DSC 96

What Happened?

Answers on Day 1

Giorgio Quer (adapted from Colin Jemmott)

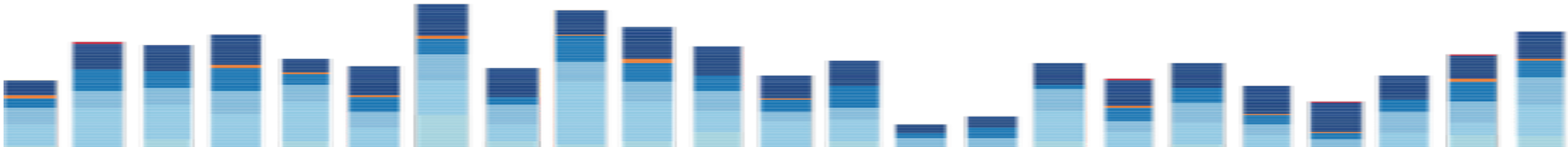
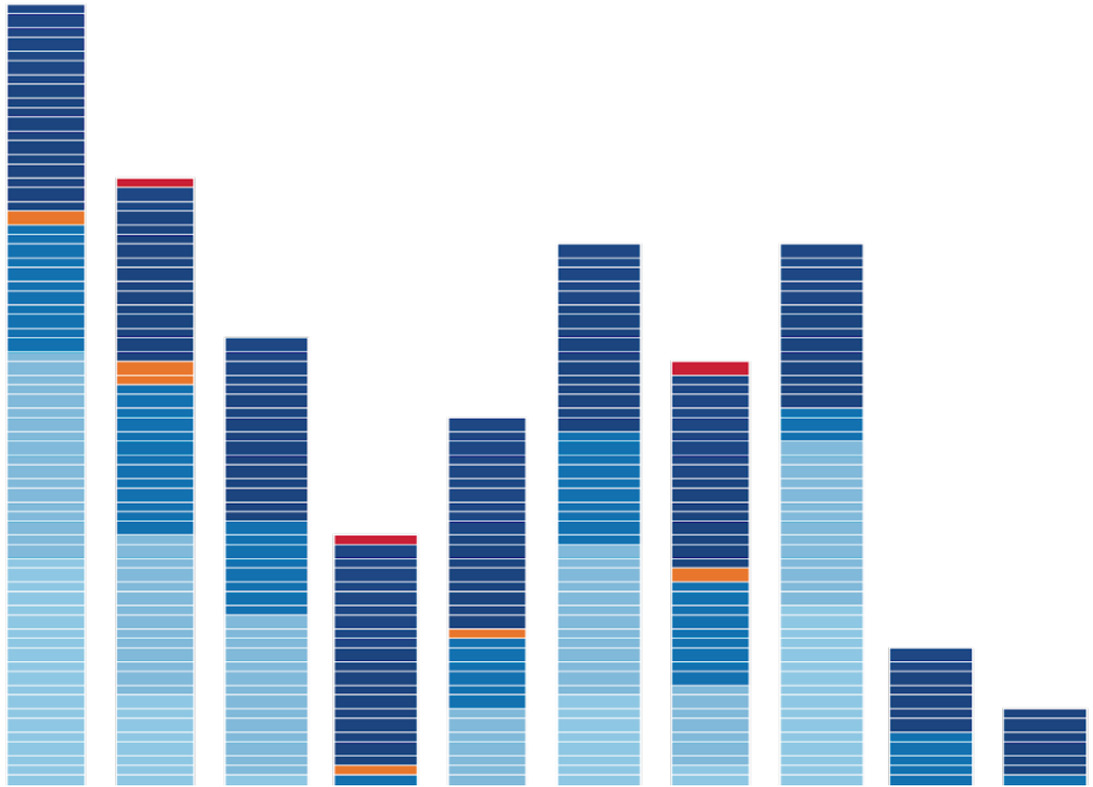


Tableau journalism basics

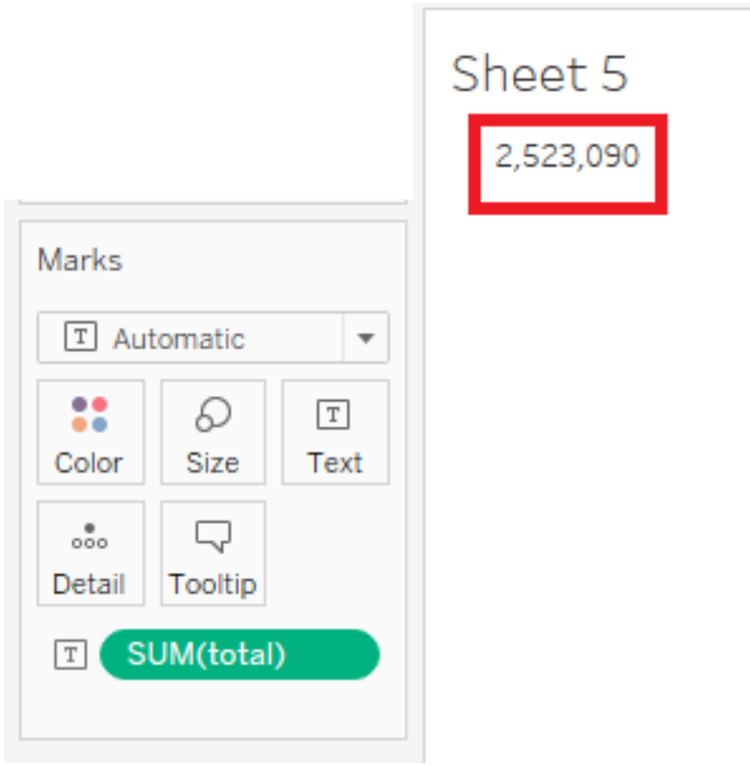


- Review of importing and data analysis
- Aggregation
- Filters
- Do we have enough to understand the story?

Basic filters and analyzing the data for story

- Drag the total-lobbying calculation from last time onto **row** and select the **tabular feature** on the **Show Me** pane to get the number: $2,523,090.

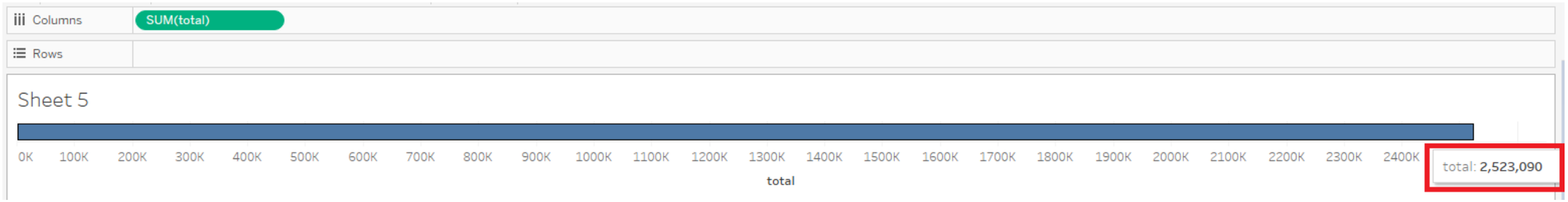
That’s the total spent by local governments for lobbyists.



