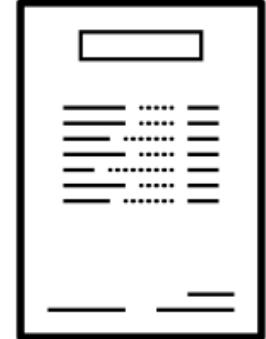


Tableau Desktop Fundamentals



Table of Contents

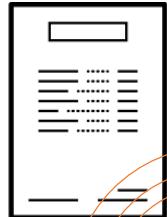


1. Taking Off with Tableau: 4
2. Connecting to Data in Tableau: 54
3. Moving Beyond Basic Visualizations: 96
4. Starting an Adventure with Calculations and Parameters: 165
5. Leveraging Level of Detail Calculations: 221
6. Diving Deep with Table Calculations: 263
7. Making Visualizations That Look Great and Work Well: 309
8. Telling a Data Story with Dashboards: 341



Table of Contents

9. Visual Analytics – Trends, Clustering, Distributions, and Forecasting: 404
10. Advanced Visualizations: 450
11. Dynamic Dashboards: 474
12. Exploring Mapping and Advanced Geospatial Features: 496
13. Understanding the Tableau Data Model, Joins, and Blends: 539
14. Structuring Messy Data to Work Well in Tableau: 592
15. Taming Data with Tableau Prep: 643
16. Sharing Your Data Story: 701



1. Taking Off with Tableau

A photograph of a person's hands typing on a silver laptop keyboard. In the background, a computer monitor displays a blurred dashboard or data visualization. The overall scene suggests a professional workspace.

Taking Off with Tableau

We will cover the following topics in this lesson

- Connecting to data
- Foundations for building visualizations
- Creating bar charts
- Creating line charts
- Creating geographic visualizations
- Using Show Me
- Bringing everything together in a dashboard



Connecting to data

- Tableau connects to data stored in a wide variety of files and databases.
- This includes flat files, such as Excel documents, spatial files, and text files; relational databases, such as SQL Server and Oracle; cloud-based data sources, such as Snowflake and Amazon Redshift; and Online Analytical Processing (OLAP) data sources, such as Microsoft SQL Server Analysis Services.

Connecting to data

1. Open Tableau. You should see the home screen with a list of connection options on the left and, if applicable, thumbnail previews of recently edited workbooks in the center, along with sample workbooks at the bottom.
2. Under Connect and To a File, click on Text File.
3. In the Open dialog box, navigate to the \Learning Tableau\lesson 01 directory and select the Superstore.csv file.



← → ⌂ ⌂

Connections

Add

Superstore

Text file

Files

...

Use Data Interpreter

Data Interpreter might be able
to clean your Text file
workbook.

Superstore.csv

New Union

Superstore

Connection

Live

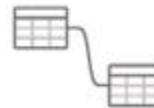
Extract

Filters

0

Add

Superstore



Need more data?

Drag tables here to relate them. [Learn more](#)



Sort fields

Data source order



Show aliases

Show hidden fields

1,000

rows

Abc Superstore.csv Category	⊕ Superstore.csv City	Abc Superstore.csv Container	# Superstore.csv Customer ID	Abc Superstore.csv Customer Name
Scissors, Rulers and T...	Washington	Small Pack	2867	Dana Teague
Storage & Organization	Oxford	Medium Box	1821	Vanessa Boyer
Telephones and Com...	Irving	Small Box	1402	Wesley Tate
Paper	Irving	Small Box	1402	Wesley Tate
Bookcases	Oxford	Jumbo Box	2747	Brian Grady

Tableau - Chapter 01 Starter

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Standard Show Me

1 2 3 4 5 6 7

Data Analytics Superstore

Search

Tables

- Order Date
- Order ID
- Order Priority
- Postal Code
- Region
- Row ID
- Ship Date
- Ship Mode
- State
- Measure Names
- Discount
- Order Quantity
- Product Base Margin
- Profit
- Sales
- Shipping Cost
- Unit Price
- Latitude (generated)
- Longitude (generated)

Pages Columns Rows

Filters

Marks

- Automatic
- Color
- Size
- Text
- Detail
- Tooltip

Sheet 1

Drop field here

Drop field here

Select or drag data
Use the Shift or Ctrl key to select multiple fields

Data Source Workbook Introduction Sheet 1 Measures and Dimensions Sales by Department Bar Chart (two levels) Bar C

The screenshot shows the Tableau desktop application with the following interface elements:

- Top Bar:** File, Data, Worksheet, Dashboard, Story, Analysis, Map, Format, Server, Window, Help.
- Standard/Show Me Buttons:** Standard (selected), Show Me.
- Left Panel (Data View):**
 - Header:** Data, Analytics, Superstore.
 - Search:** Search bar.
 - Tables Section:** Order Date, Order ID, Order Priority, Postal Code, Region, Row ID, Ship Date, Ship Mode, State, Measure Names, Discount, Order Quantity, Product Base Margin, Profit, Sales, Shipping Cost, Unit Price, Latitude (generated), Longitude (generated).
 - Filters:** Placeholder for filtering data.
 - Marks Card:** Automatic, Color, Size, Text, Detail, Tooltip.
- Center Panel (Sheet 1):** A blank canvas with two "Drop field here" placeholder areas.
- Right Panel (Visualizations):** A grid of visualization preview cards, with the first few rows visible.
- Bottom Navigation:** Data Source, Workbook Introduction, Sheet 1 (selected), Measures and Dimensions, Sales by Department, Bar Chart (two levels), Bar C.

Connecting to data

To prepare for this, please do the following:

- From the menu, select File | Exit.
- When prompted to save changes, select No.
- From the \Learning Tableau\Lesson 01 directory, open the file lesson 01 Starter.twbx. This file contains a connection to the Superstore data file and is designed to help you walk through the examples in this lesson.

Foundations for building visualizations

- When you first connect to a data source such as the Superstore file, Tableau will display the data connection and the fields in the Data pane.
- Fields can be dragged from the data pane onto the canvas area or onto various shelves such as Rows, Columns, Color, or Size.

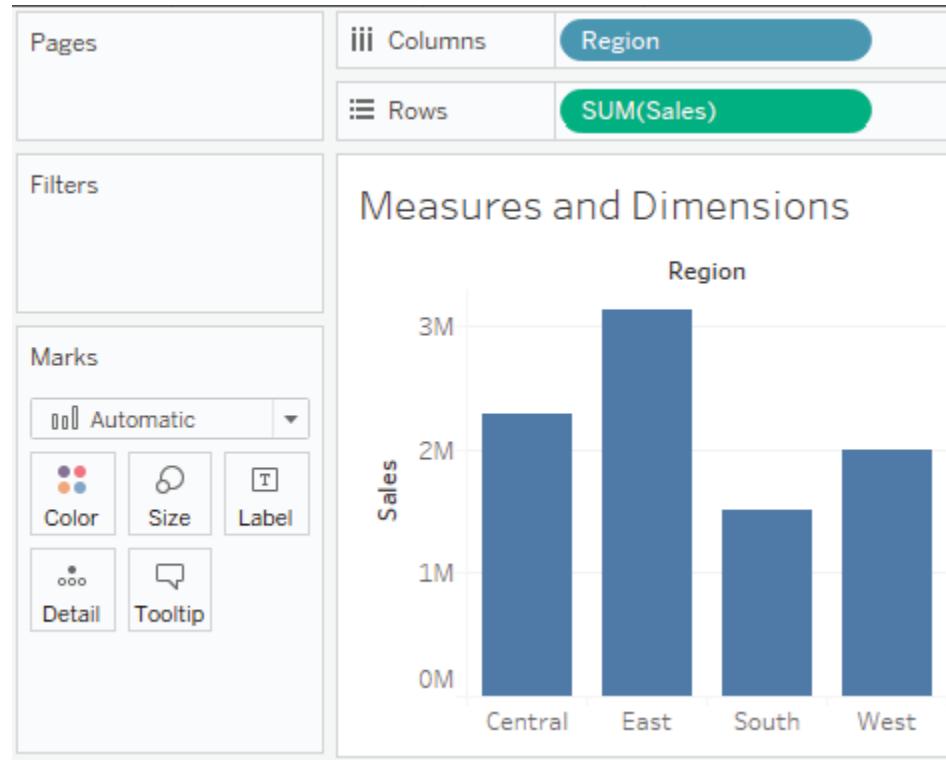
Measures and dimensions

Dimensions
are above the line

Measures
are below the line

Data	Analytics
 Superstore	
<input type="text" value="Search"/> 	 
Tables	
# Order ID	
Abc Order Priority	
 Postal Code	
Abc Region	
# Row ID	
 Ship Date	
Abc Ship Mode	
 State	
Abc <i>Measure Names</i>	
# Discount	
# Order Quantity	
# Product Base Margin	
# Profit	
# Sales	
# Shipping Cost	
# Unit Price	
 <i>Latitude (generated)</i>	
 <i>Longitude (generated)</i>	
# Number of Records	
# Superstore (Count)	
# Measure Values	

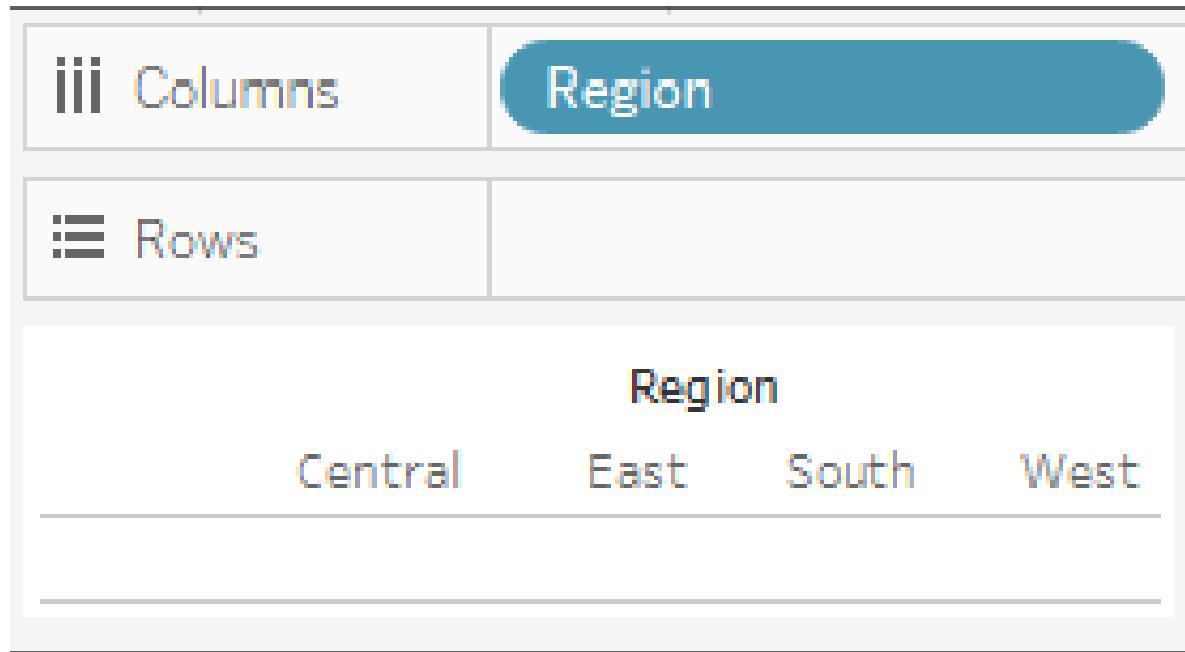
Measures and dimensions



Discrete and continuous fields

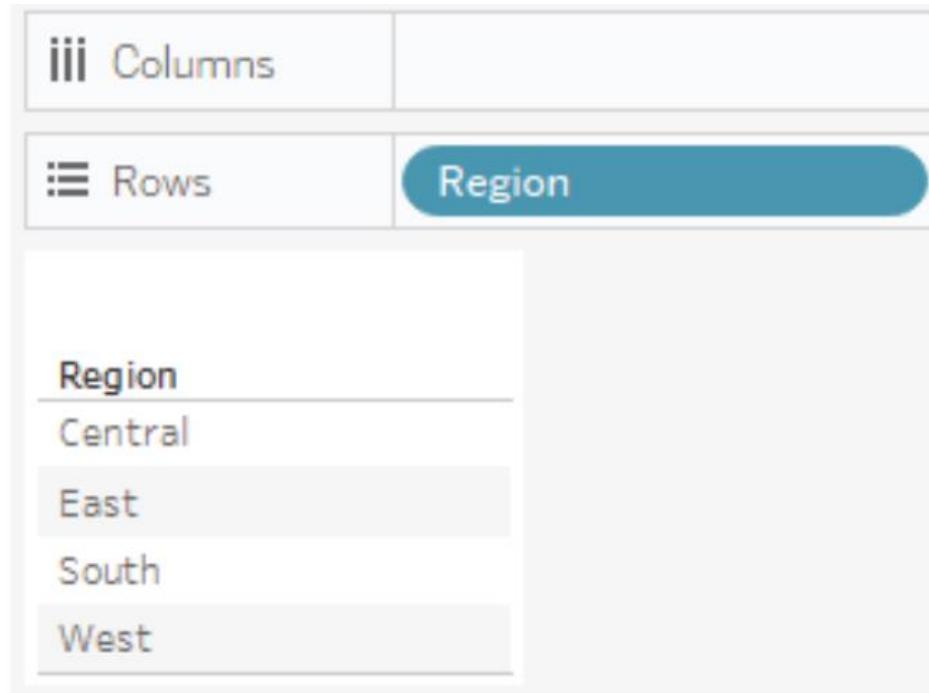
- Another important distinction to make with fields is whether a field is being used as discrete or continuous.
- Whether a field is discrete or continuous determines how Tableau visualizes it based on where it is used in the view.

Discrete fields



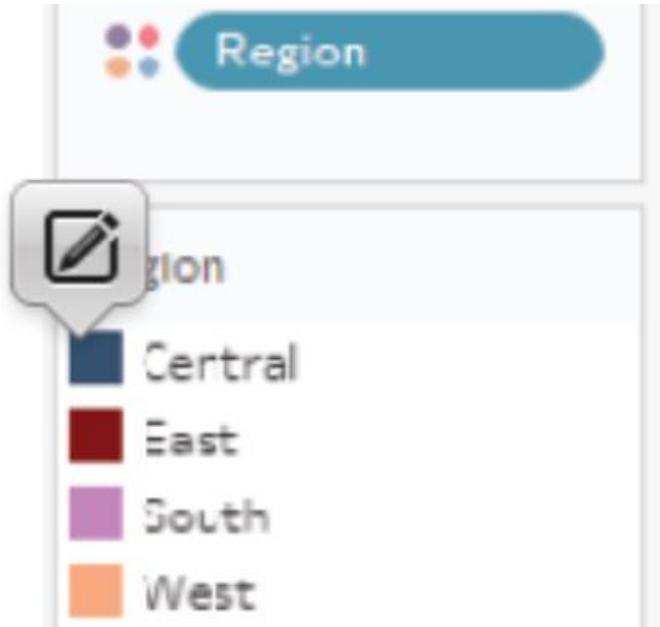
Discrete fields

- Here, it defines the row headers:



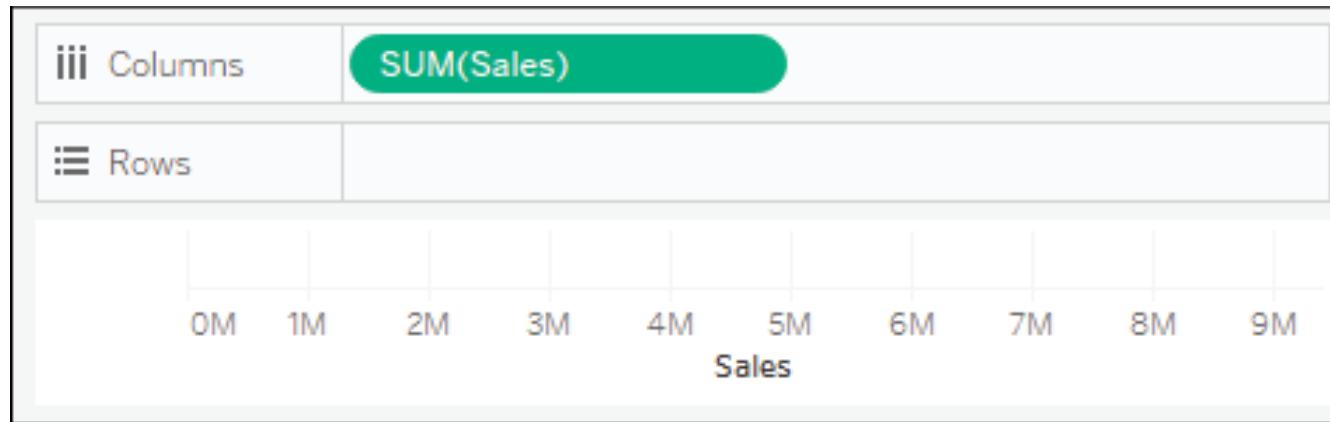
Discrete fields

- When used for Color, a discrete field defines a discrete color palette in which each color aligns with a distinct value of the field:



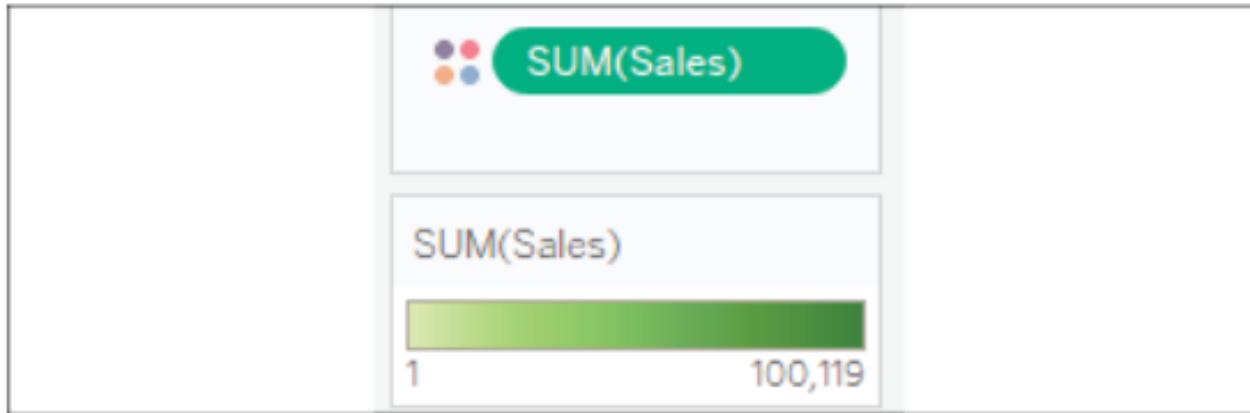
Continuous fields

- When used on Rows or Columns, a continuous field defines an axis



Continuous fields

- When used for color, a continuous field defines a gradient:



Continuous fields

While most dimensions are discrete by default, and most measures are continuous by default, it is possible to use any measure as a discrete field and some dimensions as continuous fields in the view, as shown here:

can be

	Discrete	Continuous
Dimension	Yes	If Numeric or Date
Measure	Yes	Yes

Continuous fields

In general, you can think of the differences between the types of fields as follows:

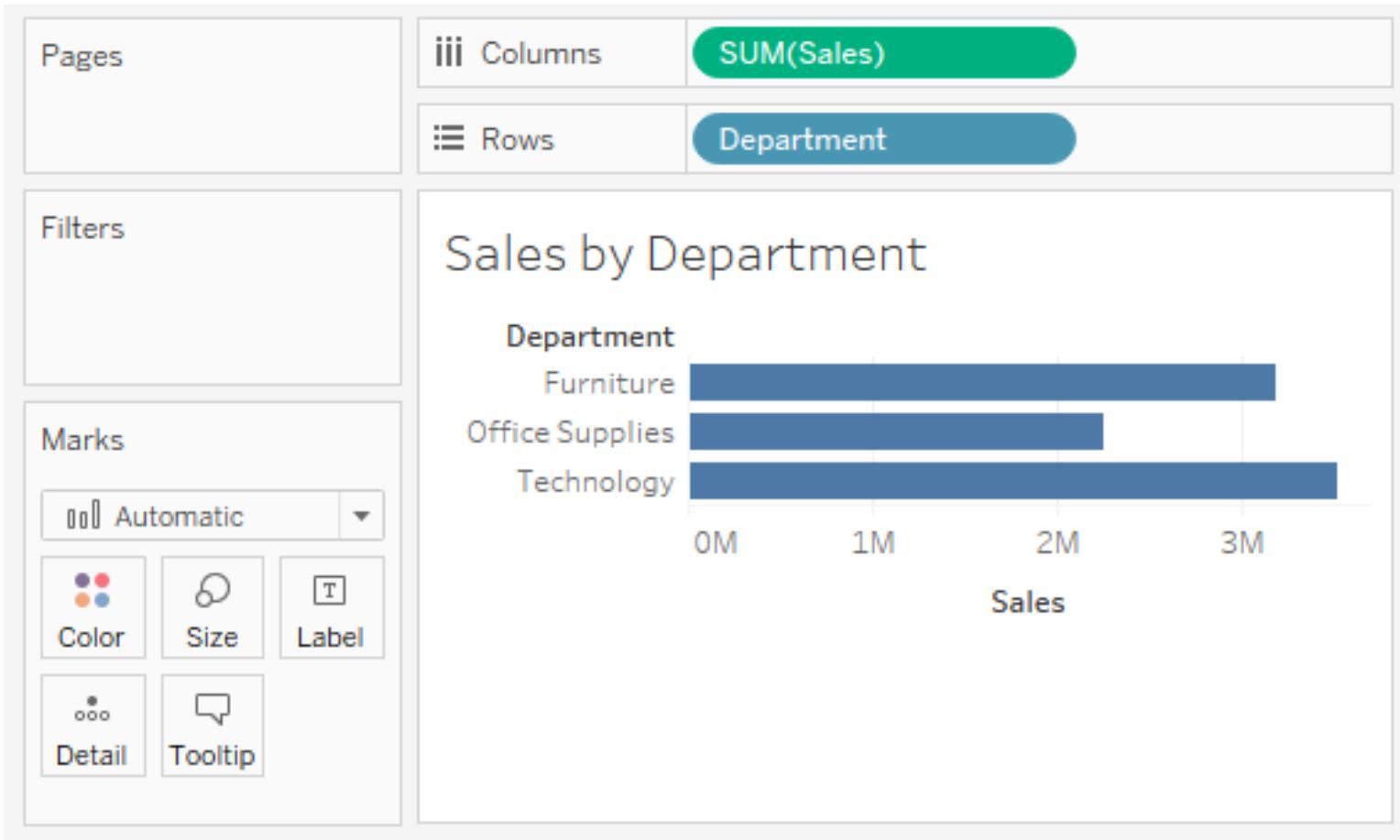
- Choosing between a dimension and measure tells Tableau how to slice or aggregate the data.
- Choosing between discrete and continuous tells Tableau how to display the data with a header or an axis and defines individual colors or a gradient.

Visualizing data

- A new connection to a data source is an invitation to explore and discover! At times, you may come to the data with very well-defined questions and a strong sense of what you expect to find.
- Other times, you will come to the data with general questions and very little idea of what you will find.

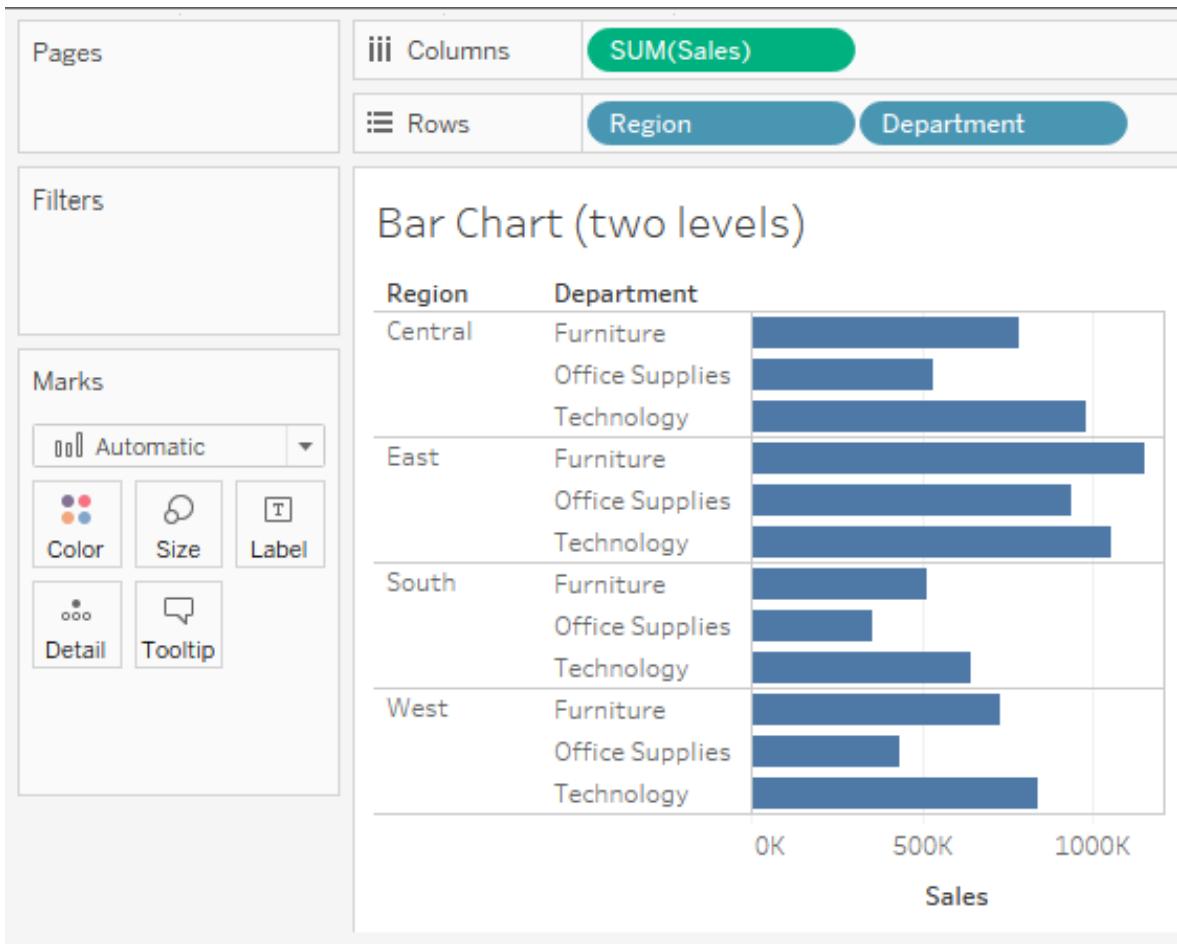
Bar charts

- Bar charts visually represent data in a way that makes the comparison of values across different categories easy.
- The length of the bar is the primary means by which you will visually understand the data.
- You may also incorporate color, size, stacking, and order to communicate additional attributes and values.

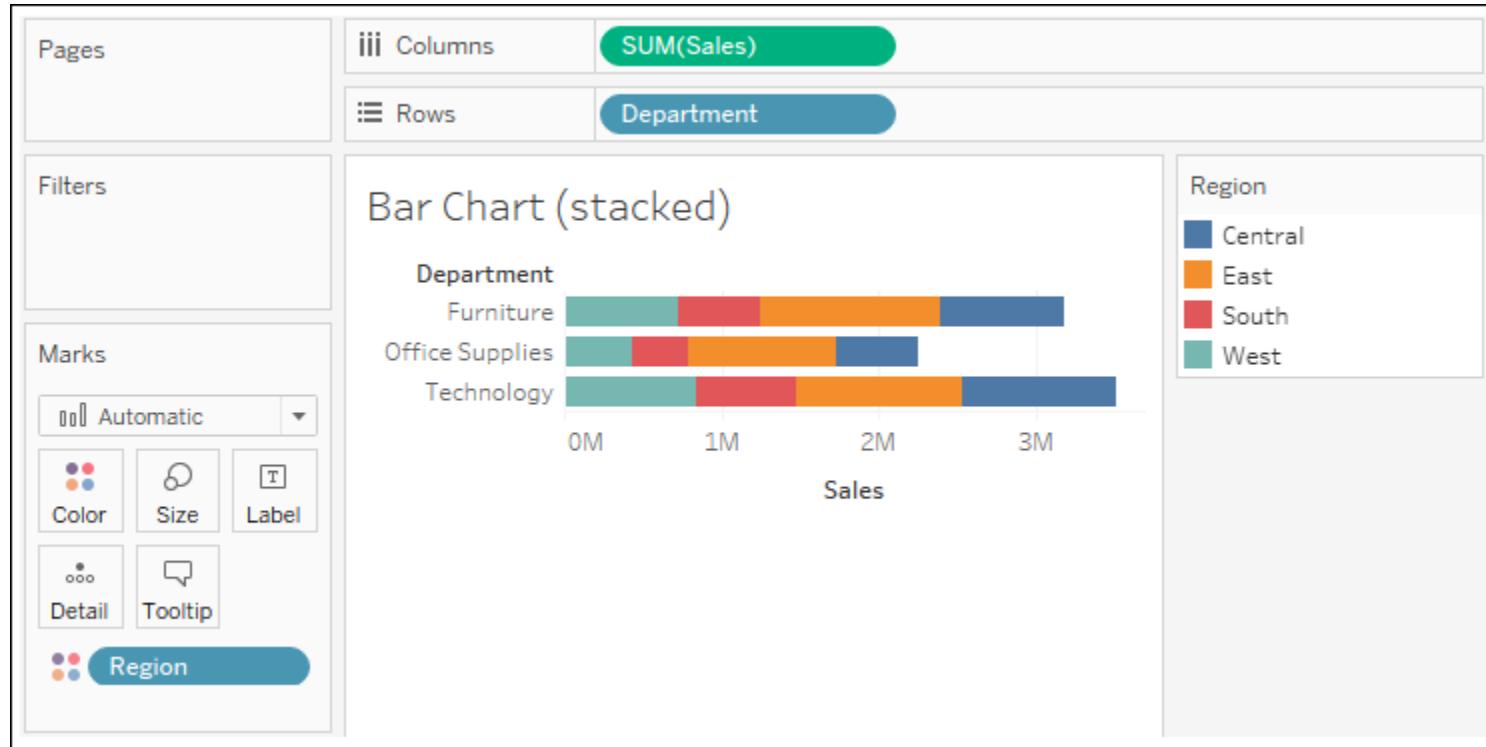


Iterations of bar charts for deeper analysis

1. Navigate to the Bar Chart (two levels) sheet, where you will find an initial view that is identical to the one you created earlier.
2. Drag the Region field from Dimensions in the Data pane to the Rows shelf and drop it to the left of the Department field already in view.

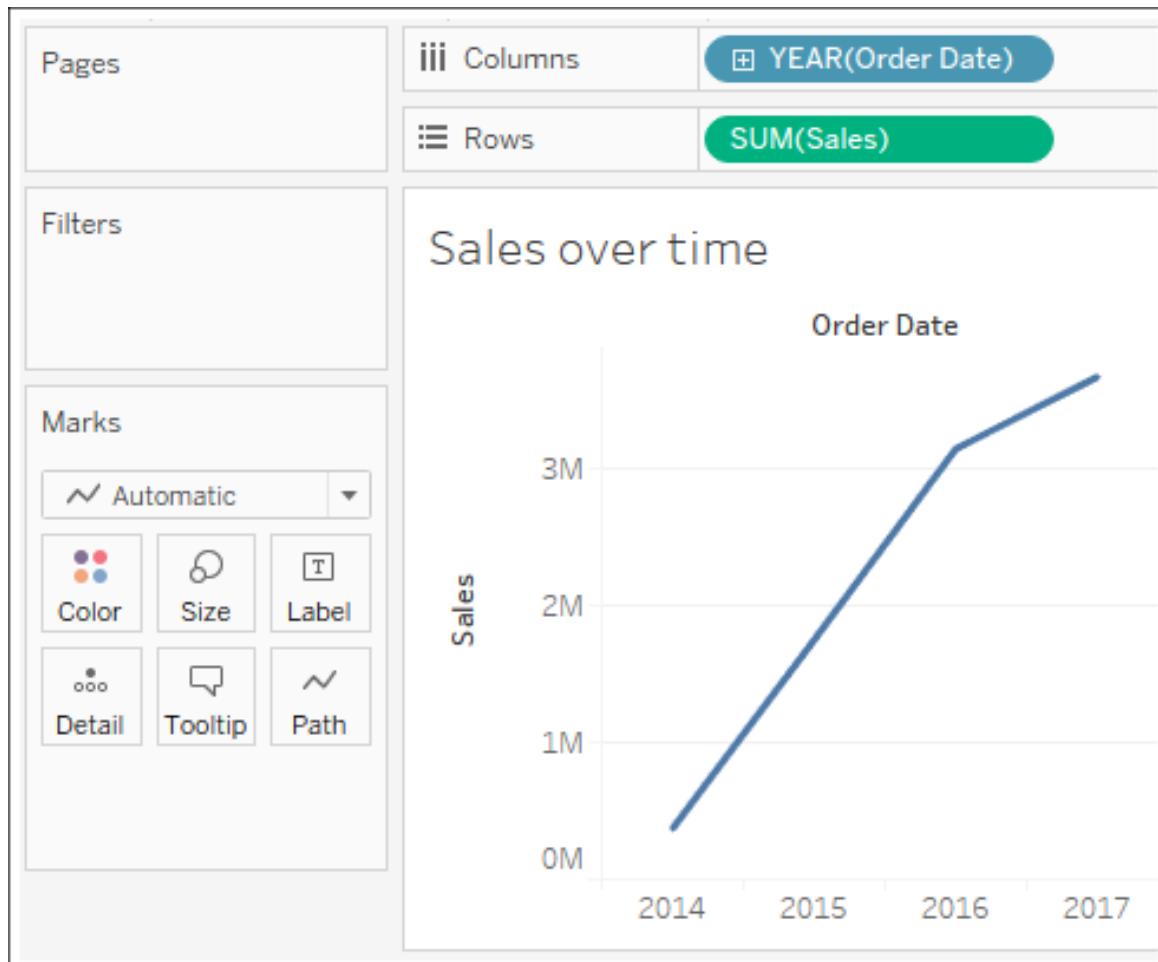


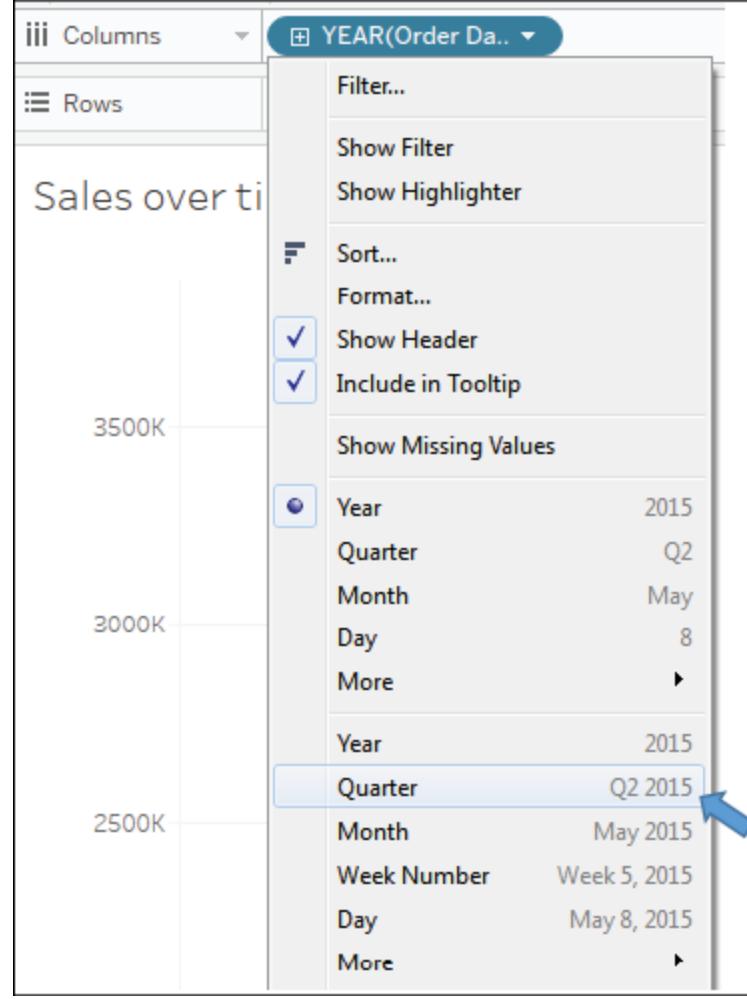
Iterations of bar charts for deeper analysis

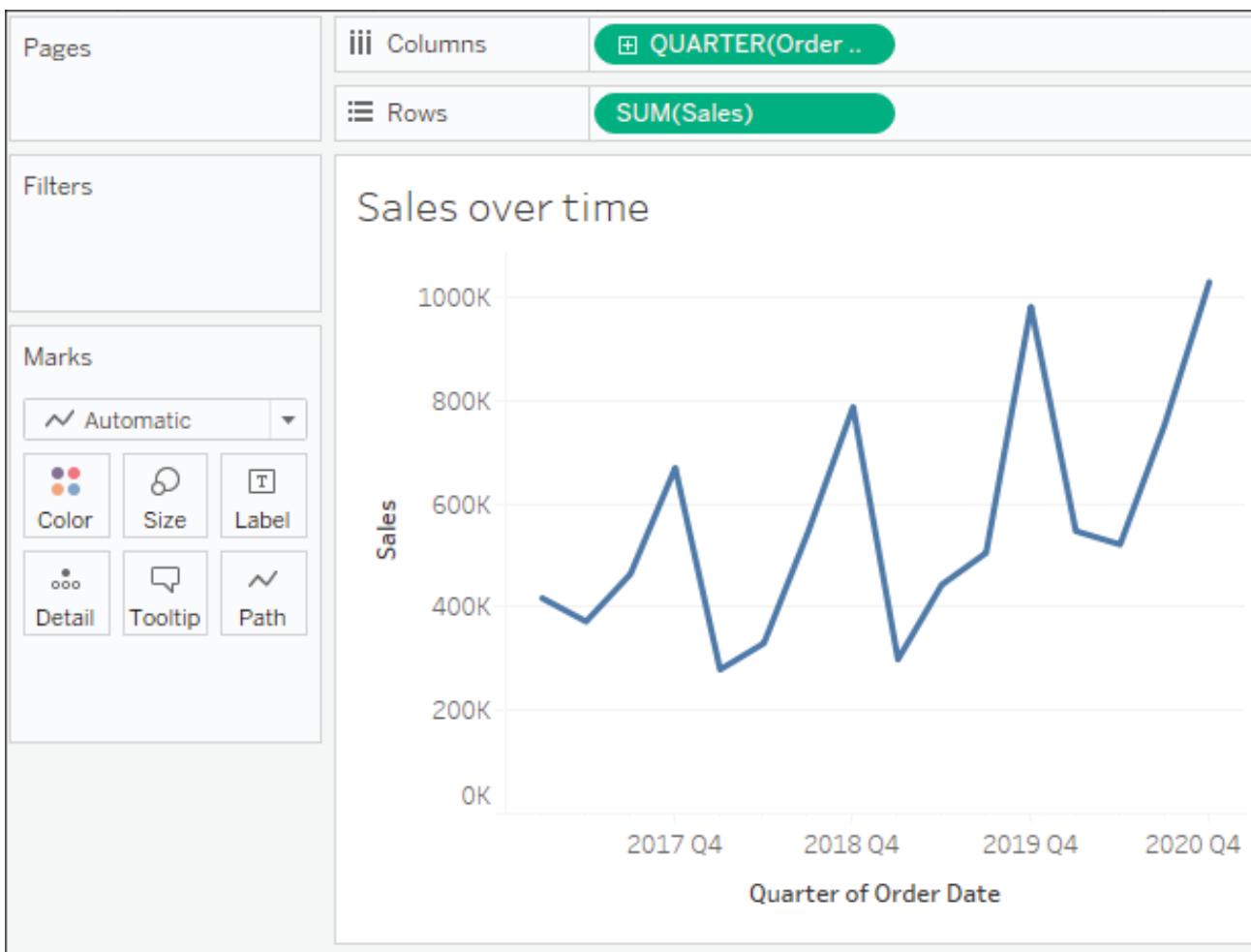


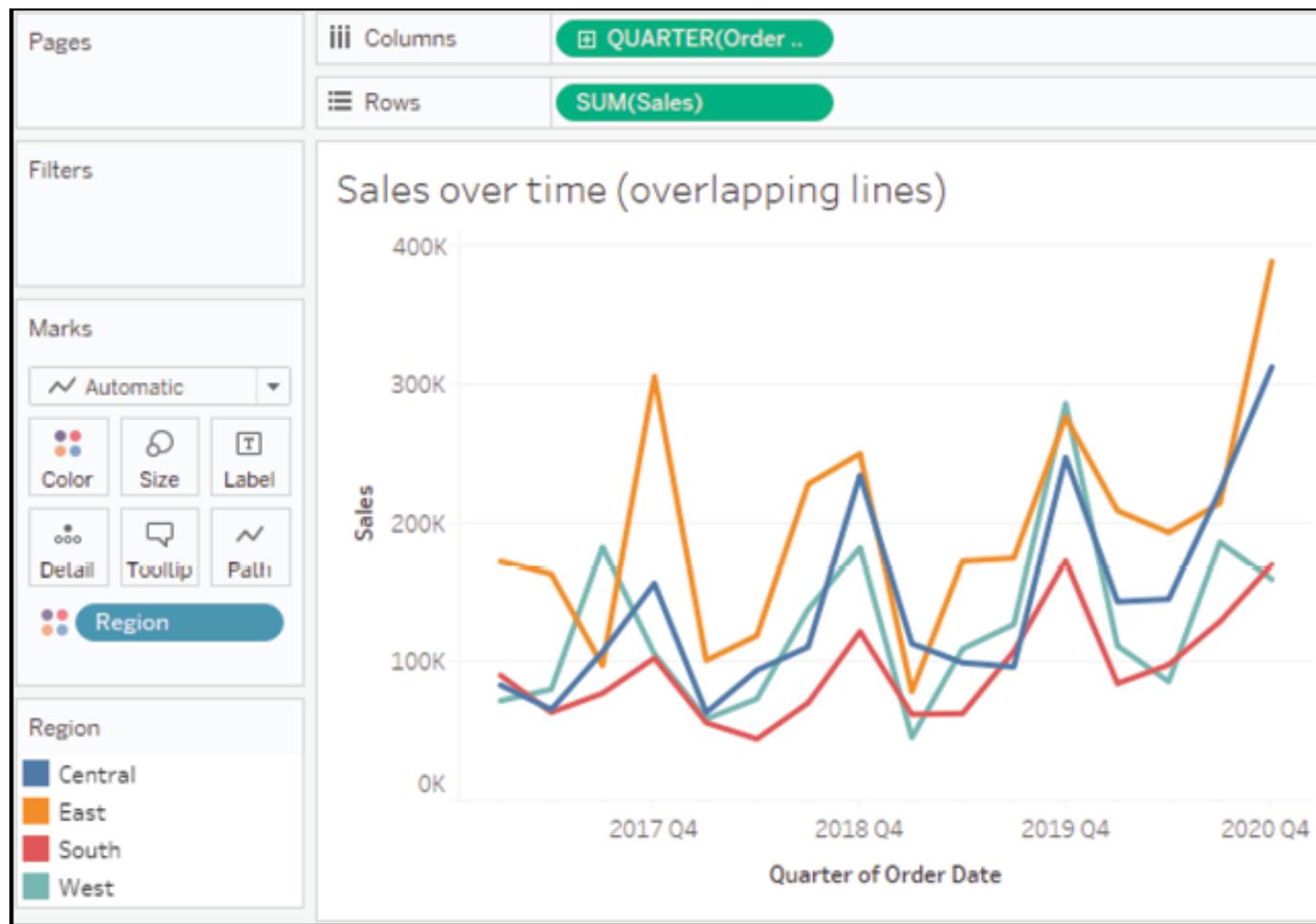
Line charts

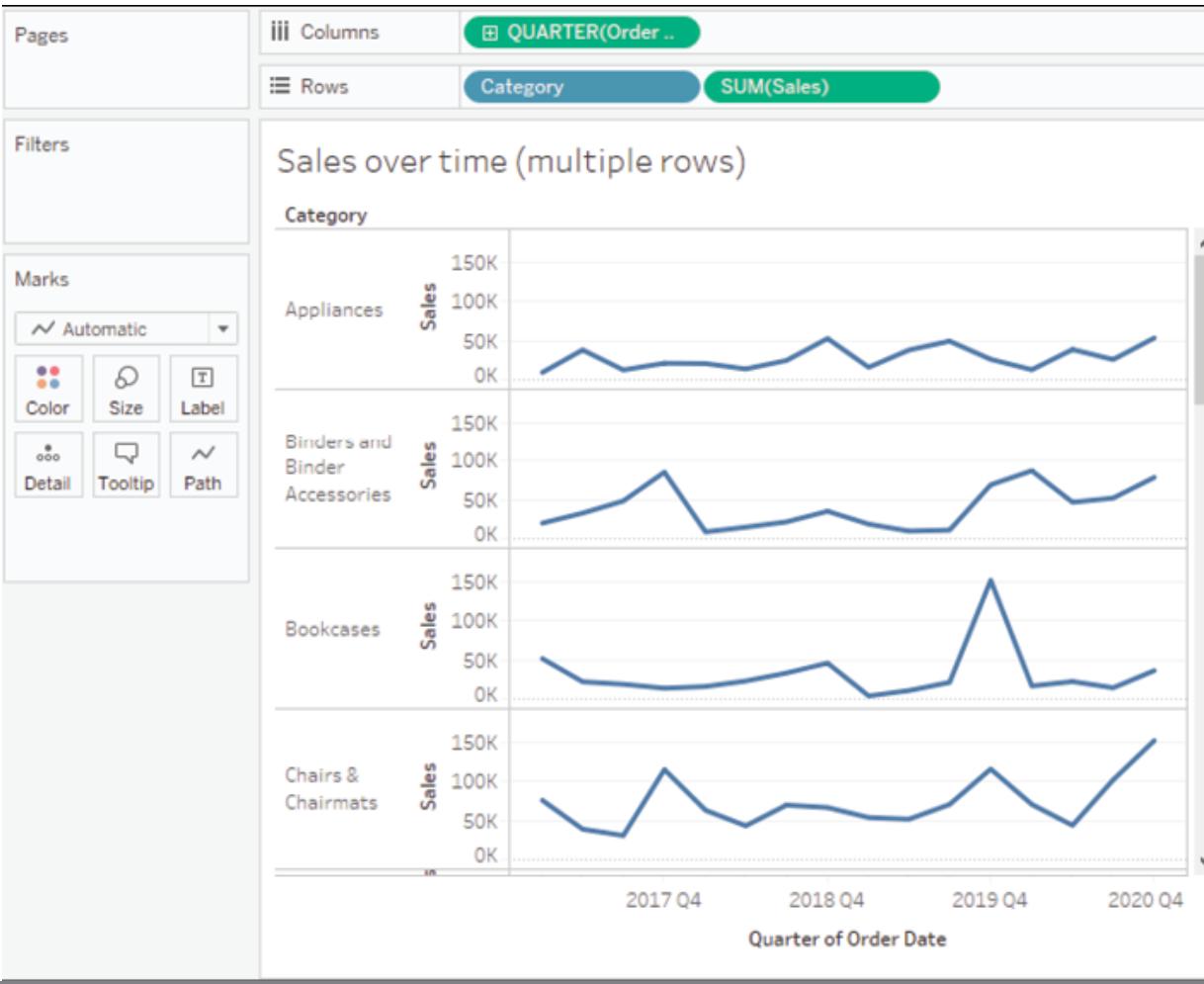
- Line charts connect related marks in a visualization to show movement or a relationship between those connected marks.
- The position of the marks and the lines that connect them are the primary means of communicating the data.
- Additionally, you can use size and color to communicate additional information.











