Information Visualization using Tableau

Erin Thompson, Ruben Canlas Jr
Questions? Email cuviz@clemson.edu.

This manual is for Tableau 10.3. Other versions will generally be similar, but no two editions are exactly alike. If you have a version other than 10.3, keep in mind that the settings and display might be slightly different.

Contents

1. Introduction
2. Data formatting and insertion into Tableau
3. Worksheet layout
4. Dashboards
5. Stories
6. More information
7. Watershed example
8. Intermediate vizzies

1. Introduction

Tableau is an information visualization program. We are visual creatures. Most of us cannot look at an excel spreadsheet and say, “Yep, the company is losing money. We need to focus on sales in South Carolina” or “Wow that’s weird, this peak is not at all where I expected it to be,” or “Hey there is a correlation between rainfall and the presence of this bacteria.” You really cannot glance at a table and predict how it will look in a year. But, with data visualization, you can.

Information visualization (aka InfoVis) is the practice of presenting abstract data in visual format to make it easier for researchers to understand their datasets. In many cases, InfoVis has also made it easier for policy makers to understand data and act. For examples of famous people in history who used InfoVis to prove their point, look up Dr. John Snow or Florence Nightingale. Dr. John Snow used bars clustered over a map of London to represent cholera deaths. This helped him trace and stop the source of the epidemic in 1848-49 (see his map here). Florence Nightingale used polar area diagrams to convince policymakers that war-related deaths were actually avoidable (read a discussion of her diagrams here).

2. Data formatting and insertion into Tableau

You cannot begin to work with your data if it is not in the correct format. Tableau uses what they like to call “long” data sources, as opposed to “wide” ones**.** Both can contain the same amount of information. For a wide dataset, multiple values or responses are held within the same row. For a long dataset, each row contains a single value. For example, if you had data from multiple years, you could either have three rows with multiple columns containing the values, one per year. That would be a wide dataset. Conversely, you could have multiple rows, each with an entry in a column specifying year and a single data point. Your headers at the top would look different. The headers for the wide dataset might read: “Value for year 1,” “Value for year2,” and “Value for year 3.” For the long dataset, the headers might only be “Year” and “Value.” The long dataset is typically longer since it has more rows. Both contain the same information. The long file might look redundant; it is not usually intuitive for us, but it will be much easier to use in Tableau.

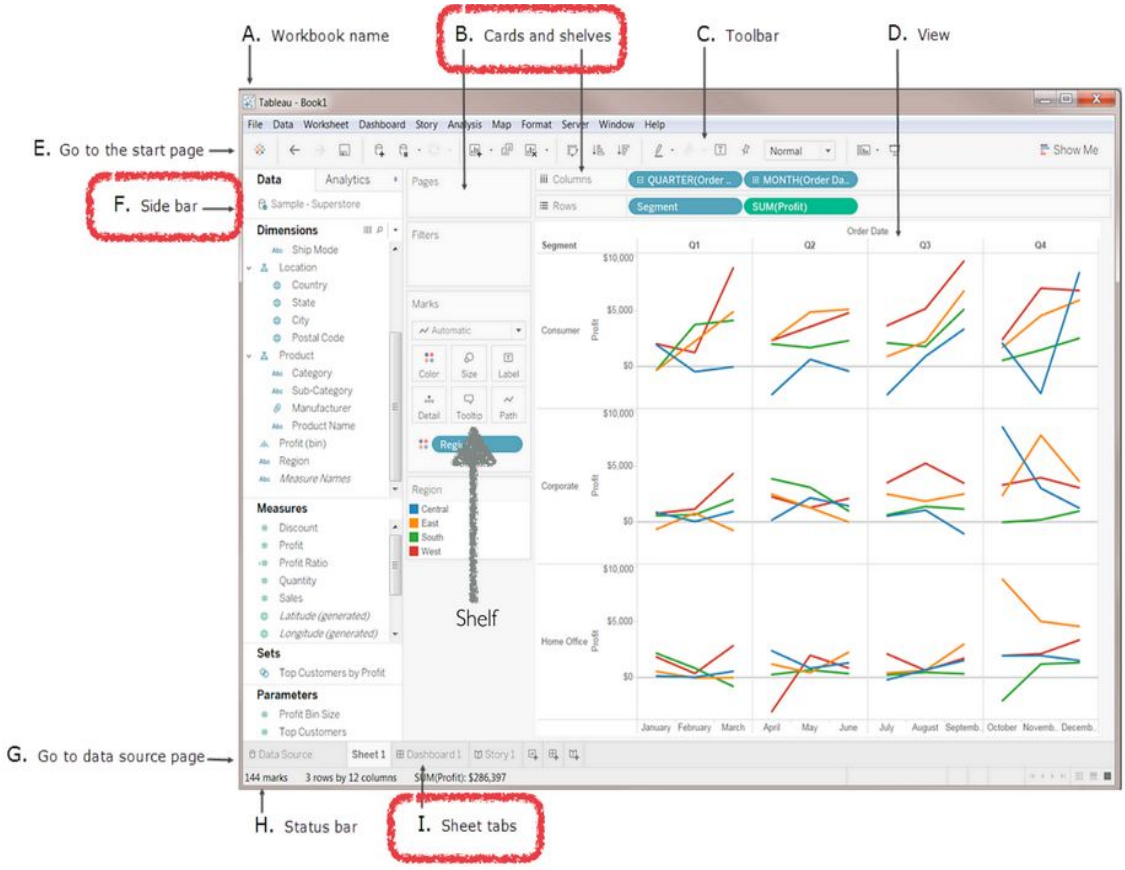
Now, let’s jump right into using Tableau. First, we must import our data. To put the data into Tableau:

1. On the Start Page under the Connect menu, click on Excel.
2. Open the Excel file.
3. Click and drag a sheet that you want into the canvas. Make sure your data is in long format.
4. Data cleaning
    a. You can rename some of the section headers if you want. Do that by double clicking the name or clicking on the triangle above the name of that category and to the right.
    b. Hide column names from view if are not going to be working with them and do not want to see them. This step is completely optional. You can go back and unhide them later.
    c. Change what type of data a column is by clicking on the symbol in the left top corner of the column name. Tableau automatically decides what type of data it thinks that column is.
        i. Types of data include: string, geographic location, number (decimal), number (whole), date and time, date, and Boolean.
        ii. Different types are processed differently; make sure that it knows what type of data you have.

3. Worksheet layout

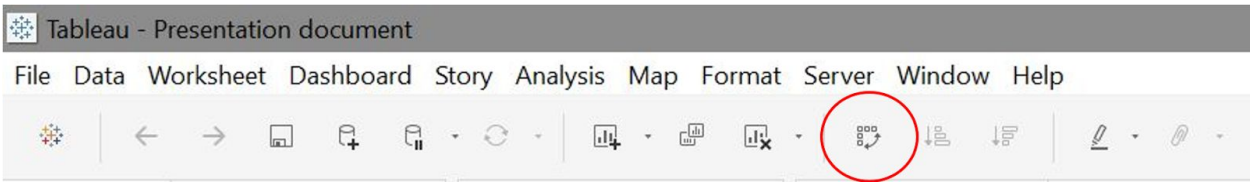
To the left of the worksheet pane, Tableau has grouped the data according to “Dimensions” and “Measures.” That’s because it looked at what you had and decided if it was text, a geographic location, numbers, or something else. This goes back to step 4c of Data cleaning under Data Formatting. Dimensions generally make axis titles while Measures are going to be your points. That is because Tableau assigns “dimension” to anything it sees as a discrete category of information. For example, a date such as May 8, a geographic location such as latitude, or a name. Measures will be lists of numbers. This includes sales, profit, counts of something, and more.

The image below shows several important points of interest of a worksheet.



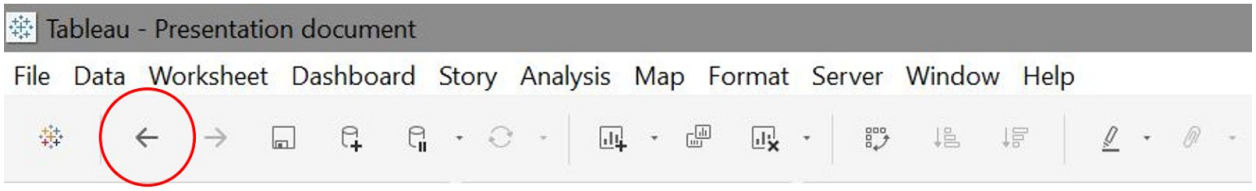
- Terminology
    - A Dimension or Measure is enclosed in a draggable, capsule-shaped container called a pill.
    - The boxes containing Marks, Filters, and Pages are called cards.
- Charts and graphs

- To create a chart, drag a Dimension to Columns and a Measure to Rows.
- Change the type of chart by clicking on the drop down box in the Marks card.
- Different charts/ graphs are used to show different things about your data. What you use should make sense and be easy for your audience to understand.
- Switch the x and y axis of your chart by pressing this button at the top.



- The Show Me feature
  - Show Me gives you some suggested diagrams that are possible given the Dimensions and Measures you have put in the Rows and Columns. For example, if you drag a Dimension in Rows and a Measure in Columns, Tableau automatically displays a graph. But if you click Show Me, it displays more possible visualizations.
  - You do not even have to drag pills into Rows and Columns. Just use Ctrl+click to select a Dimension and a Measure, then click the Show Me feature to see various charts and graphs possible based on what you have selected.
- Use the Marks card
  - Color
    - To show the data with certain aspects of a field colored, drag a pill and put it on top of the color box.
    - Tableau detects categories in your Dimension and assigns different colors to distinguish them in the graph.
    - Tableau automatically makes a legend card and puts it on the upper right side of the screen.
    - Change the colors by clicking on the Color tab in the Marks card, selecting Edit Colors, choose Select Color Palette, select one, press Assign. This changes what color palette is used. If you wanted to further change the color of each data item on a legend, click on the item name and then on the color you want.
  - Size
    - Resize objects such as points or shapes by clicking on the size option.
  - Detail
    - Adds the pill that you place there as a detail on the chart.
  - Label
    - Provides labels on your chart based on the pill that you put there.
    - These can be moved by clicking on that object/ slice/ dot/ bar that they are labeling and then clicking and dragging them to a desired location.
  - Tooltip
    - This is what appears when you hover over a point on your chart with the mouse.

        - You can change what the tooltip says and how it looks here. When you add a Dimension or Measure to your chart, it will generally also put it in the tooltip.
        - To add something to the tooltip without putting it physically on your chart, drag a pill here.
- Null values
    - Tableau alerts you if your data has null values. The alert is displayed at the bottom right of a chart. Null values can skew your results, so you can deal with them by clicking on the gray alert. This lets you set all of the nulls to a value or filter them out.
- To undo an action, click on the left arrow at the toolbar on top. The right arrow redoes an action.



- Navigating through sheets
    - Just like Excel, worksheets are organized by tabs at the bottom of the screen. You can click each tab to activate that sheet.
    - Rename a sheet by double clicking on its name in the tab.
    - The Data Source tab gives you a spreadsheet-like view of your raw data.. This is shown in the green oval below.
    - To duplicate a tab right click on it and select Duplicate.
    - Create a new worksheet by clicking on the option shown in the blue circle below. The two options to the right add a new dashboard and story, respectively.

- Filters
    - Filters can be used to show specific, related data points. For example, one of the files I worked with is given in section 7 of this document. It includes the amounts of bacteria found on specific days. If we only wanted to look at the canine bacteria count, we could filter the data so that only that was showing. The same is true for any category. Filters are also useful for removing null data. In the aforementioned document, nothing but E. Coli was sampled for prior to

September of 2016. If we wanted to look at the relevant data for canine bacteria, we should filter out the dates where there are no values.
    - To use a filter, drag a Dimension or Measure into the filters card. Filter a category by clicking on its checkbox. Unchecking it will deselect that category.
    - Edit the filter by clicking on the triangle on that pill and selecting “Edit Filter.”
    - If you click on the arrow on a filter pill, you can select show filter. This produces a filter card on the far right of the page where you can limit which items show. There are two small triangles to the right of the filter name. If you click on the one farthest from the name, you can edit the filter and even change how to limit the values. For example, you can choose between “Single Value (list),” “Single Value (dropdown),” “Multiple Values (custom list),” etc.
- Maps
    - Tableau automatically recognizes major cities and geotags them. For more accurate geotagging, you can also add latitude and longitude information on your dataset. This allows Tableau to overlay your data onto a map.
    - Tip: To make a map, drag longitude to columns and latitude to rows.
    - You can drag Measures to size or color to adjust the attributes of the circles.
    - If you are using size and the dots are too small, click on the measure values box to the right and click the drop down arrow. Edit sizes to make them more useful.
    - To show streets, highways, and more, click on Map at the very top and then Map Layers. Select what you want to show.
- Aggregate data
    - Aggregate data are data summaries and are often used for statistical analysis. They can include information from multiple measures or dimensions. This is a very general term which means that information is combined and summarized to show a specific detail.
    - Tableau usually aggregates values to provide the average or sum, but there are several other possibilities.
    - Click on the triangle on a pill. Right under Dimension and Attribute, is the Measure option where you can aggregate the data.
    - Note: clicking on either Dimension, Attribute, or Measure will temporarily change the format of that data source.
    - Create a calculated field by right clicking on a Dimension or Measure in their respective cards and selecting Create, Calculated Field. You can then type in a formula. Rename the calculation when you are done. It will appear as a separate pill for you to use.

4. Dashboards

After you have created all of your worksheets, you can organize them into dashboards. Dashboards are useful for analyzing data across several worksheets. They can include one or more worksheets and can be created such that filtering or highlighting a portion of data in one

worksheet affects the others. This is useful for comparing data and seeing multiple views at once. They rely upon the worksheets which they include; when you edit a worksheet, the change is reflected in the corresponding dashboard. A worksheet can appear in more than one dashboard if you deem it necessary.

- The button to add a dashboard is to the right of the sheet names and between the buttons for adding a new worksheet and a story.
- When you open a new dashboard, you will see all of your worksheets on the left. Click and drag them into the main page to add them to the dashboard.
- Legends
    - Some legends might not automatically appear.
    - Correct this by clicking on the second triangle at the top of the map, legends, shape legend.
    - You can also move the legends around and change where they go.
- Tile vs float
    - When you add something to the dashboard, you will have one of two options selected: Tiled or Floating.
        - If you add an object while Tiled is selected, it will arrange it in specific ways to maximize the space of each sheet. This is good if you want things laid out in squares or equally spaced. It prevents overlapping.
        - If Floating is selected, the object will go exactly where you put it.
    - You can change an object from Tiled to Floating and back again by clicking on the triangle at the top and choosing that option.
- Use as filter
    - Click on one of the little triangles at the top of a worksheet in your dashboard and select use as filter. Now, whenever the user clicks on something on that worksheet, it edits other information using that data.
    - To make a filter apply to all sheets, click on the triangle for the filter, “Apply to Worksheets,” “All using this data source.” This can be helpful when you’re looking at multiple worksheets.
- Titles
    - You can double click on titles to change them and even remove them by selecting that option in the chart’s menu.
    - Try to change legend titles that read “measure names” or “measure values.” These can also be changed in that worksheet.

