

# DSC 96

What Happened?

Answers on Day 1

Giorgio Quer (adapted from Colin Jemmott)



# 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.



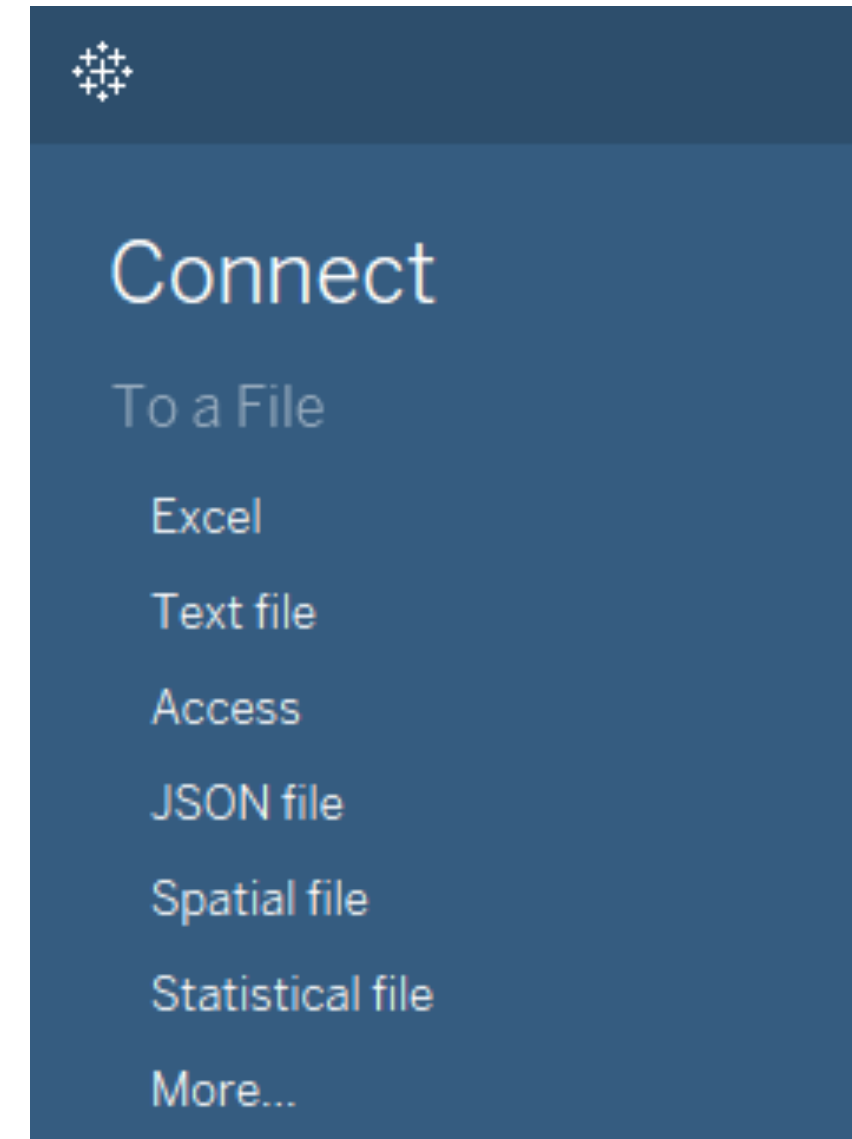
# Getting started

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.



# 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:

- Numbers (1,2,3,)
- Dates (05/26/02005)
- 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.



# How does the data look like?

---

Entity	Compensation	Expenses	entity_type
Asotin County	750.00	0.00	COUNTIES
Association Of Washington Cities	11543.13	16015.01	OTHER
Association Of Washington State Public Facilities Districts	8000.00	0.00	OTHER
Benton Public Utility District	8700.00	0.00	UTILITY DISTRICTS
Central Puget Sound Regional Transit Authority	24000.00	3.00	OTHER
Chehalis Confederated Tribes	4500.00	0.00	TRIBES
Chelan Public Utility District	35650.00	67.00	UTILITY DISTRICTS
City Of Algona	18000.00	0.00	CITIES
City Of Arlington	13687.50	0.00	CITIES
City Of Auburn	1900.00	0.00	CITIES
City Of Battle Ground	12000.00	0.00	CITIES
City Of Bellevue	19000.00	0.00	CITIES
City Of Bellingham	5733.32	4332.68	CITIES
City Of Burien	18000.00	0.00	CITIES
City Of College Place	2000.00	0.00	CITIES
City Of Covington	7166.68	0.00	CITIES
City Of Edmonds	16540.00	0.00	CITIES
City Of Everett	40200.00	1099.55	CITIES



# How does the data look like?

## Lobbying

Entity	Compensation	Expenses	entity_type
Asotin County	750	0	COUNTIES
Association Of Washington Cities	11543.13	16015.01	CTHER
Association Of Washington State Public Facilities Districts	8000	0	CTHER
Benton Public Utility District	8700	0	UTILITY DISTRICTS
Central Puget Sound Regional Transit Authority	24000	3	CTHER
Chehalis Confederated Tribes	4500	0	TRIBES
Chelan Public Utility District	35650	67	UTILITY DISTRICTS
City Of Algona	18000	0	CITIES
City Of Arlington	13687.5	0	CITIES
City Of Auburn	1900	0	CITIES
City Of Battle Ground	12000	0	CITIES
City Of Bellevue	19000	0	CITIES
City Of Bellingham	5733.32	4332.68	CITIFS
City Of Burien	18000	0	CITIES
City Of College Place	2000	0	CITIES
City Of Covington	7166.68	0	CITIES
City Of Edmonds	16540	0	CITIES
City Of Everett	40200	1099.55	CITIES

