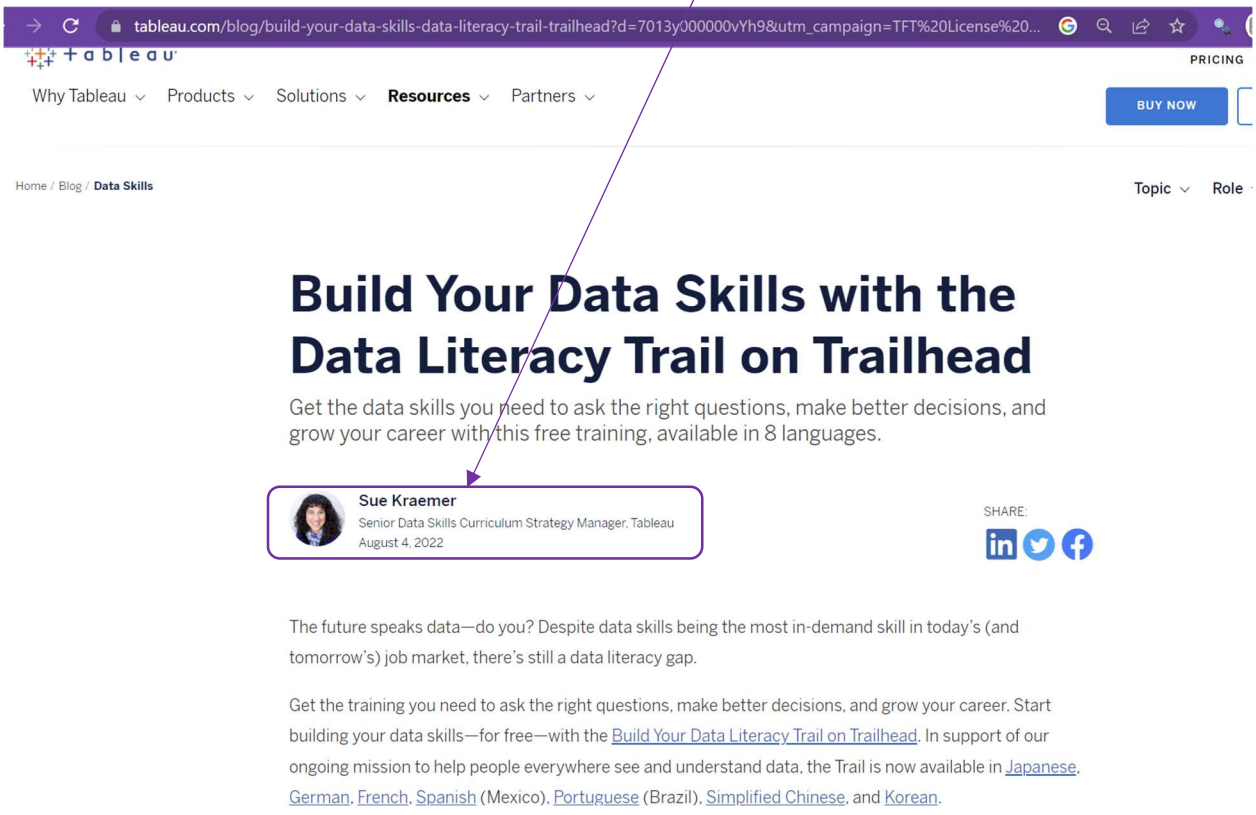
Year 2023 V6

Data Literacy, Data Analysis, and Data Visualization Skills Matter a Great Deal for Today's Undergraduate and Graduate Students: Let's explore a hands-on data visualization practice!

Given that data are being created and stored on an unprecedented scale, a strong data analytical skillset would be a notable plus for career-oriented students in the "Era of Data Ubiquity" (quotes Mitchell Stevens). Having strong skills in data literacy, data analysis, and data visualization can help both undergraduate and graduate students (who will be tomorrow's business practitioners and business leaders) make informed decisions using data. These skills can give them the ability to make "data-based decisions."

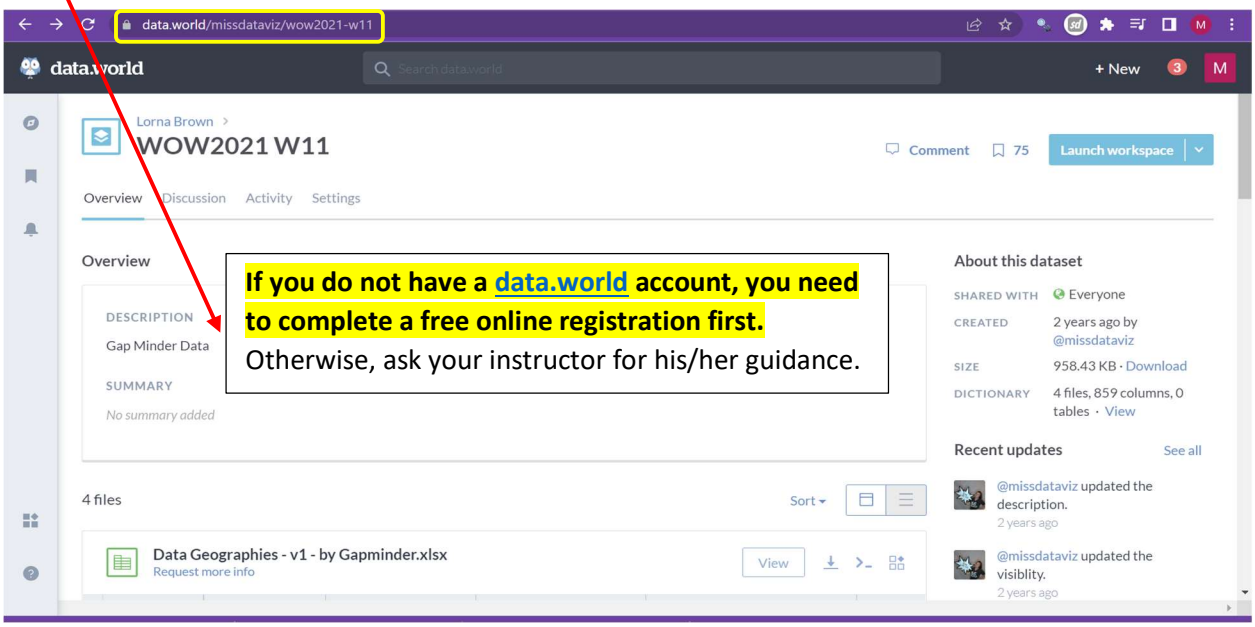
Tableau is a data visualization software tool that allows users to create interactive charts, graphs, and dashboards from their data. It is meant to help people understand and analyze their data more effectively. Tableau is used by many types of organizations to make better decisions using their data. It can be connected to various data sources (e.g., Microsoft Excel, Microsoft Access, PDF files, Statistical files) and enables users to easily create visualizations by dragging and dropping elements. The insights generated by Tableau can be shared with others. Please **watch a 2:19 minute Tableau Introduction video** if you have little idea about the power of Tableau: https://www.tableau.com/why-tableau/what-is-tableau?creative=&cq_cmp=1695532942&cq_net=g&cq_plac=#video

Who can benefit from data literacy training? According to a Tableau blog, anyone can benefit from data literacy training, especially (1) business professionals, (2) organizations, (3) informed citizens, and (4) students. Feel free to check the details on a blog posted by Sue Kraemer (https://tabsoft.co/3imKLTY).

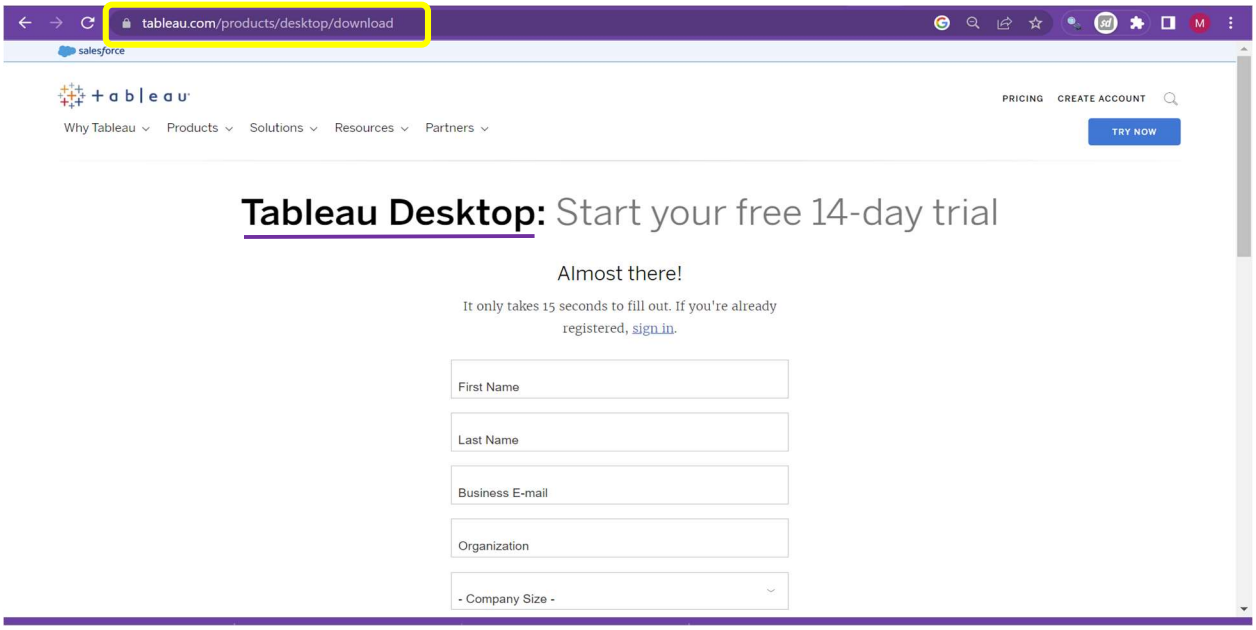


In this hands-on data visualization practice, let's **employ Tableau to visualize a dynamic (changing) relationship between income and life expectancy across over 150 nations in the past 200+ years** (from year 1800 up to year 2023).

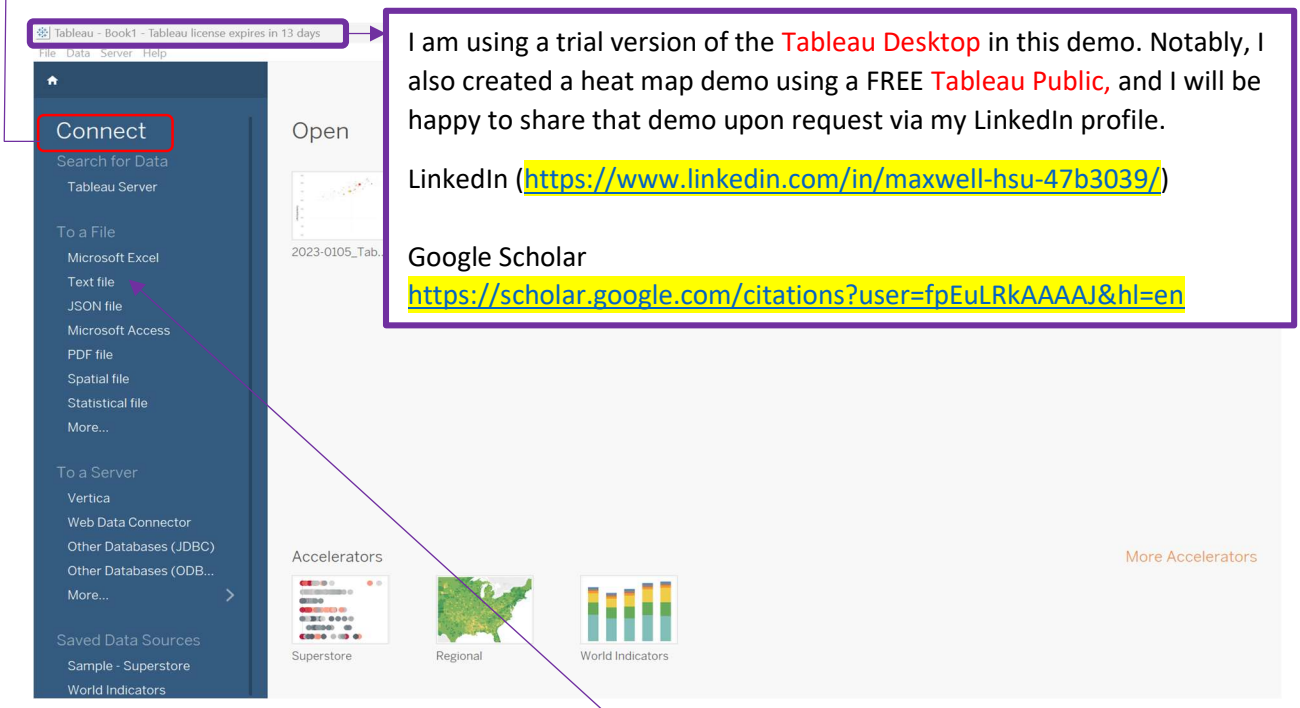
1. We will download four relevant datasets from **data.world** (https://data.world/missdataviz/wow2021-w11). Pls click the underlined link to find the data sets online, and then save the files in your hard drive.



2. Please visit Tableau's official website and then download a 14-day free trial version of the **Tableau Desktop** (https://www.tableau.com/products/desktop/download).



3. We will enable **Tableau Desktop** and then "connect" the datasets to the Tableau canvas [FYI, under **Connect**, select the relevant file types (e.g., a Microsoft Excel file, a .csv text file, or a .sav SPSS data file). In the **Open** dialog box, navigate to and select a file. Select **Open**…]

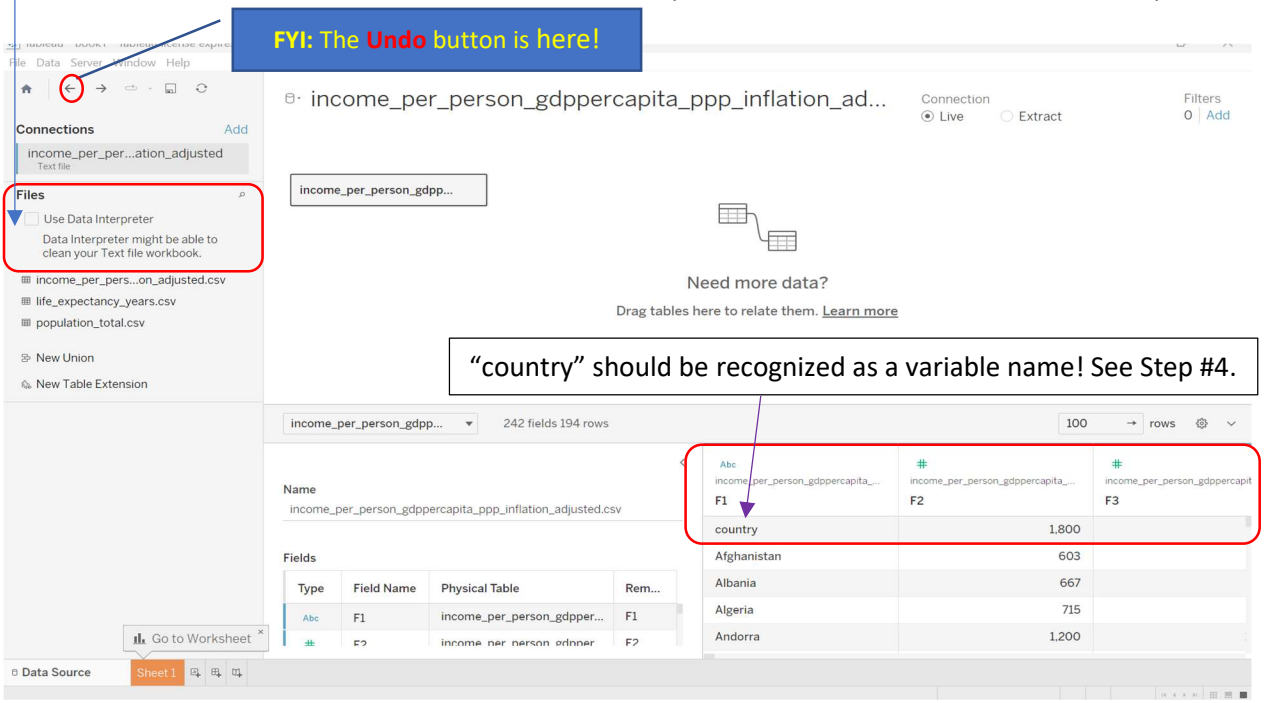


I am using a trial version of the Tableau Desktop in this demo. Notably, I also created a heat map demo using a FREE Tableau Public, and I will be happy to share that demo upon request via my LinkedIn profile.

