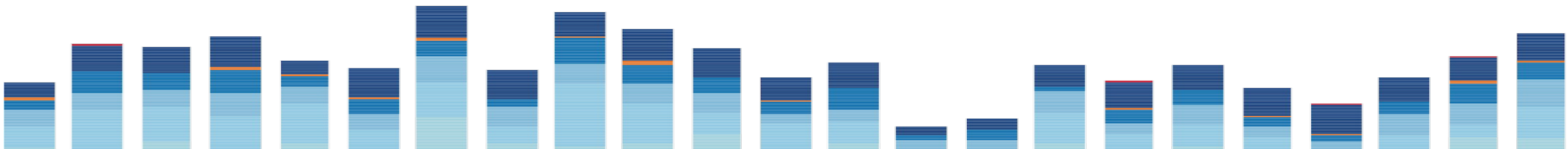# CSE 91

What Happened?

Answers on Day 1

Credited Author: Cheryl Phillips

Cheryl Phillips previously worked at The Seattle Times from 2002-2014. Her most recent position in Seattle was as Data Innovation Editor. In that role, she analyzed data for stories, facilitated online storytelling and coordinated newsroom data journalism training. She also was the deputy investigations editor, an assistant metro editor and an investigative reporter at The Seattle Times. In 2014, she was involved in coverage of a landslide that killed 43 people and was particularly focused on collecting and using data to help cover the story. That coverage received a Pulitzer Prize for breaking news. In 2009, she was the lone editor in the newsroom when four police officers were shot at a coffee shop and was integrally involved in the subsequent coverage of the shooting and 30-hour manhunt for the suspect. That work by the newsroom received a Pulitzer Prize for breaking news. She also has twice been on teams that were Pulitzer finalists. She has worked at USA Today and at newspapers in Michigan, Montana and Texas. Cheryl has taught data journalism and data visualization at the University of Washington and Seattle University. She also served for 10 years on the board of directors for Investigative Reporters and Editors, a grassroots training organization for journalists and she is a former IRE board president. She currently serves on an advisory board for Tableau Public, a data visualization software tool. Twitter: @cephillips | Faculty Profile

Data Journalism and Tableau:

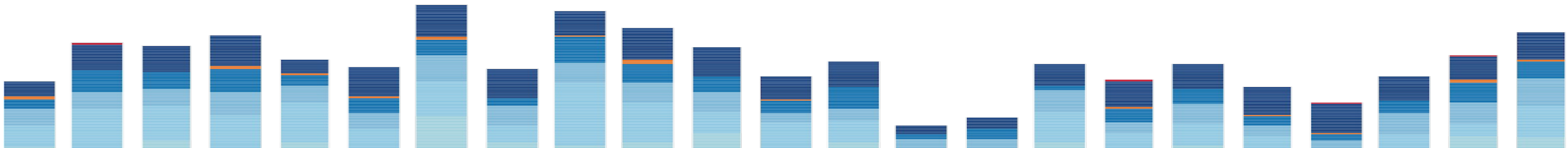## Using data for more power in storytelling

Tableau basics

- Intro to interface

- Importing

- Dimensions and Measures

- Chart types

- Finding the story

- Practice on your own

Your scenario

- You are a reporter for a metro daily news organization and you cover local politics.

- You have just downloaded data showing how much each local governmental agency in the state has paid for lobbyists to go to the state capitol and work for their causes.

- You think there's a story here, but you're not sure what it is.

- So, let's open this data up and take a look at how Tableau can help you figure that out.

- At the same time, you'll get an introduction into how Tableau works.
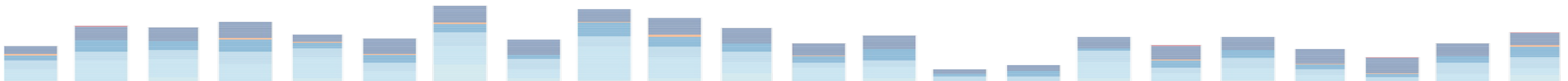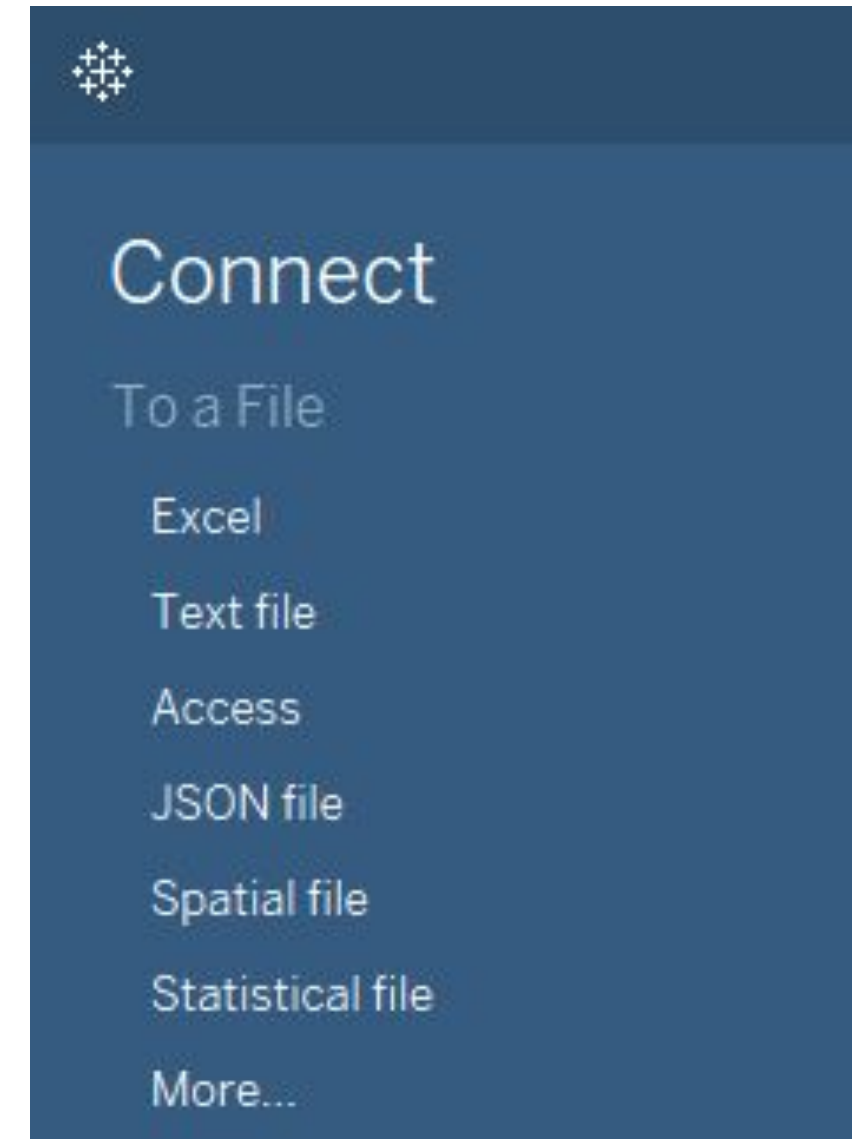
Getting started