5. Stories

Once you have all of the dashboards created, you can create a story. Stories are a specific sequence of dashboards and worksheets created to relay information to an audience. You can walk your viewer through specific parts of your data, emphasizing what is important as you go.The interactivity of Tableau dashboards and worksheets is preserved in a story, which

means that it is much more than just a PowerPoint with graphs. Drag in dashboards and worksheets, add captions, and edit them to create a really interactive presentation.

6. More information

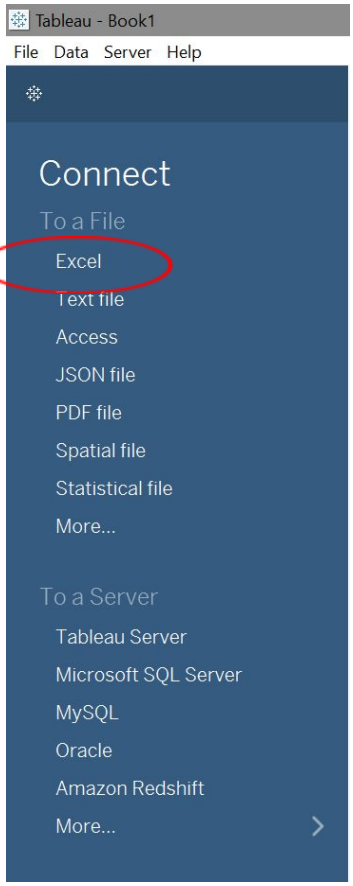
Tableau offers several online tutorials and beginner videos to help you familiarize yourself with the program. When you first open up Tableau, on the long gray banner to the right there is a link under Training which says “View All 87 Training Videos.” While you probably do not need to watch all of them, I recommend working through some if you get stuck. You are more than welcome to email me for clarification or if you have a question about something I did not mention in this document. My email is located at the top of this file.

If you want to see some examples of high quality work, go to the Public Tableau website.

Good luck!

7. Watershed Example

This is a step-by-step walkthrough to create a Tableau file using data from a project here at Clemson University. Katie Callahan, Director for the Clemson University Center for Watershed Excellence, allowed the CCIT Visualization Lab to visualize her data. In this portion of the document, you can follow along with the steps to create the file. There are screenshots of what these worksheets and dashboards should look like at the end of this section.



1. Getting the data
    a. Go to this link to open up the Google Sheet with the data.
    b. File > Download as > Microsoft Excel (.xlsx).
    c. Save the file to a location you can remember.
    d. Open Tableau. Connect to a File > Excel and find your file.
        i. The data is already in the correct format, however, testing for specific bacteria sources did not start until September 12, 2016. There are null values in columns lettered “H” through “O” until that date.
2. Sheet 1: A Map of Site Locations
    a. Drag Longitude to Columns and Latitude to Rows.
    b. Drag Site to Color in the Marks card. Color > Edit Colors… > Select Color Palette > Superfishel Stone > Assign Palette > OK.
    c. In the Marks card, change the size so that the dots are about 75% of the maximum diameter.
    d. In Map > Map Layers, check “Streets and Highways,” “County Borders,” “County Names,” and “Zip Code Boundaries” in addition to the default settings.
    e. Change the name of the sheet to “Map.”
3. Sheet 2: Positive Counts at Each Location
    a. Drag Human Presence, Swine Presence, Canine Presence, and Bovine Presence into Columns. On each one, change them from discrete to continuous and check to make sure that Measure > Sum is selected. They should now be green, not blue.
    b. Drag Site to Color.
    c. Change the name of the sheet to “Positive Counts.”
4. Sheet 3: All Data by Site
    a. Drag Date to Columns and check Day with May 8, 2015 formatting. Again, the pill should be green.

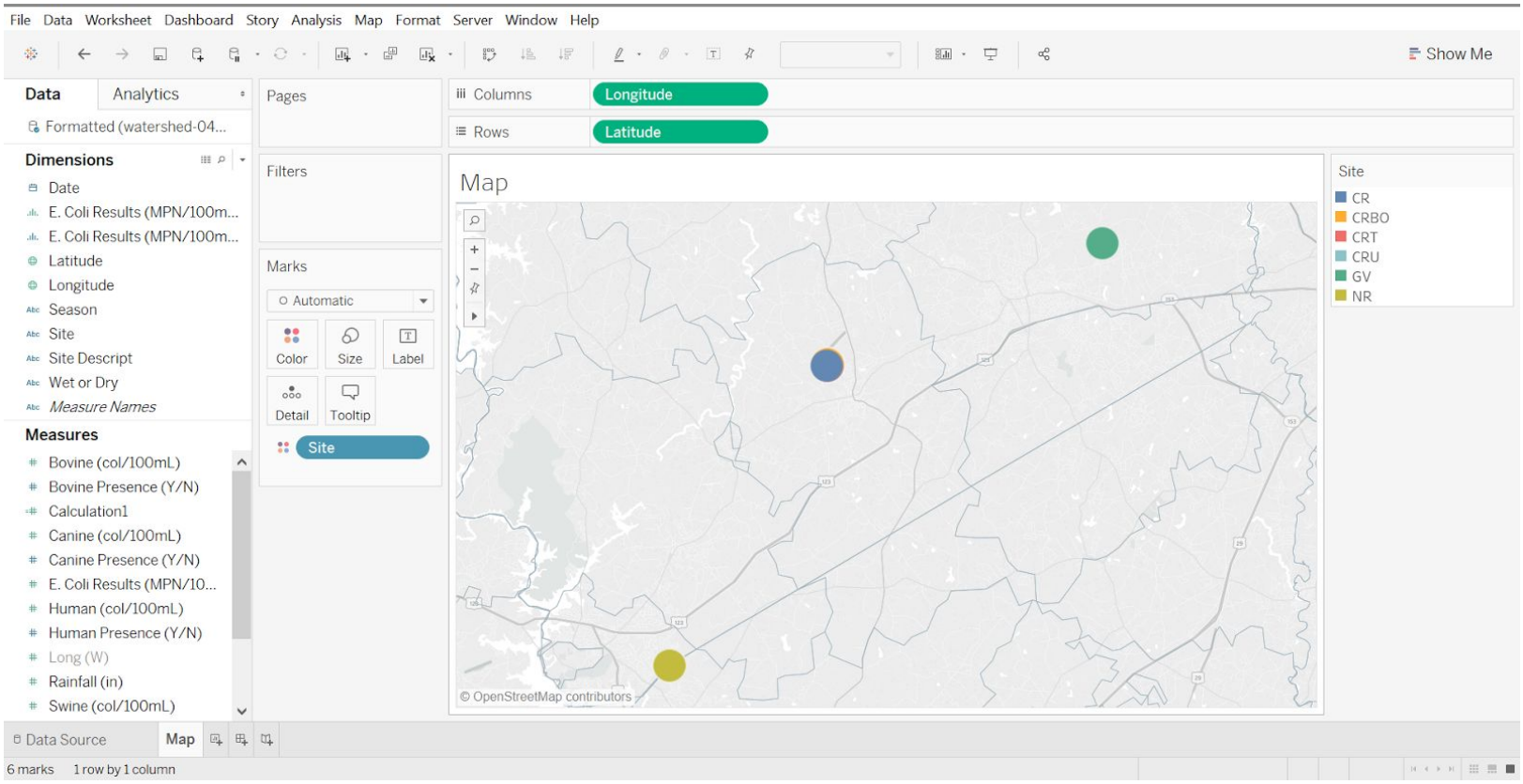
    b. Drag E. Coli Results, Bovine, Canine, Human, and Swine into Rows. For each, Measure > Average and make sure they are continuous.
    c. In the Marks card, under All, change the figure type from Automatic to Shape.
    d. Drag Site to Color and to Shape for All under the Marks card. There should be two legends to the right.
    e. Drag Site into the Filters card. Make sure that all are selected and press OK.
    f. It says there are nulls. That is okay. We already knew that sampling for the various bacteria sources did not start until later. If you want, you can right click on the gray box and select Hide Indicator.
    g. Change the name of the sheet to “Data by Site.”
5. Dashboard 1: Sites
    a. Create a new Dashboard. Make sure that Tiled is selected at the bottom of the Objects option card.
    b. Drag the worksheet titled Map into the middle of the dashboard. Put “Data by Site” right underneath it and “Positive Counts” below that. The three should be stacked with Map taking up the top half of the page and the other two sharing the bottom.
    c. Right click to hide all three titles. Right click on the Day of Date label and deselect Show Header.
    d. Decrease the height of Positive Counts and Map to increase the size of Data by Site. It should take up about half of the page. You do this so that the axis label is easier to read.
        i. It also helps if you decrease the font size. Right click on the vertical axis title and select Format. Under Axis > Default > Font, change the font size to 8.
        ii. Drag the line between the axis labels and the data to the right until you are satisfied with the information shown.
    e. Click on Data by Site then on the black triangle at the top. Filters > Site. The Site checkbox filter should now appear. Rename it “Click to filter by site” but push enter after the word “filter” and before the word “by” in order to have proper spacing in the textbox. Click on the triangle on that filter, Apply to Worksheets > Selected Worksheets > All on Dashboard > OK.
    f. There are two legends for the sites. Shorten this into one by first making both floating. Then hide the title for the shape legend. Decrease the size of the shapes textbox so that only the shapes themselves show. Move the color legend under the sites checkbox so that there is room to the left of it for the shapes. Move the shapes box there and carefully align them.
    g. Rename the dashboard “Sites.”
6. Sheet 4: E. Coli vs Rainfall
    a. Drag Rainfall into columns and E. Coli Results to Rows.
    b. Use the average of E. Coli Results and make sure it is continuous.
    c. Change Rainfall to a Dimension.
    d. In the Marks card, change the chart type to circle to a little under 50%.

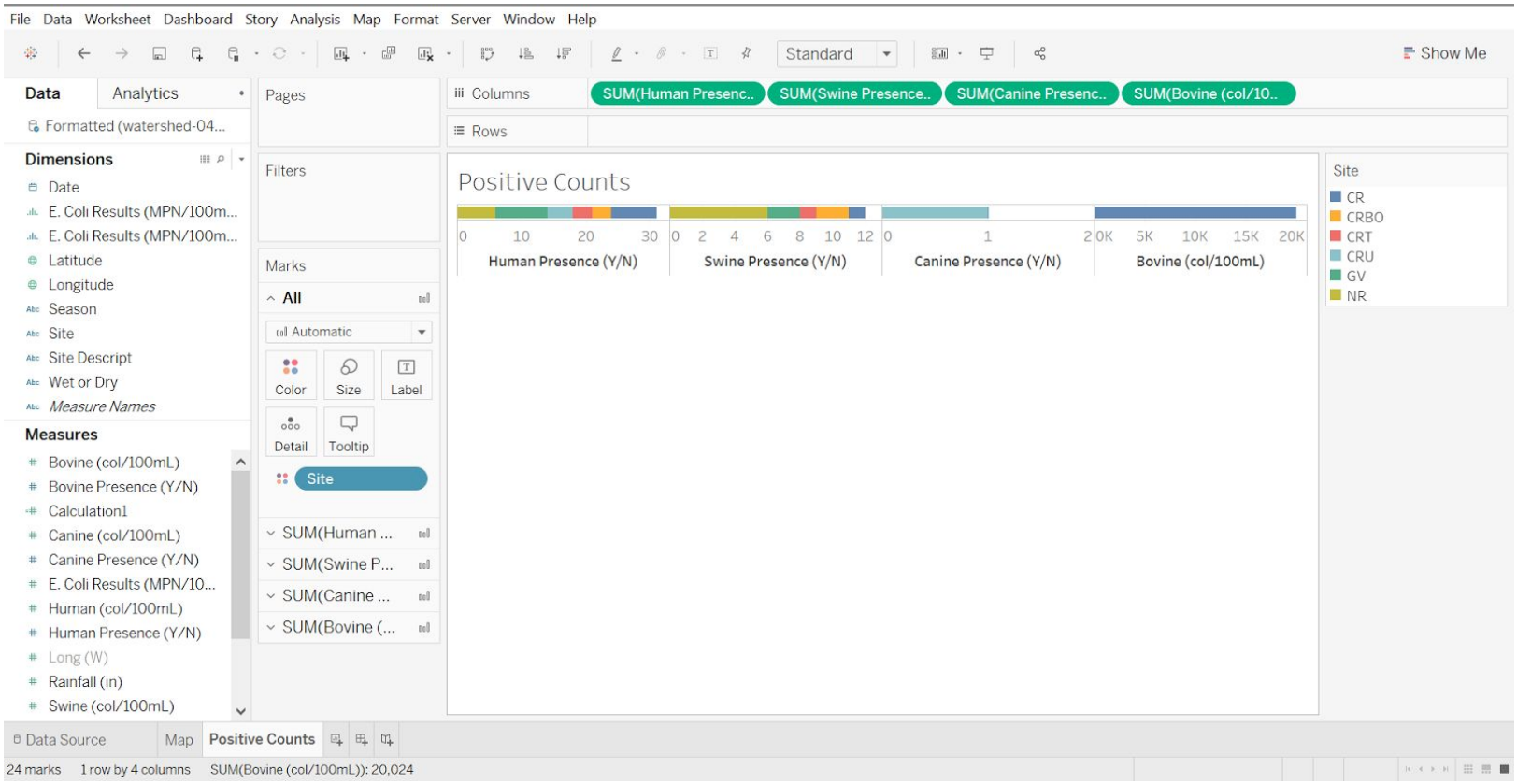
    e. Change the color of the circles to red and increase their size.
    f. Analysis > Trend Lines > Show Trend Lines
    g. Change the name to E.Coli vs Rainfall
7. Sheets 5-8: Rainfall vs Month for Specific Bacteria Types
    a. Drag Date to Columns. Press the white square box to the left of the name on the pill to expand it to quarters. Expand it again to show months.
    b. In Rows, put Rainfall and make sure it is the continuous sum.
    c. Change the type of figure from Automatic to Circle.
    d. Bring Bovine into color.
    e. Color > Edit Colors… > Stepped Color > 5 Steps > OK
    f. Drag Wet or Dry to Detail.
    g. There is only data for bovine from the end of August in 2016 to August 2017. Drag Date into filter. Range of dates > Next. Go from August 31, 2016 to the end of the data.
    h. Rename the sheet “Rain vs Month for Bovine.”
    i. Duplicate this sheet three times. Go into each one and change Bovine in the color filter to either Swine, Human, or Canine. Change the color as well. For example, make Human purple, Swine brown, Bovine blue, and Canine green.
    j. Rename each sheet to match the data presented. At the end, you should have four sheets labeled “Rain vs Month for Bovine,” “Rain vs Month for Canine,” “Rain vs Month for Human,” and “Rain vs Month for Swine.”
    k. Repeat steps (a) through (f) for E. Coli Results, skipping step (g) which would limit the time period.
8. Dashboard 2: Rainfall vs Month for the Four Bacteria Types
    a. Drag “Rain vs Month for Bovine,” “Rain vs Month for Canine,” “Rain vs Month for Human,” and “Rain vs Month for Swine” into a new dashboard. Set them up to be equally spaced squares.
    b. Show the dashboard title. In the title, type “Rainfall and Bacteria by Month.” Then decrease the font size and type “Each graph shows rainfall vs month, but the color is the concentration of each bacteria source in a sample.”
    c. Make sure that the color legends show.
    d. Rename the dashboard “Rainfall vs Month.”
    e. As a challenge, see if you can create a filter where selecting a site on the map will filter the data in this section to only show points from that site.
9. Sheets 9: Human vs Month
    a. Drag Date to Columns. Press the white square box to the left of the name on the pill to expand it to quarters. Expand the one on quarter to show months.
    b. In Rows, put Human and change it to average.
    c. Change the type of figure from Automatic to Circle.
    d. Bring Human into color. It should be the same color as the dots in “Rainfall vs Month for Human.” Make sure that it is set to average.
    e. As before, there is only data from the end of August in 2016 to August 2017. This time, those dates do not have points since they have no value in Human