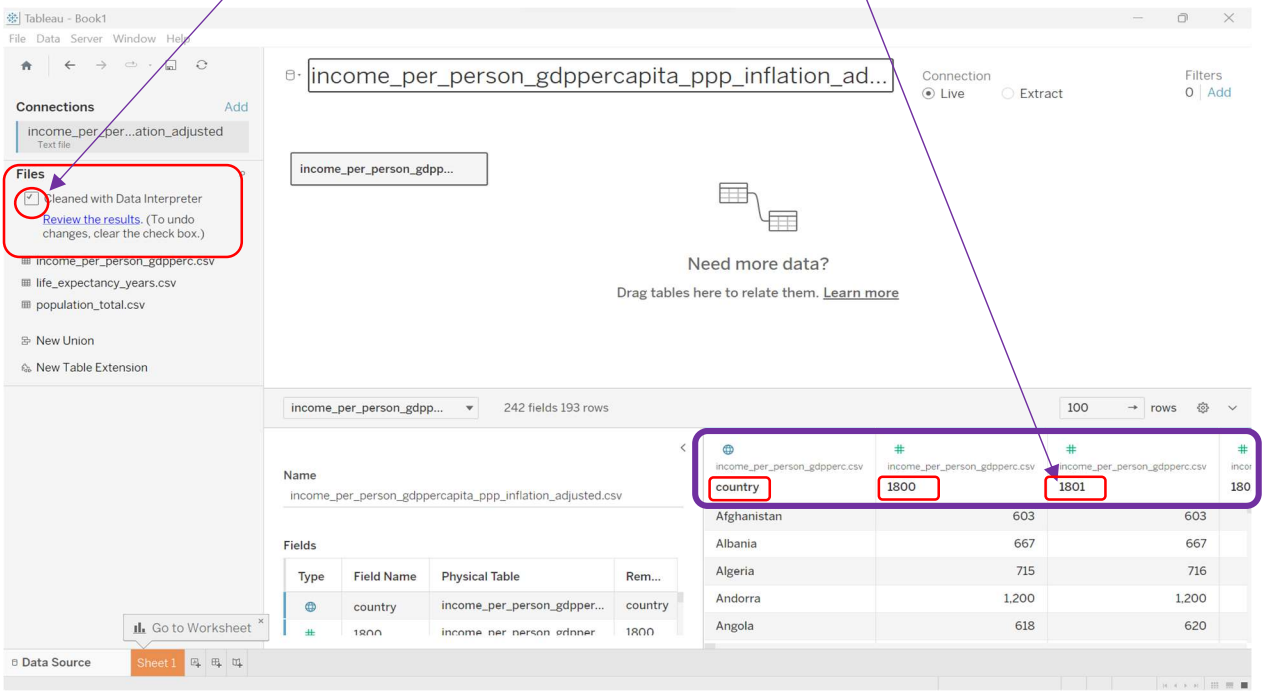
LinkedIn (https://www.linkedin.com/in/maxwell-hsu-47b3039/)

Google Scholar
https://scholar.google.com/citations?user=fpEuLRkAAAAJ&hl=en

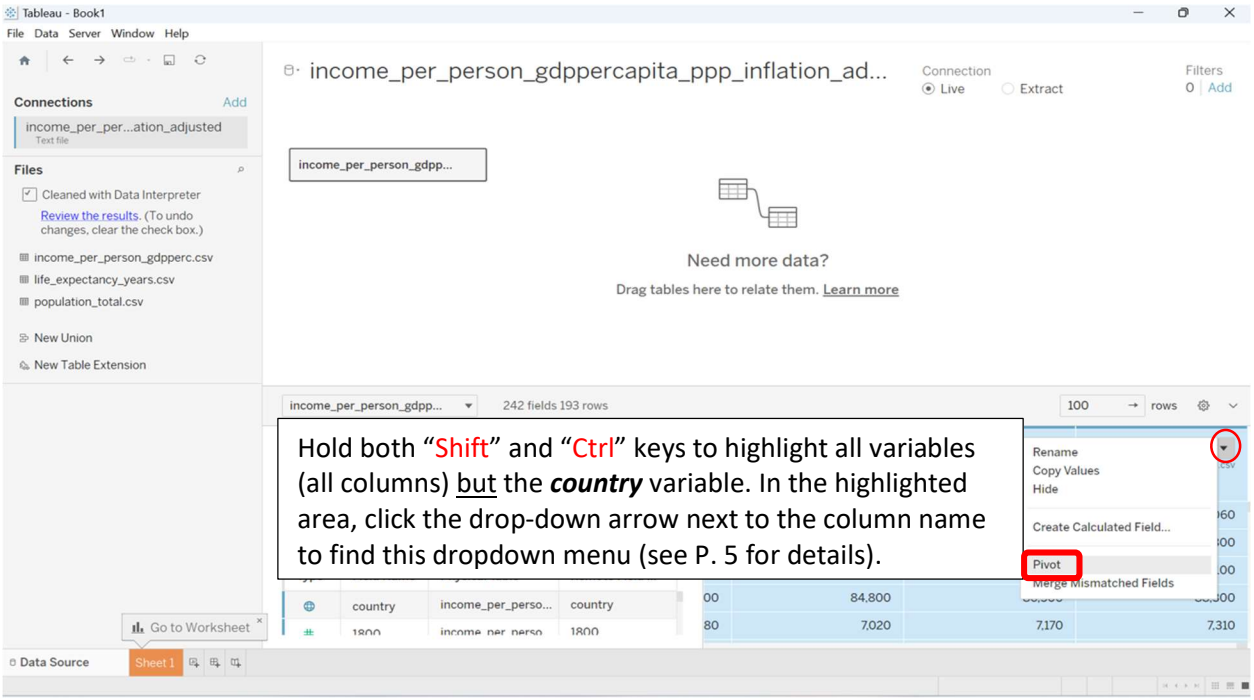
Let's start with the income per person data set. As the "*income_per-person gdppercapita_ppp_inflation_ajusted*" is a .csv file, we need to connect this file by clicking the "Text file" option. By default, the "Use Data Interpreter" option is not enabled. When we find that the 1st row data (e.g., ***country***, ***1800***) correspond to the variable names, we will check the "Use Data Interpreter" box to address the issue (see Step #4).



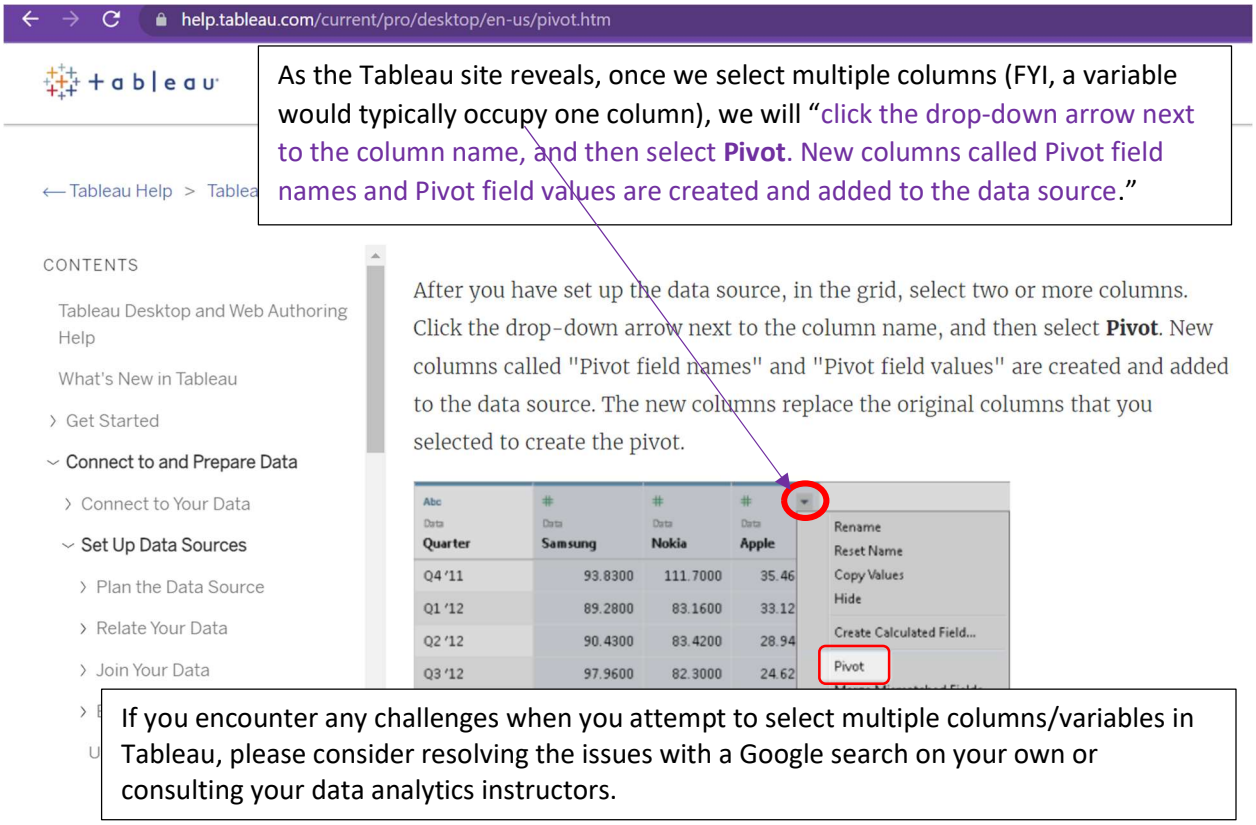
4. Let's check mark the "Cleaned with Data Interpreter" option (see the circled red rectangle box and the red circle). Subsequently, we will find that the first-row data (i.e., ***country***, ***1800***, ***1801***, etc.) are recognized as variable names. FYI, ***1800*** represents "year 1800" while ***1801*** represents "year 1801."



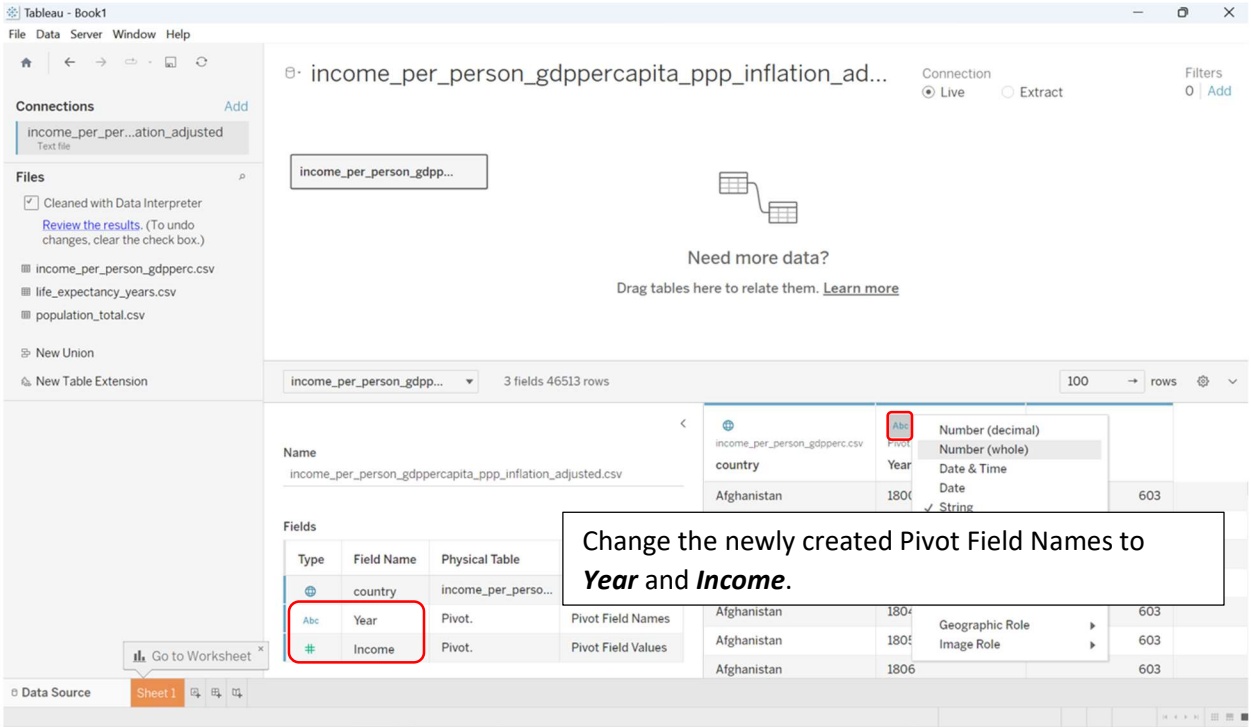
5. We will re-organize the data structure using the Pivot function (i.e., we will Pivot the income data).



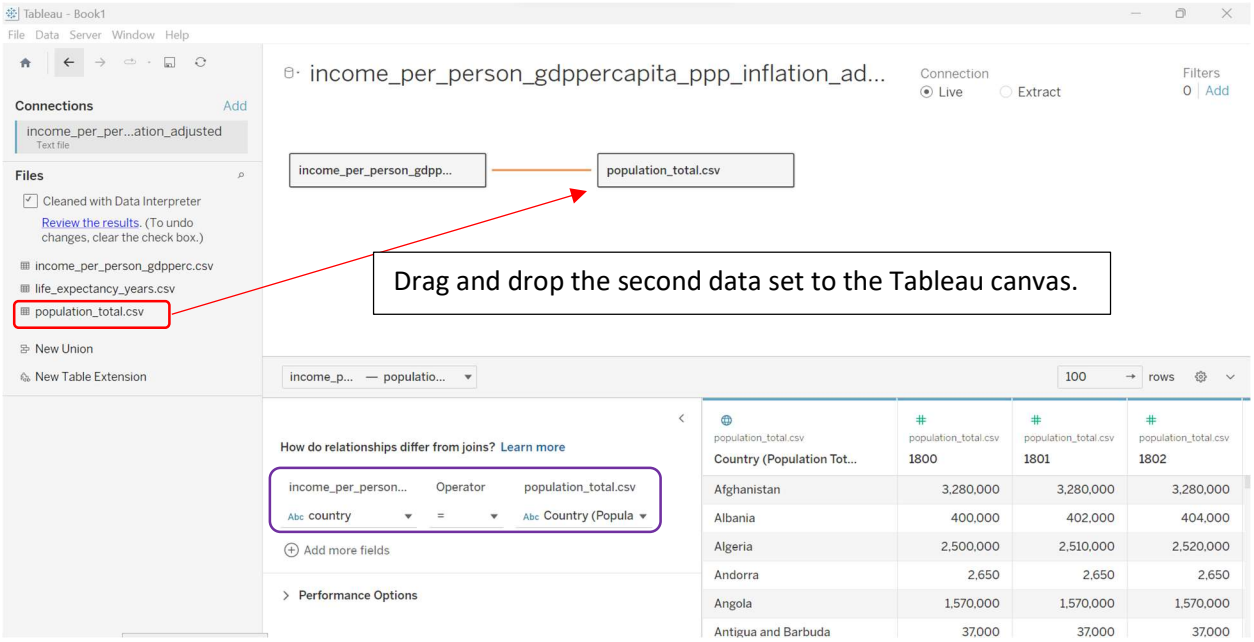
Hint: if you are not familiar with the power of the Pivot function, please check relevant information on this Tableau webpage: https://help.tableau.com/current/prep/en-us/prep_pivot.htm or this Microsoft webpage: https://bit.ly/3vMzVKd



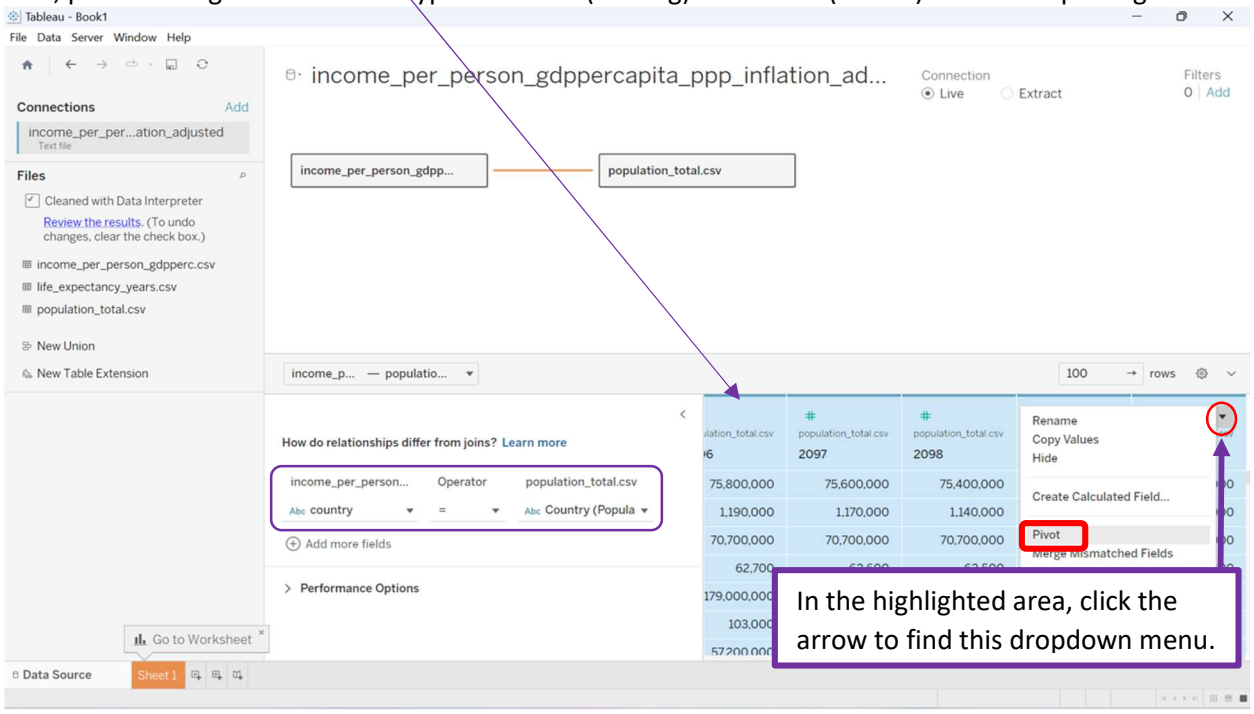
6. Double click "Pivot Field Name" and change it to ***Year***. Change "Pivot Field Values" to ***Income***. FYI, we should also click the data type icon and change variable ***Year***'s data type from ABC (string) to Number (whole) here, but I failed to do so in Step #6 → Check Step #24 to see the consequence & how to fix it.



7. We will connect the second data set (i.e., population total) to the Tableau canvas with the Drag and Drop trick.

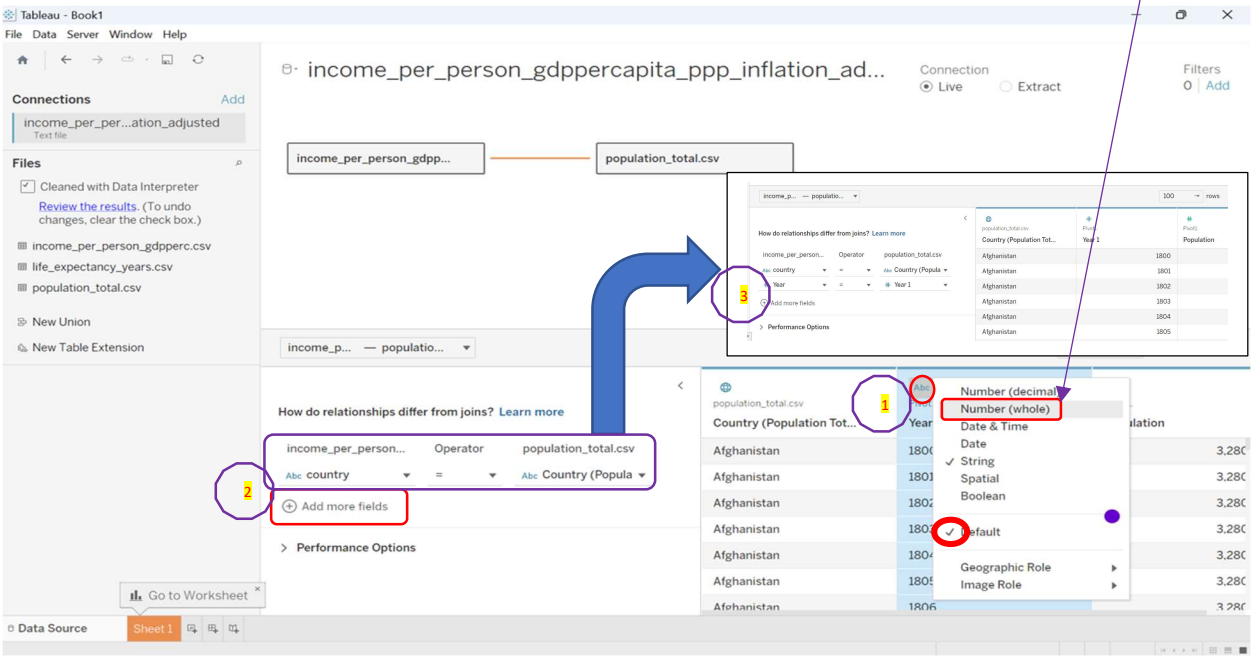


8. Please highlight all variables but the ***country*** variable. Likewise, we will re-organize the data structure using the Pivot table function. Change the variables names to "***Year 1***" and "***Population***" (like Step #6). Also, please change ***Year 1's*** data type from ABC (a string) to Number(Whole). Revisit Step #6 again.

