges.
- Click “Apply”



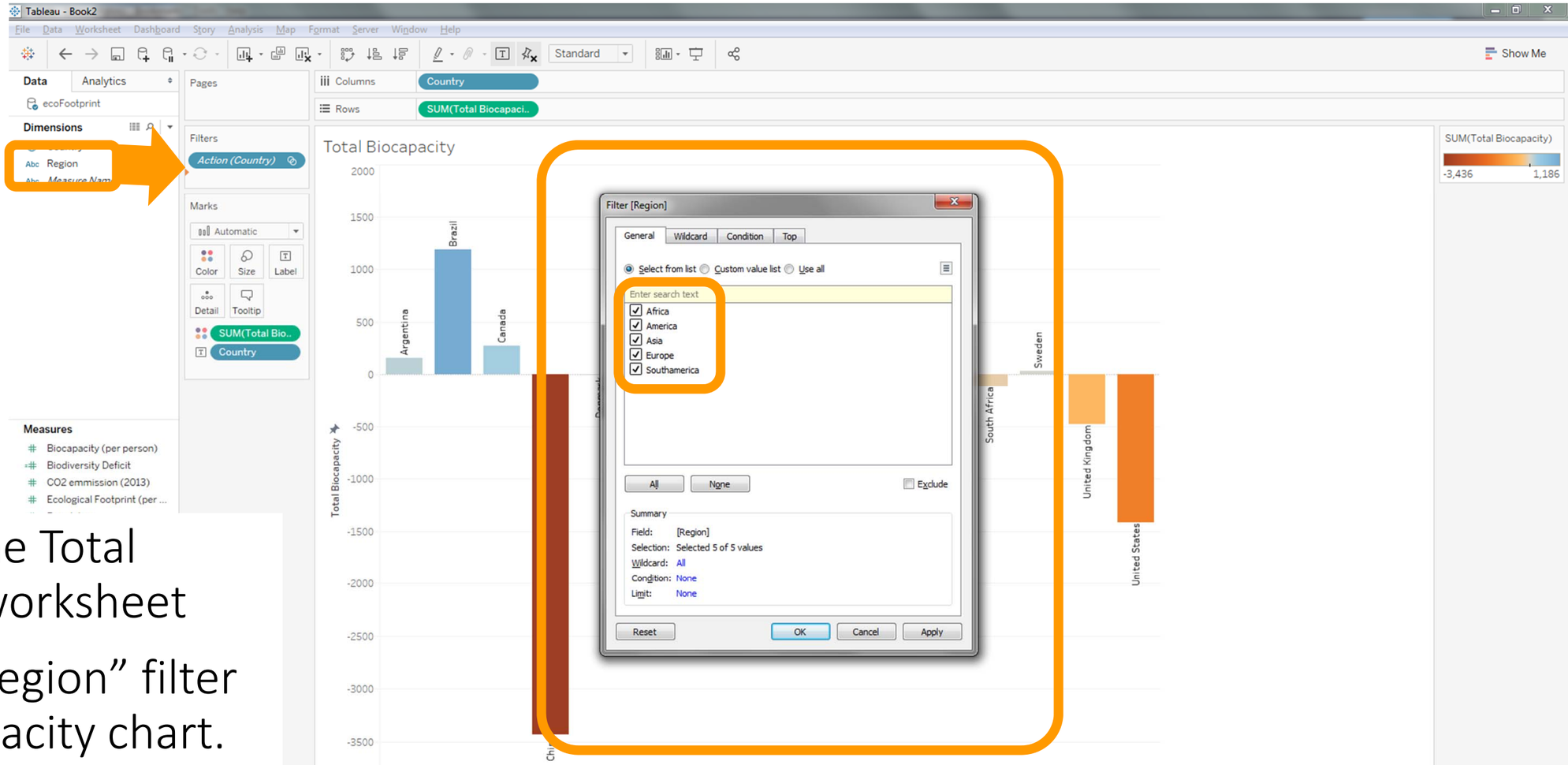
inconsistent colour scales

- 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.