Basic filters and analyzing the data for story

With Tableau, as in many software programs, there are multiple ways to do the same thing.

- Select the bar chart option on the **Show Me** feature and you'll see a bar visually illustrating the total in lobbying from all local governments:



Basic filters and analyzing the data for story

Here’s a reminder:

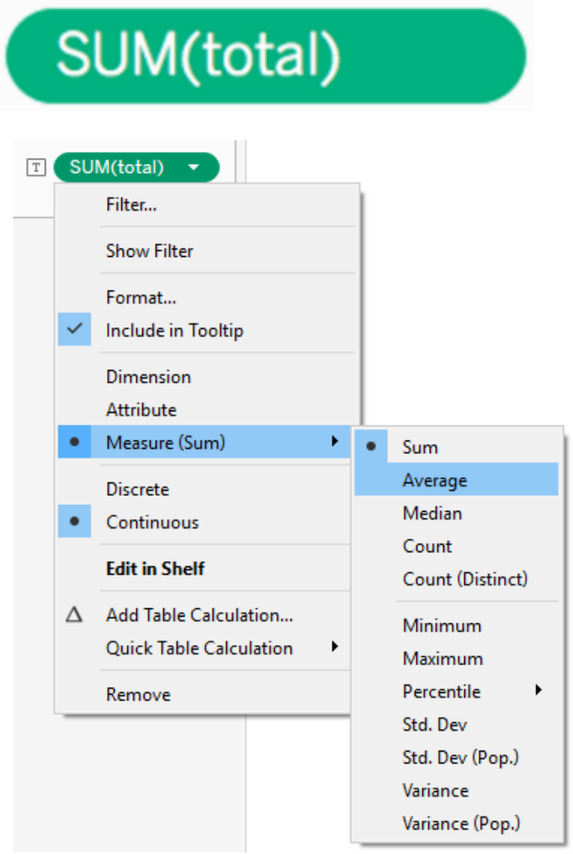
- On the menu, select **Analysis** and then **Calculated Field**.
- Give your new calculated field a name: total. Then type the calculation you want to do.
- Add **Compensation** and **Expenses**.

Now, last time you created a bar chart by the Entity and also Entity Type. What if you want just one number – the total for all lobbying money?

Basic filters and analyzing the data for story

One thing you may have noticed is that Tableau automatically aggregates your data by the sum. If you look at the green pill on the marks shelf, you'll see the word SUM in front of the total field: SUM(total)

While Tableau will automatically sum your data, that is not the only aggregation it can do. How could we see average? How about the median? It's as easy as selecting the caret, or down arrow or triangle on the right side of the pill.



Basic filters and analyzing the data for story

What if we want to get a sense of the median compared to the average? And why would that matter anyway?

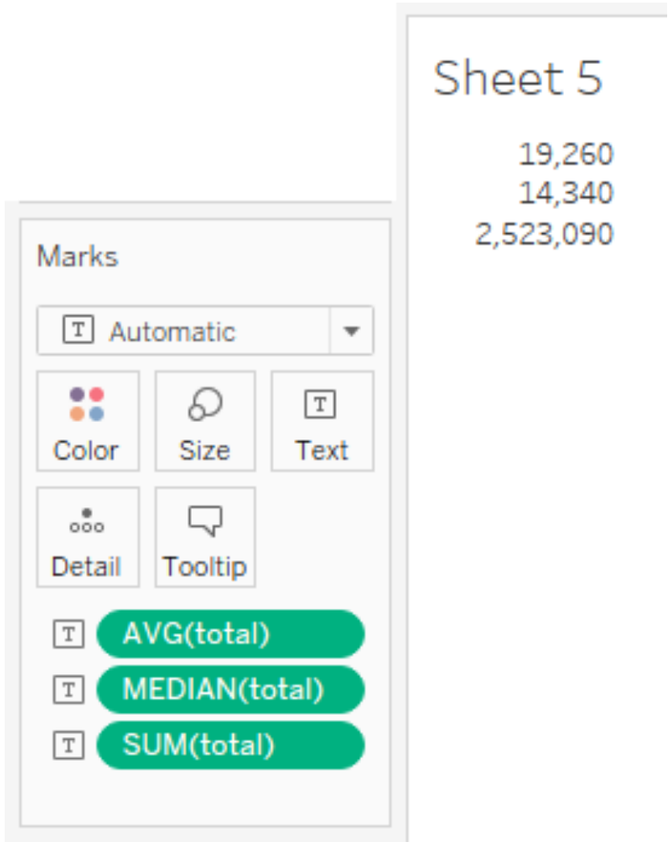
To give you a sense of how you might use this difference between the average and the median, let’s examine the type of governmental body – that’s the entity type field – that is paying lobbyists and look at the gap between the median and the average.

Basic filters and analyzing the data for story

In addition to looking at the median and average, we can also look at the total.

- Drag total onto the marks shelf three different times. Select three different aggregation levels: **SUM(total)**, **AVG(total)** and **MEDIAN(total)**.

Make sure that the mark just to the left of each pill is selected to be **Text (T)**. Then the actual numbers will show up on the workbook.



Basic filters and analyzing the data for story

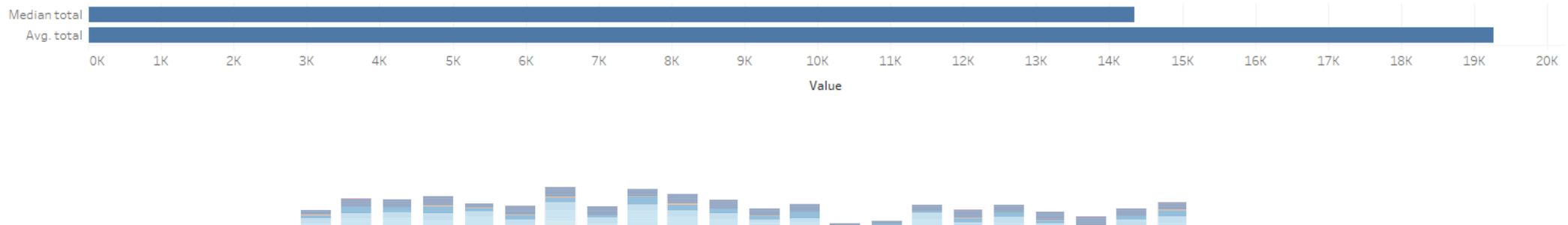
Of course, in Tableau, you can also look at your information visually. How would we do that?

Now, let's just look at the median compared to the average.