Tableau guides you with connecting to or importing data. On the left-hand side of the screen you will see the Connect bar.

The choices for the data types you can use in Tableau are listed under Connect.

What do they all mean?

Let's start with the basics.

Connect

To a File

Excel

Text file

Access

JSON file

Spatial file

Statistical file

More...

Getting started

In Tableau, we are working with structured data. A text file is usually delimited, often with tabs or commas:

Tabs: Amount **[tab]** Employee Last Name **[tab]** Employee First Name

Commas: Amount**,**Employee Last Name**,**Employee First Name

These files often have the following extensions:

.txt for text

.csv (comma-separated values)

.tsv (tab-separated values)

Text is text.

Even when it is:

o Numbers *(1,2,3,)*

o Dates *(05/26/02005)*

o Map coordinates (*Lat: 34.515610, Long:-109.687500*

A text file is just a computer file with letters and numbers organized in rows and columns of information.

Importing that lobbying data

The import window for Tableau has several features.

- On the left are the connections to data sets.

- When you click on the data type, Tableau will open a dialogue box where you can browse to the file you want.

- Once you select the file, it will appear in the top screen and the rows and columns associated with that data will appear in the bottom screen.

- And on the very bottom, toward the left, if you click on **Sheet 1** you will go to the Tableau workbook where you can begin to analyze the data.
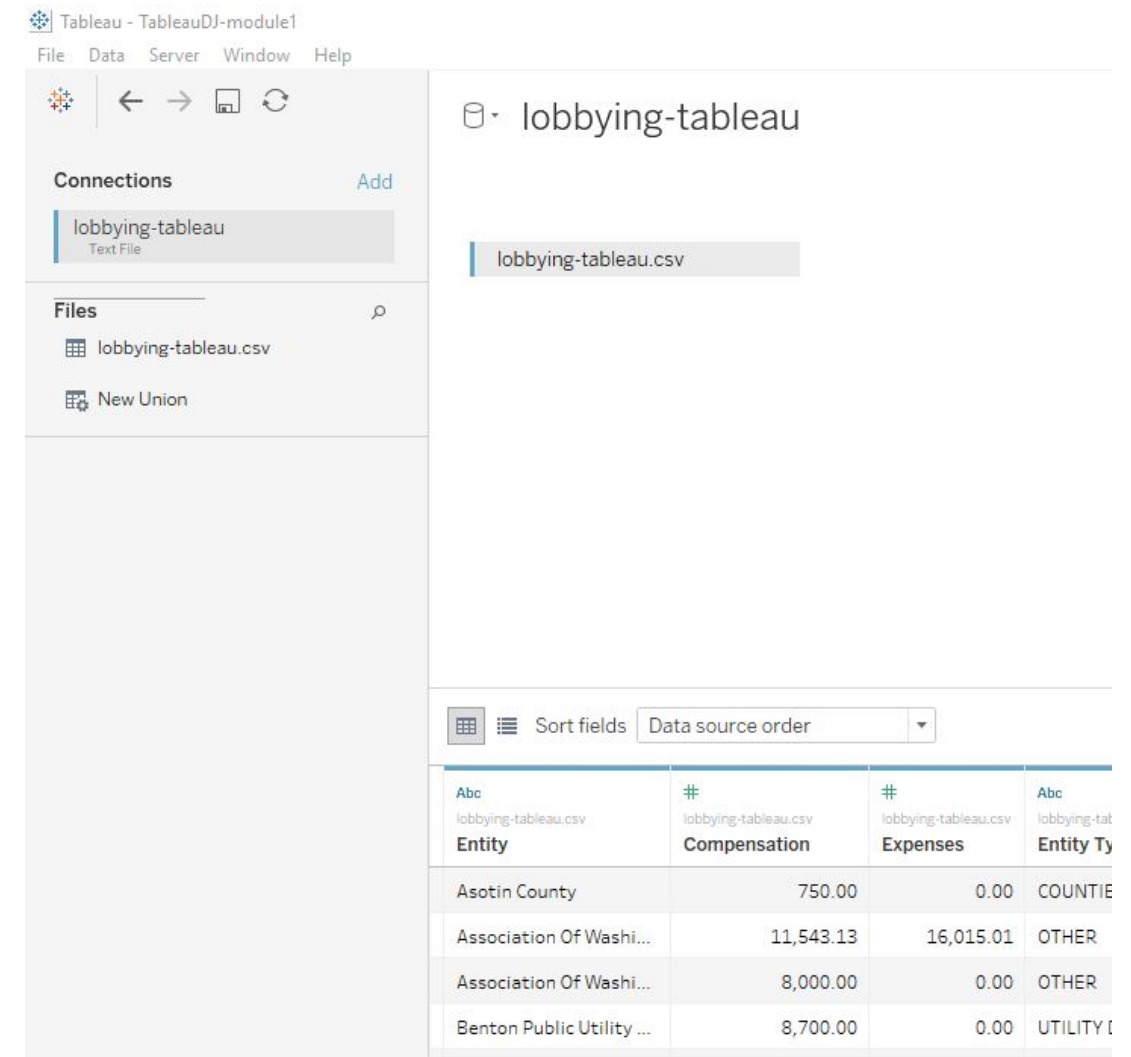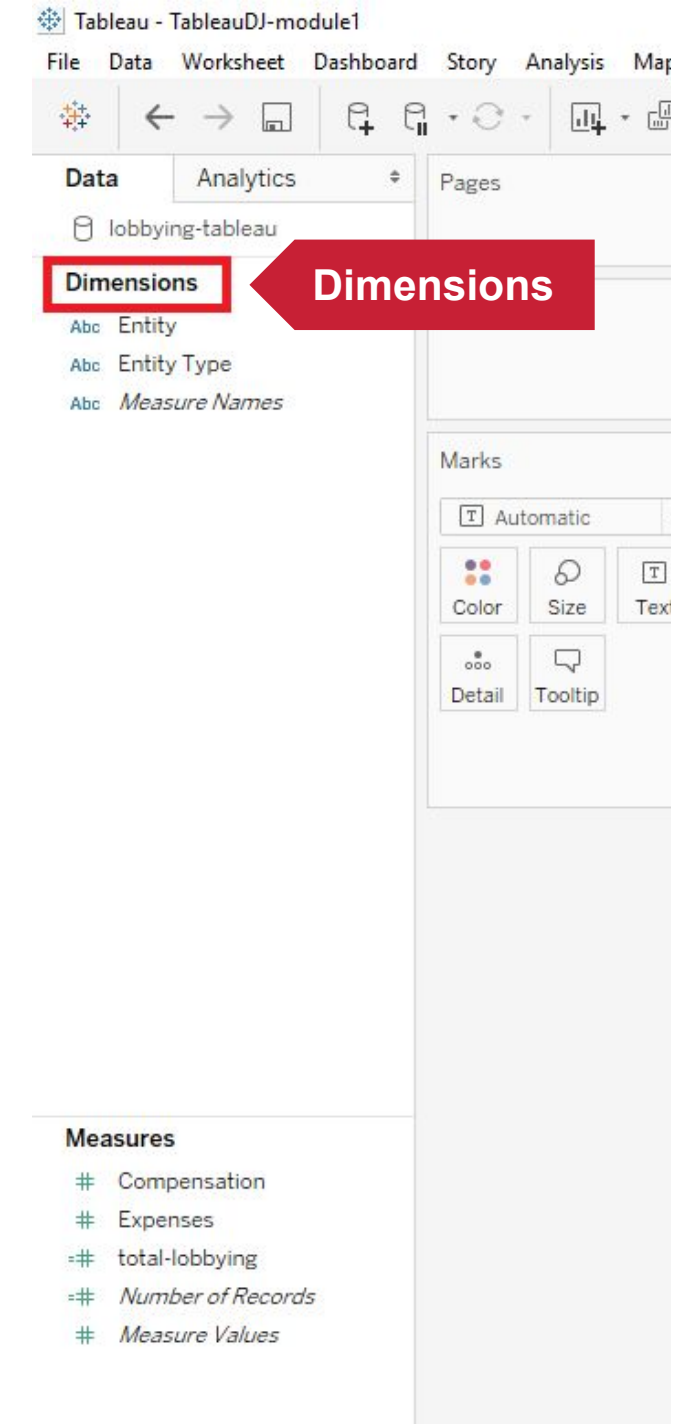
Who is paying those lobbyists?
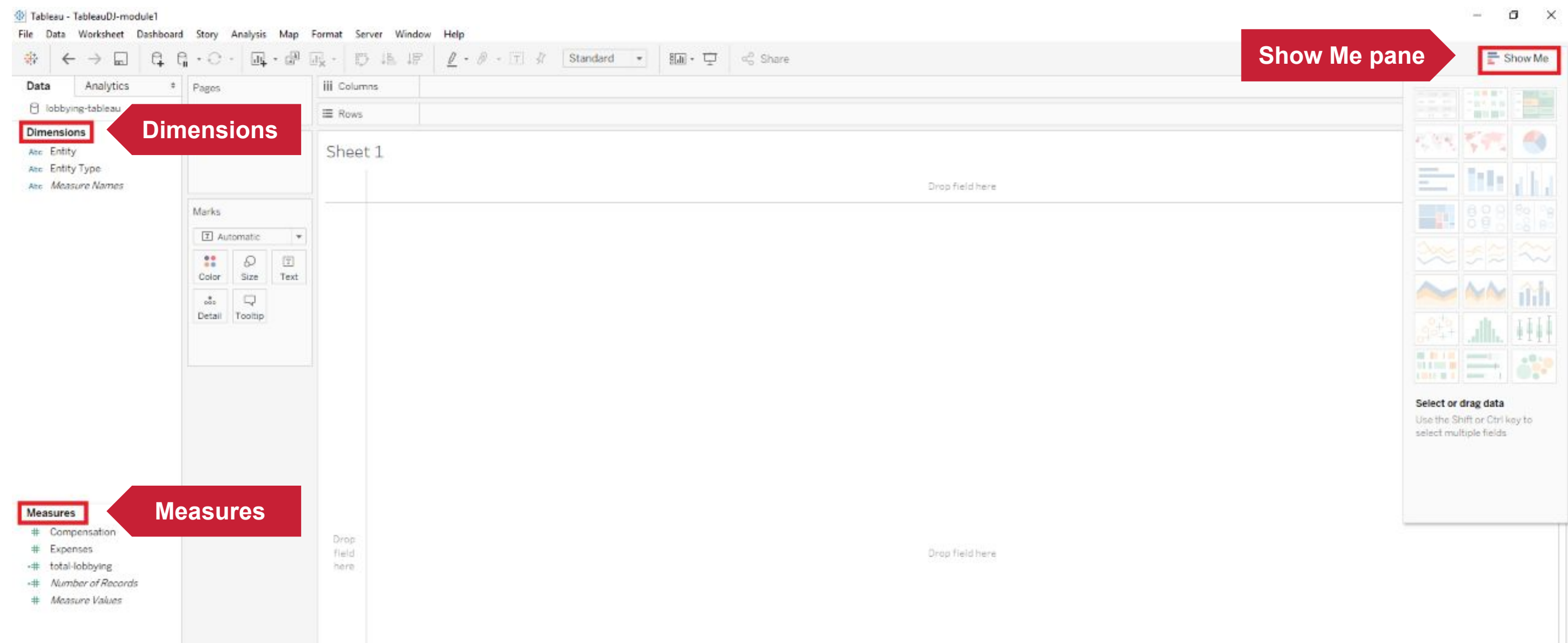
Who is paying those lobbyists?