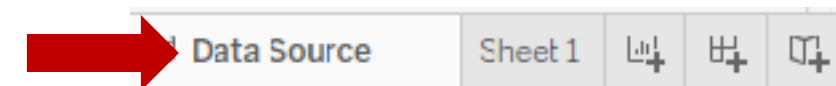
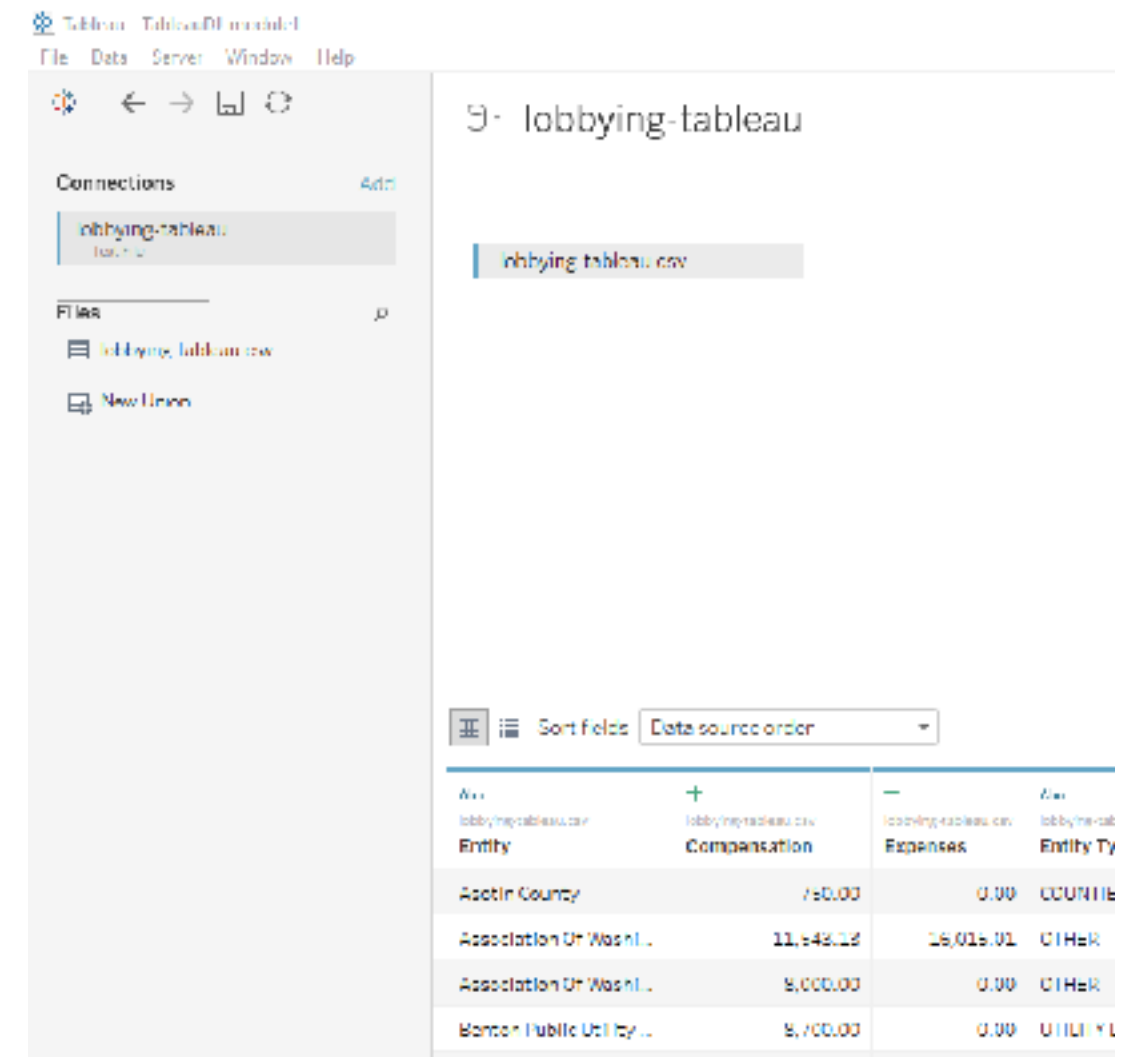