- Drag off the SUM(TOTAL) pill.

You can imagine if we left total on there, it would show us one long bar for total and two very tiny bars for median and average. That wouldn't be all that useful.

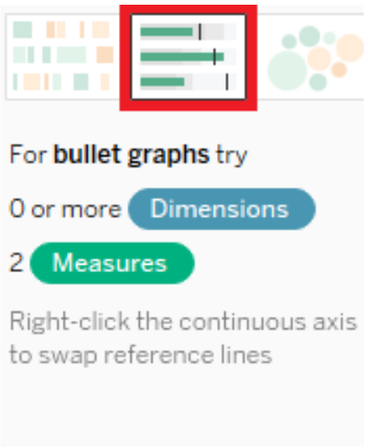
- Select the Bar Chart option on the Show Me Pane. You'll get something like this:



Basic filters and analyzing the data for story

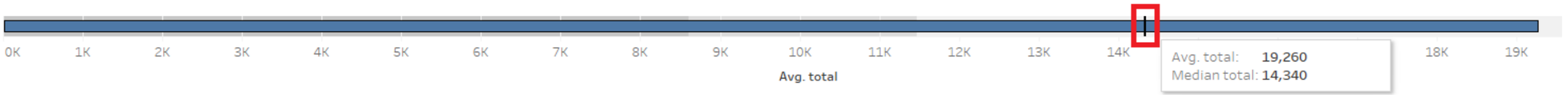
The **Show Me** pane shows you the options available to visualize your data.

If an option is greyed out, that means it's not possible to use it for the data you have. One other chart type that isn't greyed out here is a bullet chart (or graph), which often is used to look at a target VS the reality. In this case, we will tweak that traditional use of a bullet chart to look at the median compared to the average.



- Select the bullet graph option.

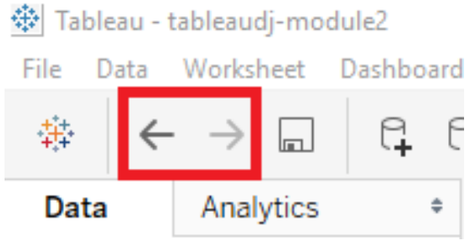
You'll see a bar chart with a line marked on it. The thin black line is the median, $14,340. And the blue bar represents the average of $19,260:



Basic filters and analyzing the data for story

Of course, just because you can use a certain chart doesn't mean you should in every instance.

But if you don't like what you've done after playing around with the data for a few minutes, just hit the **undo button** in the top left corner of your screen.



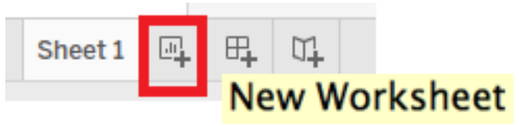
Now, we could just drag off the pills to get rid of the bullet graph with our average and median.

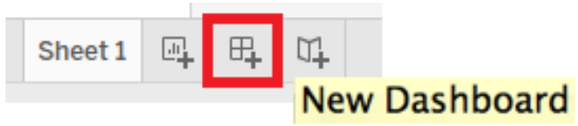
But what if we want to save that for easy reference? We can save it by leaving that viz and then adding a new sheet where we can do more analysis, or copy the sheet and do more analysis there.

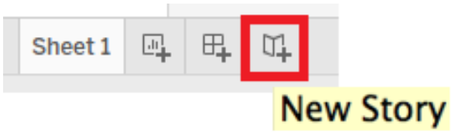
Basic filters and analyzing the data for story

You'll notice the sheet tabs, very similar to Excel, are on the bottom of the Workbook.

There is a little plus sign by three small icons to the right of the current sheet. The first one is for a new sheet. Then there is an icon that looks like window panes: that is for a new Dashboard. And the last one is for a new story. We'll get into the Dashboards and Story points a little later.





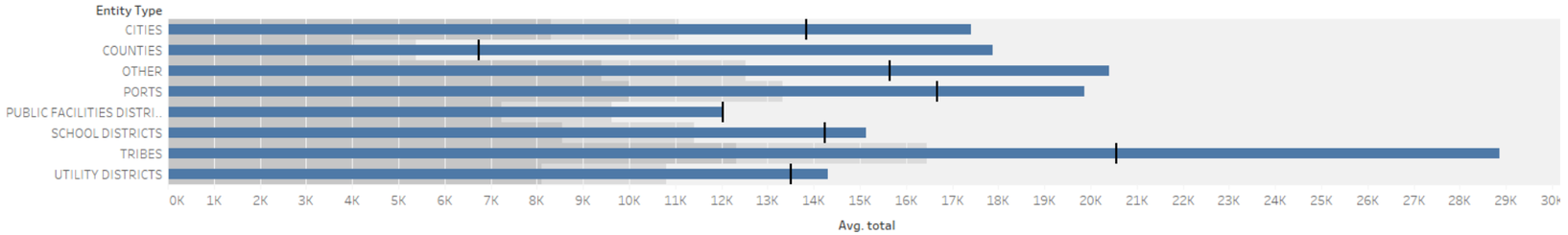


- For now, click on the plus sign for a new sheet.

Double-click on **Sheet 1** and then type the name you want for this sheet. I'm going to call it "bullet-summary." You can also copy a sheet by right-clicking or CTRL-clicking (on a Mac). Let's copy that bullet-summary sheet.

