FUNDAMENTALS OF TABLEAU

LEARNING OBJECTIVES

After this lesson, you will be able to:

1. Connect and prepare data to import into Tableau.
2. Navigate the Tableau interface.
3. Build graphs, calculations, dashboards, and stories.
4. Learn to discriminate between discrete and continuous dates in Tableau.
5. Apply your new skills to sample data.

BEFORE WE BEGIN

Please download the following files for today's lesson:

Data files:

- ❏ Divvy_Trips.zip
- ❏ Divvy_Bicycle_Stations.csv
- ❏ Move both files to **Documents > My Tableau Repository > Datasources > 2020.3 > en_GB-EU**
- ❏ Right click on Divvy_Trips.zip and select "Extract here"

FUNDAMENTALS OF TABLEAU

INTRODUCTION

Data Rock-Star

INTRODUCTION TO TABLEAU

What is Tableau?

- Tableau is a visual analytics and data discovery tool launched as the result of a Stanford research collaboration between data and visualization labs. It's based on a patent called "Polaris" that covers interactive visual data querying.

- There are a number of players in the drag-and-drop/visual data space. Power BI from Microsoft will hook into your Azure data set, and Amazon QuickSight will hook into your Redshift data. There's also Looker, Periscope Data, Qlik, and many other competitors.

- Making a decision about which to use will be a trade off between cost, ease of use, features offered, and data connections.

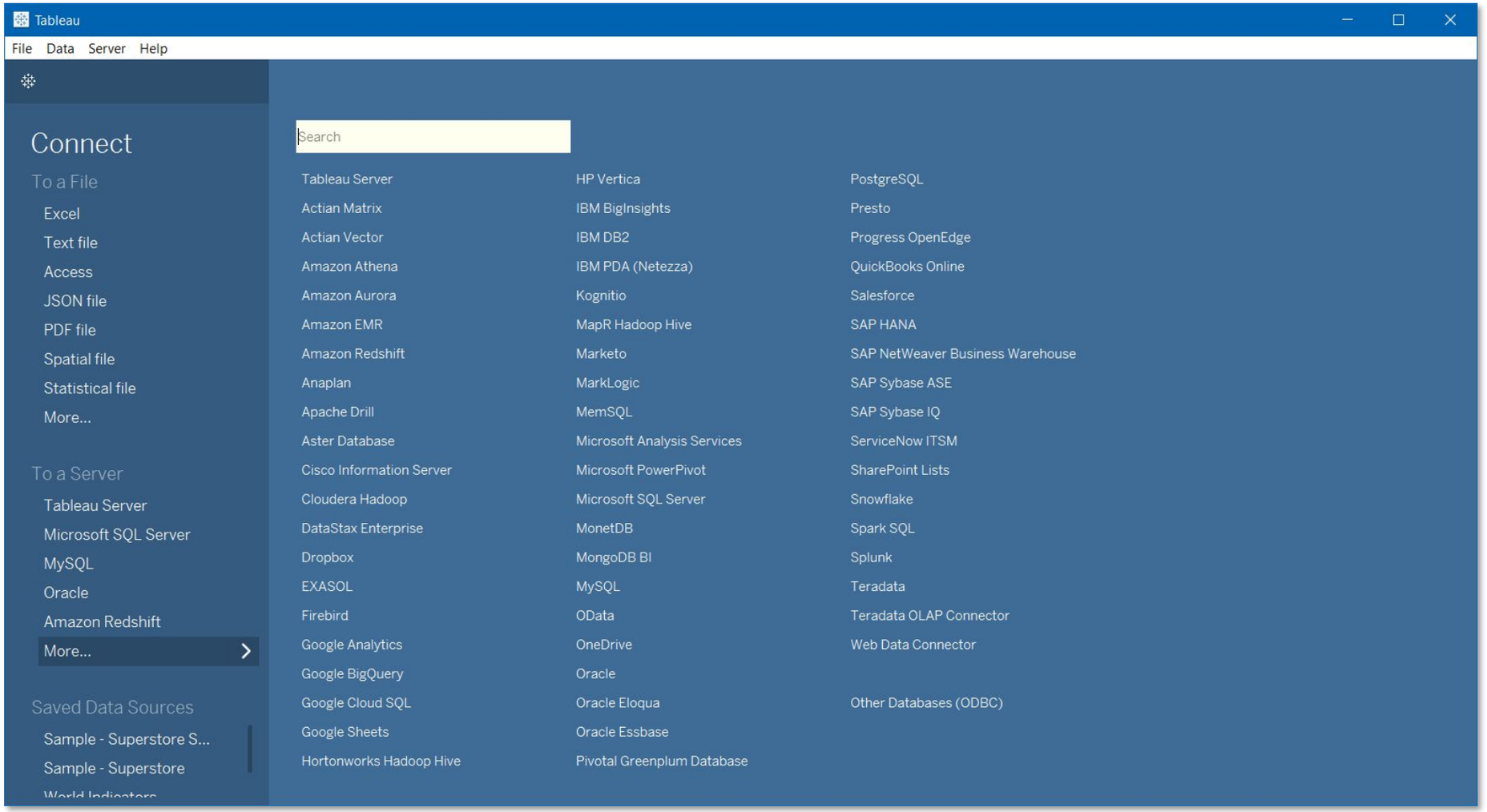
INTRODUCTION TO TABLEAU

What is Tableau?

- Tableau seamlessly creates background SQL queries called “VizQL” as you interactively click, drag, and drop data elements to explore and interrogate your data.

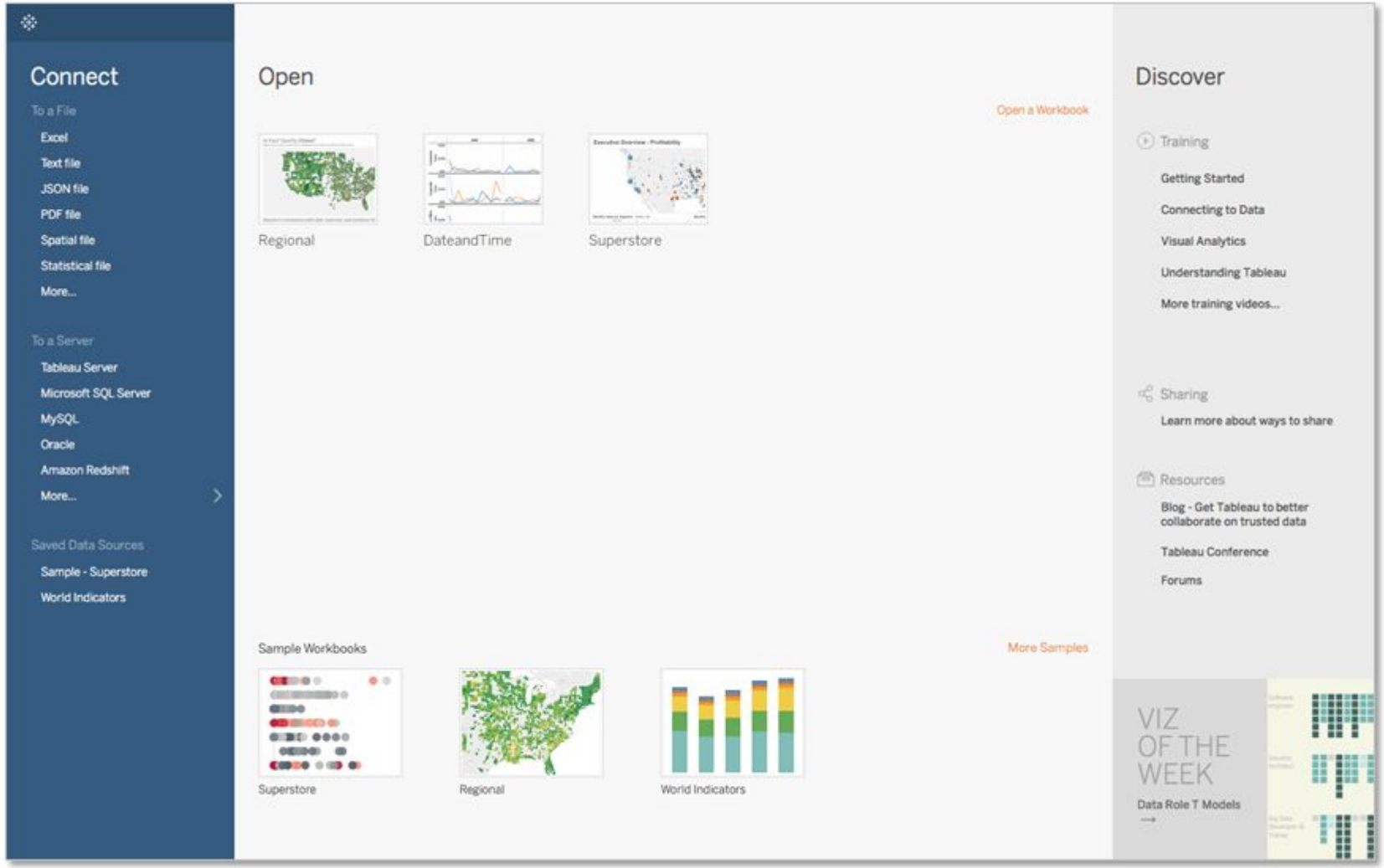
- While Tableau is designed to be used by a range of business users and non-technical audiences, it also provides access to a deep computational ability for advanced data analytics.