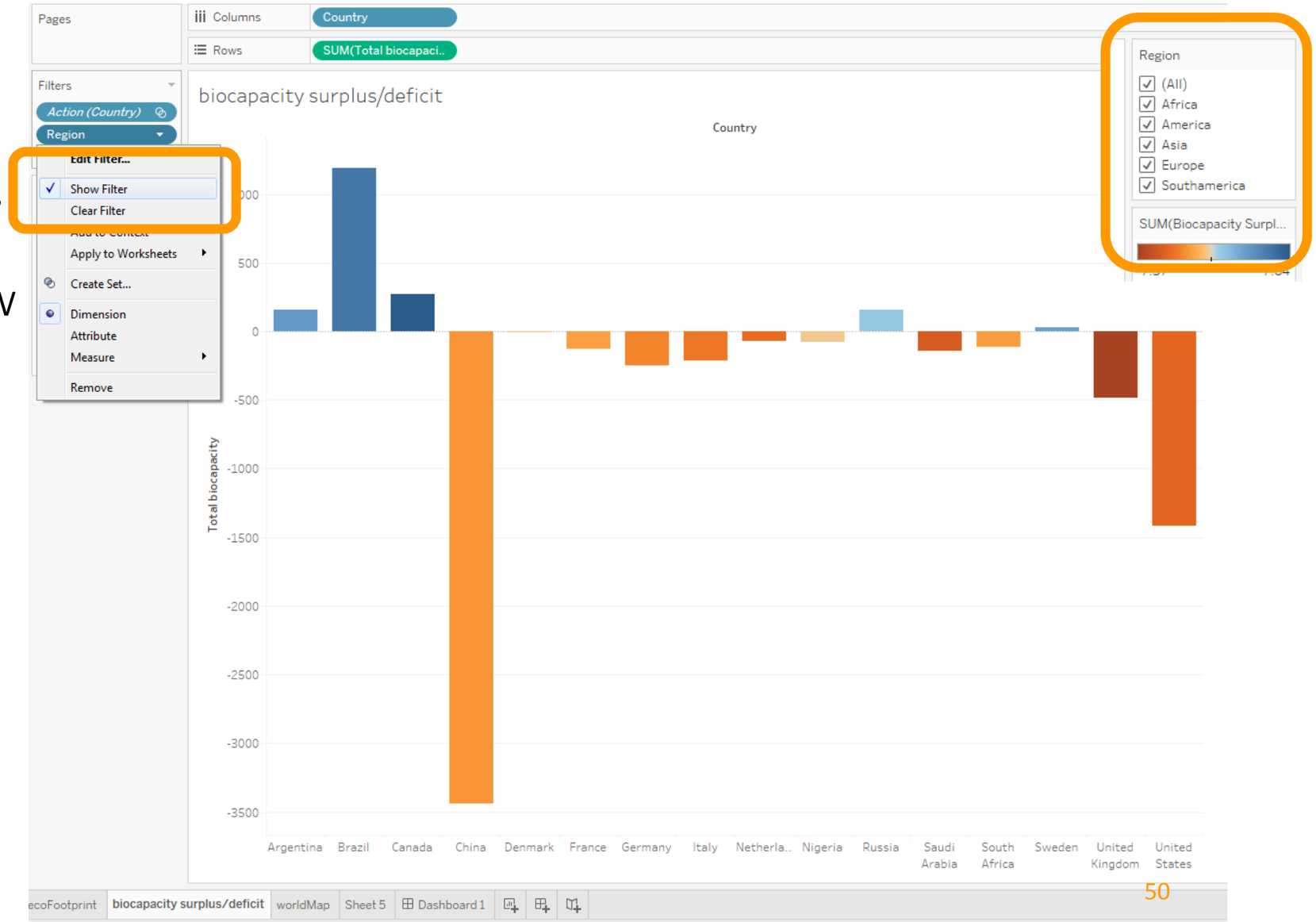
applying additional filters



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

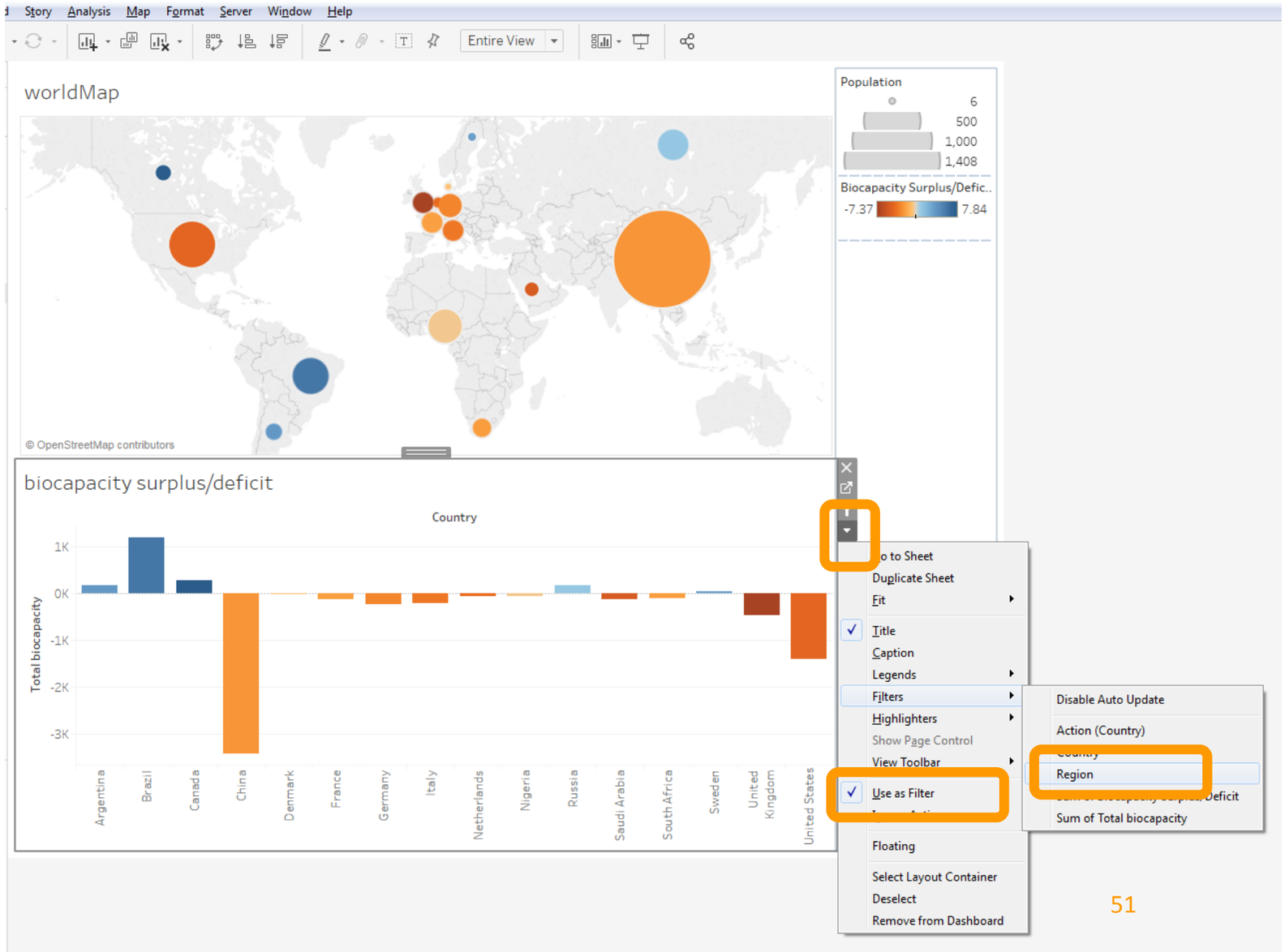
applying additional filters