Basic filters and analyzing the data for story

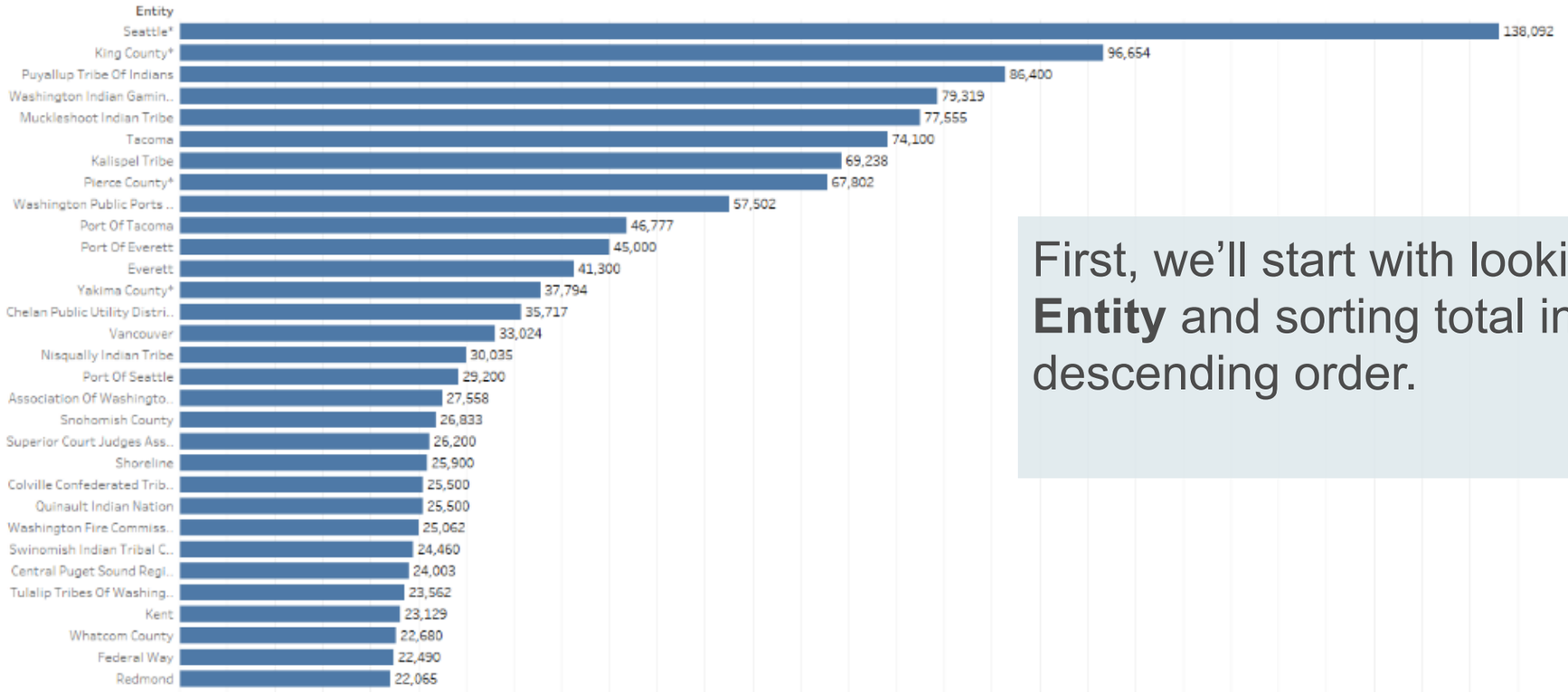
The first bullet chart gave us a useful summary of the median lobbying payment and the average, but let's take a closer look by the type of entity or government. The bullet chart in this case is not really something we would want to build out as a chart for our readers. But it does tell us which entity types we might want to examine more closely.



Where we see a big difference, that indicates that just a few local governments in that entity type accounted for much of the money going to lobbyists. For example, look at the difference between the mean and median in counties. Only two counties spent more than $60,000 on lobbying efforts. They also are the two most populous counties in the state: King County and Pierce County.

Basic filters and analyzing the data for story

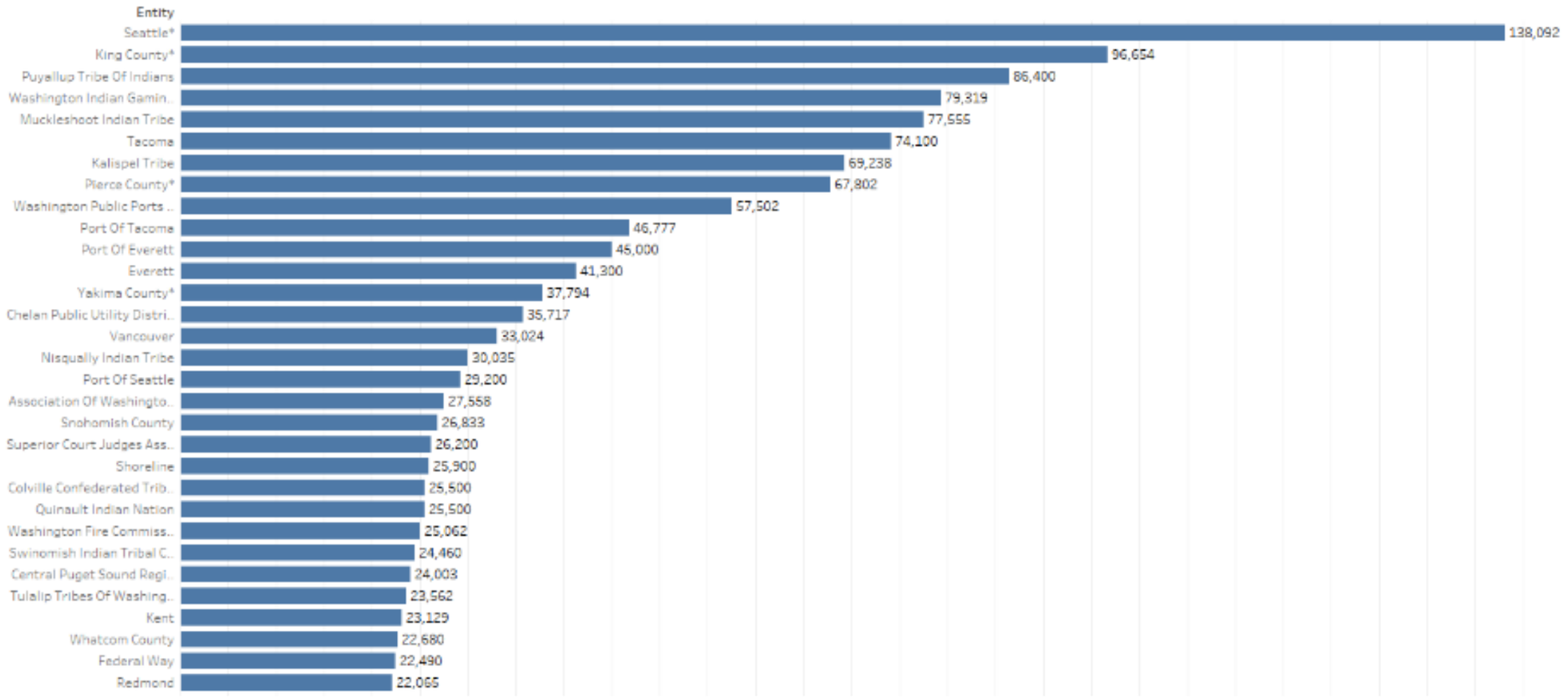
Now, what if we want to drill in more on the specific governments within each entity type and look at how much they paid lobbyists? There may be some newsworthy patterns.



First, we'll start with looking at **Entity** and sorting total in descending order.