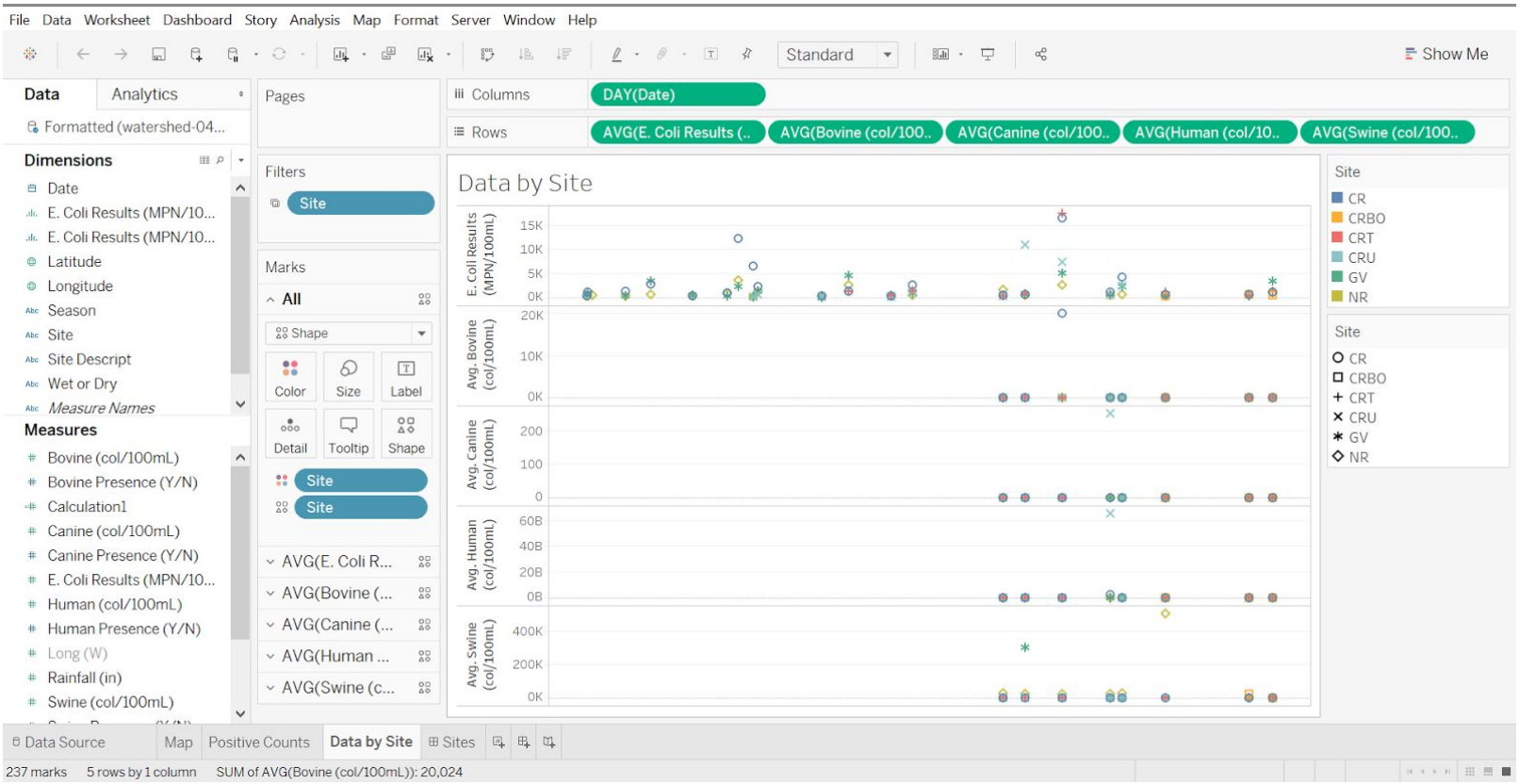
(col/100mL). At the bottom, there is a warning about null values. Click on that box and filter the data.

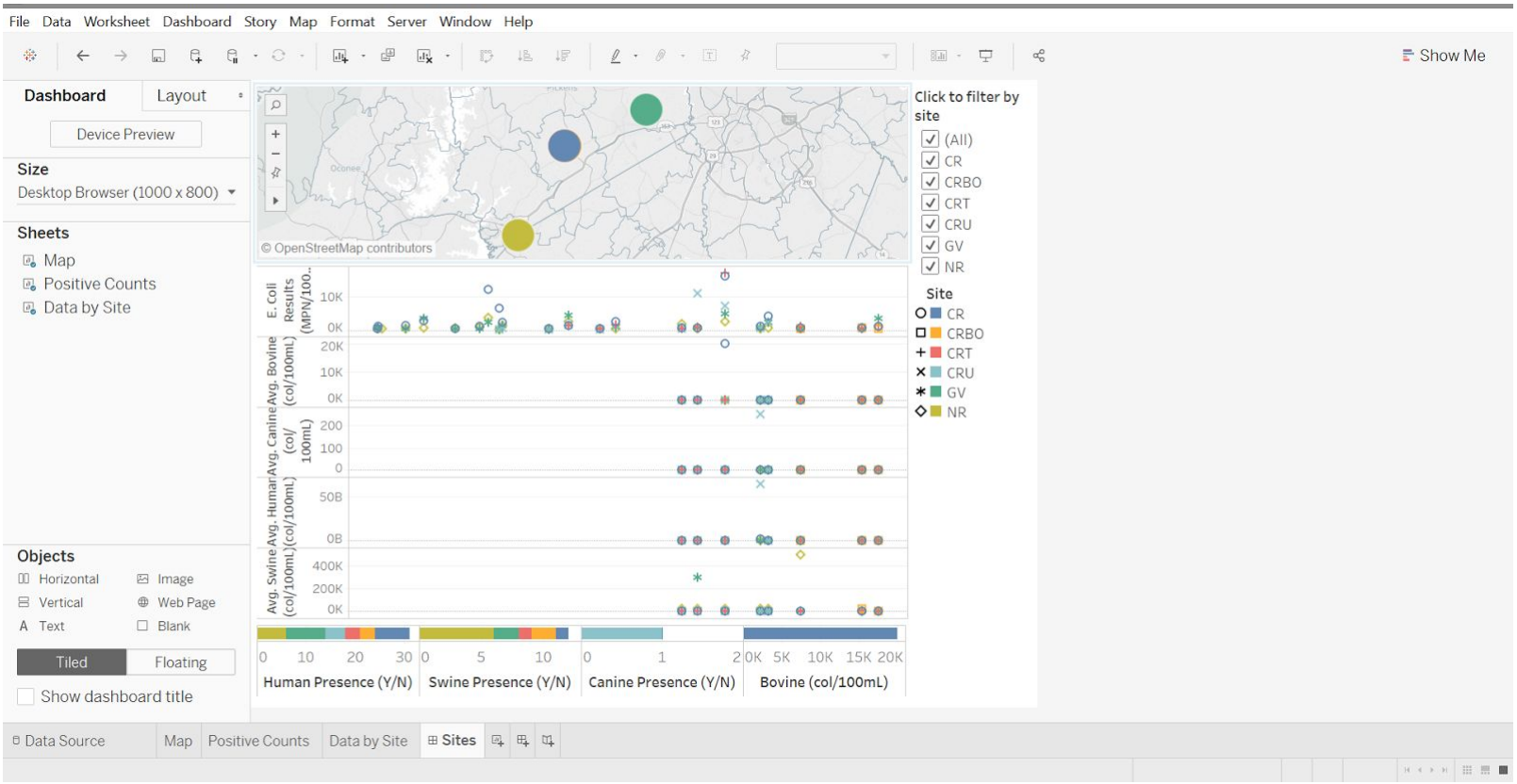
f. Rename the sheet “Human by Month.”
g. Notice that one dot stands out and the others appear to be zero. If you hover over them, you can see that this is not true. Duplicate the worksheet and rename the new version “Human by Month Excluding Outlier.”
h. In the new sheet, click on and exclude the point for January, 2017.
i. With that point gone, we can see that the data does vary.

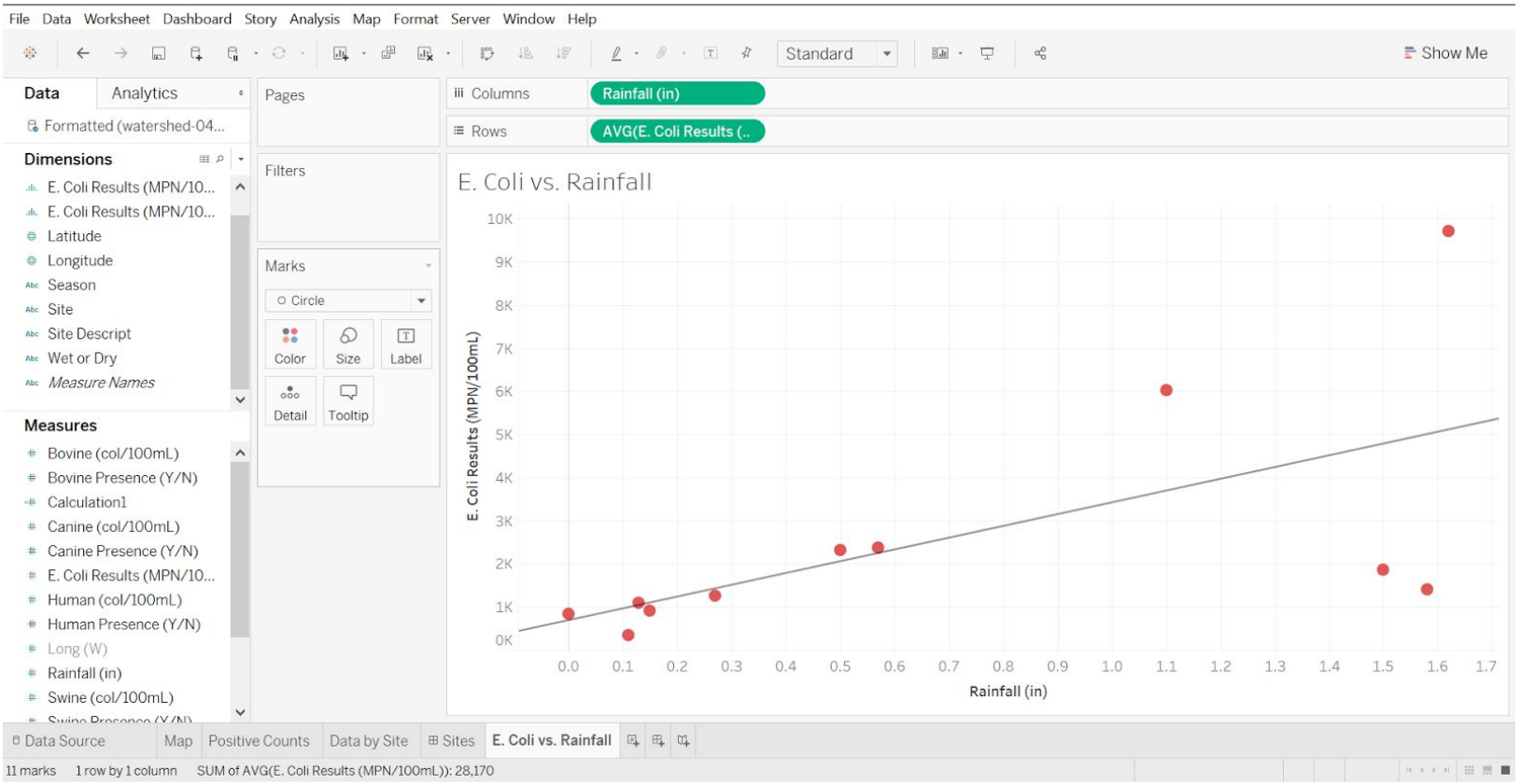
10. Dashboard 3: Studying Human Bacteria Results
    a. Add both “Human by Month” and “Human by Month Excluding Outlier” to a new dashboard. Put “Human by Month” on top.
    b. Hide the title for the worksheet “Human by Month,” but keep the title for the second.
    c. Under objects, click on text and drag it into the dashboard under the legend for the color shades. Type the following in size 9 font into the textbox:
        i. “Compare these two sheets. The bottom sheet is a duplicate of the top with one difference: the point for January is excluded. That point appears to significantly differ from the rest of the data. Without it, we can see smaller variations in the data and, perhaps eventually, a trend.”
    d. Rename the dashboard “Study of Human.”
11. Worksheet 10: Seasonal bar chart
    a. Put Site in Columns. Put the sum of Bovine Presence, Canine Presence, Human Presence, and Swine Presence in Rows. To do so, you need to make a calculated field, “Sum of presence count.” Put the average of E. Coli Results in Rows as well.
    b. Put Season on Color in Marks. Edit the colors so that they match their respective seasons. For example, you can use orange for autumn, green for summer, pink for spring, and blue for winter.
    c. Filter out the nulls.
    d. Rename the chart “Seasonal Presence.”