DATA DEMOCRACY: TABLEAU'S 60+ BUILT-IN CONNECTORS



INTRODUCTION TO TABLEAU: INTERFACE OVERVIEW

Click on the video to start. Runtime 4:57



Source: https://goo.gl/x1npZa

TABLEAU EXAMPLES

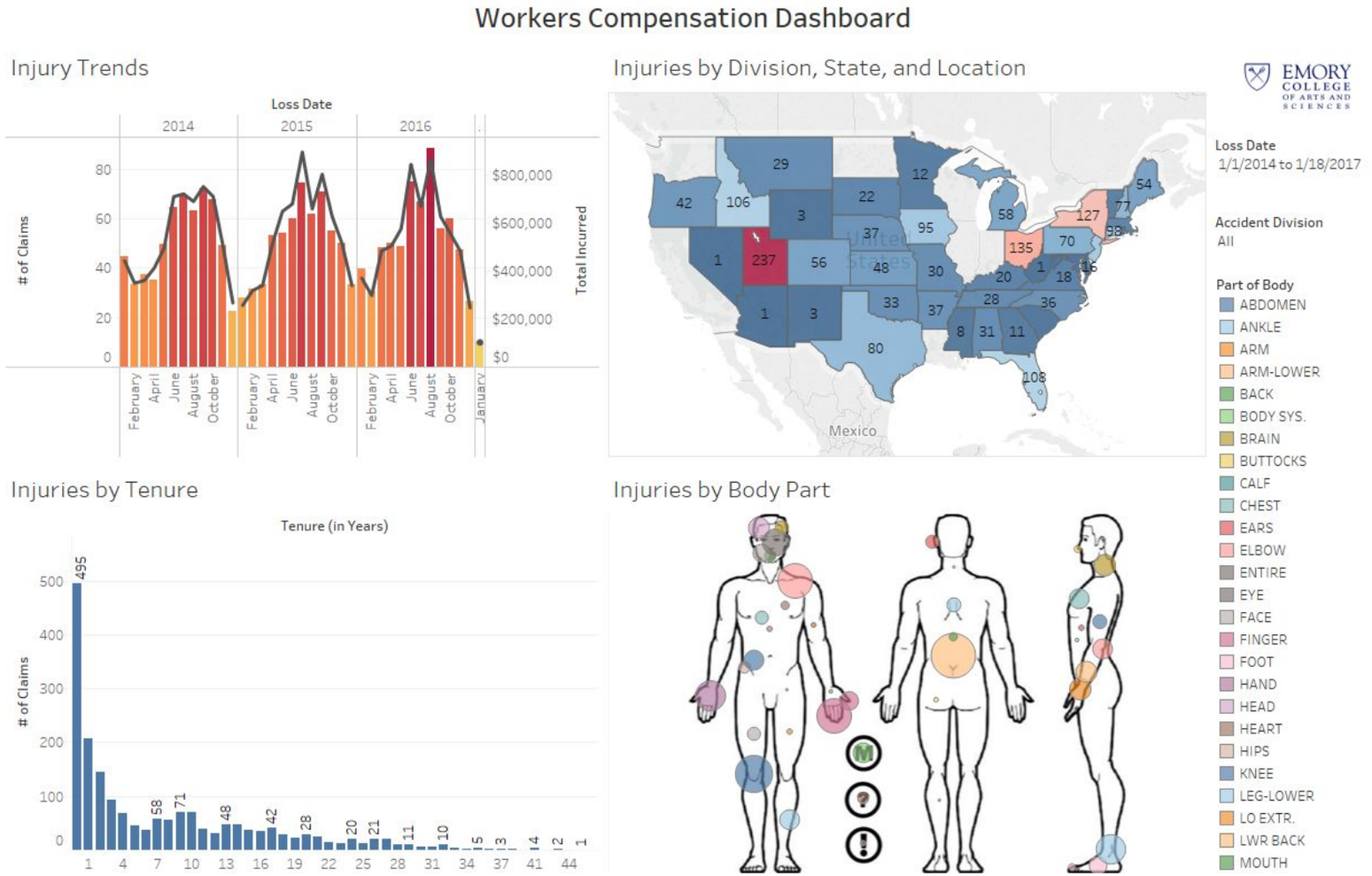


TABLEAU EXAMPLES

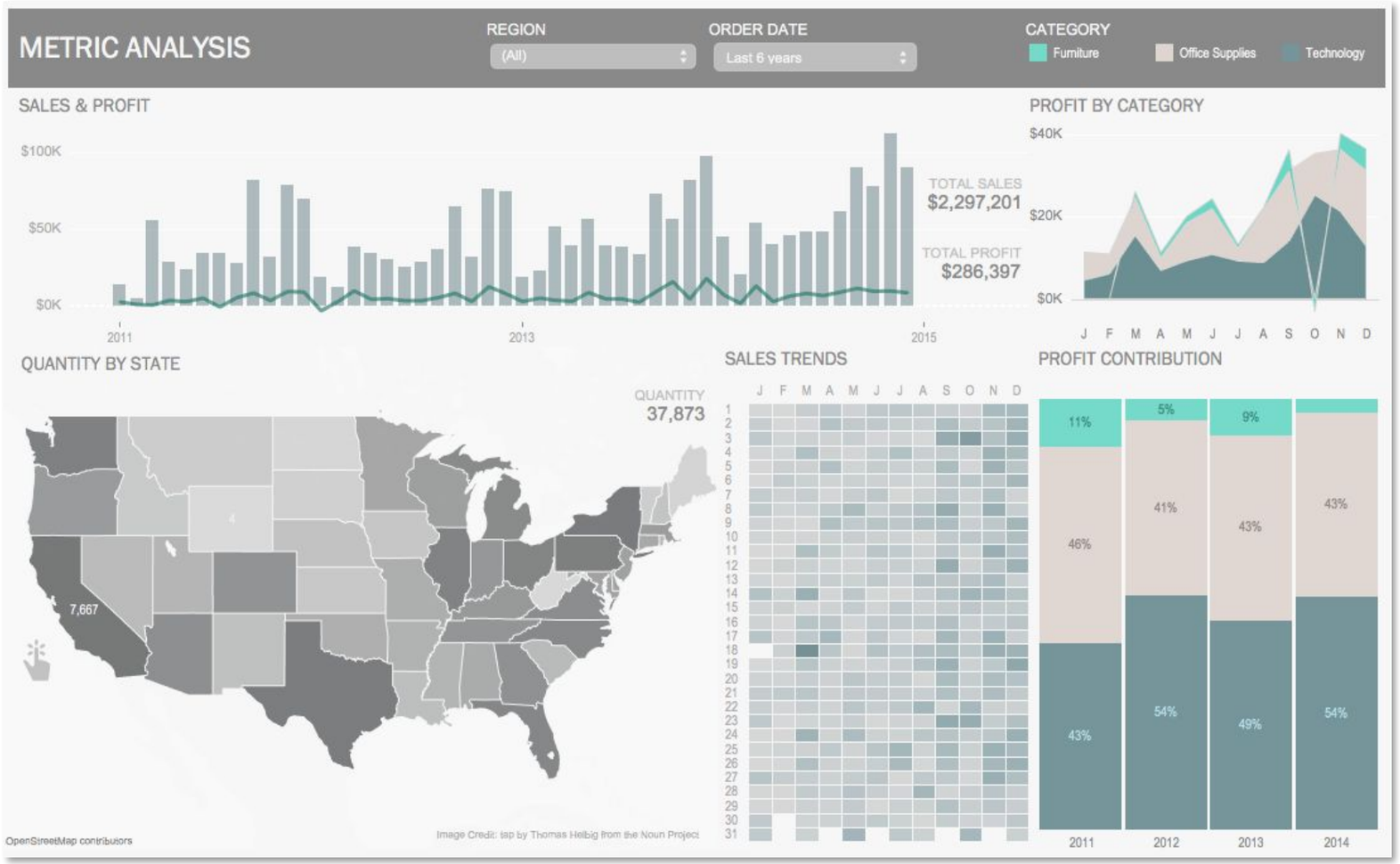


TABLEAU EXAMPLES

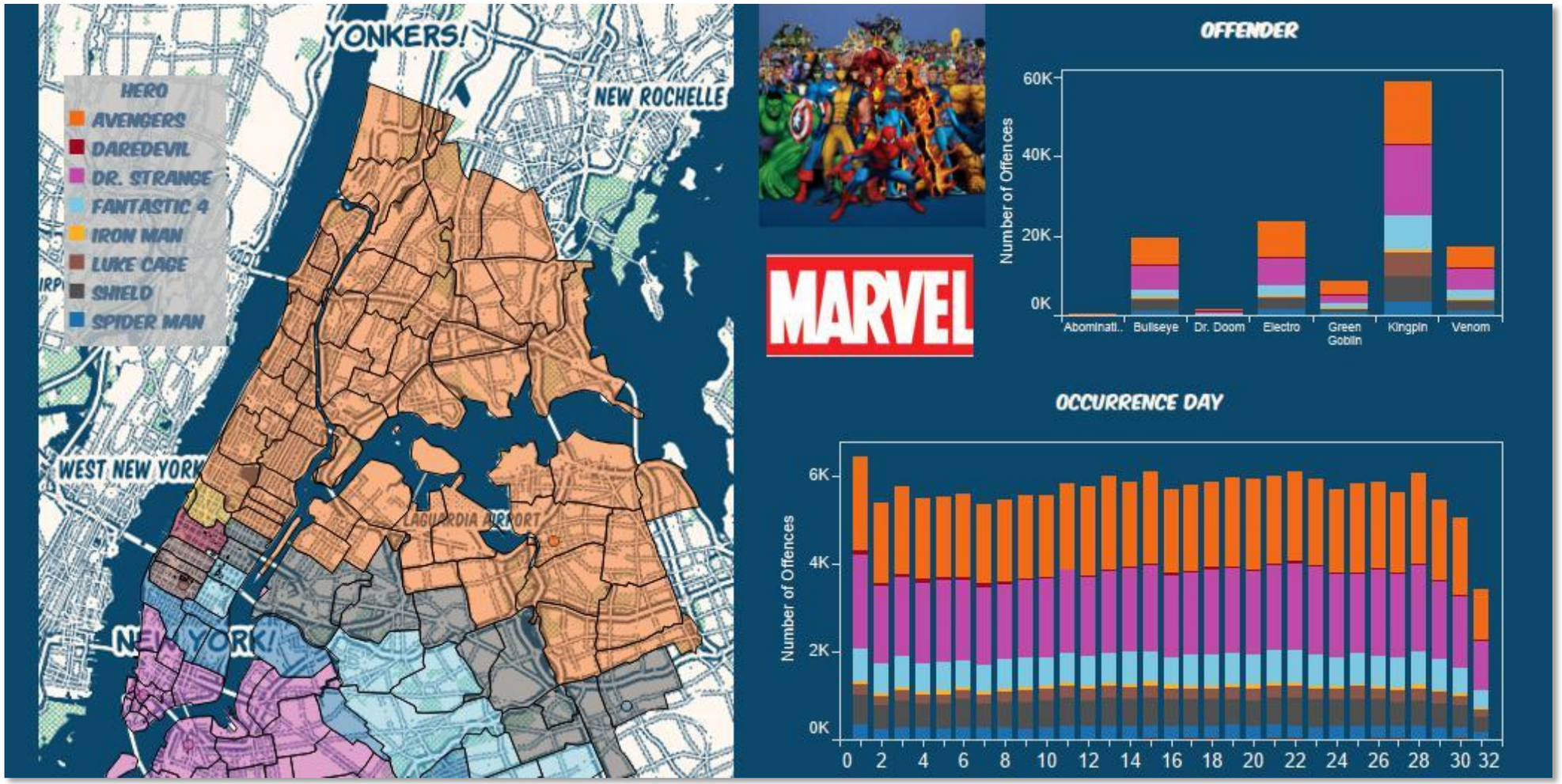
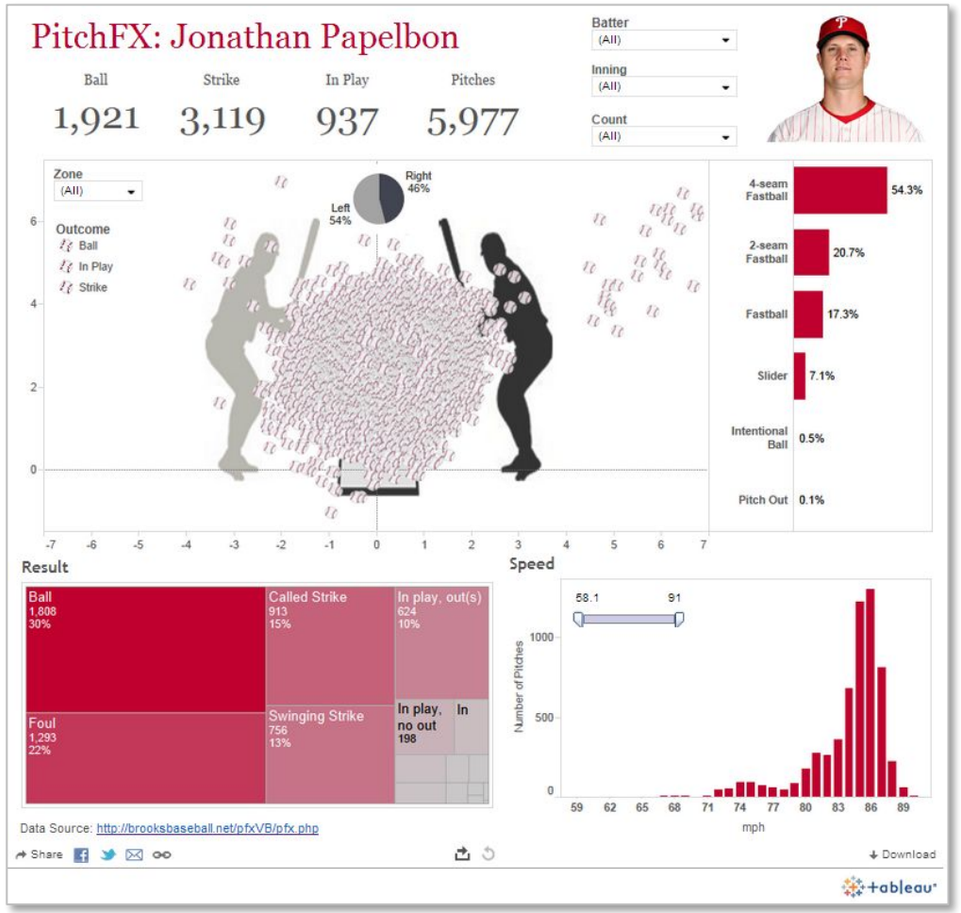