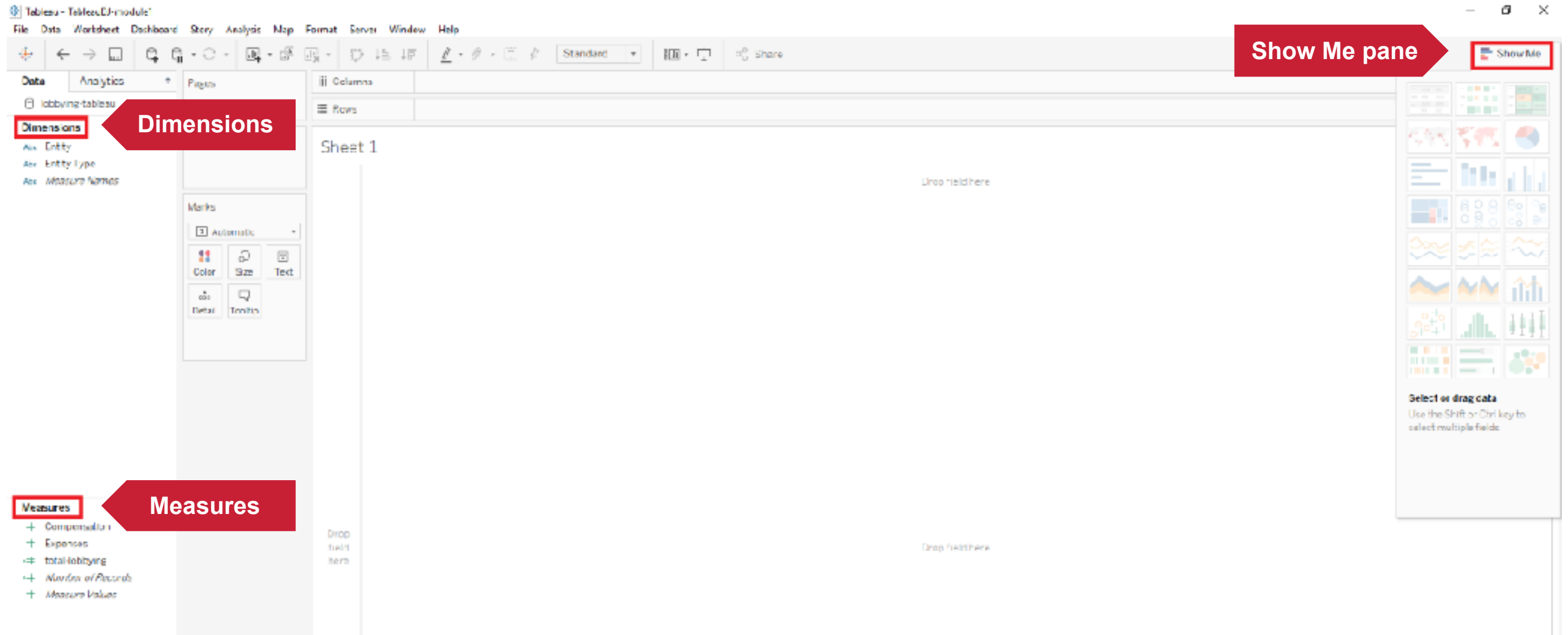
# Importing that lobbying data

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?

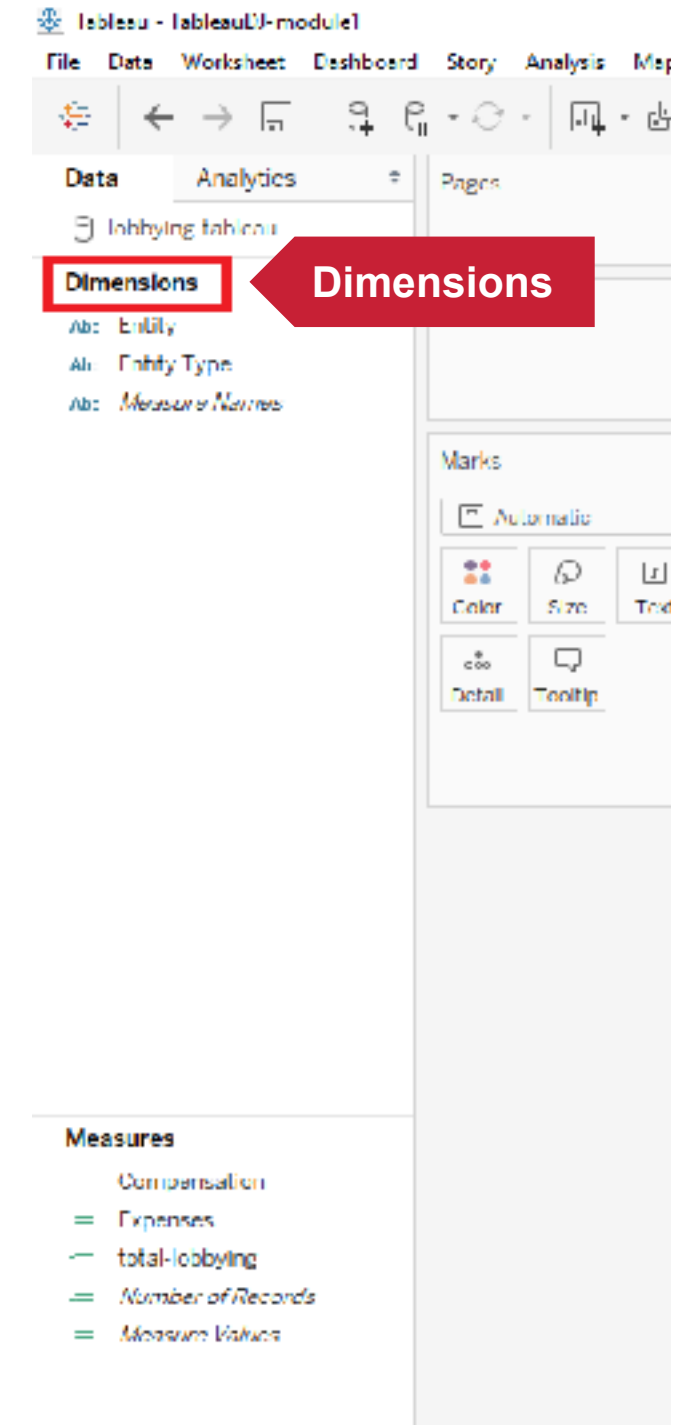


# Who is paying those lobbyists?

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.