Dimensions are typically those columns or fields that are text or that you are not going to do math on. So the name of the entity (local government) and the entity type are dimensions.

Dimensions are typically:

- Discrete data, not continuous as in numbers, but there is an exception: date fields are dimensions.

- Another type of dimension that has numbers in it is a zip code. You aren't going to add up zip codes, but you might want to know how many time each zip code appears in the data.

- Finally, information that can be mapped like city, county and state (geographic data) are also dimensions.
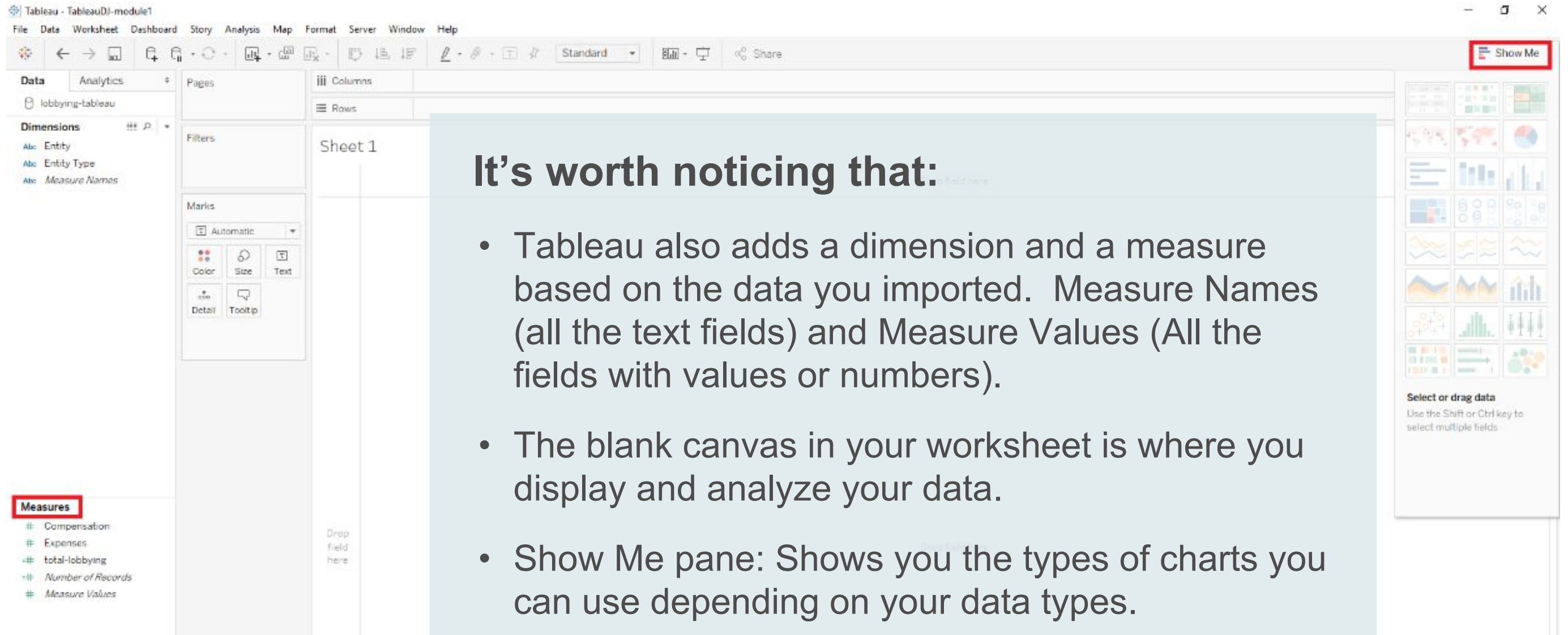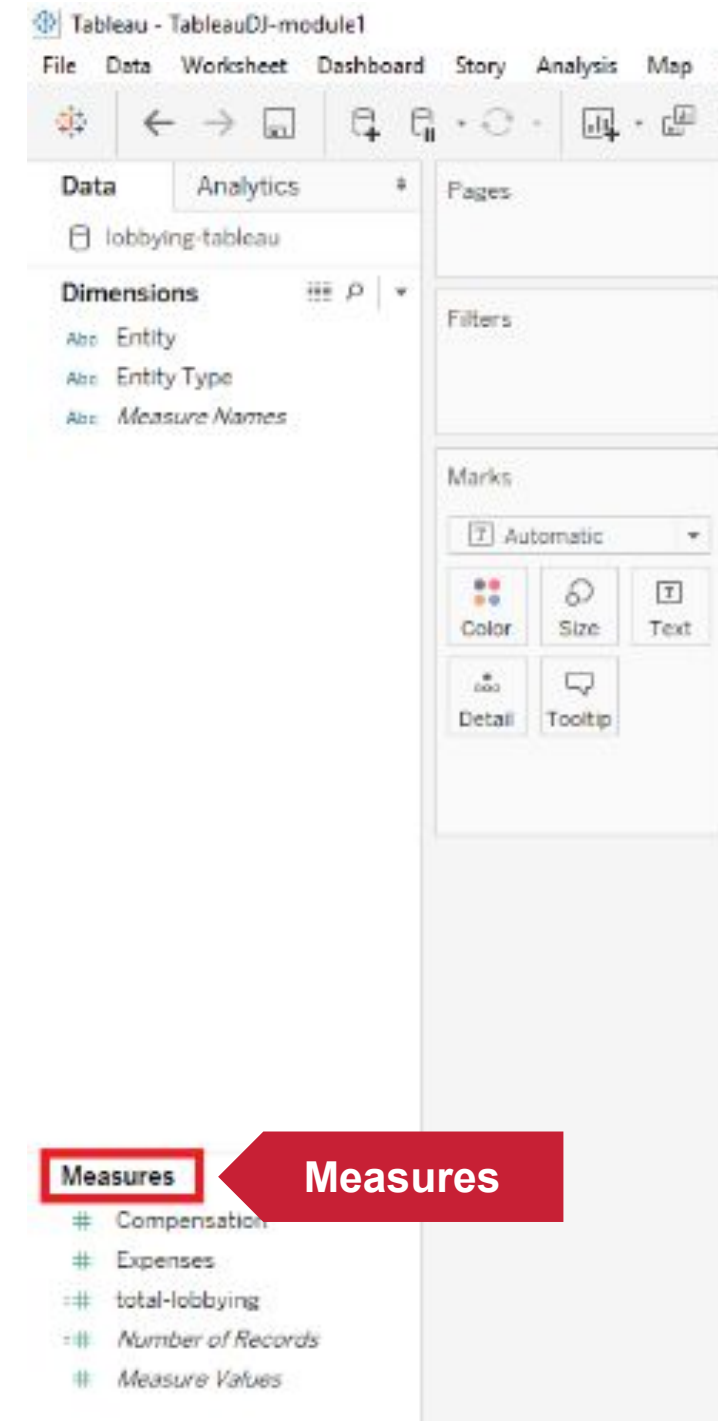
Who is paying those lobbyists?

Let's figure out what's going on here.

Measures are data you can calculate, so the compensation and expenses are measures. Measures are continuous data.

- Continuous data: quantitative data that can be measured in some way.

- In this class, the number of students is discrete data. There cannot be half of a student.

- The age of each student is continuous data. It can be any age (within the range of possible human age).

Who is paying those lobbyists?


It's worth noticing that:

- Tableau also adds a dimension and a measure based on the data you imported. Measure Names (all the text fields) and Measure Values (All the fields with values or numbers).

- The blank canvas in your worksheet is where you display and analyze your data.