TABLEAU EXAMPLES



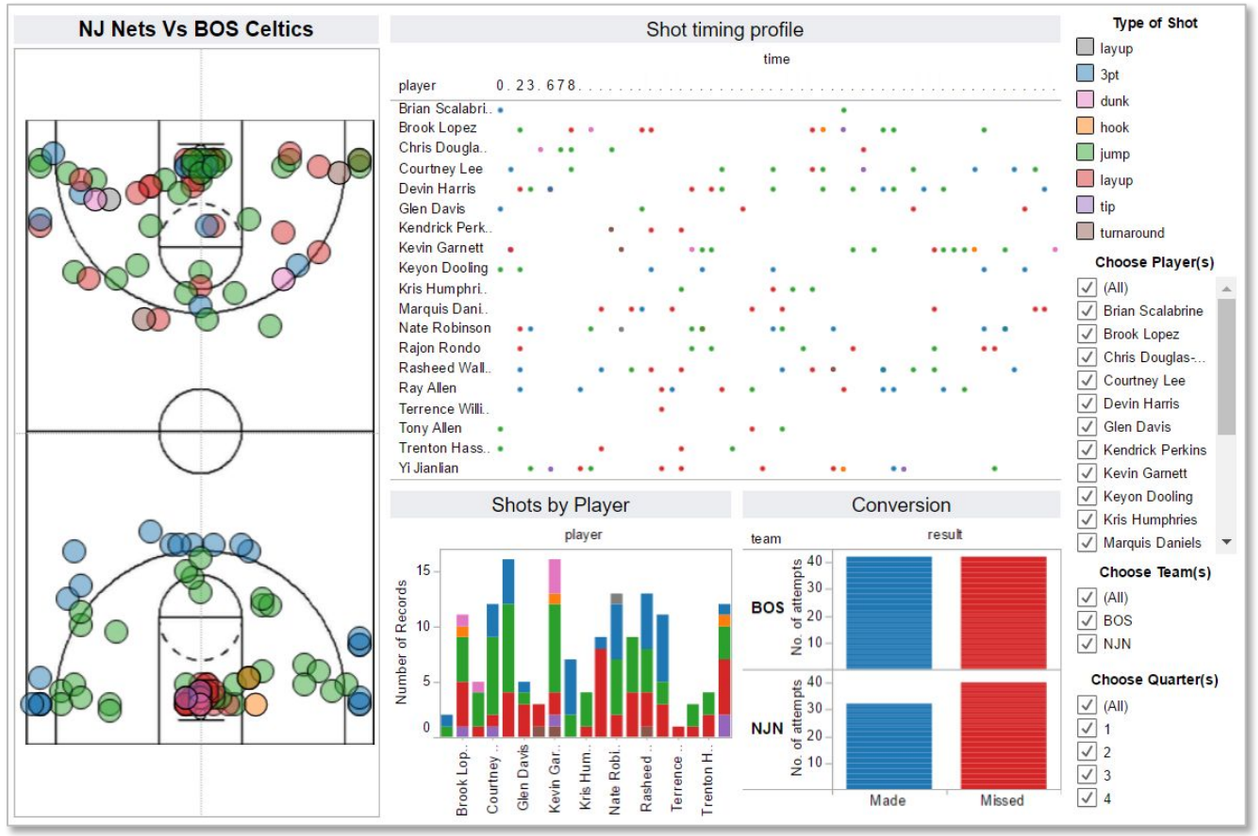


TABLEAU EXAMPLES

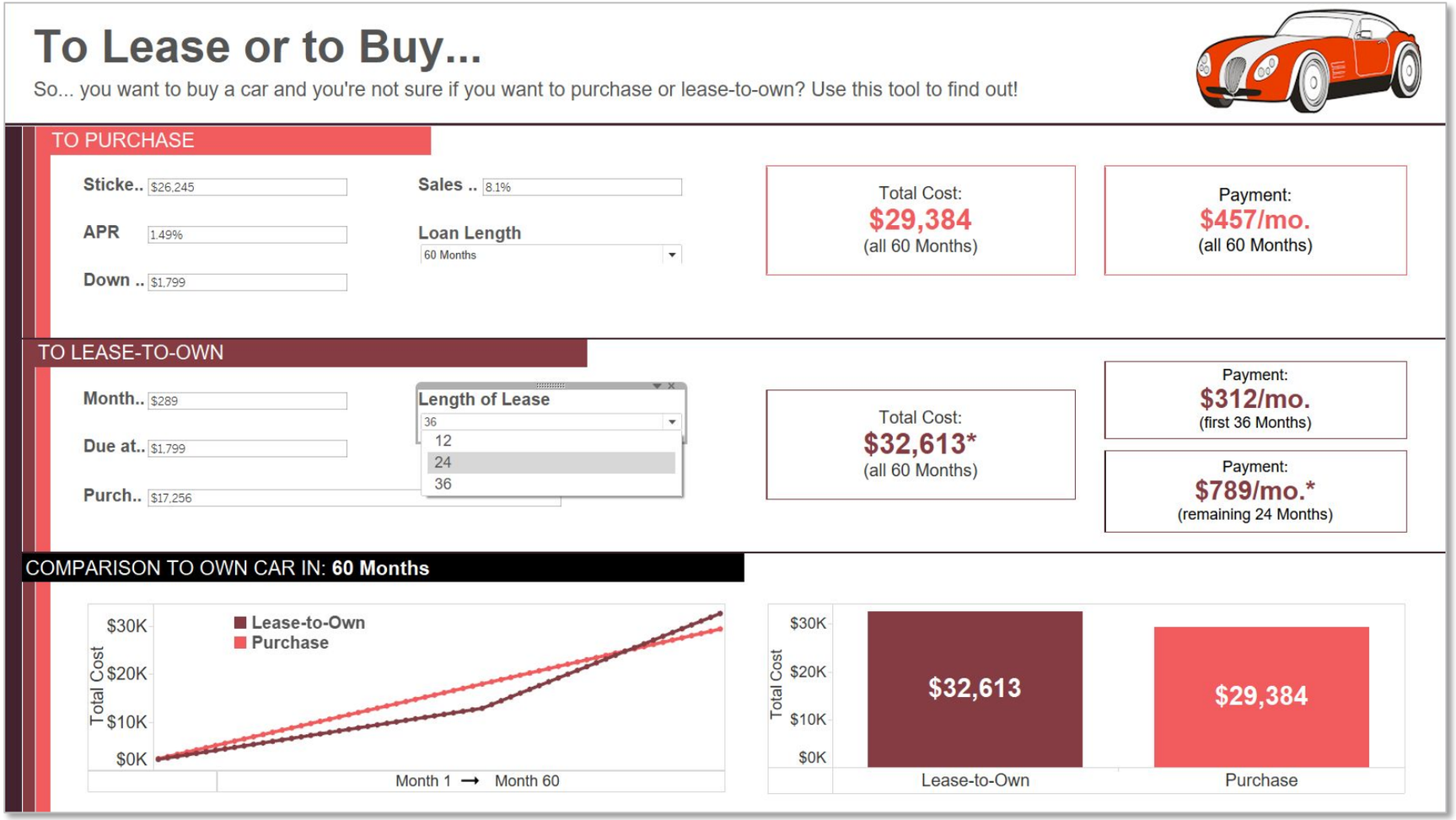


TABLEAU EXAMPLES

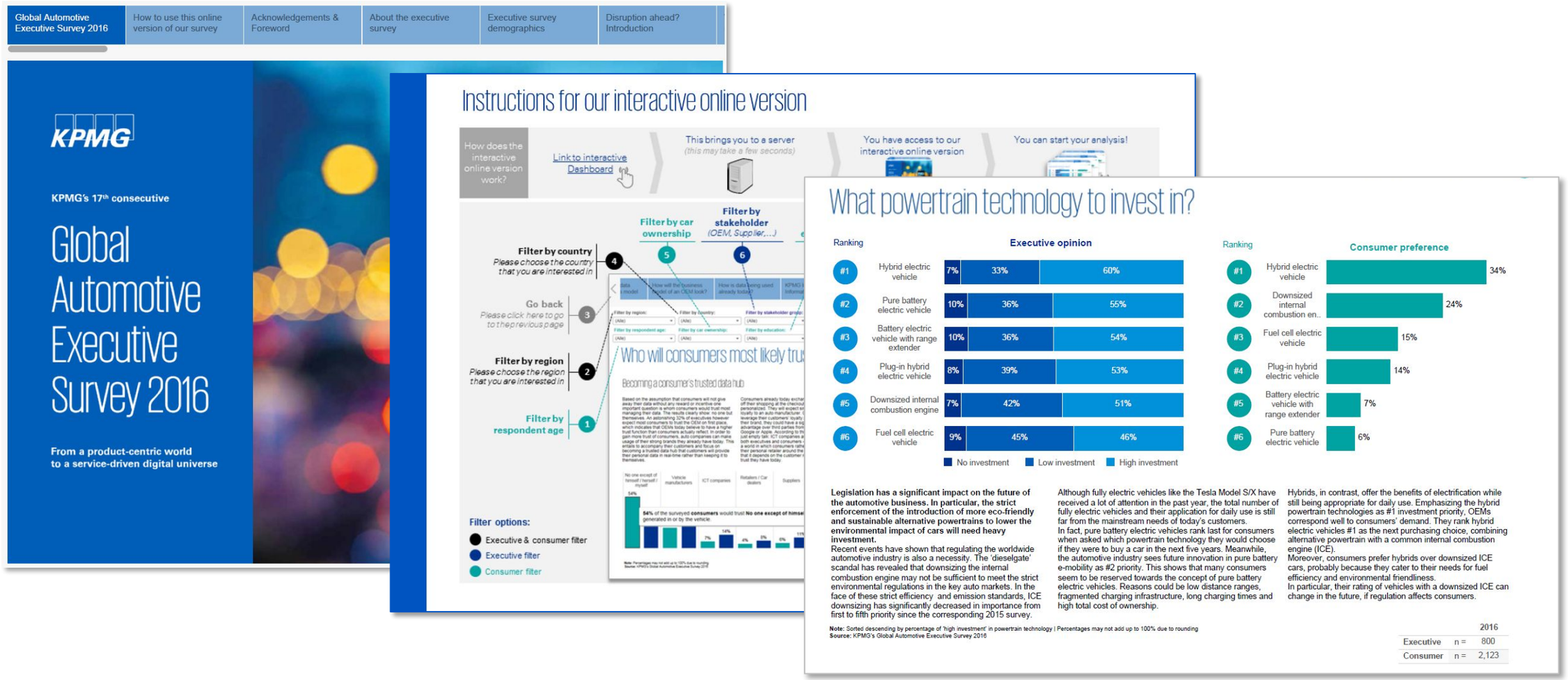
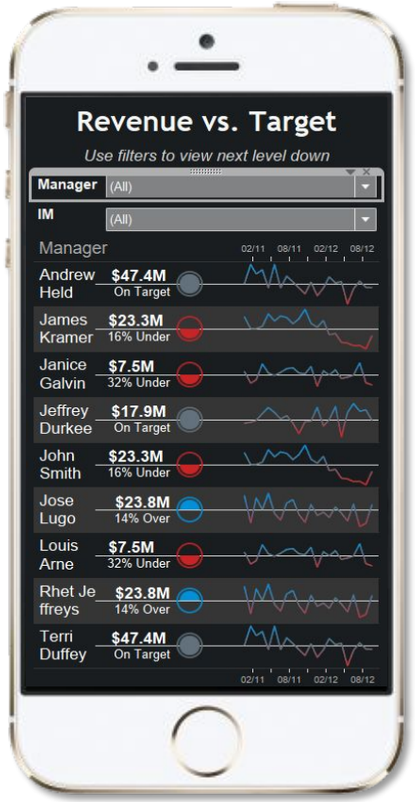


TABLEAU EXAMPLES

interworks



GUIDED PRACTICE: CONNECT TO DATA

CONNECT TO DATA AND NAVIGATE

1. For this lesson, we will be using the Divvy bike share data.
2. We'll work through the data together, asking questions as we go.
3. For context, in this data set:
    a. Every row of data is a trip.
    b. Every trip has a start and stop time.
    c. Every trip has a "from" location and "to" location.

Divvy Data

Metadata for **Divvy Trips Tables**:

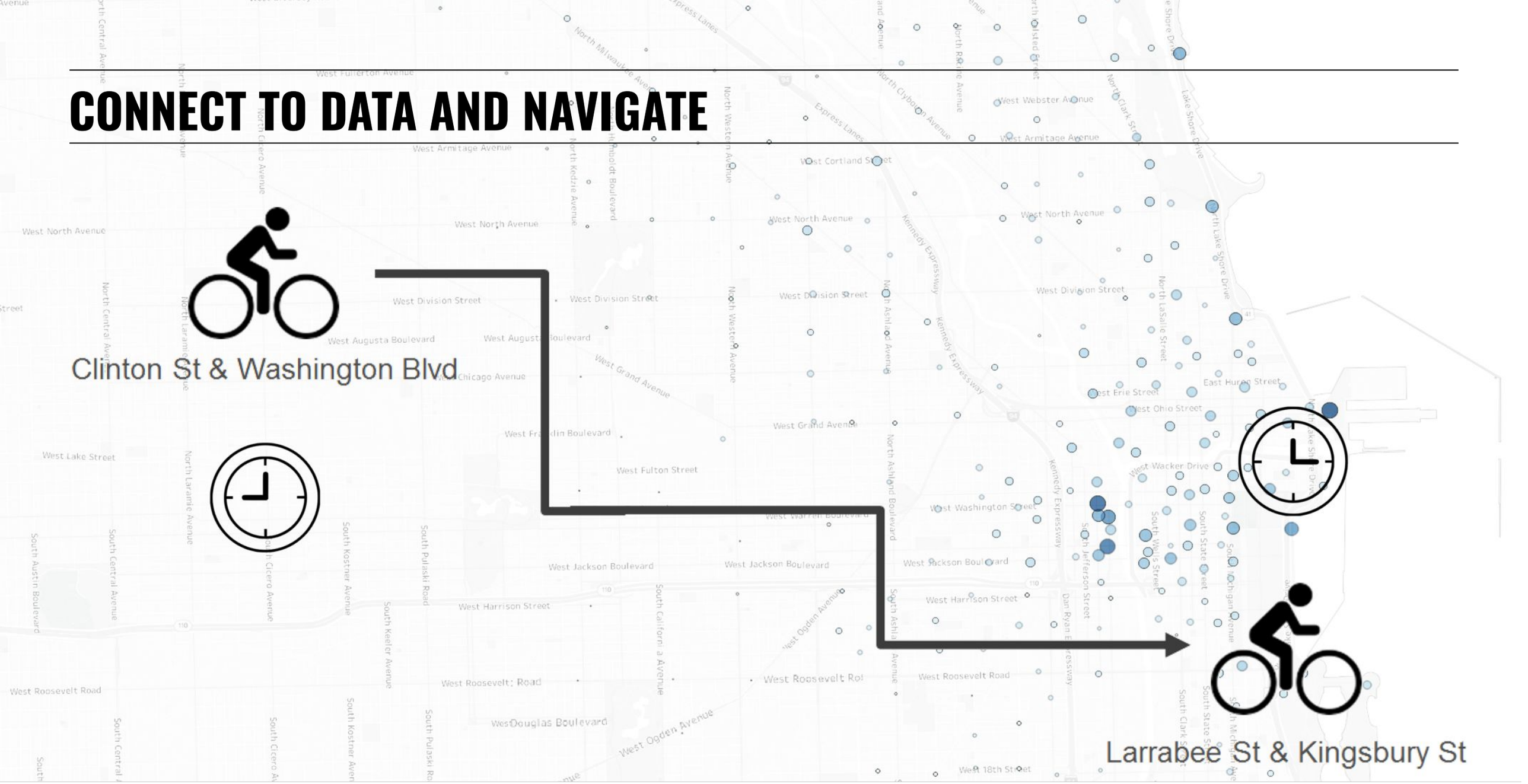
Variables:
trip_id: ID attached to each trip taken
starttime: day and time trip started, in CST
stoptime: day and time trip ended, in CST
bikeid: ID attached to each bike
tripduration: time of trip in seconds
from_station_name: name of station where trip originated
to_station_name: name of station where trip terminated
from_station_id: ID of station where trip originated
to_station_id: ID of station where trip terminated
usertype: "Customer" is a rider who purchased a 24-Hour Pass; "Subscriber" is a rider who purchased an Annual Membership; "Dependent" is a family member related to a Subscriber;
gender: gender of rider
birthyear: birth year of rider

Notes:
- First row contains column names
- Trips that did not include a start or end date were removed from original table
- Gender and birthday are *only* available for Subscribers (or their Dependents)

Metadata for **Divvy Bicycle Stations Table**:

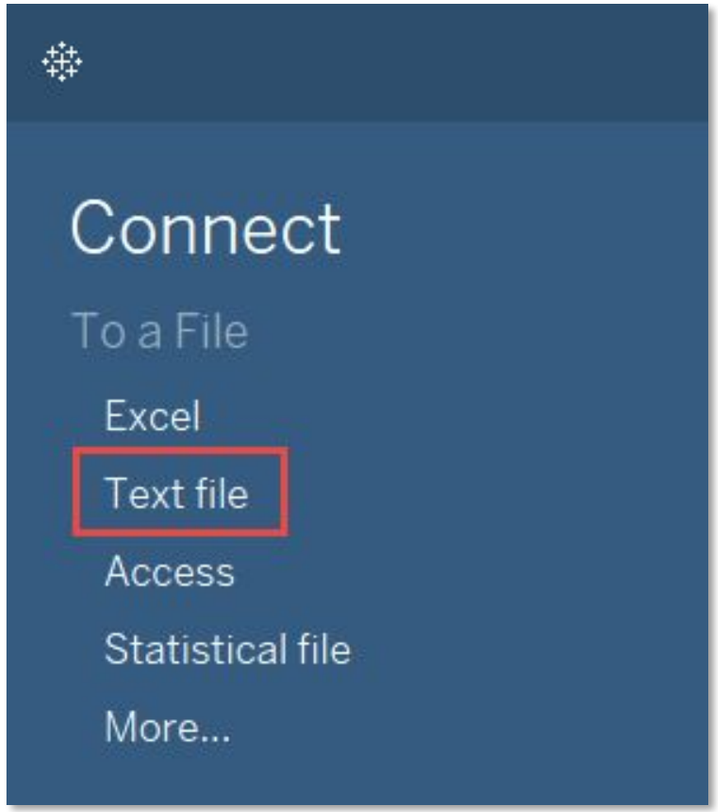
Variables:
ID: ID attached to each station
Station Name: station name
Total Docks: number of total docks at each station
Docks in Service: number of total active docks at each station
Status: In Service or Not In Service
Latitude: station latitude
Longitude: station longitude
Location: station coordinates

CONNECT TO DATA AND NAVIGATE



CONNECT TO DATA AND NAVIGATE

1. Let's get started by connecting to the data set.



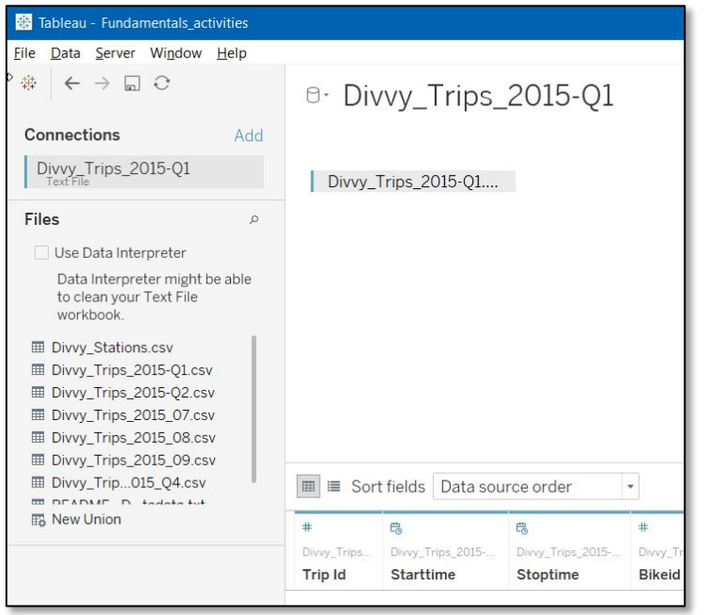
CONNECT TO DIVVY BIKE DATA

2. Navigate to the **Documents > My Tableau Repository > Datasources > 2020.3 > en_GB-EU**, where you should have already copied the Divvy Trips data files (*if not, please do so now*).

3. Locate and double click the .csv file: **Divvy_Trips_2018-Q1**.

4. Tableau will load the file into the data pane, and the related data files will show up to the left.

5. We will perform a drag-and-drop **`UNION`**, as Divvy Trips' tables are separate files.

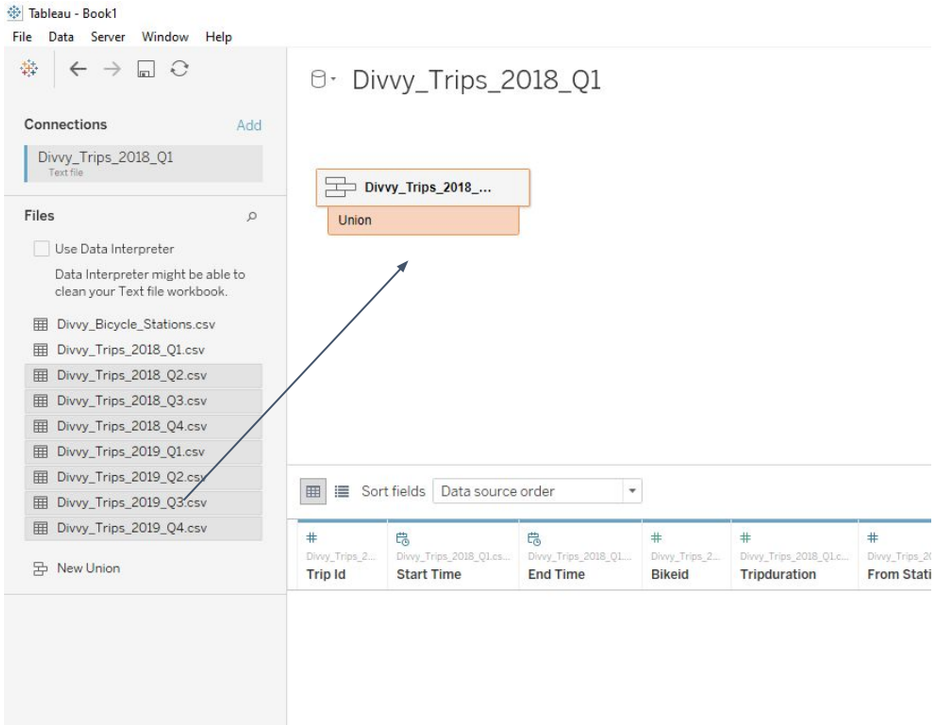


CONNECT TO DIVVY BIKE DATA: UNION THE DATA FILES

6. Recall from the SQL section that files may be combined by a **UNION**. The Divvy Trips source data comes in months and quarters, which we want to combine into one data table.