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



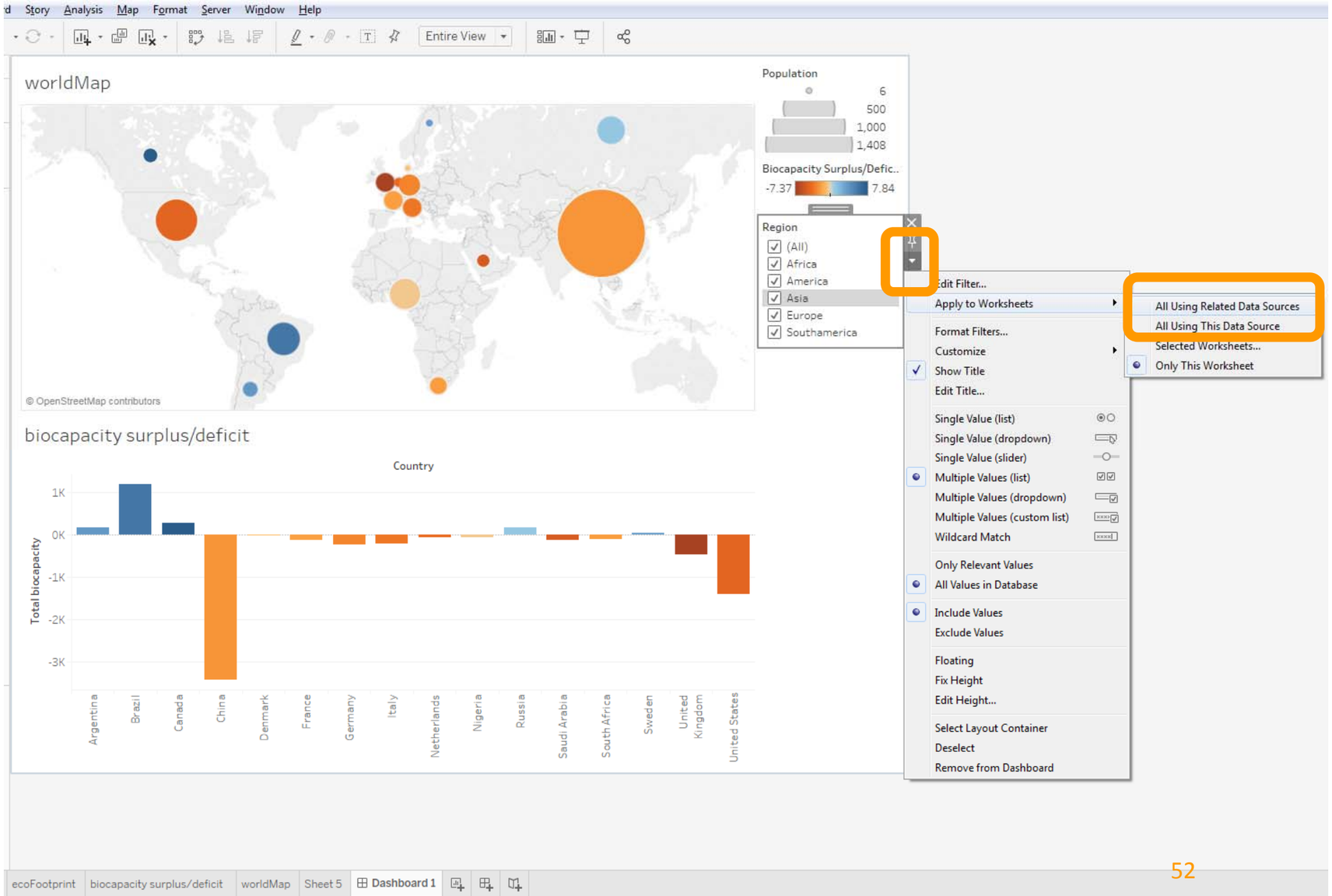
applying additional filters

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

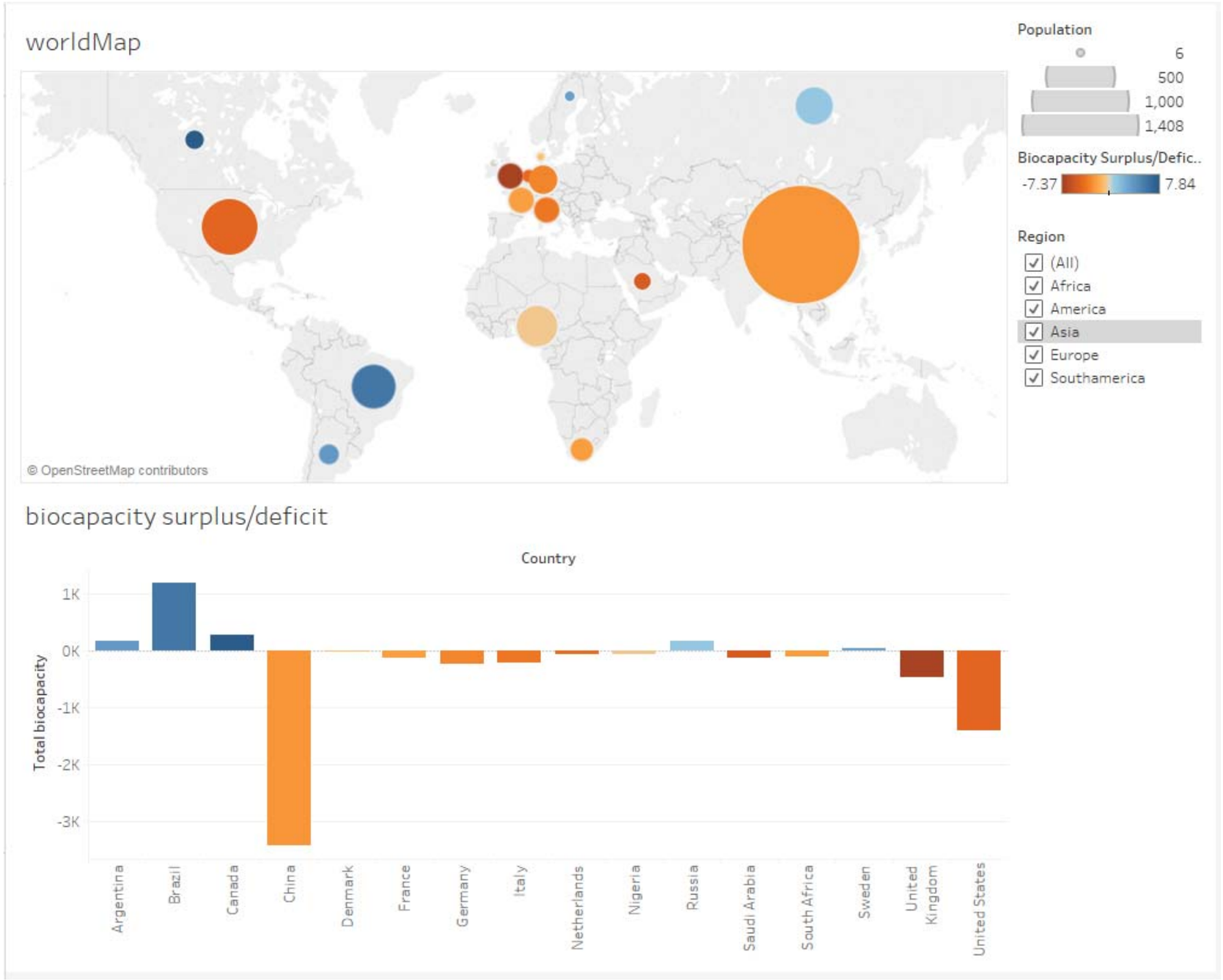


applying additional filters

- 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


worldMap
© OpenStreetMap contributors
Population
6
500
1,000
1,408
Biocapacity Surplus/Defic..
-7.37
7.84
Region
(All)
Africa
America
Asia
Europe
Southamerica
biocapacity surplus/deficit
Country
Total biocapacity
1K
0K
-1K
-2K
-3K
Argentina
Brazil
Canada
China
Denmark
France
Germany
Italy
Netherlands
Nigeria
Russia
Saudi Arabia
South Africa
Sweden
United Kingdom
United States