- Show Me pane: Shows you the types of charts you can use depending on your data types.
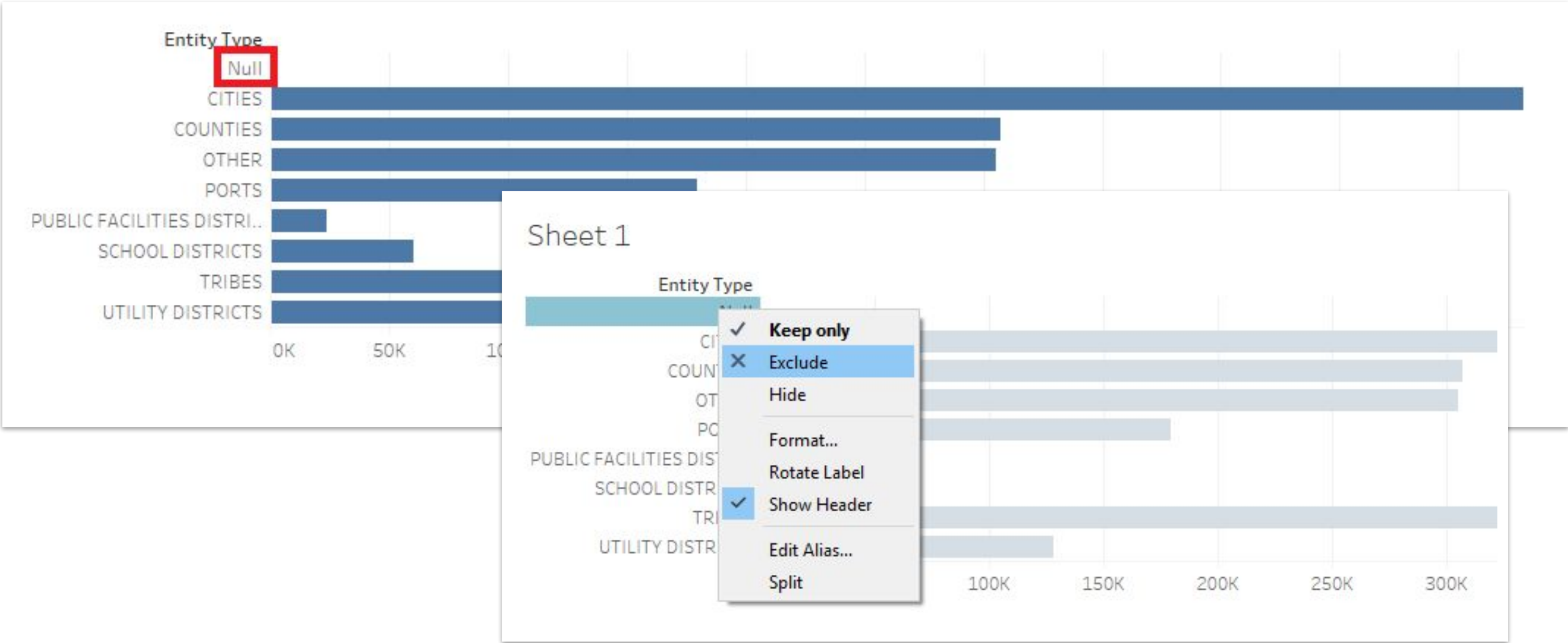
Let's get started!

Who is paying those lobbyists?

When you hover over one of the Dimensions or Measures, they will be highlighted. I call those pills.