Geographic visualizations

- In Tableau, the built-in geographic database recognizes geographic roles for fields such as Country, State, City, Airport, Congressional District, or Zip Code.
- Even if your data does not contain latitude and longitude values, you can simply use geographic fields to plot locations on a map.

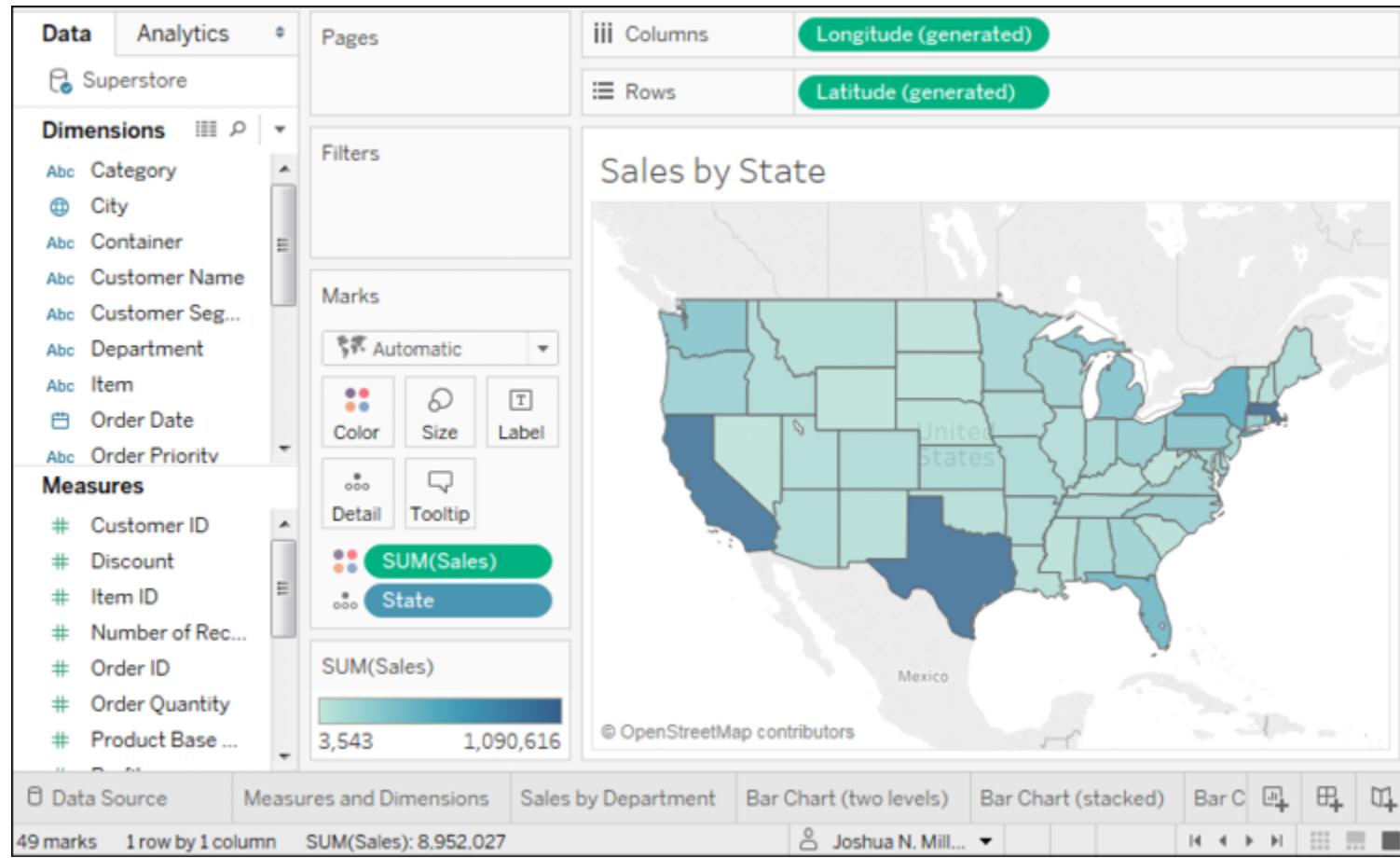
Geographic visualizations

Geographic visualization is incredibly valuable when you need to understand where things happen and whether there are any spatial relationships within the data. Tableau offers several types of geographic visualization:

- Filled maps
- Symbol maps
- Density maps

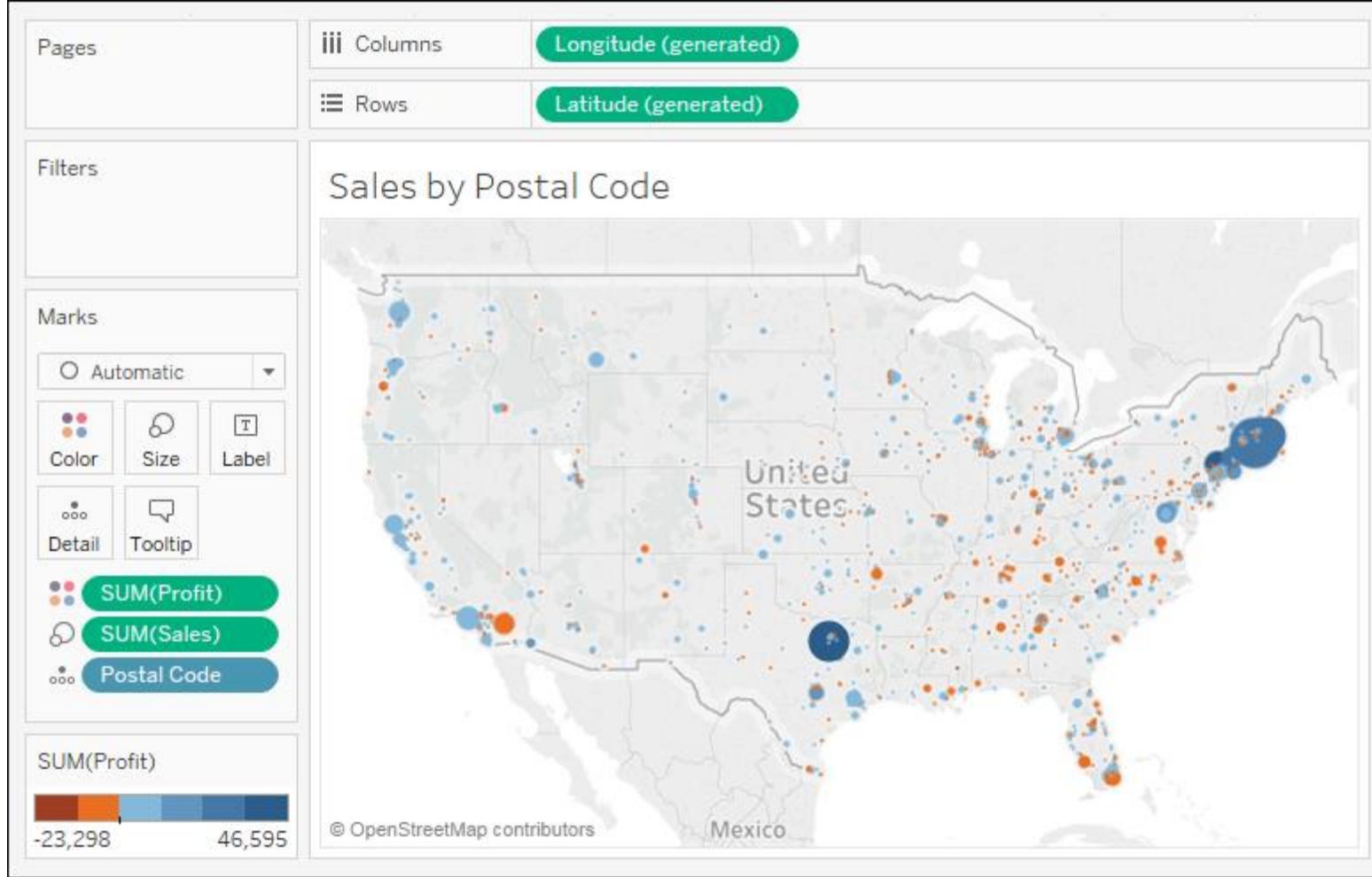
Filled maps

- Filled maps fill areas such as countries, states, or ZIP codes to show a location.
- The color that fills the area can be used to communicate measures such as average sales or population as well as dimensions such as region.
- These maps are also called choropleth maps.



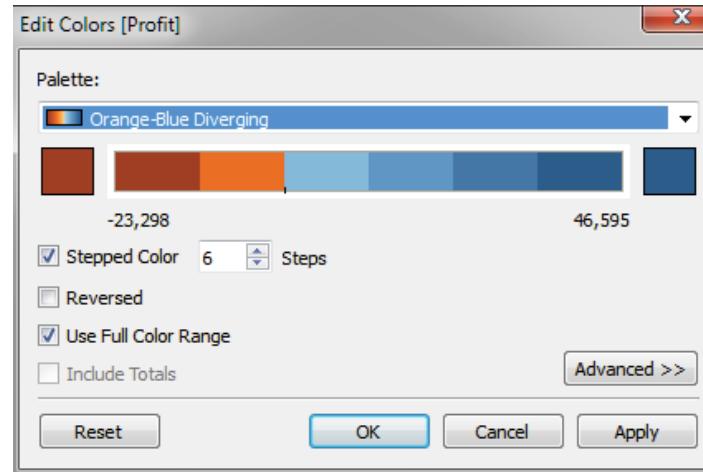
Symbol maps

- With symbol maps, marks on the map are not drawn as filled regions; rather, marks are shapes or symbols placed at specific geographic locations.
- The size, color, and shape may also be used to encode additional dimensions and measures.



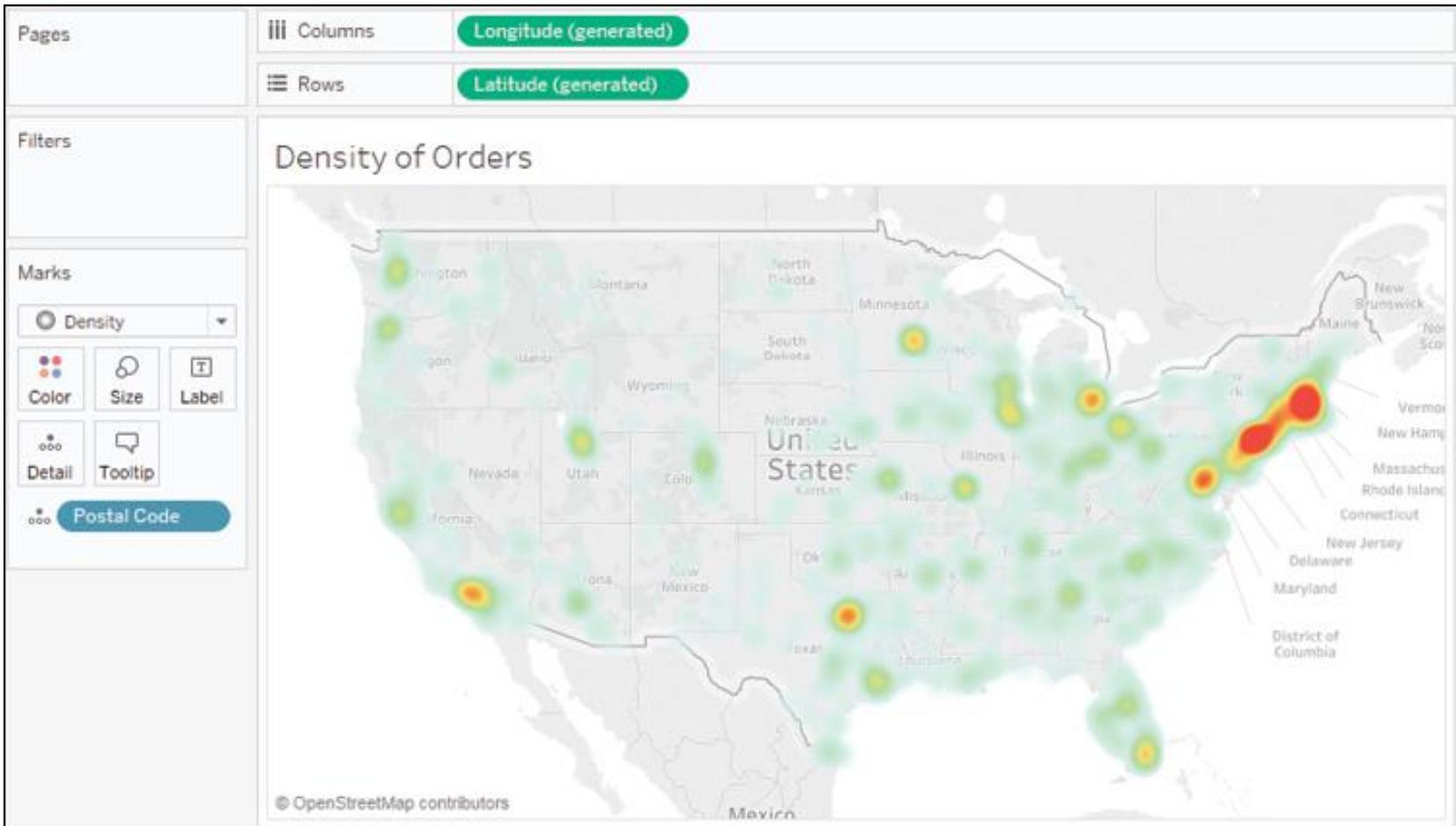
Symbol maps

- A combination of tweaking the size and using Stepped Color and Use Full Color Range, as shown here, produced the result for this example:



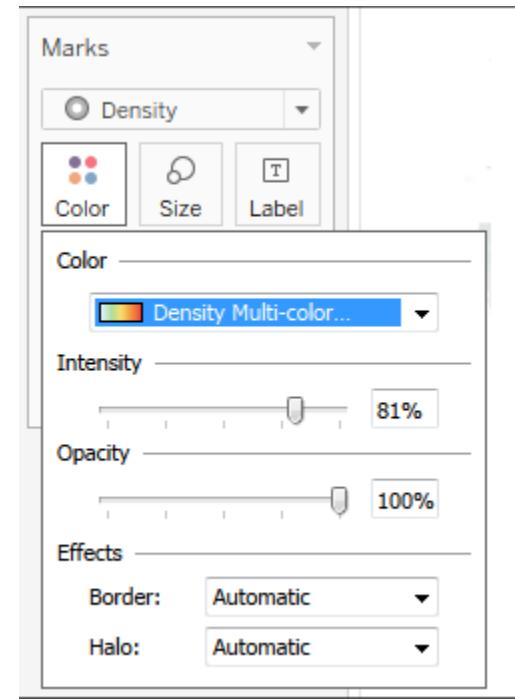
Density maps

- Density maps show the spread and concentration of values within a geographic area.
- Instead of individual points or symbols, the marks blend together, showing greater intensity in areas with a high concentration.
- You can control the Color, Size, and Intensity.



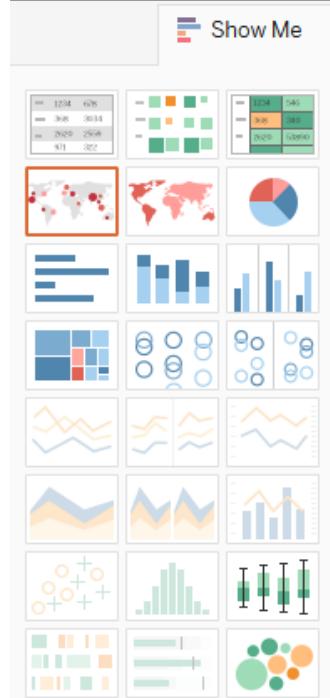
Density maps

- Try experimenting with the Color and Size options.
- Clicking on Color, for example, reveals some options specific to the Density mark type:



Using Show Me

- Show Me is a powerful component of Tableau that arranges selected and active fields into the places required for the selected visualization type.
- The Show Me toolbar displays small thumbnail images of different types of visualizations, allowing you to create visualizations with a single click.



For **symbol maps** try

1 geo Dimension

0 or more Dimensions

0 to 2 Measures

May use spatial measure in place of geo dimension

Putting everything together in a dashboard

- Often, you'll need more than a single visualization to communicate the full story of the data.
- In these cases, Tableau makes it very easy for you to use multiple visualizations together on a dashboard.
- In Tableau, a dashboard is a collection of views, filters, parameters, images, and other objects that work together to communicate a data story.

Dashboard Layout

Default Phone

Device Preview

Size

Desktop Browser (1000 x 8... ▾)

Sheets

- Measures and Dime...
- Sales by Department
- Bar Chart (two level...)
- Bar Chart (stacked)
- Bar Chart (experim...
- Sales over time
- Sales over time (ov...
- Sales over time (m...
- Sales by State
- Sales by Postal Code
- Density of Orders
- Show Me

Objects

	Horizontal		Web Page
	Vertical		Blank
	Text		Button
	Image		Extension

Tiled Floating

Show dashboard title

The dashboard interface

Grip

Sales by Department

Department

Department	Sales
Furniture	~3M
Office Supplies	~2.5M
Technology	~3M

0M 1M 2M 3M

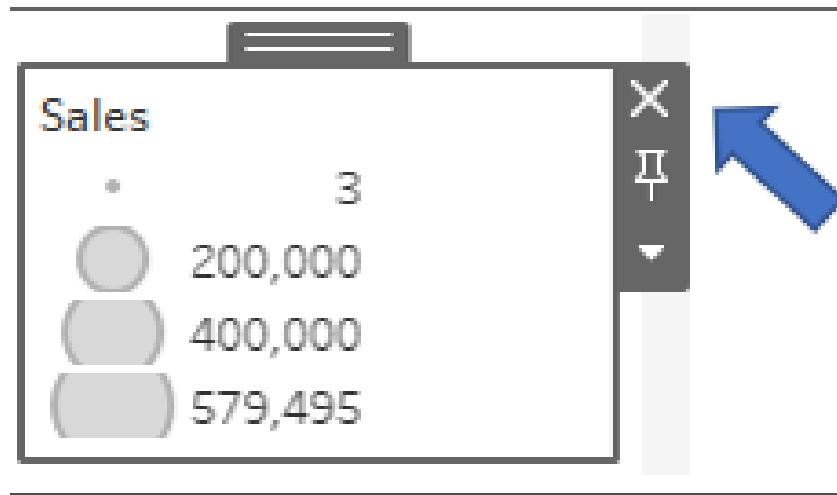
Sales

- Remove from dashboard
- Navigate to the sheet
- Use the sheet as a filter on the dashboard
- Open the dropdown menu

- Go to Sheet
- Duplicate Sheet
- Fit
- Title
- Caption
- Legends
 - Color Legend (Sales)
 - Shape Legend
 - Size Legend (Sales)
 - Map Legend
- Filters
- Highlighters
- Show Page Control
- View Toolbar
- Use as Filter
- Ignore Actions
- Floating
 - Floating Order
- Deselect
- Remove from Dashboard
- Rename Dashboard Item...

Building your dashboard

- Select the Sales size legend by clicking on it. Use the X option to remove the legend from the dashboard:



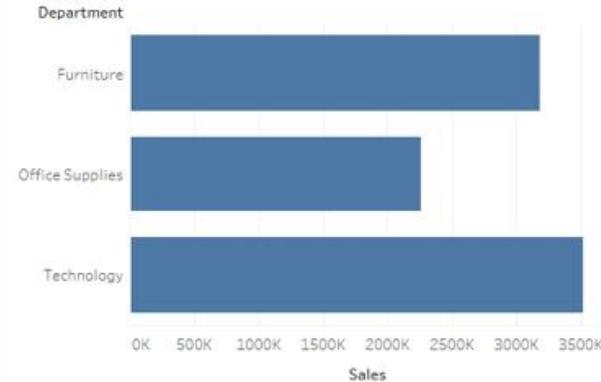
Building your dashboard

- For each view (Sales by Department, Sales by Postal Code, and Sales over time), select the view by clicking on an empty area in the view.
- Then, click on the Use as Filter option to make that view an interactive filter for the dashboard:

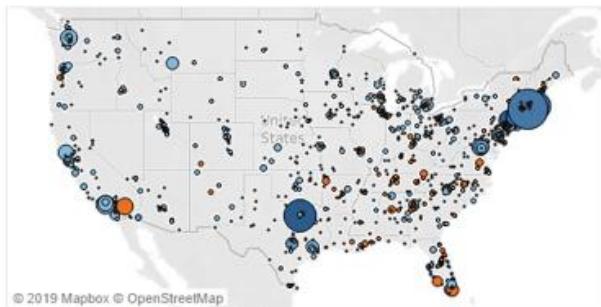


Superstore Sales

Sales by Department



Sales by Postal Code

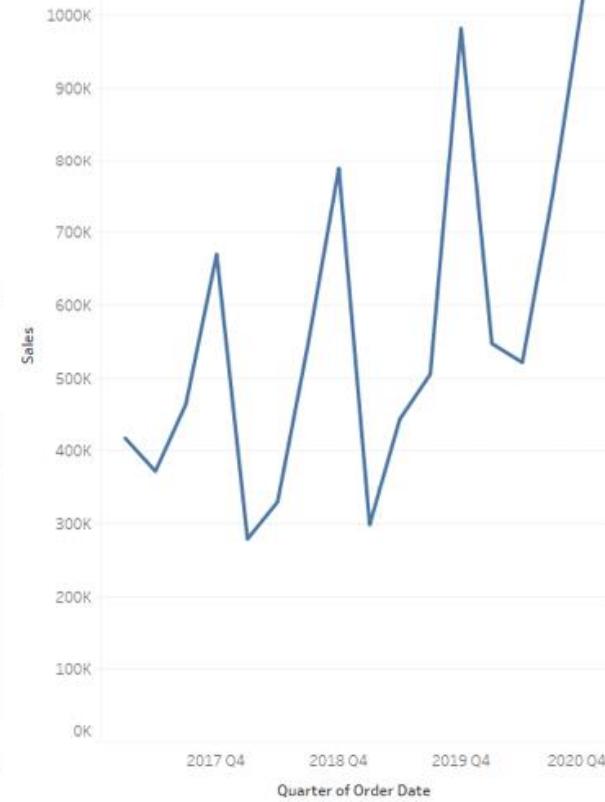


Profit



-23,298 | 46,595

Sales over time



Summary

- Tableau's visual environment allows a rapid and iterative process of exploring and analyzing data visually.
- You've taken your first steps toward understanding how to use the platform.
- You connected to data and then explored and analyzed the data using some key visualization types such as bar charts, line charts, and geographic visualizations.

COMPLETE LAB 1

2. Connecting to Data in Tableau

Connecting to Data in Tableau

We'll cover the following topics:

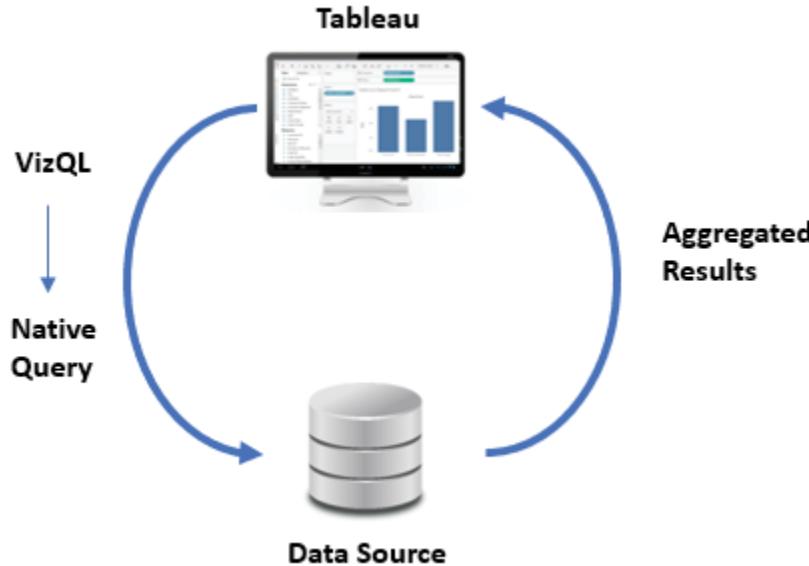
- The Tableau paradigm
- Connecting to data
- Managing data source metadata
- Working with extracts instead of live connections
- Filtering data

The Tableau paradigm

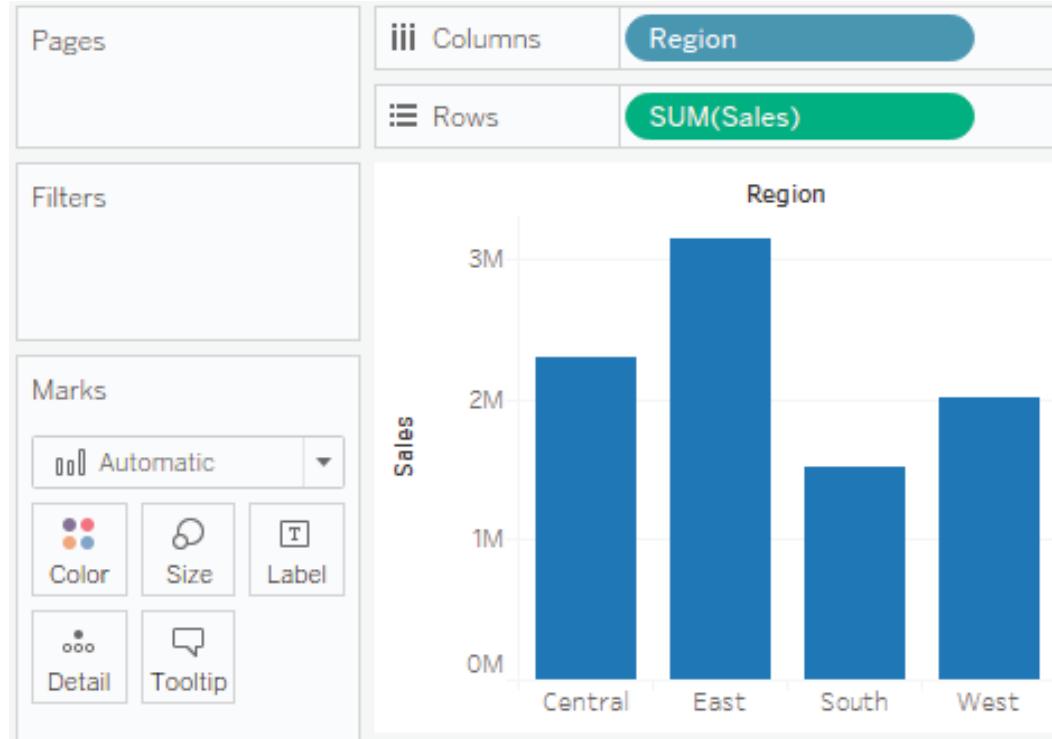
- The unique and exciting experience of working with data in Tableau is a result of VizQL (Visual Query Language).
- VizQL was developed as a Stanford University research project, focusing on the natural ways that humans visually perceive the world and how that could be applied to data visualization.
- We naturally perceive differences in size, shape, spatial location, and color.

The Tableau paradigm

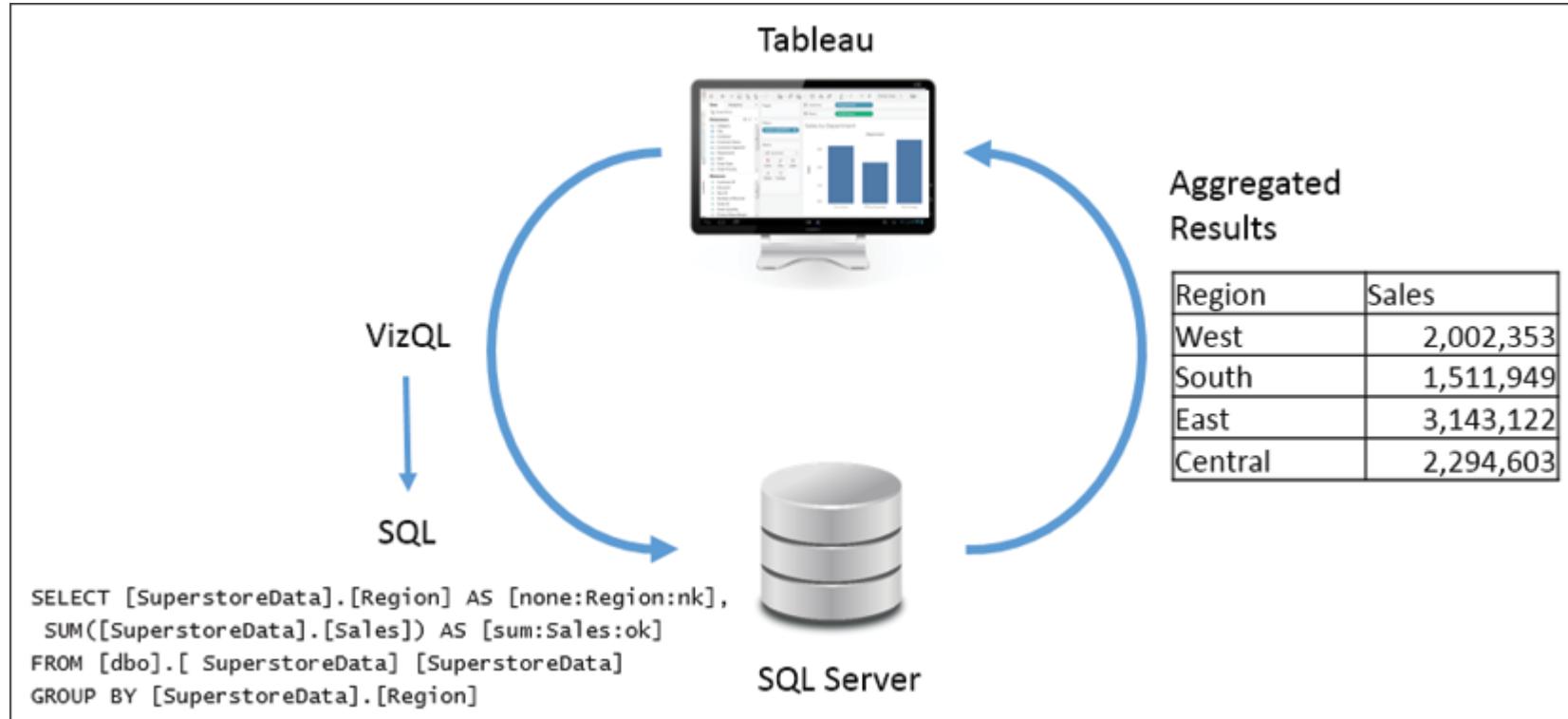
- In its simplest form, the Tableau paradigm of working with data looks like the following diagram:



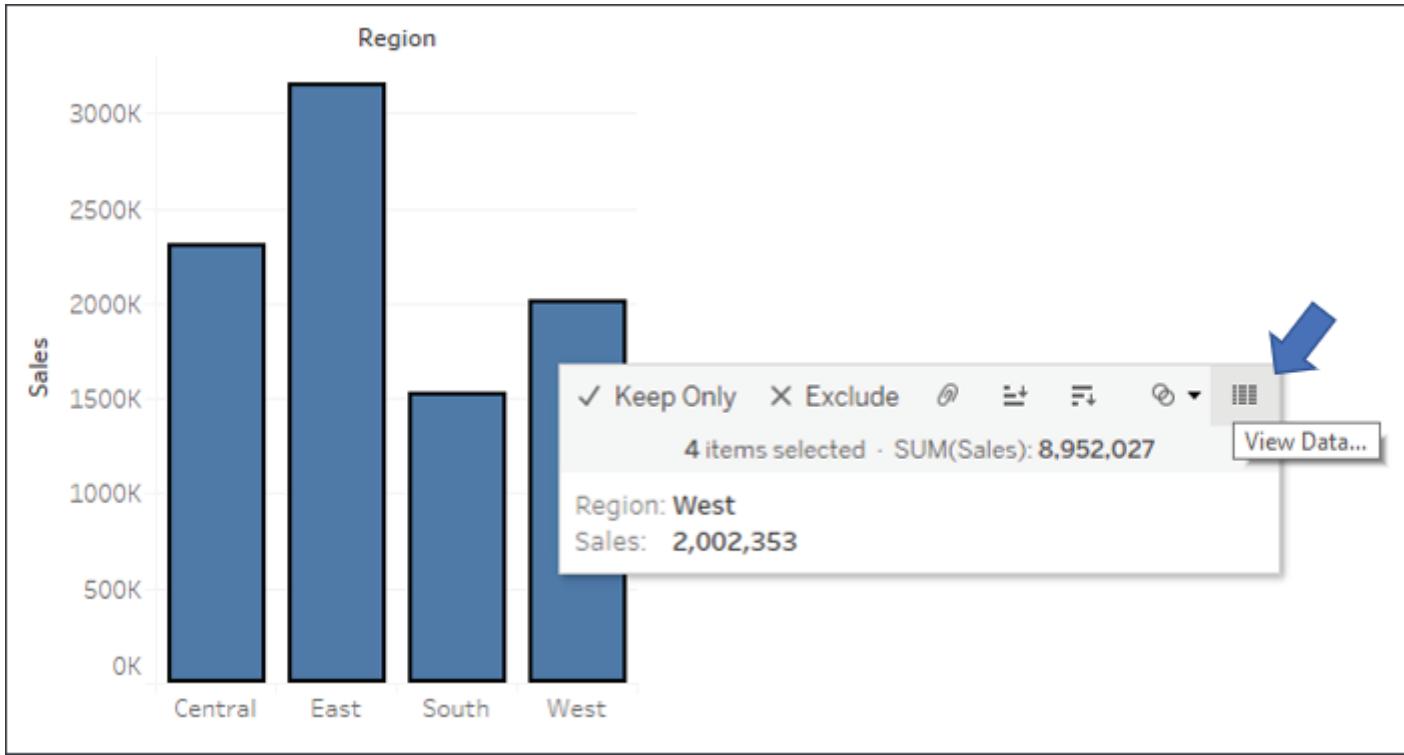
A simple example



A simple example



A simple example



A simple example

- This will reveal a View Data window:

The screenshot shows a modal window titled "View Data: Tableau Paradigm". At the top left is a checkbox labeled "Show aliases" which is checked. To its right are two buttons: "Copy" and "Export All". The main area contains a table with four rows of data. The columns are labeled "Region" and "Sales". The data is as follows:

Region	Sales
West	2,002,353
South	1,511,949
East	3,143,122
Central	2,294,603

At the bottom of the window, there are three buttons: "Summary", "Full Data", and "4 rows".

A simple example

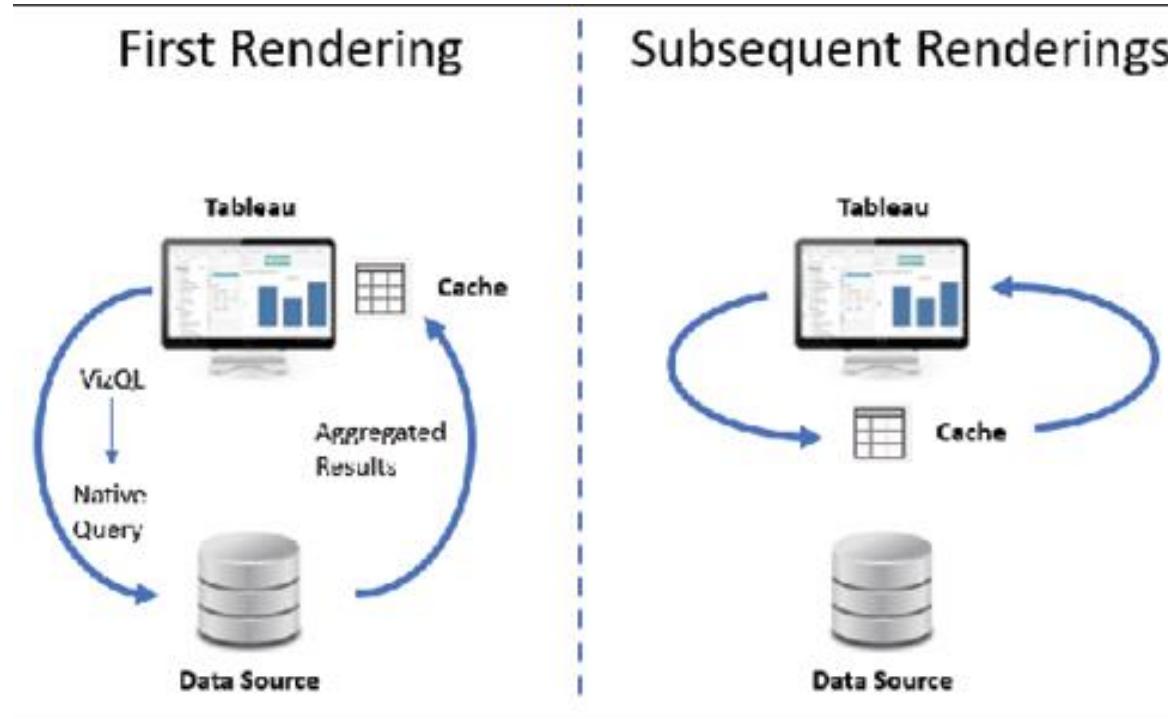
- In this case, there are 9,426 underlying records, as indicated on the status bar in the lower-right corner of the following screenshot:

The screenshot shows a software interface titled "View Data". At the top, it displays "9,426 rows" and has checkboxes for "Show aliases" and "Show all fields", along with "Copy" and "Export All" buttons. The main area is a table with the following data:

Category	City	Container	Customer ID	Customer Name
Appliances	Lancaster	Small Box	1890	Jonathan Drummey
Paper	Oxford	Small Box	607	Joe Mako
Paper	Boston	Small Box	608	Chuck Hooper
Paper	Conway	Small Box	3073	Shawn Wallwork

At the bottom, there are buttons for "Summary" and "Full Data", and the status bar shows "9,426 rows".

A simple example



Connecting to data

- There is virtually no limit to the data that Tableau can visualize! Almost every new version of Tableau adds new native connectors.
- Tableau continues to add native connectors for cloud-based data.
- The web data connector allows you to write a connector for any online data you wish to retrieve.
- The Tableau Hyper API allows you to programmatically read and write extracts of data, enabling you to access data from any source and write it to a native Tableau format.

Connecting to data in a file

File-based data includes all sources of data where the data is stored in a file. File-based data sources include the following:

- Extracts: A .hyper or .tde file containing data that was extracted from an original source.
- Microsoft Access: An .mdb or .accdb database file created in Access.
- Microsoft Excel: An .xls, .xlsx, or .xlsm spreadsheet created in Excel. Multiple Excel sheets or sub-tables may be joined or unioned together in a single connection.