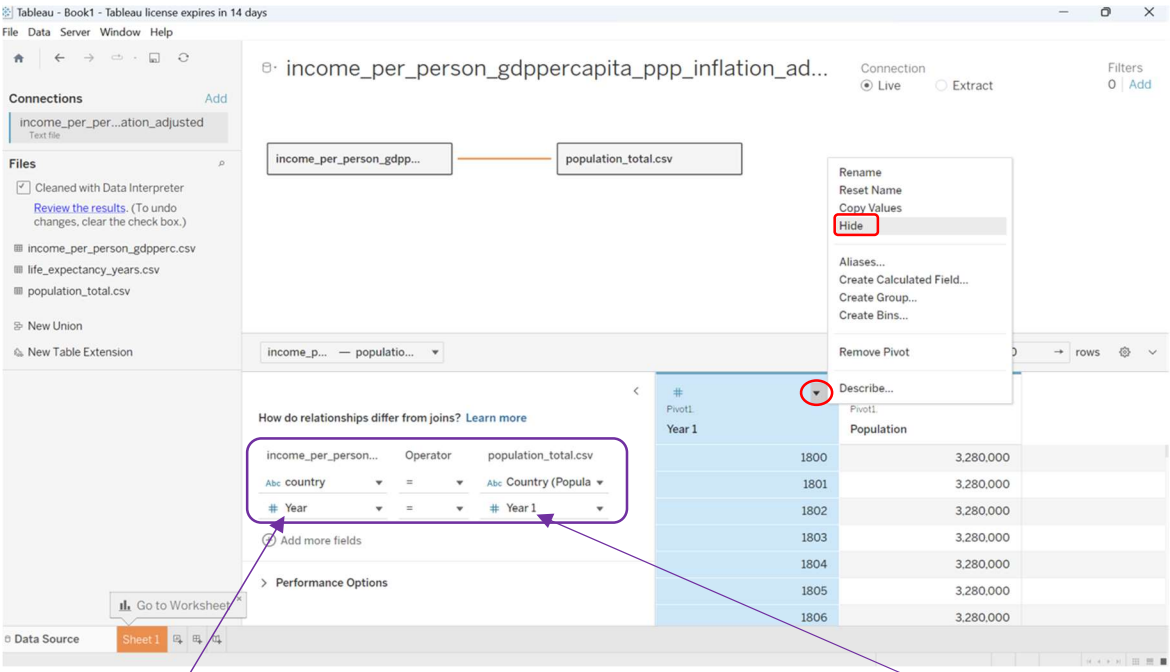


Hint: Do NOT highlight the ***country*** variable when we attempt to generate a Pivot varible … Also, please change the newly created variable name to ***Year 1*** (or any variable name but "***Year***" because this variable name has been used and the ***Year*** variable could be found in the income per person data file).

9. We will recognize the relationship between variable ***Year*** (in the income per person data set) and variable ***Year 1*** (in the poppulation data set). Remember to edit variable ***Year 1***'s data type (to Number).

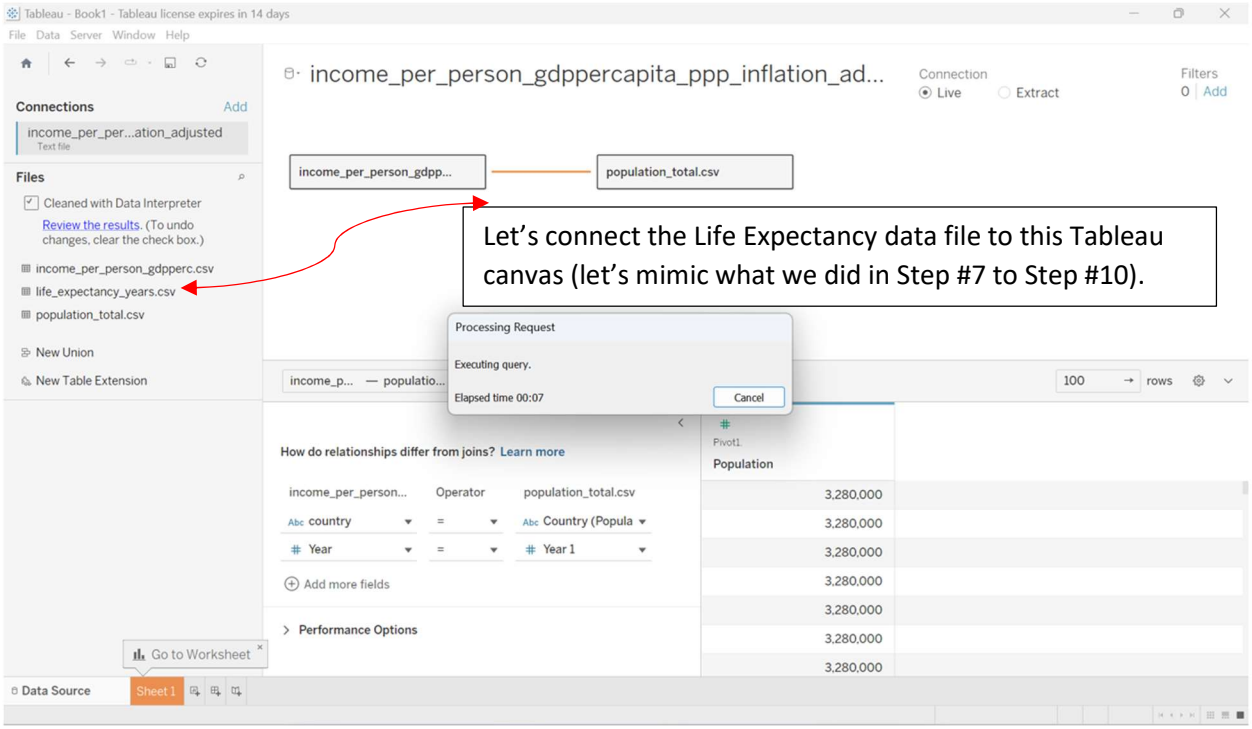


10. We will hide all variables but the ***Population*** variable (Hint: it's okay if you skip the hinding step, as not hinding other variables won't cause any issues when we create a bubble chart).

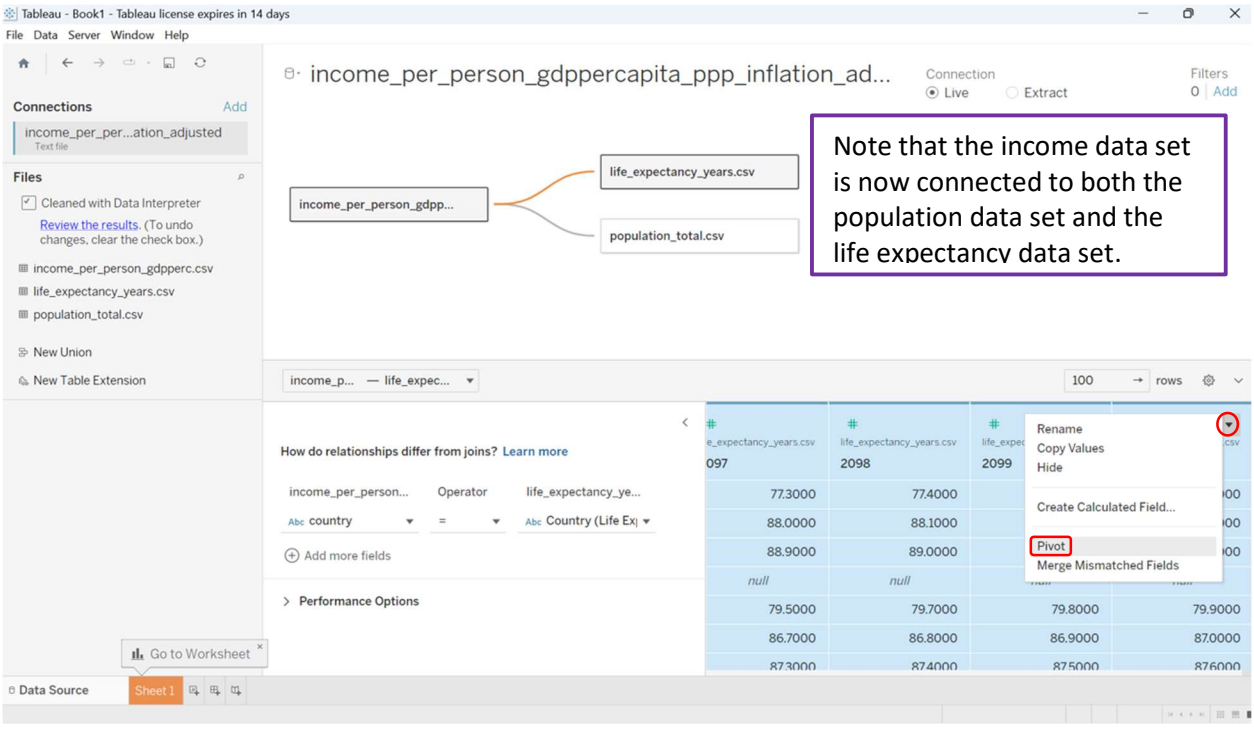


Note: the ***Year*** variable in the income per person data set corresponds to the ***Year 1*** variable in the population total data set.

11. Now, we will connect the third data set (i.e., life expectancy) to this Tableau project.

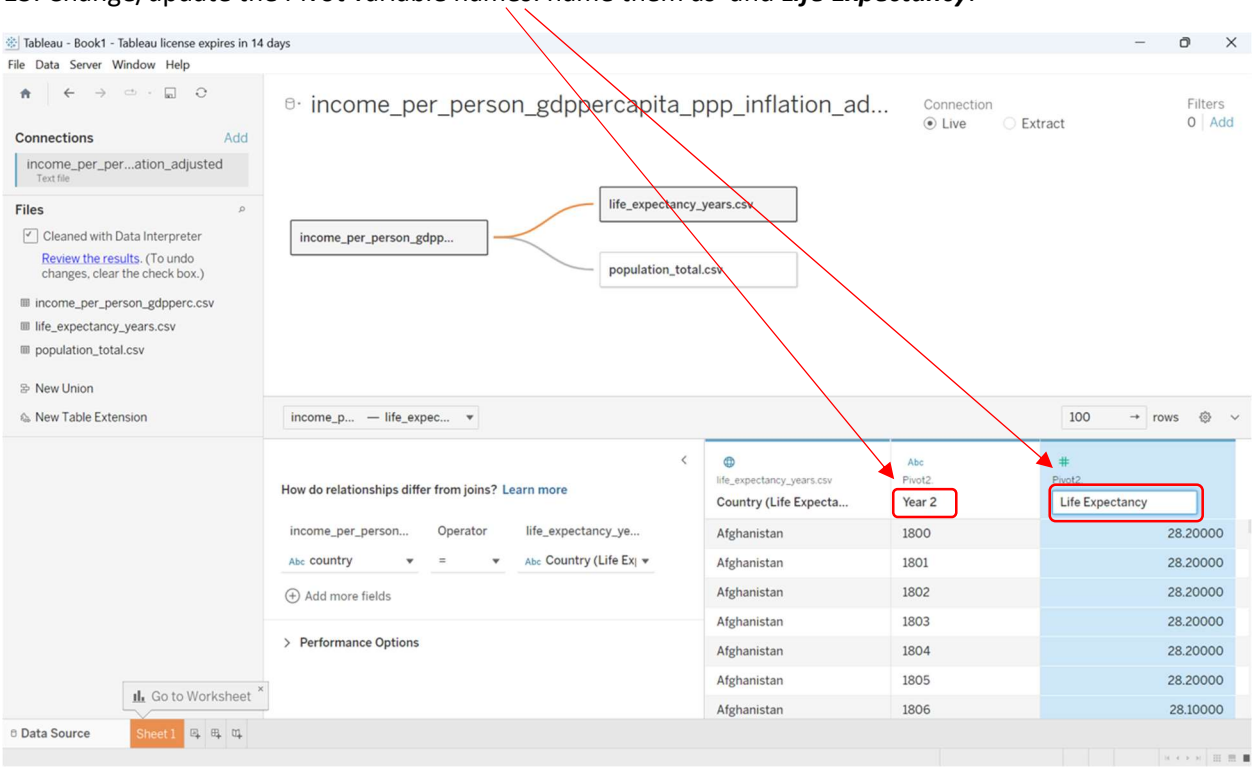


12. Similarly, we will Pivot the life expectancy data (FYI, you may want to revisit Step #8).

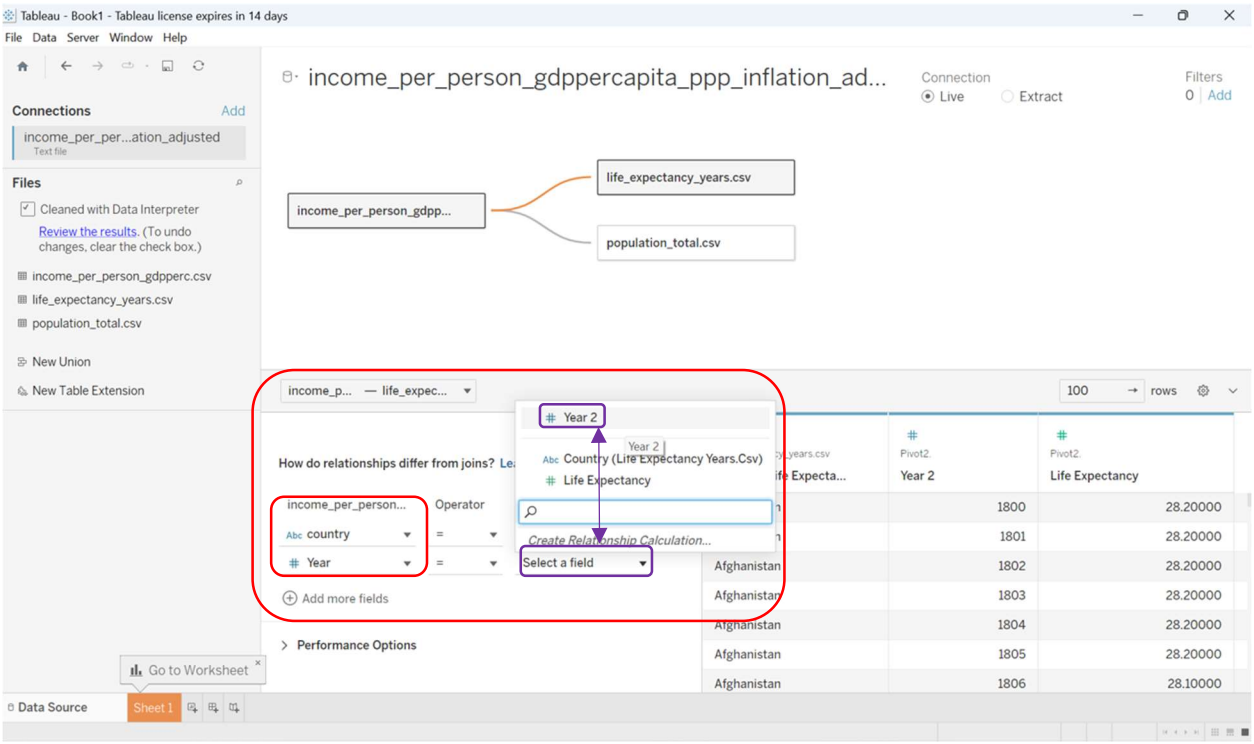


Hint: Leave the ***country*** variable intact (i.e., the ***country*** variable should not be included in the Pivoting task). As a reminder, hold both the **Ctrl** and **Shift** keys when you attempt to highlight multiple columns.

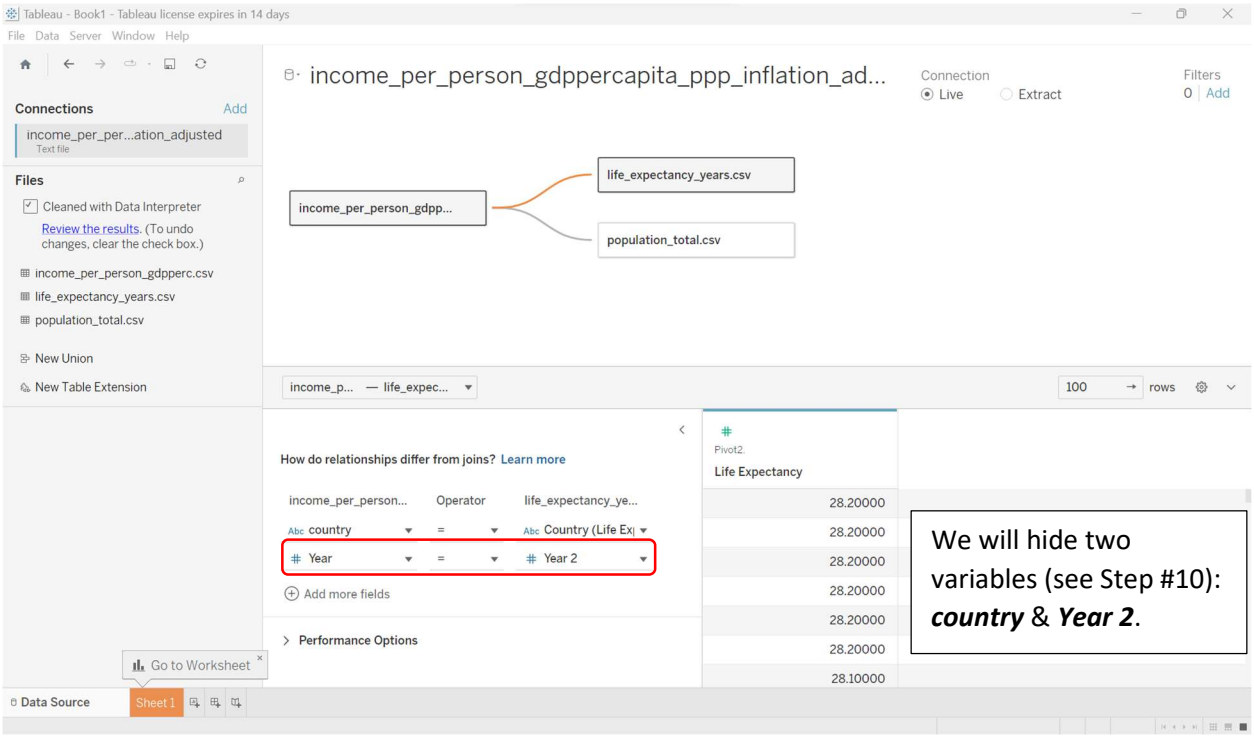
13. Change/update the Pivot variable names: name them as and ***Life Expectancy***.