You can click on the pills and drag and drop them to visually explore the data.
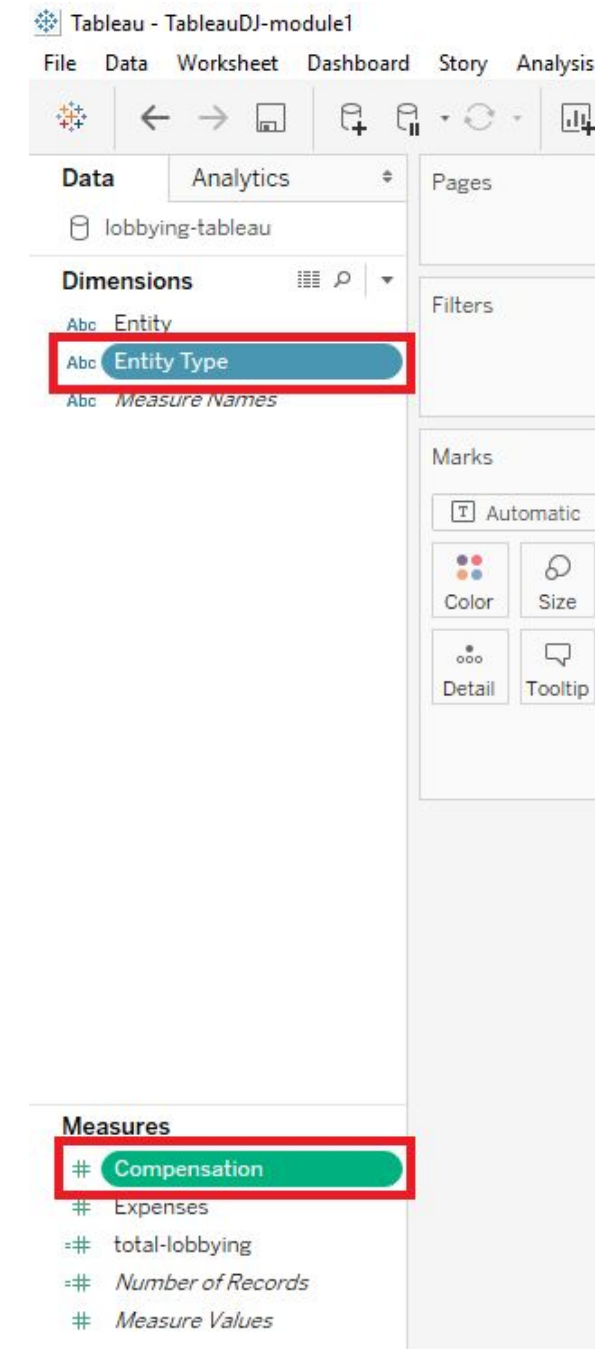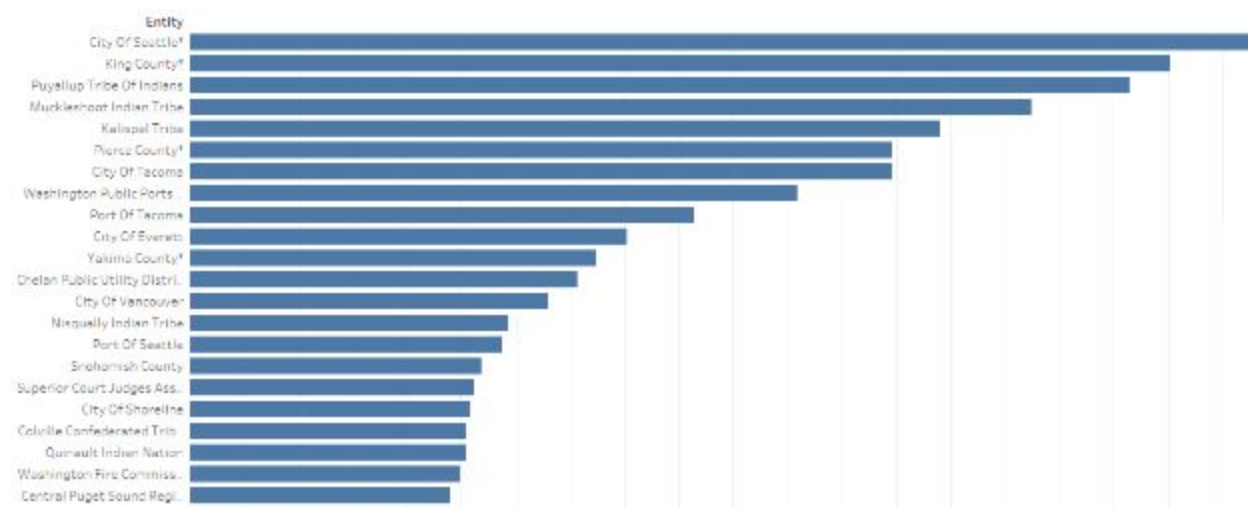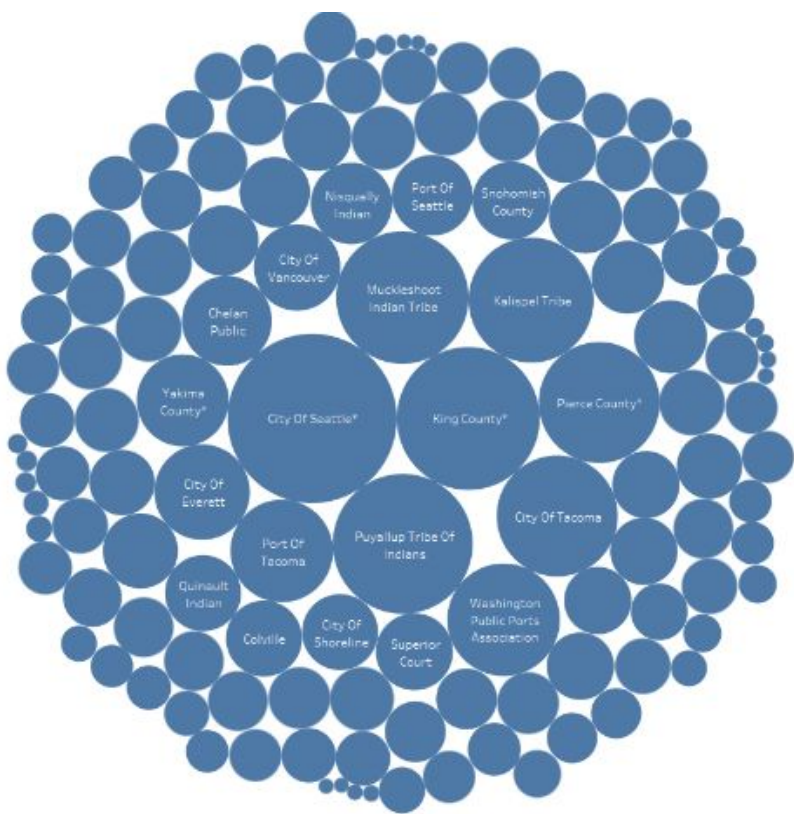
Who is paying those lobbyists?
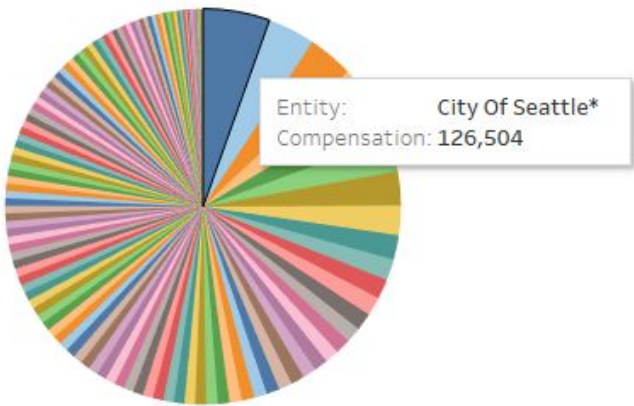
Who is paying those lobbyists?

Bar Chart

Pie Chart

Entity: City Of Seattle*
Compensation: 126,504

Bubble Chart

Highlights Chart

Entity	
City Of Seattle*	126,504
King County*	90,159
Puyallup Tribe Of Indians	86,400
Muckleshoot Indian Tribe	77,372
Kalispel Tribe	69,000
Pierce County*	64,602
City Of Tacoma	64,600
Washington Public Ports ..	55,882
Port Of Tacoma	46,425
City Of Everett	40,200
Yakima County*	37,333
Chelan Public Utility Distr..	35,650
City Of Vancouver	33,024
Nisqually Indian Tribe	29,333
Port Of Seattle	28,700
Snohomish County	26,833
Superior Court Judges As..	26,200
City Of Shoreline	25,900
Colville Confederated Trib..	25,500
Quinault Indian Nation	25,500
Washington Fire Commiss..	24,907
Central Puget Sound Regi..	24,000
Swinomish Indian Tribal C..	24,000
Tulalip Tribes Of Washing..	23,500
City Of Federal Way	22,000