7. To do so, simply shift-select the remaining five Divvy_Trips data files, then drag and release them on the bottom edge of Divvy_Trips_2018-Q1.

8. The data pane will now reflect Divvy_Trips as a **UNION**.

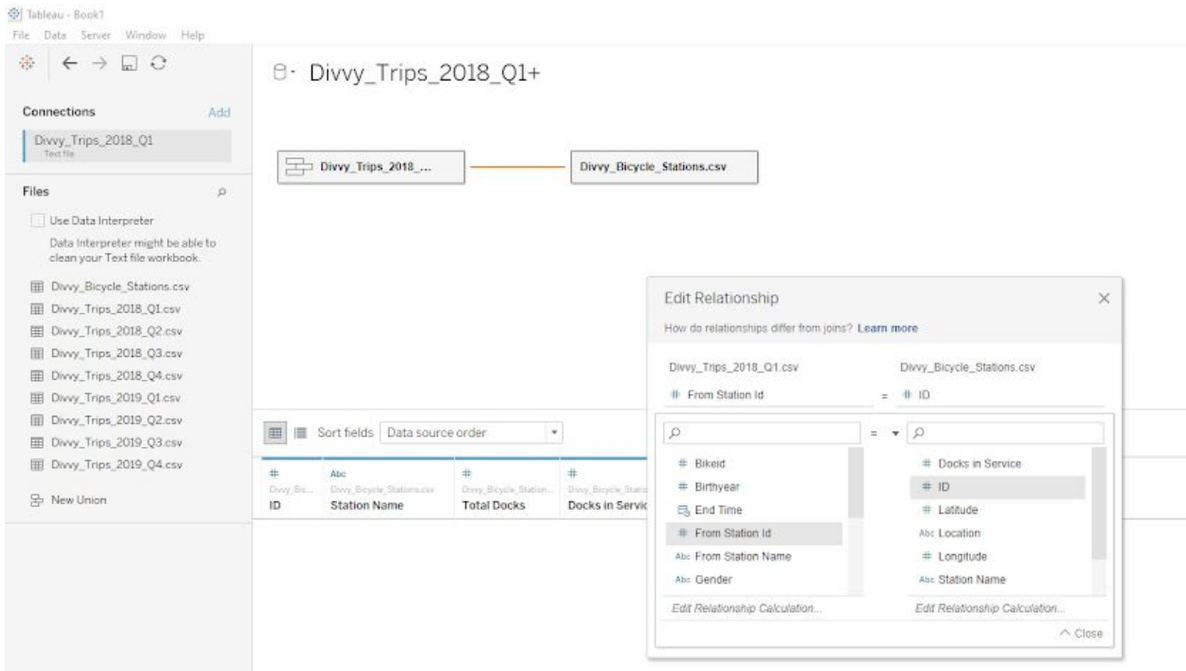


CONNECT TO DIVVY BIKE DATA: JOIN THE DATA FILES

6. Now we will perform a drag-and-drop **JOIN:** select Divvy_Bicycle_Stations.csv and drop besides Divvy_Trips_2018_…

7. Tableau will prompt a “Edit Relationship” window;

8. Select *From Station ID* on the left side and *ID* on the right side



CONNECT TO DIVVY BIKE DATA: CREATE A TABLEAU EXTRACT

9. Data may be accessed in Tableau either **live** or as an **extract**.
   a. Live connections are best for relatively small or very fast data sources.
   b. Data extracts bring the live data into a Tableau columnar database, which is highly optimized for speed and visual data interaction.

10. Click on the Extract button under "Connection."
    a. Now select the "Sheet 1" worksheet tab.
    b. Tableau will take up to two minutes to build the extract.
    c. Saving Tableau as a "Packaged Workbook" will include data from the extract.

CONNECT TO DIVVY BIKE DATA: ALTERNATIVE ROUTE

- If you have any issues downloading the datasets or loading it to Tableau, download the Tableau Packaged Workbook which includes the pre-loaded dataset.

- Fundamentals_Packaged_Workbook.twbx

GUIDED PRACTICE: BUILDING DIVVY BIKES VISUALIZATIONS

CONNECT TO DATA AND NAVIGATE

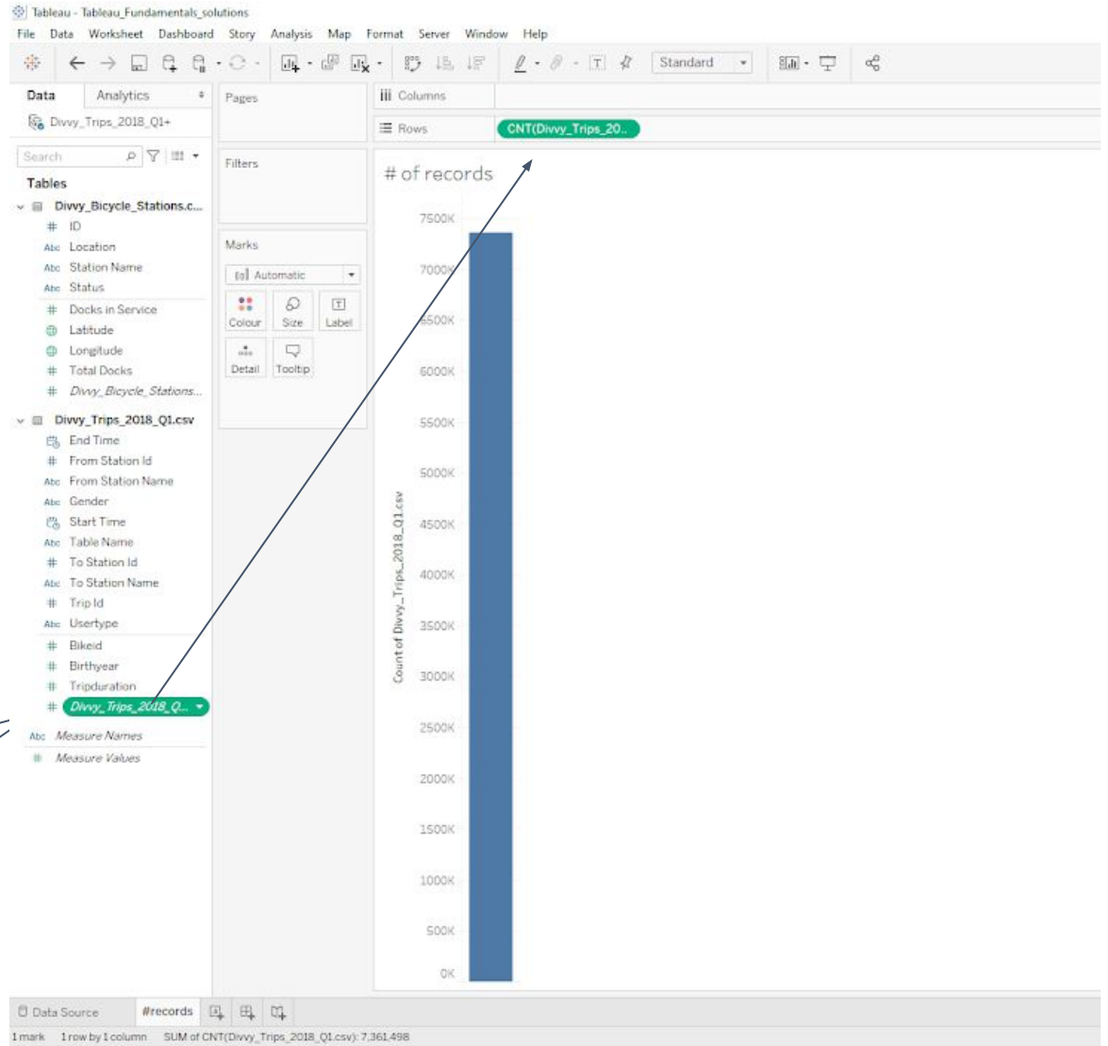
Let's build a few visualizations to explore the Divvy Bike tables.

Here are some questions we'll use visualizations to answer:

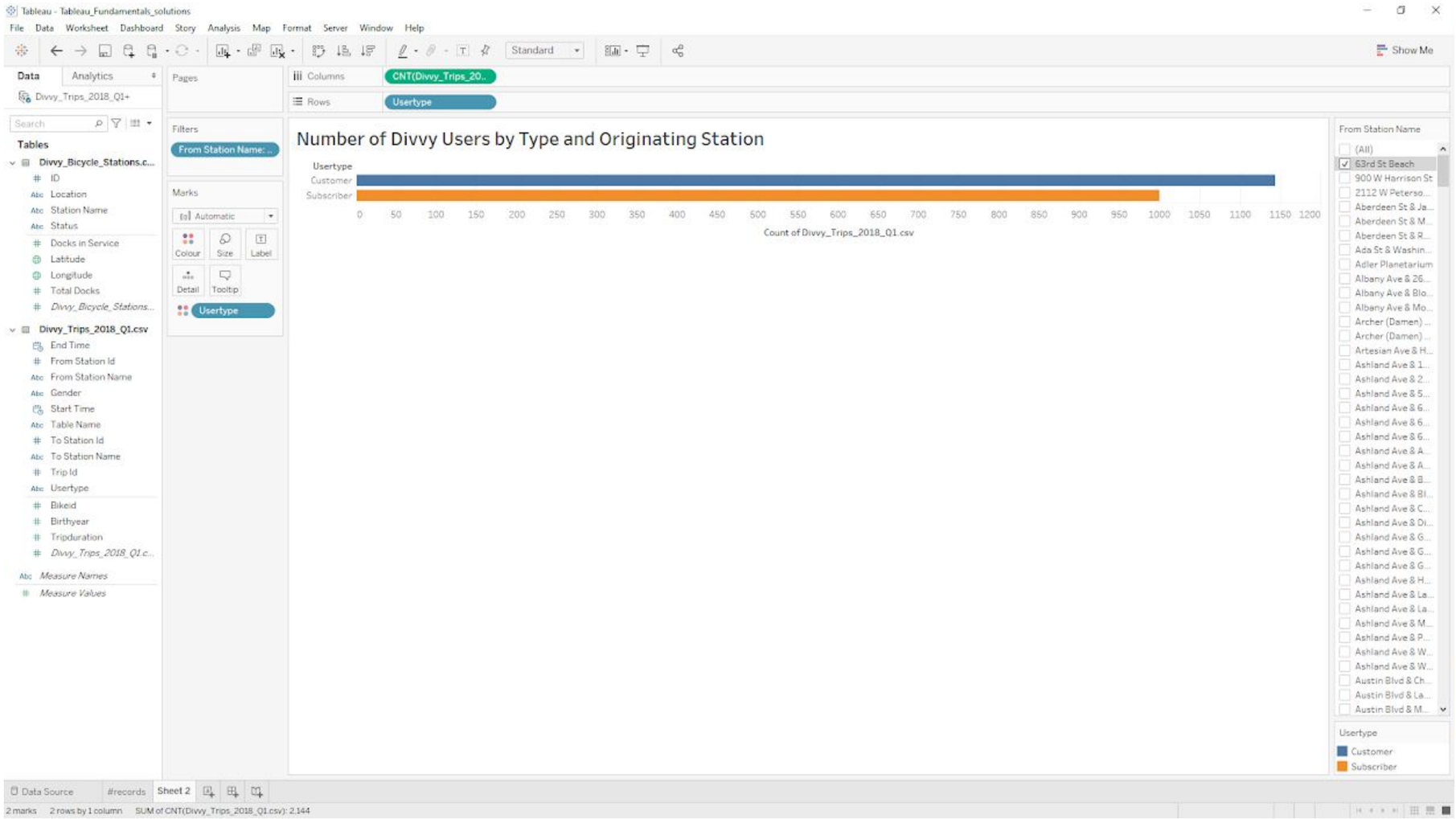
- The number of records in the dataset
- The number of Divvy users by type and originating station [stacked bar].
- The bikes used for less than 100 trips (by Bikeid) [stacked bar chart].
- The age of Divvy Bike users (year of birth) [using a histogram].
- The average trip times for each Divvy Bike location.

of Records

some older versions of Tableau might name this variable as Number of Records

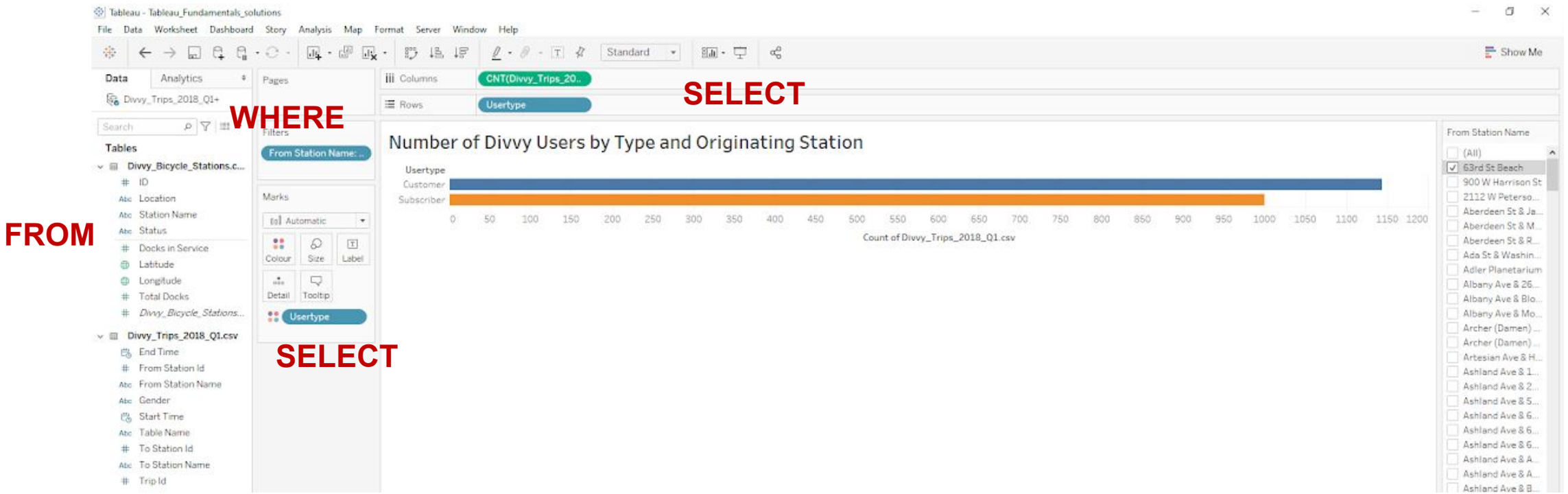


Number of Divvy users by type and originating station [stacked bar]