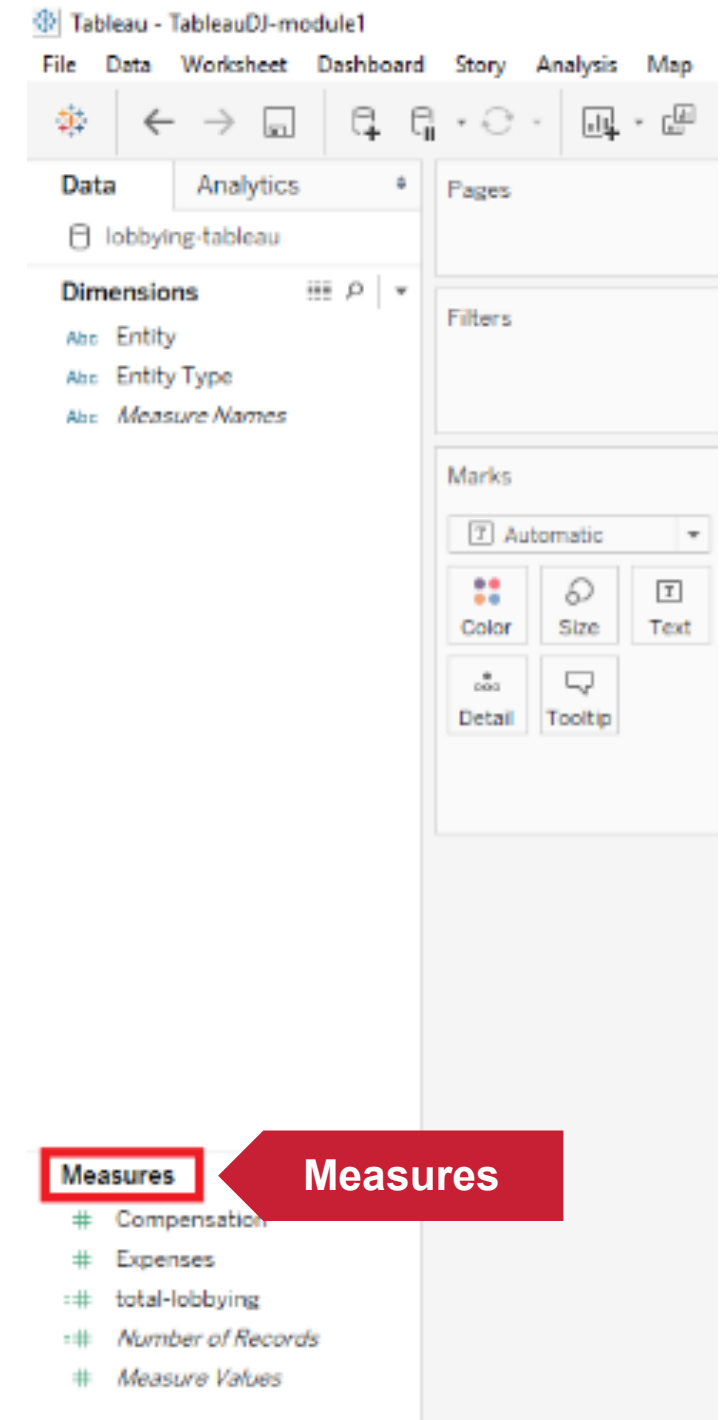


# Who is paying those lobbyists?

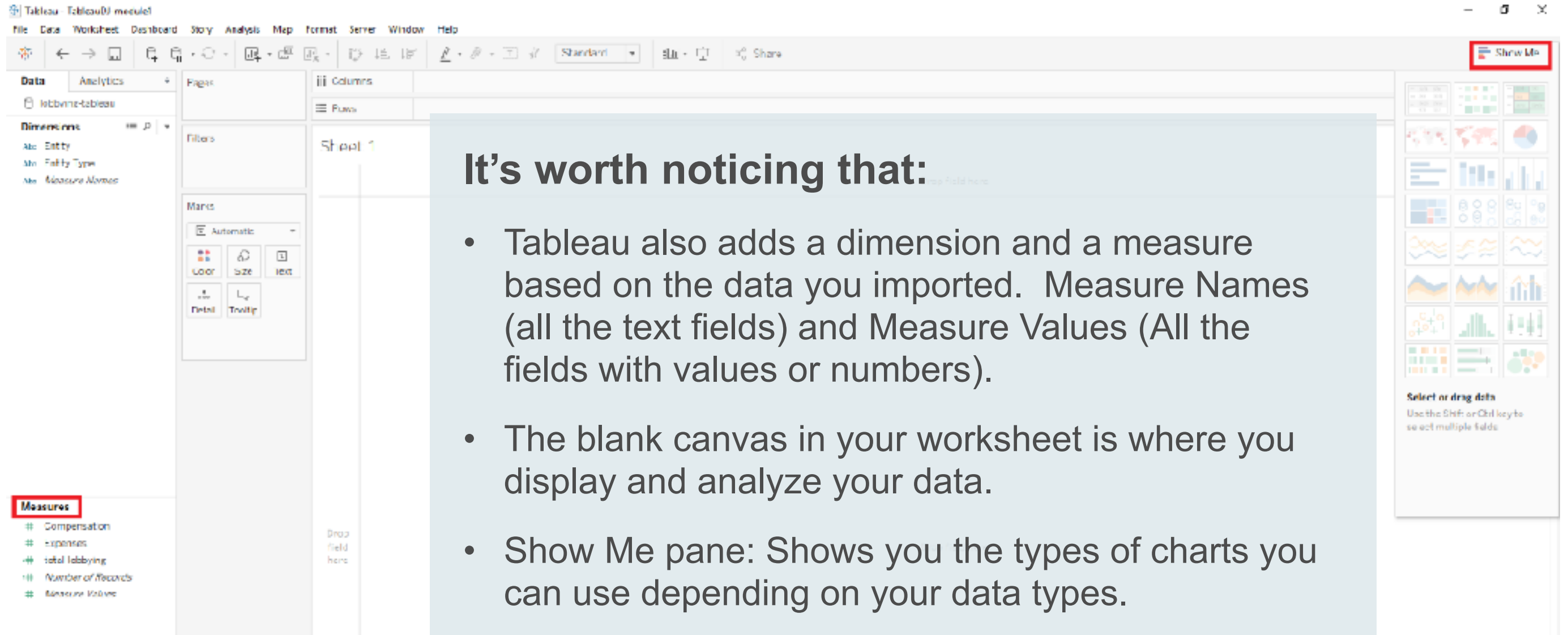
Let's figure out what's going on here.