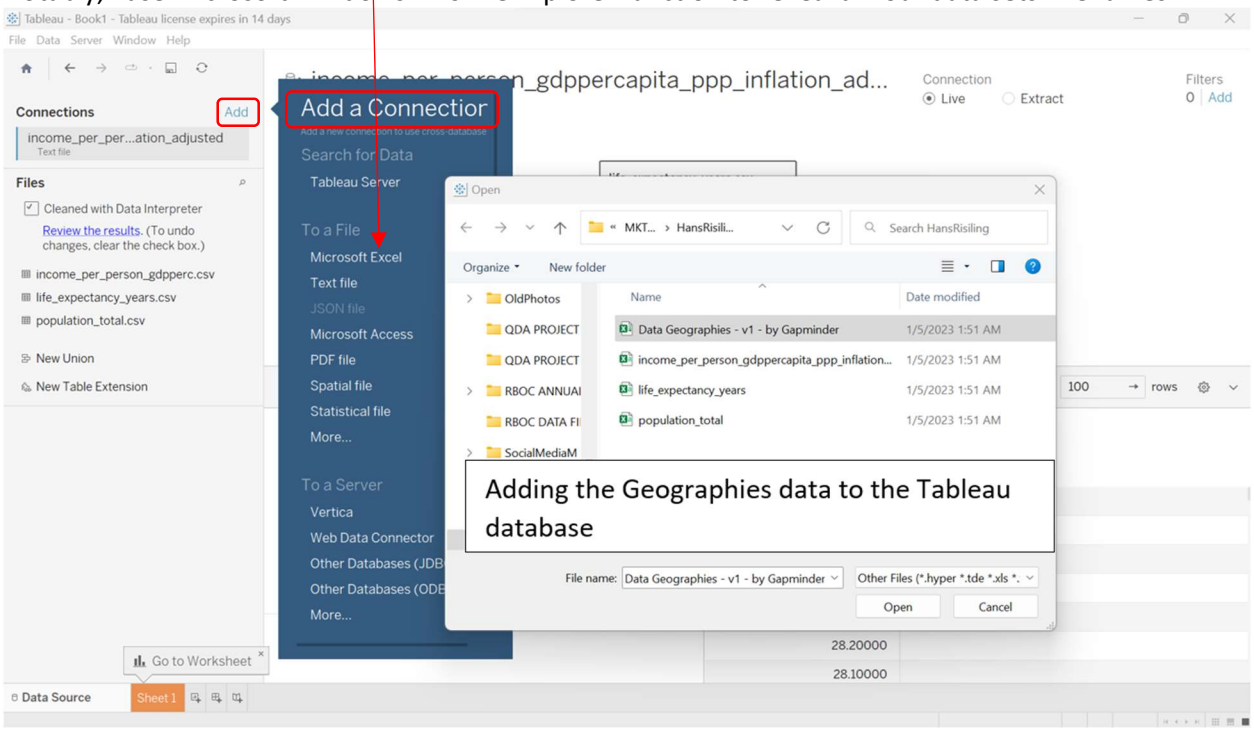
14. It's about time to instruct Tableau to recognize the additional relationships between the data sets.



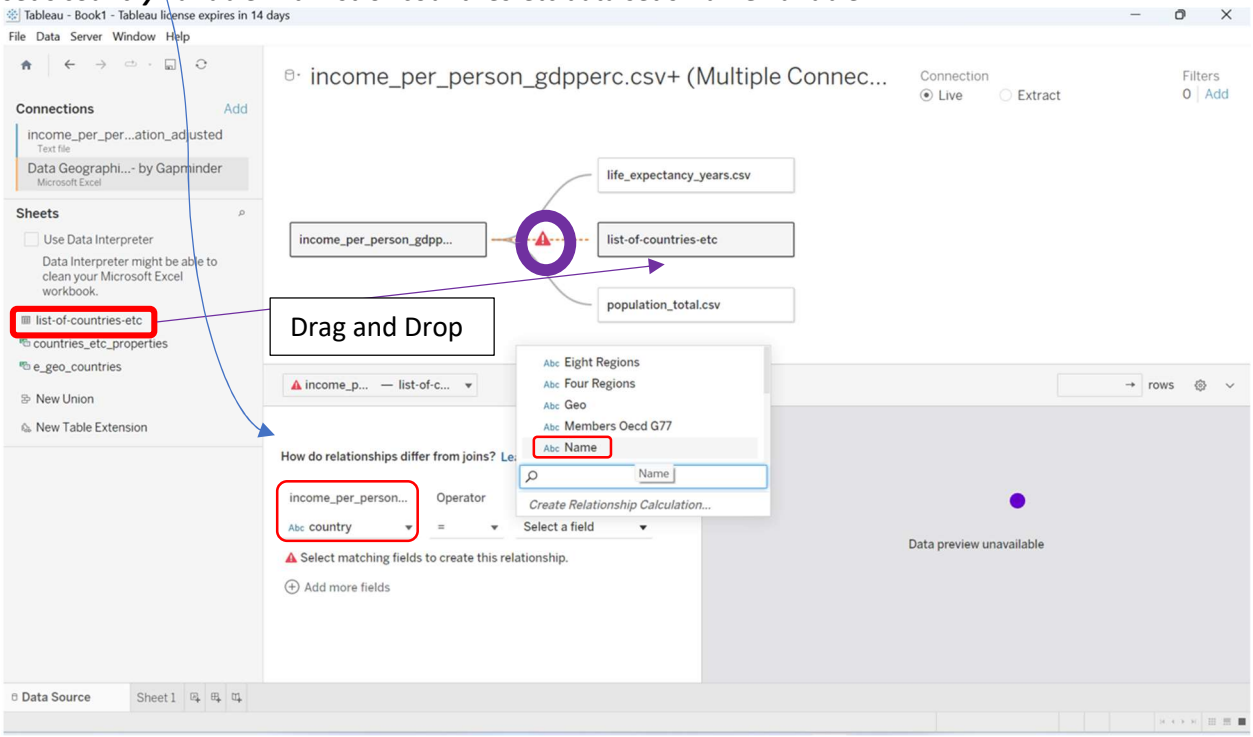
15. After we hide variables ***country*** and ***Year 2***, the only "visible" variable in the life expectancy data set will be "Life Expectancy". Check the red box between Step #14 and Step #15 (a relationship between ***Year*** in the income data set and ***Year 2*** in the life expectancy data set has been established).



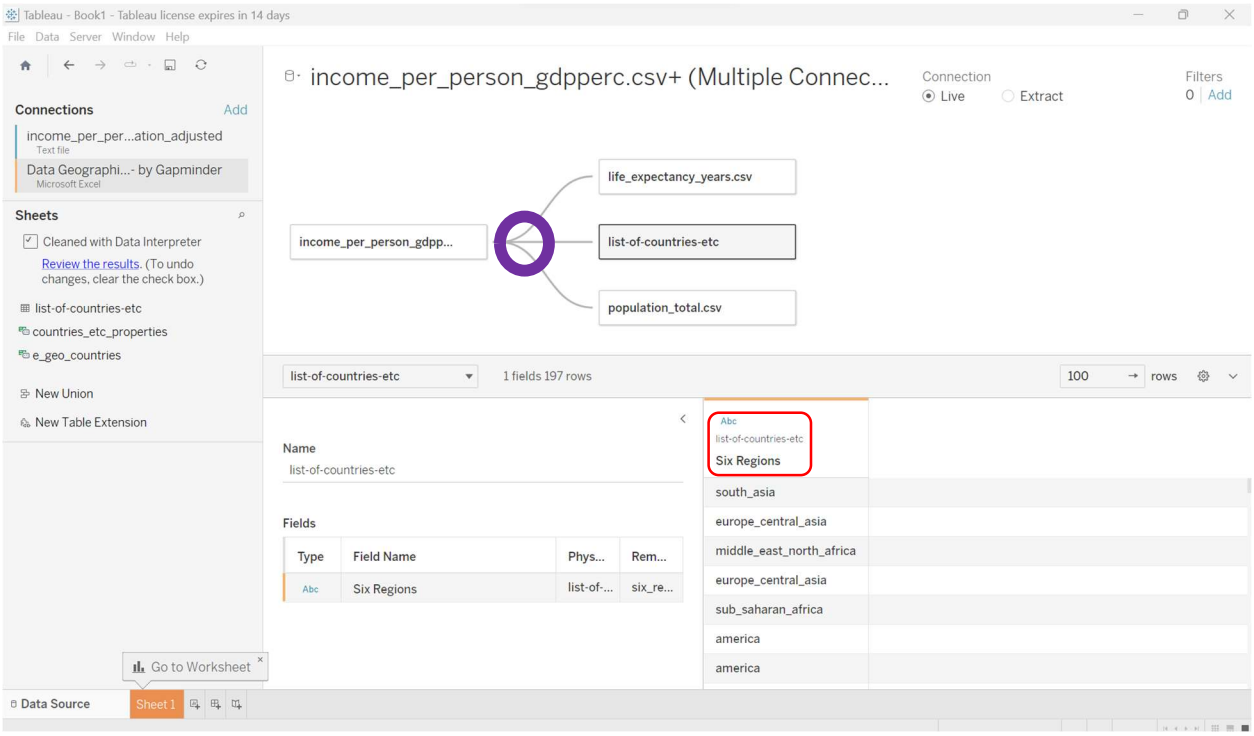
16. We will connect the last data set (i.e., the Geographies) to the current Tableau project. Since this data file is an Excel type, click "Microsoft Excel" and Open "Data Geographies-v1-by Gapminder." Notably, I use Microsoft Windows 11's File Explorer function to reveal all four data sets' filenames.



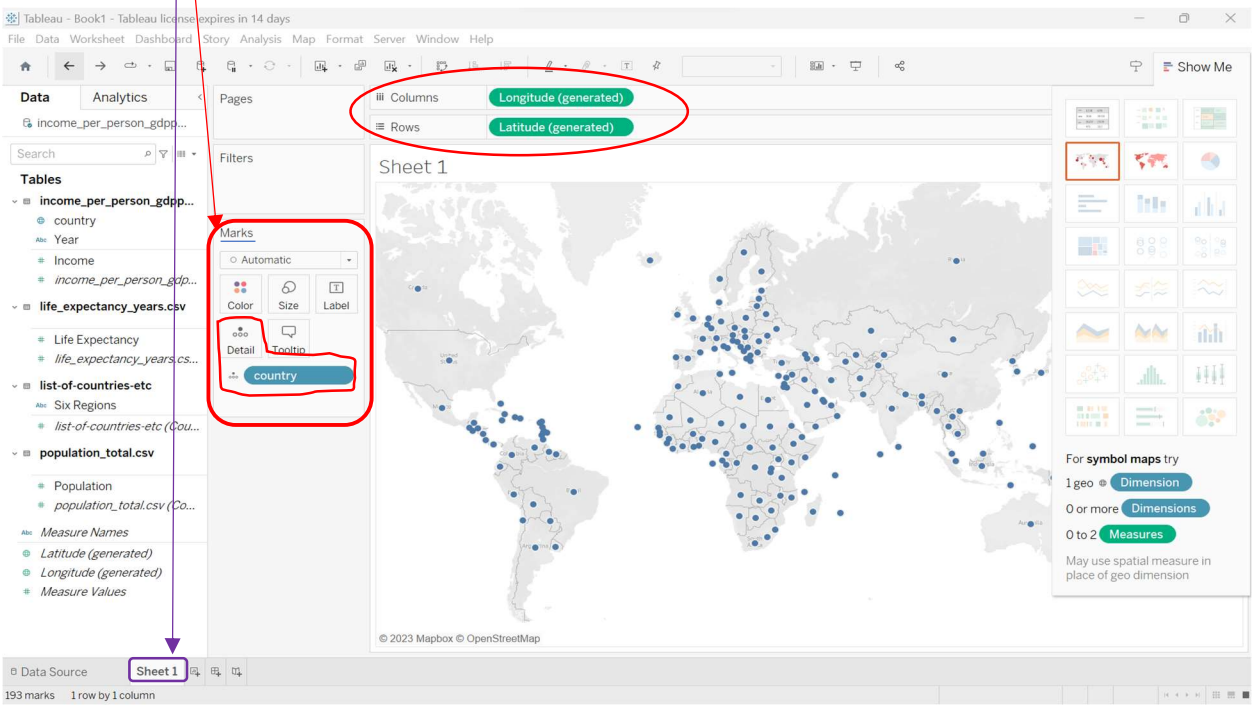
17. Now, let's connect the "list-of-countries-etc" sheet to the income per person data set shown on the Tableau canvas. Establish the relationship between two data sets (i.e., income & list of countries), and the exclamation mark (see the purple circle) will disappear. Note that we need to connect Income data set's ***country*** variable with list-of-countries-etc data set's ***Name*** variable.



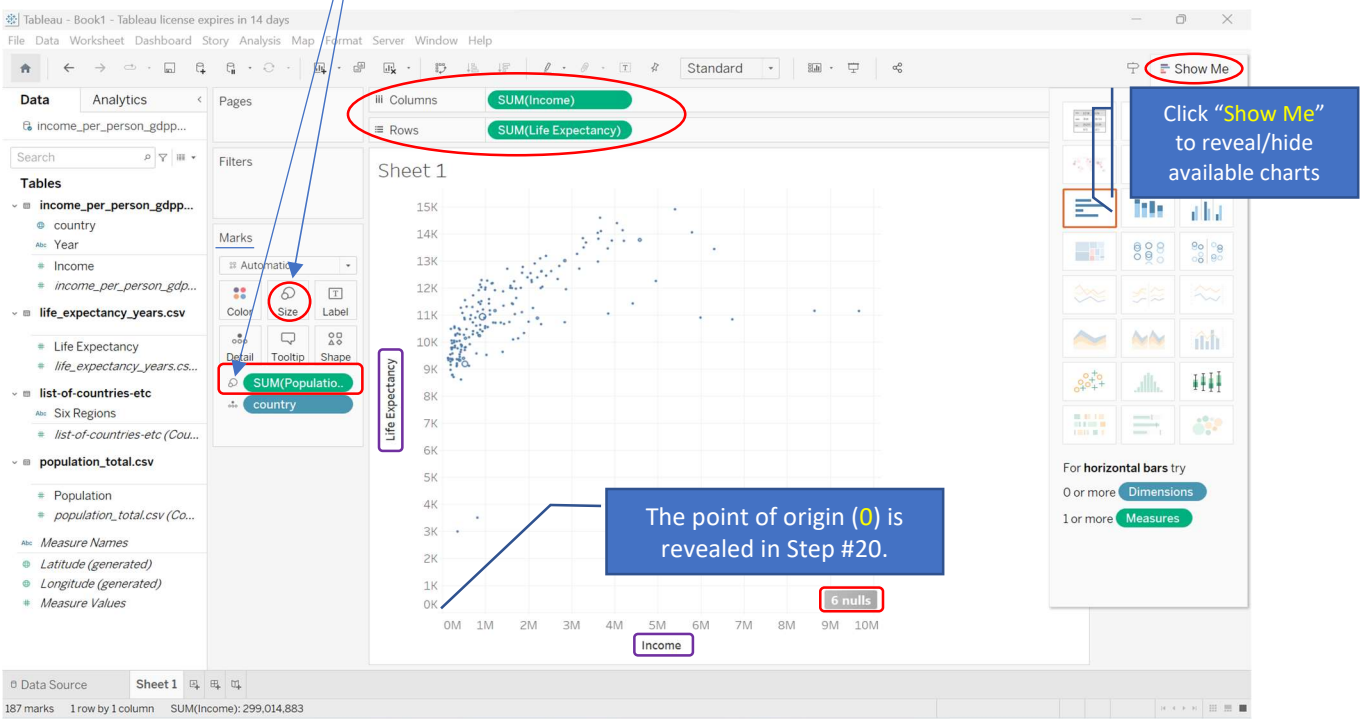
18. Again, we will hide all variables but the "***Six Regions***" variable in the list of countries data set.



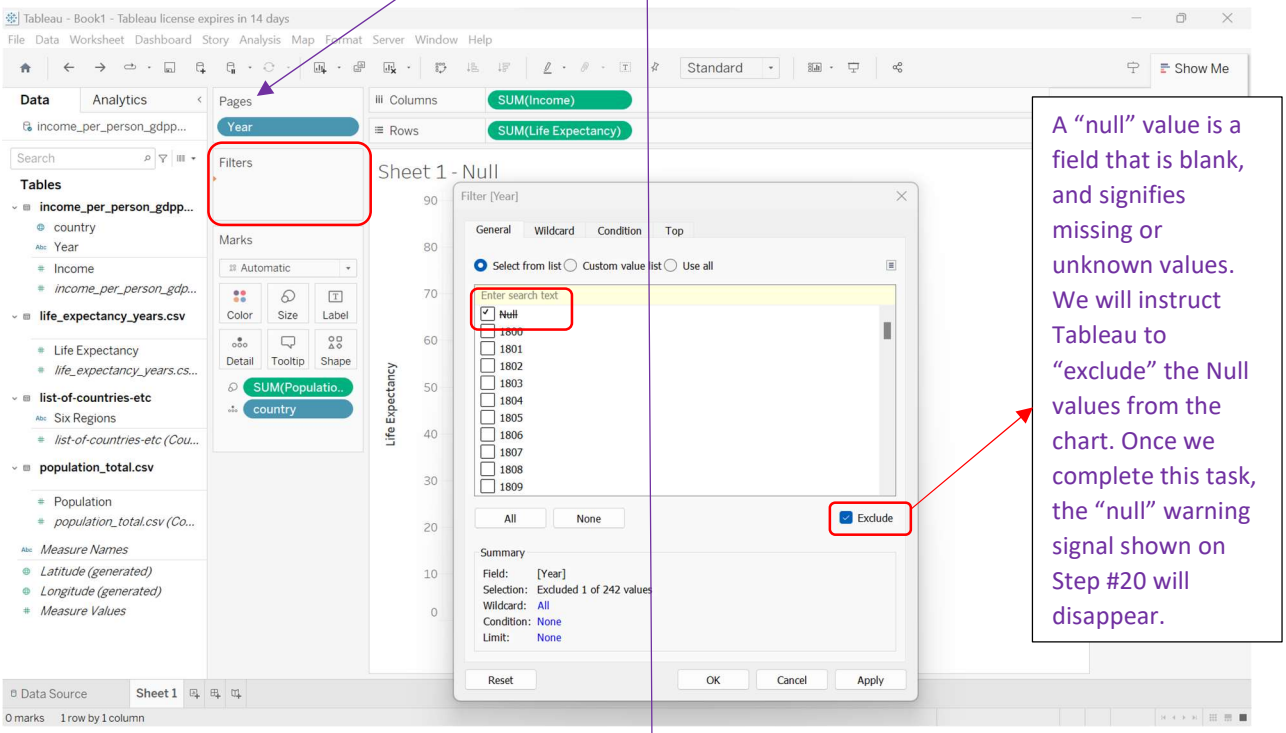
19. We are now ready to build a “static” bubble chart. Specifically, please click "Sheet1" at the bottom. After clicking “**Sheet 1**” (on the bottom of the Tableau window), we will need to drag the ***country*** variable to the Marks card (i.e., the Detail).