1 2 **Connections** Add
Superstore Microsoft Excel

3 **Sheets**
Orders Returns

4 **Orders and Returns**

5 Orders ————— Returns

6 Connection
 Live Extract

7 Filters 0 | Add

8 Sort fields Data source order ▾ Show aliases Show hidden fields 1,000 ↗ rows

9

Category	City	Container	Customer ID	Customer Name	Customer Segment
Paper	Ponca City	Small Box	3035	Larry Harris	Home Office
Paper	Ponca City	Wrap Bag	3035	Larry Harris	Home Office
Pens & Art Supplies	Stillwater	Wrap Bag	3385	J.B. Bond	Corporate
Binders and Binder Ac...	Desoto	Small Box	3133	Kurt Krohn	Corporate
Rubber Bands	Desoto	Wrap Bag	3133	Kurt Krohn	Corporate
Storage & Organization	Argyle	Small Box	1697	Mark Piland	Home Office
Tables	Grand Prairie	Jumbo Box	1603	Bill Eubanks	Small Business

Pages

iii Columns

CNT>Returns

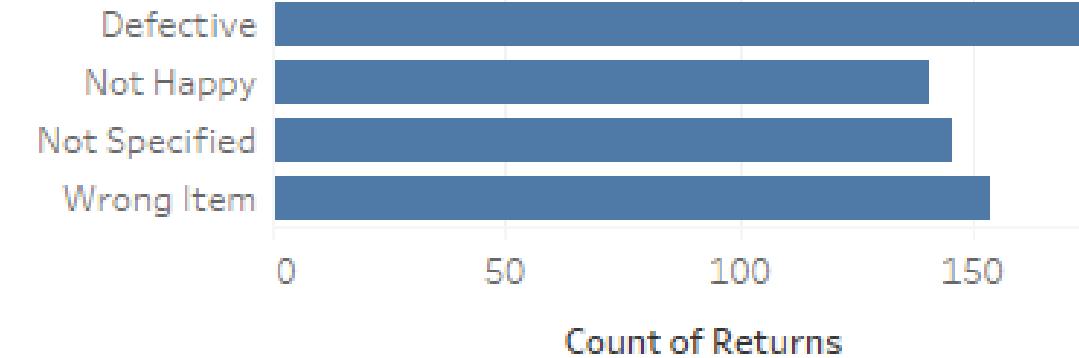
Rows

Return Reason

Filters

Connect to Excel

Return Reason



Marks

Automatic



Color



Size



Label



Detail



Tooltip

Connecting to data on a server

- Database servers, such as SQL Server, Snowflake, Vertica, and Oracle, host data on one or more server machines and use powerful database engines to store, aggregate, sort, and serve data based on queries from client applications.
- Tableau can leverage the capabilities of these servers to retrieve data for visualization and analysis.

- As soon as the Microsoft SQL Server connection is selected, the interface displays options for some initial configuration as follows:

Microsoft SQL Server

Server: TDS-W541-JM\AGAPE

Database: Optional

Enter information to sign in to the database:

Use Windows Authentication (preferred)

Use a specific username and password:

Username:

Password:

Require SSL

Read uncommitted data

Initial SQL... Sign In

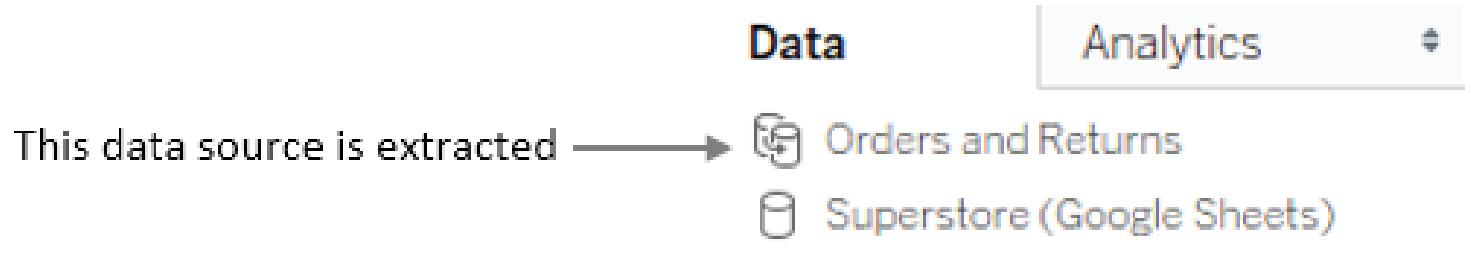
Connecting to data on a server

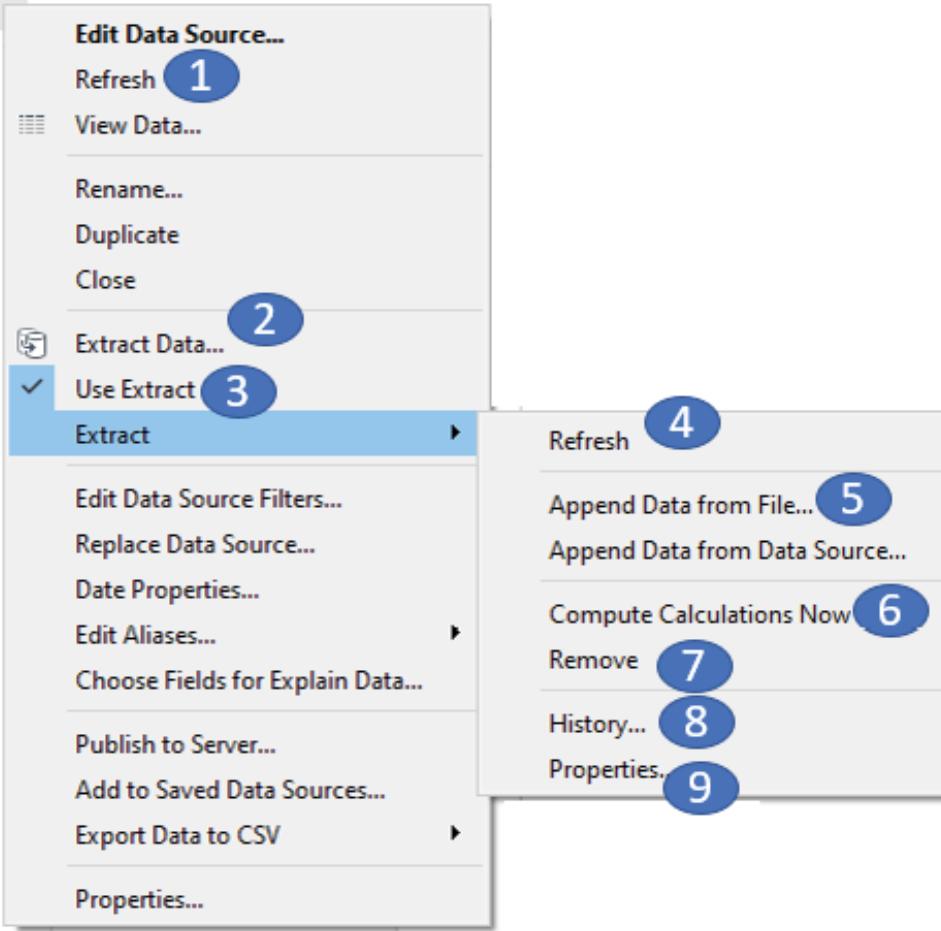
- Once you click the orange Sign In button, you will see a screen that is very similar to the connection screen you saw for Excel.
- The main difference is on the left, where you have an option for selecting a Database, as shown in the following screenshot:

The screenshot shows a software interface for managing database connections. At the top, there's a 'Connections' section with an 'Add' button. Below it, a list shows a single entry: 'TDS-W541-JM\TDS Microsoft SQL Server'. Underneath this, there's a 'Database' section with a dropdown menu set to 'Hospital'. On the left side, there's a tree view labeled 'Table' containing various database objects: Discharge Details, Hospital Visit, Hospital_Visits, Patient, Patient Visit, Primary Physician, v_Patients, and v_Visits. Below the table section are two more items: 'New Custom SQL' and 'New Union'. At the bottom, there's a 'Stored Procedures' section listing three entries: Get_Active_Doctors, Get_Patient_Diagnoses, and Get_Patient_Treatments.

Using extracts

- Any data source that is using an extract will have a distinctive icon that indicates the data has been pulled from an original source into an extract, as shown in the following screenshot:





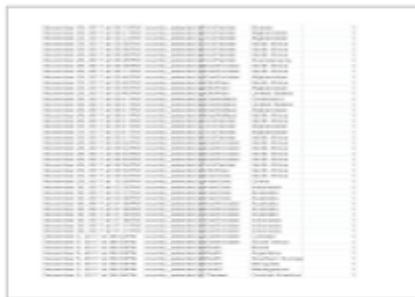
Connecting to data in the cloud

- Certain data connections are made to data that is hosted in the cloud.
- These include Amazon RDS, Google BigQuery, Microsoft SQL Azure, Snowflake, Salesforce, Google Sheets, and many others

Select Your Google Sheet

Signed in as milligan. [Sign Out](#)

Name	Owned by	Last Opened By Me
Greek New Testament	Joshua Milligan	Mar 12, 2018
Polygonic Hex Map.xlsx	Joshua Milligan	Dec 20, 2017
SMS received	Joshua Milligan	Sep 25, 2017
Teknion Sign Up Sheet August – December 2017		Aug 11, 2017
Superstore	Joshua Milligan	Jul 10, 2017
Company Profit	Joshua Milligan	Jun 27, 2017
Test_1	Joshua Milligan	Dec 5, 2017
SMS received	Joshua Milligan	Sep 25, 2017
Assimilations	Borg Queen	Aug 11, 2017
Superstore	Joshua Milligan	Jul 10, 2017
Company Profit	Joshua Milligan	Jun 27, 2017



Greek New Testament

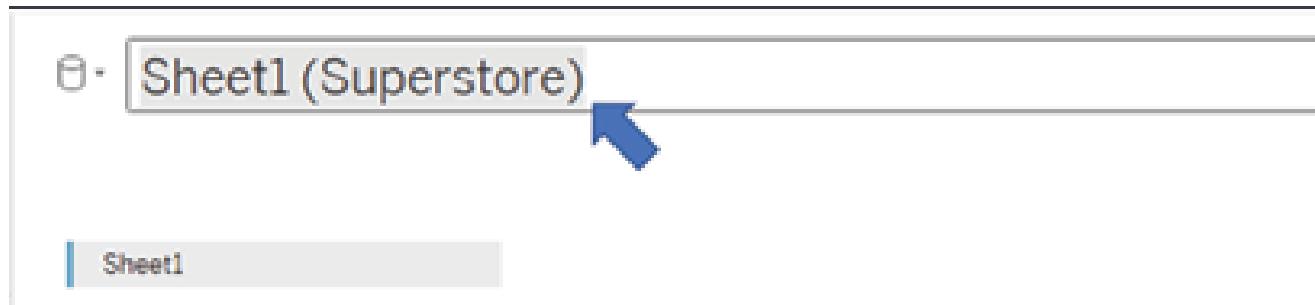
Last Modified On Dec 7, 2017

Last Modified By Joshua Milligan

[Open in Google Drive](#)

Connecting to data in the cloud

- Click the Data Source name to rename it to Superstore (Google Sheets):

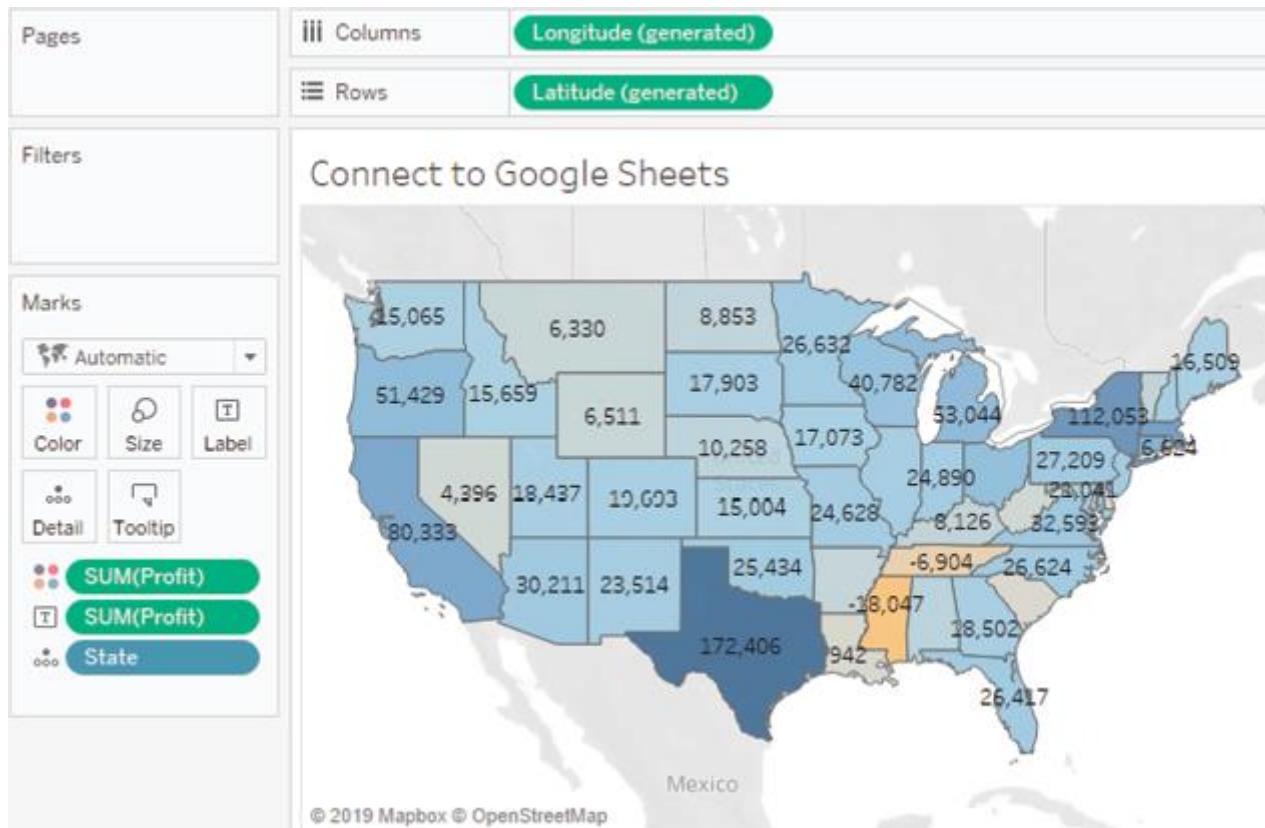


Connecting to data in the cloud

- For the purpose of this example, switch the connection option from Live to Extract.
- When connecting to your own Google Sheets data, you may choose either Live or Extract:



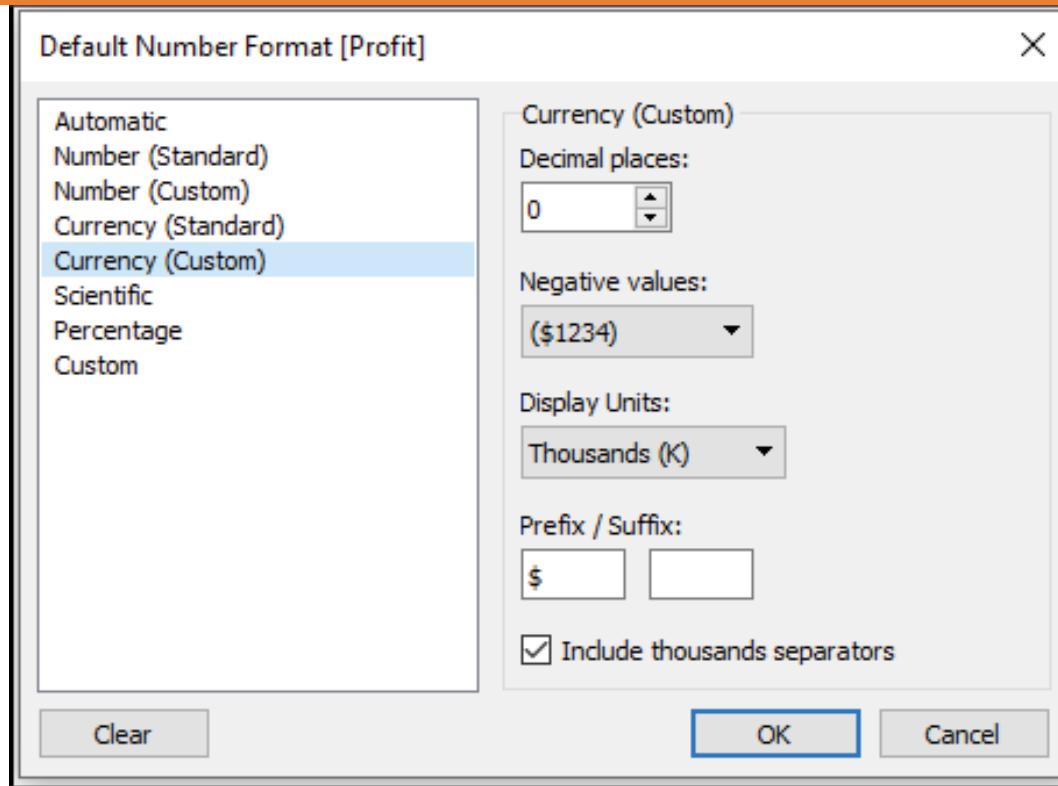
- Create a filled map of Profit by State, with Profit defining the Color and Label:



Managing data source metadata

- Data sources in Tableau store information about the connection(s).
- In addition to the connection itself (for example, database server name, database, and/or filenames), the data source also contains information about all the fields available (such as field name, data type, default format, comments, and aliases).
- Often, this data about the data is referred to as metadata.

Managing data source metadata



Pages

iii Columns

Longitude (generated)

Rows

Latitude (generated)

Filters

Connect to Google Sheets

Marks

Automatic

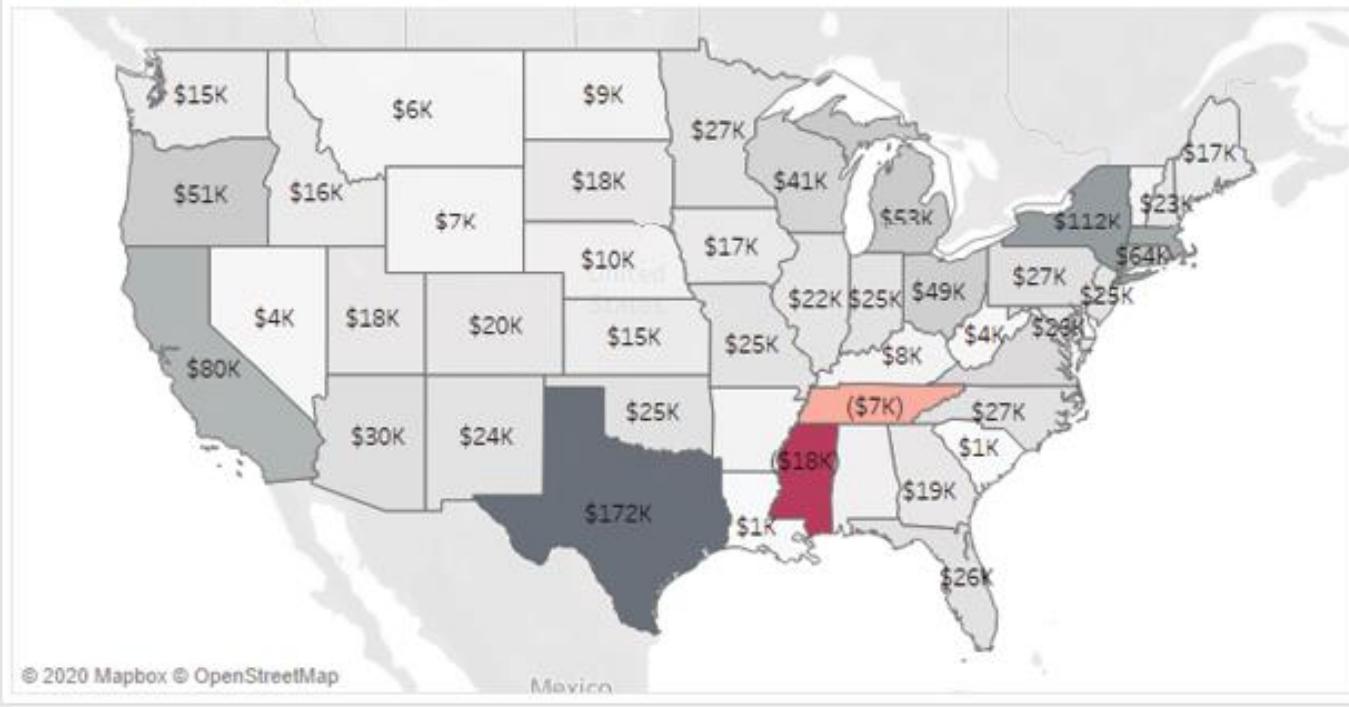
Color
Size
LabelDetail
Tooltip

SUM(Profit)

SUM(Profit)

State

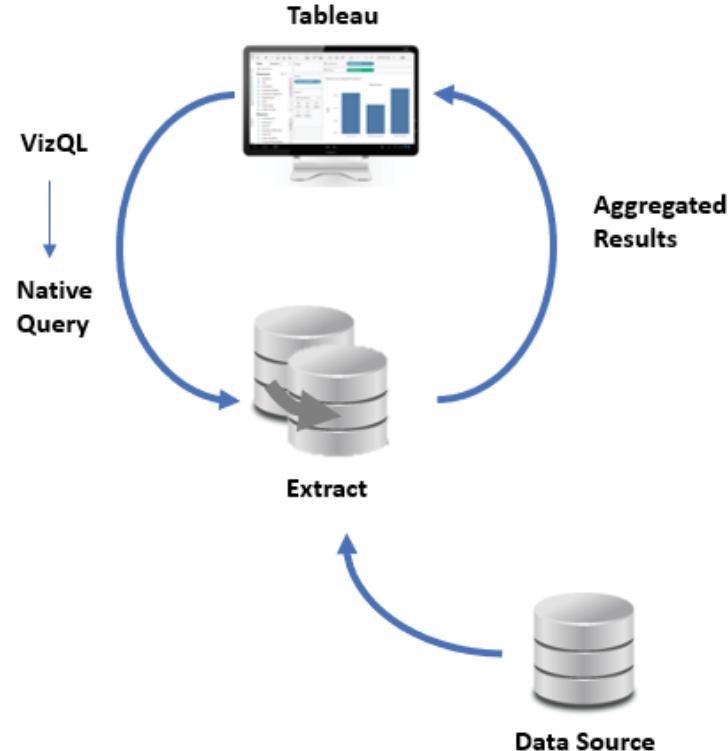
SUM(Profit)



© 2020 Mapbox © OpenStreetMap

Working with extracts instead of live connections

- Extracts extend the way in which Tableau works with data.
- Consider the following diagram:



Creating extracts

- Select Extract on the Data Source screen as follows.
The Edit... link will allow you to configure the extract:

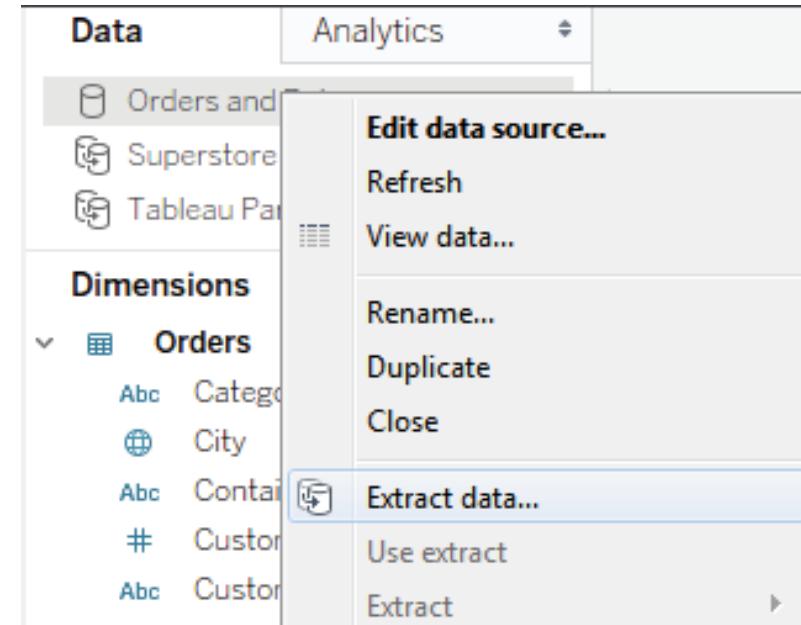
Connection

[Edit...](#)

Extract will include all data.

Creating extracts

- Select the data source from the Data menu, or right-click the data source on the data pane and select Extract data.
- You will be given a chance to set configuration options for the extract, as demonstrated in the following screenshot:



Extract Data X

Specify how to store data in the extract:

Data Storage

 Logical Tables Physical TablesStore data using one table for each logical table. [Learn more](#)

Use this option if you need to use extract filters, aggregation, or other extract settings.

Specify how much data to extract:

Filters (optional)

Filter	Details
Region	keeps Central and South
Category	keeps Office Machines

[Add...](#)[Edit...](#)[Remove](#)

Aggregation

 Aggregate data for visible dimensions Roll up dates to Year

Number of Rows

 All rows Incremental refresh Top: Orders by rows Sample: Orders by rows[History...](#)[Hide All Unused Fields](#)[Extract](#)[Cancel](#)

Performance

There are two types of extracts in Tableau:

- Tableau Data Extracts (.tde files): prior to Tableau 10.5, these were the only type of extract available.
- Hyper (.hyper files) are available in Tableau 10.5 or later.

Portability and security

- Let's say that your data is hosted on a database server accessible only from inside your office network.
- Normally, you'd have to be onsite or using a VPN to work with the data.
- Even cloud-based data sources require an internet connection.
- With an extract, you can take the data with you and work offline.

When to use an extract

- You should consider various factors when determining whether to use an extract.
- In some cases, you won't have an option (for example, OLAP requires a live connection and some cloud-based data sources require an extract).
- In other cases, you'll want to evaluate your options.

Filtering data

- Data Source Filters are applied before all other filters and are useful when you want to limit your analysis to a subset of data. These filters are applied before any other filters.
- Extract Filters limit the data that is stored in an extract (.tde or .hyper). Data source filters are often converted into extract filters if they are present when you extract the data.

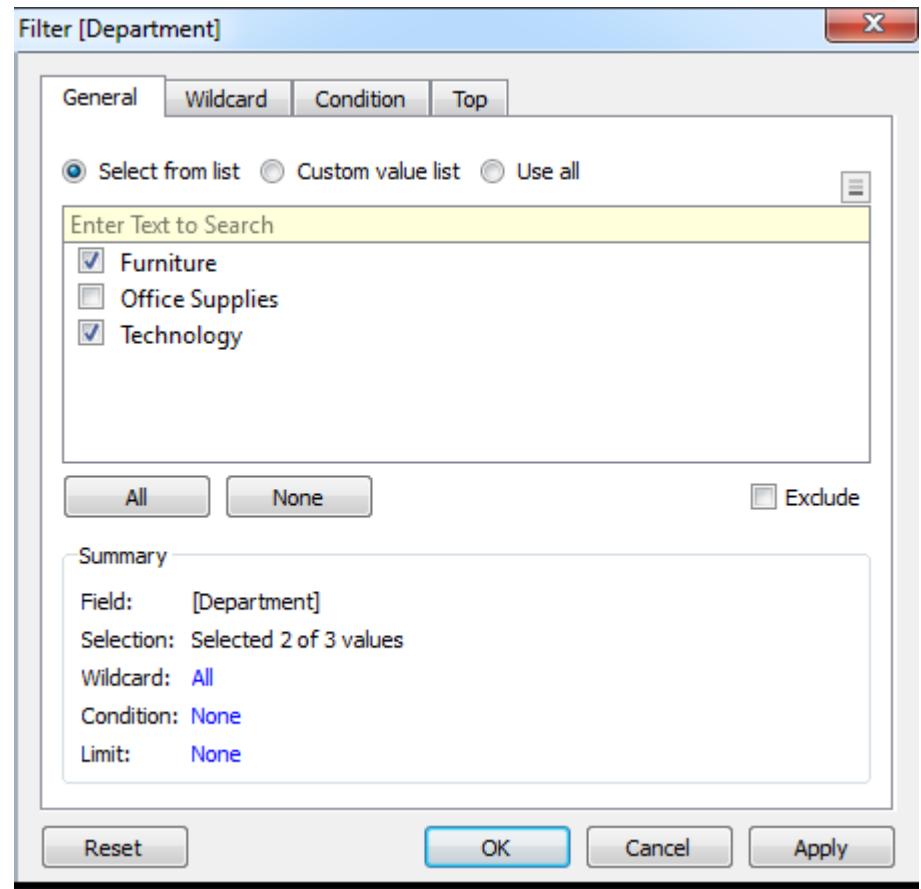
Filtering data

Additionally, you can apply filters to one or more views using one of the following techniques:

- Drag and drop fields from the data pane to the Filters shelf.
- Select one or more marks or headers in a view and then select Keep Only or Exclude, as shown here:

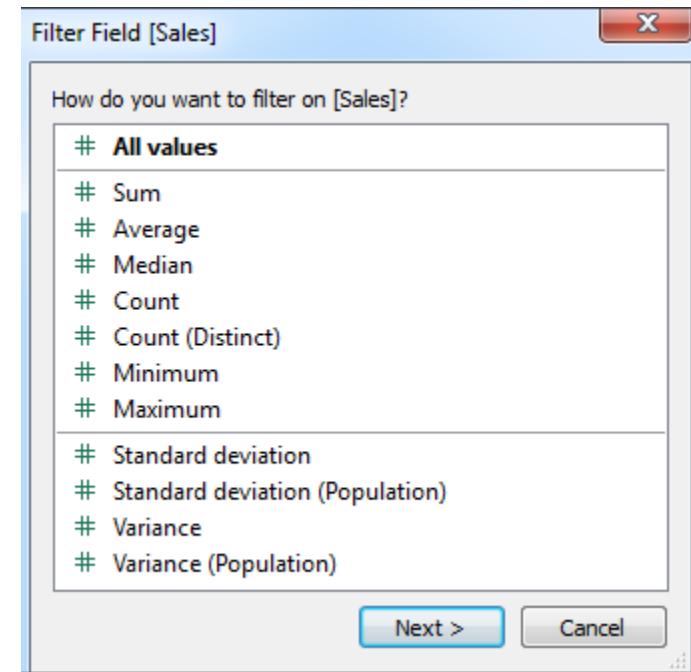


Filtering discrete (blue) fields



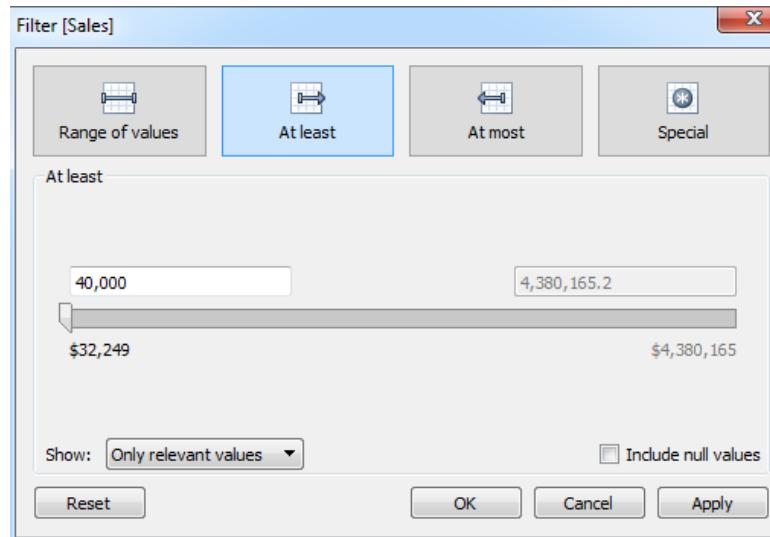
Filtering continuous (green) fields

- If you drop a continuous dimension onto the Filters shelf, you'll get a different set of options.
- Often, you will first be prompted as to how you want to filter the field, as follows:

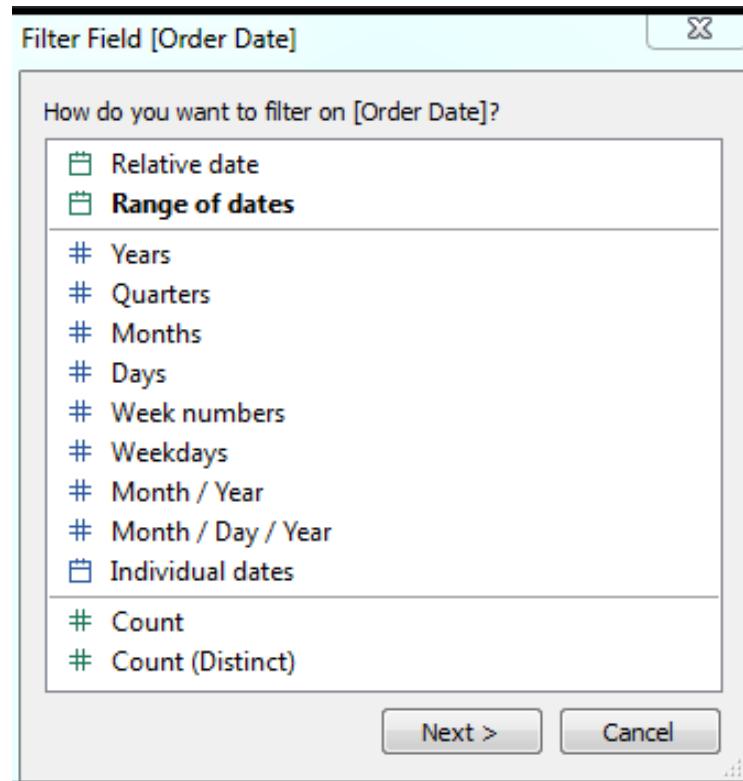


Filtering continuous (green) fields

- Once you've made a selection (or if the selection wasn't applicable for the field selected), you will be given another interface for setting the actual filter, as follows:



Filtering dates



Summary

- This lesson covered key concepts of how Tableau works with data.
- Although you will not usually be concerned with what queries Tableau generates to query underlying data engines, having a solid understanding of Tableau's paradigm will greatly aid you as you analyze data.

COMPLETE LAB 2

3. Moving Beyond Basic Visualizations

Moving Beyond Basic Visualizations

In this lesson, visualizations will fall under the following major categories:

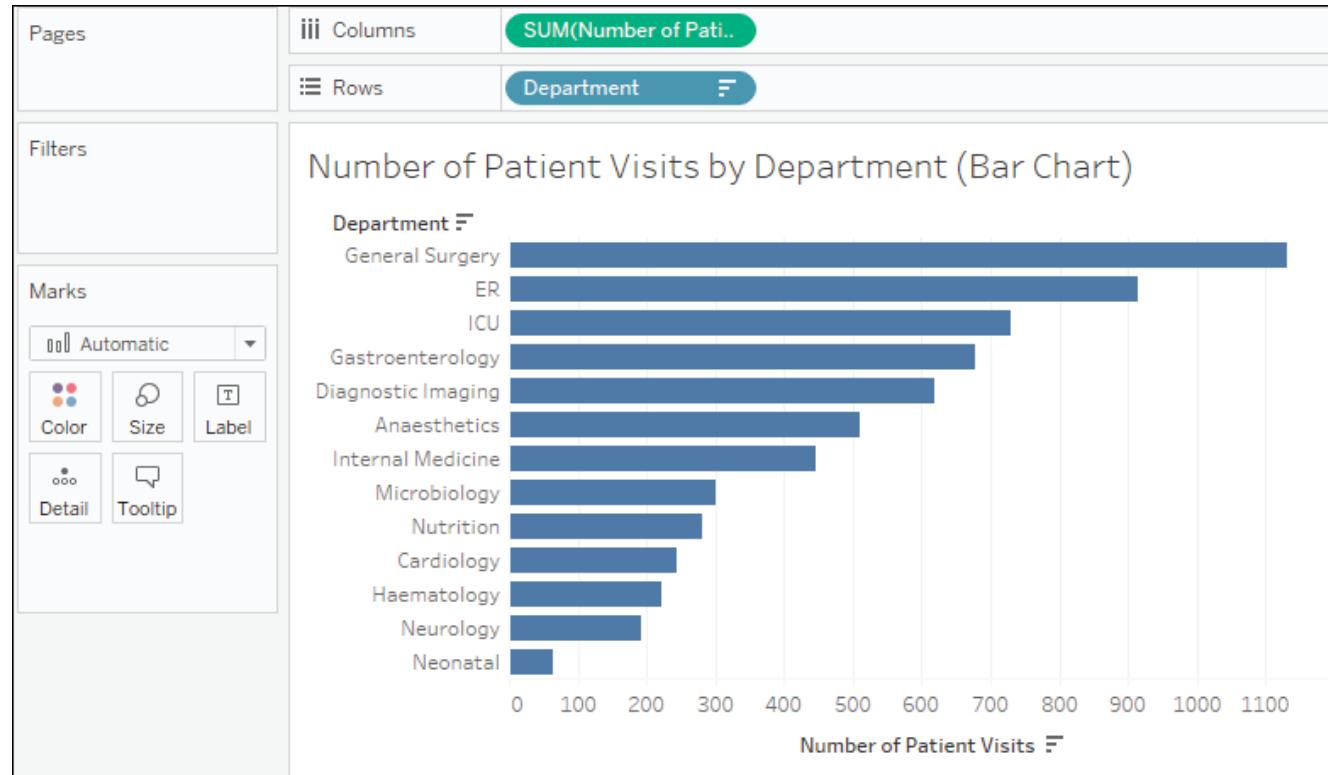
- Comparison
- Dates and times
- Relating parts of the data to the whole
- Distributions
- Multiple axes

Comparing values

Often, you will want to compare the differences between measured values across different categories. You might find yourself asking the following questions:

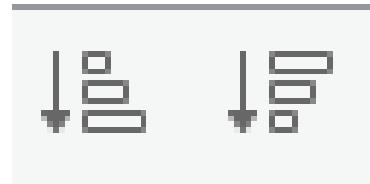
- How many customers did each store serve?
- How much energy did each wind farm produce?
- How many patients did each doctor see?