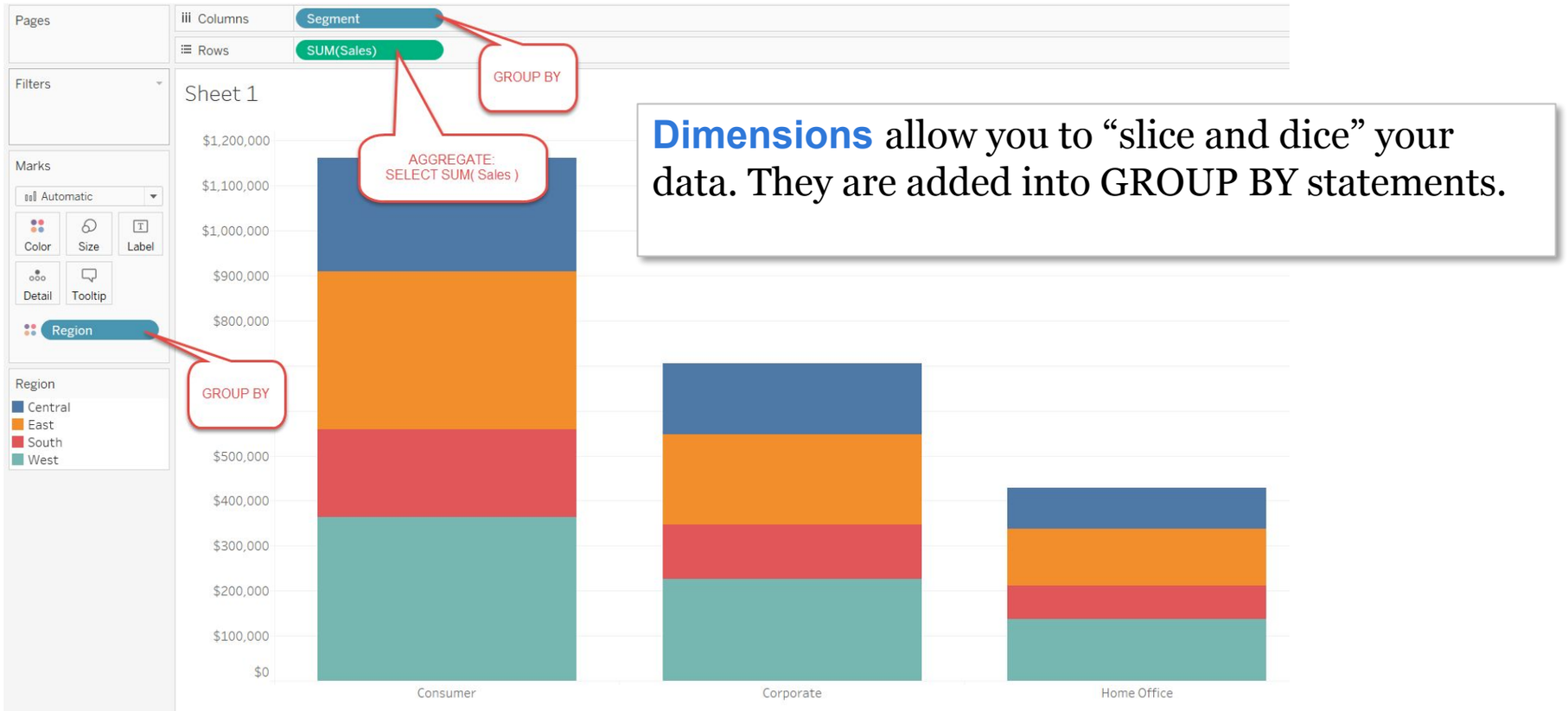
VizQL → SQL ANALOGS

Tableau internally generates VizQL based on where data are placed on the view.



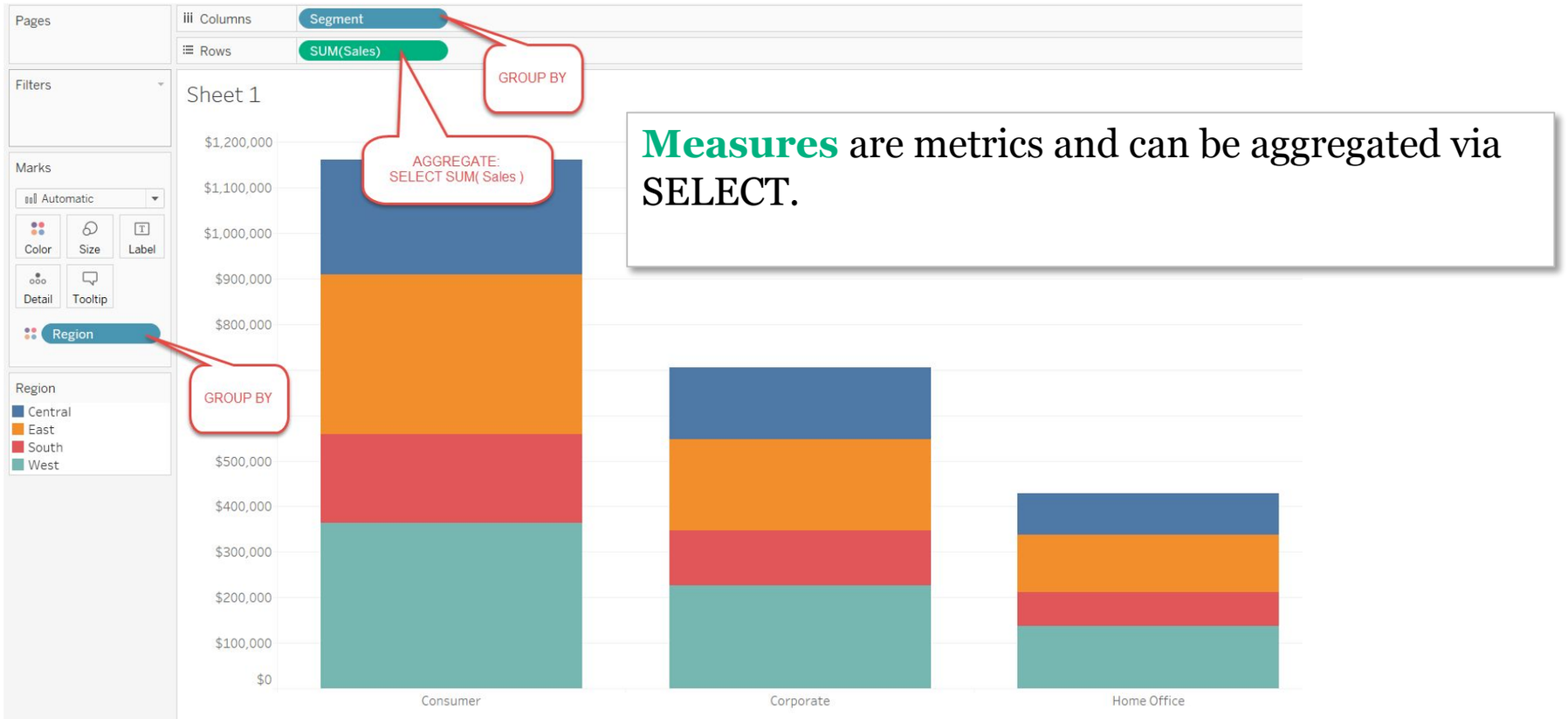
TABLEAU's VizQL to SQL ANALOGS

Dimensions are fields that contain discrete qualitative data. Examples of dimensions include dates, customer names, and customer segments. They are shown as blue.



TABLEAU's VizQL to SQL ANALOGS

Measures (shown as green) are fields that contain continuous numerical data that can be aggregated. Examples of measures include sales, profit, number of employees, temperature, frequency, pressure, etc.

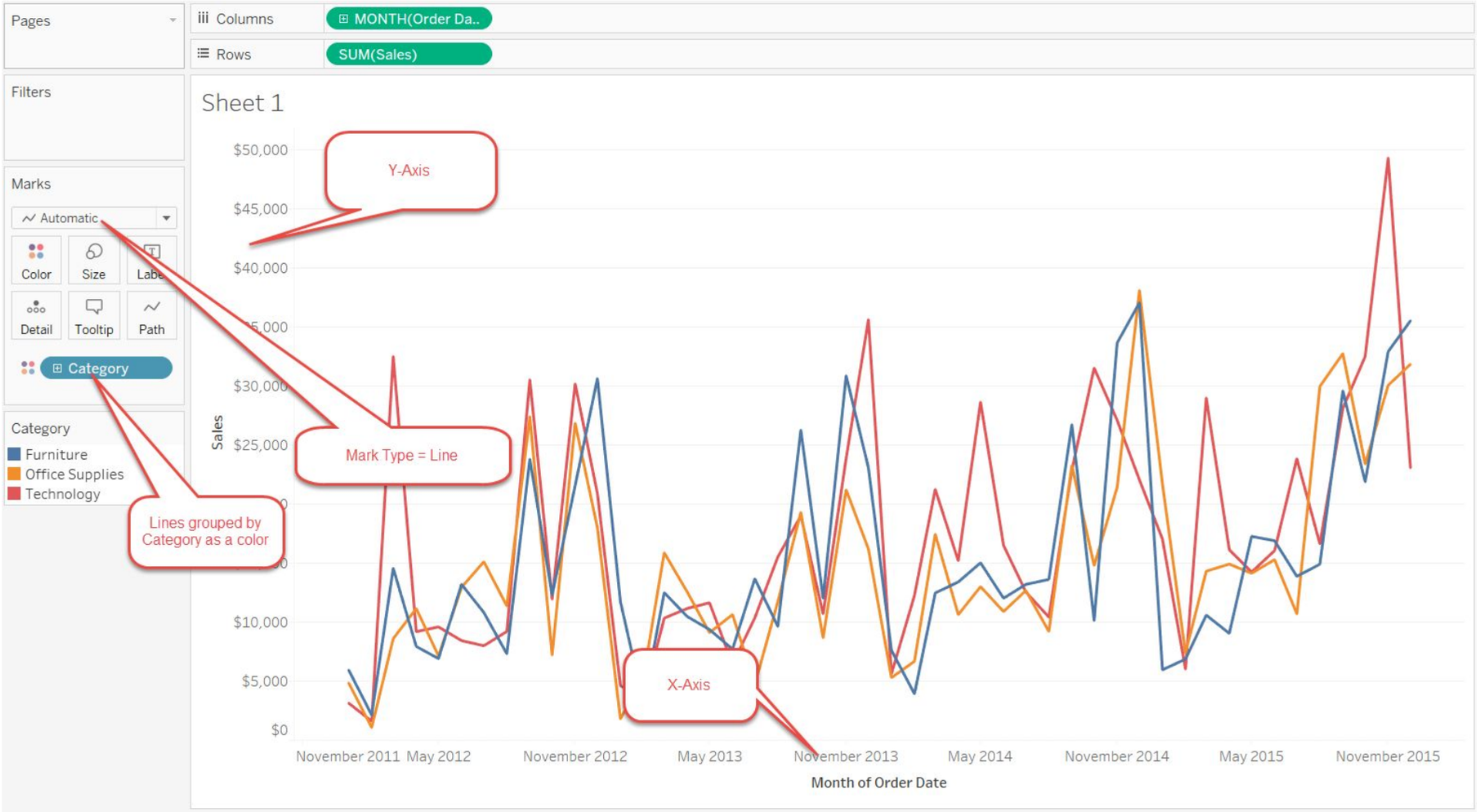


VISUALIZATION TYPES

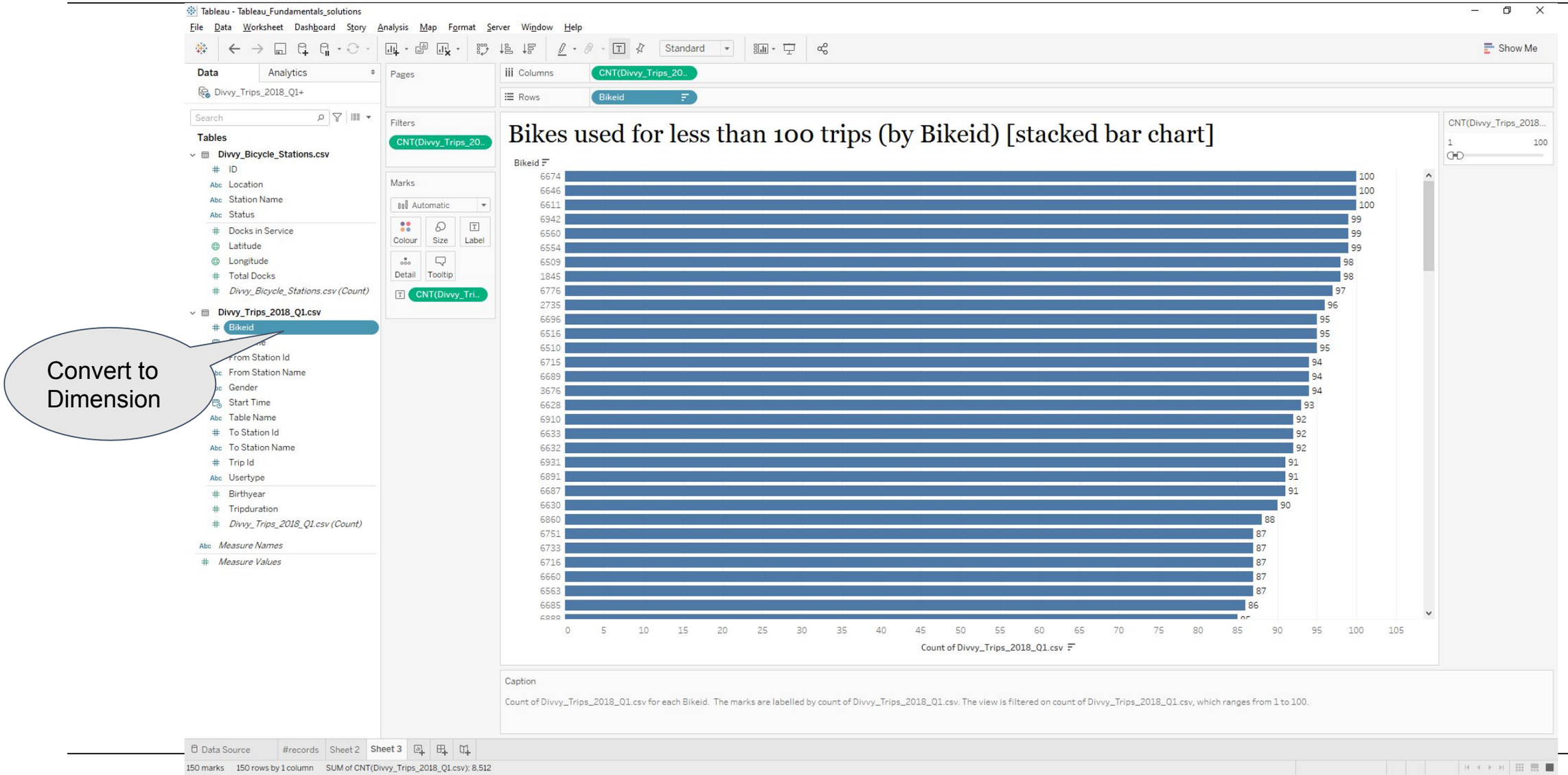
The four primary elements of a Tableau visualization (or “Viz”) are:

1. **Columns**: x axis.
2. **Rows**: y axis.
3. **Mark card**: color, size, and text.
4. **Mark type**: line, bar, square, circle, shape, area, etc.