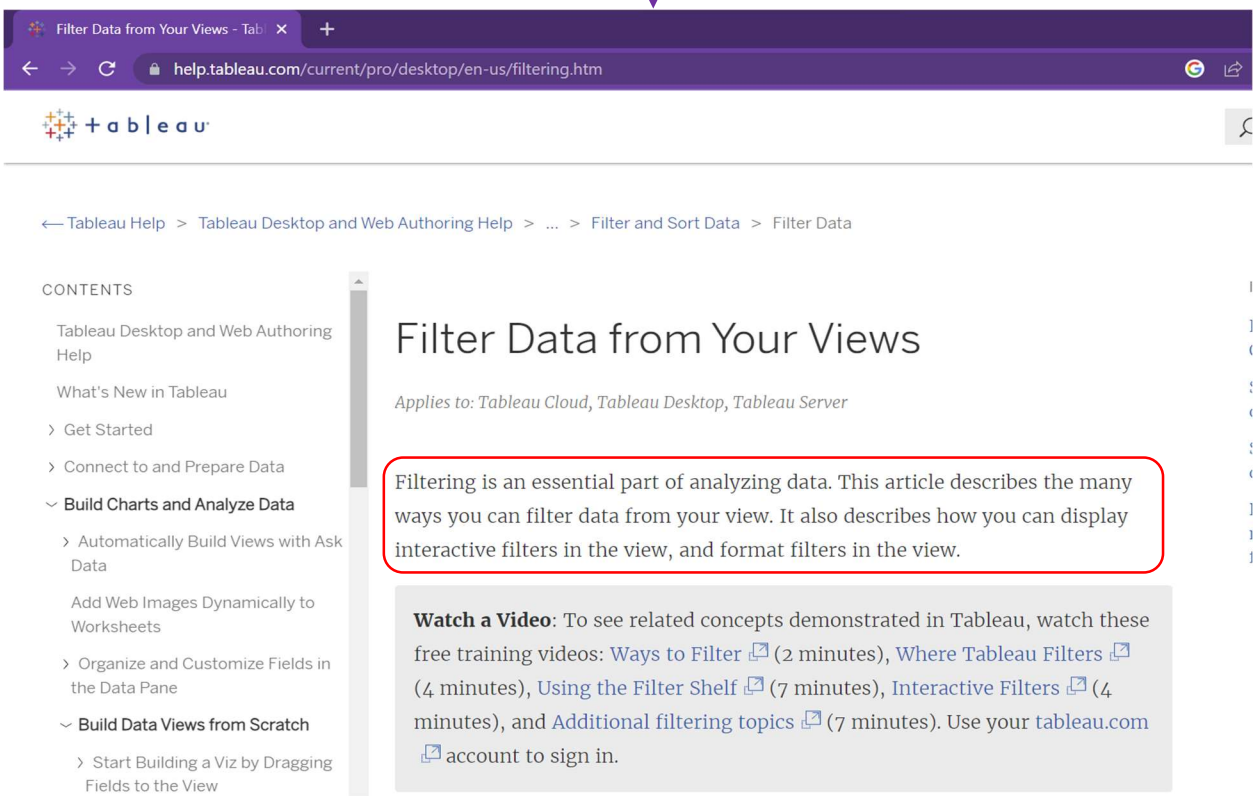
20. Let’s remove the Longitude and Latitude information from the chart. Next, let’s associate the ***income*** variable with the X-axis (i.e., columns) and associate the ***life expectancy*** variable with the Y-axis (i.e., rows). Notably, the darkness of the dot’s color on the charts corresponds to each country’s population.



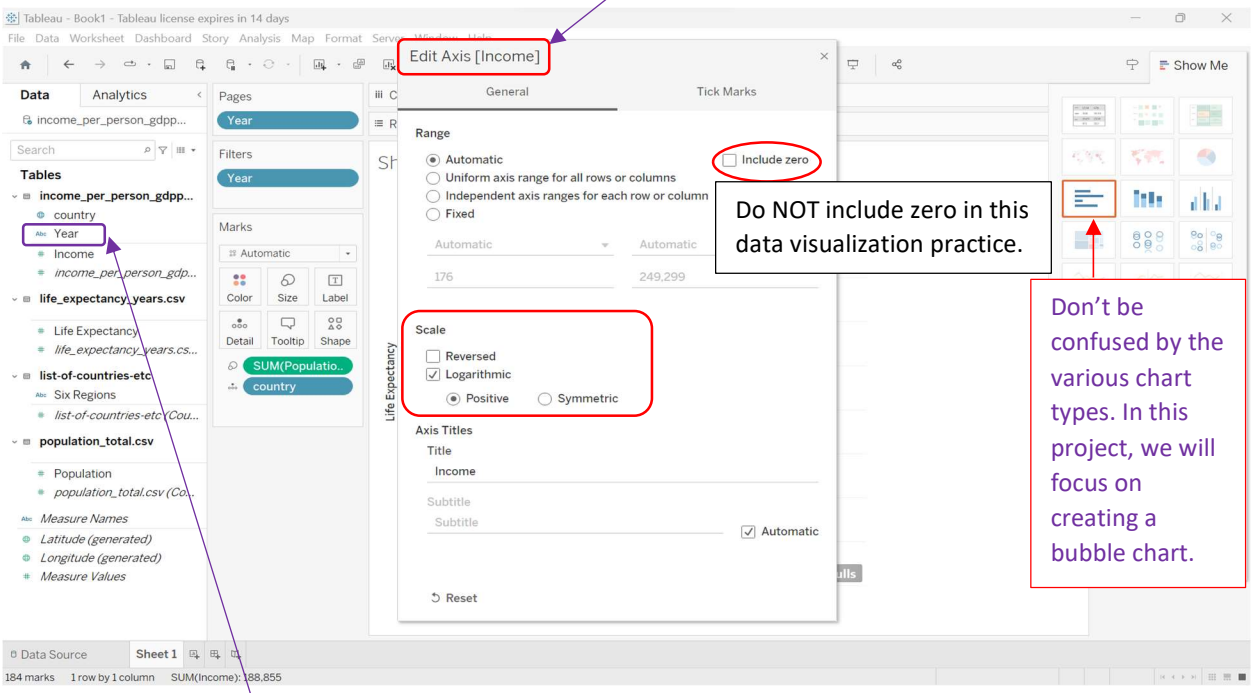
21. We will drag the ***Year*** variable to the “Pages” box (as a dimension). In addition, we will drag the same ***Year*** variable to the "Filter” box and then exclude the missing data (i.e., Null) from the data visualization task. More details about Tableau’s filtering function could be found online.



A “null” value is a field that is blank, and signifies missing or unknown values. We will instruct Tableau to “exclude” the Null values from the chart. Once we complete this task, the “null” warning signal shown on Step #20 will disappear.

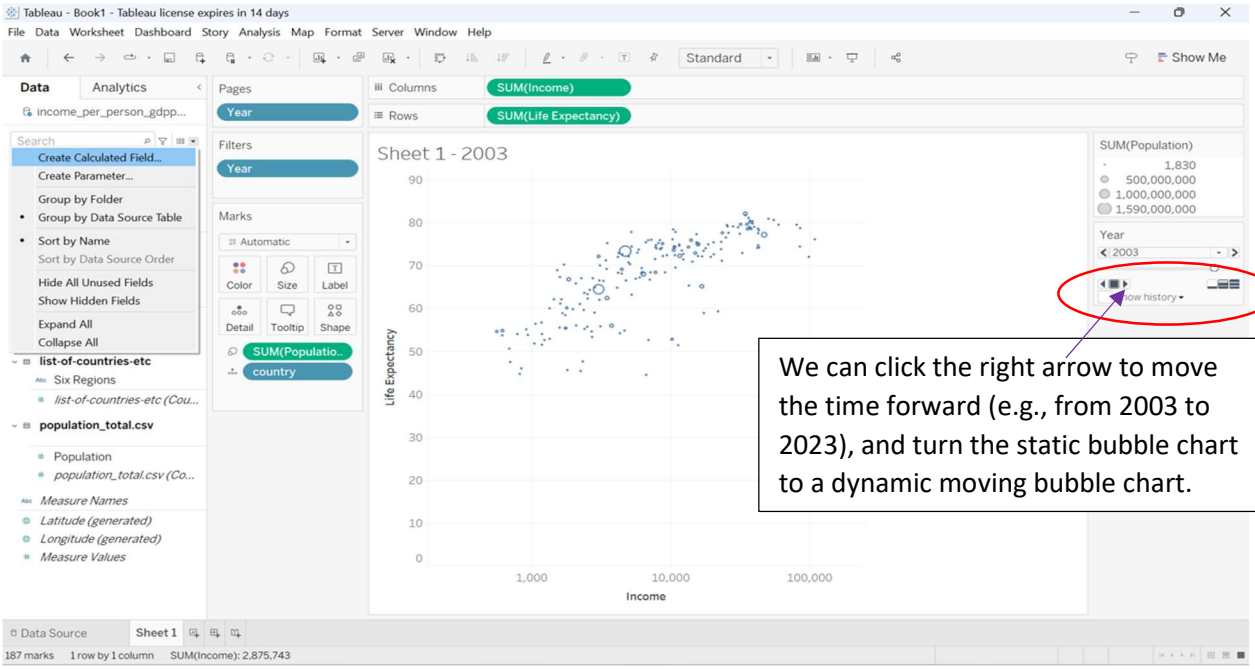


22. We will modify the scale. The way to show the "Edit Axis [income]" is to double click ***income*** at the bottom of the chart. Once we are in the Edit Axis [Income] pop-up window, check mark “Logarithmic” and uncheck “Include zero.”

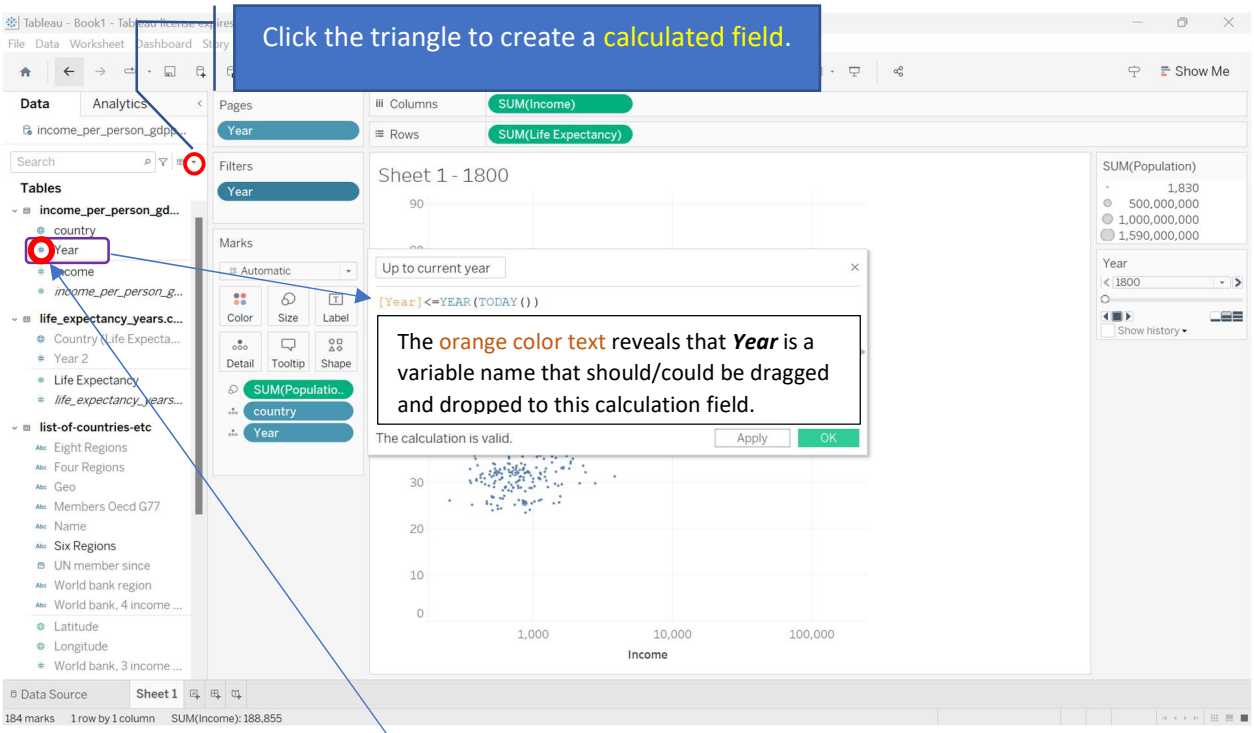


Hint: The variable “***year***” could be an issue (as it appears the data type is still a string, with an ABC icon)… and we will deal with this issue in Step #24.

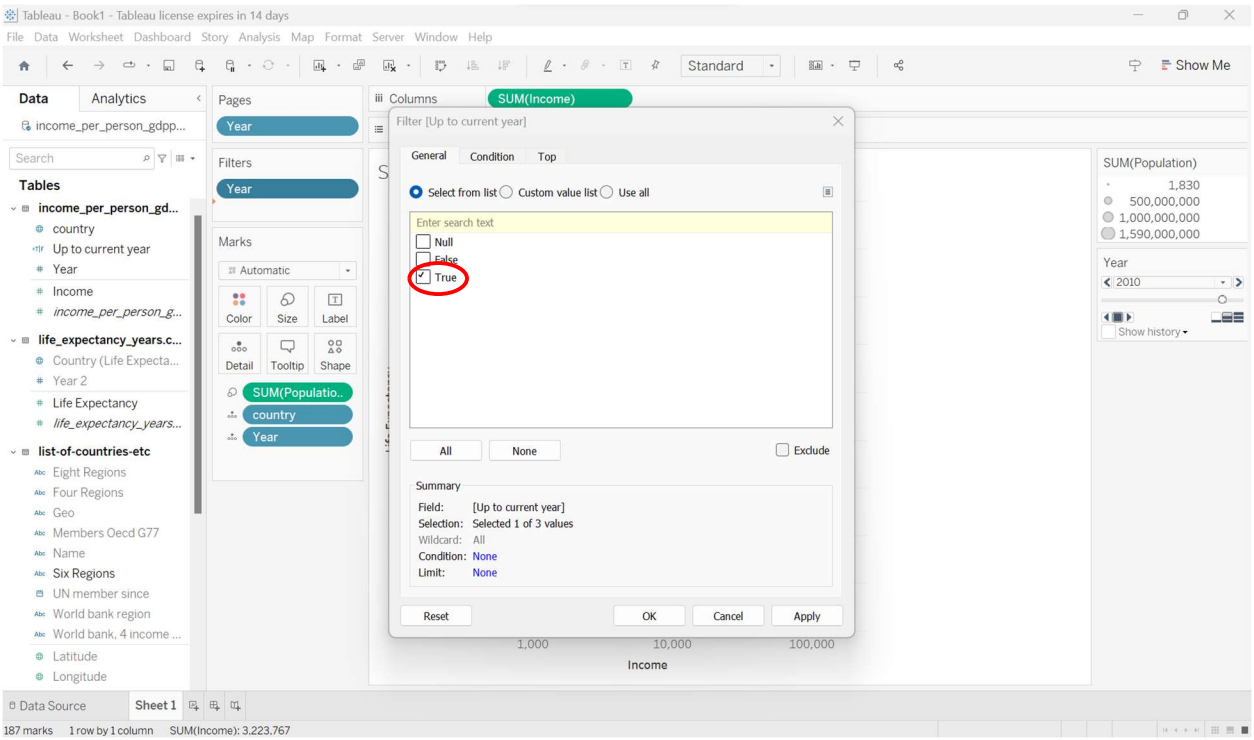
23. The data sets include income and life expectancy information all the way to year 2040.



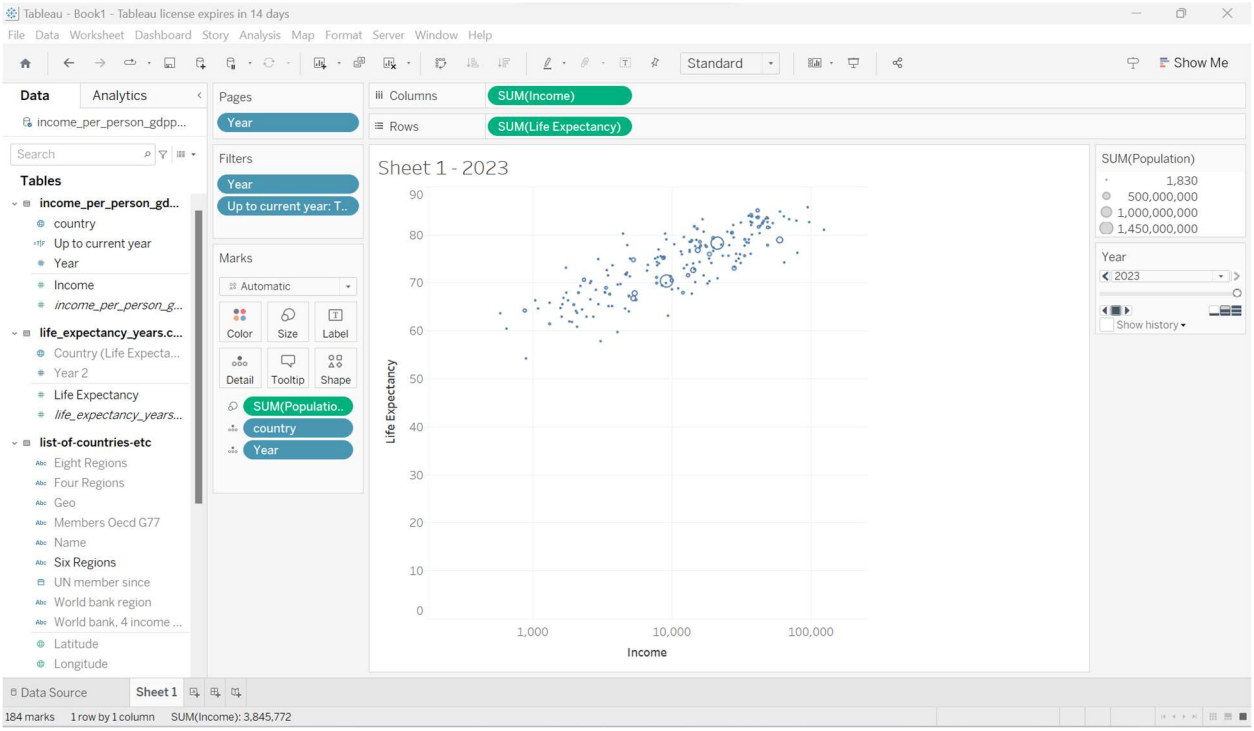
24. We will create a functional filter to reveal the moving bubble charts to the current year (i.e., 2023). Specifically, we will create a calculated field to set the time ceiling to be the current year (e.g., 2023). FYI, the to-be-typed formula in the calduated field box is [Year] <= YEAR(TODAY())