**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?



The screenshot shows the Tableau Desktop interface. The 'Data' pane on the left lists dimensions: 'Entity', 'Fifty Type', and 'Measure Names'. The 'Measures' pane at the bottom left lists: 'Compensation', 'expenses', 'total lobbying', 'Number of Records', and 'Measure Values'. The 'Columns' shelf is empty. The 'Rows' shelf is empty. The 'Marks' card is set to 'Automatic'. The 'Show Me' pane on the right displays various chart types. The 'Show Me' button in the top right corner is highlighted with a red box. The 'Measures' list at the bottom left is also highlighted with a red box.

**It's worth noticing that:**

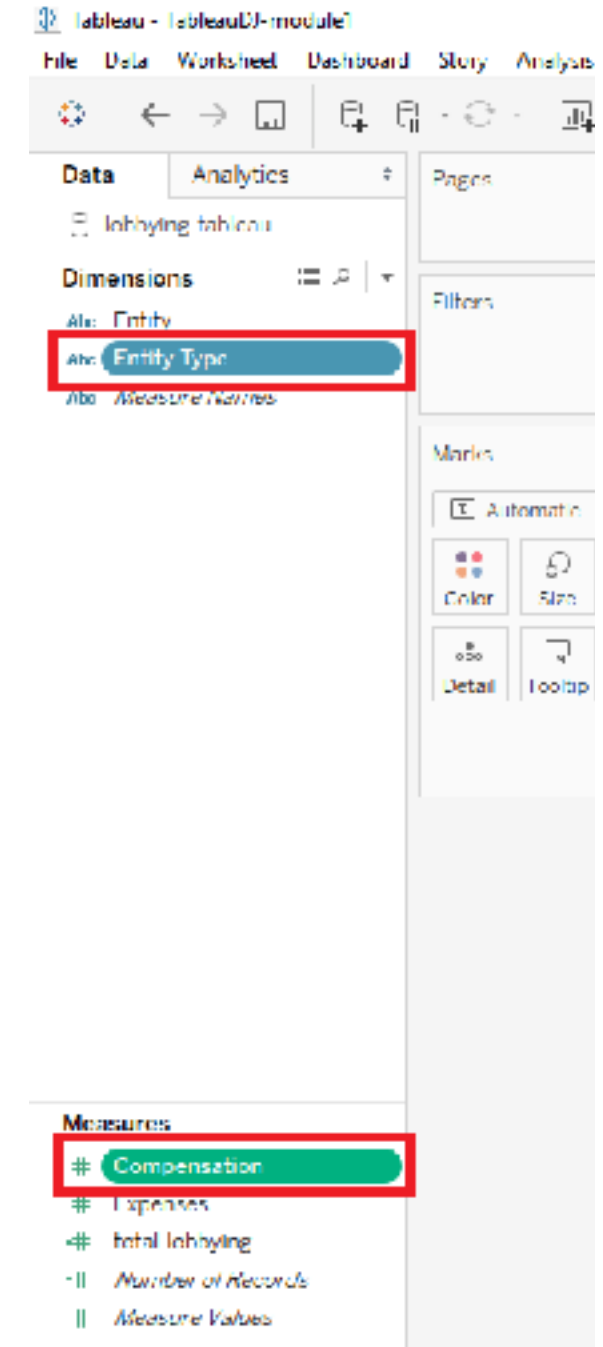
- Tableau also adds a dimension and a measure based on the data you imported. Measure Names (all the text fields) and Measure Values (All the fields with values or numbers).
- The blank canvas in your worksheet is where you display and analyze your data.
- Show Me pane: Shows you the types of charts you can use depending on your data types.

Let's get started!