TABLEAU VISUALIZATION PANE



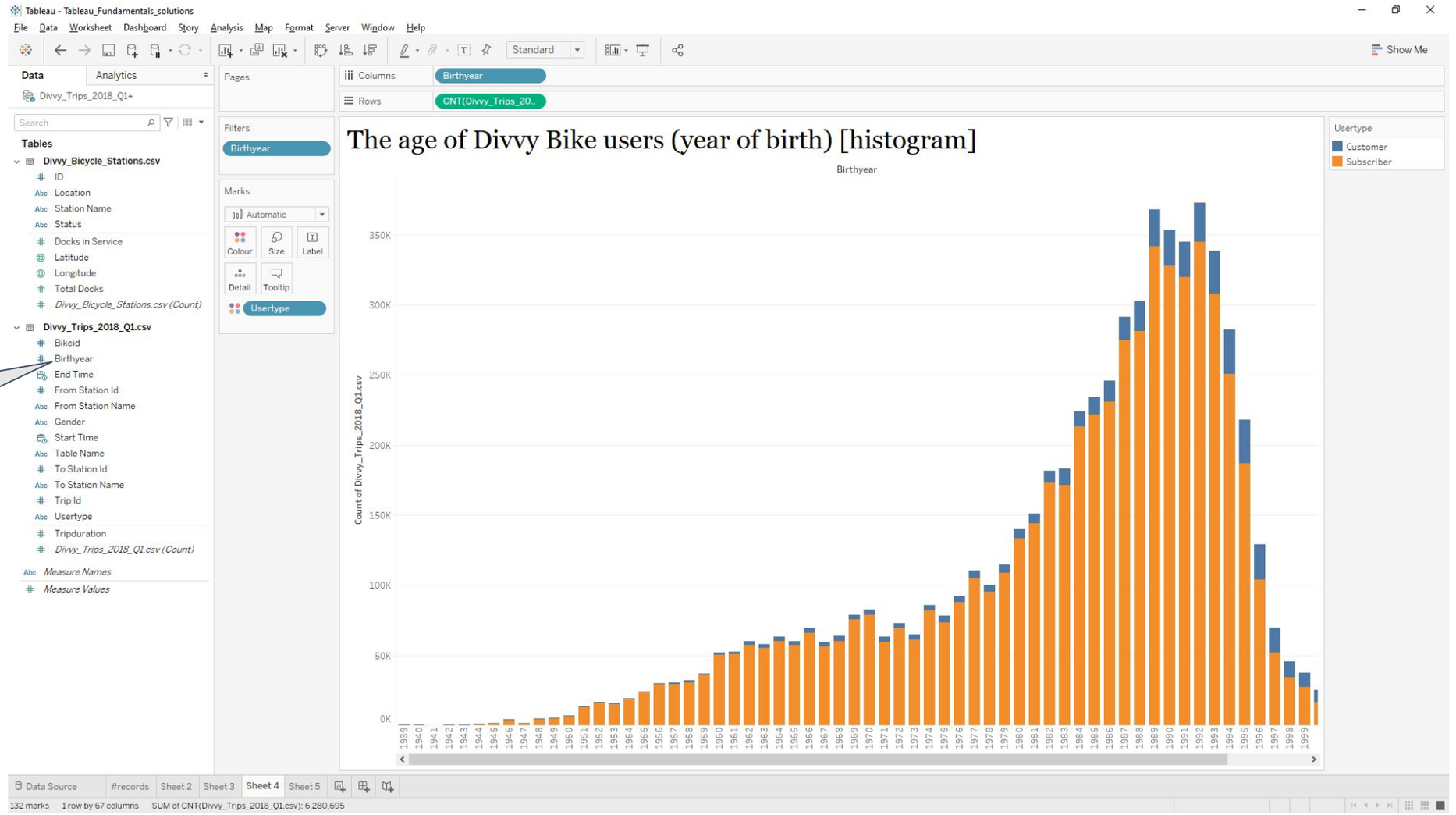
Bikes used for less than 100 trips (by Bikeid) [stacked bar chart]



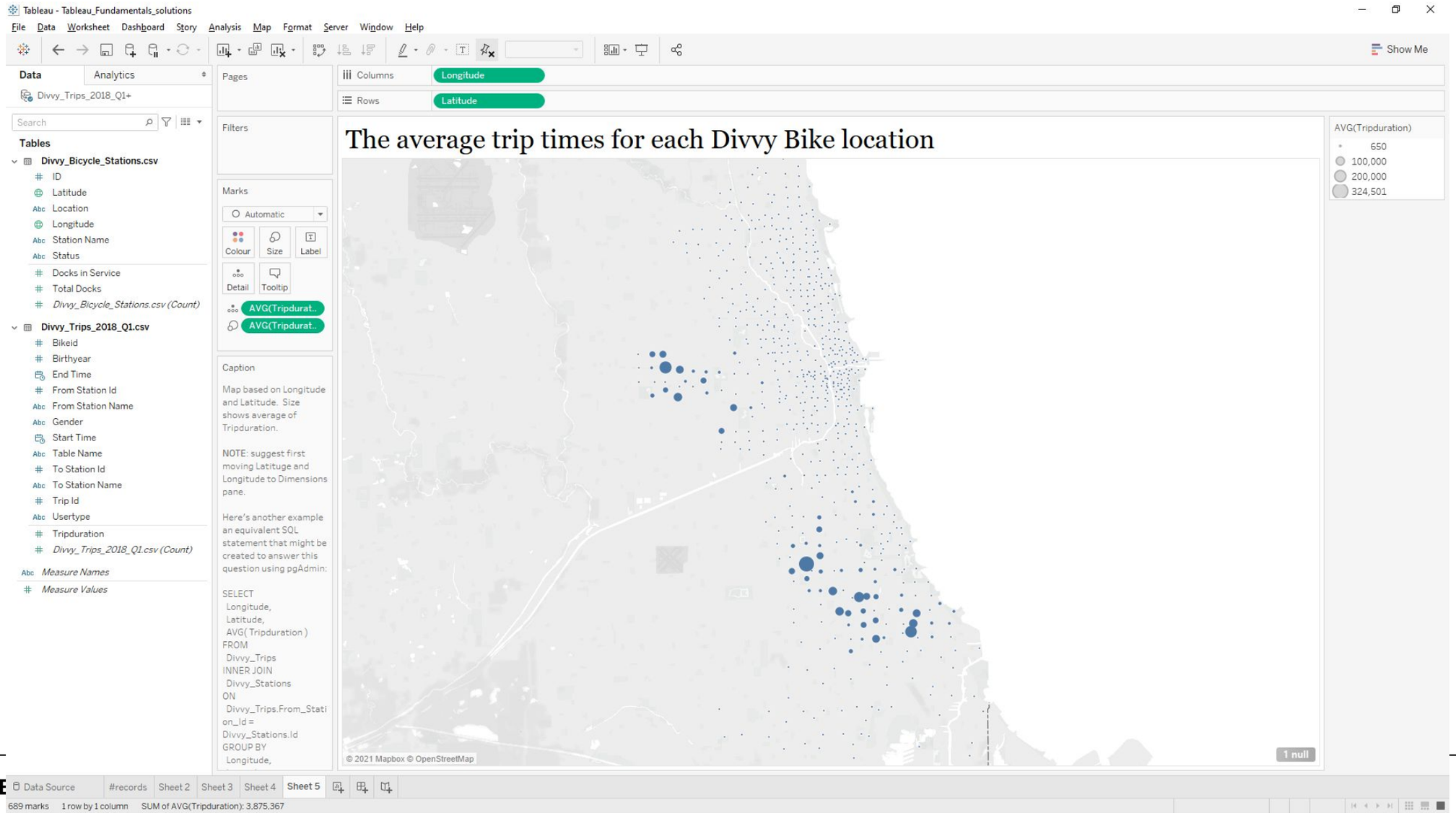
Convert to Dimension

The age of Divvy Bike users (year of birth) [histogram].

Convert to Dimension



The average trip times for each Divvy Bike location



INDEPENDENT PRACTICE: BUILD GRAPHS TO ANSWER QUESTIONS

ACTIVITY: INDEPENDENT PRACTISE

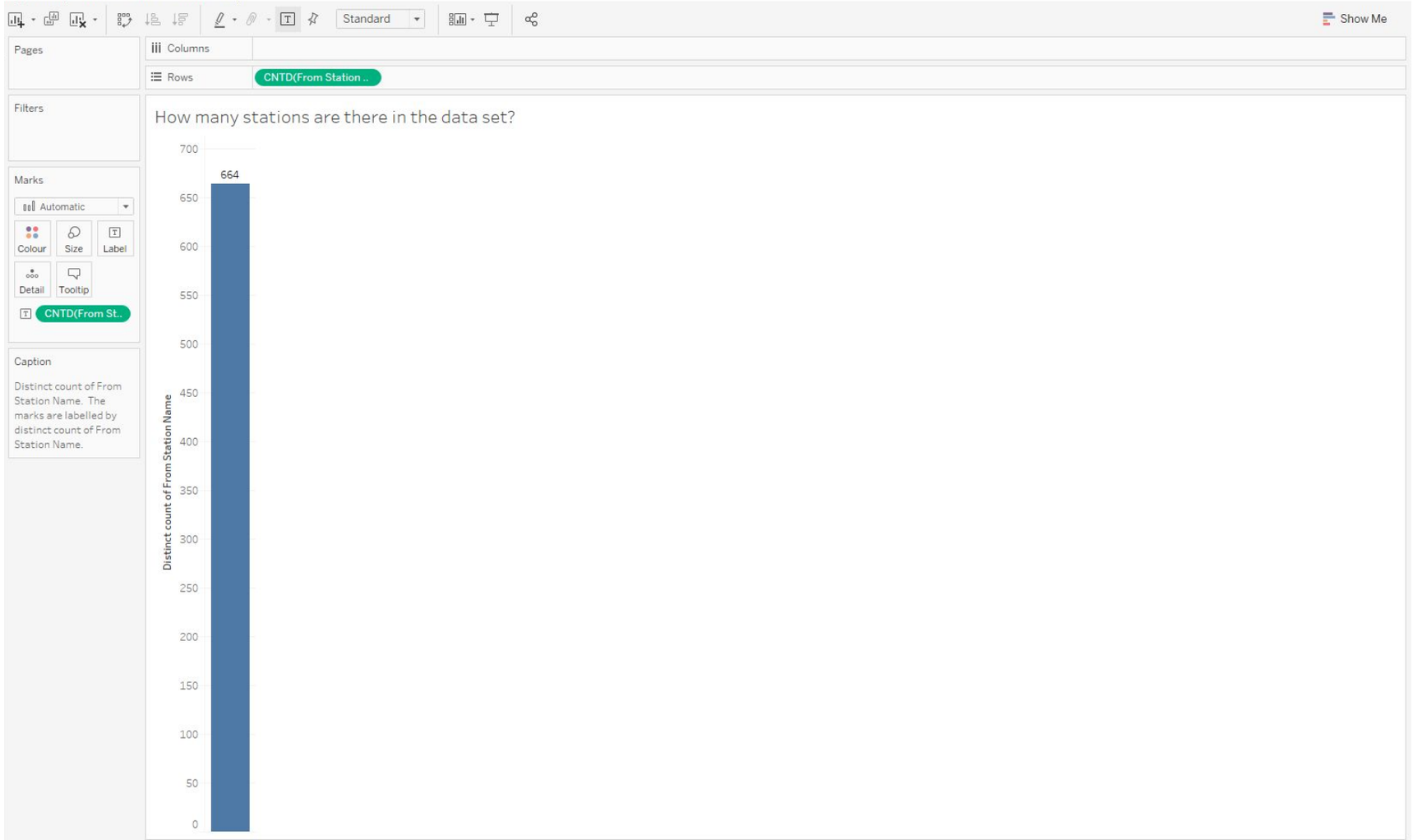
DIRECTIONS

1. Take the next 10 minutes to build out visualizations to answer the following questions.

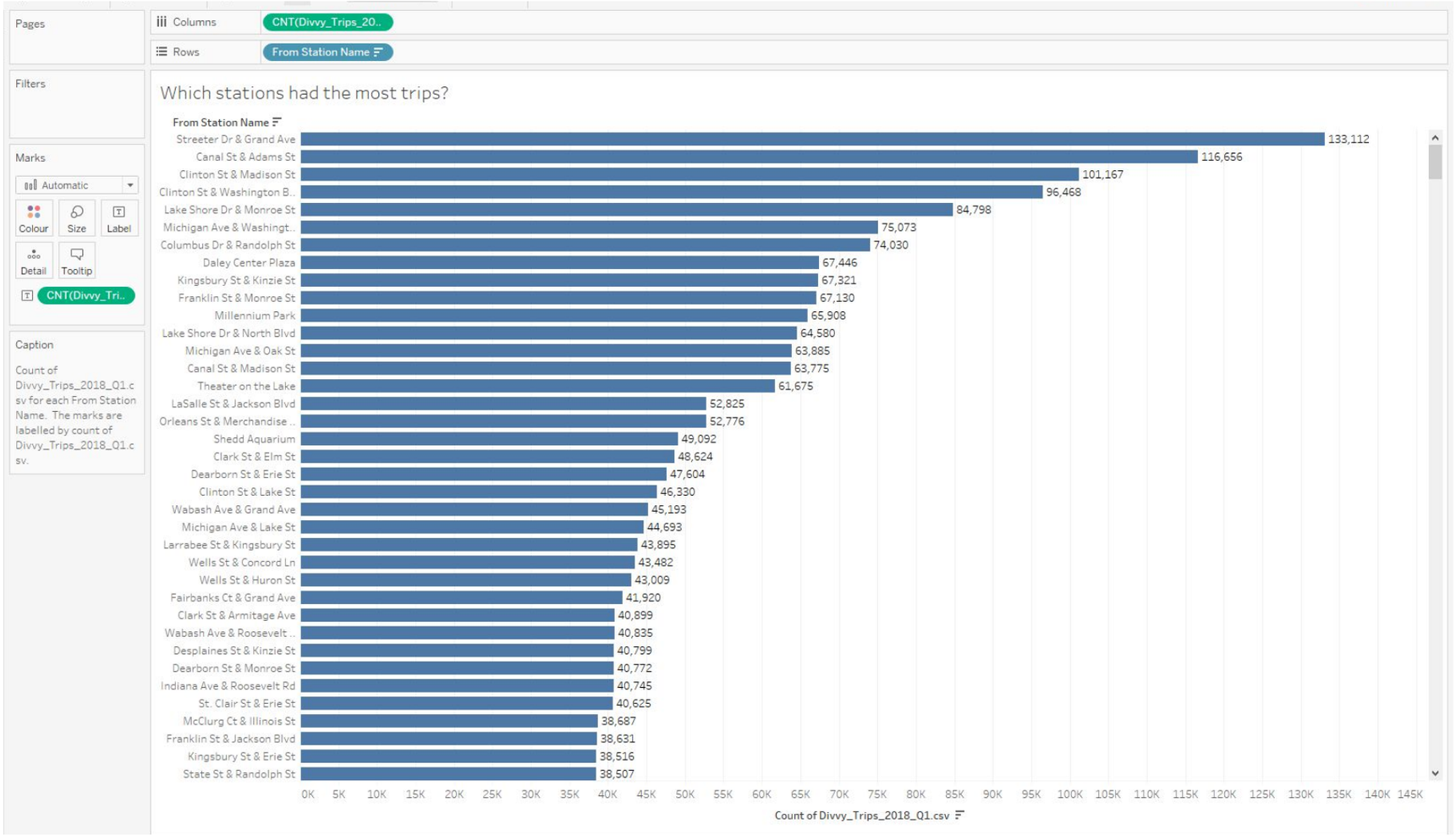
2. Use **From Station Name** for this exercise.
   a. How many stations are there in the data set?
   b. Which station had the most trips?
   c. Which station had the longest recorded trip duration?