Basic filters and analyzing the data for story

Now, you have everything sorted by total:

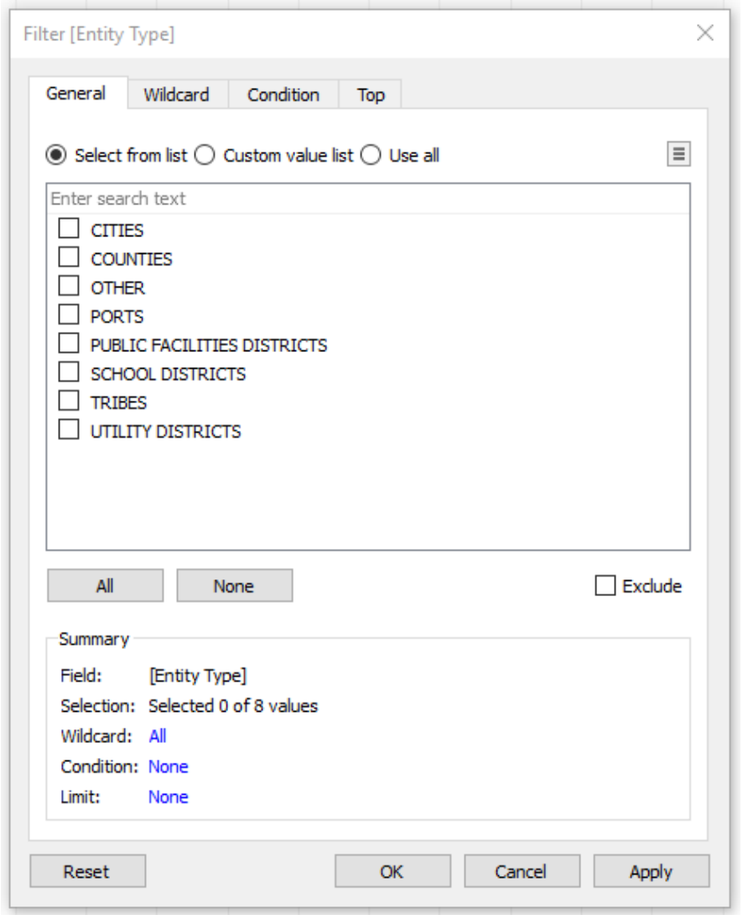


Basic filters and analyzing the data for story

But what if you want to just look at counties, or cities or tribes?

We're going to use the power of filtering next.

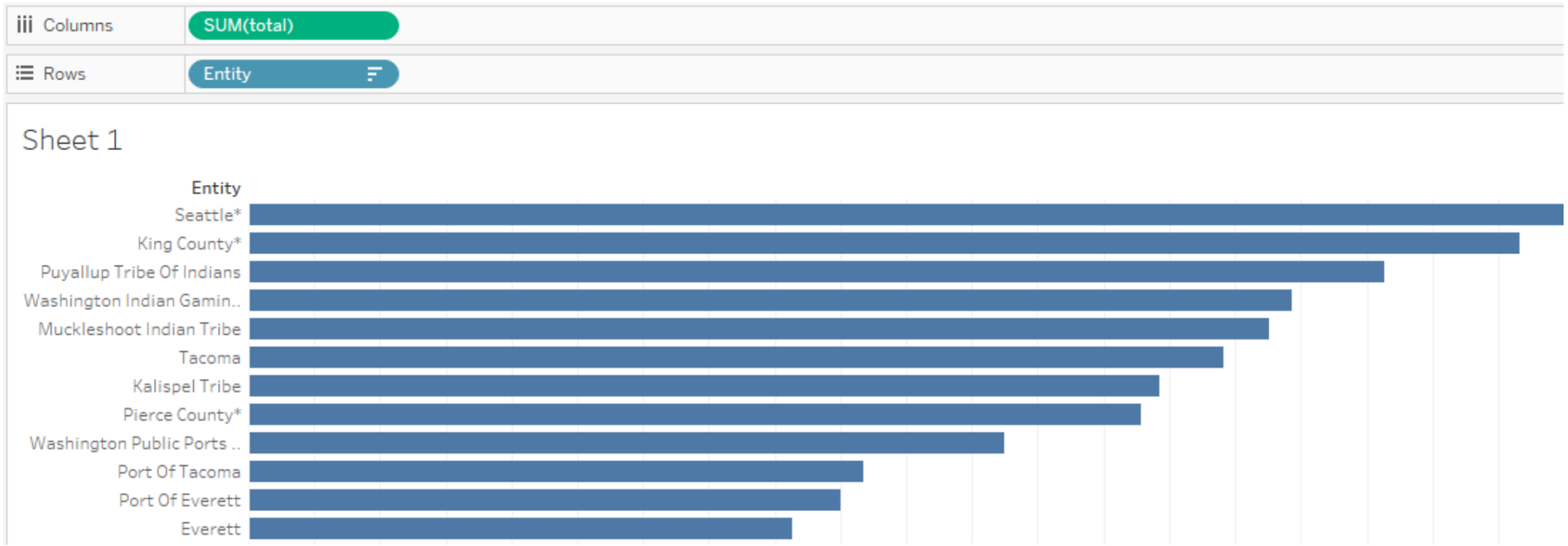
First, we'll show you a demo of how to set up filters.



Basic filters and analyzing the data for story

Filtering (step 1):

- Drag the total onto the columns shelf and entity onto the rows shelf. You can then practice sorting it in descending order. It will look like this:

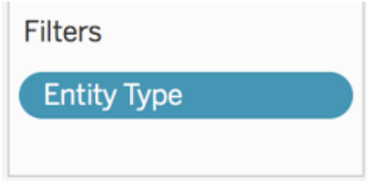


Basic filters and analyzing the data for story

Filtering (step 2):

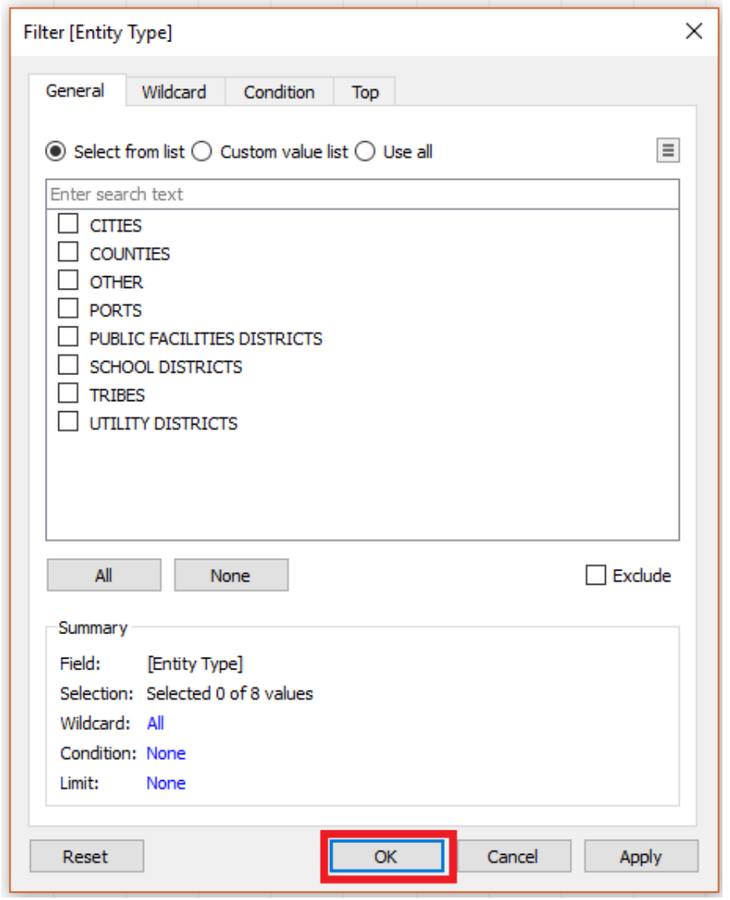
Now, this is somewhat useful. But we really want to compare entities that are like one another. Cities compared to cities and tribes to tribes. So, let's filter by entity type.