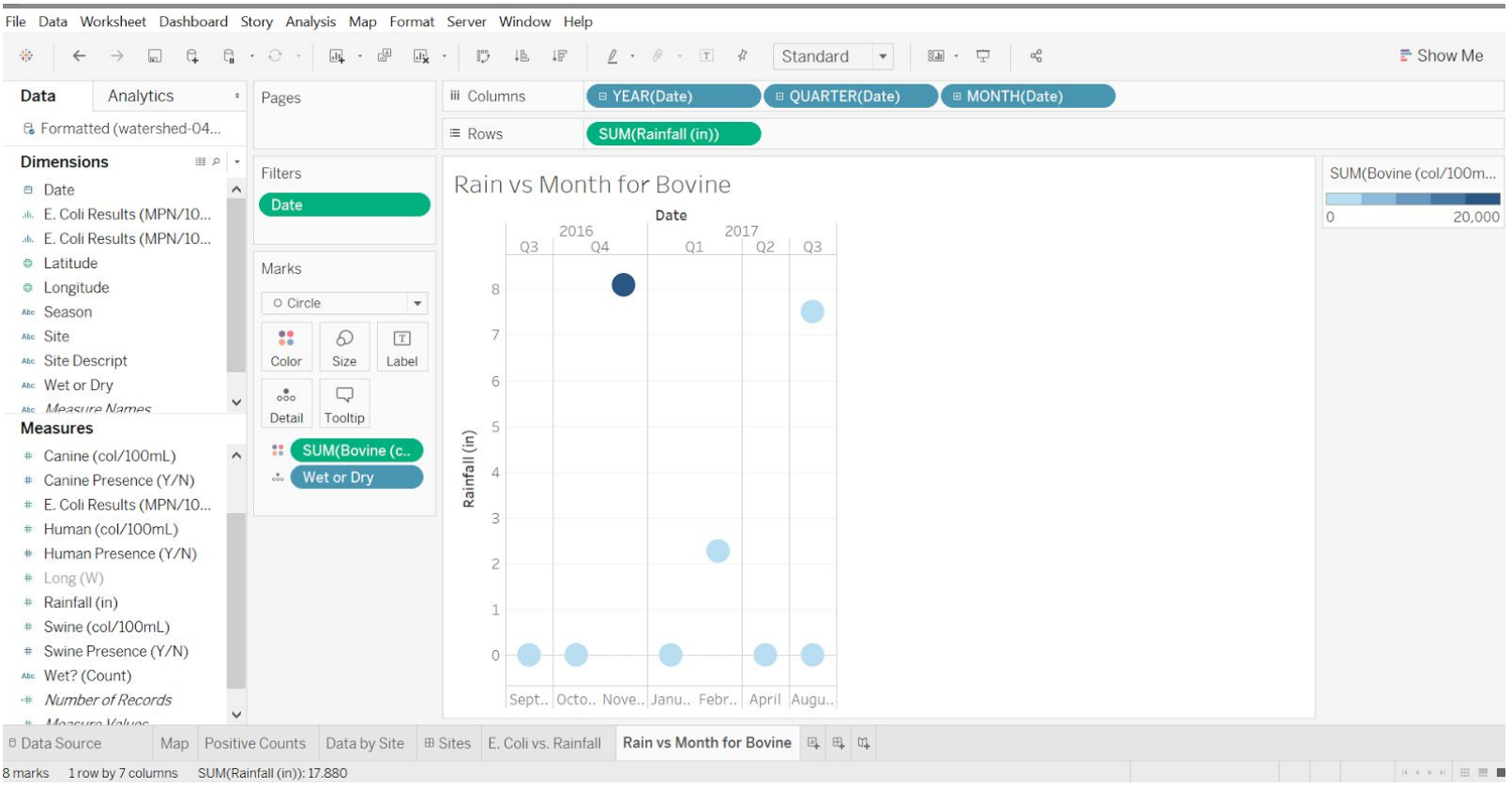

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Show Me
Data
Analytics
Formatted (watershed-04...
Dimensions
Date
E. Coli Results (MPN/100m...
E. Coli Results (MPN/100m...
Latitude
Longitude
Season
Site
Site Descript
Wet or Dry
Measure Names
Measures
Bovine (col/100mL)
Bovine Presence (Y/N)
Calculation1
Canine (col/100mL)
Canine Presence (Y/N)
E. Coli Results (MPN/10...
Human (col/100mL)
Human Presence (Y/N)
Long (W)
Rainfall (in)
Swine (col/100mL)
Pages
Columns
Longitude
Rows
Latitude
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Site
Map
© OpenStreetMap contributors
Site
CR
CRBO
CRT
CRU
GV
NR
Data Source
Map
6 marks
1 row by 1 column

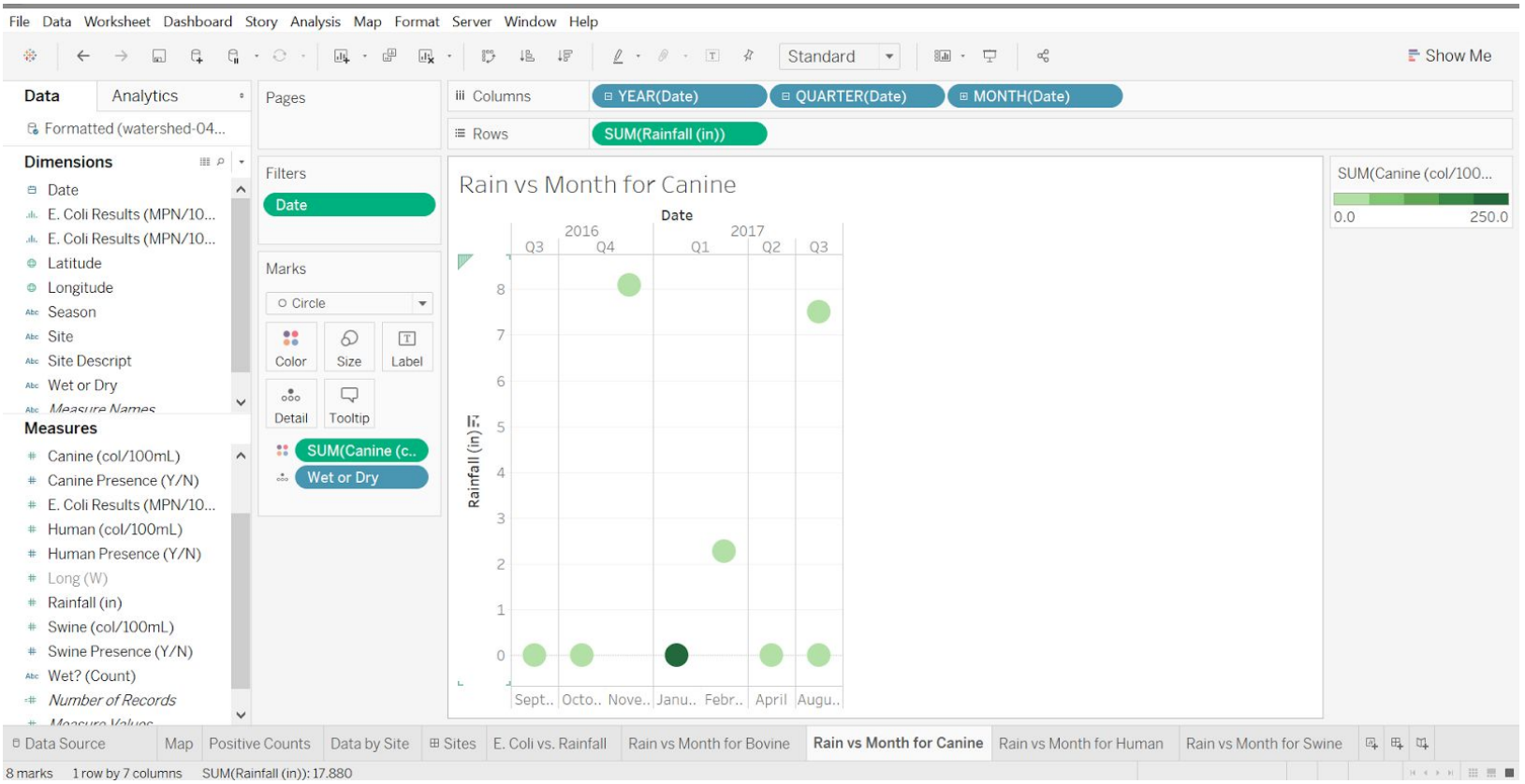

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Formatted (watershed-04...
Dimensions
Date
E. Coli Results (MPN/100m...
E. Coli Results (MPN/100m...
Latitude
Longitude
Season
Site
Site Descript
Wet or Dry
Measure Names
Measures
Bovine (col/100mL)
Bovine Presence (Y/N)
Calculation1
Canine (col/100mL)
Canine Presence (Y/N)
E. Coli Results (MPN/10...
Human (col/100mL)
Human Presence (Y/N)
Long (W)
Rainfall (in)
Swine (col/100mL)
Pages
Columns
SUM(Human Presenc..
SUM(Swine Presence..
SUM(Canine Presenc..
SUM(Bovine (col/10..
Rows
Filters
Marks
All
Automatic
Color
Size
Label
Detail
Tooltip
Site
SUM(Human ...
SUM(Swine P...
SUM(Canine ...
SUM(Bovine (...
Positive Counts
0 10 20 30
Human Presence (Y/N)
0 2 4 6 8 10 12
Swine Presence (Y/N)
0 1 2
Canine Presence (Y/N)
0K 5K 10K 15K 20K
Bovine (col/100mL)
Site
CR
CRBO
CRT
CRU
GV
NR
Data Source
Map
Positive Counts
24 marks
1 row by 4 columns
SUM(Bovine (col/100mL)): 20,024

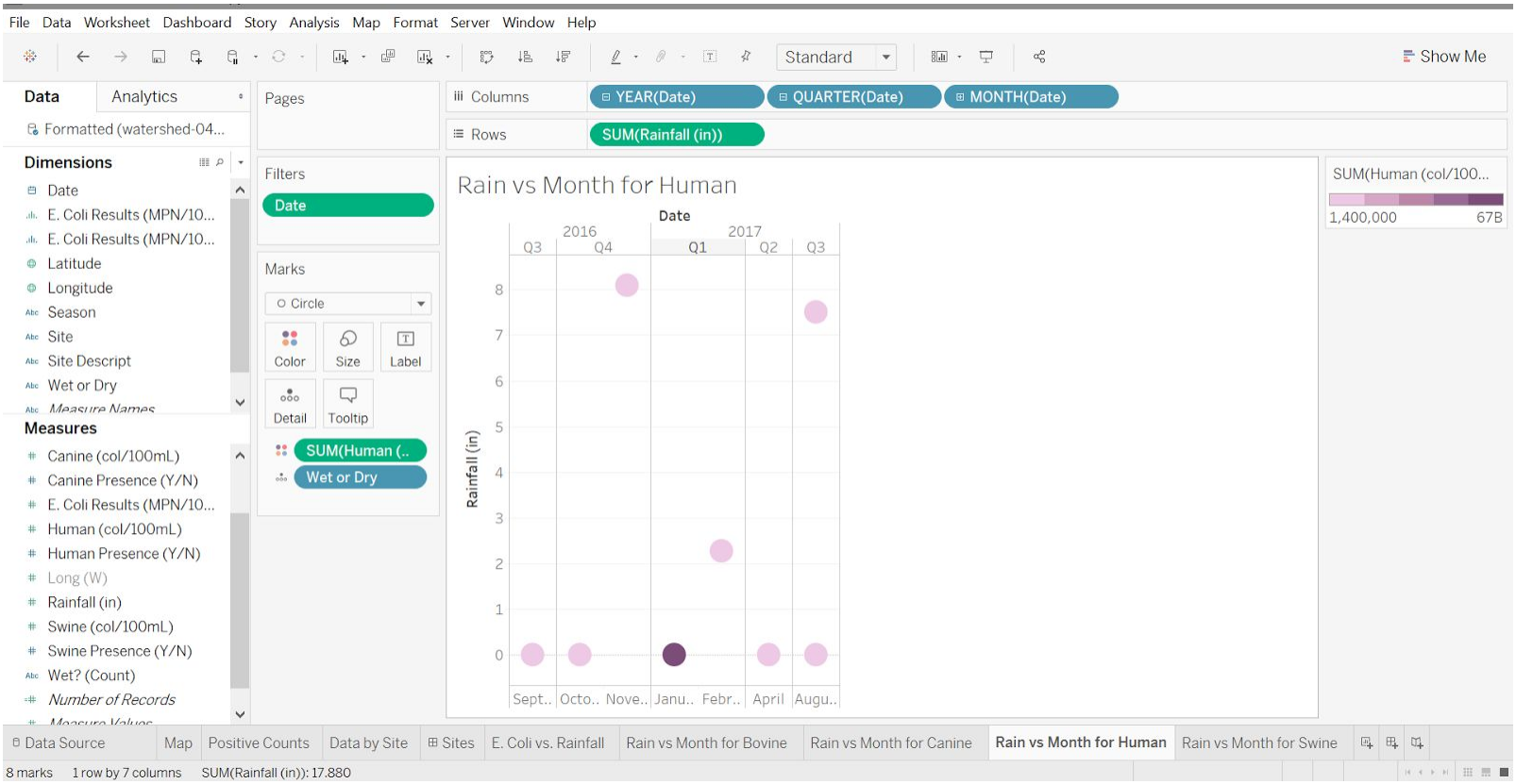

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Formatted (watershed-04...
Dimensions
Date
E. Coli Results (MPN/10...
E. Coli Results (MPN/10...
Latitude
Longitude
Season
Site
Site Descript
Wet or Dry
Measure Names
Measures
Bovine (col/100mL)
Bovine Presence (Y/N)
Calculation1
Canine (col/100mL)
Canine Presence (Y/N)
E. Coli Results (MPN/10...
Human (col/100mL)
Human Presence (Y/N)
Long (W)
Rainfall (in)
Swine (col/100mL)
Pages
Filters
Site
Marks
All
Shape
Color
Size
Label
Detail
Tooltip
Shape
Site
Site
AVG(E. Coli R...
AVG(Bovine (...
AVG(Canine (...
AVG(Human ...
AVG(Swine (c...
Columns
DAY(Date)
Rows
AVG(E. Coli Results (..
AVG(Bovine (col/100..
AVG(Canine (col/100..
AVG(Human (col/10..
AVG(Swine (col/100..
Data by Site
E. Coli Results (MPN/100mL)
Avg. Bovine (col/100mL)
Avg. Canine (col/100mL)
Avg. Human (col/100mL)
Avg. Swine (col/100mL)
Site
CR
CRBO
CRT
CRU
GV
NR
Site
CR
CRBO
CRT
CRU
GV
NR
Data Source
Map
Positive Counts
Data by Site
Sites
237 marks
5 rows by 1 column
SUM of AVG(Bovine (col/100mL)): 20,024