# Who is paying those lobbyists?

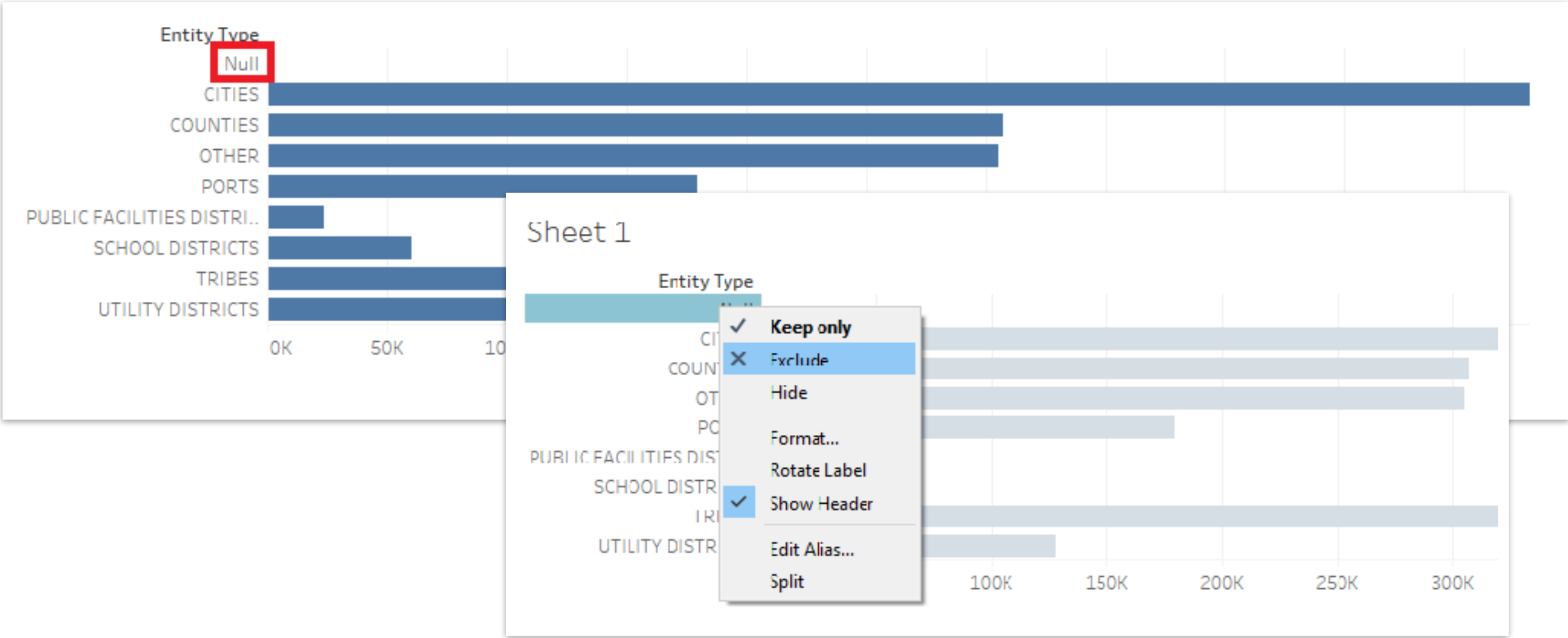
When you hover over one of the Dimensions or Measures, they will be highlighted. I call those pills.

You can click on the pills and drag and drop them to visually explore the data.