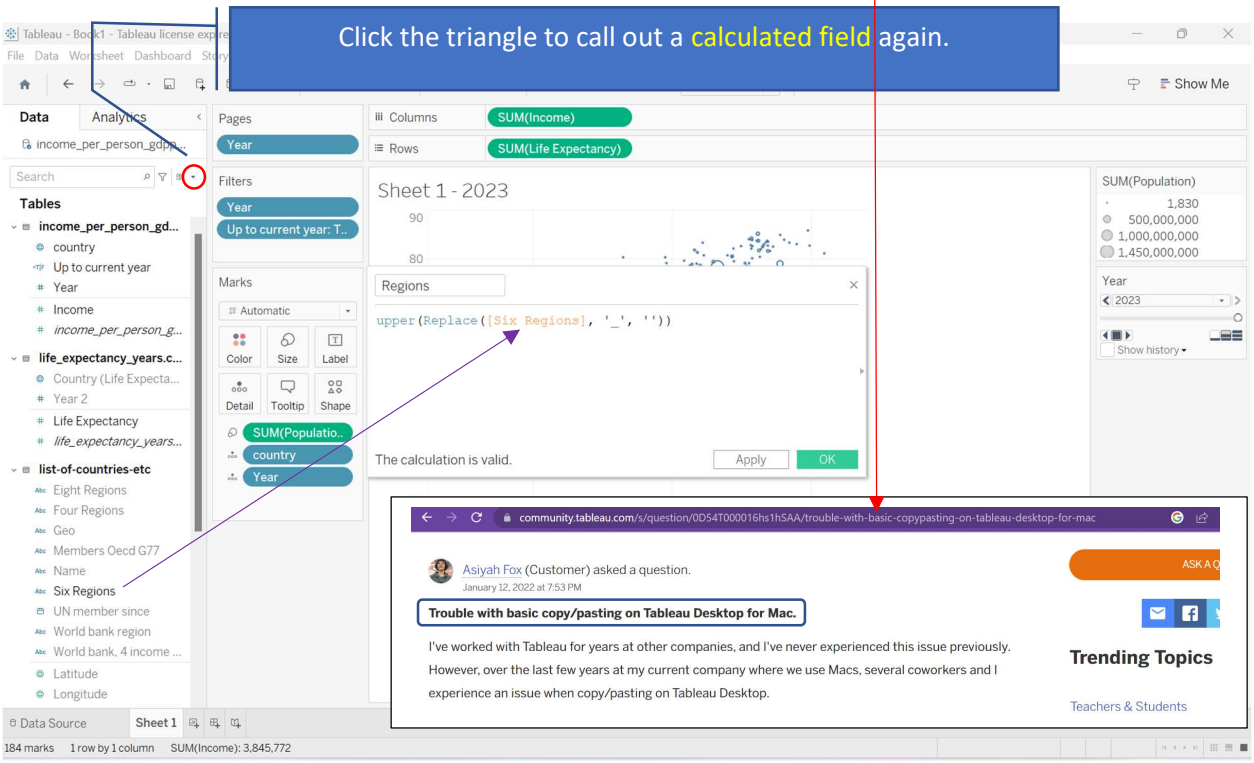
Hint: Check the variable ***Year***'s data type. If we do not update variable ***Year***'s data type (from ABC to Whole #), we won't be able to adequately create a "calculated field." Thus, we should double click the ***Year*** variable and then change its data type from ABC (string) to Number (whole).

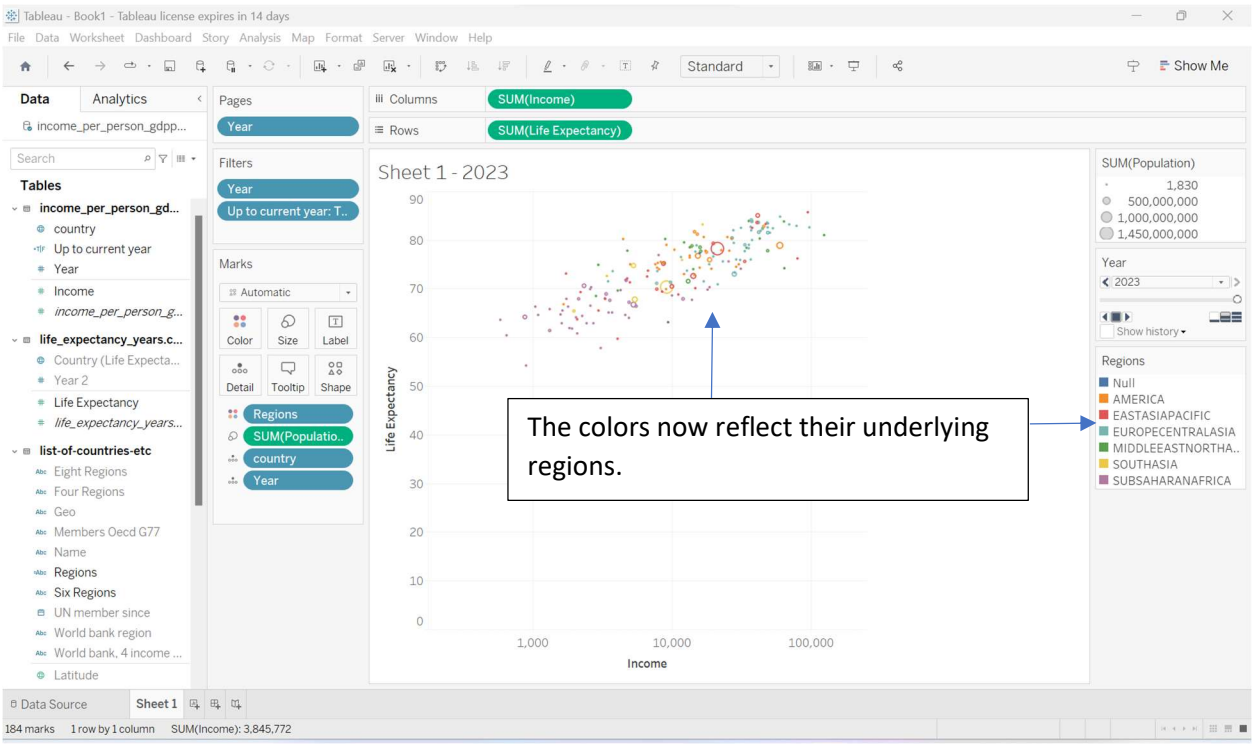


25. The current year shown on this revised bubble chart is year 2023.

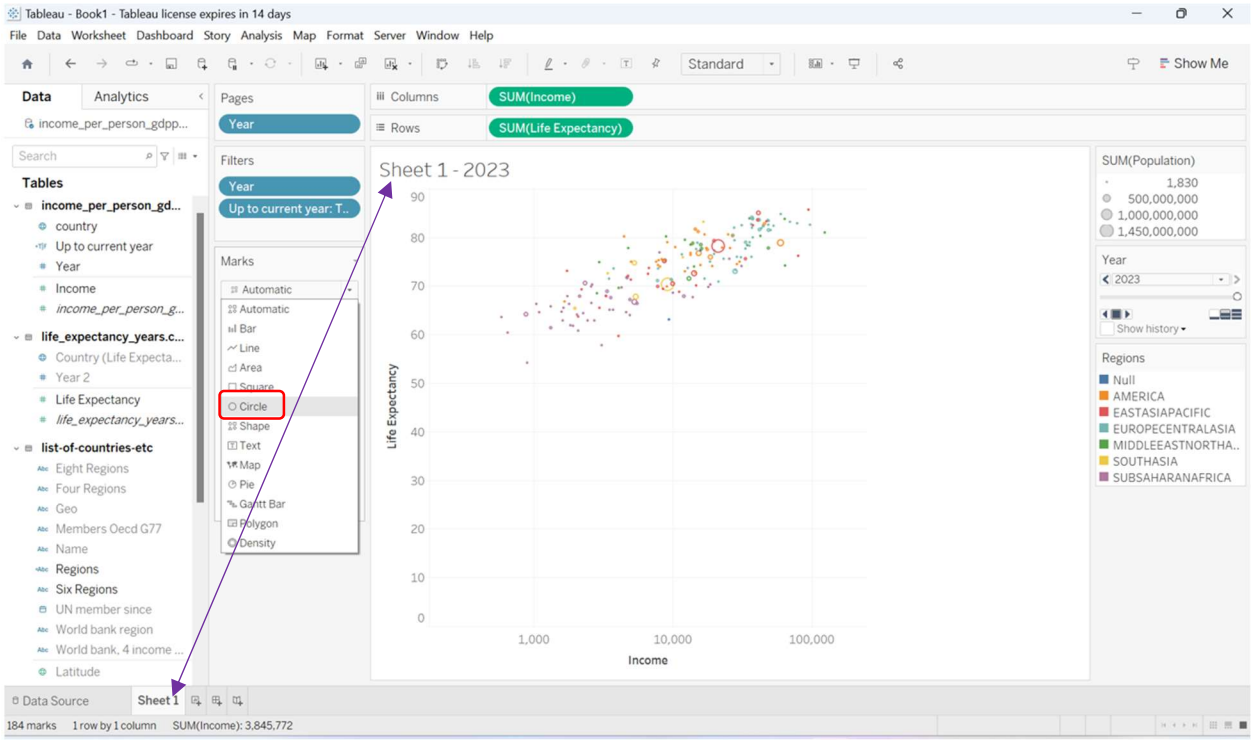


26. We will bring the "region" information to the chart. FYI, check the Appendix (p. 20) for Tableau's Replacement function; the to-be-**typed** formula is UPPER(REPLACE([Six Regions], '_', ' ')) → It's not recommended to copy and paste the formula onto the following box… or you might encounter errors.

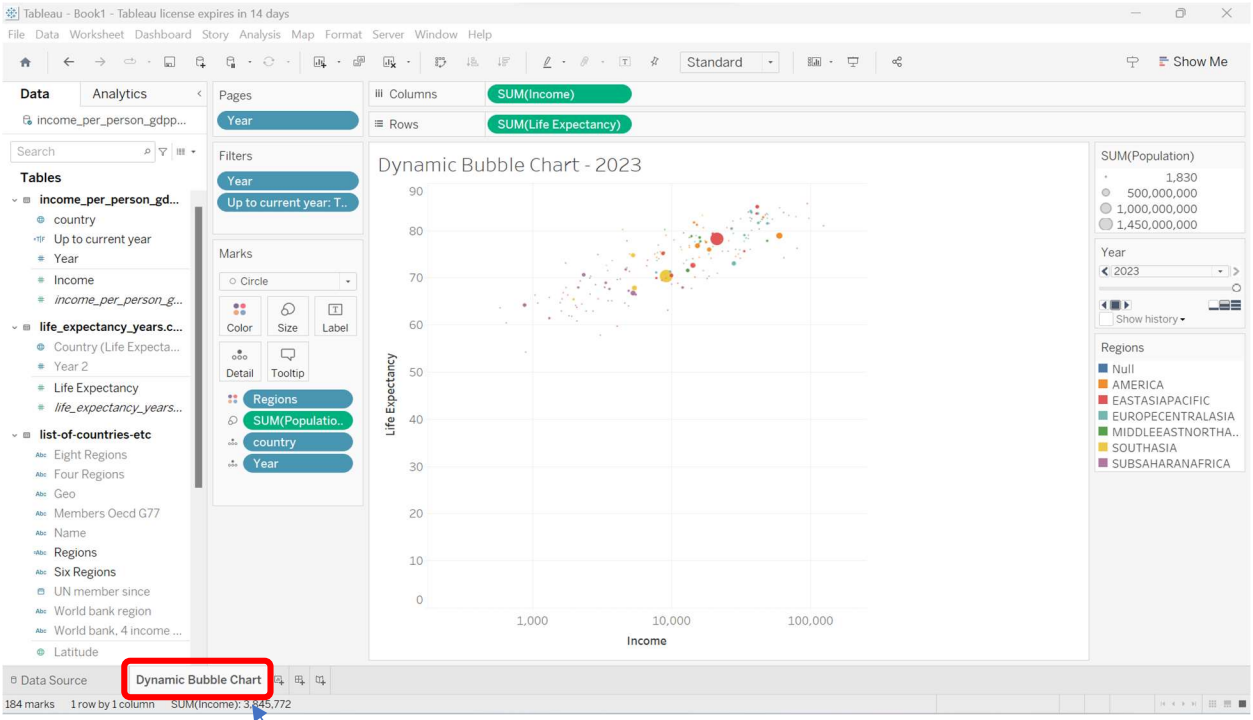




27. We can replace the dots with circles.

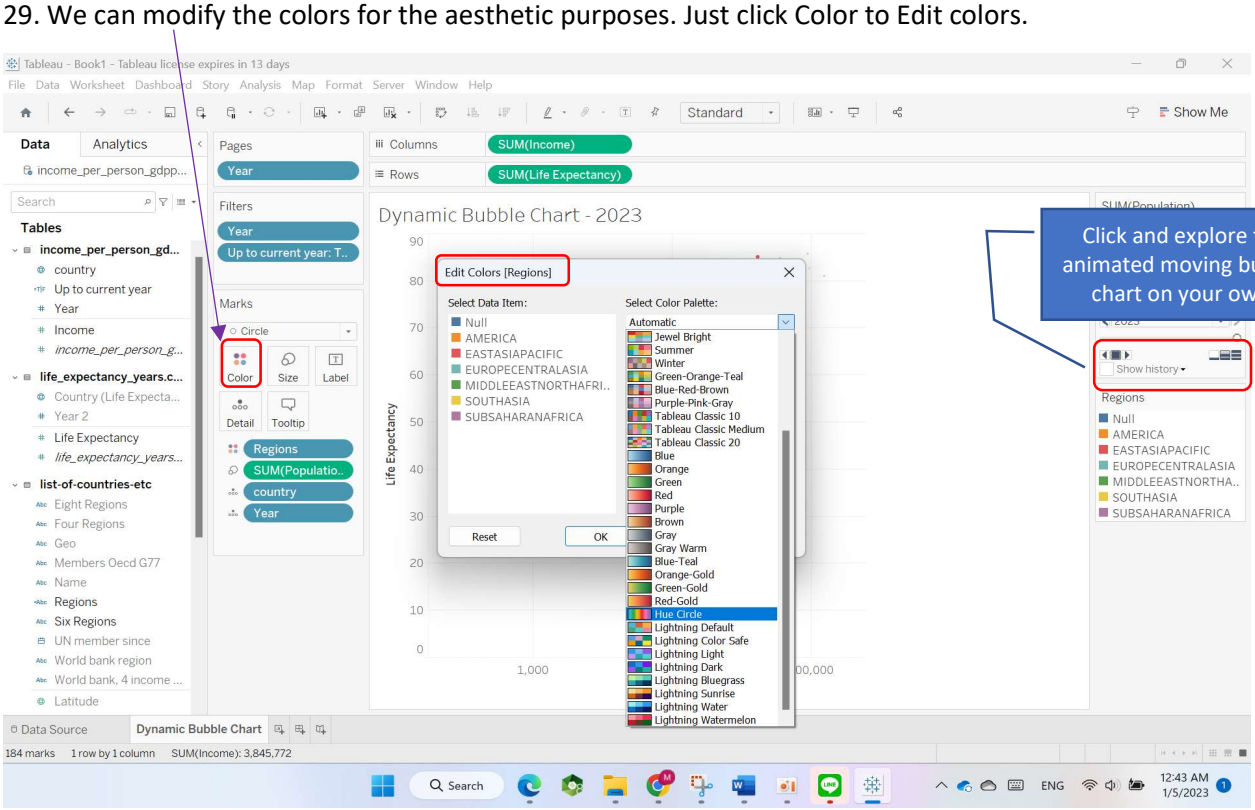


28. We will update the chart name (from “Sheet 1” to “Dynamic Bubble Chart”).

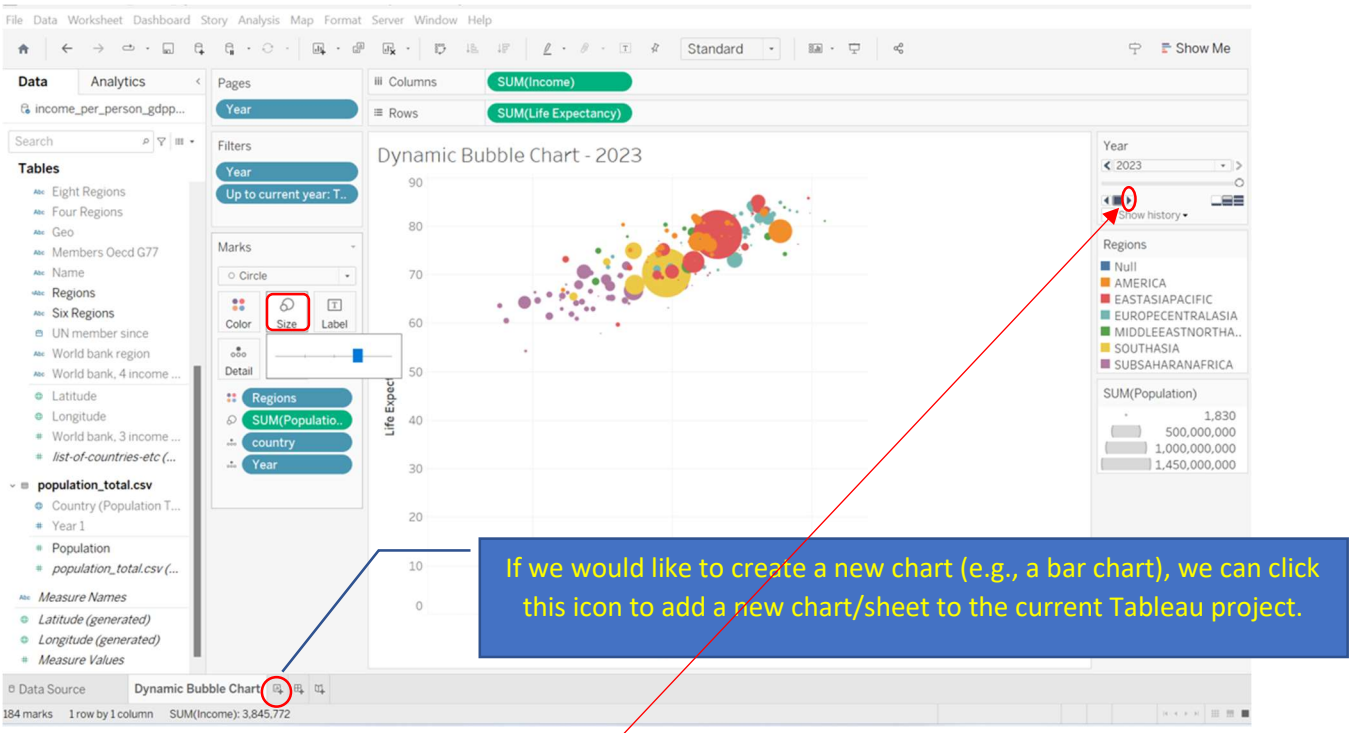


Hint: just click the name of the sheet and we can change its name.