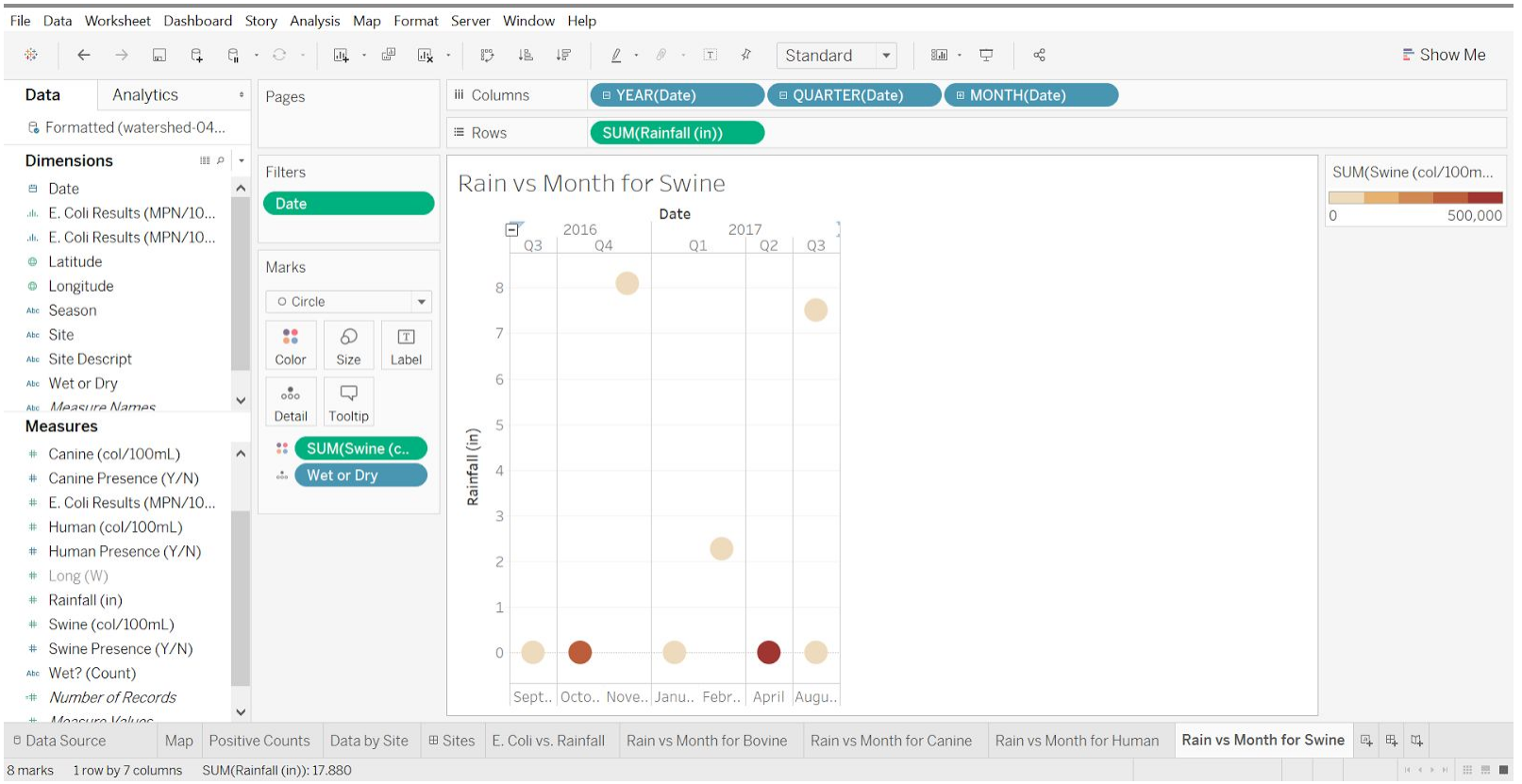

File Data Worksheet Dashboard Story Map Format Server Window Help
Show Me
Dashboard
Layout
Device Preview
Size
Desktop Browser (1000 x 800)
Sheets
Map
Positive Counts
Data by Site
Objects
Horizontal
Vertical
Text
Image
Web Page
Blank
Tiled
Floating
Show dashboard title
© OpenStreetMap contributors
Click to filter by site
(All)
CR
CRBO
CRT
CRU
GV
NR
Site
CR
CRBO
CRT
CRU
GV
NR
Human Presence (Y/N)
Swine Presence (Y/N)
Canine Presence (Y/N)
Bovine (col/100mL)
Data Source
Map
Positive Counts
Data by Site
Sites

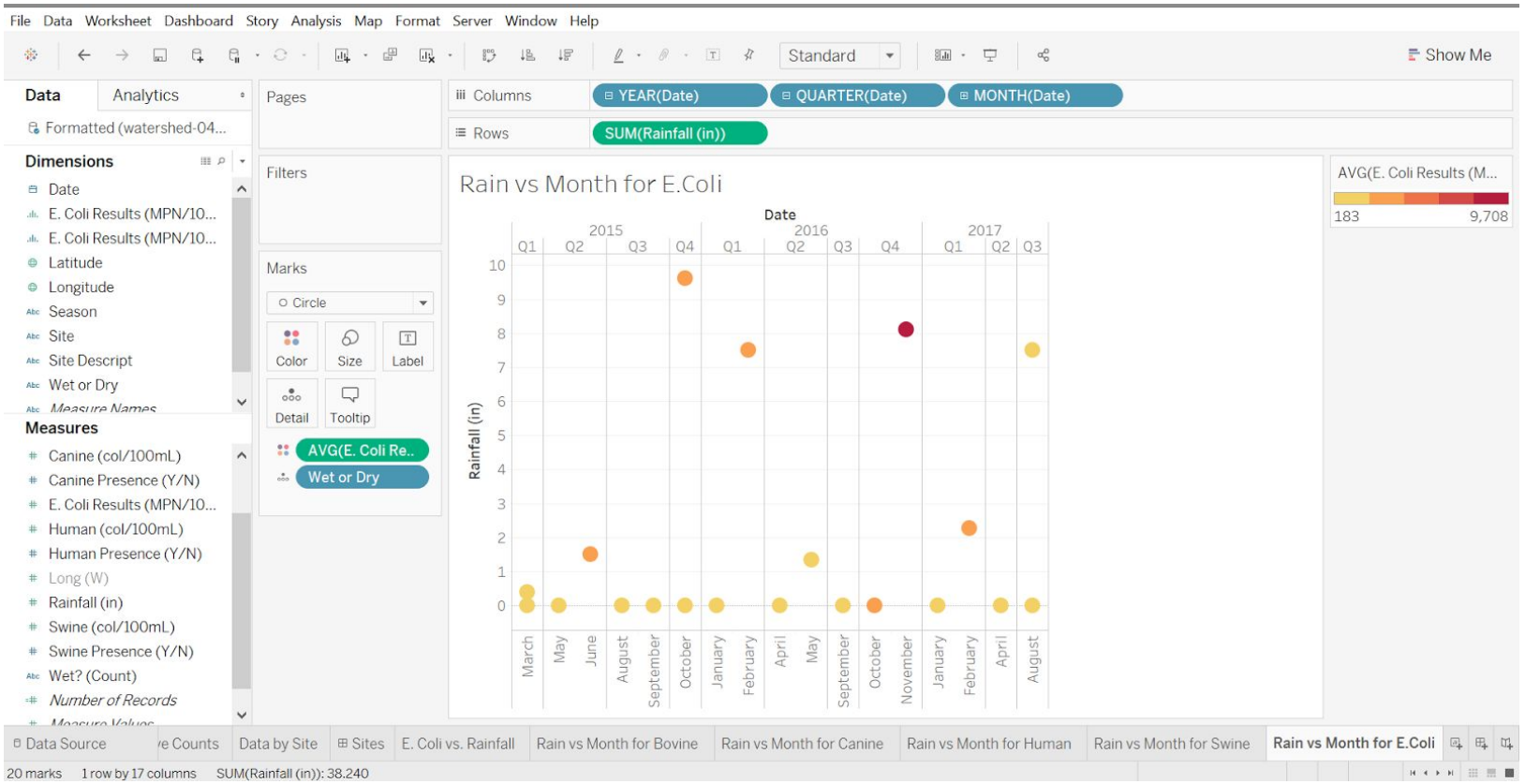

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Formatted (watershed-04...
Dimensions
E. Coli Results (MPN/10...
E. Coli Results (MPN/10...
Latitude
Longitude
Season
Site
Site Descript
Wet or Dry
Measure Names
Measures
Bovine (col/100mL)
Bovine Presence (Y/N)
Calculation1
Canine (col/100mL)
Canine Presence (Y/N)
E. Coli Results (MPN/10...
Human (col/100mL)
Human Presence (Y/N)
Long (W)
Rainfall (in)
Swine (col/100mL)
Pages
Filters
Marks
Circle
Color
Size
Label
Detail
Tooltip
Columns
Rainfall (in)
Rows
AVG(E. Coli Results (..
E. Coli vs. Rainfall
E. Coli Results (MPN/100mL)
10K
9K
8K
7K
6K
5K
4K
3K
2K
1K
0K
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4 1.5 1.6 1.7
Rainfall (in)
Data Source
Map
Positive Counts
Data by Site
Sites
E. Coli vs. Rainfall
11 marks
1 row by 1 column
SUM of AVG(E. Coli Results (MPN/100mL)): 28,170

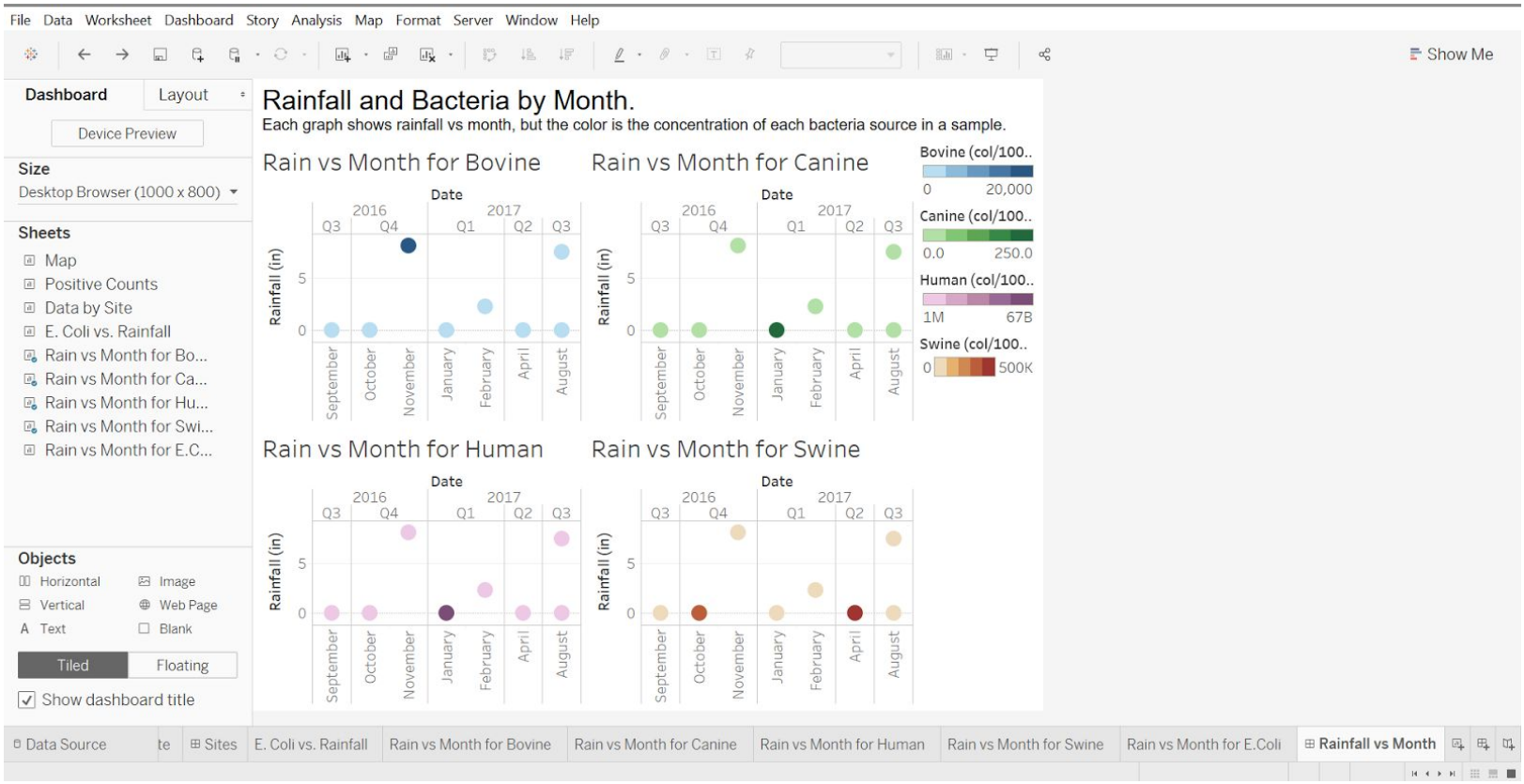

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Formatted (watershed-04...
Dimensions
Date
E. Coli Results (MPN/10...
E. Coli Results (MPN/10...
Latitude
Longitude
Season
Site
Site Descript
Wet or Dry
Measure Names
Measures
Canine (col/100mL)
Canine Presence (Y/N)
E. Coli Results (MPN/10...
Human (col/100mL)
Human Presence (Y/N)
Long (W)
Rainfall (in)
Swine (col/100mL)
Swine Presence (Y/N)
Wet? (Count)
Number of Records
Pages
Filters
Date
Marks
Circle
Color
Size
Label
Detail
Tooltip
SUM(Bovine (c..
Wet or Dry
Columns
YEAR(Date)
QUARTER(Date)
MONTH(Date)
Rows
SUM(Rainfall (in))
Rain vs Month for Bovine
Date
2016
2017
Q3
Q4
Q1
Q2
Q3
Rainfall (in)
8
7
6
5
4
3
2
1
0
Sept..
Octo..
Nove..
Janu..
Febr..
April
Augu..
SUM(Bovine (col/100m...
0
20,000
Data Source
Map
Positive Counts
Data by Site
Sites
E. Coli vs. Rainfall
Rain vs Month for Bovine
8 marks
1 row by 7 columns
SUM(Rainfall (in)): 17.880