3. After that's completed, discuss which graphs are most effective for answering each question.

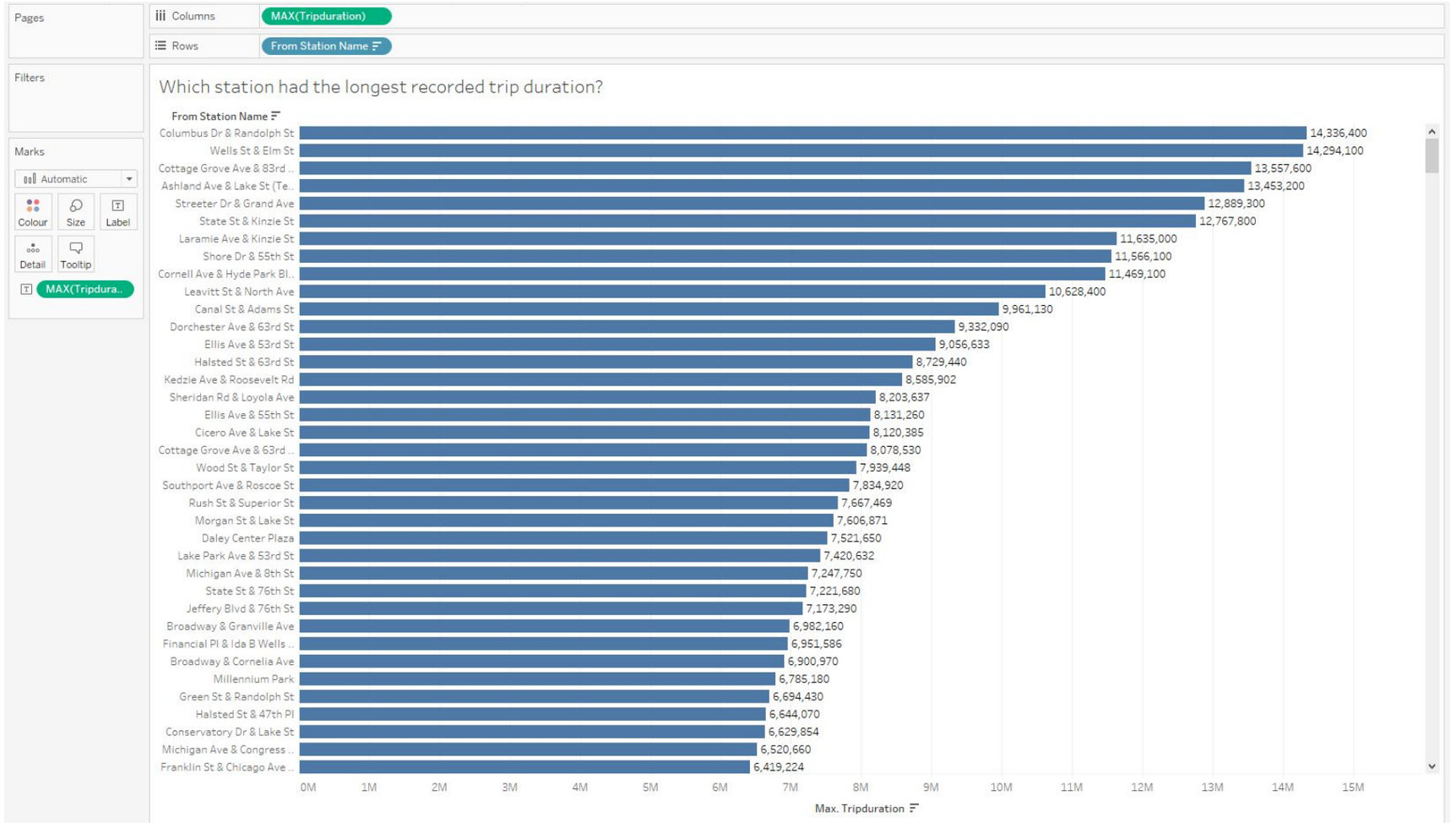
How many stations are there in the data set?



Which stations had the most trips?



Which station had the longest recorded trip duration?



GUIDED PRACTICE: WORKING WITH DATES

WORKING WITH DATES

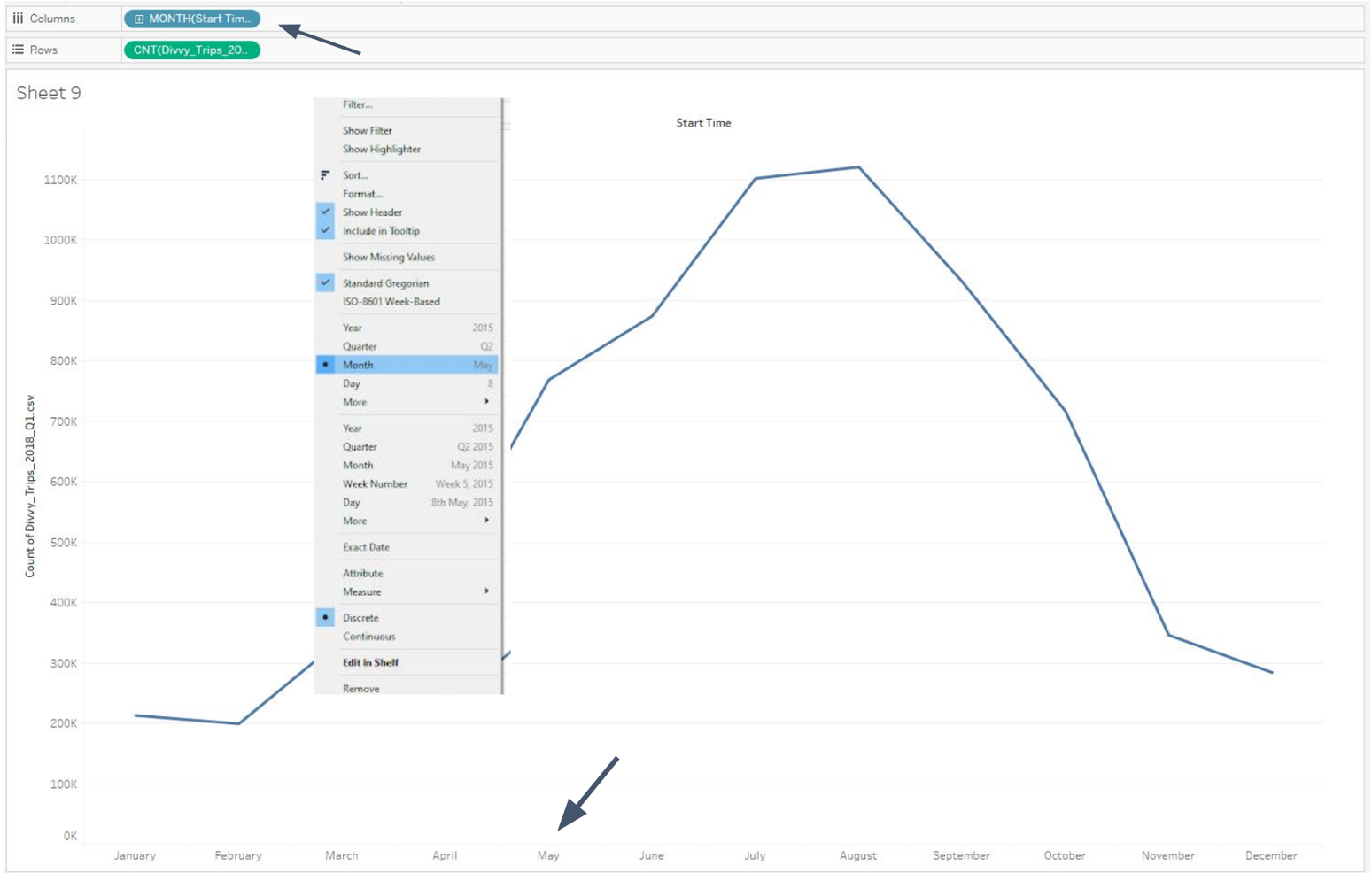
Dates can be used in a variety of ways in Tableau. The most important concept to understand is continuous versus discrete uses for dates in Tableau.

Discrete dates create bins, while **continuous** dates create axes.

- If you visualize each month as a discrete date, it will not distinguish between years. However, you can sort the time axis.
- If month is used as a continuous date, then an axis that separates each month by year will be visualized. You cannot sort this axis.

DISCRETE DATE AGGREGATION

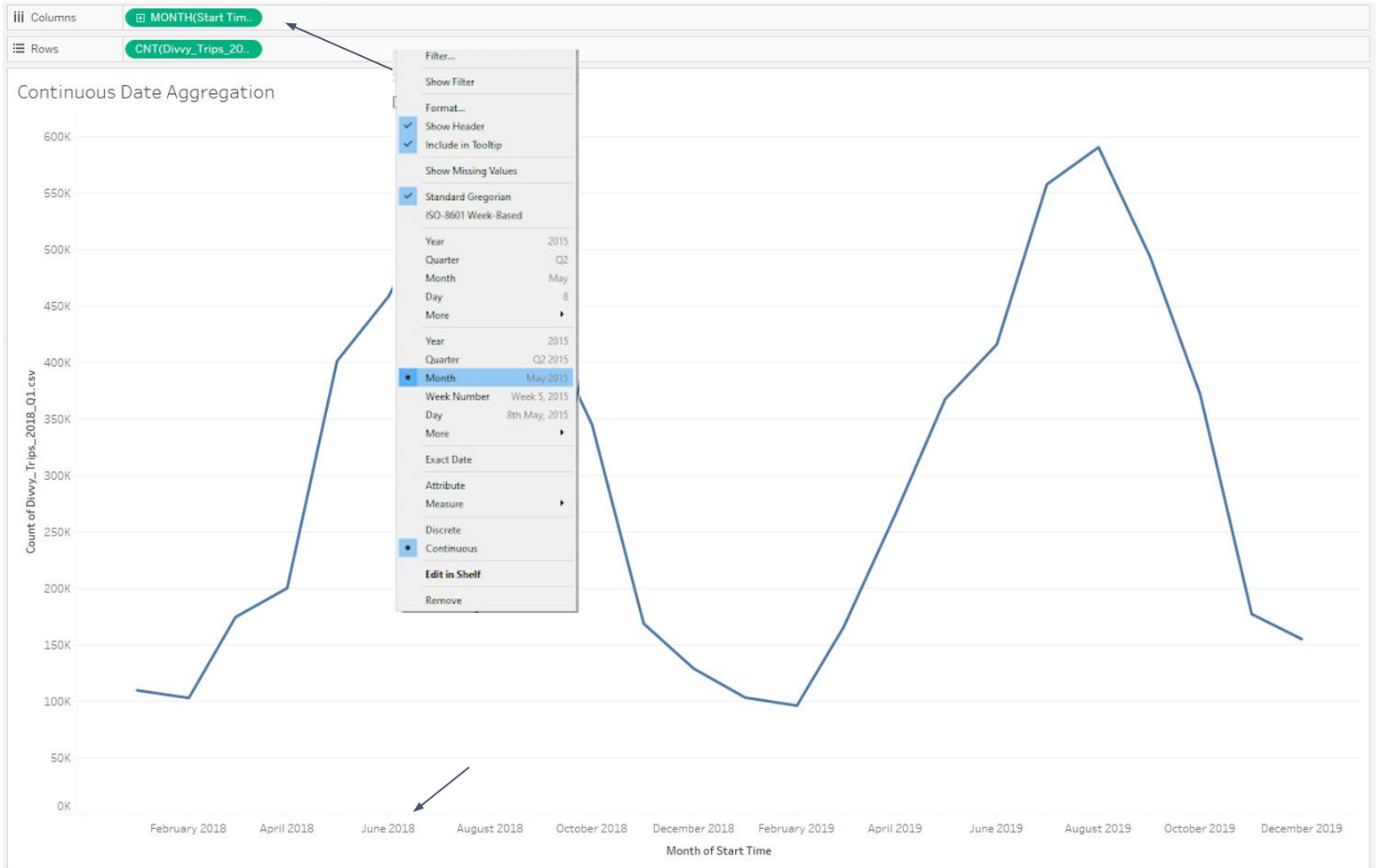
If we choose "Month" where the example is listed as May, then it will aggregate across all years.



CONTINUOUS DATE AGGREGATION

If we select “Month” where the date is formatted as May 2015, then our dates will be aggregated to year and month.

You can also drill down to week, day, etc.