Bar charts



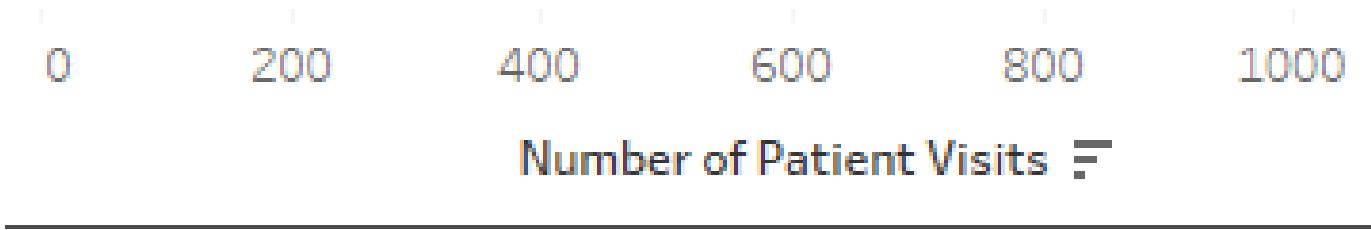
Bar charts

- Click one of the sort icons on the toolbar: This results in an automatic sort of the dimension based on the measure that defined the axis.
- Changes in data or filtering that result in a new order will be reflected in the view:



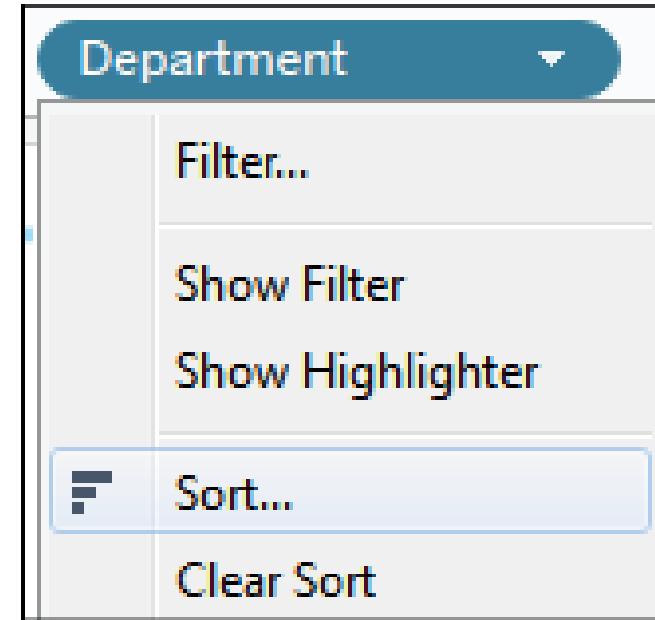
Bar charts

- Click the sort icon on the axis: The option icon will become visible when you hover over the axis and will then remain in place when you enable the sort.
- This will also result in automatic sorting:



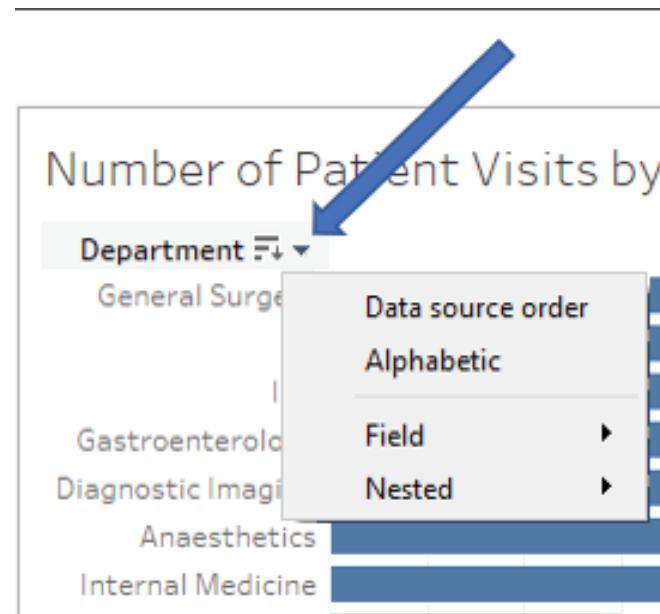
Bar charts

- Use the drop-down on the active dimension field and select Sort to view and edit the sorting options.
- You can also select Clear Sort to remove any sorting:



Bar charts

- Use the drop-down on the field label for rows and select the desired sorting option:

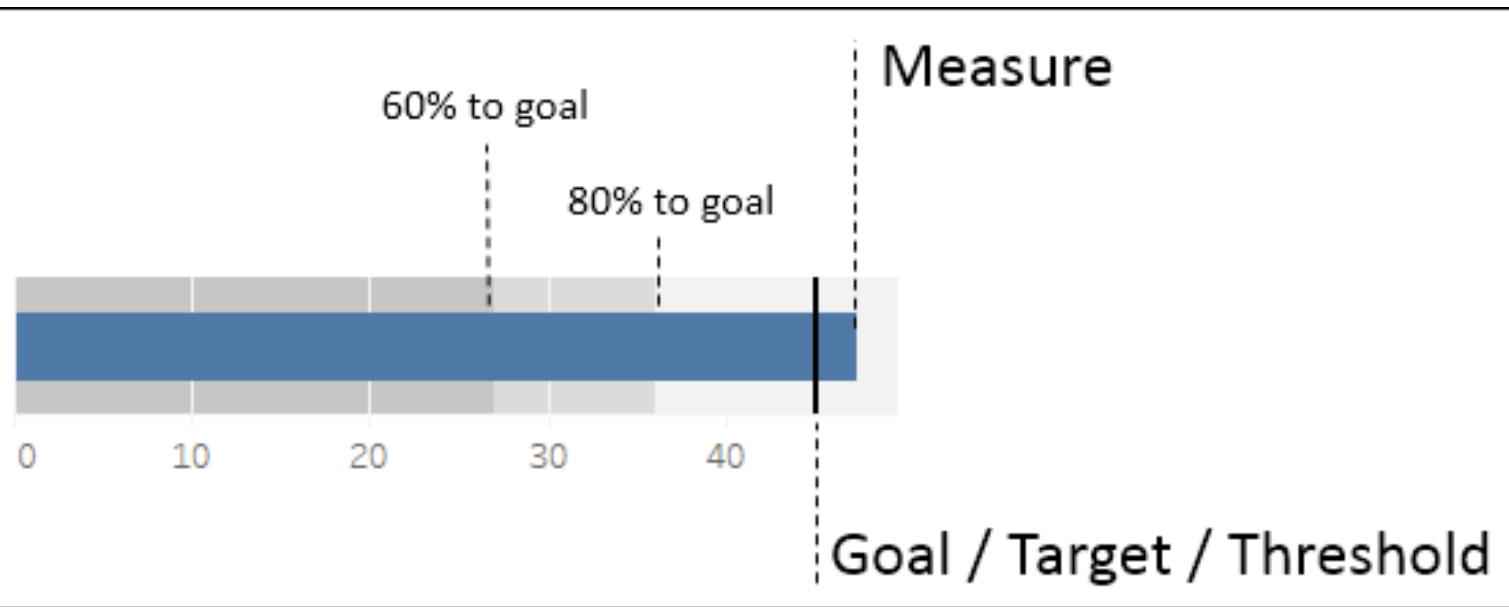


Bar chart variations

A basic bar chart can be extended in many ways to accomplish various objectives. Consider the following variations:

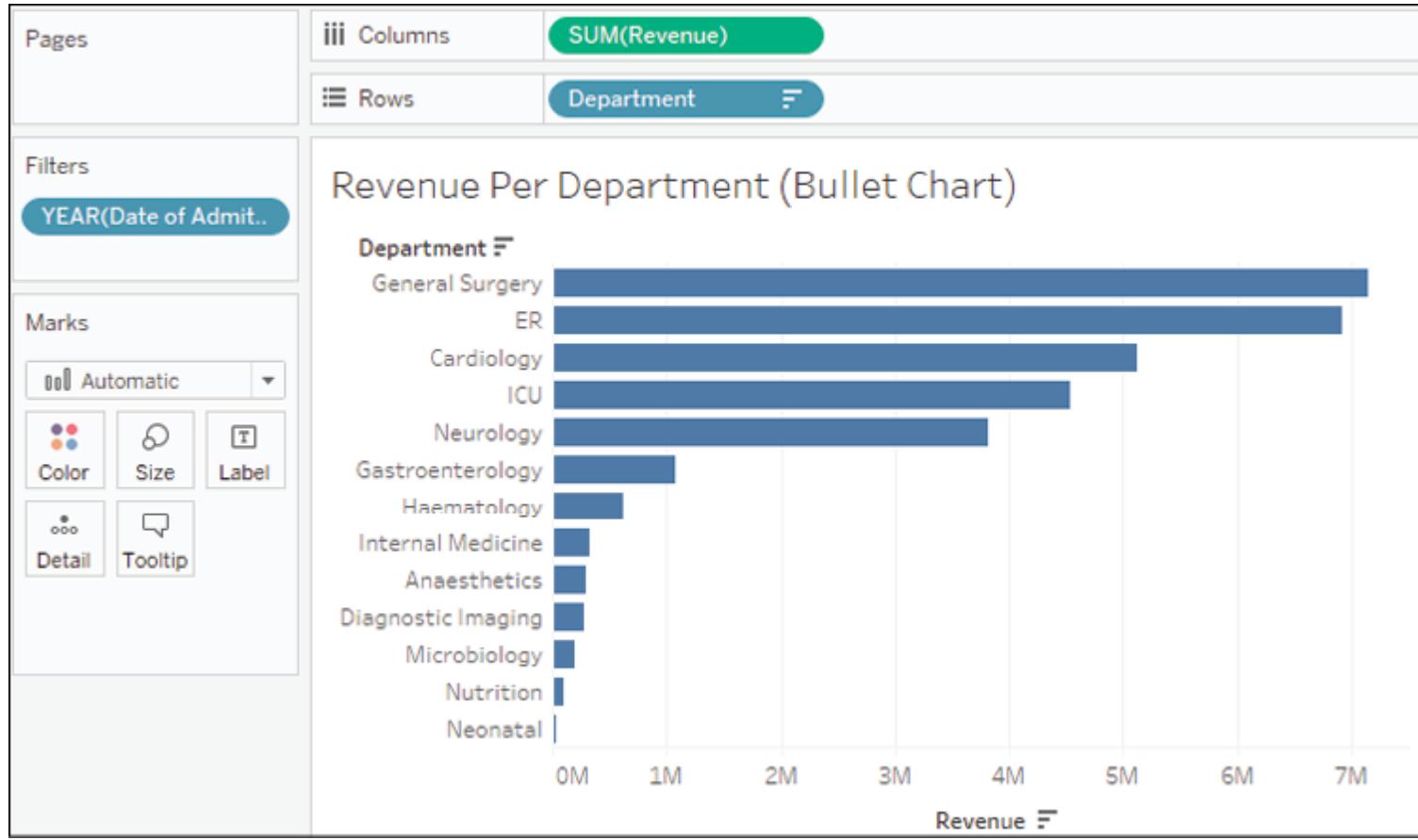
- Bullet chart to show progress towards a goal, target, or threshold
- Bar-in-bar chart to show progress toward a target or compare two specific values within a category
- Highlighting categories of interest

Bullet chart

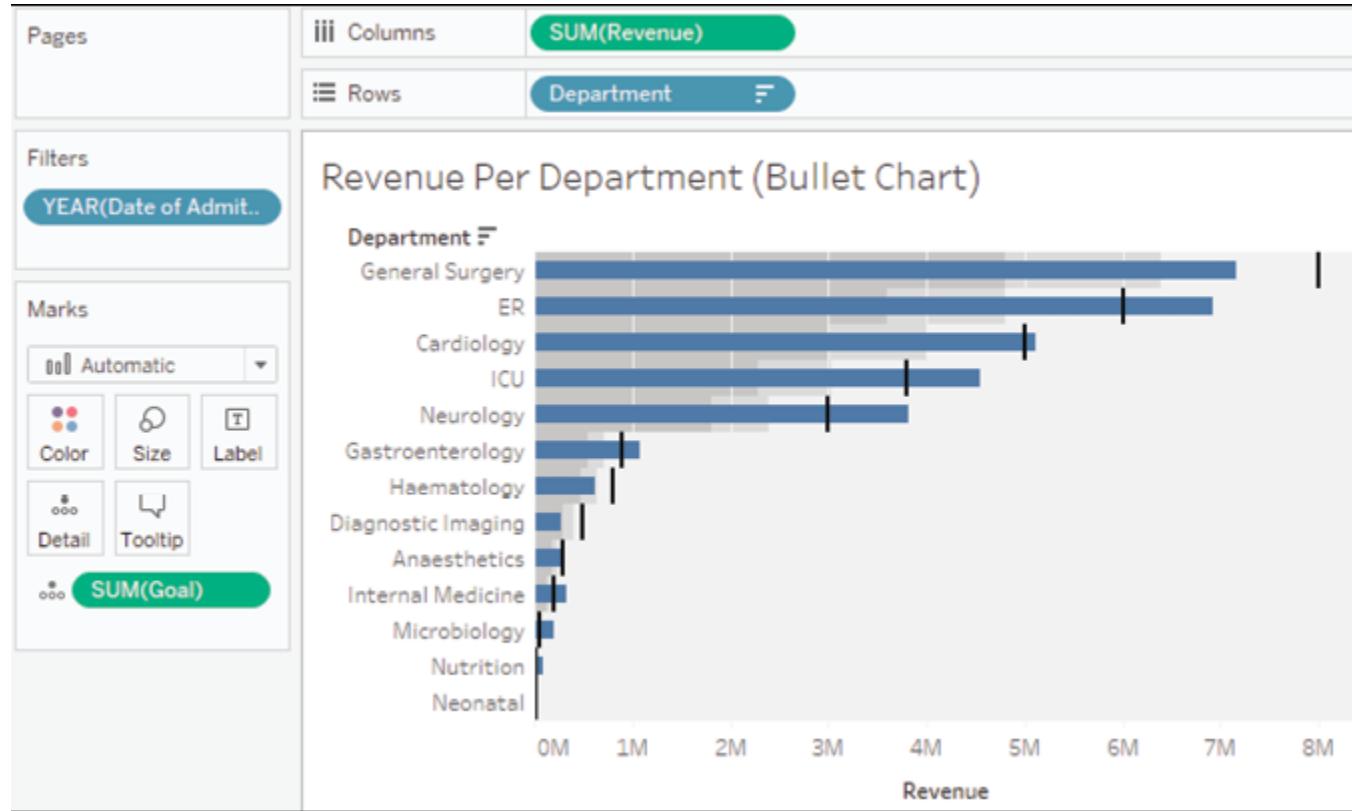


Bullet chart

Department	Goal
Anaesthetics	\$300,000
Cardiology	\$5,000,000
Diagnostic Imaging	\$500,000
ER	\$6,000,000
Gastroenterology	\$900,000
General Surgery	\$8,000,000
Haematology	\$800,000
ICU	\$3,800,000
Internal Medicine	\$200,000
Microbiology	\$50,000
Neonatal	\$10,000
Neurology	\$3,000,000
Nutrition	\$10,000

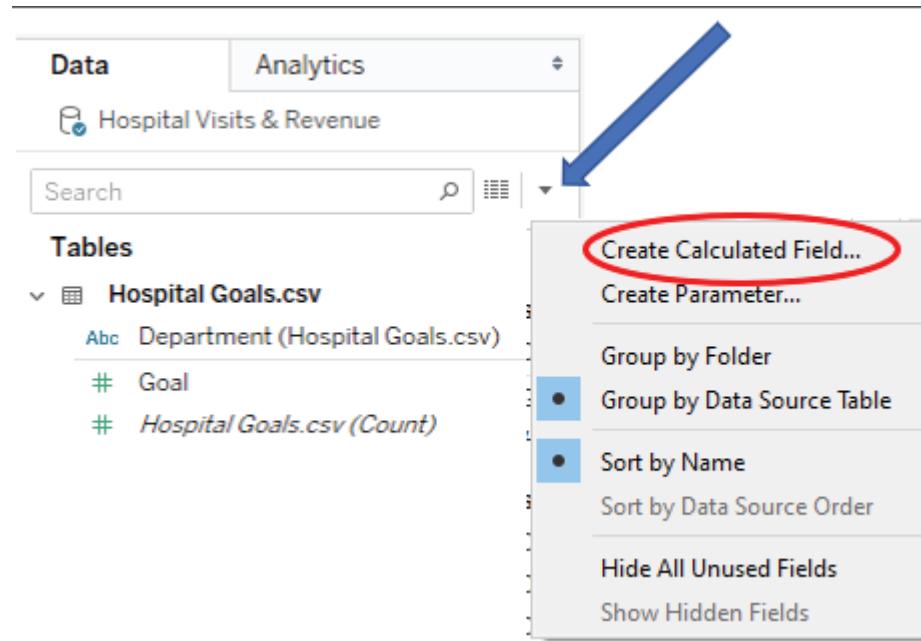


- The completed bullet chart should look like the following:



Calling out thresholds

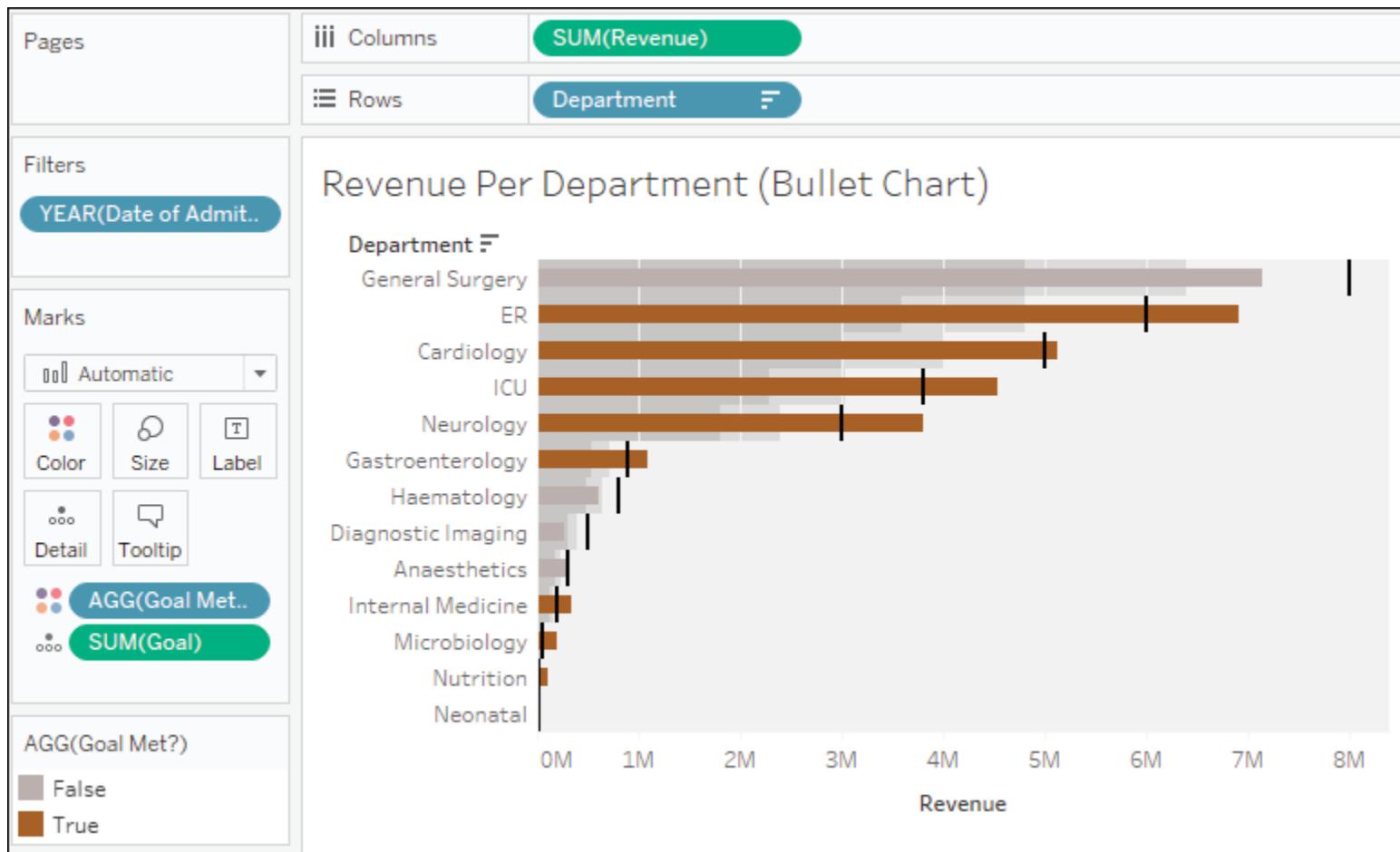
- Use the drop-down arrow in the Data pane and select Create Calculated Field...:

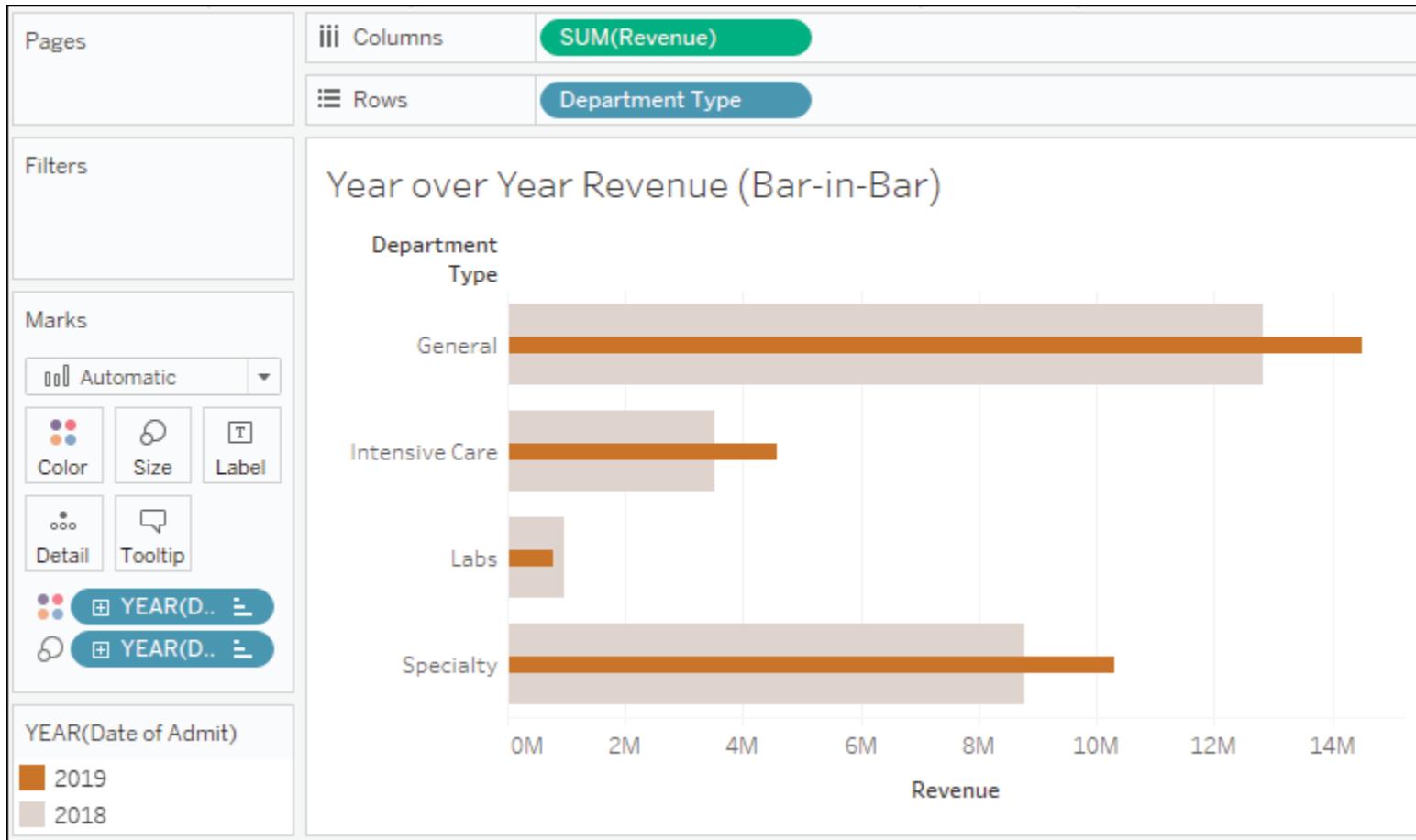


Calling out thresholds

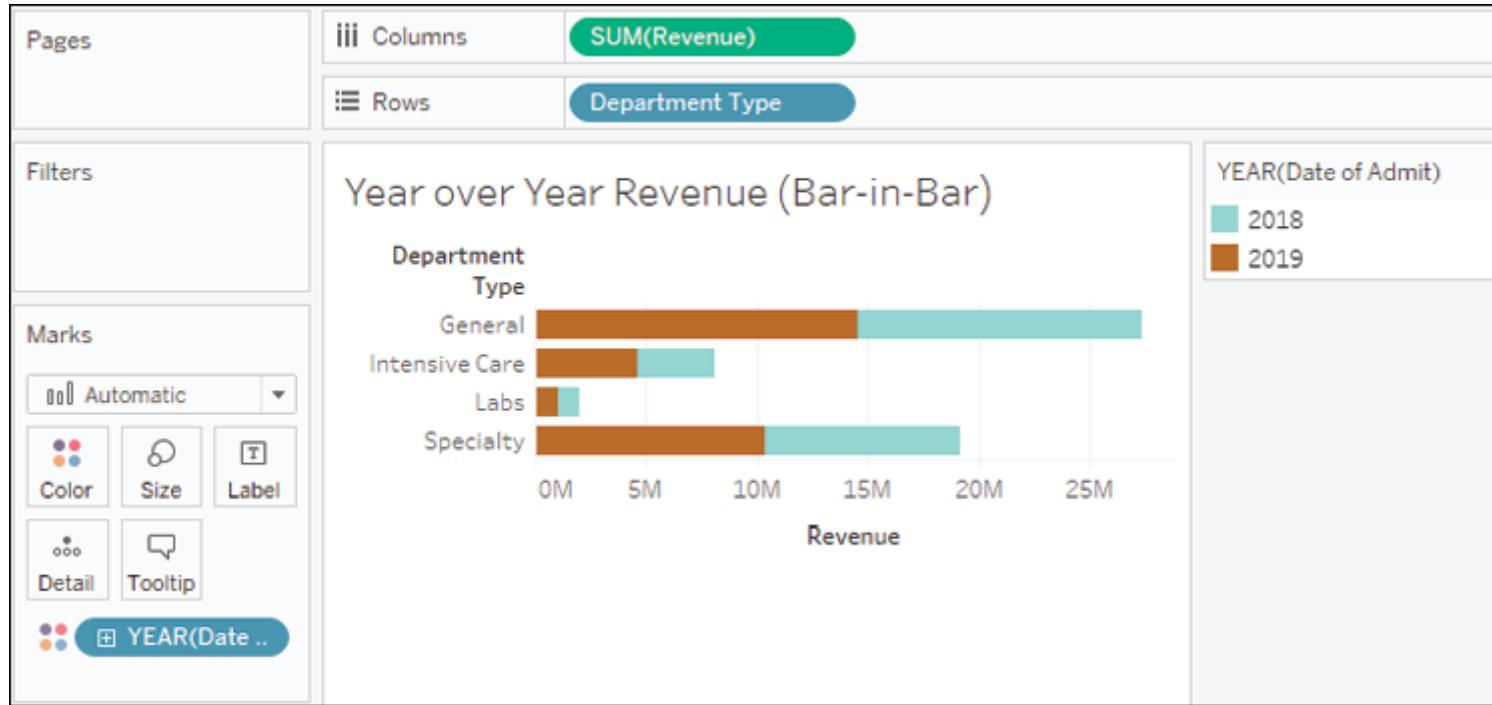
- Name the calculated field named Goal Met? with the following code:

$\text{SUM}([\text{Revenue}]) \geq \text{SUM}([\text{Goal}])$



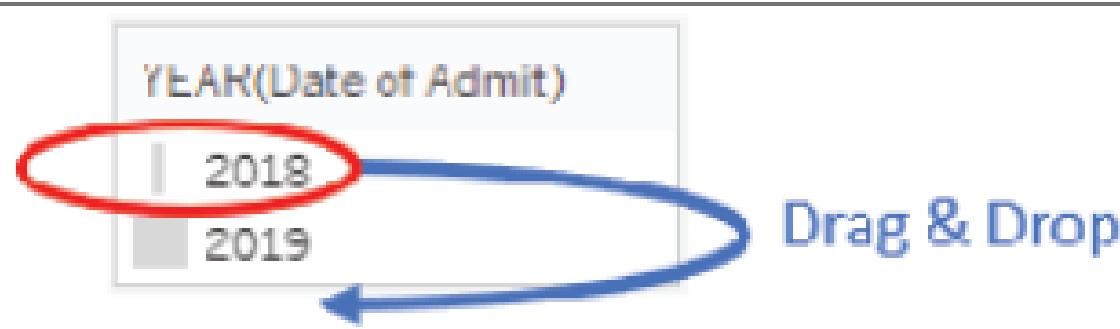


Bar-in-bar chart



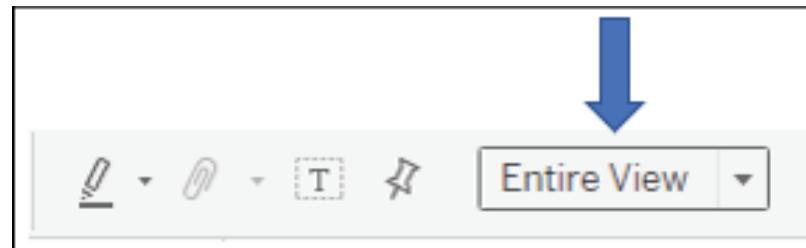
Bar-in-bar chart

- We want 2019 to be in front and 2018 to be in the background, so drag and drop 2019 within the Size legend to reorder the values so that 2018 comes after 2019:

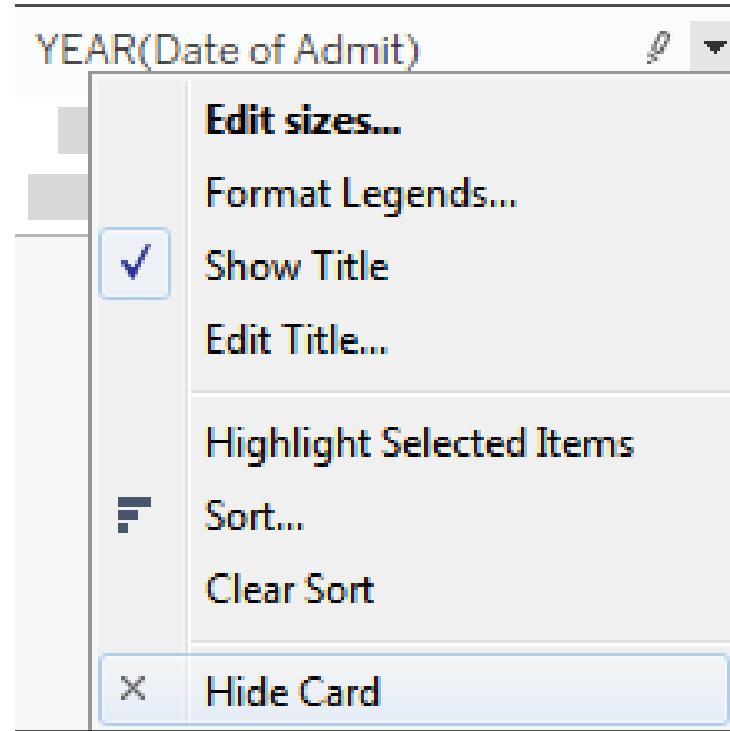


Bar-in-bar chart

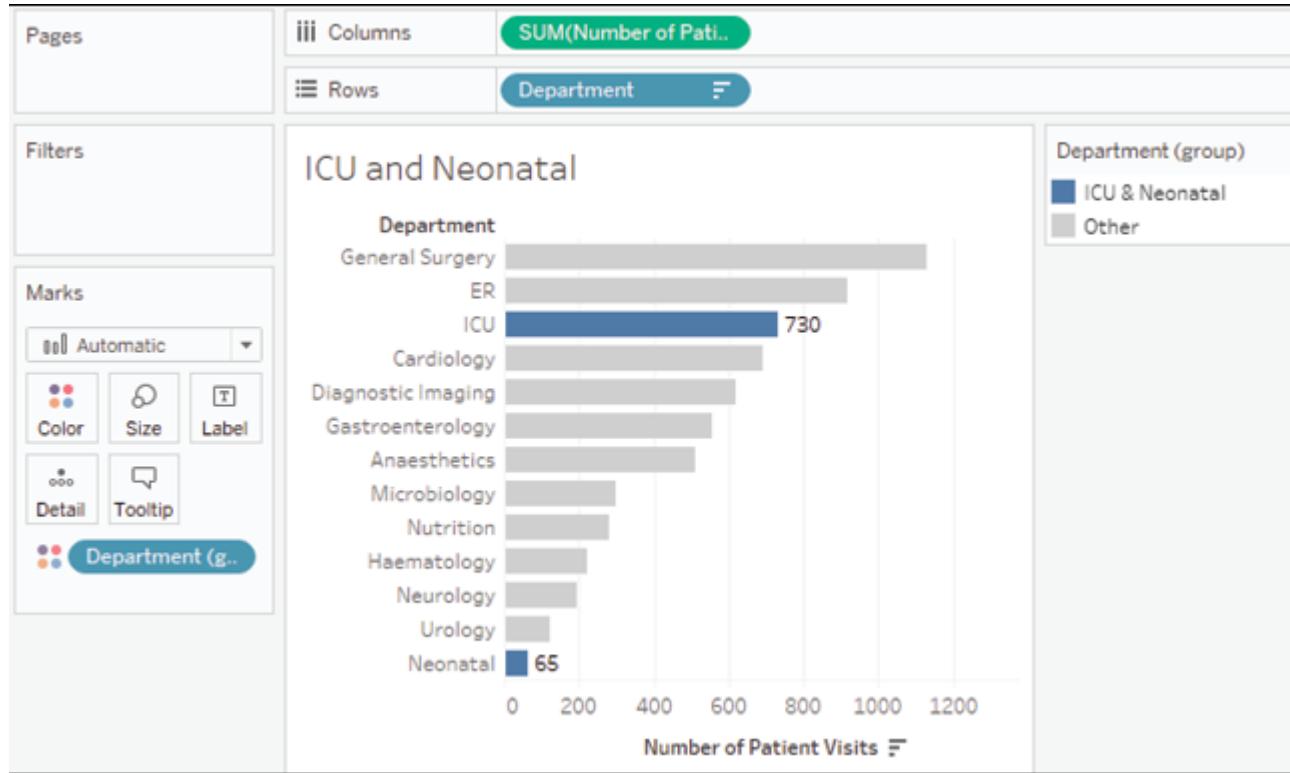
- Adjusting the sizing of the view. Accomplish this by hovering over the canvas, just over the bottom border, until the mouse cursor changes to a sizing cursor, and then click and drag to resize the view.
- You may also want to adjust how the view fills the space.
- Use the drop-down on the toolbar and experiment with the options:



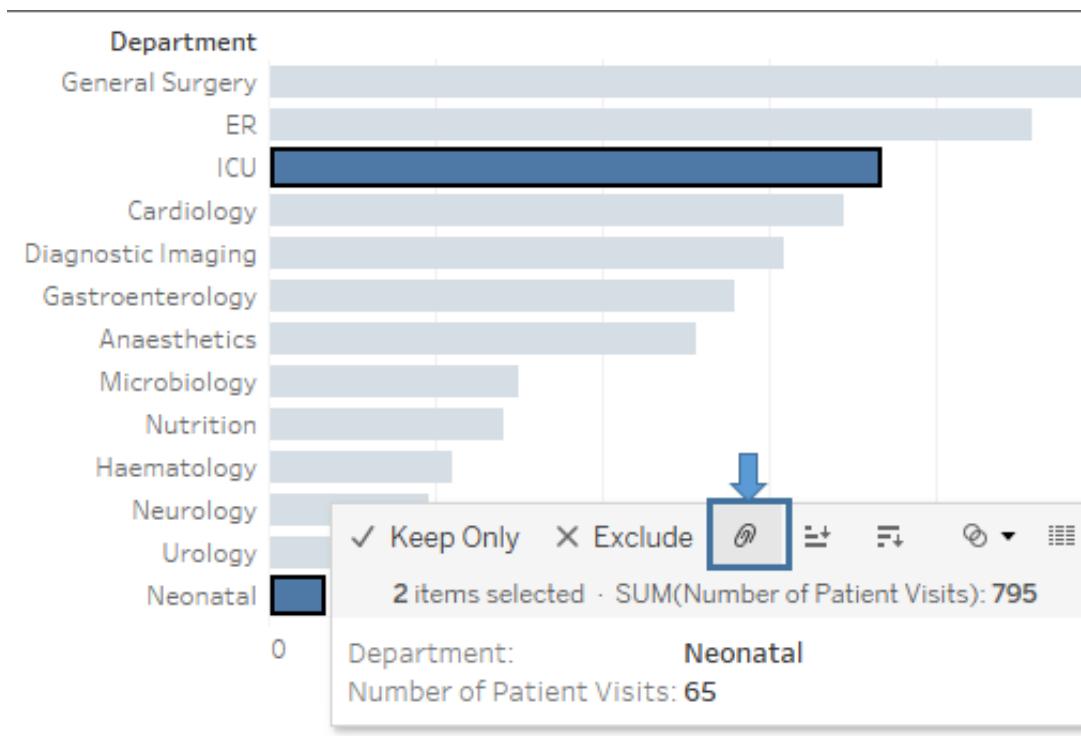
Bar-in-bar chart



Highlighting categories of interest



Highlighting categories of interest



Visualizing dates and times

In your analysis, you will often want to understand when something happened. You'll ask questions like the following:

- When did we gain the most new customers?
- Is profit trending up or down?
- What times of day have the highest call volume?
- What kinds of seasonal trends do we see in sales?

Date parts, date values, and exact dates

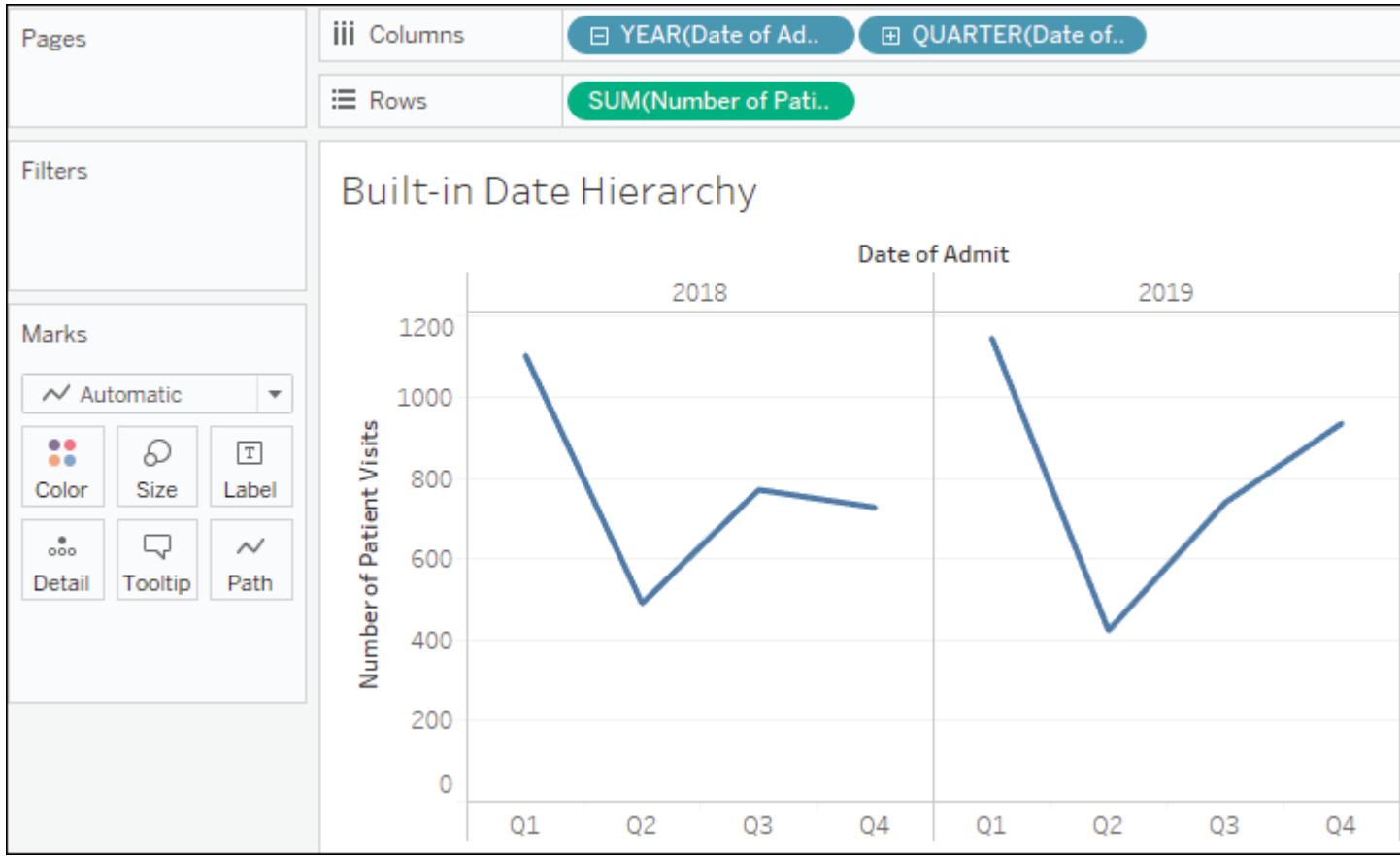
- To see this in action, continue with the lesson 3 workbook, navigate to the Built-in Date Hierarchy sheet, and create a view similar to the one that was shown by dragging and dropping Number of Patient Visits to Rows and Date of Admit to Columns.
- The YEAR(Date of Admit) field on Columns will have a plus sign indicator, like this:

[+] YEAR(Date of Ad..)