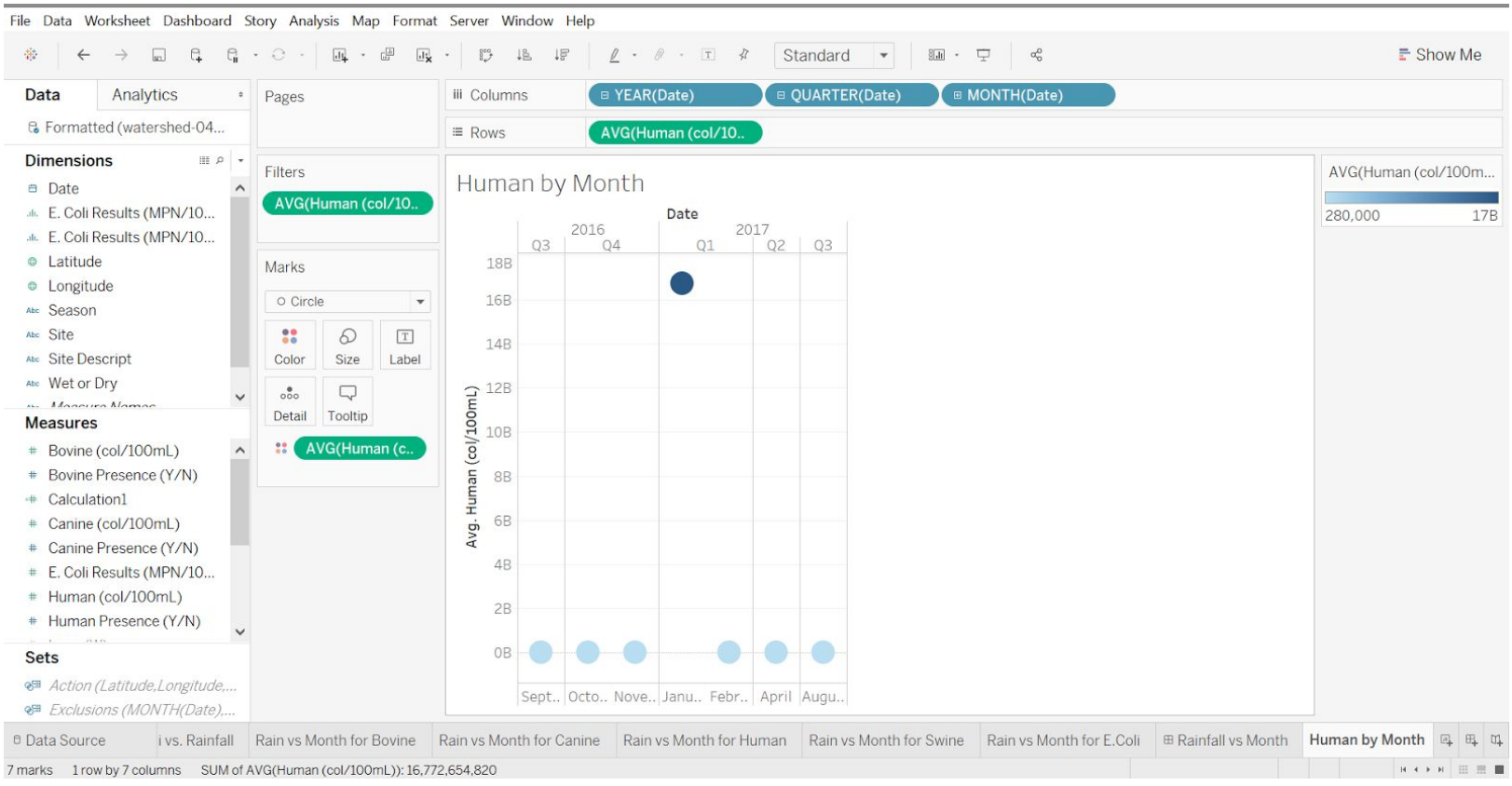

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Formatted (watershed-04...
Dimensions
Date
E. Coli Results (MPN/10...
E. Coli Results (MPN/10...
Latitude
Longitude
Season
Site
Site Descript
Wet or Dry
Measure Names
Measures
Canine (col/100mL)
Canine Presence (Y/N)
E. Coli Results (MPN/10...
Human (col/100mL)
Human Presence (Y/N)
Long (W)
Rainfall (in)
Swine (col/100mL)
Swine Presence (Y/N)
Wet? (Count)
Number of Records
Pages
Filters
Date
Marks
Circle
Color
Size
Label
Detail
Tooltip
SUM(Canine (c..
Wet or Dry
Columns
YEAR(Date)
QUARTER(Date)
MONTH(Date)
Rows
SUM(Rainfall (in))
Rain vs Month for Canine
Date
2016
2017
Q3
Q4
Q1
Q2
Q3
Rainfall (in)
0
1
2
3
4
5
6
7
8
Sept..
Octo..
Nove..
Janu..
Febr..
April
Augu..
SUM(Canine (col/100...
0.0
250.0
Data Source
Map
Positive Counts
Data by Site
Sites
E. Coli vs. Rainfall
Rain vs Month for Bovine
Rain vs Month for Canine
Rain vs Month for Human
Rain vs Month for Swine
8 marks
1 row by 7 columns
SUM(Rainfall (in)): 17.880

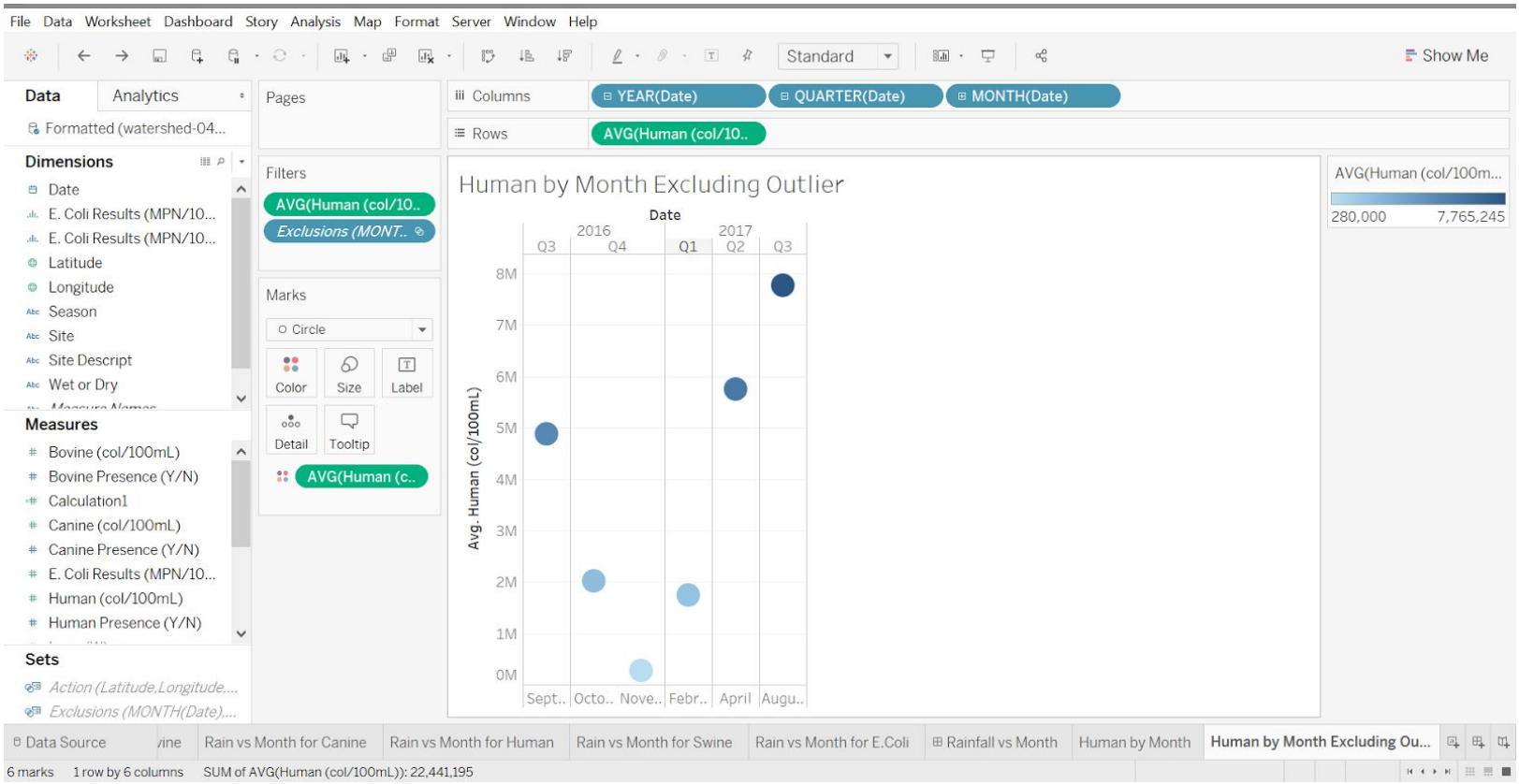

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Formatted (watershed-04...
Dimensions
Date
E. Coli Results (MPN/10...
E. Coli Results (MPN/10...
Latitude
Longitude
Season
Site
Site Descript
Wet or Dry
Measure Names
Measures
Canine (col/100mL)
Canine Presence (Y/N)
E. Coli Results (MPN/10...
Human (col/100mL)
Human Presence (Y/N)
Long (W)
Rainfall (in)
Swine (col/100mL)
Swine Presence (Y/N)
Wet? (Count)
Number of Records
Pages
Filters
Date
Marks
Circle
Color
Size
Label
Detail
Tooltip
SUM(Human (..
Wet or Dry
Columns
YEAR(Date)
QUARTER(Date)
MONTH(Date)
Rows
SUM(Rainfall (in))
Rain vs Month for Human
Date
2016
2017
Q3
Q4
Q1
Q2
Q3
Rainfall (in)
0
1
2
3
4
5
6
7
8
Sept..
Octo..
Nove..
Janu..
Febr..
April
Augu..
SUM(Human (col/100...
1,400,000
67B
Data Source
Map
Positive Counts
Data by Site
Sites
E. Coli vs. Rainfall
Rain vs Month for Bovine
Rain vs Month for Canine
Rain vs Month for Human
Rain vs Month for Swine
8 marks
1 row by 7 columns
SUM(Rainfall (in)): 17.880

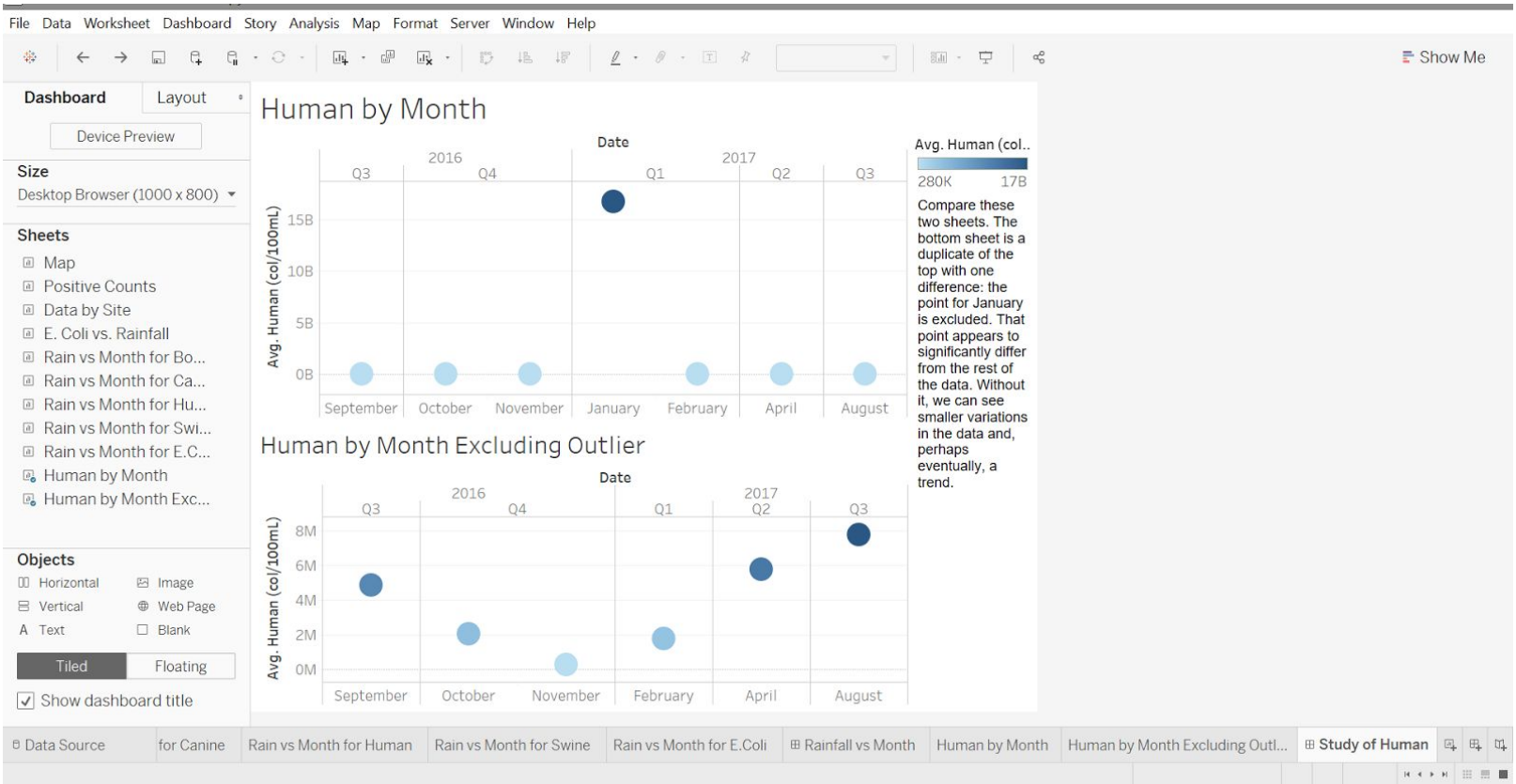

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Formatted (watershed-04...
Dimensions
Date
E. Coli Results (MPN/10...
E. Coli Results (MPN/10...
Latitude
Longitude
Season
Site
Site Descript
Wet or Dry
Measure Names
Measures
Canine (col/100mL)
Canine Presence (Y/N)
E. Coli Results (MPN/10...
Human (col/100mL)
Human Presence (Y/N)
Long (W)
Rainfall (in)
Swine (col/100mL)
Swine Presence (Y/N)
Wet? (Count)
Number of Records
Measure Values
Pages
Filters
Date
Marks
Circle
Color
Size
Label
Detail
Tooltip
SUM(Swine (c..
Wet or Dry
Columns
YEAR(Date)
QUARTER(Date)
MONTH(Date)
Rows
SUM(Rainfall (in))
Rain vs Month for Swine
Date
2016
2017
Q3
Q4
Q1
Q2
Q3
Rainfall (in)
0
1
2
3
4
5
6
7
8
Sept..
Octo..
Nove..
Janu..
Febr..
April
Augu..
SUM(Swine (col/100m...
0
500,000
Data Source
Map
Positive Counts
Data by Site
Sites
E. Coli vs. Rainfall
Rain vs Month for Bovine
Rain vs Month for Canine
Rain vs Month for Human
Rain vs Month for Swine
8 marks
1 row by 7 columns
SUM(Rainfall (in)): 17.880