- Drag **Entity Type** onto the filter shelf:



When you do that, a dialogue box will pop up.

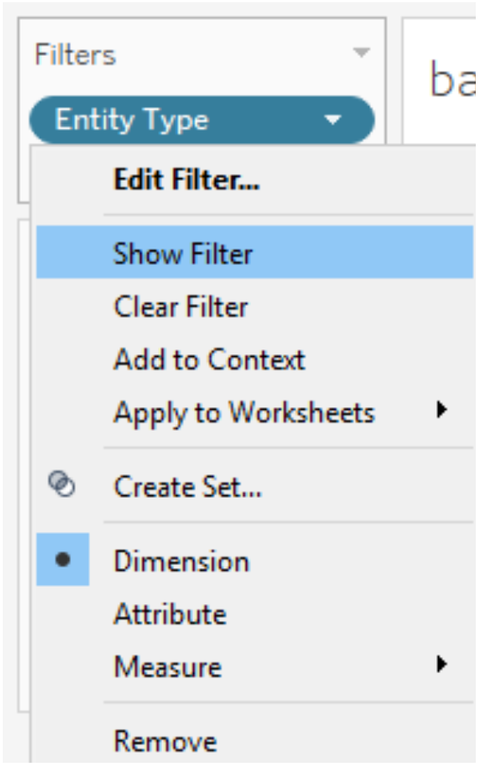
- Click **OK**



Basic filters and analyzing the data for story

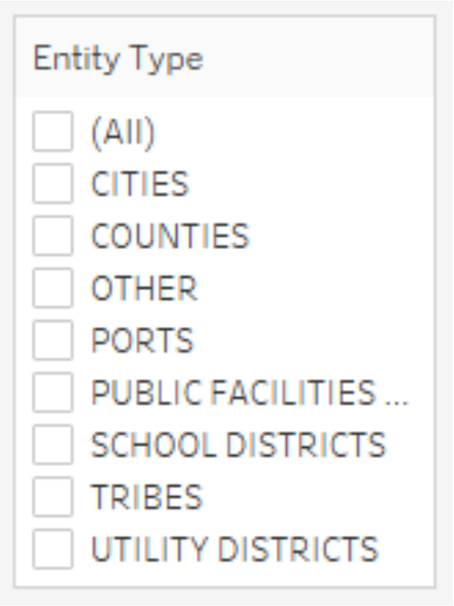
Filtering (step 3):

- Click on the caret on the right side of the **Entity Type** pill. The drop-down will offer you some choices.
- Select **Show Filter**:



When you do that, a box will appear on the right-hand side of your workspace.

By default, Tableau will provide empty check boxes. You can select all or any of the options.



Go ahead and practice to see what happens.

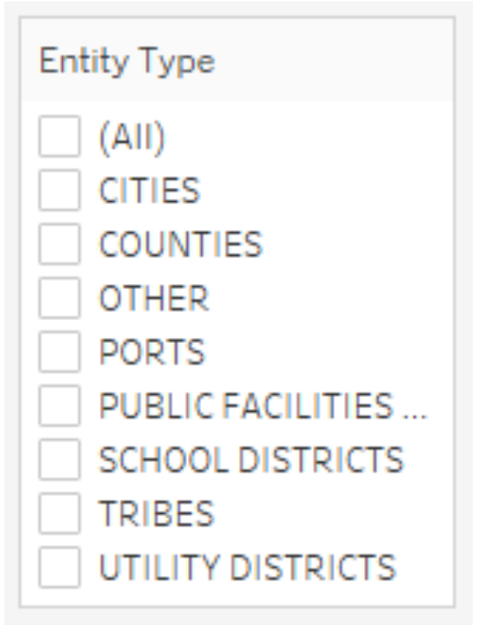
Basic filters and analyzing the data for story

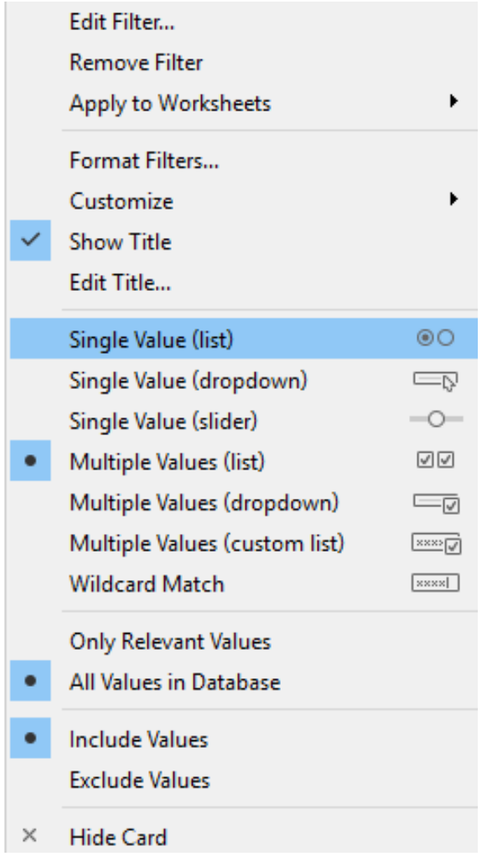
Filtering (step 4):

Now, we're going to modify our filter options a bit.

- Click on the caret on the right side of your filter, to the right of entity type.

When you do that, a box will appear. Tableau gives us a lot of options.