Date parts, date values, and exact dates

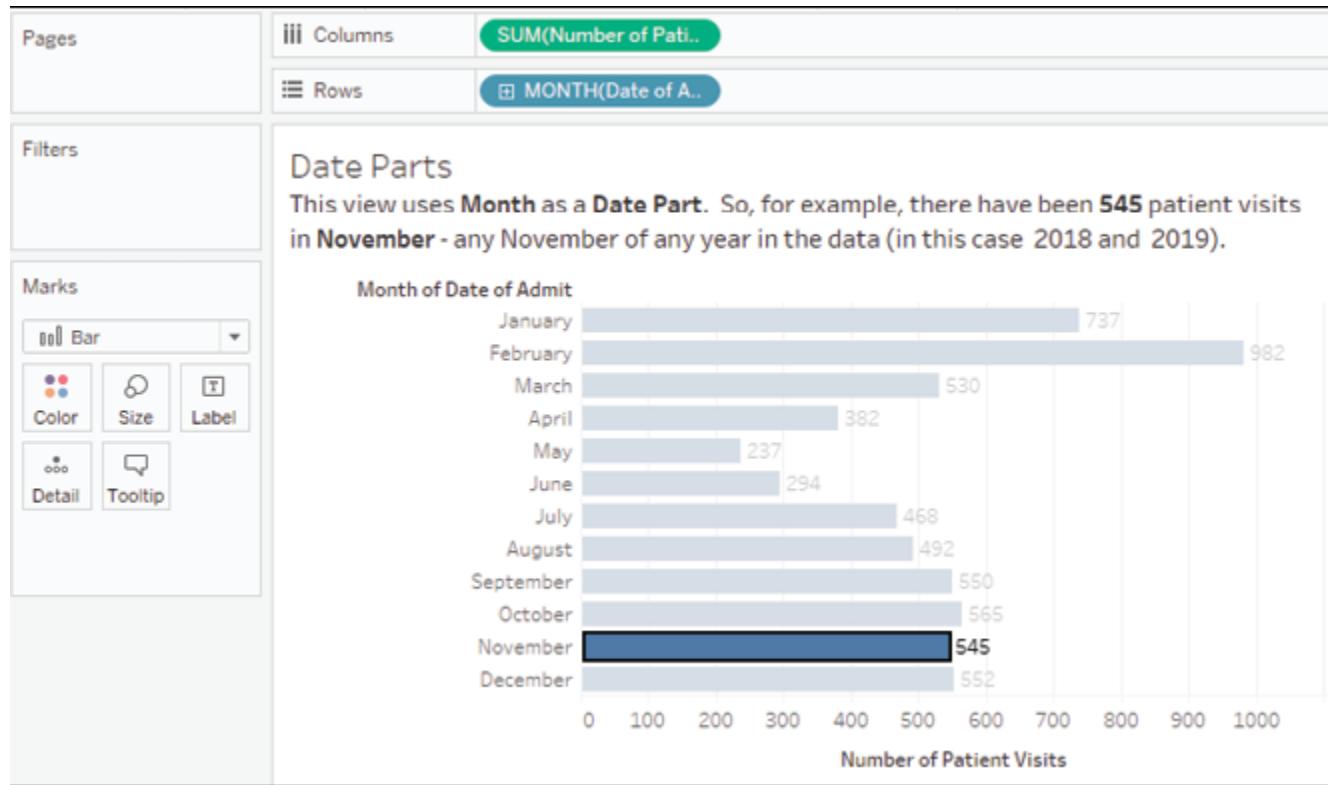
- You'll also find a plus or minus indicator as you hover over headers, like this:

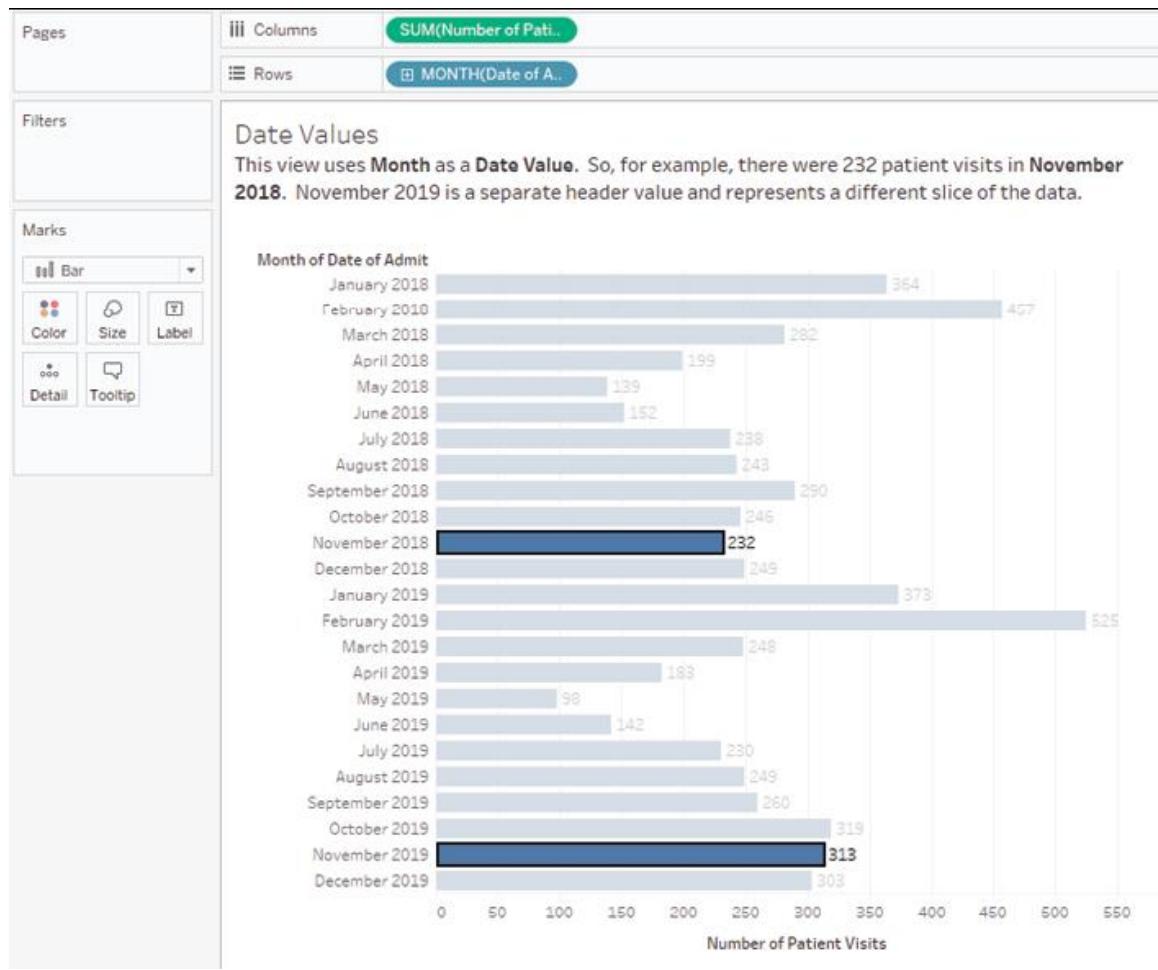




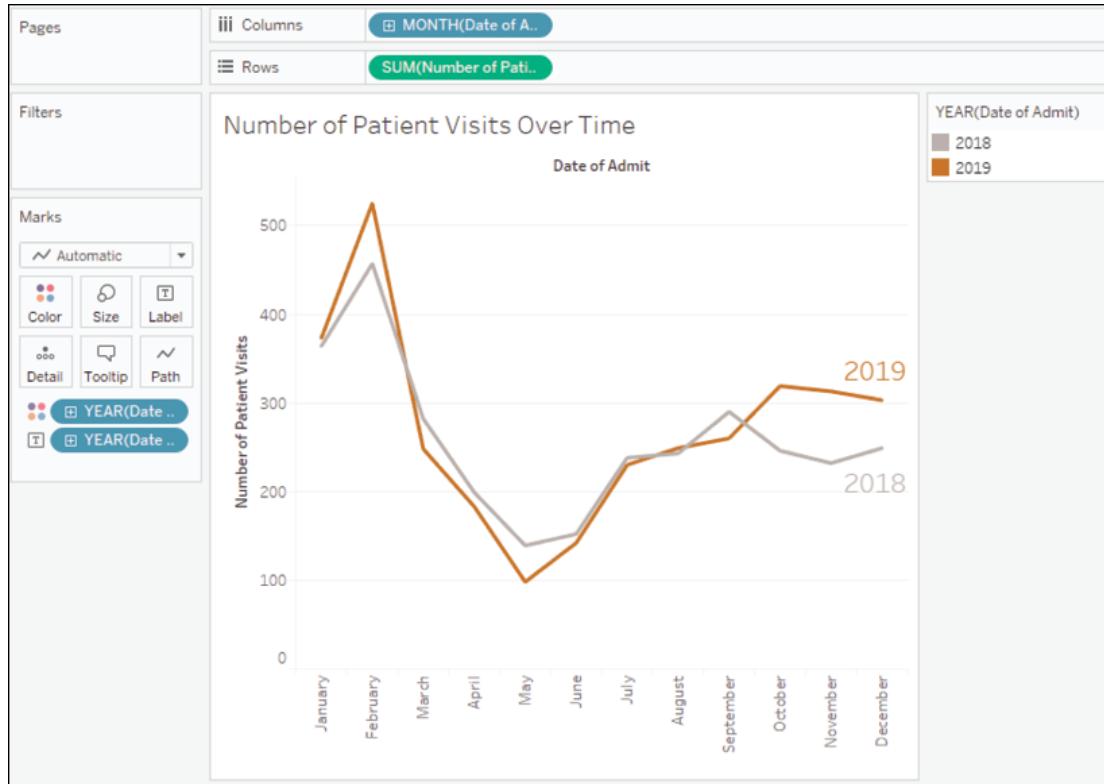
Show Missing Values	Show dates that are "skipped" in the data
<input checked="" type="radio"/> Year 2011	
Quarter Q2	
Month May	
Day 8	
More ▾	
Year 2011	
Quarter Q2 2011	
Month May 2011	
Week Number Week 5, 2011	
Day May 8, 2011	
More ▾	
Exact Date	Exact Value of Field
Attribute	
Measure ▾	
<input checked="" type="radio"/> Discrete	Switch between Discrete and Continuous
Continuous	

Date parts, date values, and exact dates

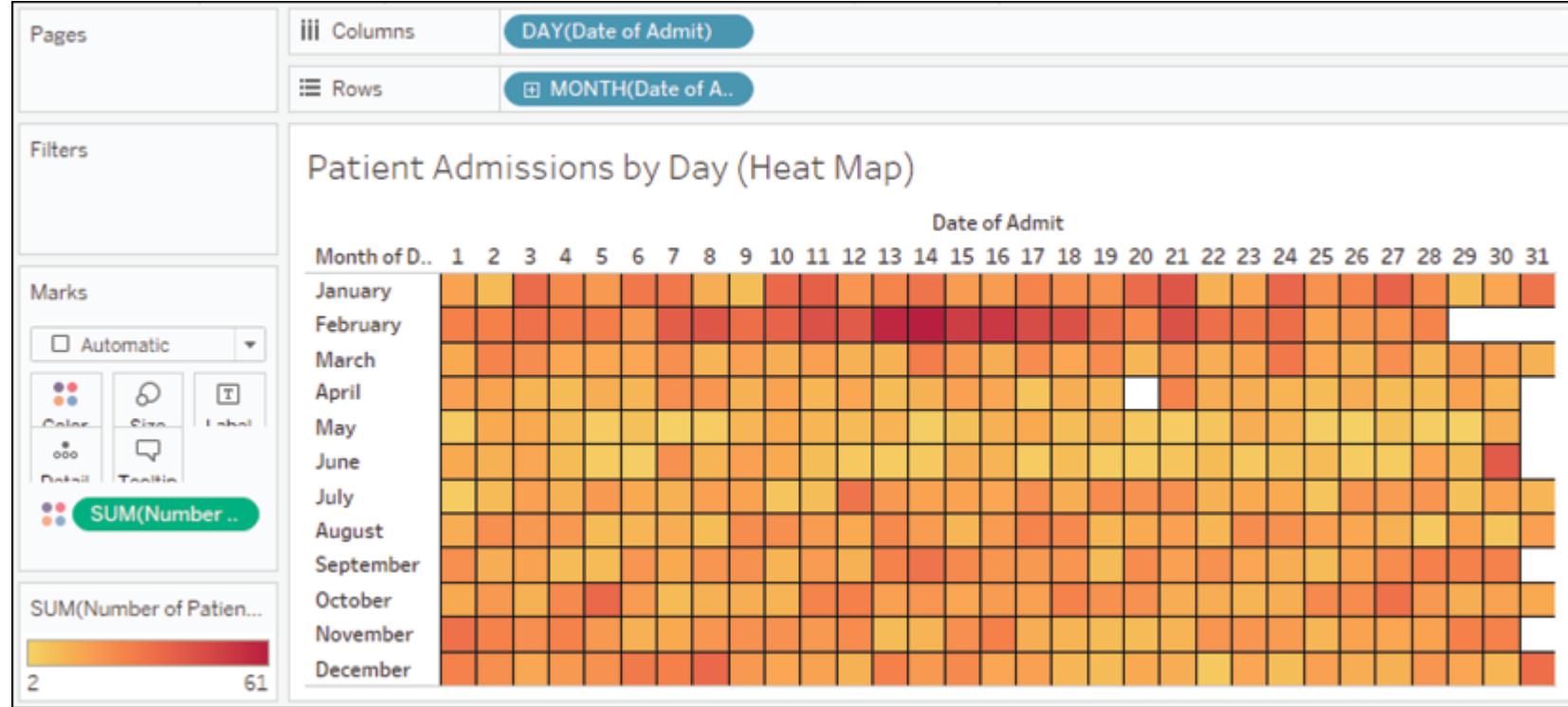




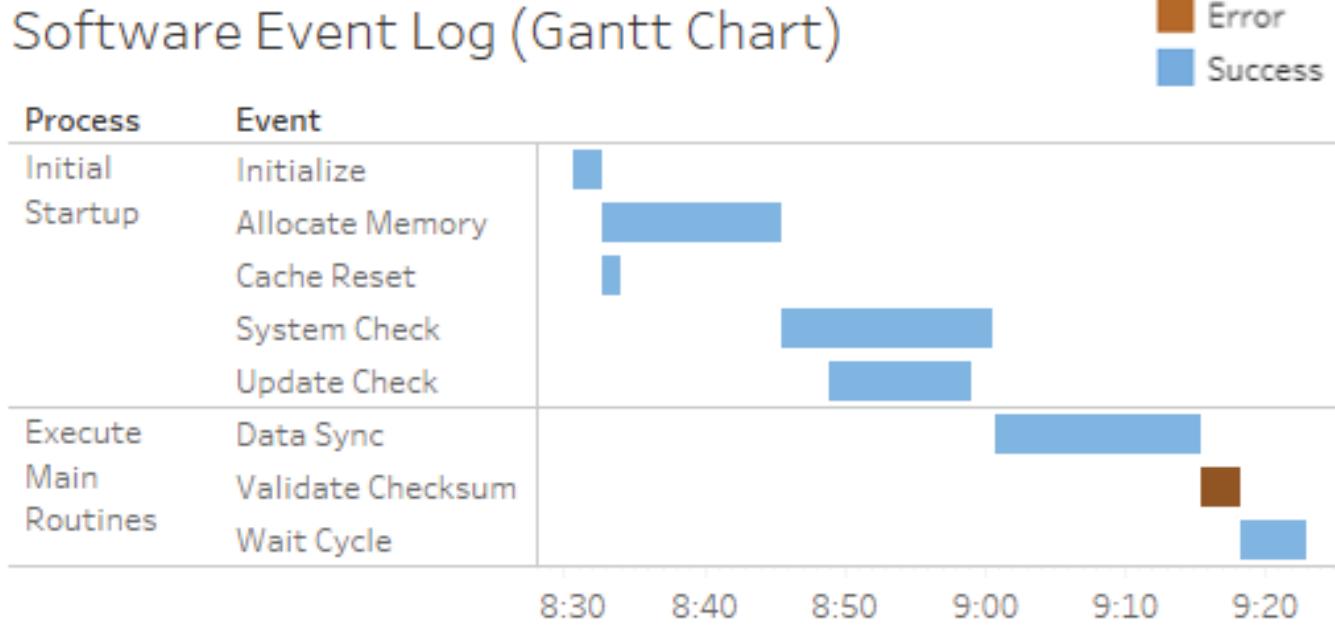
Variations of date and time visualizations



Date parts, date values, and exact dates

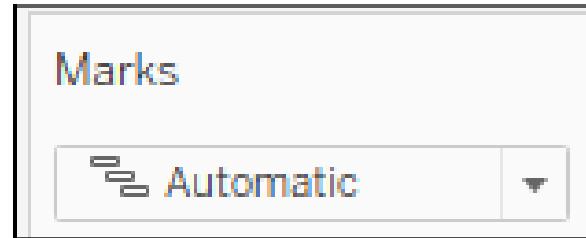


Gantt charts



Gantt charts

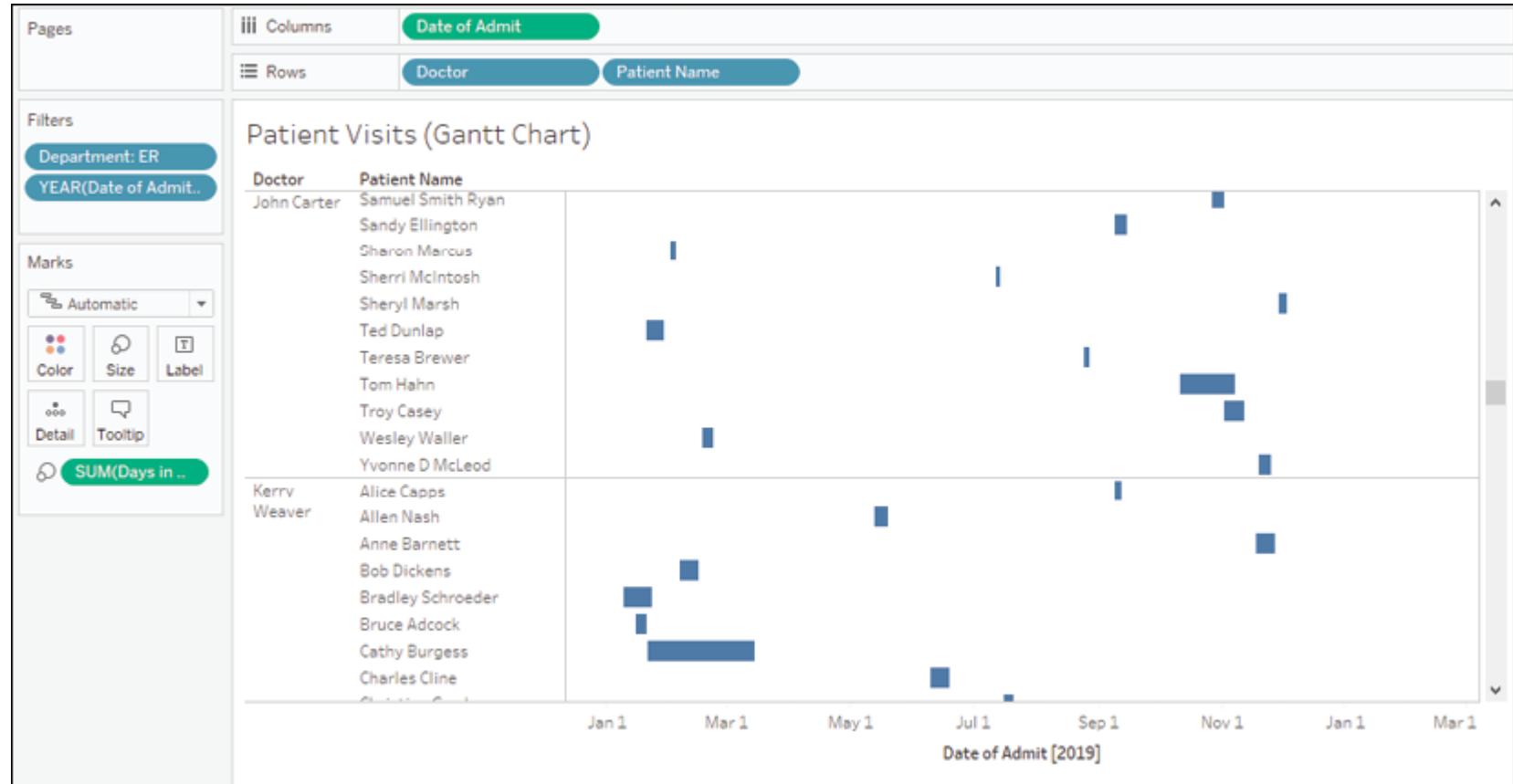
- Place Department on Filters and keep only ER.
- Place Date of Admit on Filters, select Years as the option for filtering, and keep only 2019.
- Place Date of Admit on Columns as a continuous Exact Date or as a Day value (not Day part) & Notice that Tableau's automatic default for the mark type is Gantt bars:



Gantt charts

- The length of the Gantt bar is set by placing a field with a value of duration on the Size shelf. There is no such field in this dataset.
- However, we have the Date of Discharge, and we can create a calculated field for the duration. We'll cover calculations in more detail in the next lesson.
- For now, select Analysis from the menu and click Create Calculated Field.... Name the field Days in the Hospital and enter the following code:

DATEDIFF('day', [Date of Admit], [Date of Discharge])

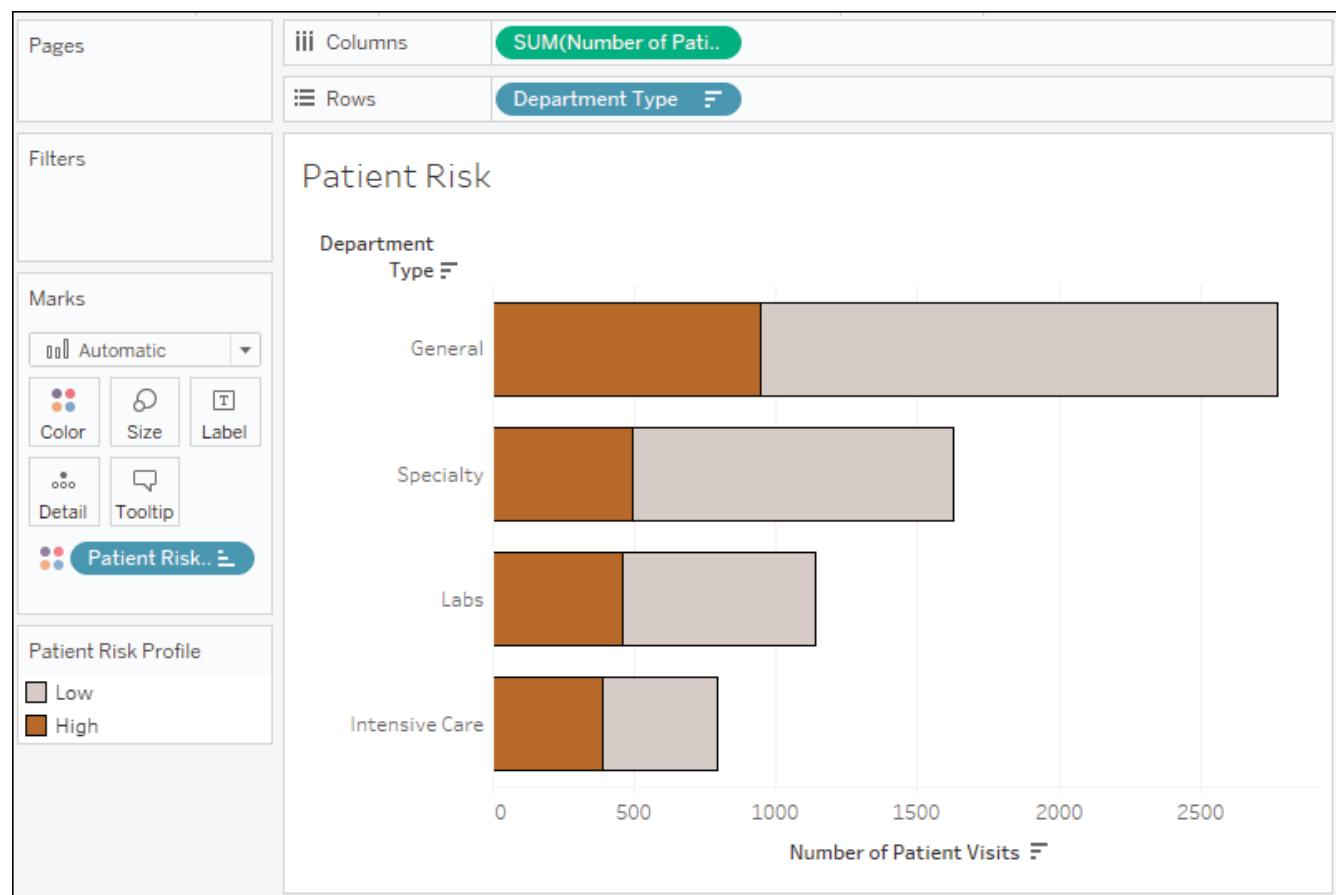


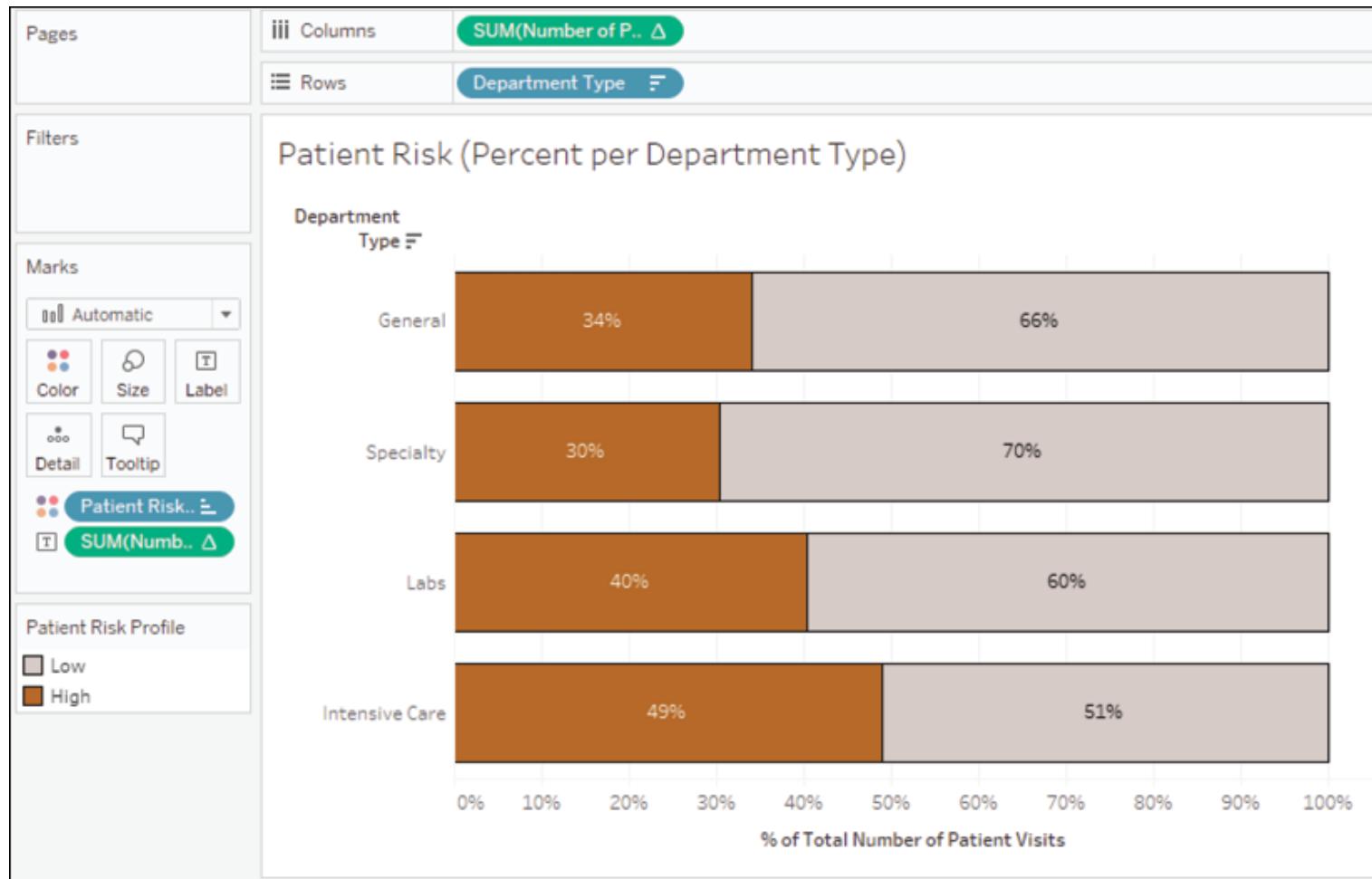
Relating parts of the data to the whole

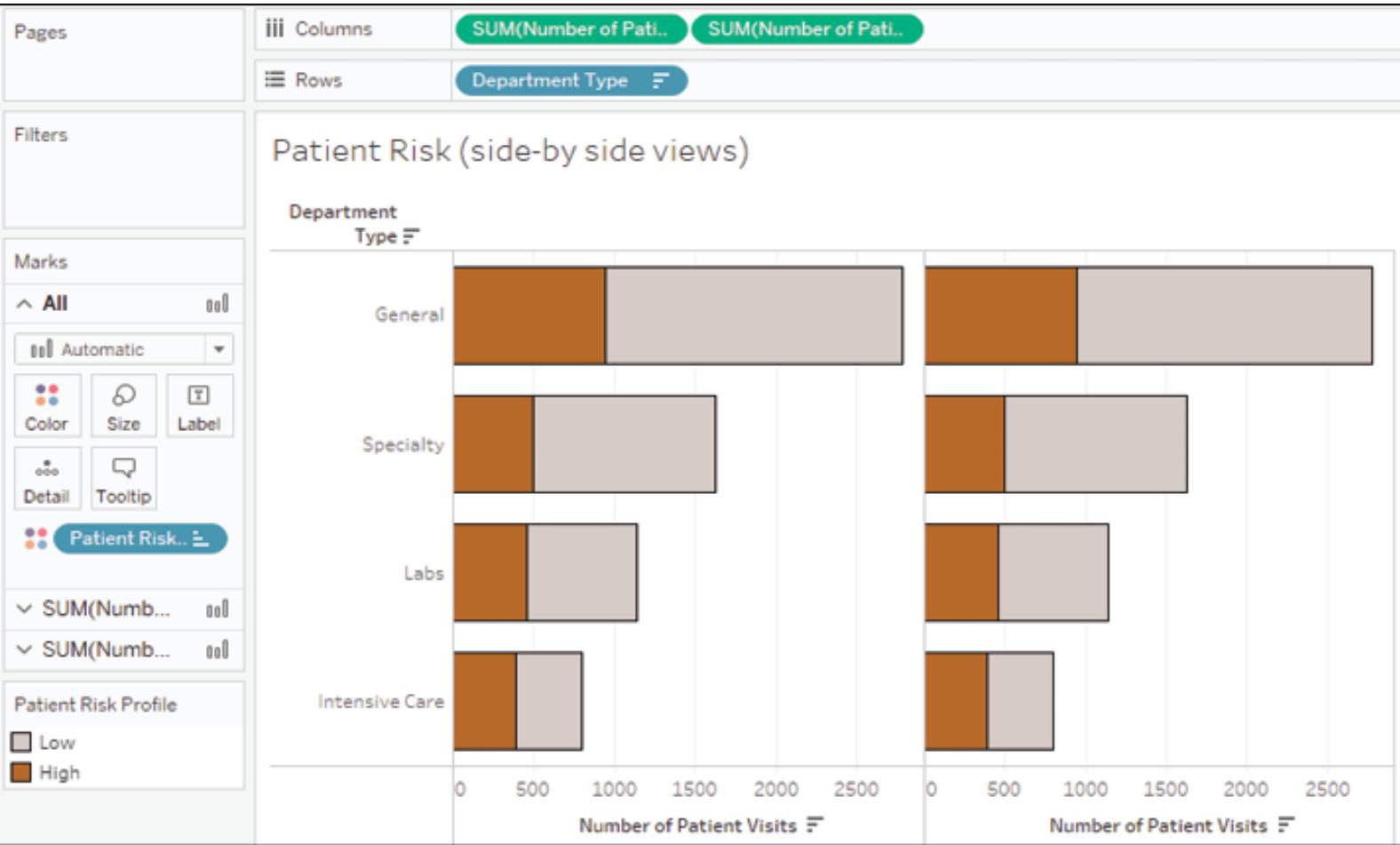
As you explore and analyze data, you'll often want to understand how various parts add up to a whole. For example, you'll ask questions such as the following:

- How much does each electric generation method (wind, solar, coal, and nuclear) contribute to the total amount of energy produced?
- What percentage of total profit is made in each state?
- How much space does each file, subdirectory, and directory occupy on my hard disk?

Stacked bars

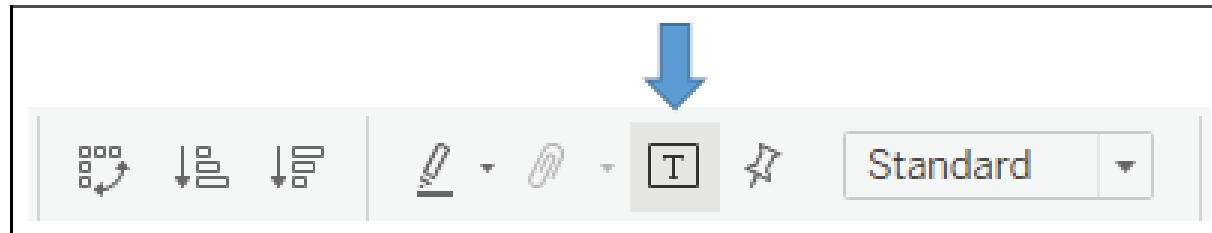






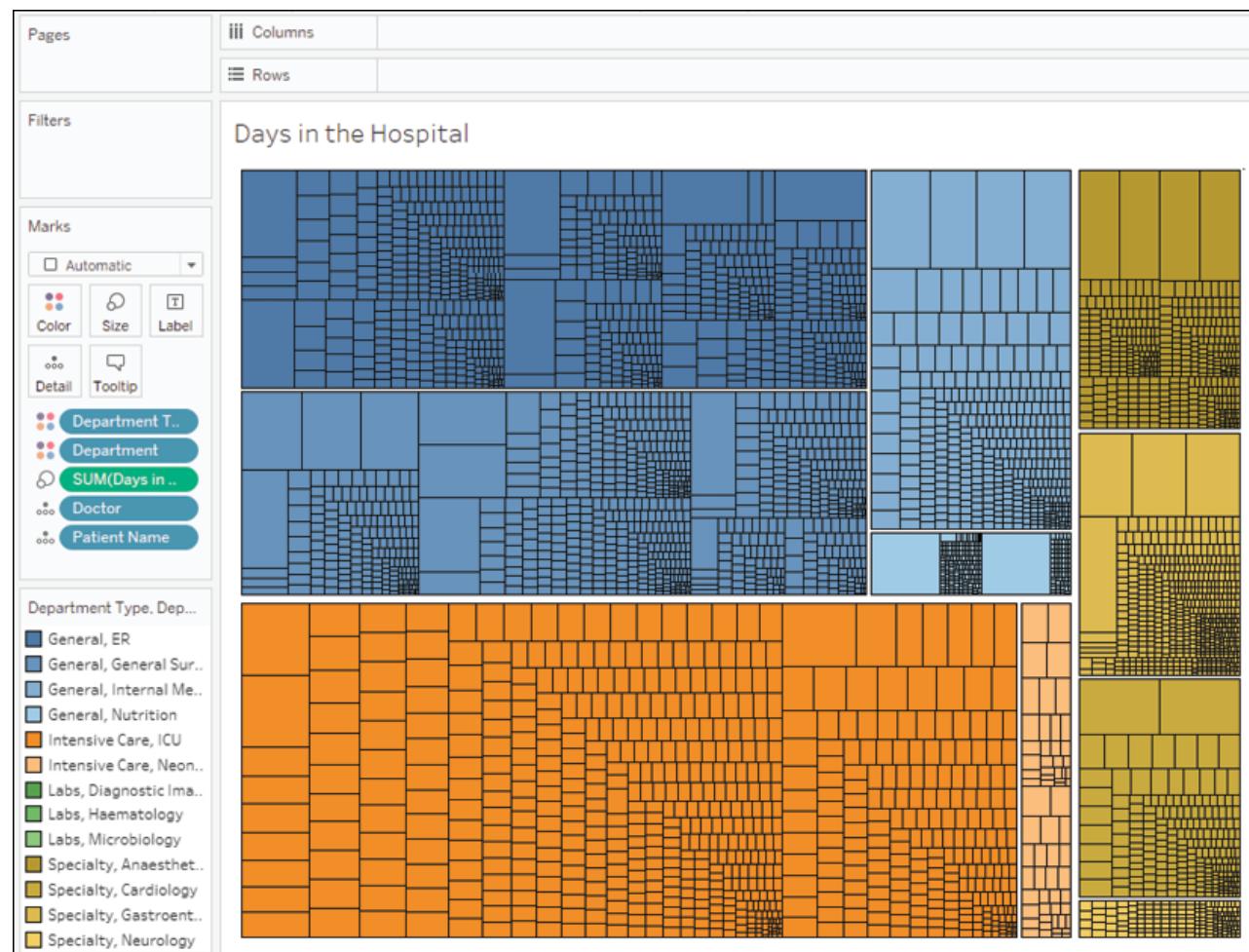
Stacked bars

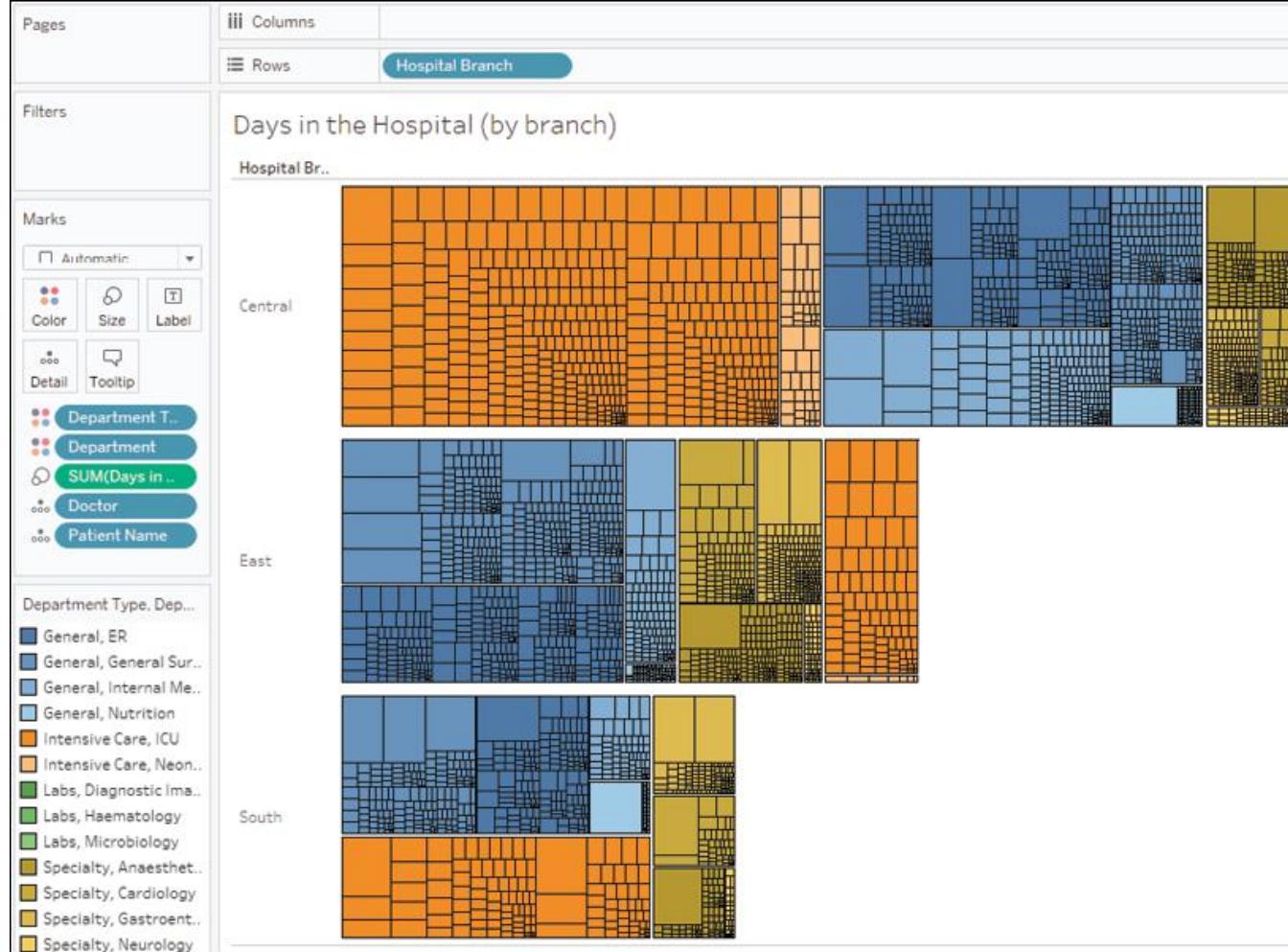
- Turn on labels by clicking the T button on the top toolbar.
- This turns on default labels for each mark:





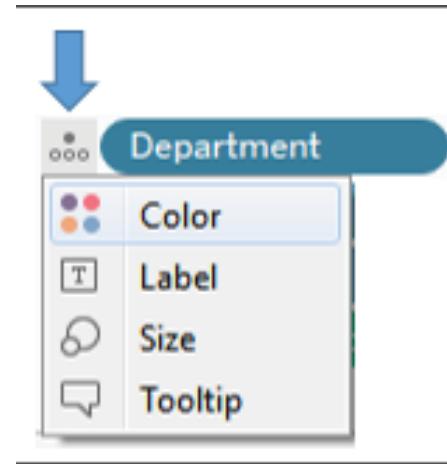
Treemaps



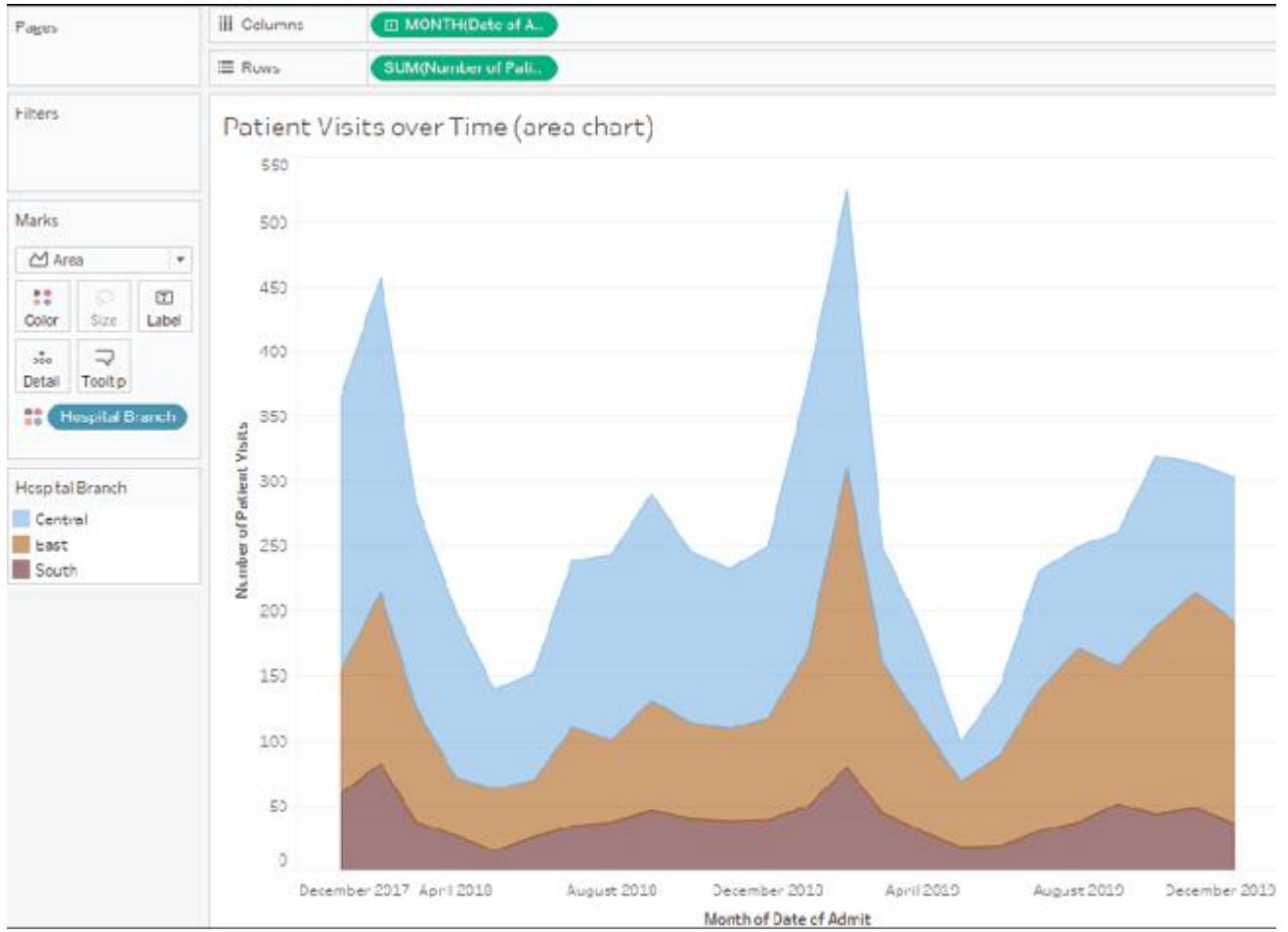


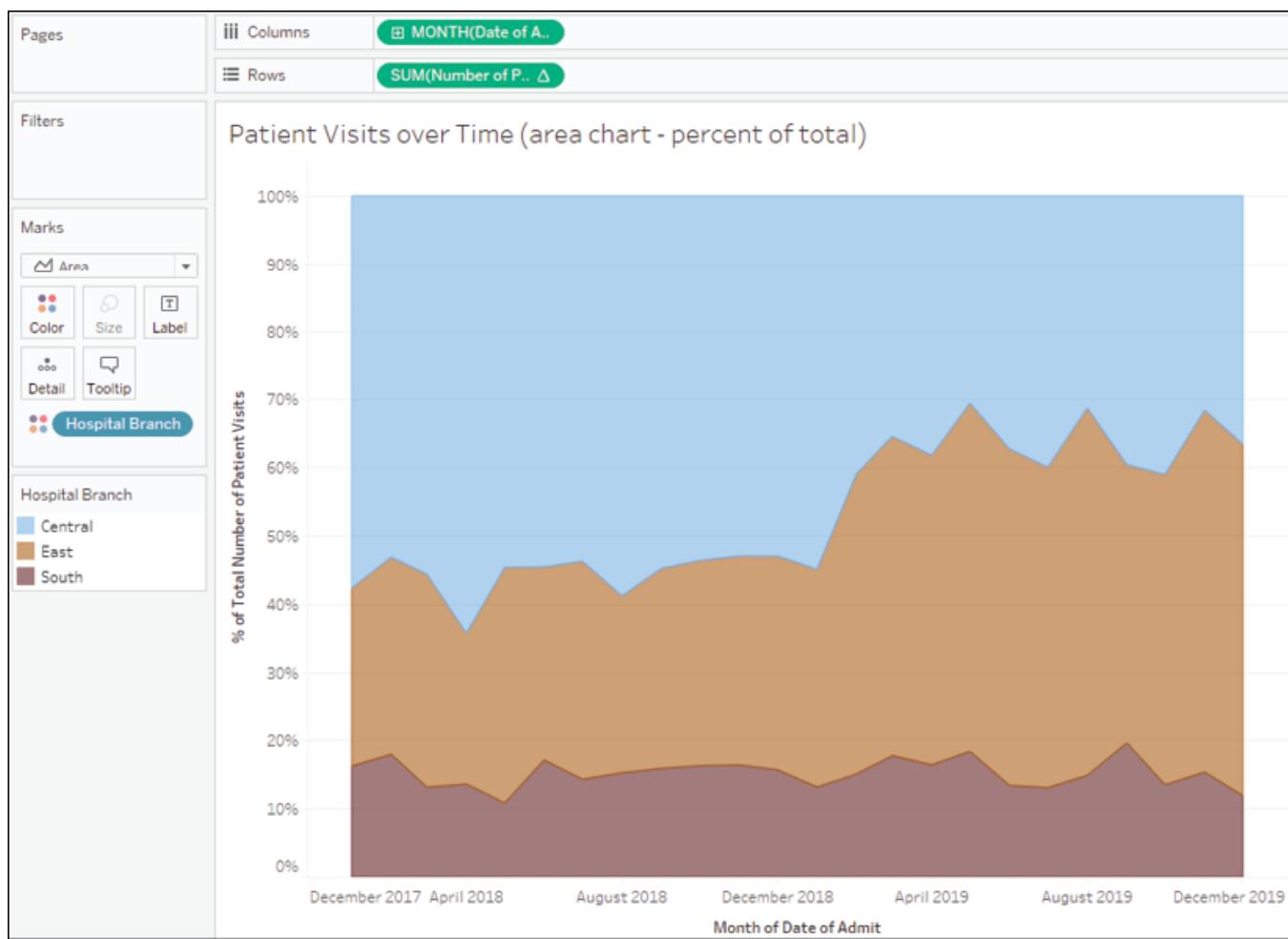
Treemaps

- Alternatively, the icon or space to the left of each field on the Marks card can be clicked to change which shelf is used for the field:

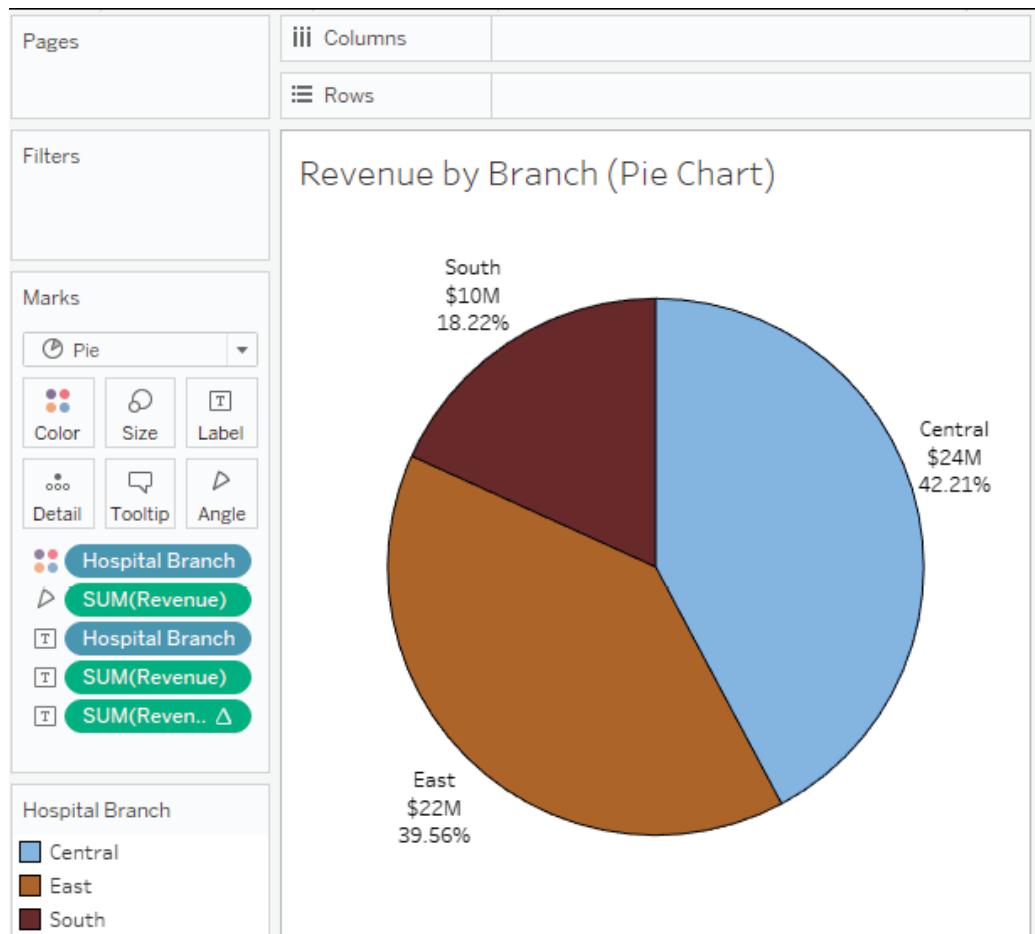


Area charts





Pie charts



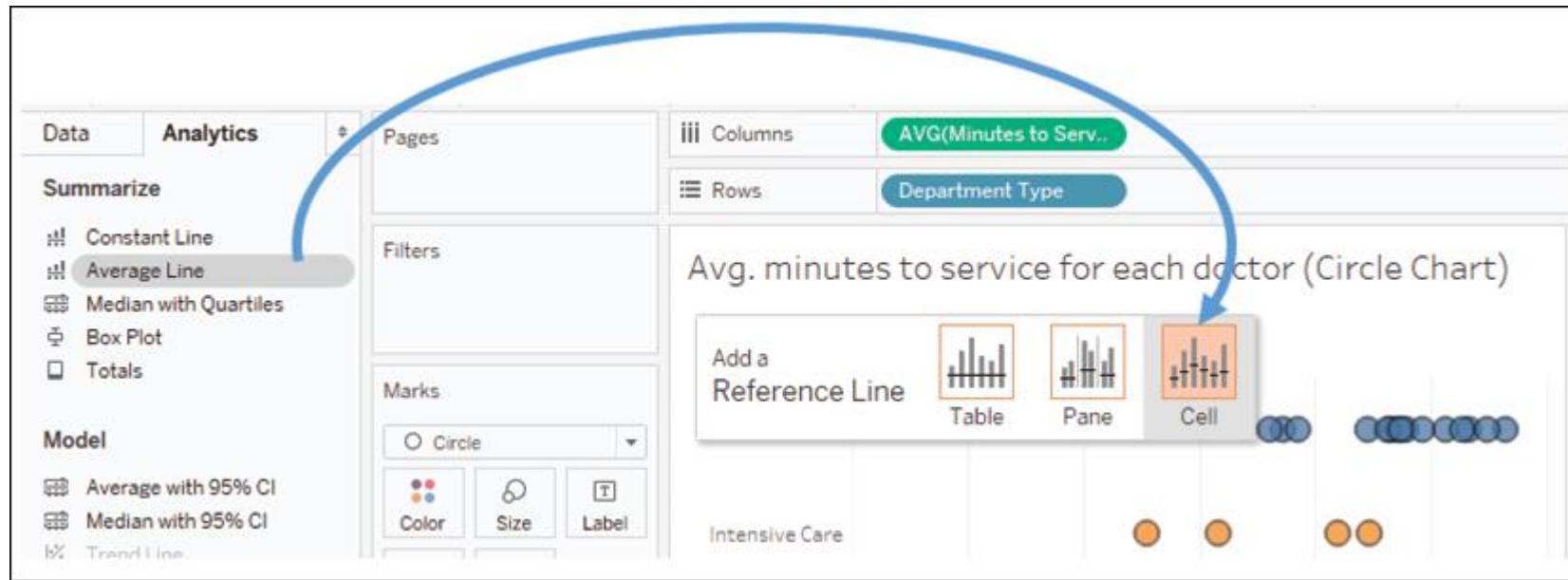
Visualizing distributions

- Often, simply understanding totals, sums, and even the breakdown of part-to-whole only gives a piece of the overall picture.
- Most of the time, you'll want to understand where individual items fall within a distribution of all similar items.

Circle charts



Circle charts



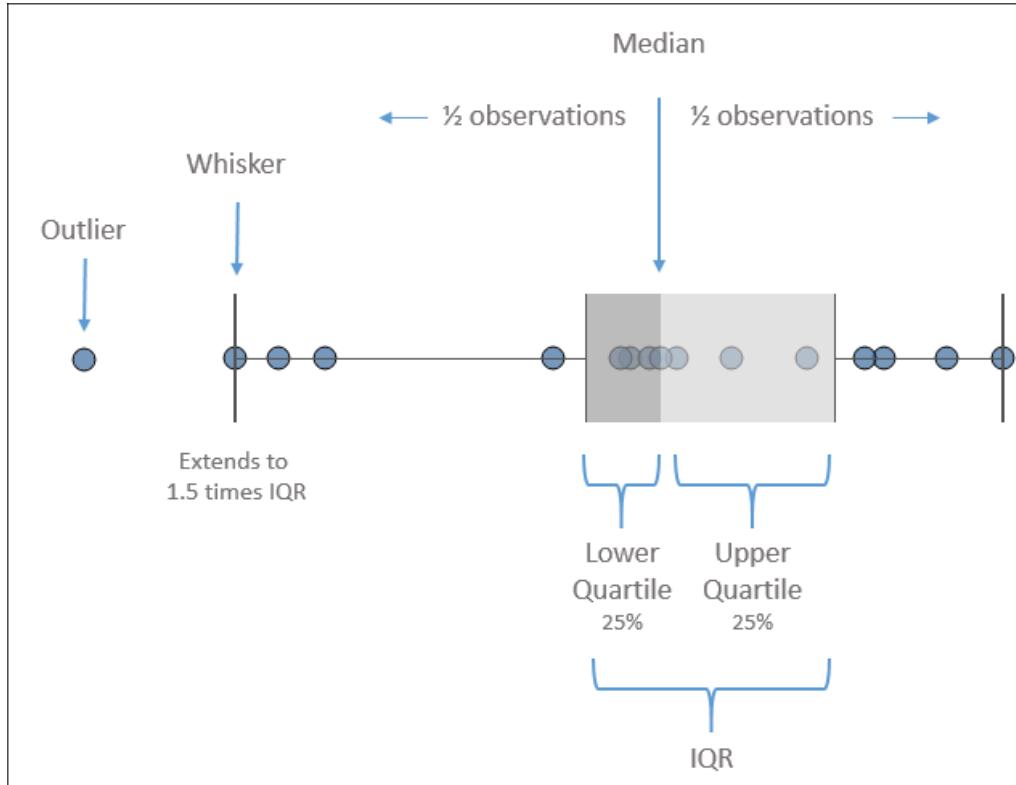
You may also click one of the resulting average lines and select Edit to find fine-tuning options, such as labeling.

Jittering

- When using views like circle plots or other similar visualization types, you'll often see that marks overlap, which can lead to obscuring part of the true story.
- Do you know for certain, just by looking, how many doctors there are in Intensive Care who are above average? How many are below?

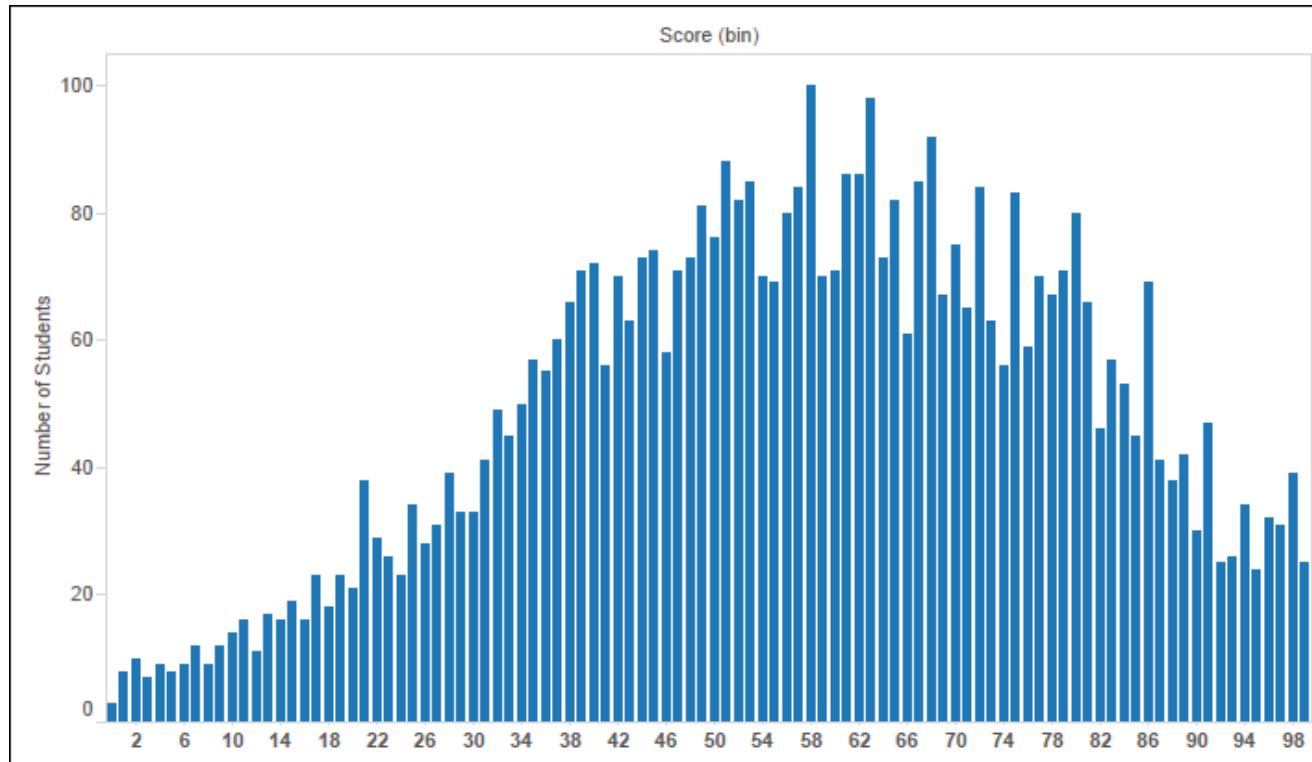


Box and whisker plots



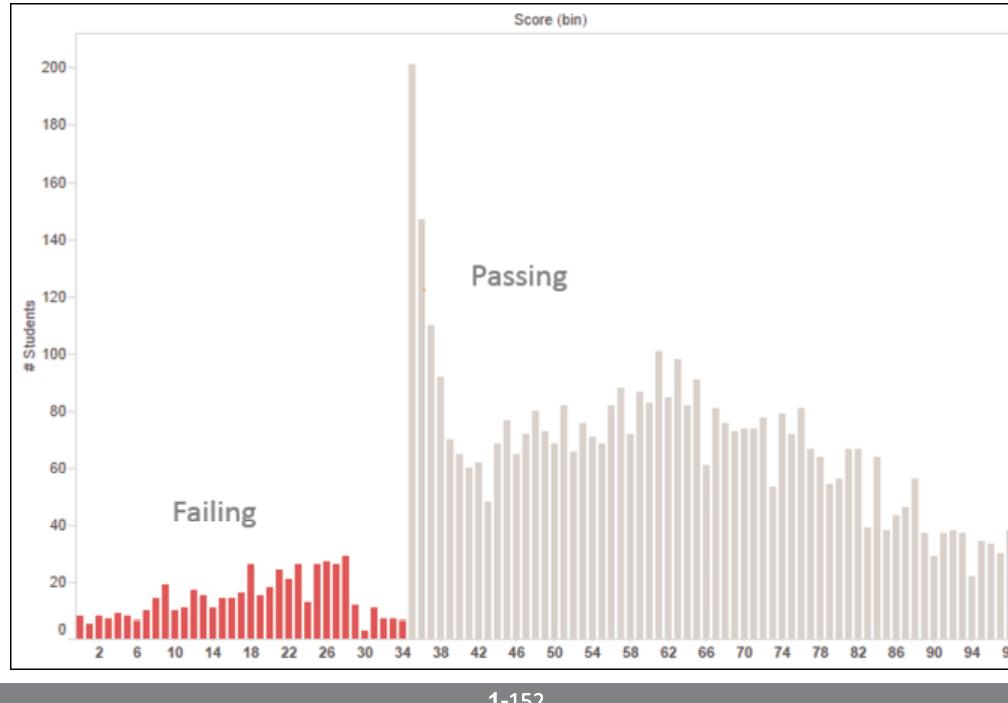


Histograms



Histograms

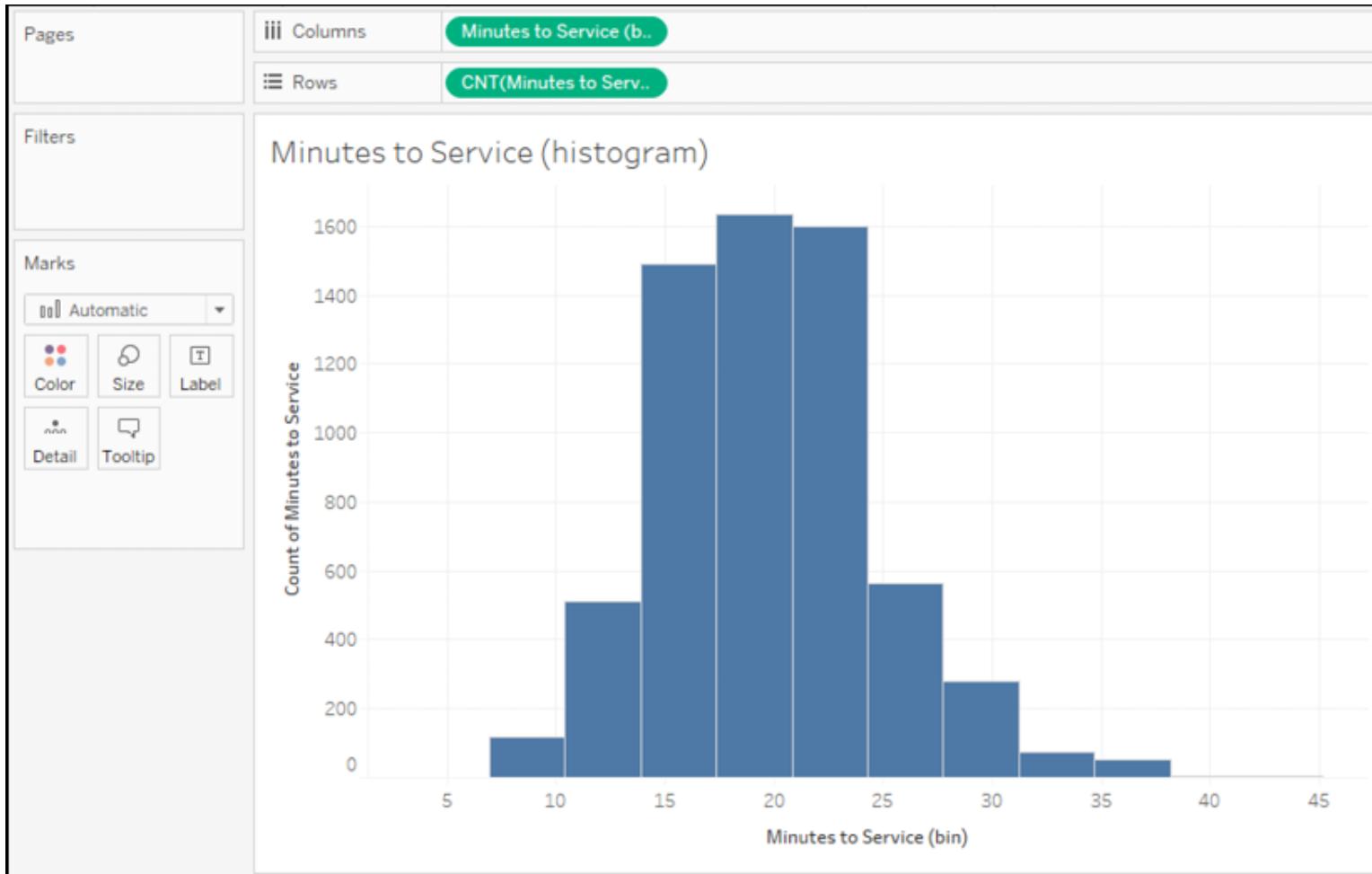
What if auditors saw something like this?



Histograms

You might start with a blank view follow steps like these:

1. Click to select the Minutes to Service field under Measures in the data pane.
2. Expand Show Me if necessary and select the histogram.



Histograms

Edit Bins [Minutes to Service] X

New field name:

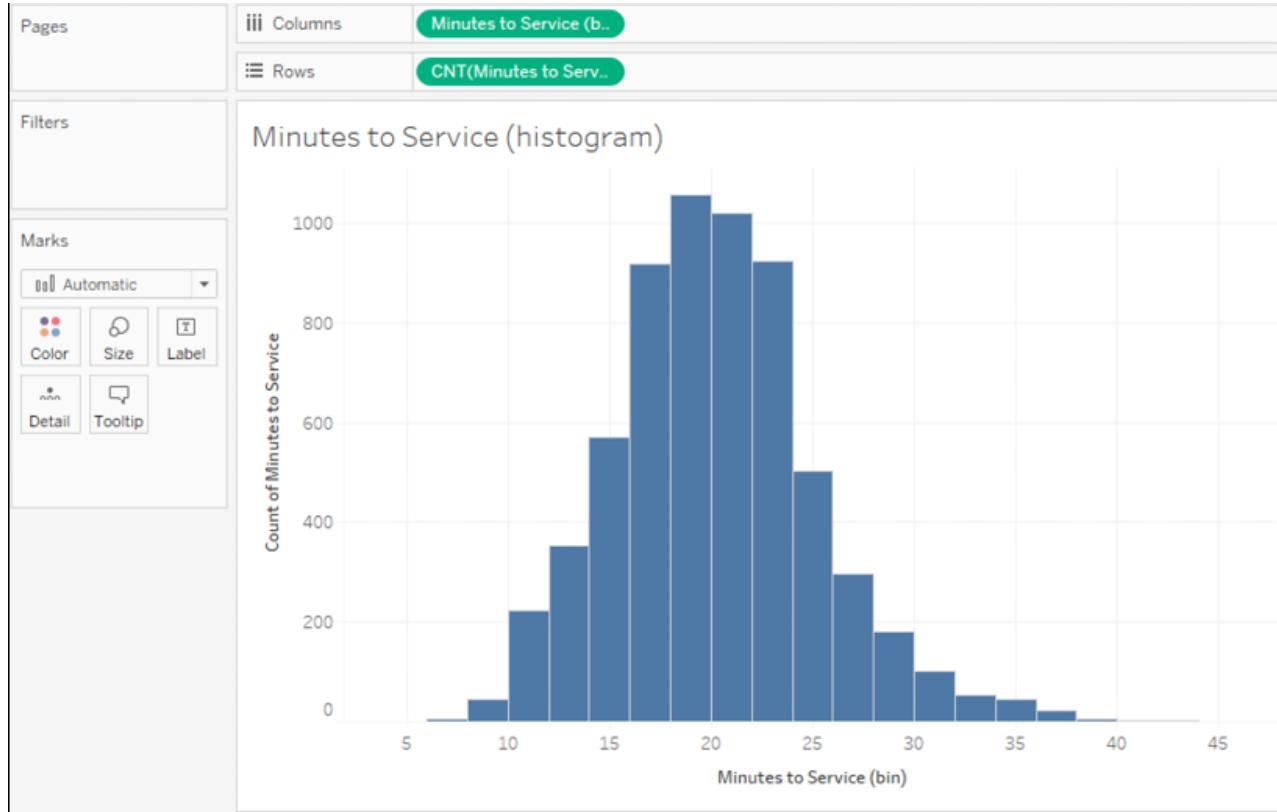
Size of bins: [Suggest Bin Size](#)

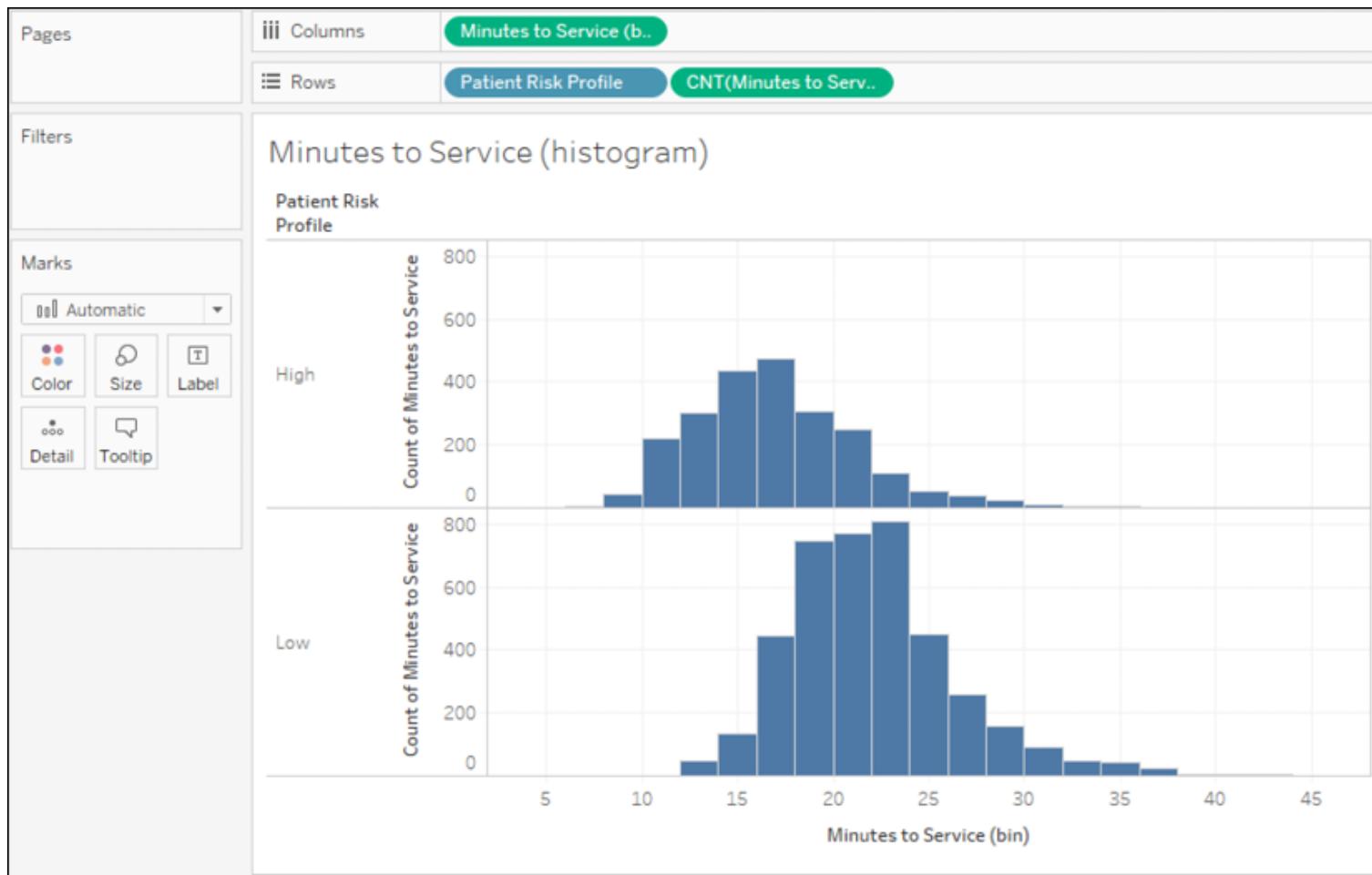
Range of Values:

Min:	<input type="text" value="4"/>	Diff:	<input type="text" value="40"/>
Max:	<input type="text" value="44"/>	CntD:	<input type="text" value="40"/>

[OK](#) [Cancel](#)

- Here, for example, is the same histogram with each bin sized to 2 minutes:

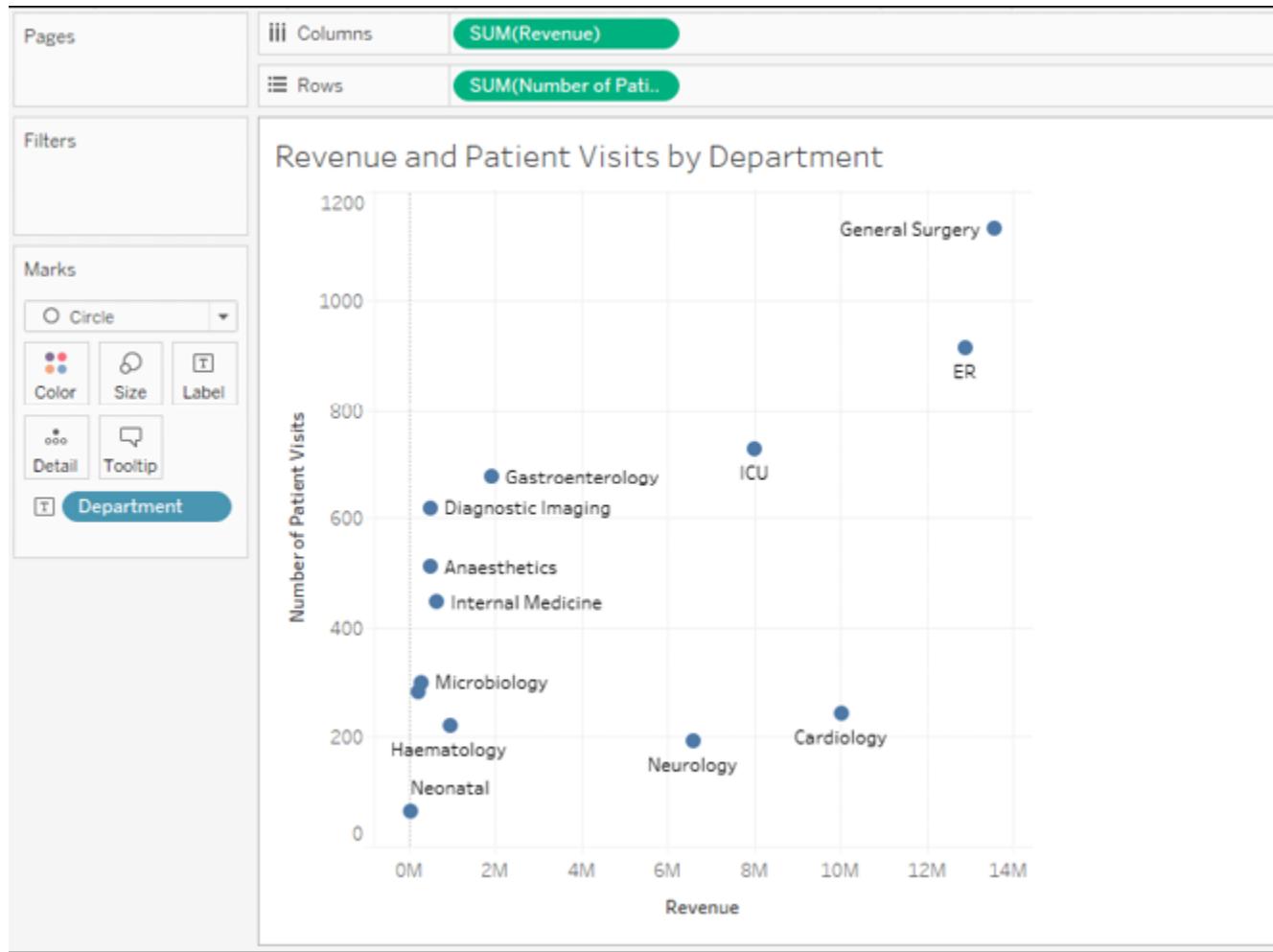




Scatterplot

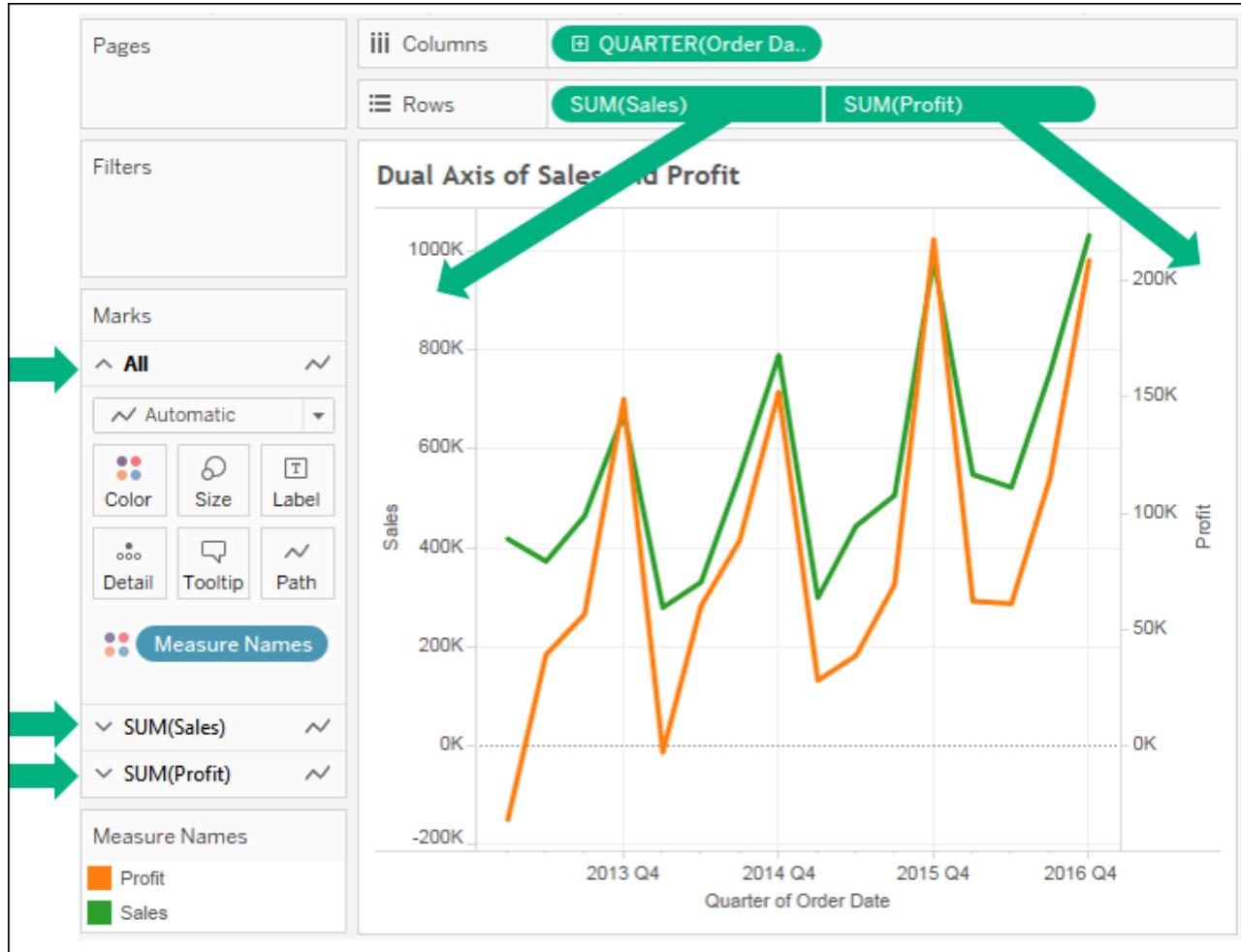
A scatterplot is an essential visualization type for understanding the relationship between two measures. Consider a scatterplot when you find yourself asking questions like the following:

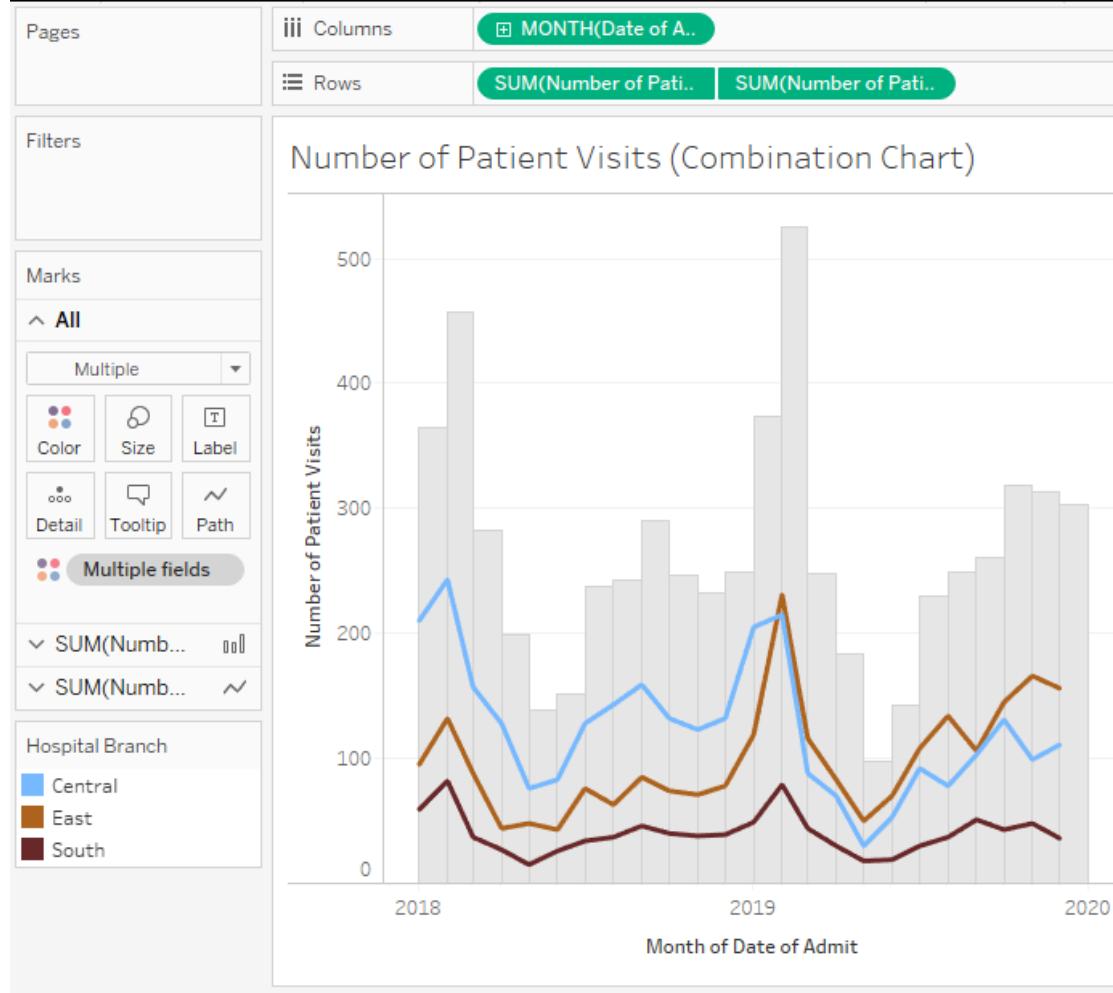
- Does how much I spend on marketing really make a difference on sales?
- How much does power consumption go up with each degree of heating/cooling?
- Is there any correlation between hours of study and test performance?



Dual-axis and combination charts

- One very important feature in Tableau is the ability to use a dual axis.
- Scatterplots use two axes, but they are X and Y.
- You also observed in the stacked bar example that placing multiple continuous (green) fields next to each other on Rows or Columns results in multiple side-by-side axes.



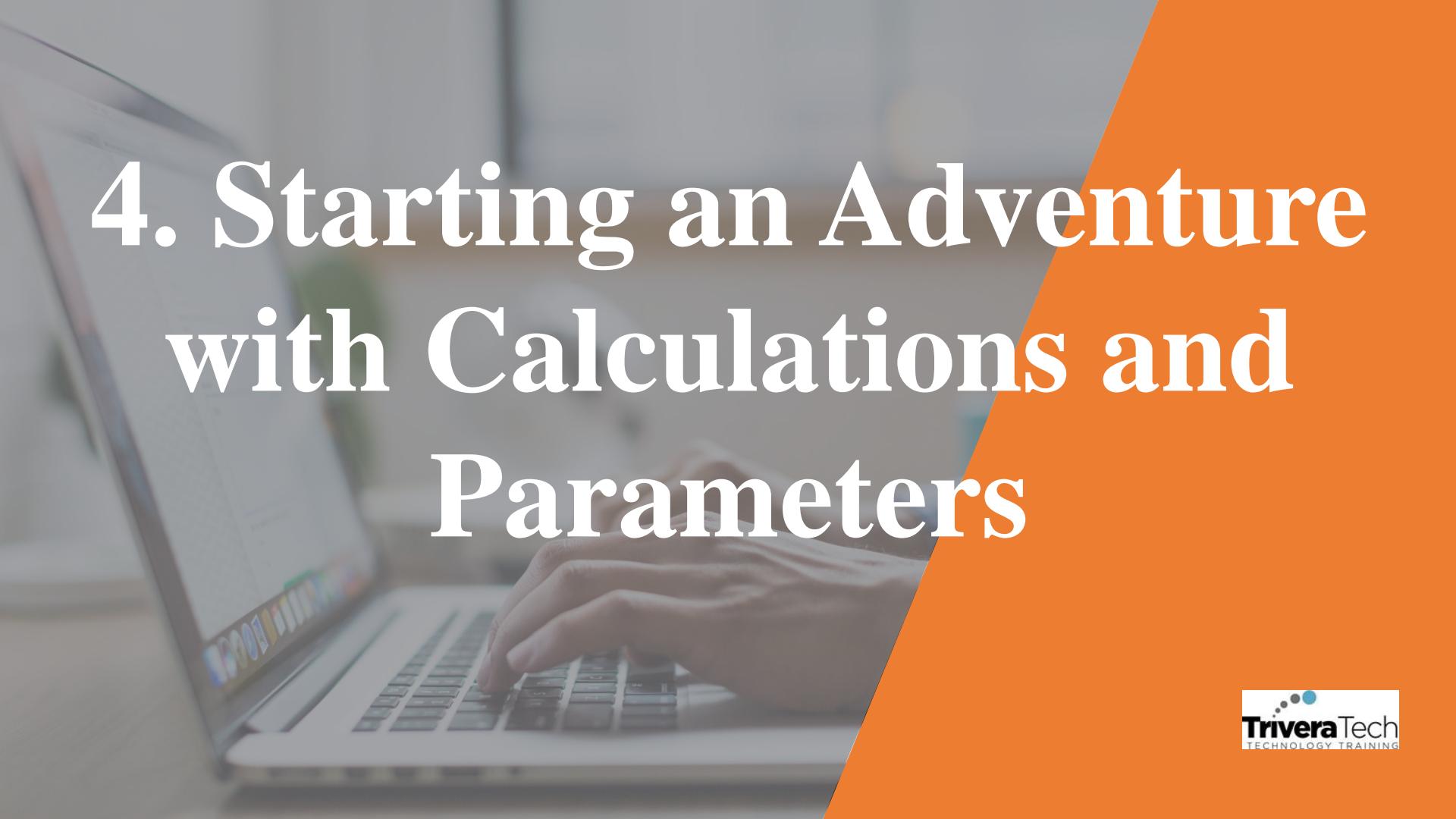


Summary

- We've covered quite a bit of ground in this lesson! You should now have a good grasp of when to use certain types of visualizations.
- The types of questions you ask about data will often lead you to a certain type of view.
- You've explored how to create these various types and how to extend basic visualizations using a variety of advanced techniques, such as calculated fields, jittering, multiple mark types, and dual axis.

COMPLETE LAB 3

4. Starting an Adventure with Calculations and Parameters

A blurred background image of a person's hands typing on a laptop keyboard, set against a dark grey gradient overlay.