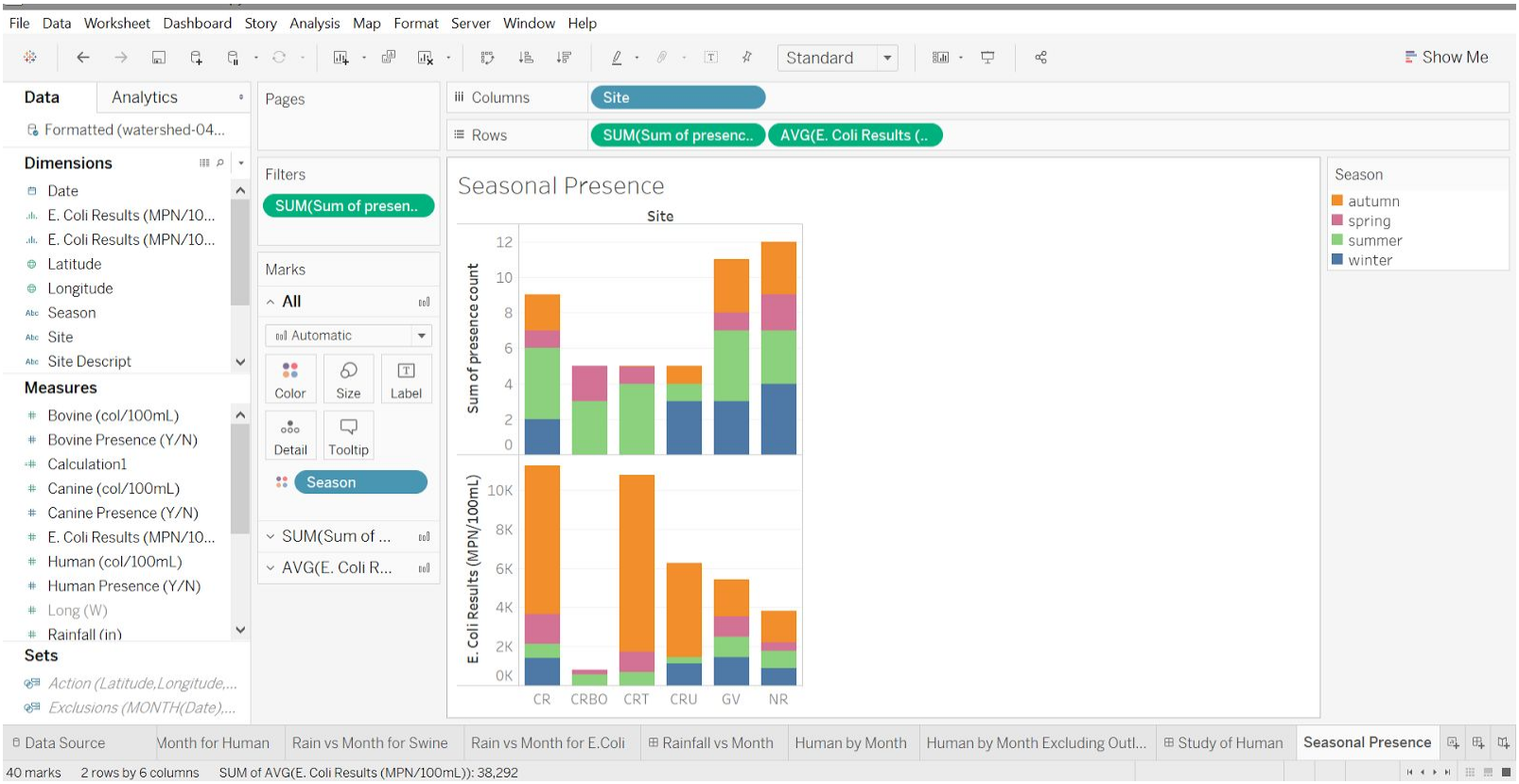

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Formatted (watershed-04...
Dimensions
Date
E. Coli Results (MPN/10...
E. Coli Results (MPN/10...
Latitude
Longitude
Season
Site
Site Descript
Wet or Dry
Measure Names
Measures
Canine (col/100mL)
Canine Presence (Y/N)
E. Coli Results (MPN/10...
Human (col/100mL)
Human Presence (Y/N)
Long (W)
Rainfall (in)
Swine (col/100mL)
Swine Presence (Y/N)
Wet? (Count)
Number of Records
Measure Values
Pages
Filters
Marks
Circle
Color
Size
Label
Detail
Tooltip
AVG(E. Coli Re..
Wet or Dry
Columns
YEAR(Date)
QUARTER(Date)
MONTH(Date)
Rows
SUM(Rainfall (in))
Rain vs Month for E.Coli
Date
2015
2016
2017
Q1
Q2
Q3
Q4
Rainfall (in)
0
1
2
3
4
5
6
7
8
9
10
March
May
June
August
September
October
January
February
April
May
September
October
November
January
February
April
August
AVG(E. Coli Results (M...
183
9,708
Data Source
Data by Site
Sites
E. Coli vs. Rainfall
Rain vs Month for Bovine
Rain vs Month for Canine
Rain vs Month for Human
Rain vs Month for Swine
Rain vs Month for E.Coli
20 marks
1 row by 17 columns
SUM(Rainfall (in)): 38.240


File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Show Me
Dashboard
Layout
Device Preview
Size
Desktop Browser (1000 x 800)
Sheets
Map
Positive Counts
Data by Site
E. Coli vs. Rainfall
Rain vs Month for Bo...
Rain vs Month for Ca...
Rain vs Month for Hu...
Rain vs Month for Swi...
Rain vs Month for E.C...
Objects
Horizontal
Image
Vertical
Web Page
Text
Blank
Tiled
Floating
Show dashboard title
Rainfall and Bacteria by Month.
Each graph shows rainfall vs month, but the color is the concentration of each bacteria source in a sample.
Rain vs Month for Bovine
Rain vs Month for Canine
Rain vs Month for Human
Rain vs Month for Swine
Date
2016
2017
Q3
Q4
Q1
Q2
Q3
Rainfall (in)
0
5
September
October
November
January
February
April
August
Bovine (col/100..
0
20,000
Canine (col/100..
0.0
250.0
Human (col/100..
1M
67B
Swine (col/100..
0
500K
Data Source
Sites
E. Coli vs. Rainfall
Rain vs Month for Bovine
Rain vs Month for Canine
Rain vs Month for Human
Rain vs Month for Swine
Rain vs Month for E.Coli
Rainfall vs Month


File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Formatted (watershed-04...
Dimensions
Date
E. Coli Results (MPN/10...
E. Coli Results (MPN/10...
Latitude
Longitude
Season
Site
Site Descript
Wet or Dry
Measures
Bovine (col/100mL)
Bovine Presence (Y/N)
Calculation1
Canine (col/100mL)
Canine Presence (Y/N)
E. Coli Results (MPN/10...
Human (col/100mL)
Human Presence (Y/N)
Sets
Action (Latitude,Longitude,...
Exclusions (MONTH(Date),...
Pages
Filters
AVG(Human (col/10..
Marks
Circle
Color
Size
Label
Detail
Tooltip
AVG(Human (c..
Columns
YEAR(Date)
QUARTER(Date)
MONTH(Date)
Rows
AVG(Human (col/10..
Human by Month
Date
2016
2017
Q3
Q4
Q1
Q2
Q3
Avg. Human (col/100mL)
18B
16B
14B
12B
10B
8B
6B
4B
2B
0B
Sept..
Octo..
Nove..
Janu..
Febr..
April
Augu..
AVG(Human (col/100m...
280,000
17B
Data Source
vs. Rainfall
Rain vs Month for Bovine
Rain vs Month for Canine
Rain vs Month for Human
Rain vs Month for Swine
Rain vs Month for E.Coli
Rainfall vs Month
Human by Month
7 marks
1 row by 7 columns
SUM of AVG(Human (col/100mL)): 16,772,654,820


File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Formatted (watershed-04...
Dimensions
Date
E. Coli Results (MPN/10...
E. Coli Results (MPN/10...
Latitude
Longitude
Season
Site
Site Descript
Wet or Dry
Measures
Bovine (col/100mL)
Bovine Presence (Y/N)
Calculation1
Canine (col/100mL)
Canine Presence (Y/N)
E. Coli Results (MPN/10...
Human (col/100mL)
Human Presence (Y/N)
Sets
Action (Latitude,Longitude,...
Exclusions (MONTH(Date),...
Pages
Filters
AVG(Human (col/10..
Exclusions (MONT..
Marks
Circle
Color
Size
Label
Detail
Tooltip
AVG(Human (c..
Columns
YEAR(Date)
QUARTER(Date)
MONTH(Date)
Rows
AVG(Human (col/10..
Human by Month Excluding Outlier
Date
2016
2017
Q3
Q4
Q1
Q2
Q3
Avg. Human (col/100mL)
8M
7M
6M
5M
4M
3M
2M
1M
0M
Sept..
Octo..
Nove..
Febr..
April
Augu..
AVG(Human (col/100m...
280,000
7,765,245
Data Source
Rain vs Month for Canine
Rain vs Month for Human
Rain vs Month for Swine
Rain vs Month for E.Coli
Rainfall vs Month
Human by Month
Human by Month Excluding Ou...
6 marks
1 row by 6 columns
SUM of AVG(Human (col/100mL)): 22,441,195


File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Show Me
Dashboard
Layout
Device Preview
Size
Desktop Browser (1000 x 800)
Sheets
Map
Positive Counts
Data by Site
E. Coli vs. Rainfall
Rain vs Month for Bo...
Rain vs Month for Ca...
Rain vs Month for Hu...
Rain vs Month for Swi...
Rain vs Month for E.C...
Human by Month
Human by Month Exc...
Objects
Horizontal
Image
Vertical
Web Page
Text
Blank
Tiled
Floating
Show dashboard title
Human by Month
Date
2016
2017
Q3
Q4
Q1
Q2
Q3
Avg. Human (col/100mL)
15B
10B
5B
0B
September
October
November
January
February
April
August
Avg. Human (col..
280K
17B
Compare these two sheets. The bottom sheet is a duplicate of the top with one difference: the point for January is excluded. That point appears to significantly differ from the rest of the data. Without it, we can see smaller variations in the data and, perhaps eventually, a trend.
Human by Month Excluding Outlier
Date
2016
2017
Q3
Q4
Q1
Q2
Q3
Avg. Human (col/100mL)
8M
6M
4M
2M
0M
September
October
November
February
April
August
Data Source
for Canine
Rain vs Month for Human
Rain vs Month for Swine
Rain vs Month for E.Coli
Rainfall vs Month
Human by Month
Human by Month Excluding Outl...
Study of Human


File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Formatted (watershed-04...
Dimensions
Date
E. Coli Results (MPN/10...
E. Coli Results (MPN/10...
Latitude
Longitude
Season
Site
Site Descript
Measures
Bovine (col/100mL)
Bovine Presence (Y/N)
Calculation1
Canine (col/100mL)
Canine Presence (Y/N)
E. Coli Results (MPN/10...
Human (col/100mL)
Human Presence (Y/N)
Long (W)
Rainfall (in)
Sets
Action (Latitude,Longitude,...
Exclusions (MONTH(Date),...
Pages
Filters
SUM(Sum of presen..
Marks
All
Automatic
Color
Size
Label
Detail
Tooltip
Season
SUM(Sum of ...
AVG(E. Coli R...
Columns
Site
Rows
SUM(Sum of presenc..
AVG(E. Coli Results (..
Seasonal Presence
Site
Sum of presence count
12
10
8
6
4
2
0
E. Coli Results (MPN/100mL)
10K
8K
6K
4K
2K
0K
CR
CRBO
CRT
CRU
GV
NR
Season
autumn
spring
summer
winter
Data Source
Month for Human
Rain vs Month for Swine
Rain vs Month for E.Coli
Rainfall vs Month
Human by Month
Human by Month Excluding Outl...
Study of Human
Seasonal Presence
40 marks
2 rows by 6 columns
SUM of AVG(E. Coli Results (MPN/100mL)): 38,292

8. Intermediate Vizzies

This part is about more advanced visualizations and Tableau tools.
The sections are as follows:

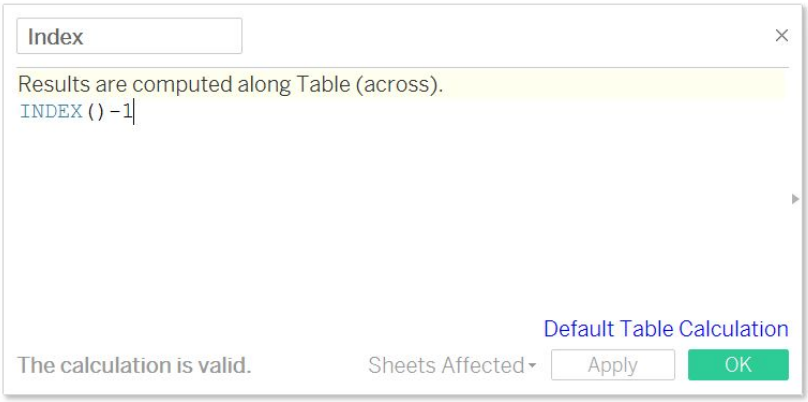
1. Radial bar chart
2. Types of joins
3. Hex map
4. Inserting images

1. Radial bar chart
    a. The information for this came from here and was not originally mine. However, I altered it slightly. This means that if you try to follow along and compare both this manual and that website, you might get confused.
    b. The data I will use comes from the applicant data for the Research Experience for Undergraduates in Collaborative Data Visualization Applications. The image below shows the data. Copy it into Excel. When doing so, I recommend using formulas to create the Decimal and Percent categories. There were 228 total applicants.

	A	B	C	D	E
1	Ethnicity	Path	Number	Decimal	Percent
2	White	1	113	0.495614	49.5614
3	African American	1	50	0.219298	21.92982
4	Asian	1	41	0.179825	17.98246
5	Hispanic/Latino	1	13	0.057018	5.701754
6	Other	1	9	0.039474	3.947368
7	Prefer not to specify	1	2	0.008772	0.877193
8	White	270	113	0.495614	49.5614
9	African American	270	50	0.219298	21.92982
10	Asian	270	41	0.179825	17.98246
11	Hispanic/Latino	270	13	0.057018	5.701754
12	Other	270	9	0.039474	3.947368
13	Prefer not to specify	270	2	0.008772	0.877193

    c. There are some specific things to note about the data. First, there are 6 ethnicities which are repeated so that they carry two path values, 1 and 270. When creating your own dataset for a radial bar chart, be sure to do this step.
    d. Right click on Path. Create > Bins… and keep the field name Path (bin). Make sure size of bins is one before clicking okay.
    e. The next few steps involve creating several calculated fields. For each one, make sure you carefully type in the formula so that your calculation is valid.

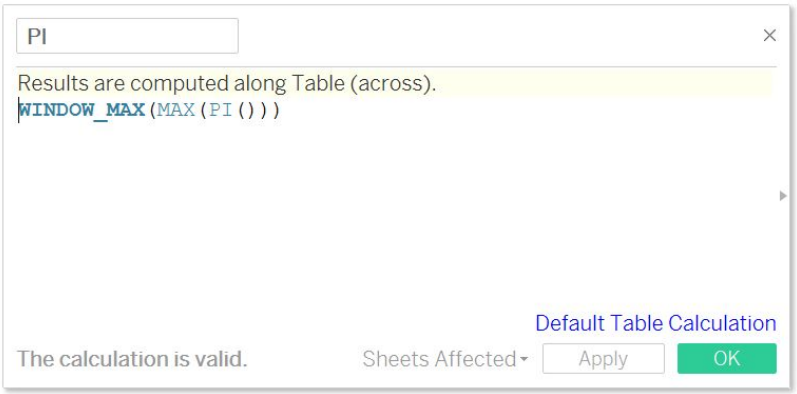
f.


Index
Results are computed along Table (across).
INDEX()-1
Default Table Calculation
The calculation is valid.
Sheets Affected
Apply
OK

g.

Max Value
Results are computed along Table (across).
WINDOW_MAX(SUM([Percent]))
Default Table Calculation
The calculation is valid.
Sheets Affected
Apply
OK

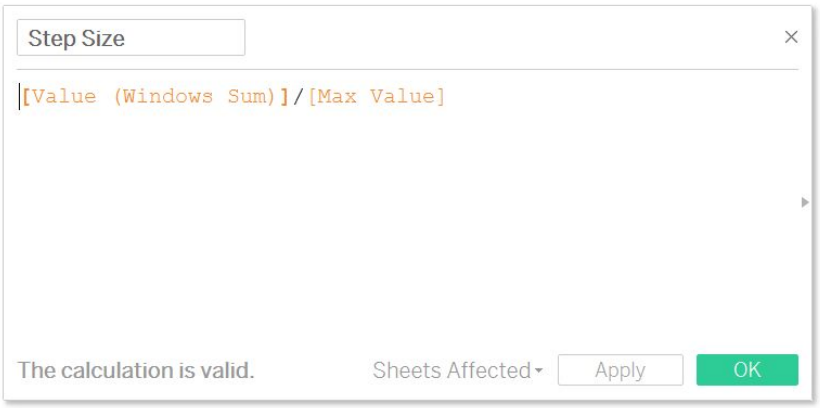
h.