additional exercises, tutorials, data, resources

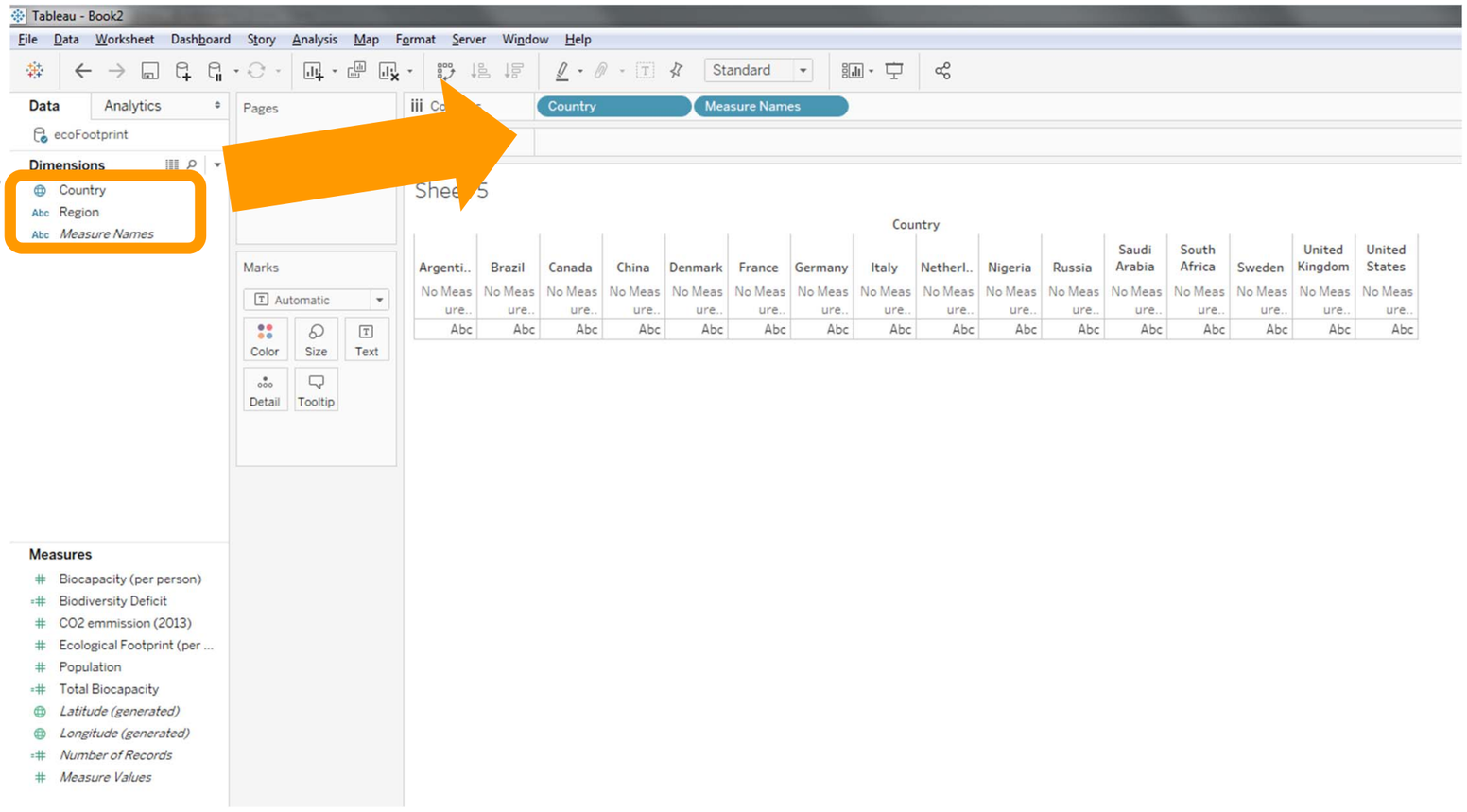
- Tableau_visExercise_01
  - Try to visualise this small dataset in your own time
- How-to instructions in on studres: CS5044/Tutorials/Tableau/tableau_howTo
  - Manipulating data
  - Interactive elements
  - Visualising geospatial data
- More general Tableau tutorials on resource list
- Lots of resources on the web!