Who is paying those lobbyists?
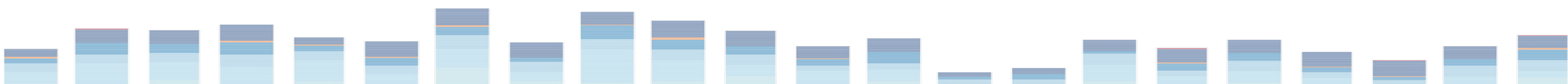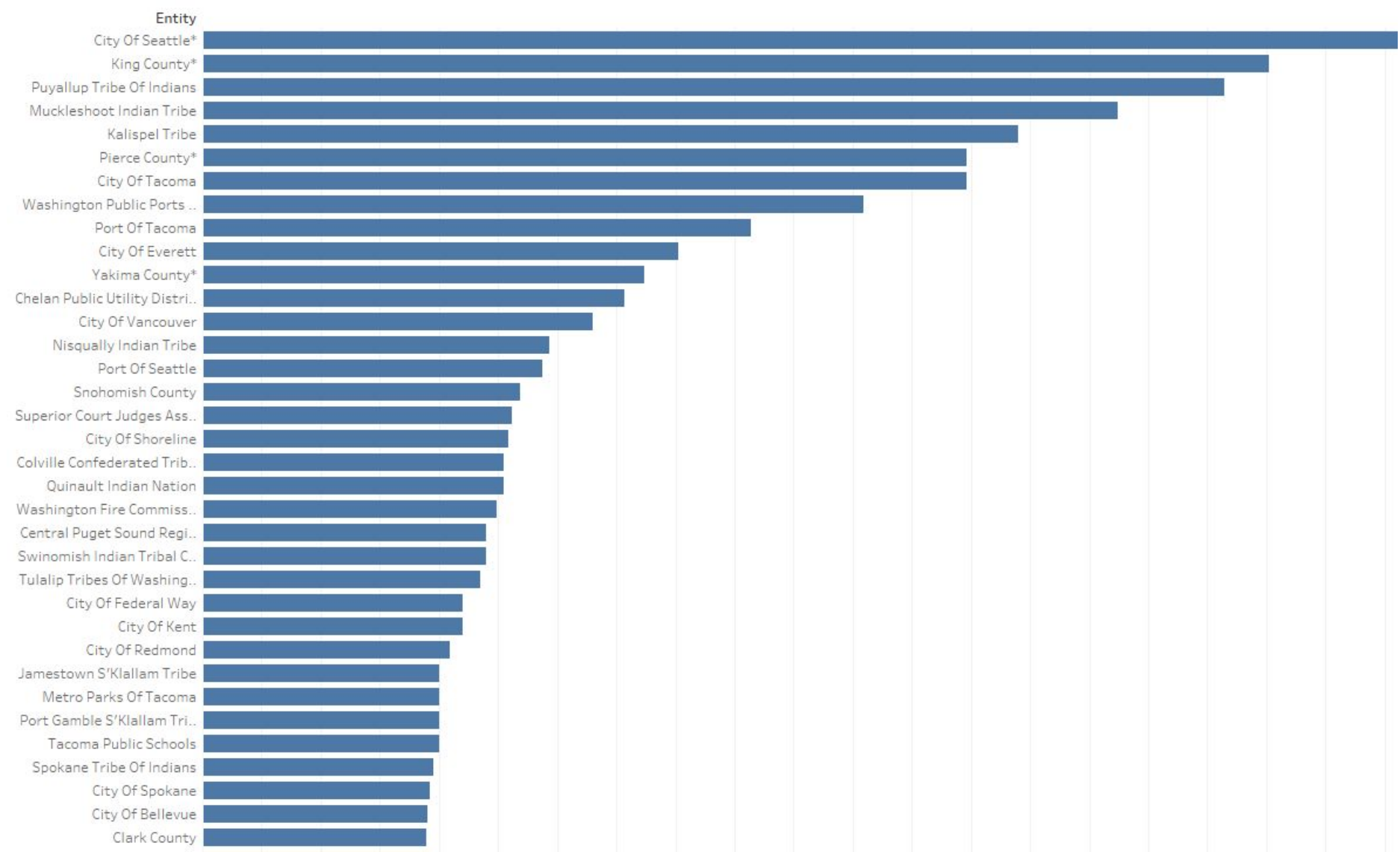
# Who is paying those lobbyists?

Let's do a calculation:

o   On the menu bar at the top, select **Analysis** and then **Calculated Field**.

o   Give your new calculated field a name: **total-lobbying**.

o   Then type the **calculation** you want to do. *Tableau offers guidance as you go along and will tell you if your calculation is valid at the bottom of the window.*

total-lobbying

[Compensation] + [Expenses]

The calculation is valid.   Apply   OK

In this case, we want to add the **Compensation** and the **Expenses**:

o   Now, go ahead and create a bar chart, but with the total-lobbying field. **Sort**.

Which government spent the most? How much?
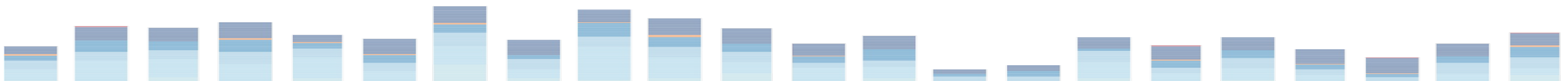
Practice

Now you try in pairs with the Titanic data.

- How many people were on the ship?
- What was the total of all fares paid?

Think of other questions you might want to answer.