PI
Results are computed along Table (across).
WINDOW_MAX(MAX(PI()))
Default Table Calculation
The calculation is valid.
Sheets Affected
Apply
OK

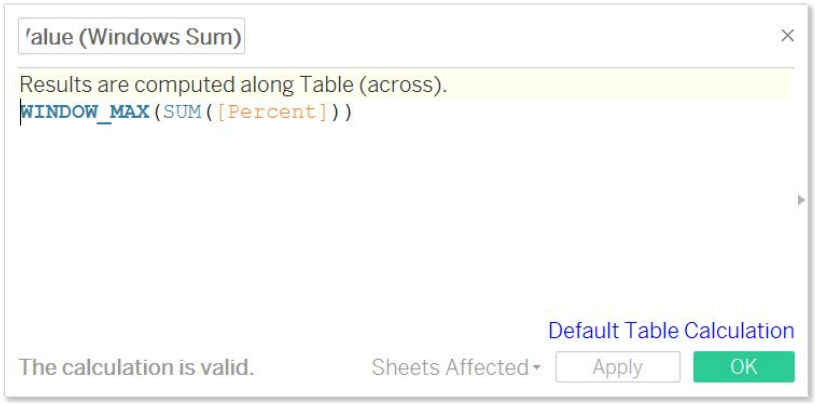
i.

Rank
Results are computed along Table (across).
RANK_UNIQUE([Value (Windows Sum)], 'asc')
Default Table Calculation
The calculation is valid.
Sheets Affected
Apply
OK

j.



k.

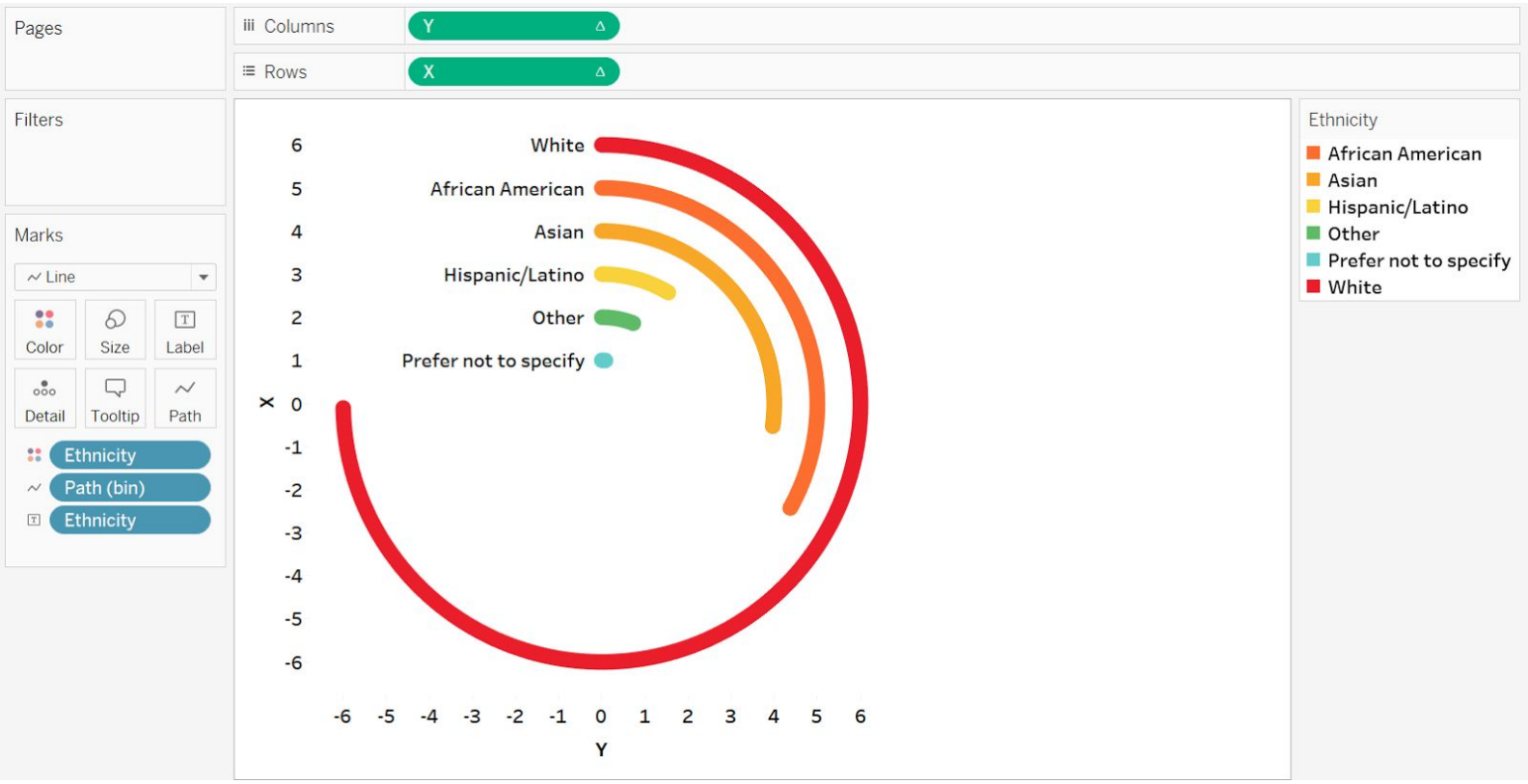


i. Note that the name for this one is “Value (Windows Sum).”

l.

m.

n. Drag Ethnicity to Color. Drag Value (Windows Sum) to Label. Drag Ethnicity to what used to be Label and now shows up as Text. Change the Marks type from automatic to line. Drag the dimension Path (bin) into the Marks Path. Drag Y to Columns and X to Rows.
o. Open the drop-down menu for the Y pill in Columns. Compute using > Path (bin). Do the same for X.
p. Click on Y again and go to Edit Table Calculation…
    i. In Nested Calculation, choose Max Value. Make sure both Path (bin) and Ethnicity are selected under Compute Using. Path (bin) should be above Ethnicity and “At the level” should be set to Deepest.
    ii. In Nested Calculation, choose Rank. Make sure both Path (bin) and Ethnicity are selected but this time Ethnicity should be above Path (bin), so you have to move it there. At the level should be Ethnicity.
q. Click on X and Edit Table Calculations…
    i. In Nested Calculation, choose Max Value. Make sure both Path (bin) and Ethnicity are selected under Compute Using. Path (bin) should be above Ethnicity and “At the level” should be set to Deepest.
    ii. In Nested Calculation, choose Rank. Make sure both Path (bin) and Ethnicity are selected but this time Ethnicity should be above Path (bin), so you have to move it there. At the level should be Ethnicity.
r. Increase size in the Marks card.
s. To make it more visually appealing, I hid the headers and edited the tooltip to only show Ethnicity.

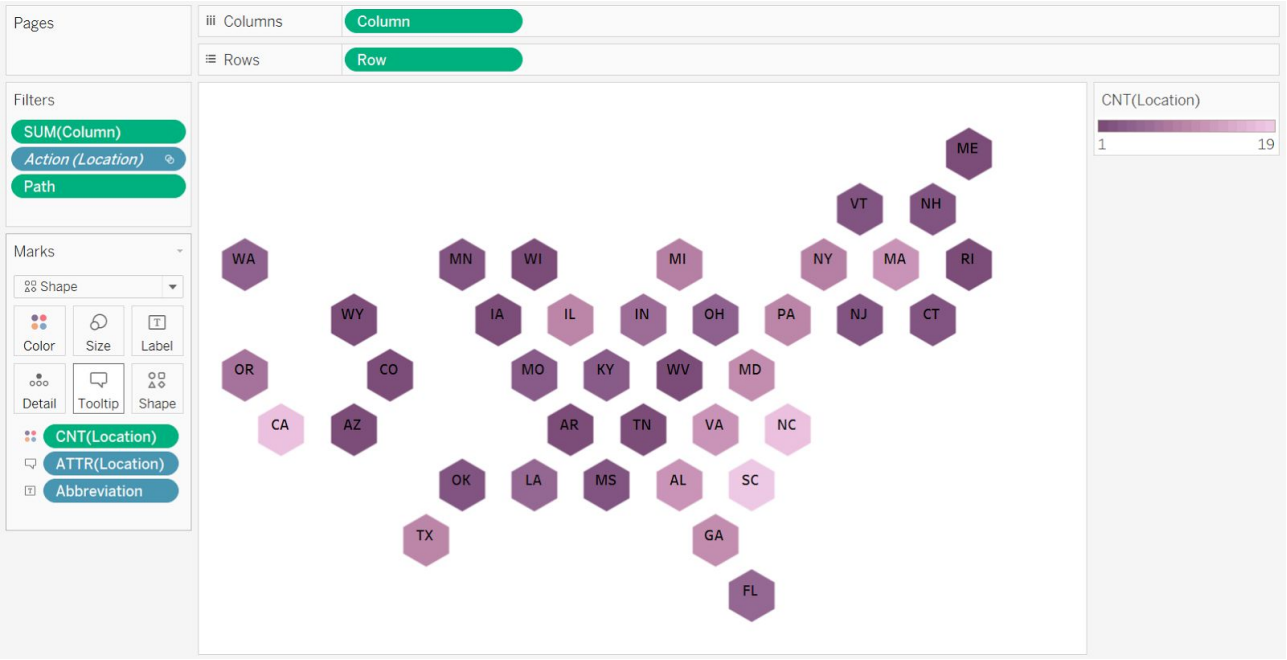


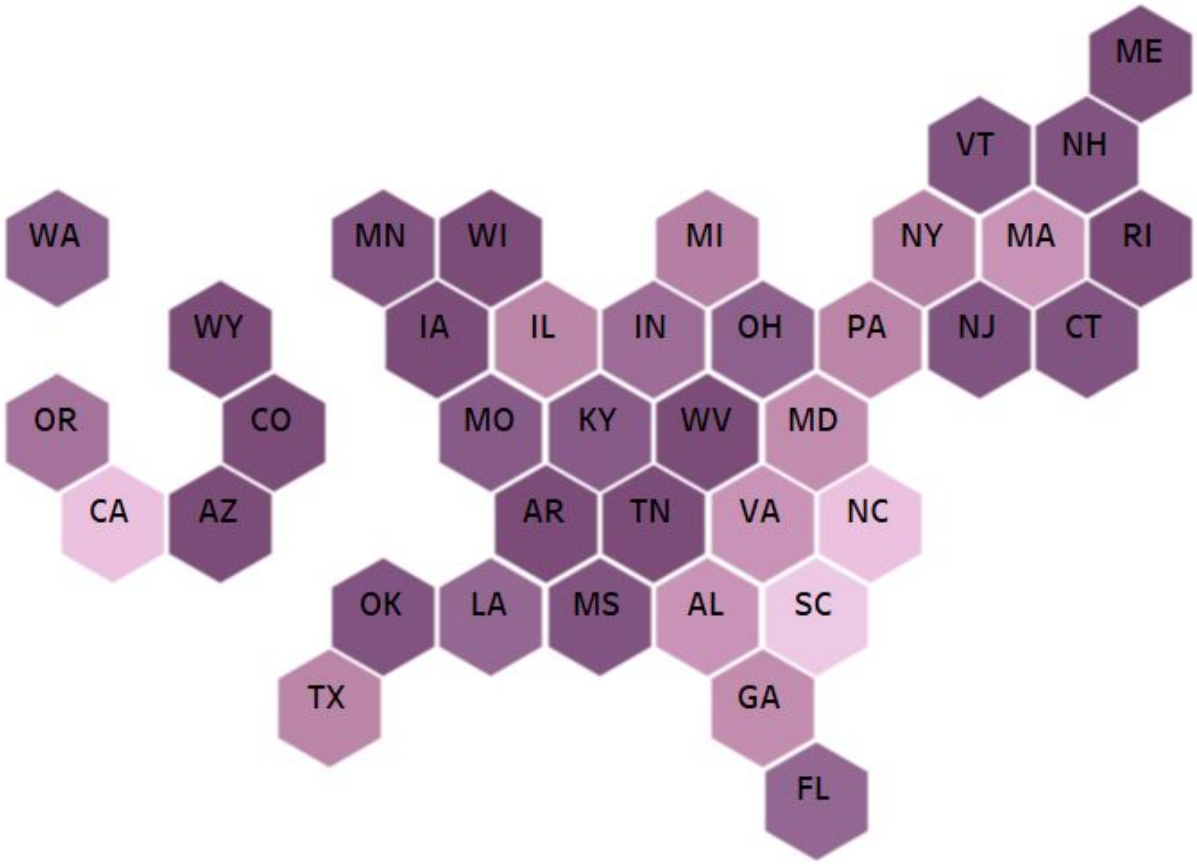
2. Joins
    a. If you had another sheet that you also wanted to include, you can click and drag it into your file as well. When you do, it will “join” itself to your previous data with a symbol looking like a Venn diagram. If you click on the Venn diagram in the middle, it asks you about types of joins you want to do.
    b. Full outer join means that all of the data will be lumped together. This might make null values if the two lists do not match up perfectly.
    c. Left join means to take the data from the right that matches something from the left and add it in, keeping everything on the left and only some of the stuff from the right. The reverse is true for a right join.
    d. An inner join keeps only the stuff that they both have in common. Under the gray data source and letter, select what should be the same for both to equate the fields.

3. Hex map
    a. The information originally came from here.
    b. The first step is to find or create a hexagonal shape and save it to your shapes repository. Save it to your computer in C:\Users\<username>\Documents\My

Tableau Repository/Shapes. I made a folder named “Other shapes” and saved it there.

c. Next, go here and copy the spreadsheet into a form you can use in Tableau. I recommend copying everything into Excel. Again, this was not originally mine; it was created by Matt Chambers who deserves credit for his work.
d. Drag the Row Measures pill into Rows. The way this spreadsheet was created required you to reverse the scale on the Rows axis. Right click on the x-axis and select Edit Axis. In the Scale box, select Reversed.
e. Add the Column Measure to Columns.
f. The marks type should be Shapes and you need to change the Shape type to the hexagonal shape you added earlier.
g. I recommend hiding the axis titles and scales. Matt Chambers also recommended adding Abbreviation to label. If you do that, change the vertical and horizontal alignment to center. You will also want to allow overlap with other marks or some of the labels will not appear.
h. Now you can add insights from your own data. Try dragging a pill to Color.
i. A major part of this worksheet involves making sure the spacing is correct. When you add it to a dashboard, the white space will change. Increase the Size in Marks and the size of the sheet on the dashboard to get it to look correct.
j. The first of the two images below shows the worksheet; the second shows what the same worksheet looks like on a dashboard. Please note that the filters were specific to my dataset and will not necessarily be required for your worksheet.





4. Inserting images
    a. Inserting images to a dashboard is not intuitive.
    b. First, save the image in the right place.
        i. Copy the image file to a folder in C:\Users\<username>\Documents\My Tableau Repository/Shapes.
        ii. You will need to create a folder titled “Images,” or something similar.
    c. Go back into Tableau and create a calculated field with the formula: "image.”
        i. Right click on any Dimension, Create, Calculated Field.
        ii. Delete any text within the body of the formula and type “image” – including quotation marks. Change the title of the calculated field to Image. Click OK.

d. Drag that calculated field to the Rows shelf. Usually you would do this on a blank Worksheet. You do this so that you can drag that Worksheet into your Dashboard to enhance it.
e. In the Marks card drop down menu (which usually starts at Automatic), select Shape. Click on the Shape button in the Marks card and click on More Shapes. Select the Shape Palette and find the Image folder (this is the folder you named in step 4bii above, so if you used a different name, choose that), then select your image as a shape.
f. It will start out really, really small. Change its size by using the Size button on the Marks card. Put it to about halfway and then grab the image's corners and expand it until it is the size you want.