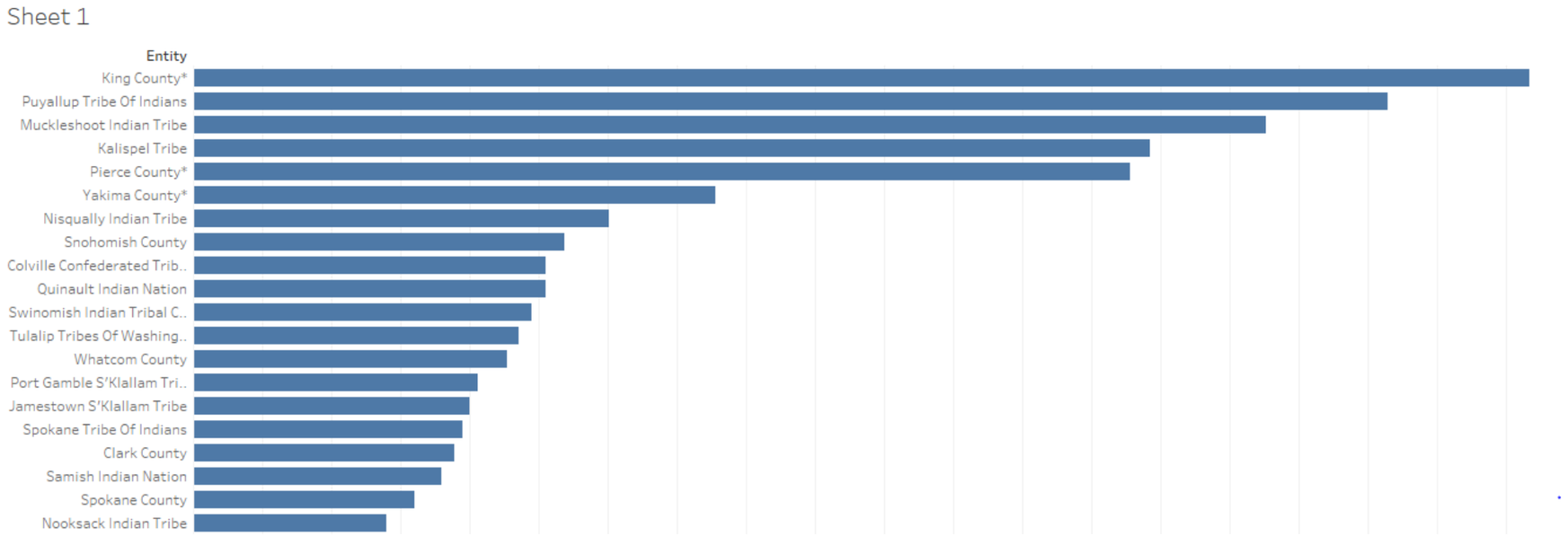


- For now, we're going to select **Single Value (list)**.

Basic filters and analyzing the data for story

Filtering (step 5):

Now, practice selecting and re-selecting different entity types. You can easily see the amount of lobbying by the various local governments but filtered by the type of government. Let’s practice that for a minute.



Basic filters and analyzing the data for story

Now, as additional practice, how many entity types do I have?

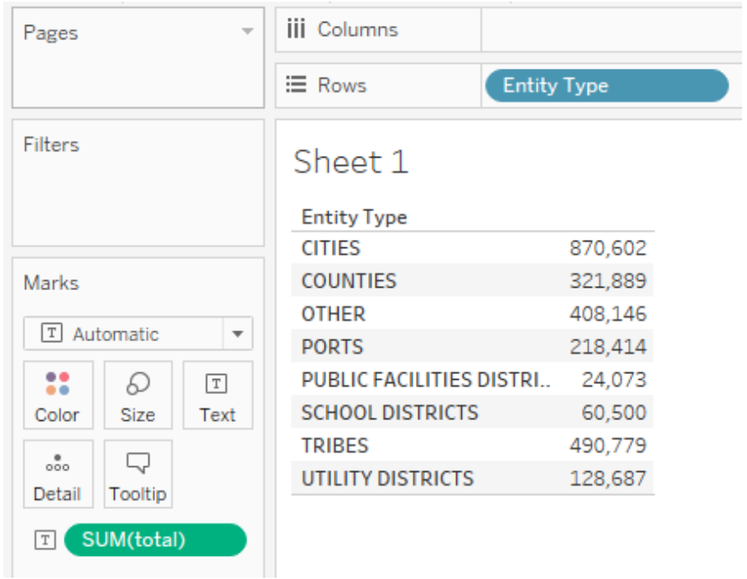
Which type of entity accounted for the most in lobbying spending, and how much was that?

But what if I want to get a sense of how much each of these gave as a percentage out of the total? Should I do another calculated field? What other ways, based on what we learned in the last module, could I contemplate exploring that?

Basic filters and analyzing the data for story

Tree Map (step 1):

Tableau also has some nice shortcuts. Instead of dragging the pill out to a certain shelf. Try double-clicking total. Then double-click **Entity Type**. You should end up with a tabular chart as a default.



The **Show Me** pane will show you the other visualization options.

- Select **Tree Map**.

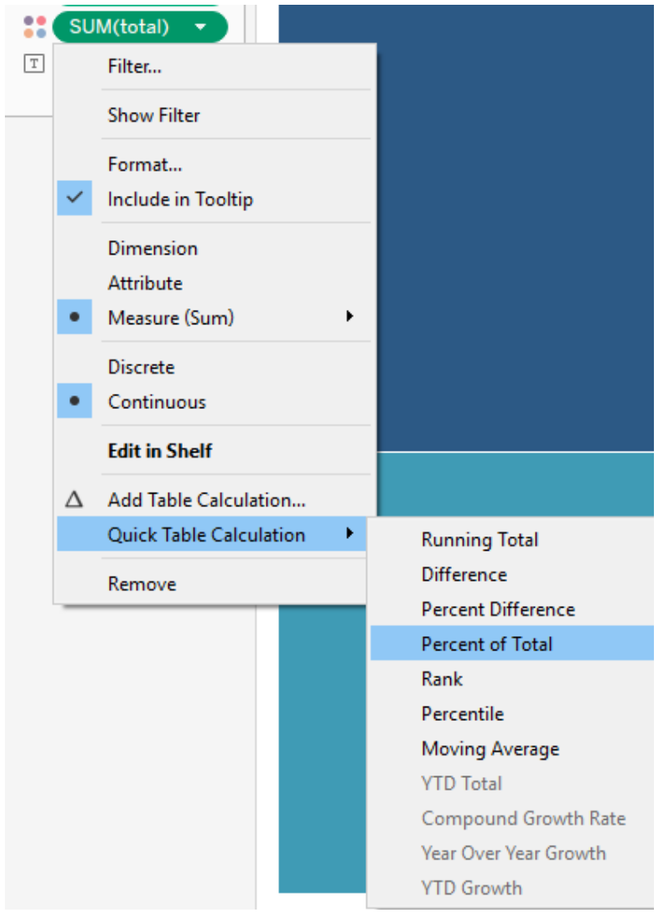
Basic filters and analyzing the data for story

Tree Map (step 2):

If we hover over each section, we can see the total in lobbying to that entity type.

But we also can add a calculation – the percentage out of the total would be useful to know.

Then, when you hover over the tree map, you will see the type of government, the total and the percentage of the total.