- Next week Monday (Week 4) – Tableau tutorial II

Additional material

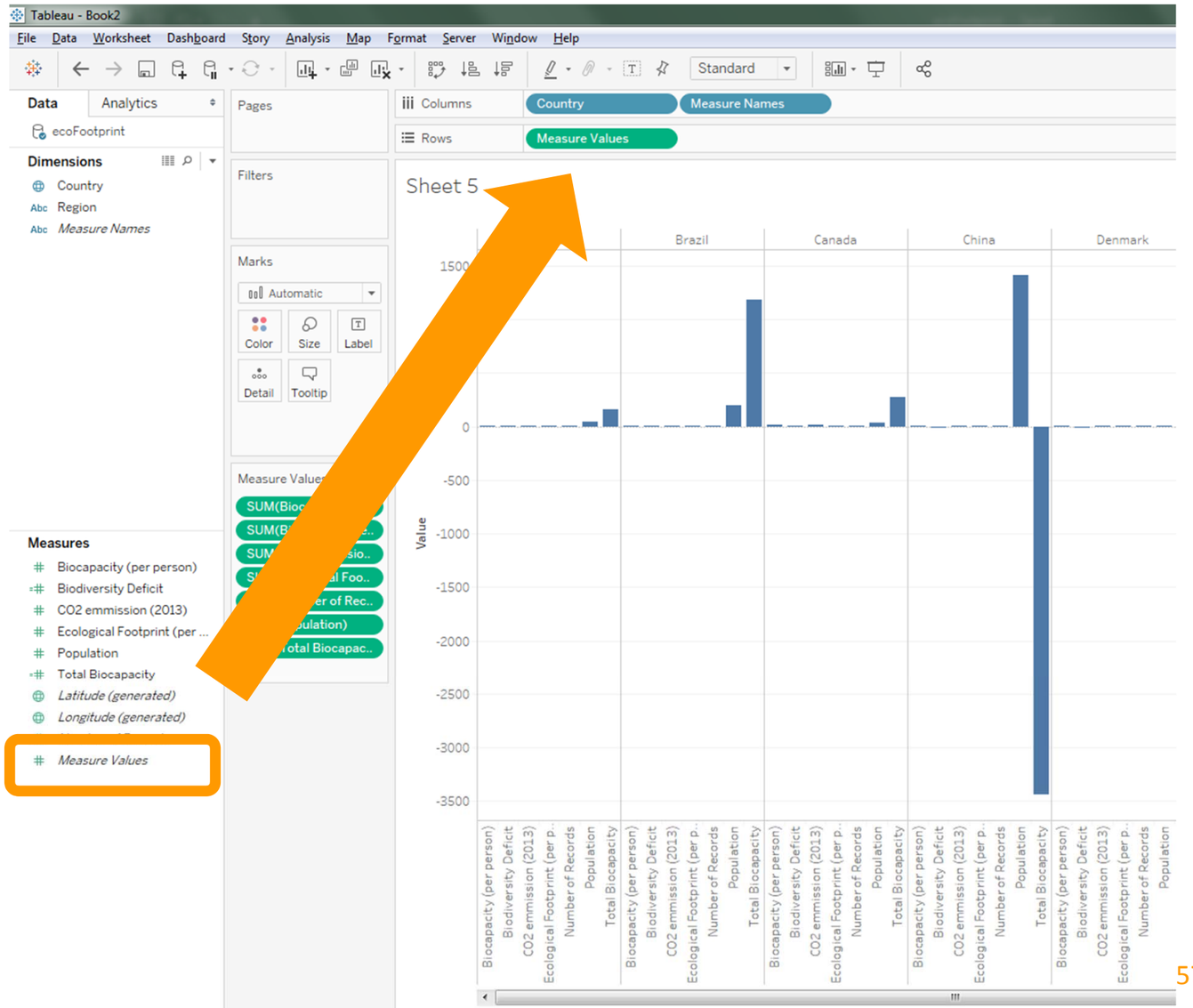
Creating a grouped bar chart manually
→ eco footprint vs. biocapacity