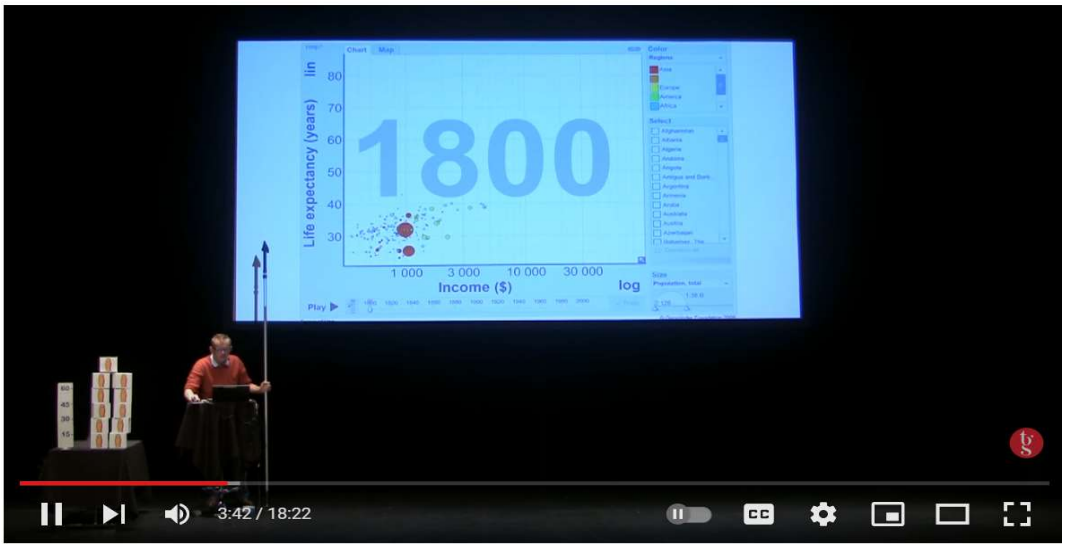
29. We can modify the colors for the aesthetic purposes. Just click Color to Edit colors.



30. We can also click the Size icon to further modify the size of the circles in this bubble chart.



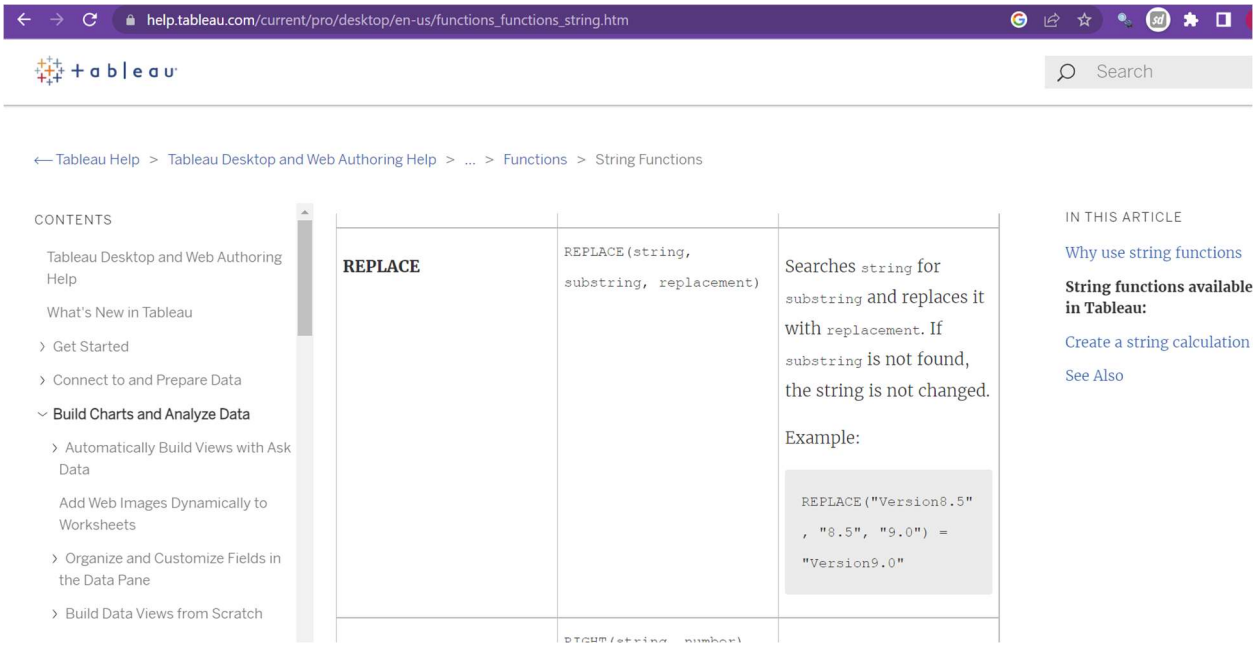
Finally, we are ready to demonstrate an animated moving bubble chart and tell our story by clicking a forward triangle button (see a red circle ◯ shown on the above). If you wonder how to give an impressive live presentation using the dynamic charts, you might find **Prof. Hans Rosling**'s presentation of relevance. Please click the underlined link for a MUST SEE presentation: Please do watch the first four-minute video, and then feel free to briefly scan through the rest of this YouTube video.
https://www.ted.com/talks/hans_rosling_new_insights_on_poverty?referrer=playlist-the_best_hans_rosling_talks_yo&autoplay=true



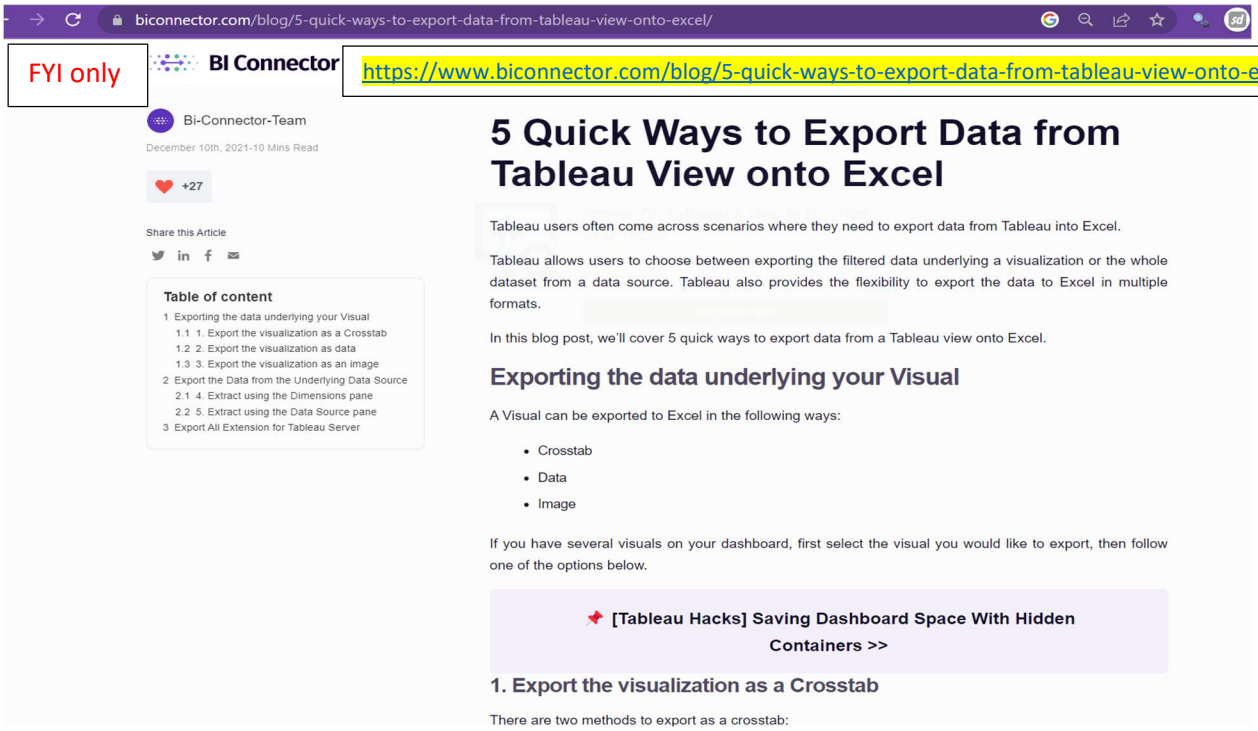
Correlating income and life expectancy throughout history | Hans Rosling | TGS.ORG

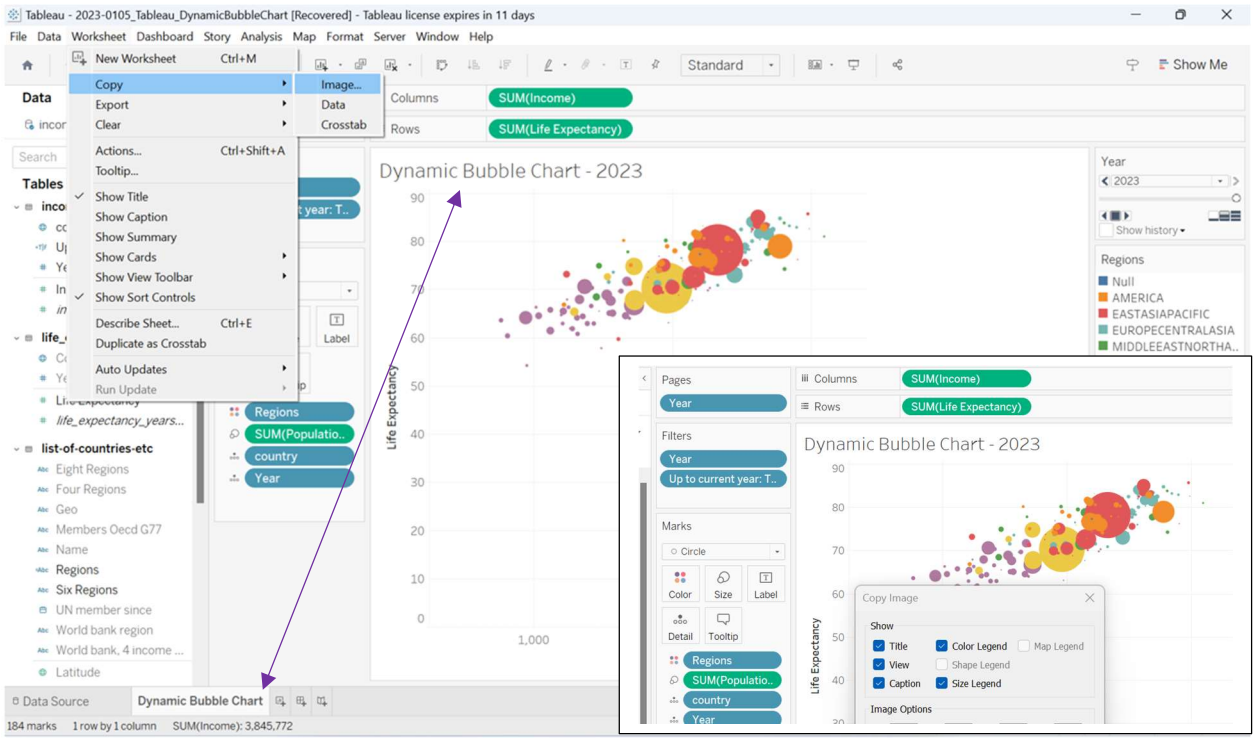
Appendix:

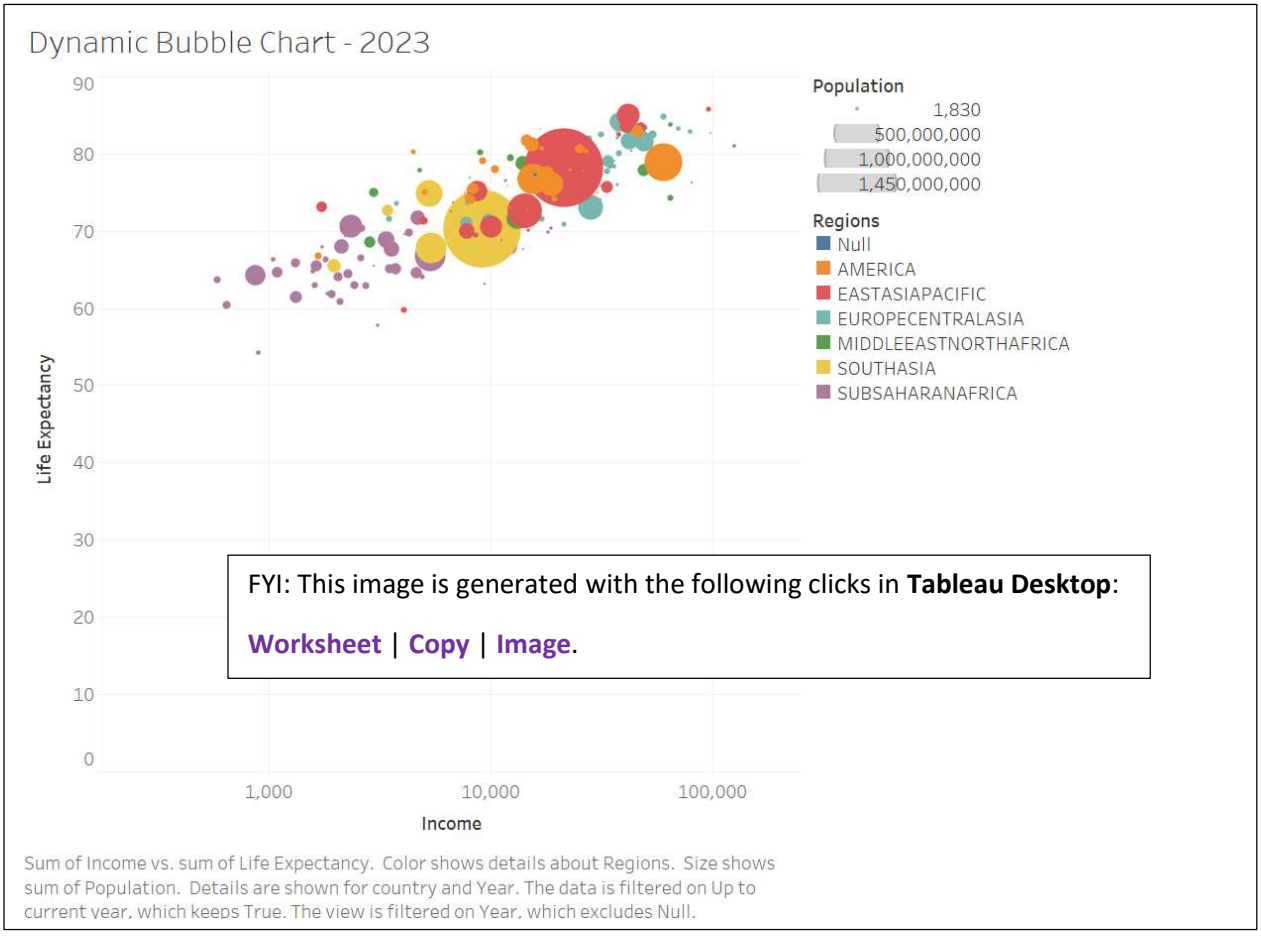
1. Tableau's main functions (e.g., Replacement) could be found online: https://help.tableau.com/current/pro/desktop/en-us/functions.htm



2. Feel free to Google search for additional Tableau tips (e.g., how to export data from Tableau View onto Excel or a relational database).



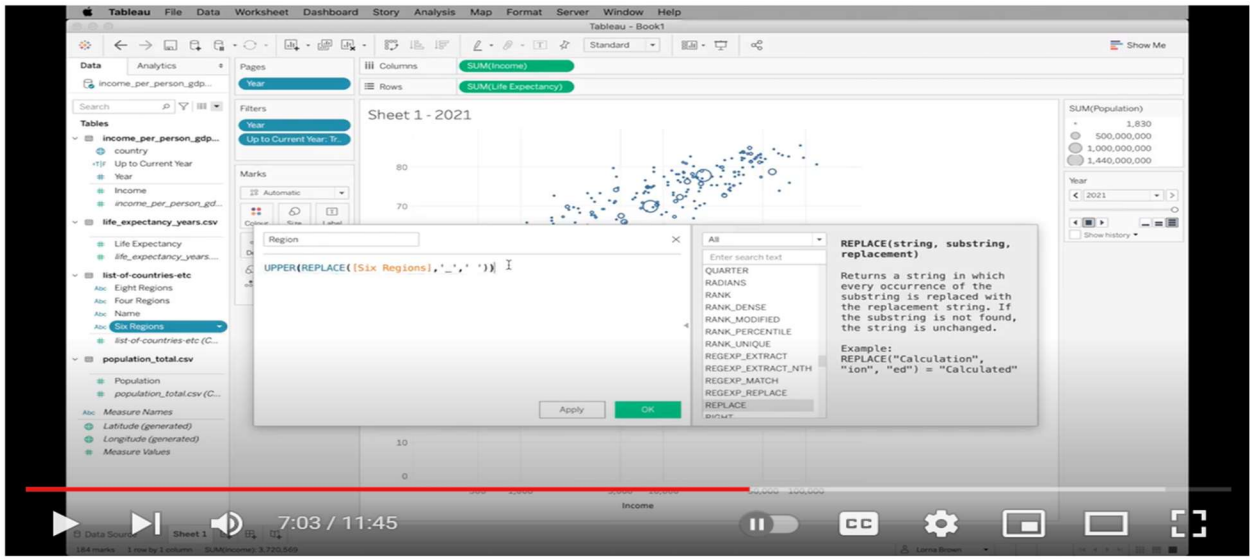




FYI: This image is generated with the following clicks in **Tableau Desktop**:

Worksheet | **Copy** | **Image**.

3. If you (the audience) prefer to watch a Tableau bubble chart creation video, you may find the following two YouTube videos of help.