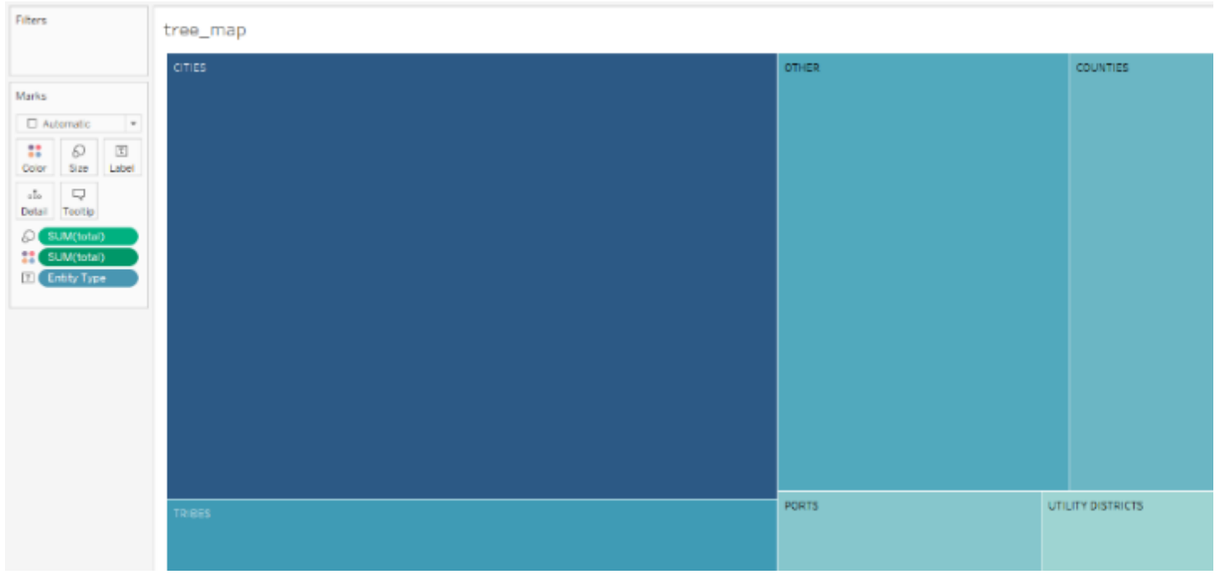


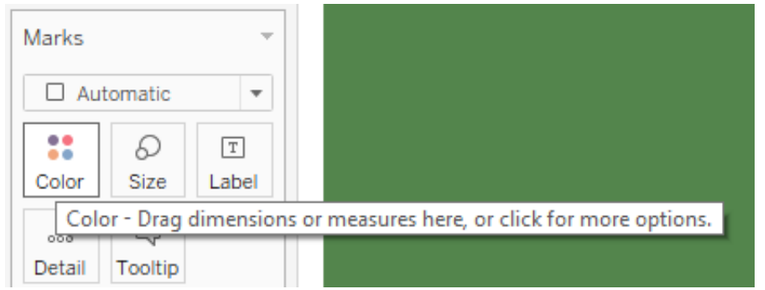
Basic filters and analyzing the data for story

Tree Map (step 3):

Adjustments in Color. You can change the color scheme of your tree map by double-clicking on color and selecting edit. Go ahead and try out some other options for the tree map.

After you've played around a bit, change the color scheme to green.



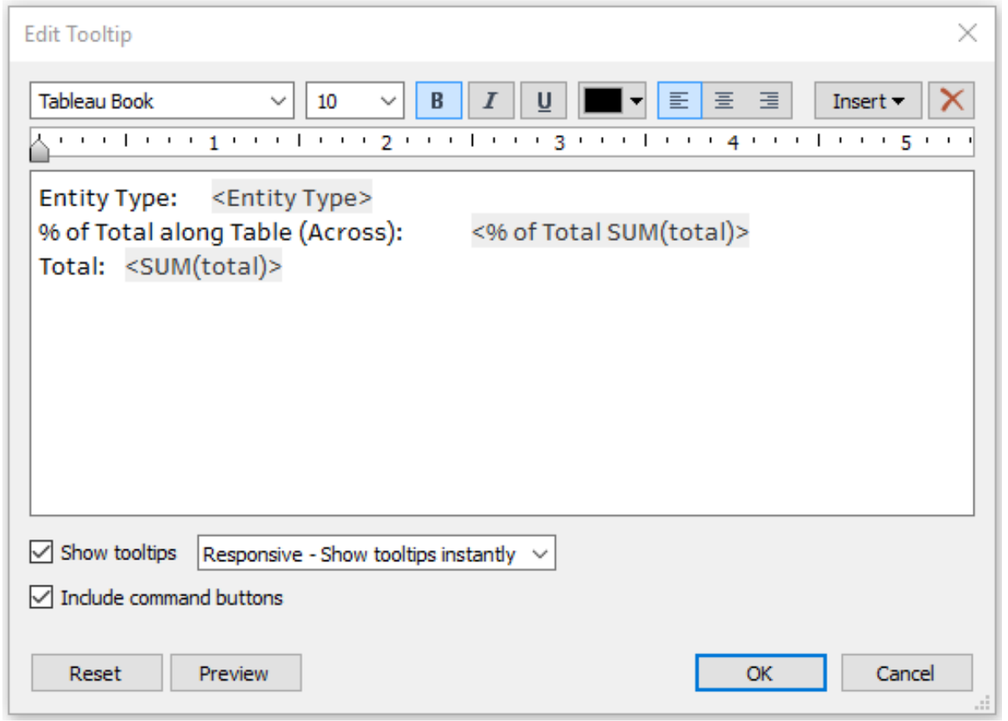


Basic filters and analyzing the data for story

Tree Map (step 4):

Adjustments in tooltip styles. Since we are using the tooltip with this tree map, let's take the opportunity to learn how to edit a tooltip.

- Double-click on the **tooltip icon** in the marks shelf. A box will open. You can edit the font and text right from the tooltip box.
- Go ahead and change **Entity Type** to **Government**, and **% of Total along Table (Across)** to **% of Total SUM (total).**
- Click **OK** and look at your tooltip. We'll do more with tooltips in future modules.



Basic filters and analyzing the data for story

- Now, as additional practice, how many entity types do I have?
- Which type of entity accounted for the most in lobbying spending?
- And how much was that?

Basic filters and analyzing the data for story

- So, is it newsworthy if the most populous areas, such as King County, or the city of Seattle, are responsible for most of the lobbying to the legislature?
- What about smaller governmental bodies and how much are they paying lobbyists?
- How could we look at that in a way that would be enlightening?
- What additional data might we need?

Titanic!

Titanic dataset and cool things you can do with it: https://www.kaggle.com/c/titanic

Tableau official training: https://www.tableau.com/learn

Tableau examples with Titanic data: https://public.tableau.com/search/all/titanic

Titanic Questions

- Can you attribute survival to a single primary trait? In other words, can you attempt to tease apart the confounding effects of pclass/age/sex, when assessing one?

- Can you explain why the distribution of fares by pclass seems off? why are some 3rd class tickets more expensive than first class?

- Does group size have an effect on survival rate?