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



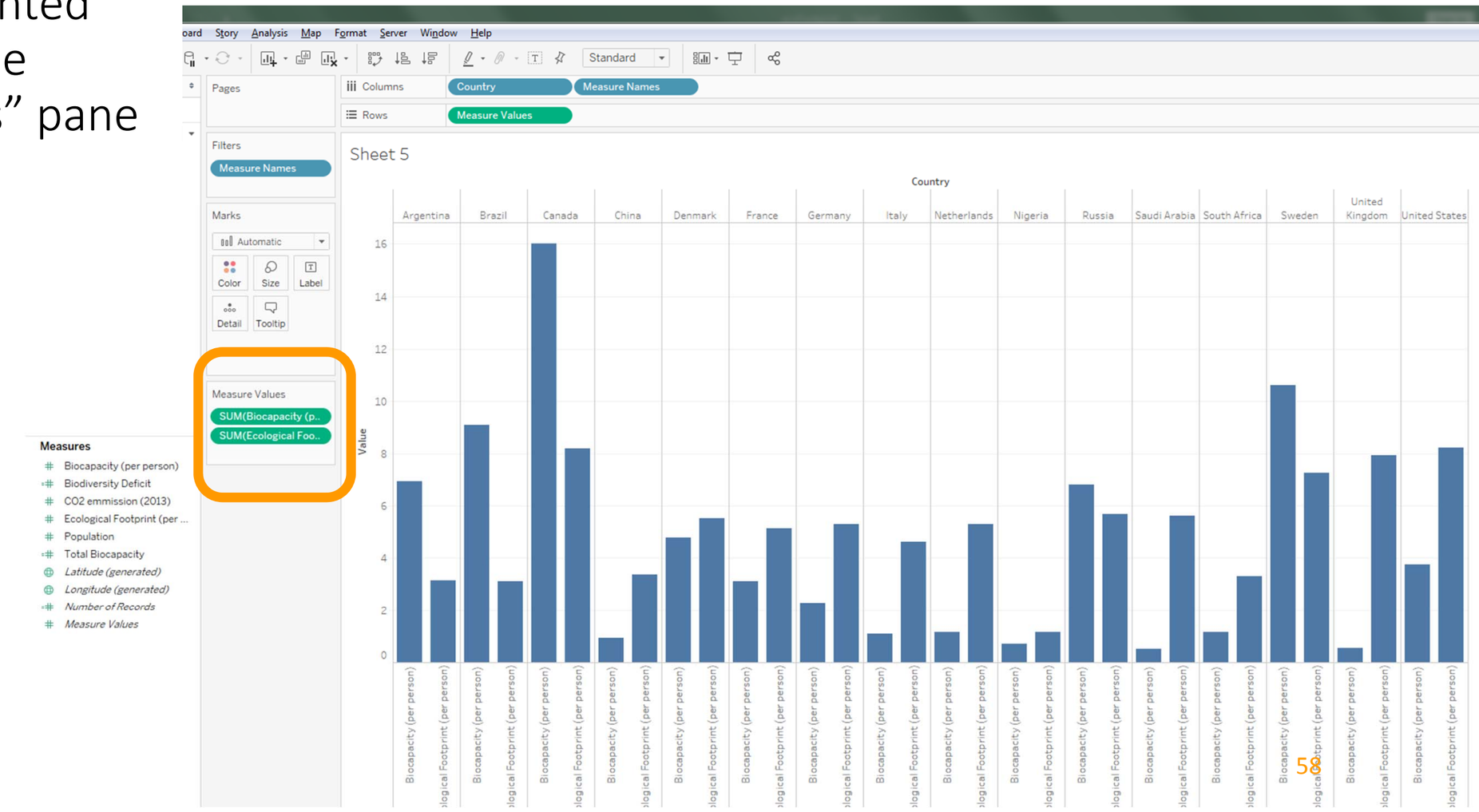
eco footprint vs. biocapacity

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”