Starting an Adventure with Calculations & Parameters

The topics we will study here are as follows:

- Overview of the four main types of calculations
- Creating and editing calculations
- Row-level calculation examples
- Aggregate calculation examples
- Parameters
- Practical examples
- Performance considerations

Introduction to calculations

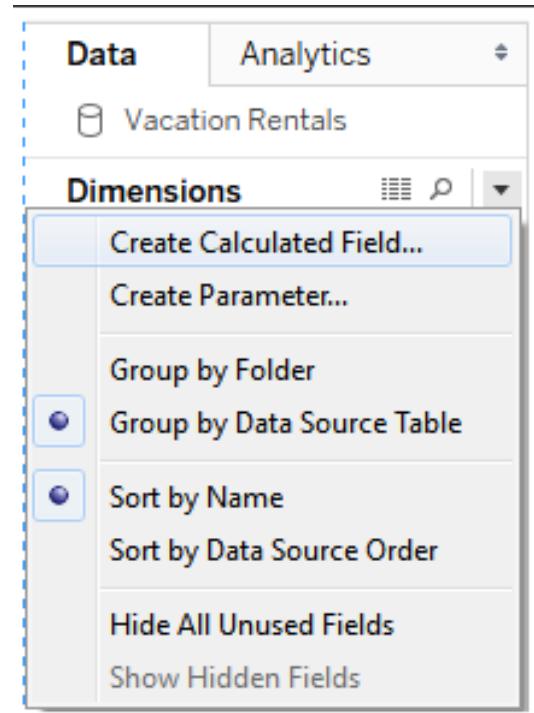
- A calculation is often referred to as a Calculated Field in Tableau because, in most cases, when you create a calculation, it will show up as either a new measure or dimension in the data pane.
- Calculations consist of code that's made up of functions, operations, and references to other fields, parameters, constants, groups, or sets.

The four main types of calculations

- **Row-level calculations:** These calculations are performed for every row of underlying data.
- **Aggregate calculations:** These calculations are performed at an aggregate level, which is usually defined by the dimensions used in the view.
- **Level of detail calculations:** These special calculations are aggregations that are performed at a specified level of detail, with the results available at the row level.
- **Table calculations:** These calculations are performed on the table of aggregate data that has been returned by the data source to Tableau.

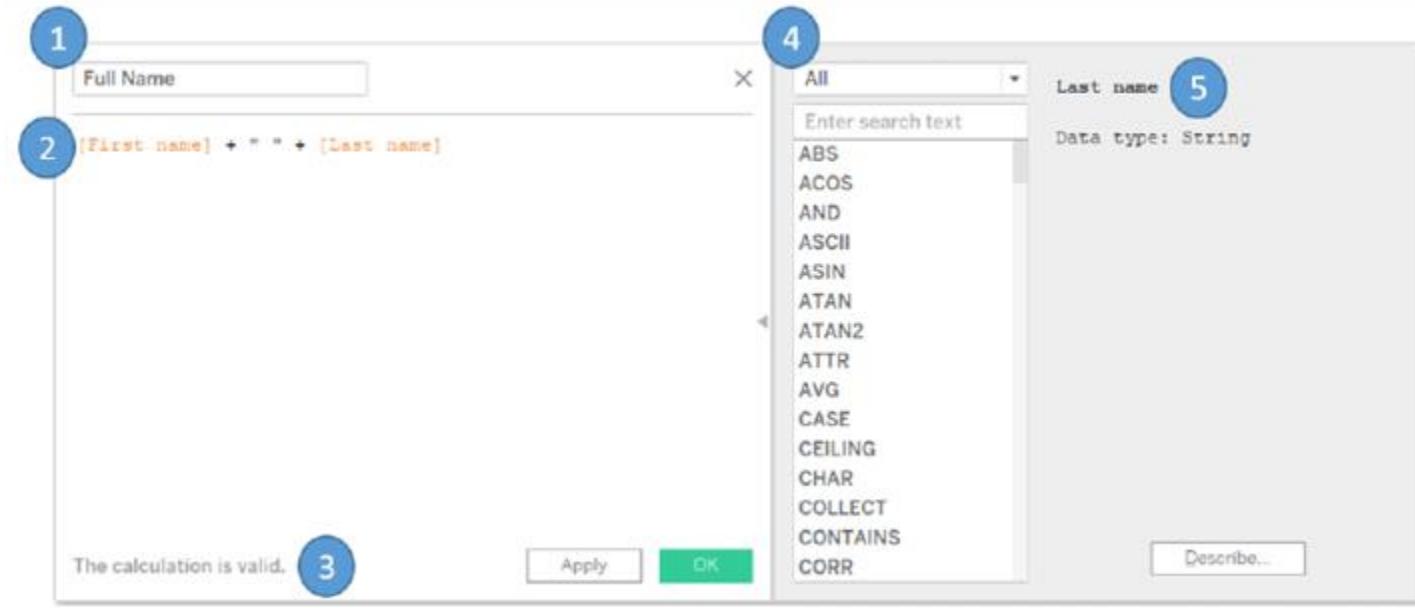
Creating and editing calculations

- Select Analysis | Create Calculated Field... from the menu.
- Use the drop-down menu next to Dimensions in the Data pane:



Creating and editing calculations

- The interface for creating and editing calculations looks like this:



Data types

- Fundamental to the concept of calculations are data types, which describe the kind of information stored by a field, parameter, or returned by a function.
- Tableau distinguishes six types of data:
- Number (decimal), Number (whole), Date and Time, Date, String, Boolean & Spatial etc.

Additional functions and operators

Operator / Keyword	Description
AND	Logical <i>and</i> between two Boolean (true/false) values or statements
OR	Logical <i>or</i> between two Boolean values or statements
NOT	Logical <i>not</i> to negate a Boolean value or statement
= or ==	Logical <i>equals to</i> to work course test the equality of two statements or values (single or double equal signs are equivalent in Tableau's syntax)
+	Addition of numeric or date values or the concatenation of strings
-	Subtraction of numeric or date values
*	Multiplication of numeric values
/	Division of numeric values
^	Raise to a power with numeric values
()	Parentheses to define the order of operations or enclose function arguments
[]	Square brackets to enclose field names
{ }	Curly braces to enclose the level of detail calculations
//	Double slash to start a comment

Example data

Rental Property	First	Last	Start	End	Discount	Rent	Tax per Night
112-Asbury Atoll	Mary	Slessor	Dec 2	Dec 9	150	1,500	15
112-Asbury Atoll	Amy	Carmichael	Dec 9	Dec 15	0	1,500	15
155-Beach Breeze	Charles	Ryrie	Dec 2	Dec 9	260	1,300	10
155-Beach Breeze	Dwight	Pentecost	Dec 16	Dec 23	280	1,400	10
207-Beach Breeze	Lewis	Chafer	Dec 9	Dec 23	280	2,800	10
207-Beach Breeze	John	Walvoord	Dec 2	Dec 9	60	1,500	10

Row-level calculations

- We'll walk through several examples of row-level calculations in this section.
- You'll find the completed calculations in the Complete workbook, but you might prefer to start from scratch in the Starter workbook.

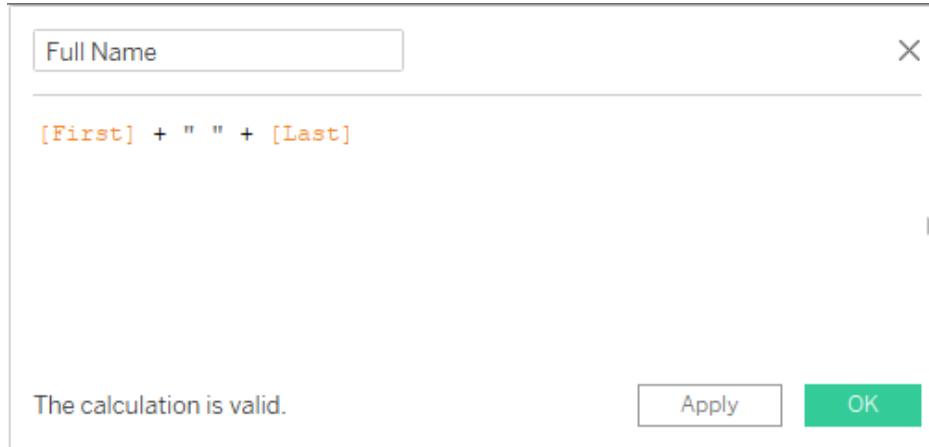
Simple example

- We'll start with a very simple example and then build up in complexity.
- In the lesson 04 workbook, create a new calculated field called Full Name with the following code:

[First] + " " + [Last]

Simple example

- Previous code concatenates the strings of First and Last with a space in-between them.
- Your calculation editor should look something like the following:

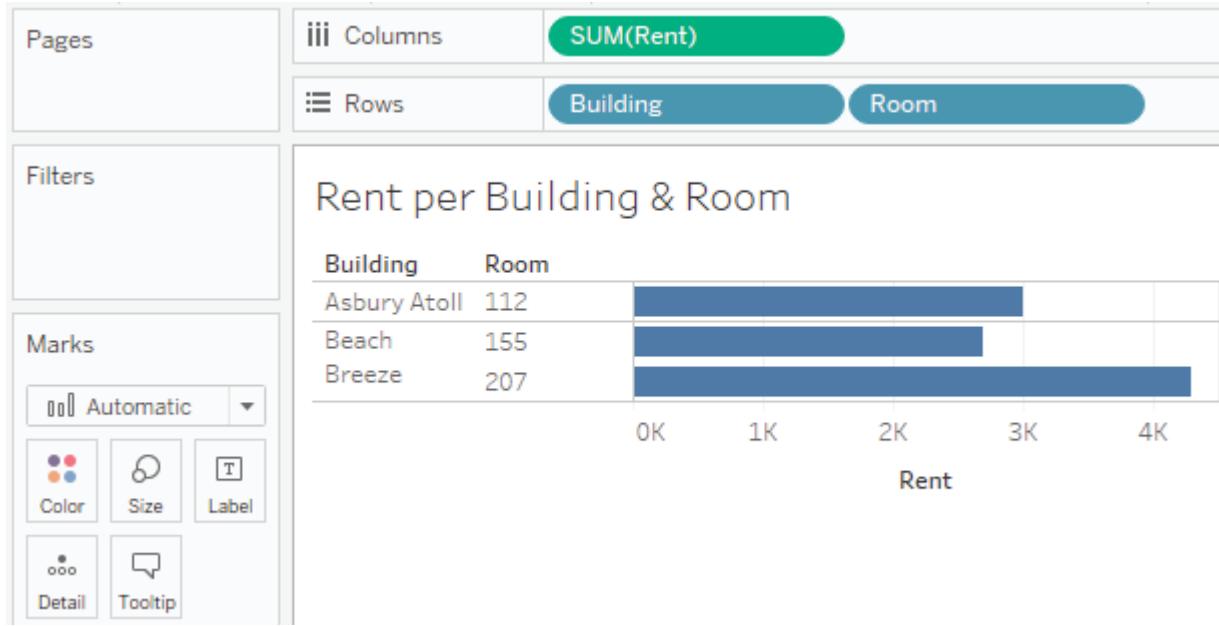


More complex examples

- Name the first Room with the following code:
SPLIT([Rental Property], "-", 1)
- Then, create another calculated field named Building with the following code:
SPLIT([Rental Property], "-", 2)

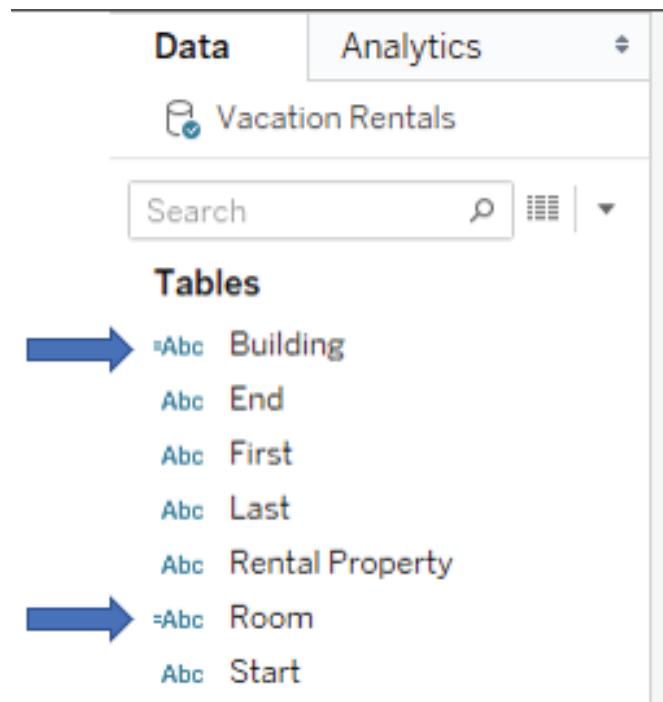
More complex examples

- Using the two calculated fields, create a bar chart of Rent per Building & Room, like this:



More complex examples

- Note that Tableau adds a small equals sign to the icon of the fields in the data pane to indicate that they are calculated fields:



The screenshot shows the Tableau Data pane. At the top, there are tabs for "Data" and "Analytics", with "Data" selected. Below the tabs, there's a search bar and a filter icon. The main area is titled "Tables" and lists several fields:

- =Abc Building
- Abc End
- Abc First
- Abc Last
- Abc Rental Property
- =Abc Room
- Abc Start

Two blue arrows point from the text "Note that Tableau adds a small equals sign to the icon of the fields in the data pane to indicate that they are calculated fields:" to the entries for "Building" and "Room".

More complex examples

- Here, the new fields of Building and Unit, along with the row-level values, can be clearly seen:

View Data: Vacation Rentals

6 rows Show aliases Copy Export All

Building	End	First	Last	Rental Property	Room	Start	Discount	Rent	Tax per Night
Asbury Atoll	9-Dec	Mary	Slessor	112-Asbury Atoll	112	2-Dec	150	1,500.00	15
Asbury Atoll	15-Dec	Amy	Carmichael	112-Asbury Atoll	112	9-Dec	0	1,500.00	15
Beach Breeze	9-Dec	Charles	Ryrie	155-Beach Breeze	155	2-Dec	130	1,300.00	10
Beach Breeze	23-Dec	Dwight	Pentecost	155-Beach Breeze	155	16-Dec	280	1,400.00	10
Beach Breeze	23-Dec	Lewis	Chafer	207-Beach Breeze	207	9-Dec	280	2,800.00	10
Beach Breeze	9-Dec	John	Walvoord	207-Beach Breeze	207	2-Dec	60	1,500.00	10

Extending the example

- We'll extend the example a bit more and assume you know that the floor of a room is indicated by its number.
- Rooms 100 through 199 are on the first floor, and 200 through 299 are on the second.
- You'd like to have that information available for analysis.

More complex examples

- Instead, you can create a row-level calculation in Tableau to extend the data.
- To do so, create a calculated field named Floor with the following code:

```
IF LEFT([Room], 1) = "1"  
THEN "First Floor"  
ELSEIF LEFT([Room], 1) = "2"  
THEN "Second Floor"  
END
```

Planning for data variations

- To account for additional cases, we might simplify our calculation to the following:

`LEFT([Room], 1)`

- This code simply returns the leftmost character of the room number.

Planning for data variations

- We'll get 3 for 306 and 8 for 822. But what if we have room numbers such as 1056 on the tenth floor, and 1617 on the sixteenth?
- We'd have to consider other options, such as the following:

`MID([Room], 0, LEN([Room]) - 2)`

Aggregate calculations

- We've already considered aggregations such as sum, min, and max in Tableau.
- Often, you'll use fields as simple aggregations in the view.
- But sometimes, you'll want to use aggregations in more complex calculations.
- For example, you might be curious to explore the percentage of the rent that was discounted, There is no such field in the data.

Aggregate calculations

- Let's create a calculation named Discount % with the following code:

$\text{SUM}([\text{Discount}]) / \text{SUM}([\text{Rent}])$

Pages

Columns Measure Names

Rows Building Room Full Name Start End

Filters

Measure Names

Marks

Automatic

Color Size Text

Detail Tooltip

Measure Values

SUM(Rent)

SUM(Discount)

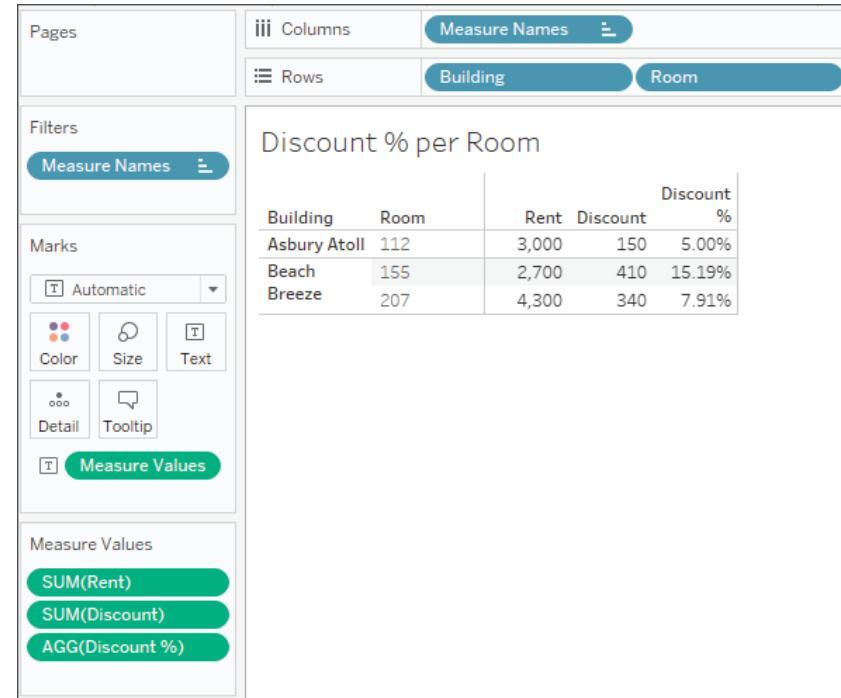
AGG(Discount %)

Discount % per Rental

Building	Room	Full Name	Start	End	Rent	Discount	Discount %
Asbury Atoll	112	Amy Carmichael	9-Dec	15-Dec	1,500	0	0.00%
		Mary Slessor	2-Dec	9-Dec	1,500	150	10.00%
Beach Breeze	155	Charles Ryrie	2-Dec	9-Dec	1,300	130	10.00%
		Dwight Pentec..	16-Dec	23-Dec	1,400	280	20.00%
	207	John Walvoord	2-Dec	9-Dec	1,500	60	4.00%
Lewis Chafer		9-Dec	23-Dec	2,800	280	10.00%	

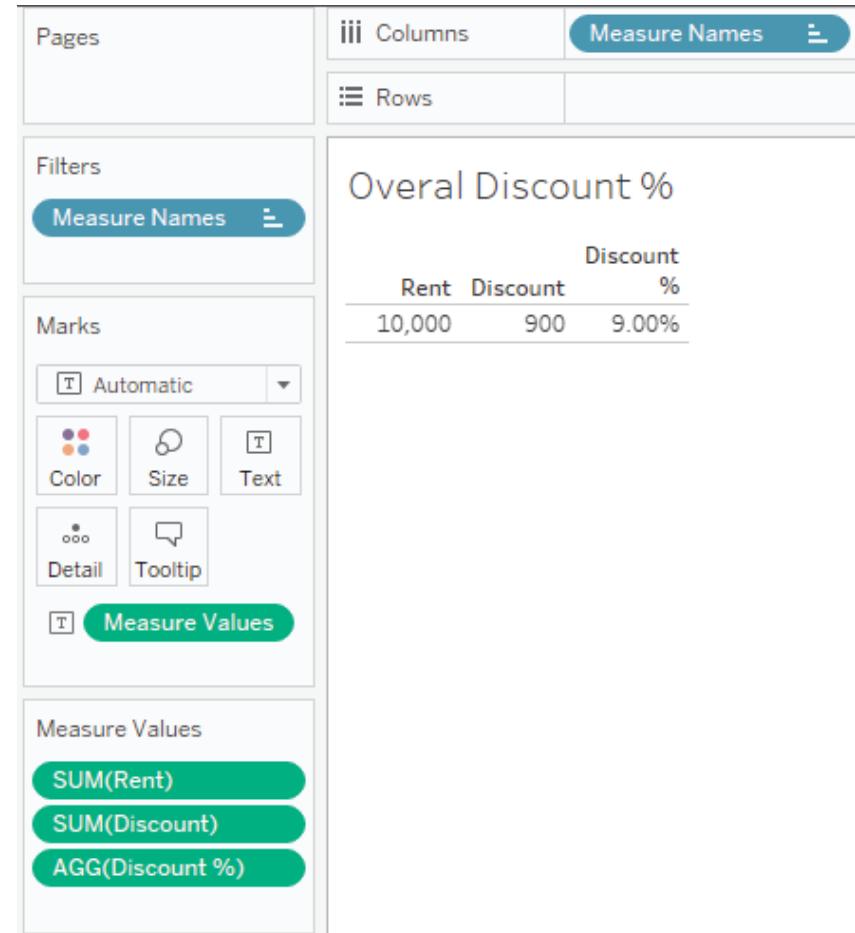
Aggregate calculations

- You can see the percentage given by way of discount for each rental period.
- However, notice how the values change when you remove all fields except Building and Room:



Aggregate calculations

- Notice that the values change again, as expected, if you look at the overall dataset without slicing by any dimensions:



Why the row level versus aggregate difference matters

- Let's say you created a Discount % (row level) calculation with the following code:
[Discount]/[Rent]
- The code differs from the aggregate calculation you created previously, which had the following code:
SUM([Discount])/SUM([Rent])

Pages

Columns Measure Names

Rows

Filters

Measure Names

Marks

Automatic

Color

Size

Text

Detail

Tooltip

Measure Values

SUM(Rent)

SUM(Discount)

AGG(Discount %)

SUM(Discount % (ro..

Overall Discount % (Aggregate v. Row Level)

Rent	Discount	Discount %	Discount % (row level)
10,000	900	9.00%	54.00%

Aggregate
 $\text{SUM}([\text{Discount}] / \text{SUM}([\text{Rent}]))$

Row Level
 $[\text{Discount}] / [\text{Rent}]$

Why the row level versus aggregate difference matters

- In fact, the row-level calculation and the final aggregation is performed like this:

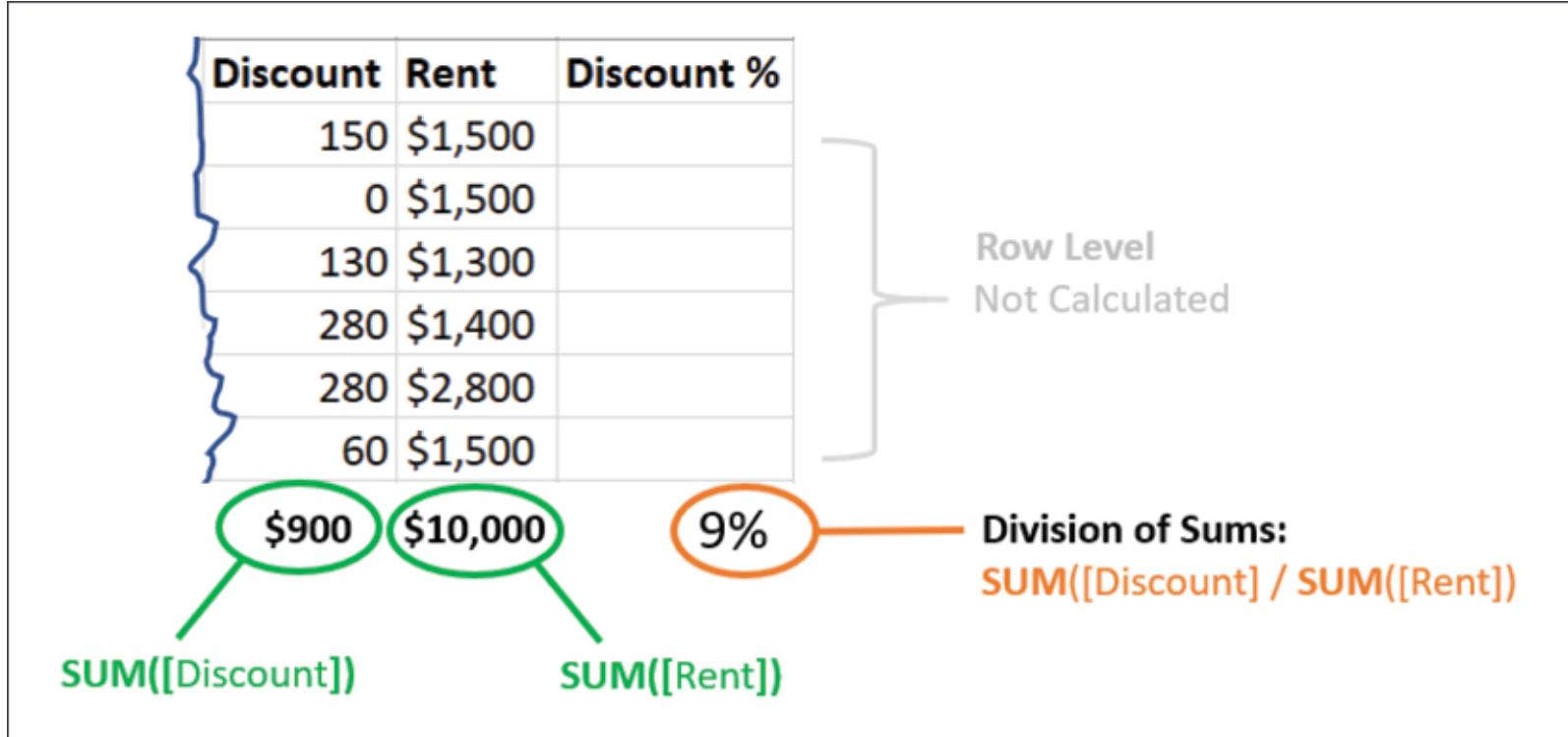
Rental Property	First	Discount	Rent	Discount %
112-Asbury Atoll	Mary	150	\$1,500	10%
112-Asbury Atoll	Amy	0	\$1,500	0%
155-Beach Breeze	Charles	130	\$1,300	10%
155-Beach Breeze	Dwight	280	\$1,400	20%
207-Beach Breeze	Lewis	280	\$2,800	10%
207-Beach Breeze	John	60	\$1,500	4%

Row Level
[Discount] / [Rent]
Results calculated for each row

Final Aggregation
Sum of row level results

54%

Why the row level versus aggregate difference matters



Parameters

- Before moving to some additional examples of row-level and aggregate calculations, let's take a little side trip to examine parameters, given that they can be used in incredible ways in calculations.
- A parameter in Tableau is a placeholder for a single, global value such as a number, date, or string.
- Parameters may be shown as controls (such as sliders, drop-down lists, or type-in text boxes) to end users of dashboards or views, giving them the ability to change the current value of the parameter.

Parameters

- Alter the results of a calculation
- Change the size of bins
- Change the number of top or bottom items in a top n filter or top n set
- Set the value of a reference line or band
- Change the size of bins
- Pass values to a custom SQL statement that's used in a data source

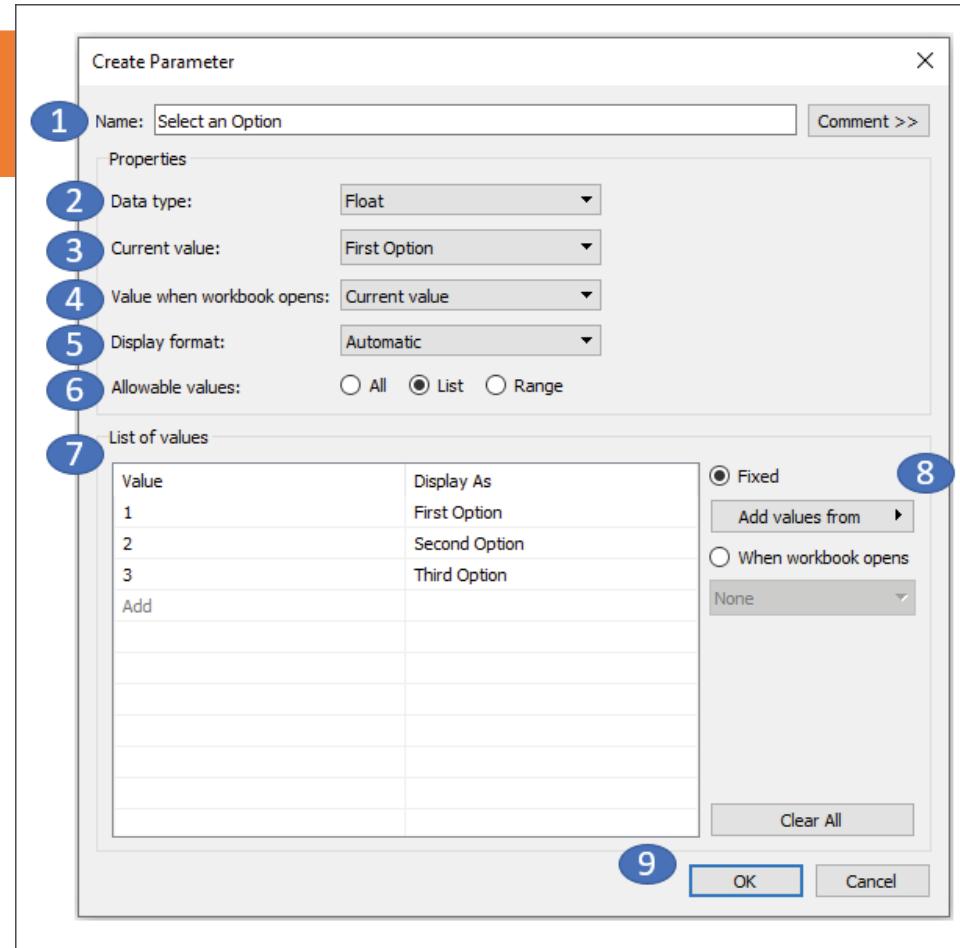
Creating parameters

Creating a parameter is similar to creating a calculated field & there are multiple ways to create a parameter in Tableau:

- Use the drop-down menu next to Dimensions in the data pane and select Create Parameter.
- Right-click an empty area in the data pane and select Create Parameter.
- Use the drop-down menu on a field, set, or parameter already in the data pane and select Create | Parameter....

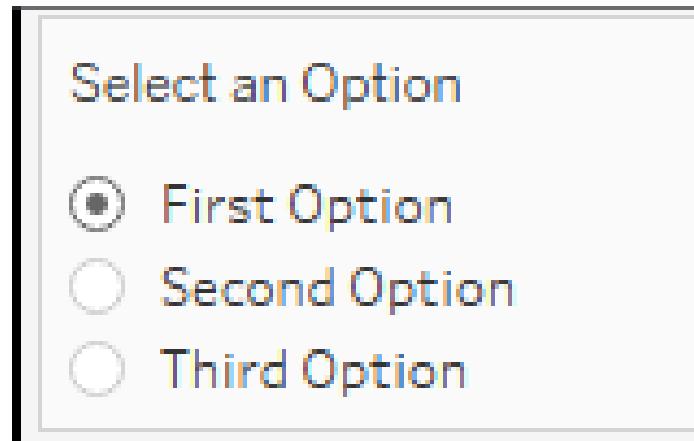
Creating parameters

- When you first create a parameter (or subsequently edit an existing parameter), Tableau will present an interface like this:



Creating parameters

- The parameter control, shown as a single value list, for the parameter we created earlier:



Practical examples of calculations and parameters

- Let's turn our attention to some practical examples of row-level and aggregate calculations.
- The goal is to learn and understand some of what is possible with calculations.
- You will be able to build on these examples as you embark on your analysis and visualization journey.

Fixing data issues

- Often, data is not entirely clean. That is, it has problems that need to be corrected before meaningful analysis can be accomplished.
- For example, dates may be incorrectly formatted, or fields may contain a mix of numeric values and character codes that need to be separated into multiple fields.
- Calculated fields can often be used to fix these kinds of issues.

Fixing data issues

- We'll continue working with the Vacation Rentals data.
- You'll recall that the start and end dates looked something like this:

Start	End
Dec 2	Dec 9
Dec 9	Dec 15
Dec 16	Dec 23

Fixing data issues

- Here is the code for getting the start date:
DATE([Start] + ", 2020")
- And here is the code for getting the end date:
DATE([End] + ", 2020")

Fixing data issues

- A quick check in Tableau reveals the expected results:

The screenshot shows a Tableau interface with the following components:

- Pages:** Shows "Corrected Date Values".
- Columns:** Shows "Rental Prop..", "Room", "Start", "End", "Start Date", and "End Date".
- Rows:** Shows data for three rental properties: "112-Asbury Atoll", "155-Beach Breeze", and "207-Beach Breeze".
- Filters:** Shows "Automatic" under Marks.
- Marks:** Shows options for Color, Size, and Text, along with Detail and Tooltip buttons.

The data table displays the following information:

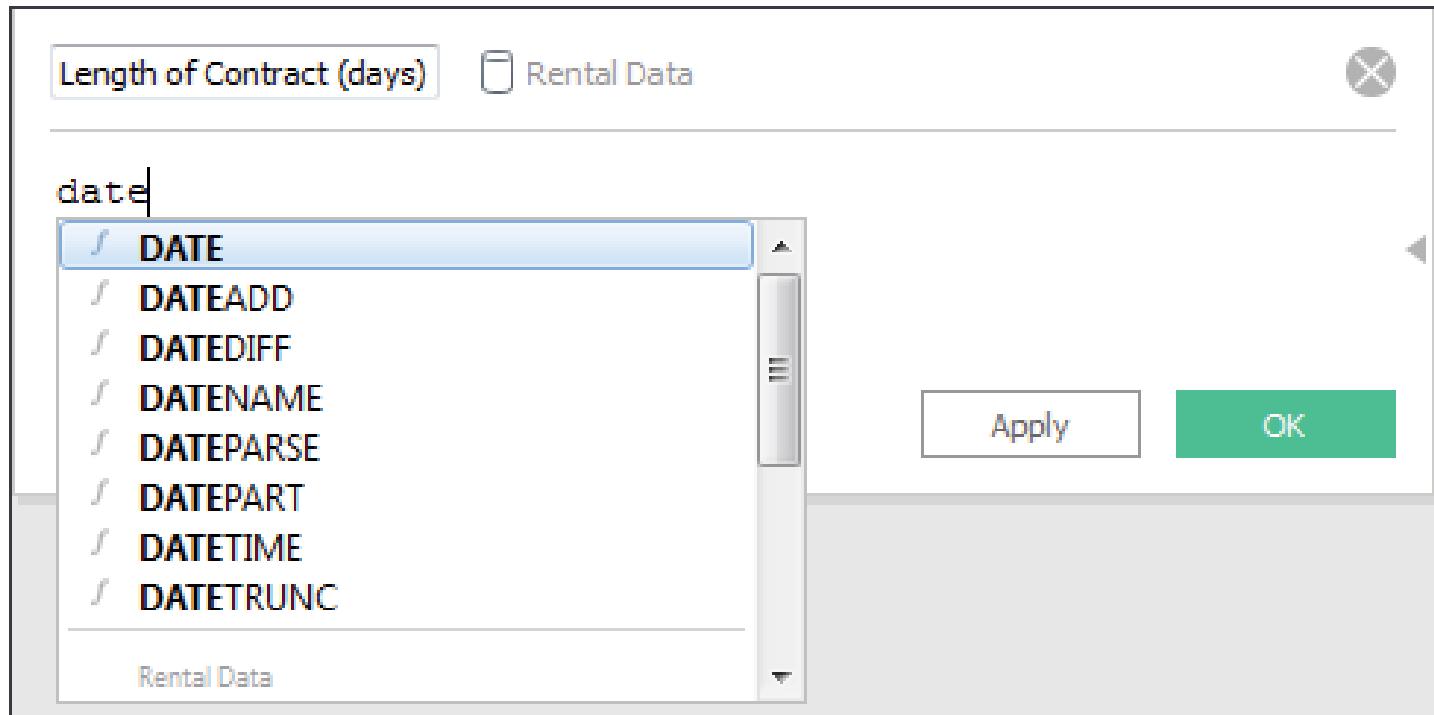
Rental Prop..	Room	Start	End	Start Date	End Date	
112-Asbury Atoll	112	2-Dec	9-Dec	12/2/2020	12/9/2020	Abc
155-Beach Breeze	155	2-Dec	9-Dec	12/2/2020	12/9/2020	Abc
207-Beach Breeze	207	16-Dec	23-Dec	12/16/2020	12/23/2020	Abc
		2-Dec	9-Dec	12/2/2020	12/9/2020	Abc
		9-Dec	23-Dec	12/9/2020	12/23/2020	Abc

Extending the data

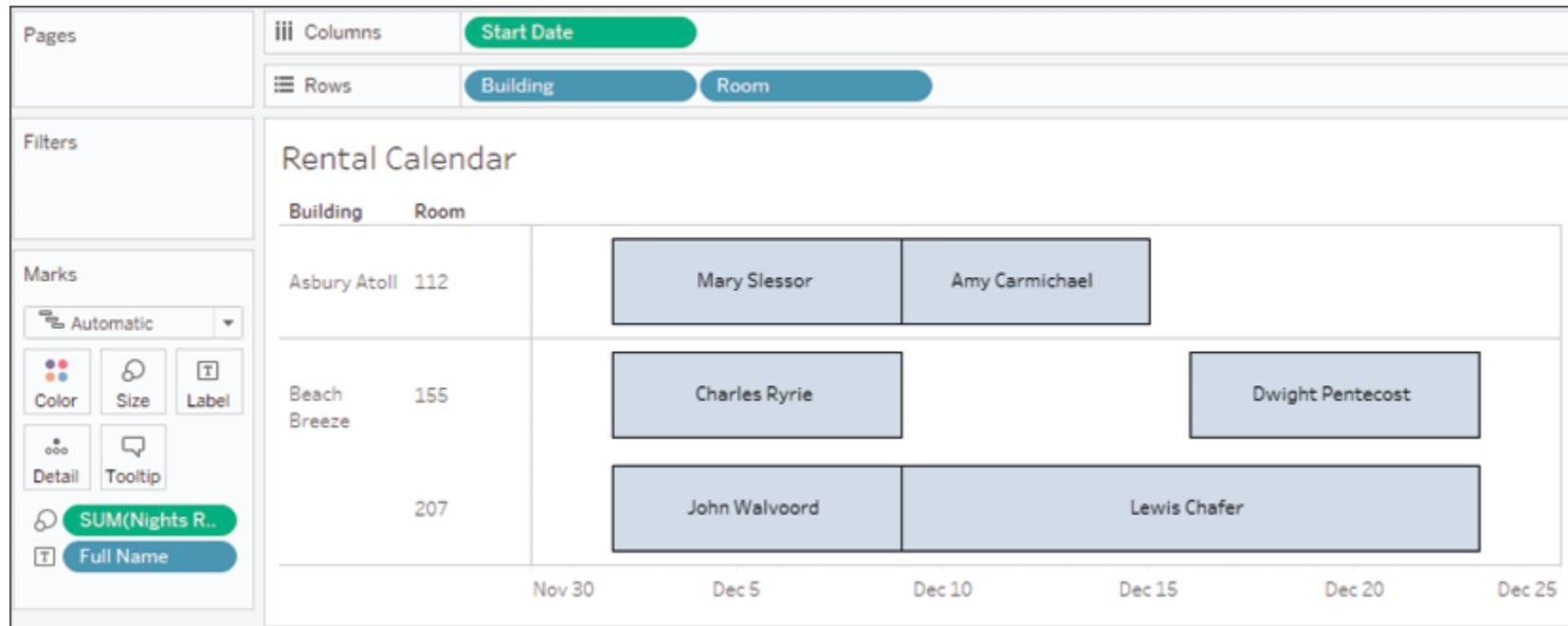
- Create a calculated field named Nights Rented with the following code:

DATEDIFF('day', [Start Date], [End Date])

Extending the data



Extending the data



Enhancing user experience, analysis, and visualizations

- Calculations and parameters can greatly enhance the user experience, the analysis, and the visualizations.
- Let's say we want to give the vacation condo manager the ability to do some what-if analysis.
- Every year, she offers a free night during the month of December.

Data **Analytics**

Summarize

- Constant Line
- Average Line
- Median with Quartiles
- Box Plot
- Totals

Model

- Average with 95% CI
- Median with 95% CI
- Trend Line
- Forecast
- Cluster

Custom

- Reference Line
- Reference Band
- Distribution Band
- Box Plot

Pages

Columns Start Date

Rows Building Room

Filters

Marks

Automatic

Color Size Label

Detail Tooltip

Gets Free Night
SUM(Days Re...
Full Name

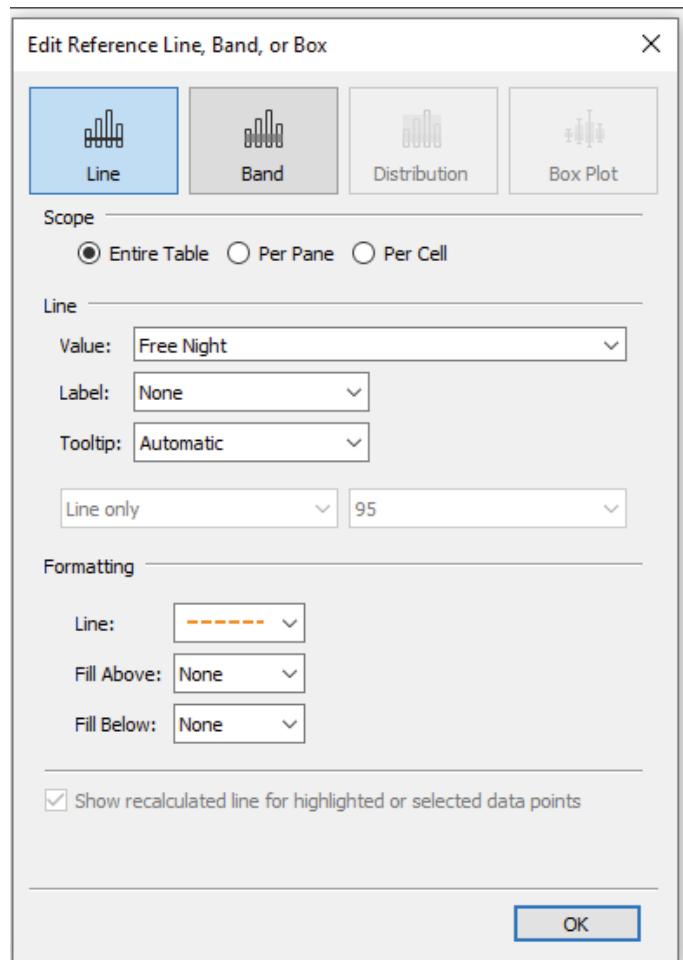
Free Night Analysis

Add a Reference Line

Table Pane Cell

Beach Breeze	155	Charles Ryrie
	207	John Walvoord
	Nov 30	Dec 5

- In the resulting dialog box, set Line Value to Free Night.
- You may wish to set the Label to None, or Custom with the text Free Night.
- You may also wish to adjust the formatting of the line:



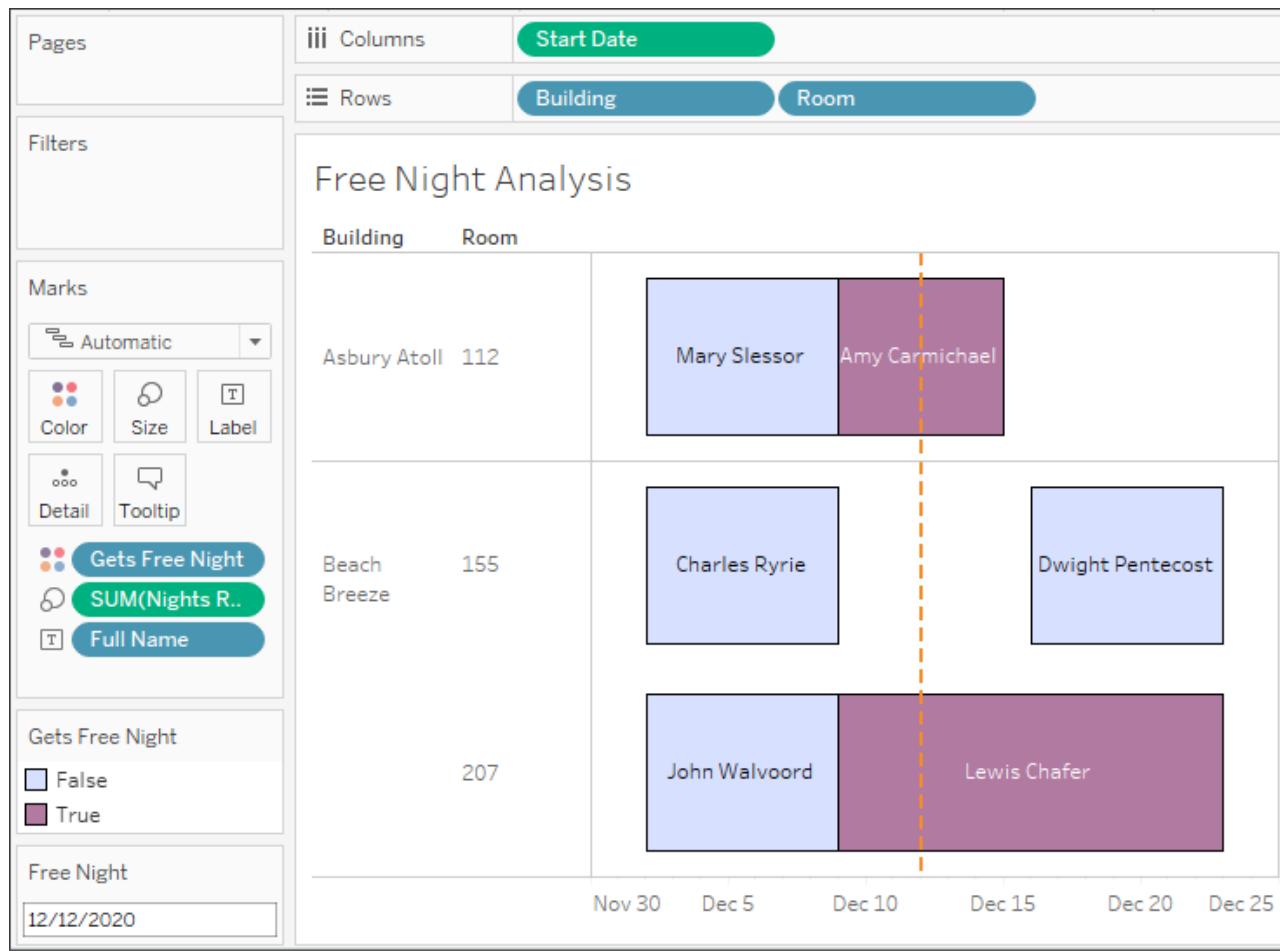
Enhancing user experience, analysis, and visualizations

- Create a calculated field called Gets Free Night that returns a true or false value, depending on whether the free night falls within the rental period:

[Free Night] >= [Start Date]

AND

[Free Night] <= [End Date]



Meeting business requirements

- Sometimes, data doesn't exactly match what your organization wants.
- For example, the measures in the data may not be the exact metrics required to make key business decisions.
- Or dimension values may need to be grouped according to a certain set of rules.