Dimensions and Measures -Using annotate tool tick the right bucket

Characteristics	Dimensions	Measures
Qualitative - usually discrete		
Can be aggregated		
Quantitative - usually continuous		
Usually color green		
Usually color coded blue		
Used to group and segment data		

INDEPENDENT PRACTICE: WORKING WITH DATES

ACTIVITY: WORKING WITH DATES

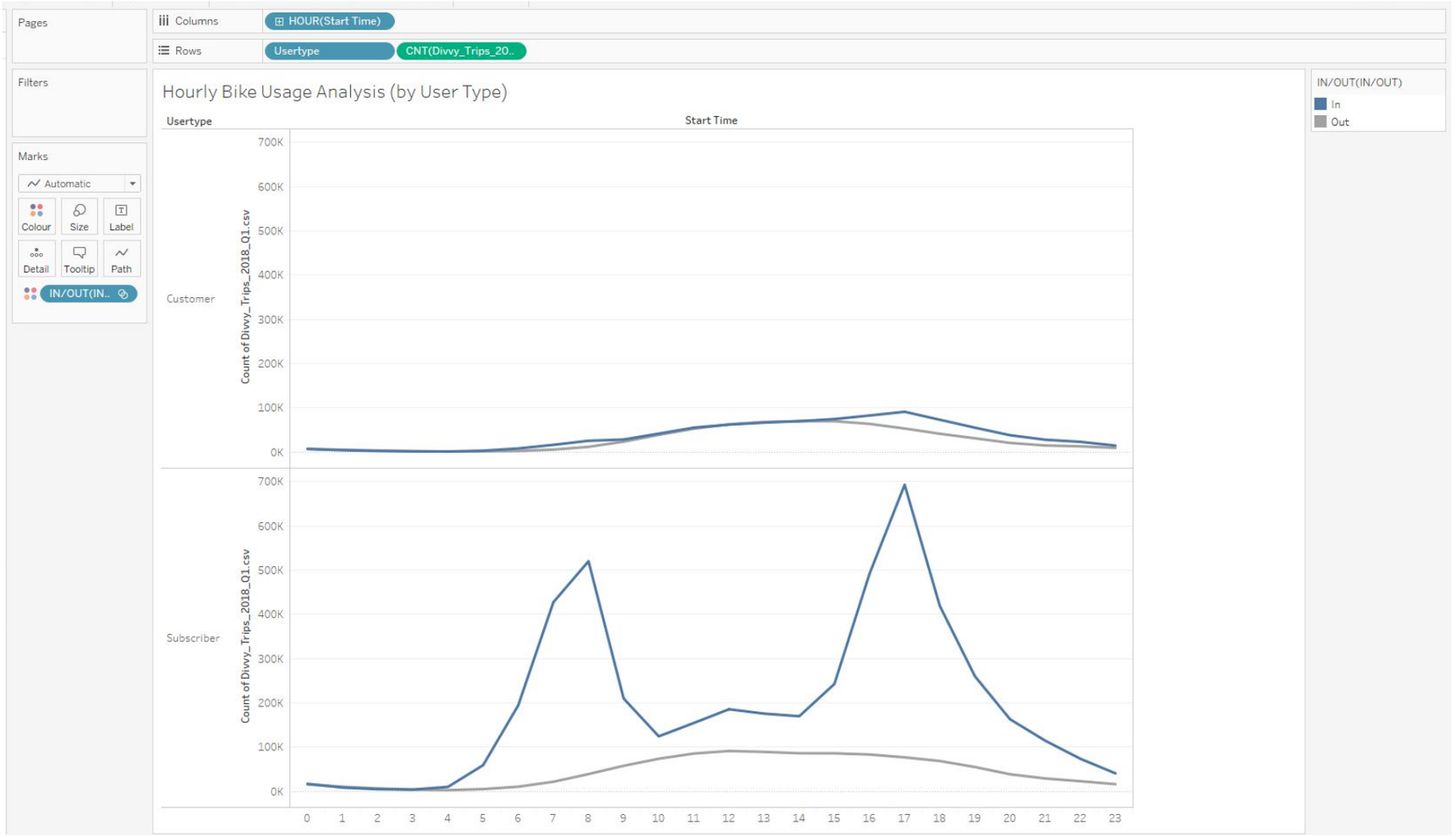
DIRECTIONS

1. Create a set using weekdays considering IN/OUT (weekday in/out)

2. Given the question — “What is the hourly bike usage analysis by user type?” — determine whether you will need to use discrete or continuous dates? Why?

Note: Use the “weekday in/out” dataset you previously created.

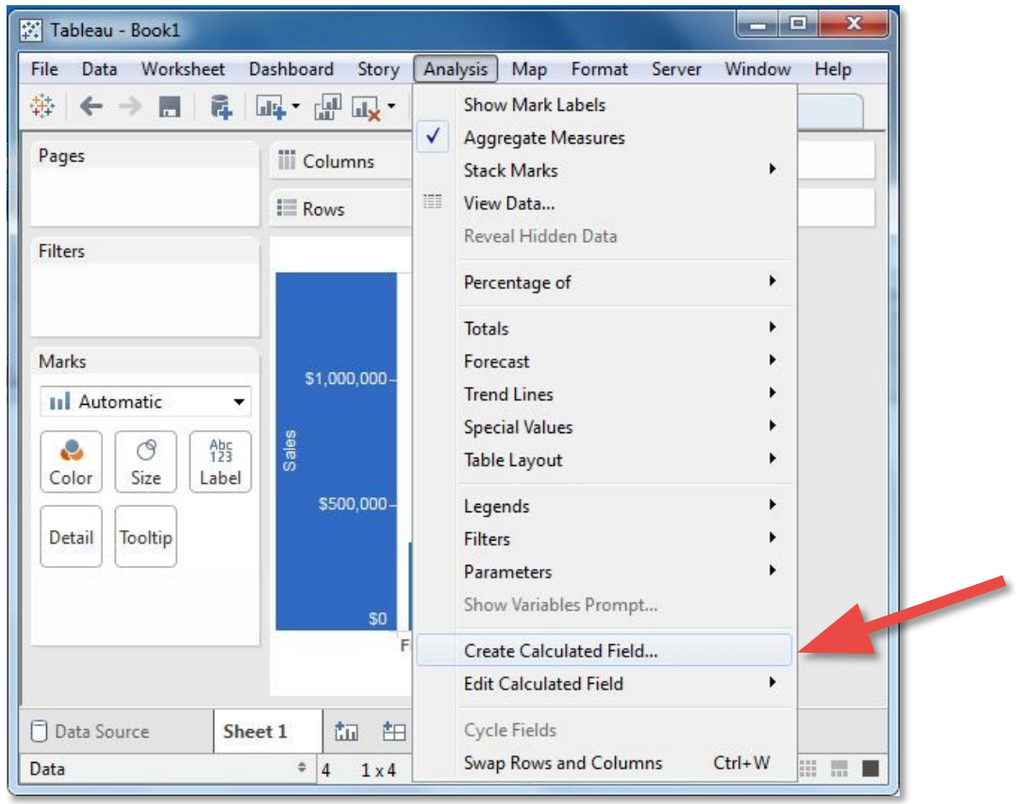
Hourly Bike Usage Analysis (by User Type)



INTRODUCTION: CALCULATED FIELDS

INTRODUCTION TO CALCULATIONS

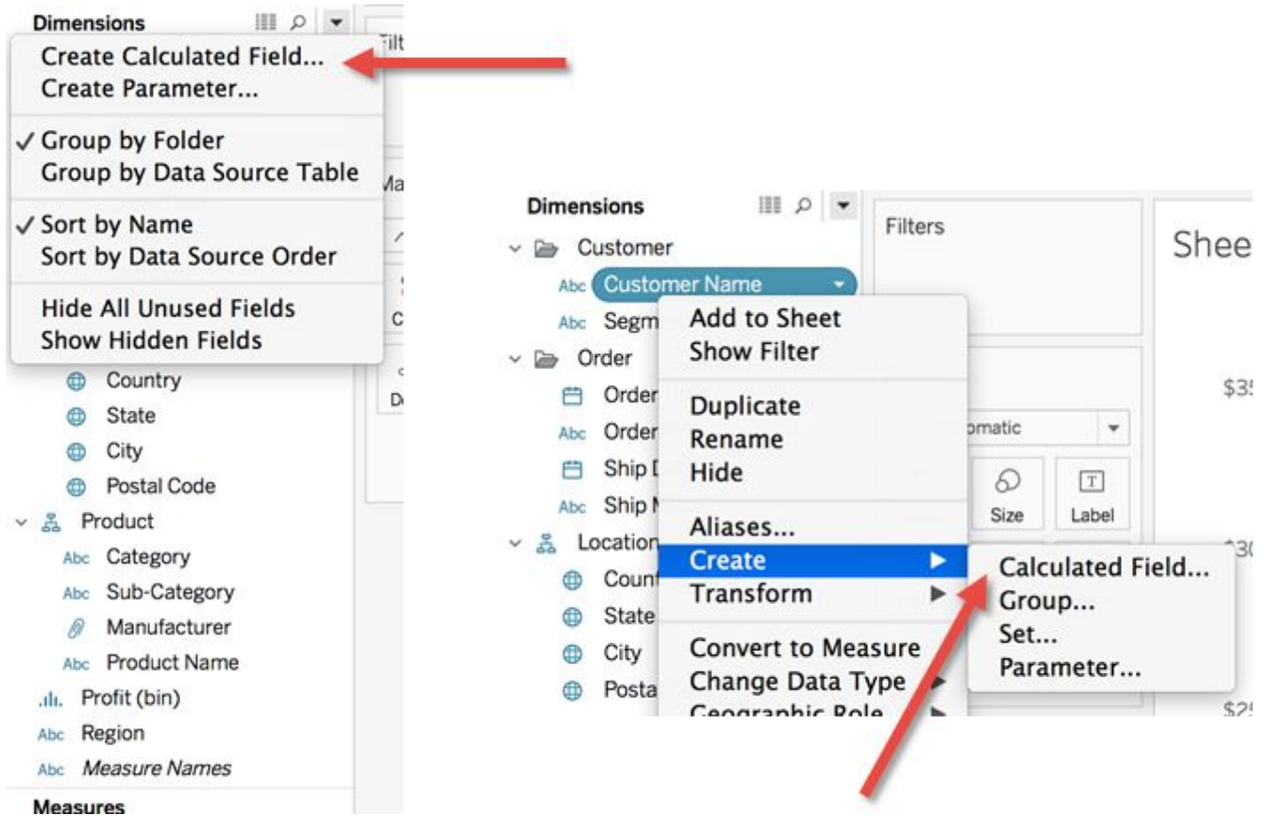
Calculations can be created in Tableau from the “Analysis” tab.



INTRODUCTION TO CALCULATIONS

Calculations can also be created by clicking on the triangle next to dimensions and selecting "**Create Calculated Field…**"

If you know what specific dimension or measure you'll use to create a calculated field, you can right click on that pill and select "**Create > Calculated Field…**"



INTRODUCTION TO CALCULATIONS

A **calculated field** is one that you calculate; it gives you a large amount of flexibility on how you want to see your data.

Calculated fields can be used against text to combine multiple fields.

Pro Tip: Calculations function like an `IF` statement in programming, which is similar to a `CASE` statement in SQL.

INTRODUCTION TO CALCULATIONS

Calculated fields also give us the ability to perform mathematical operations on different measures.

For example, if you have *sales* as one measure and *population* as another, you can determine *sales per capita* as a new measure.

There are too many field types for us to cover in this class, but if you want to learn more about fields we don't discuss, you can find more tutorials on the Tableau reference site.

GUIDED PRACTICE: DASHBOARDS & STORY POINTS

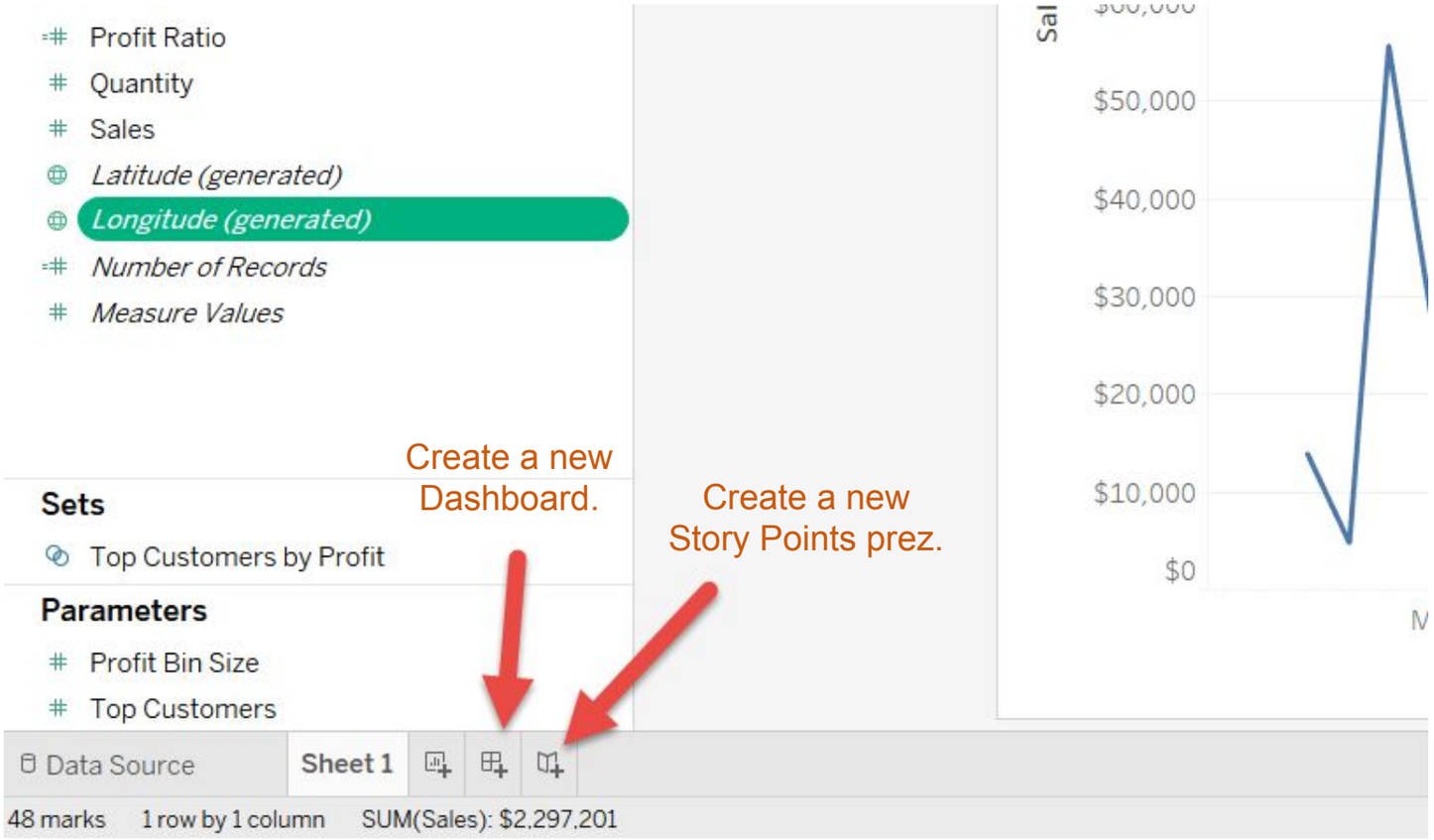
DASHBOARDS AND STORY POINTS

Dashboards allow you to combine multiple visualizations in a single view and add interactions between those visualizations.

Story Points combine multiple dashboards or visualizations into an interactive full-screen presentation.

DASHBOARDS AND STORY POINTS

Create dashboards and Story Points using tabs at the bottom of the screen.



INDEPENDENT PRACTICE: DASHBOARDS & STORY POINTS

ACTIVITY: DASHBOARDS AND STORIES

DIRECTIONS

1. In groups, build your own graphs, dashboards, or stories that show the key differences between ***subscribers*** and ***customers*** (user type).

Pro Tip: Under Tableau's worksheet menu, locate and select "Describe Sheet" to see how each of the example stretch activity's visualizations are constructed.

DELIVERABLES

As a team, present your visualizations and findings to the class.

CONCLUSION

RECAP

In this lesson, we learned how to:

1. Connect and prepare data to import into Tableau.
2. Explore the basics of the Tableau interface.
3. Build graphs, calculated fields, dashboards, and stories.
4. Work with discrete and continuous dates in Tableau.

Q&A

RESOURCES

- Visual Analysis Best Practices (Tableau): https://www.tableau.com/learn/whitepapers/tableau-visual-guidebook

- Which Chart or Graph is Right for You? (Tableau): https://www.tableau.com/sites/default/files/media/which_chart_v6_final_0.pdf

- Tableau “Starter Kits” (role-based tutorials): https://www.tableau.com/learn/starter-kits