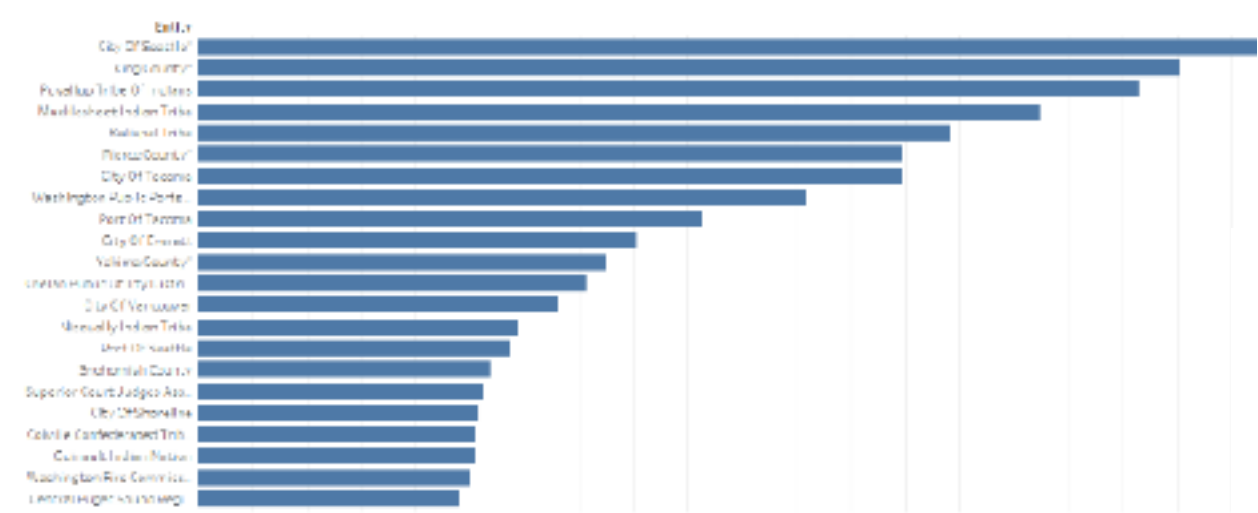


# Who is paying those lobbyists?

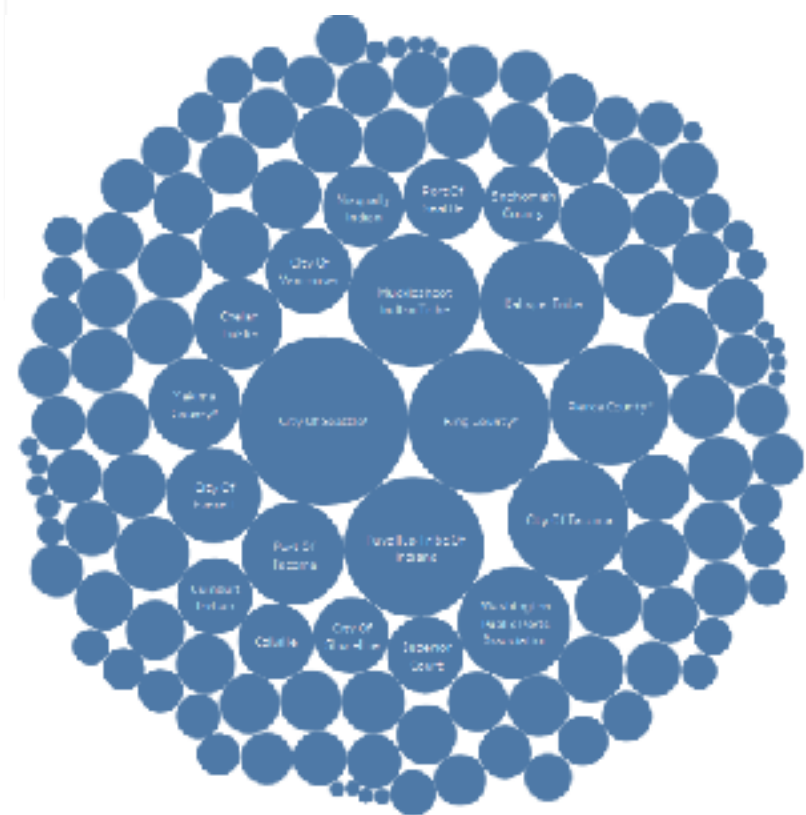


# Who is paying those lobbyists?

Bar Chart



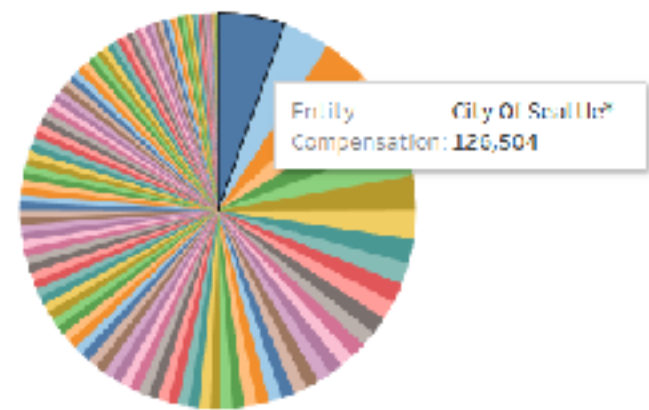
Bubble Chart



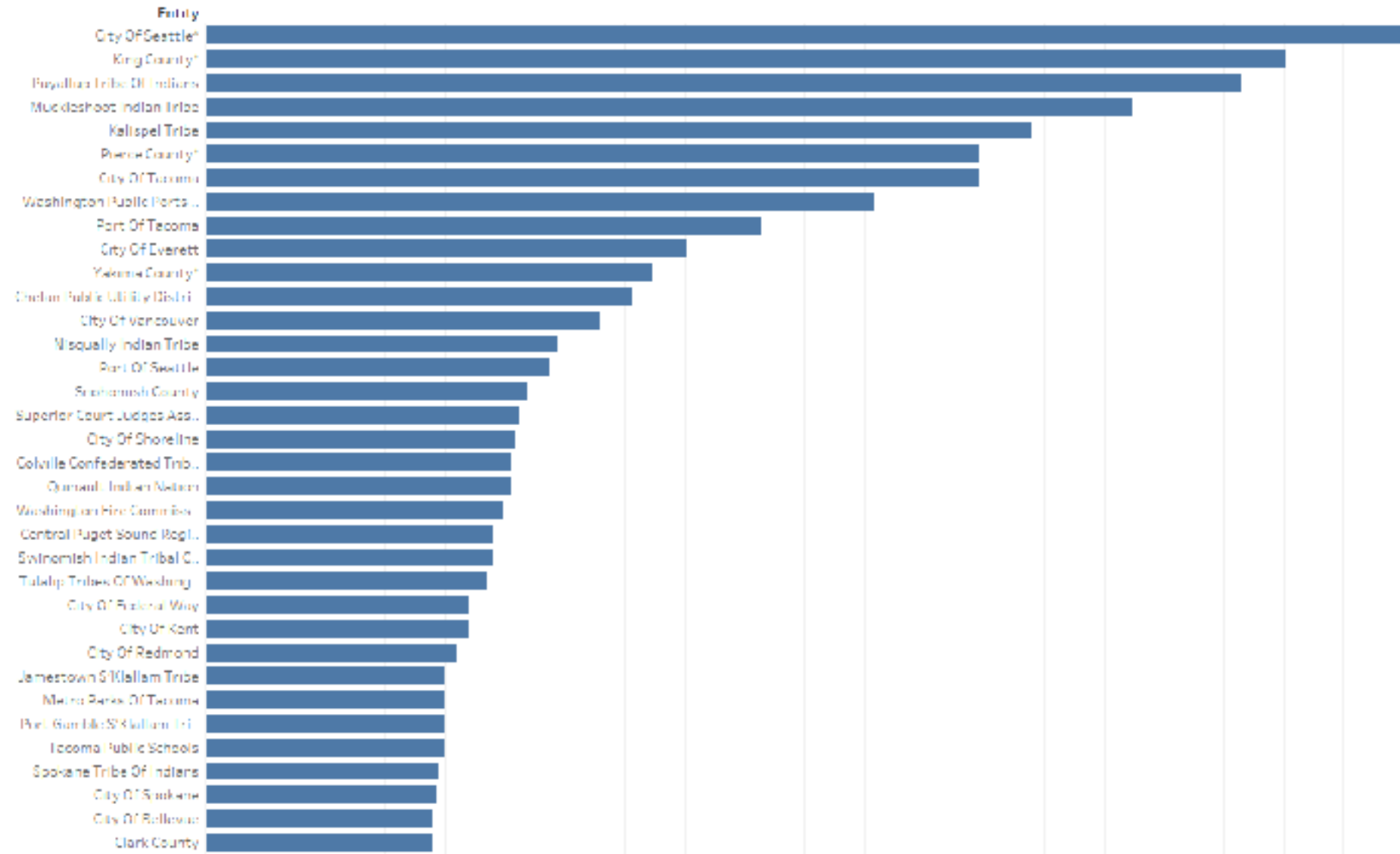
Highlights Chart

Entity	Compensation
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*	64,600
City Of Tacoma	55,882
Washington Public Ports ..	46,425
Port Of Tacoma	40,200
City Of Everett	37,333
Yakima County*	35,650
Chelan Public Utility Distr	33,021
City Of Vancouver	29,333
Nisqually Indian Tribe	28,700
Port Of Seattle	26,033
Snohomish County	26,200
Superior Court Judges As..	25,900
City Of Shoreline	25,500
Colville Confederated Trib	25,500
Quinault Indian Nation	24,907
Washington Fire Commiss	24,000
Central Puget Sound Regi..	23,500
Swinomish Indian Tribal C..	22,000
Tulalip Tribes Of Washing..	
City Of Federal Way	

Pie Chart



# Who is paying those lobbyists?

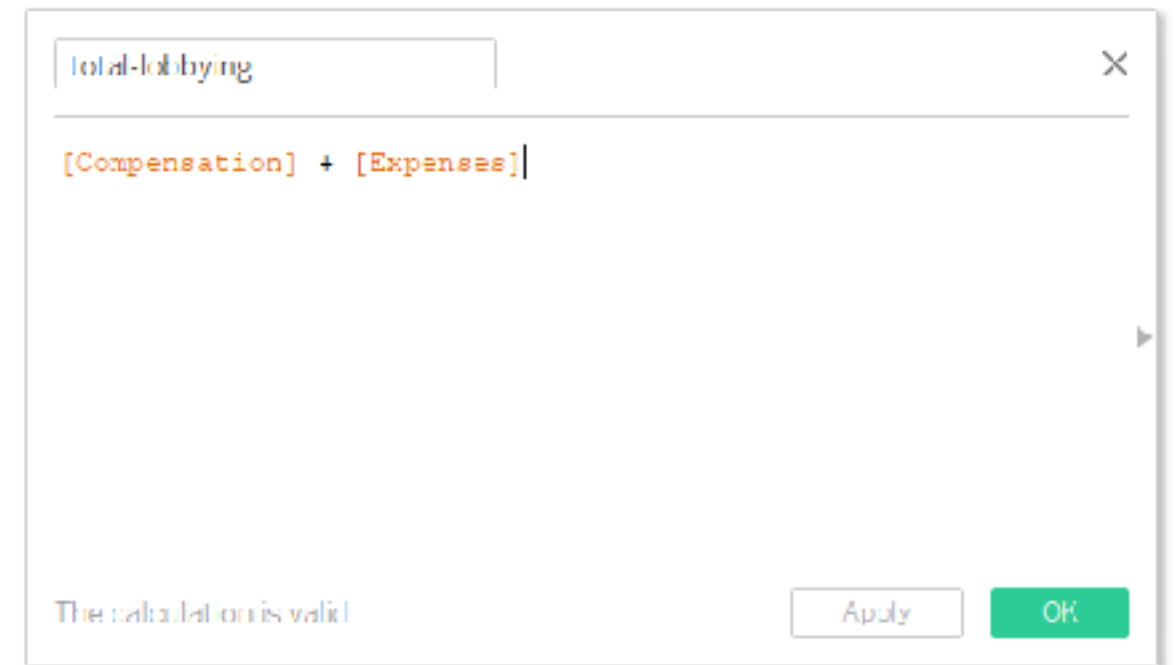


# Who is paying those lobbyists?

Let's do a calculation:

- On the menu bar at the top, select **Analysis** and then **Calculated Field**.
- Give your new calculated field a name: **total-lobbying**.
- 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.*

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



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

Which government spent the most? How much?



# Practice 00

Now you try in pairs with the Lobbyist data.

- open the link (received via email)
- Enter Tableau Online
- Enter your Section
- Create your worksheet
  - 00\_lobbyist\_FirstName\_LastName
- Check if you can do what we described during the Lecture.



# Practice 01

Now you try in pairs with the Titanic data.

- Enter Tableau Online
- Create your worksheet
  - 00\_titanic\_FirstName\_LastName
- How many people were on the ship?
- What was the total of all fares paid?

Think of other questions you might want to answer.