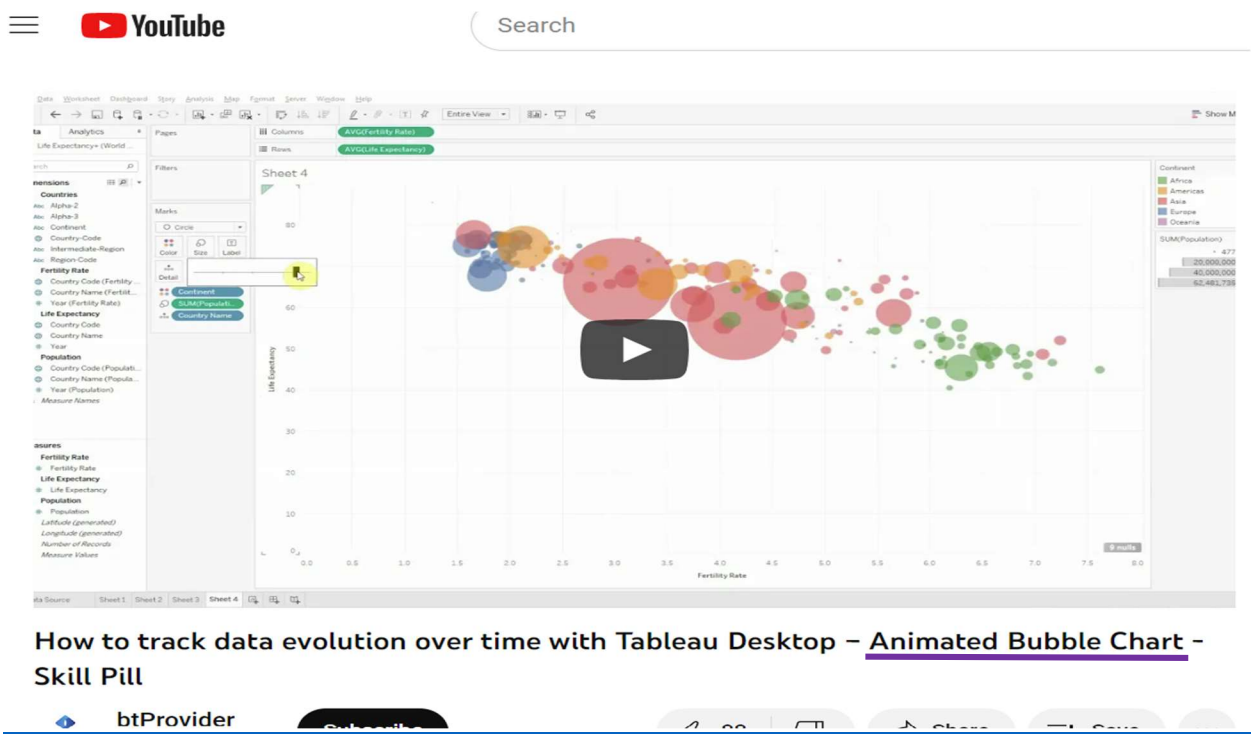


#WOW2021 Week 11 | Tableau : Can you recreate the work of Hans Rosling?

Workout Wednesday
1.93K subscribers
Subscribe
11
Share

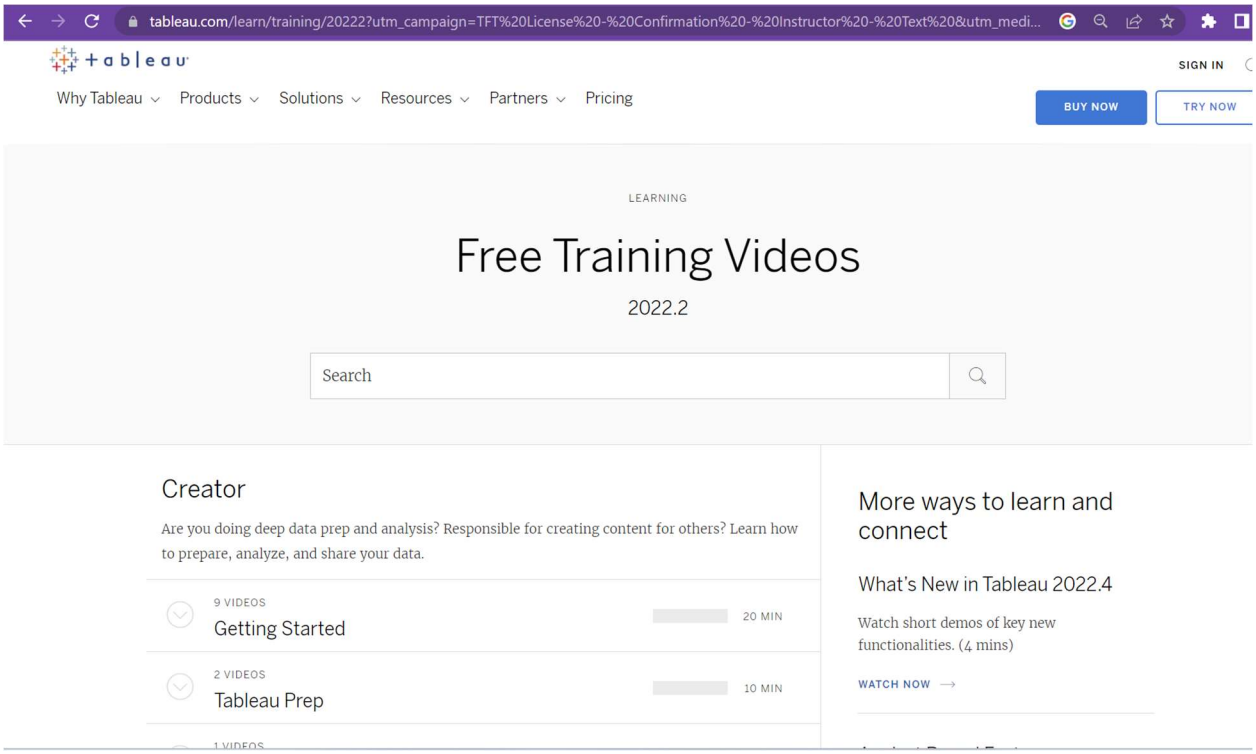
https://www.youtube.com/watch?v=LaCprFuwmHk

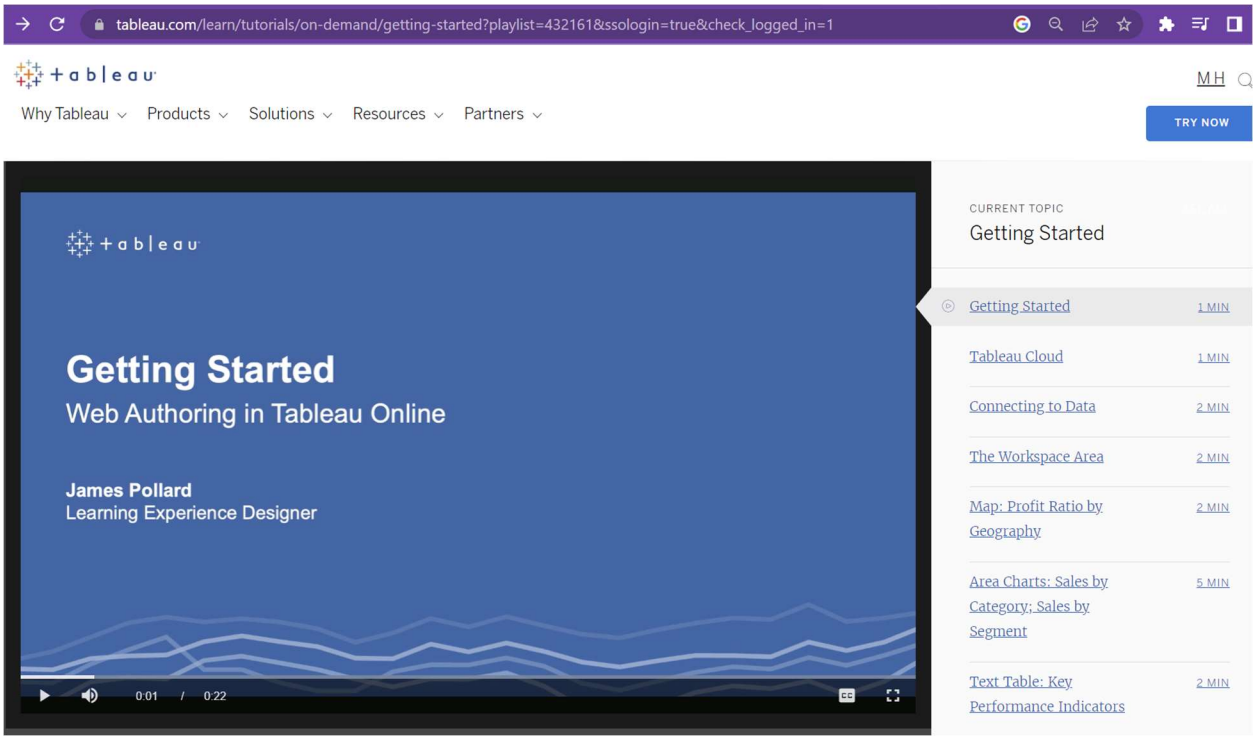
****



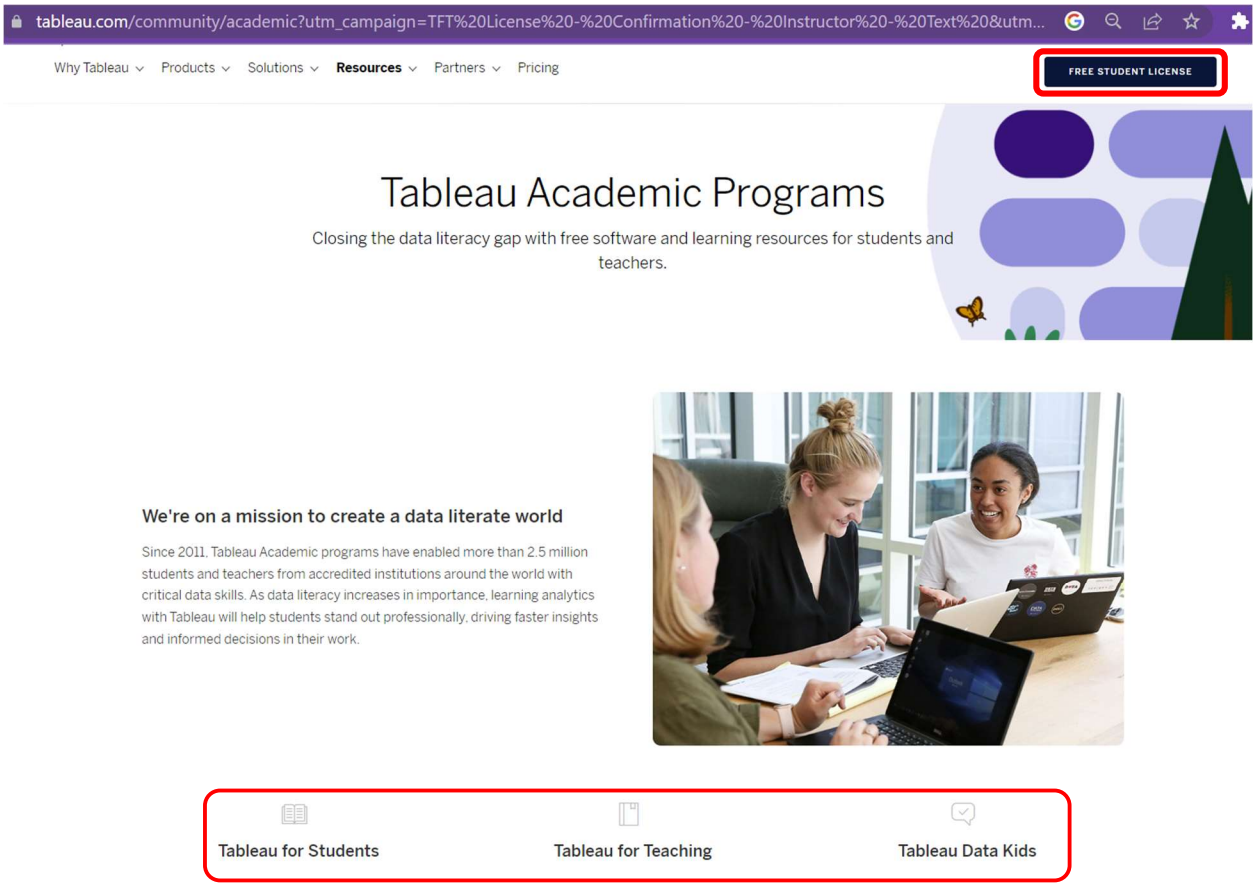
https://www.youtube.com/watch?v=lK6AE3u6LGs

4. Tableau also offers quite a few training videos (https://tabsoft.co/3QuMLGp). **Happy learning!**

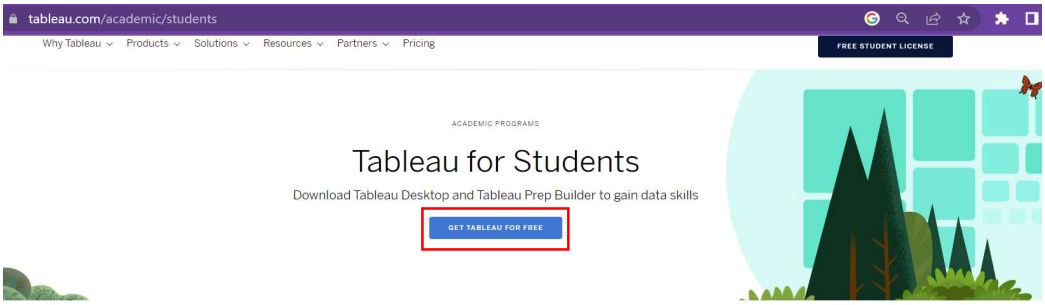




5. To learn more about available Tableau programs, check out its Academic Community page: https://tabsoft.co/3vOkq4m



6. A student license is FREE. Check here for details: https://www.tableau.com/academic/students



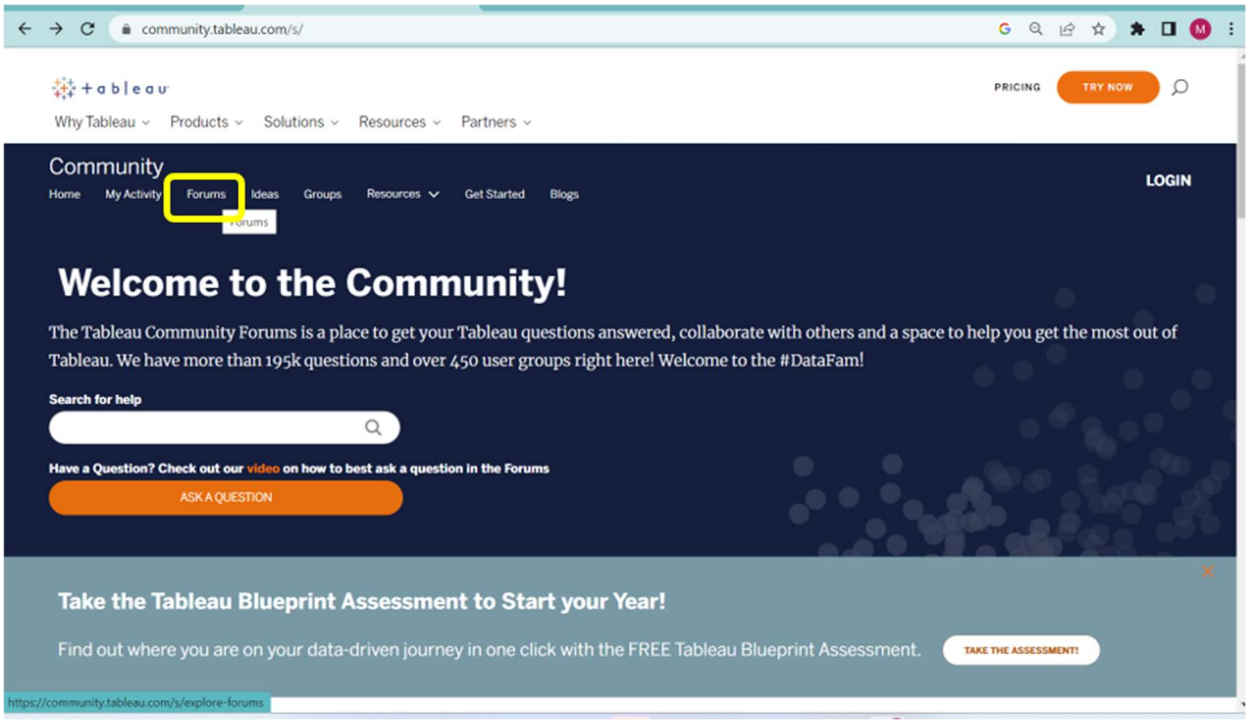
We offer free one-year Tableau licenses to students at accredited academic institutions through our Tableau for Students program. Receive access to our entire eLearning suite once verified.

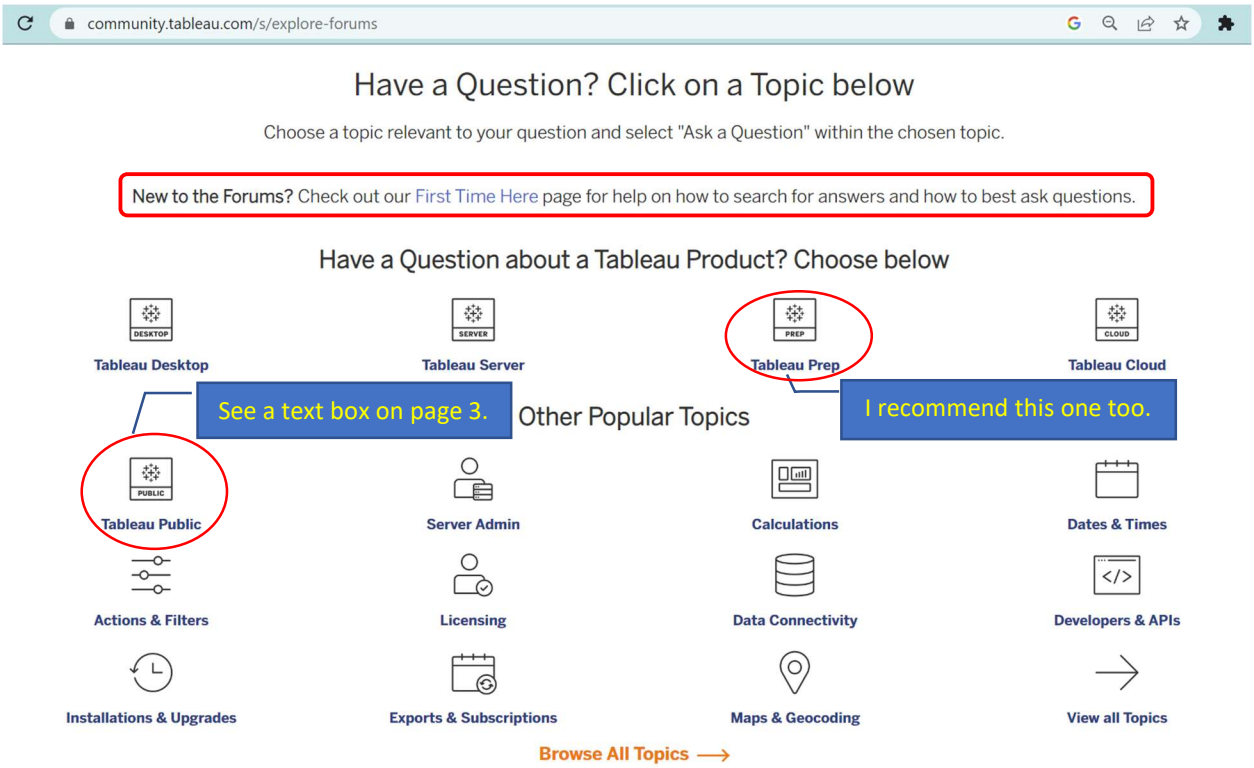
FEATURED CONTENT

Student Stories

This summer, we reached out to our student community and asked them to tell us about their journey with data, and to see what advice they have for students interested in learning data analytics. And as we head into

7. Explore the Tableau Community (https://community.tableau.com/s/). For example, click the Forums and then check the **Tableau Public**!





Acknowledgement: Special thanks to Prof. Gary H.T. Chao for his constructive comments toward a draft version of this 30-step Tableau Bubble chart learning tutorial.