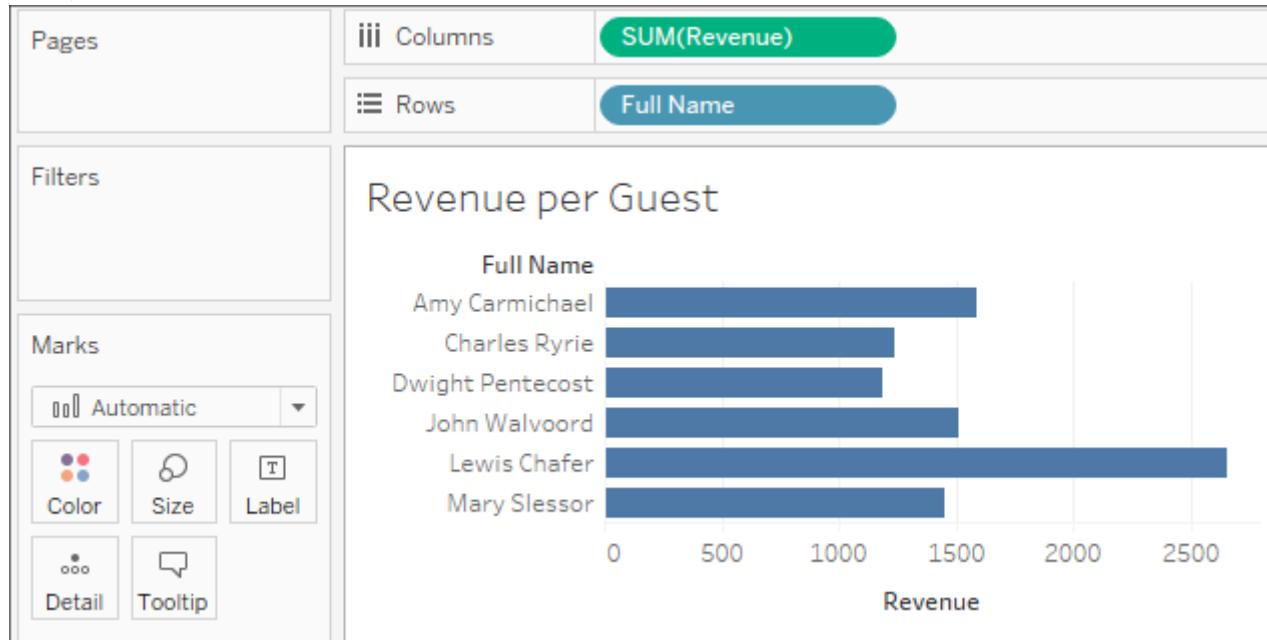
Meeting business requirements

- In this example, consider that the measure Rent is simply the base rent and does not include the discount or taxes.
- Those are separate fields in the data.
- If you needed to analyze the total Revenue, you'd need to calculate it. That calculation might look something like this:

[Rent] - [Discount] + ([Tax per Night] * [Nights Rented])

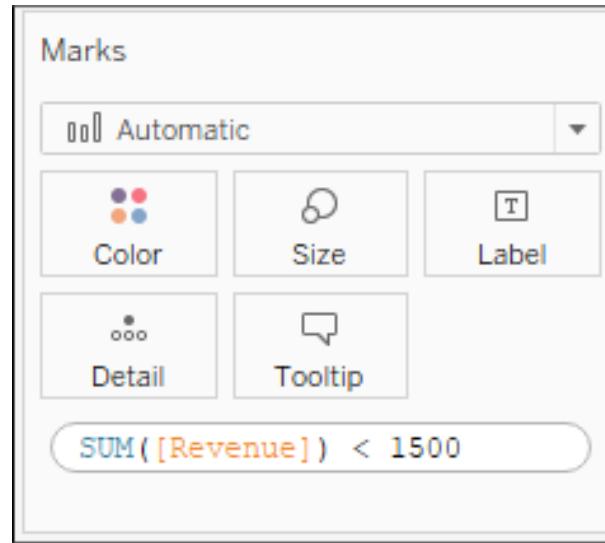
Ad hoc calculations

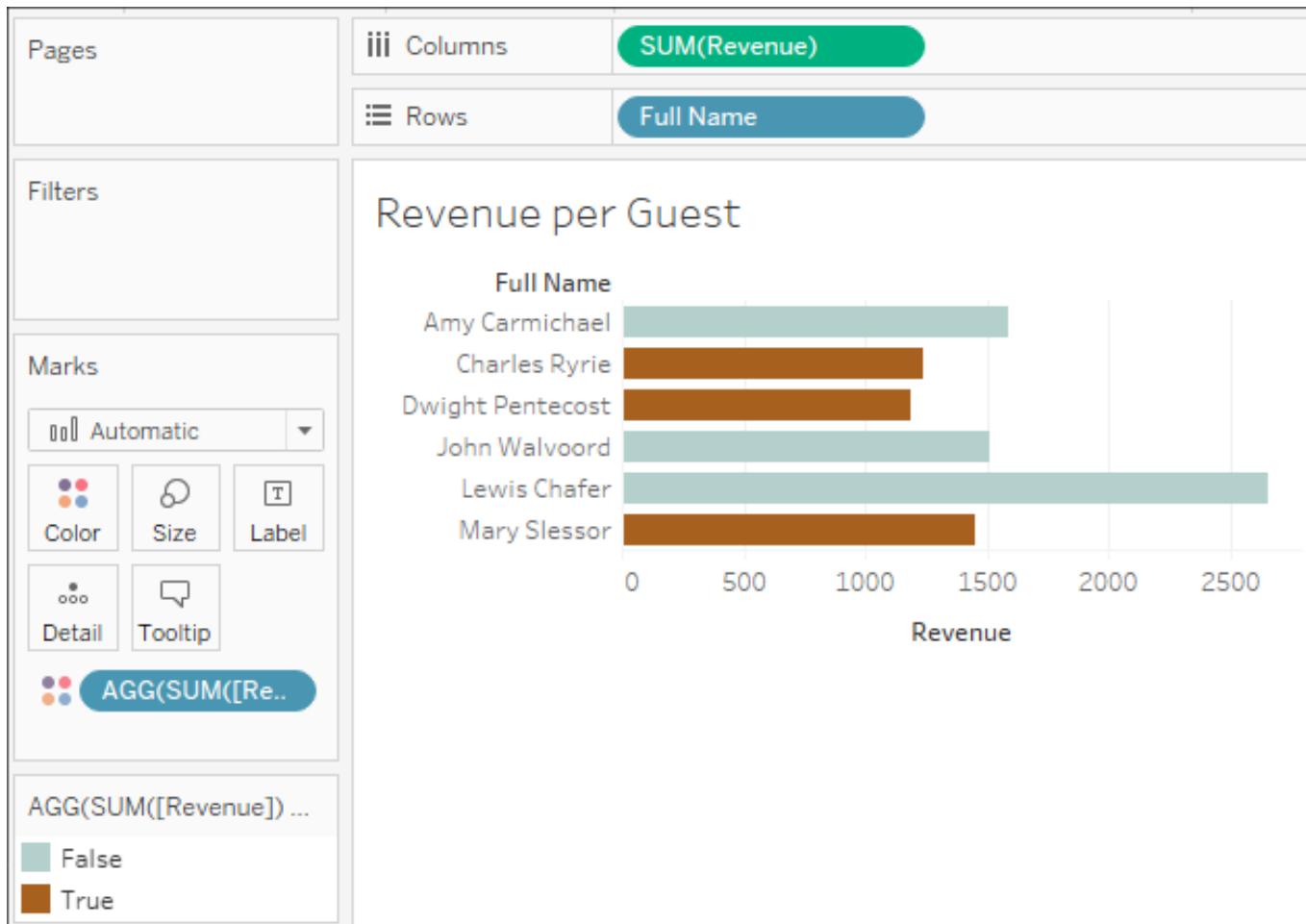
- Let's say you have a simple view that shows the Revenue per Guest, like this:



Ad hoc calculations

- In this example, we've double-clicked the empty space on the Marks shelf:





Performance considerations

- When working with a small dataset and an efficient database, you often won't notice inefficient calculations.
- With larger datasets, the efficiency of your calculations can start to make a difference to the speed at which a view is rendered.

Performance considerations

//This is potentially less efficient...

IF [Type] = "Dog" AND [Age] < 1 THEN "Puppy"

ELSEIF [Type] = "Cat" AND [Age] < 1 THEN "Kitten"

END

//...than this code:

IF [Age] < 1 THEN

 IF [Type] = "Dog" THEN "Puppy"

 ELSEIF [Type] = "Cat" THEN "Kitten"

END

END

Summary

- Calculations open amazing possibilities in Tableau.
- You are no longer confined to the fields in the source data.
- With calculations, you can extend the data by adding new dimensions and measures, fix bad or poorly formatted data, and enhance the user experience with parameters for user input and calculations that enhance the visualizations.

COMPLETE LAB 4

5. Leveraging Level of Detail Calculations

Leveraging Level of Detail Calculations

In this lesson, we'll cover the following:

- Overview of level of detail
- Level of detail calculation syntax and variations
- Examples of FIXED level of detail calculations
- Examples of INCLUDE level of detail calculations
- Examples of EXCLUDE level of detail calculations

Overview of level of detail

What does the term level of detail mean? A lot depends on the context in which the term is used. Within Tableau, we'll distinguish several levels of detail, each of which is vitally important to understand in order to properly analyze data:

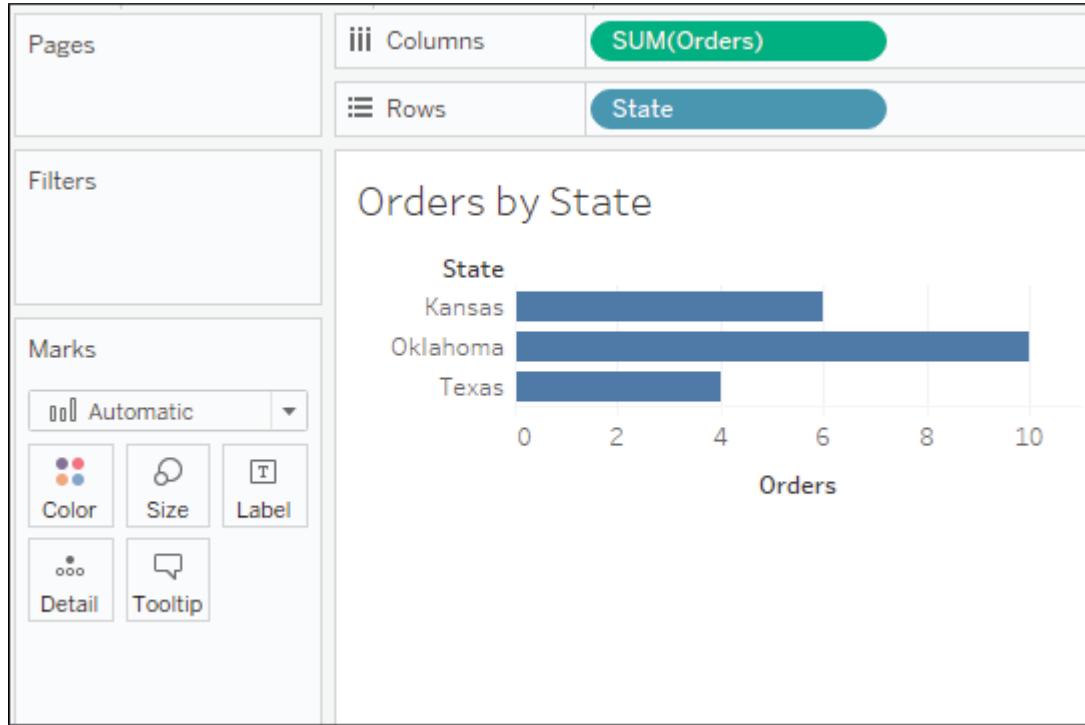
- Data level of detail
- View level of detail
- Calculated level of detail

Overview of level of detail

- Consider the following data set, with a data level of detail of one record per customer:

Customer	State	Membership Date	Membership Level	Orders
Neil	Kansas	2009-05-05	Silver	1
Jeane	Kansas	2012-03-17	Gold	5
George	Oklahoma	2016-02-01	Gold	10
Wilma	Texas	2018-09-17	Silver	4

- Now consider a view created from the data with a view level of detail of state:



Overview of level of detail

Based on that view, we might want to enhance our understanding by asking additional questions, such as the following:

- Which customer was the first member of each state in the view?
- How does the number of orders per state compare to the average number of orders for all states?
- Which membership level had the highest or lowest number of orders per state?

Level of detail calculations

Level of detail syntax

- Level of detail calculations follow this basic pattern of syntax:

{FIXED|INCLUDE|EXCLUDE [Dim 1],[Dim 2] :
AGG([Field])}

Level of detail types

- Three types of level of detail calculations are used in Tableau: **FIXED**, **INCLUDE**, and **EXCLUDE**.

FIXED

- Fixed level of detail expressions work at the level of detail that's specified by the list of dimensions in the code, regardless of what dimensions are in the view.
- For example, the following code returns the average orders per state, regardless of what other dimensions are in the view:

```
{FIXED [State] : AVG([Orders])}
```

FIXED

- You may include as many dimensions as needed or none.
- The following code represents a fixed calculation of the average orders for the entire set of data from the data source:

{**FIXED** : AVG([Orders])}

- Alternately, you might write the calculation in the following way with identical results:

{**AVG**([Orders])}

INCLUDE

- For example, the following code calculates the average orders at the level of detail that's defined by dimensions in the view, but includes the dimension Membership Level, even if Membership Level is not in the view:

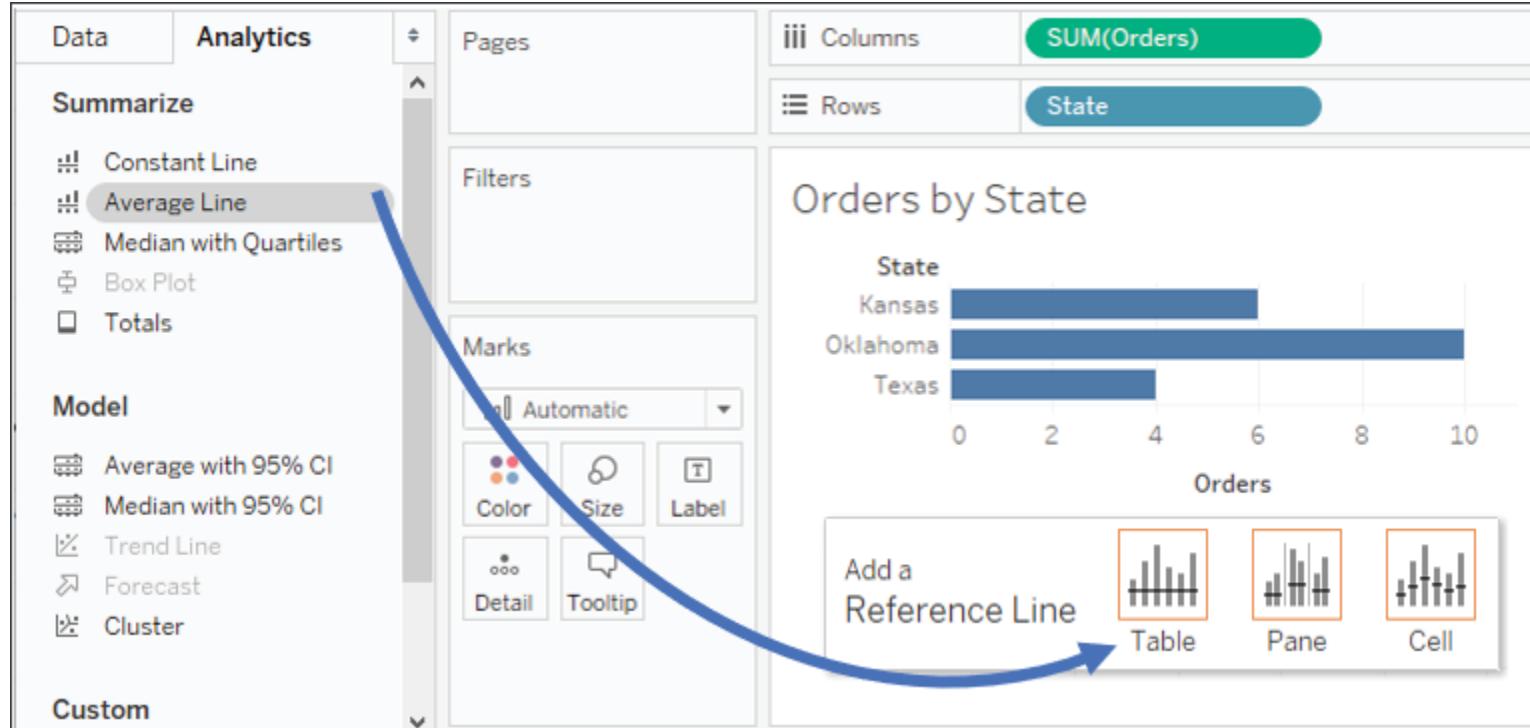
```
{INCLUDE [Membership Level] : AVG([Orders])}
```

EXCLUDE

- For example, the following code calculates the average number of orders at the level of detail defined in the view, but does not include the Customer dimension as part of the level of detail, even if Customer is in the view:

```
{EXCLUDE [Customer] : AVG([Orders])}
```

An illustration of the difference level of detail can make



An illustration of the difference level of detail can make

- You'll end up with an average line that looks like this:

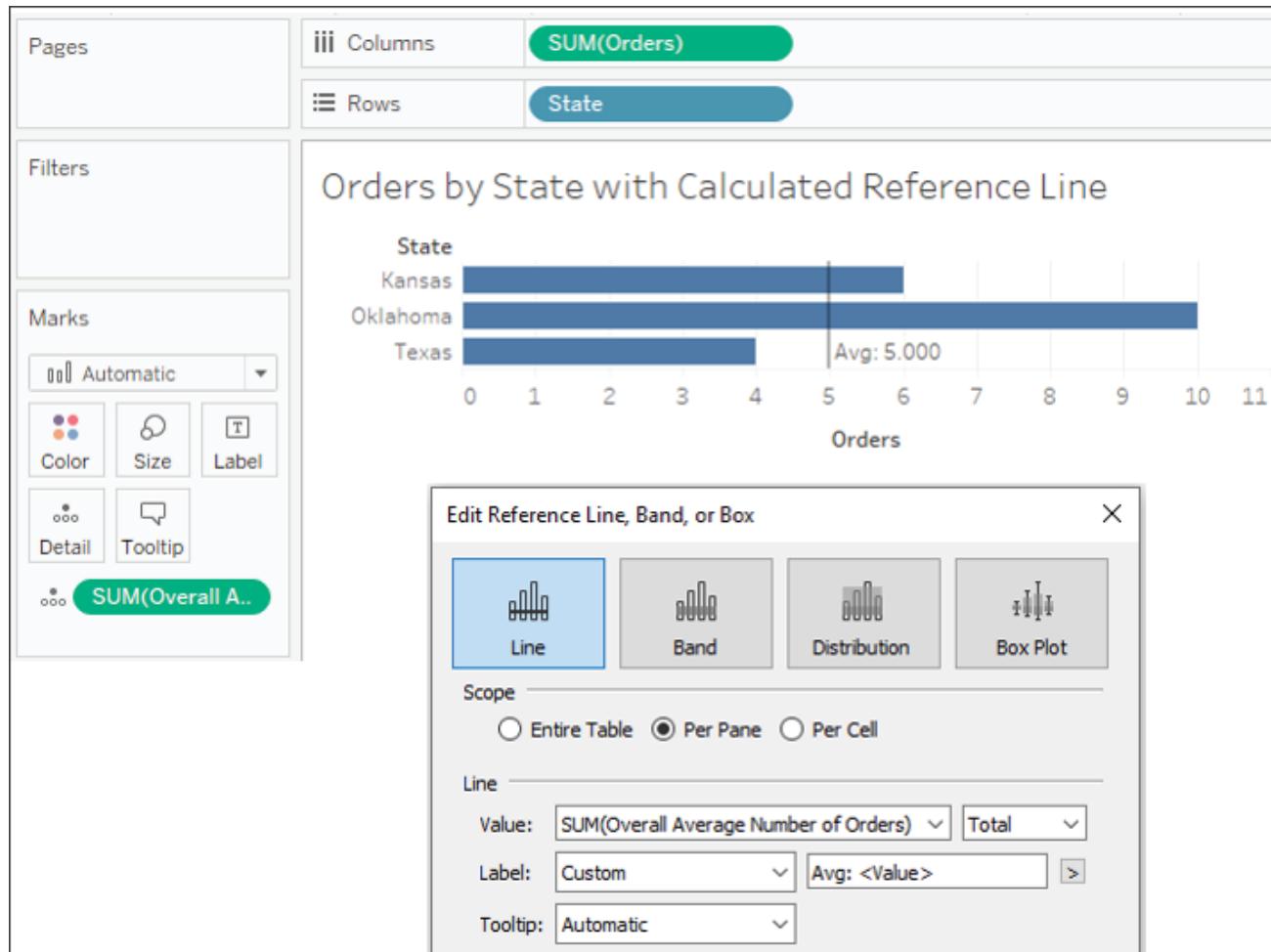
Orders by State



An illustration of the difference level of detail can make

- To get the average number of orders present in the entire data set, we might consider creating a calculation named Overall Average Number of Orders and using a fixed level of detail calculation like this:

```
{FIXED : AVG([Orders])}
```



An illustration of the difference level of detail can make

- You'll recall that the original data set had four records, and a quick check validates the result:

$$(1 + 5 + 10 + 4) / 4 = 5$$

Examples of fixed level of detail calculations

Date	Portfolio	Loan Type	Balance	Open Date	Member Name	Credit Score	Age	State
3/1/2020	Auto	New Auto	15987	9/29/2018	Samuel	678	37	California
7/1/2020	Mortgage	1st Mortgage	96364	8/7/2013	Lloyd	768	62	Ohio
3/1/2020	Mortgage	HELOC	15123	4/2/2013	Inez	751	66	Illinois
3/1/2020	Mortgage	1st Mortgage	418635	9/30/2015	Patrick	766	60	Ohio
5/1/2020	Auto	Used Auto	1151	10/22/2018	Eric	660	44	Pennsylvania
...
...

Was a member ever at risk?

Member ID	Member Name	Loan Type	Date	Balance	Credit Score
158	Vicki Modzelewski	Used Auto	1/1/2020	10,615	712
			2/1/2020	10,441	712
			3/1/2020	10,285	699
			4/1/2020	10,108	699
			5/1/2020	9,891	699
			6/1/2020	9,736	717
			7/1/2020	9,556	717
479	Thomas Villareal	Used Auto	2/1/2020	7,407	526
			3/1/2020	7,191	526
			4/1/2020	6,984	563
			5/1/2020	6,771	563
			6/1/2020	6,551	563
			7/1/2020	6,334	591
			8/1/2020	6,115	591
576	Charles Reeves	Used Auto	1/1/2020	28,145	610
			2/1/2020	27,187	610
			3/1/2020	26,226	535
			4/1/2020	25,267	535
			5/1/2020	24,302	535
			6/1/2020	23,337	530
			7/1/2020	22,366	530

Was a member ever at risk?

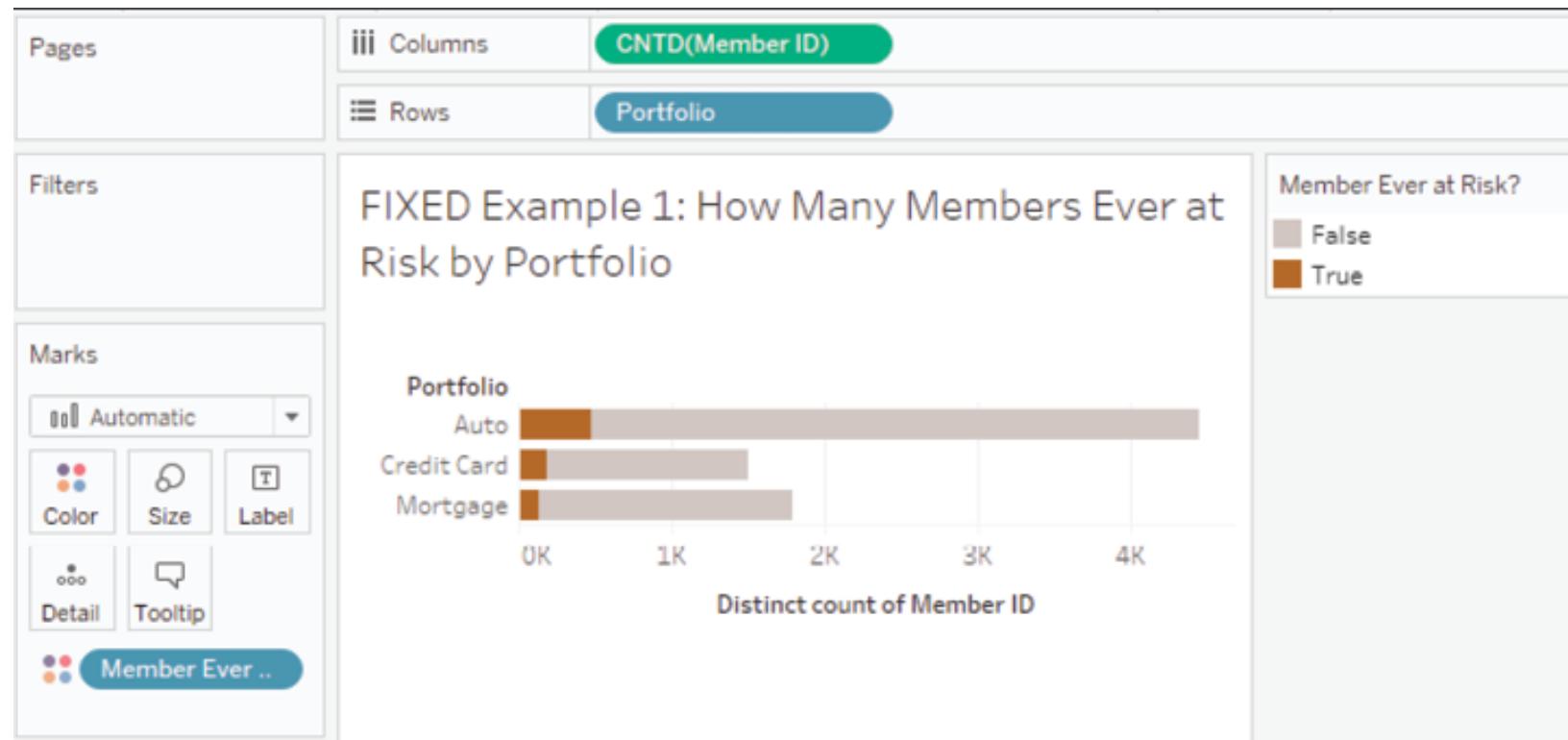
- We want every record for a given member to be TRUE if any of the records for that member are below the threshold and FALSE if none of the records are below the threshold.
- One solution is to use a level of detail calculation, which we'll name Member Ever at Risk?, with the code:

{FIXED [Member ID] : MIN([Credit Score])} < 550

FIXED Example 1: Member Ever at Risk?

Member ID	Member Name	Loan Type	Date	Member Ever at Risk?	Balance	Credit Score
158	Vicki Modzelewski	Used Auto	1/1/2020	False	10,615	712
			2/1/2020	False	10,441	712
			3/1/2020	False	10,285	699
			4/1/2020	False	10,108	699
			5/1/2020	False	9,891	699
			6/1/2020	False	9,736	717
			7/1/2020	False	9,556	717
			8/1/2020	True	9,381	717
479	Thomas Villareal	Used Auto	1/1/2020	True	7,407	526
			2/1/2020	True	7,191	526
			3/1/2020	True	6,984	563
			4/1/2020	True	6,771	563
			5/1/2020	True	6,551	563
			6/1/2020	True	6,334	591
			7/1/2020	True	6,115	591
			8/1/2020	True	5,898	591
576	Charles Reeves	Used Auto	1/1/2020	True	28,145	610
			2/1/2020	True	27,187	610
			3/1/2020	True	26,226	535
			4/1/2020	True	25,267	535
			5/1/2020	True	24,302	535
			6/1/2020	True	23,337	530
			7/1/2020	True	22,366	530
			8/1/2020	True	21,394	530

Was a member ever at risk?



Latest balance for a member

Many data sets contain a series of events or a history of transactions. You may find yourself asking questions such as:

- What diagnoses are common for a patient's first visit to the hospital?
- What was the last reported status of each computer on the network?
- How much did each customer spend on their last order?
- How much did the first trade of the week make compared to the last?

Loan					
Member ID	Member Name	Number	Loan Type	Date	
14827	Kelly Wooldridge	1	New Auto Plus	1/1/2020	21,684
				2/1/2020	21,348
				3/1/2020	21,001
				4/1/2020	21,001
				5/1/2020	20,327
				6/1/2020	19,987
				7/1/2020	19,646
16024	Joseph Clark	1	Used Auto	2/1/2020	19,043
				3/1/2020	18,656
				4/1/2020	18,263
				5/1/2020	17,873
				6/1/2020	17,479
				7/1/2020	17,087
				8/1/2020	16,691
16070	Gerald Quinney	1	1st Mortgage	3/1/2020	144,138
				4/1/2020	140,943
				5/1/2020	137,737
				6/1/2020	134,520
				7/1/2020	131,293
		2	Used Auto	3/1/2020	6,809
				4/1/2020	6,636
				5/1/2020	6,460
				6/1/2020	6,285
				7/1/2020	6,107
				8/1/2020	5,929
				9/1/2020	5,749

Latest balance for a member

- What if you want to identify only the records that represent the latest known balance for a member?
- You might consider using a fixed level of detail calculation called Latest Date per Member/Loan with code such as this:

{FIXED [Member ID],[Loan Number] : MAX([Date])} = [Date]

FIXED Example 2: Most Recent Balance

Member ID	Member Name	Loan Number	Loan Type	Date	Latest Date per Member/Loan
14827	Kelly Wooldridge	1	New Auto Plus	1/1/2020	False 21,684
				2/1/2020	False 21,348
				3/1/2020	False 21,001
				4/1/2020	False 21,001
				5/1/2020	False 20,327
				6/1/2020	False 19,987
				7/1/2020	True 19,646
16024	Joseph Clark	1	Used Auto	2/1/2020	False 19,043
				3/1/2020	False 18,656
				4/1/2020	False 18,263
				5/1/2020	False 17,873
				6/1/2020	False 17,479
				7/1/2020	False 17,087
				8/1/2020	True 16,691
16070	Gerald Quinney	1	1st Mortgage	3/1/2020	False 144,138
				4/1/2020	False 140,943
				5/1/2020	False 137,737
				6/1/2020	False 134,520
				7/1/2020	True 131,293
		2	Used Auto	3/1/2020	False 6,809
				4/1/2020	False 6,636
				5/1/2020	False 6,460
				6/1/2020	False 6,285
				7/1/2020	False 6,107
				8/1/2020	False 5,929
				9/1/2020	True 5,749

- You can see the results of the calculation here:

Example of include level of detail expressions

- Include level of detail calculations can be very useful when you need to perform certain calculations at levels of detail that are lower (more detailed) than the view level of detail.
- Let's take a look at an example.

Pages

Columns

Longitude (generated)

Filters

Rows

Latitude (generated)

Marks

Map

Color

Size

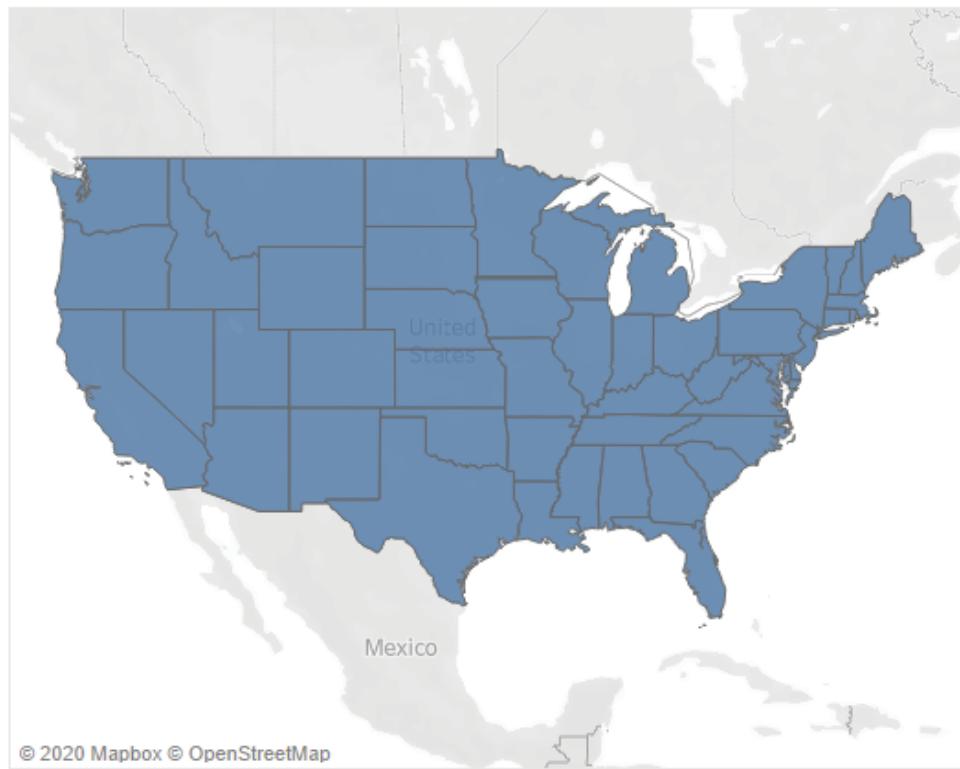
Label

Detail

Tooltip

State

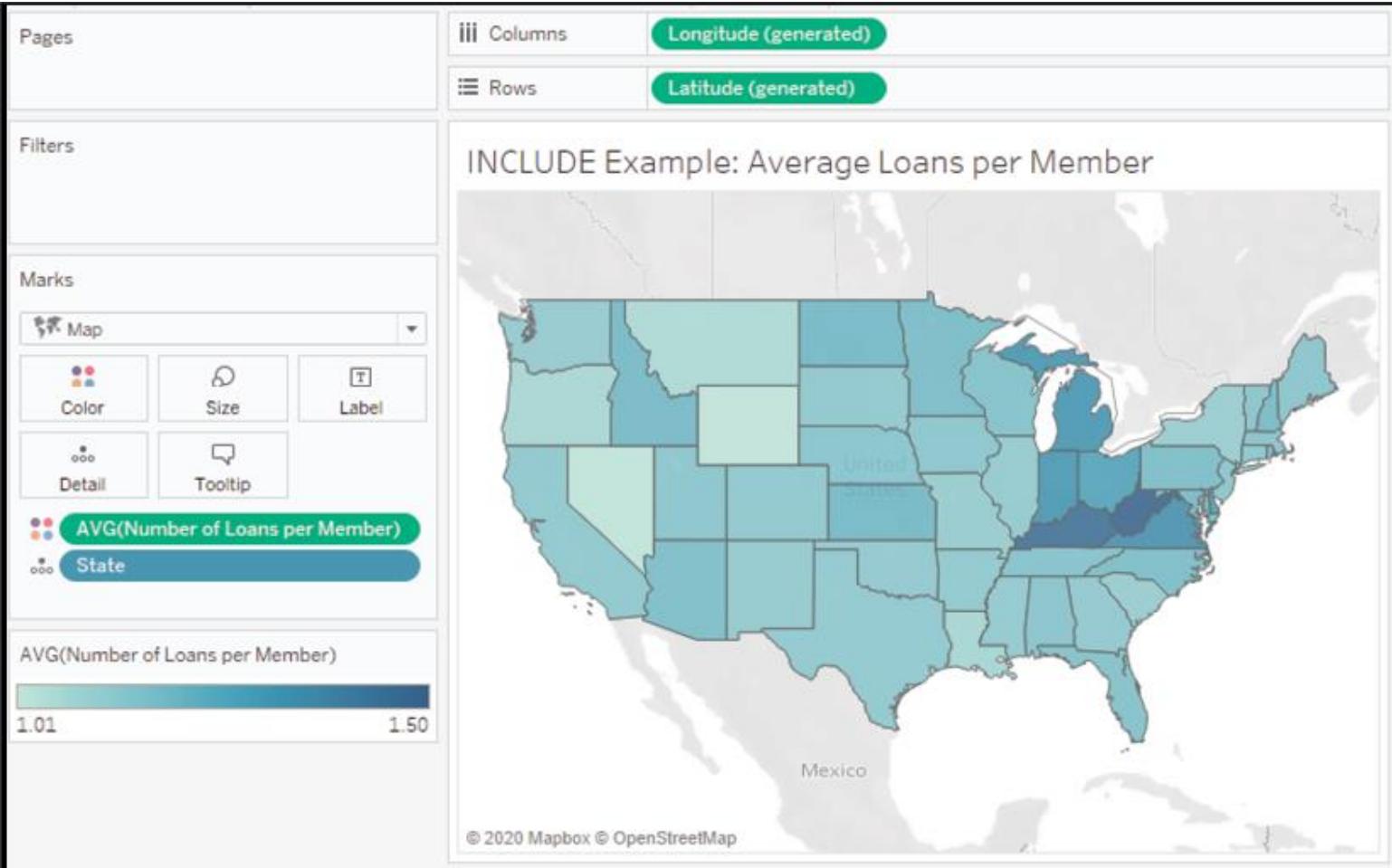
INCLUDE Example: Average Loans per Member



Average loans per member

- While there are several possible approaches to solving this kind of problem, here we'll consider using the following level of detail expression named Number of Loans per Member:

```
{INCLUDE [Member ID] : COUNTD([Loan Number])}
```



Pages

iii Columns

Rows State Member ID

Filters

State: North Dakota

Marks

Automatic

Color Size Text

Detail Tooltip

CNTD(Loan Number)

Crosstab to illustrate

State	Member ID
North	1576
Dakota	2557
	2660
	3548
	4491
	4656
	5365
	6460
	6517
	9601
	10147
	11258
	11895
	19140
	28925

Alternative approaches

- It's worth noting that the above dataset actually allows you to use MAX([Loan Number]) instead of COUNTD([Loan Number]) as the number simply increments for each member based on how many loans they have.
- The highest number is identical to the number of loans for that member.

Alternative approaches

- There are also a few other approaches to solving this problem, such as the calculation.
- For example, you could write the following code:

```
COUNTD(STR([Member ID]) + "_" + STR([Loan  
Number]))  
/  
COUNTD([Member ID])
```

Alternative approaches

- Another approach would be to use a fixed level of detail expression, such as:

```
{FIXED [State],[Member ID] : COUNTD([Loan Number])}
```

Example of exclude level of detail calculations

- Exclude level of detail calculations are useful when you want to perform certain calculations at higher (less detailed) levels than the view level of detail.
- The following example will demonstrate how we can leverage this functionality.

Average credit score per loan type

The screenshot shows a Tableau dashboard with the following components:

- Pages:** A section labeled "Pages" with a "Columns" button.
- Filters:** A section labeled "Filters".
- Marks:** A section labeled "Marks" containing controls for "Automatic", "Color", "Size", "Text", "Detail", and "Tooltip". Below these is a green button labeled "AVG(Credit Score)".
- Rows:** A section labeled "Rows" with buttons for "Portfolio" (highlighted in blue) and "Loan Type".
- Table:** A data table titled "Average Credit Scores by Loan Type" with the following data:

Portfolio	Loan Type	Avg Credit Score
Auto	New Auto	676.85
	New Auto Plus	722.77
	Used Auto	680.36
	Used Auto Plus	700.56
Credit Card	Gold Card	668.81
	Platinum Card	724.99
	Standard Card	665.71
Mortgage	1st Mortgage	704.53
	HELOC	710.02

Average credit score per loan type

- What if we wanted to compare the average credit score of each loan type with the overall average credit score for the entire portfolio?
- We could accomplish this with an exclude level of detail calculation that looks like this:

{EXCLUDE [Loan Type] : AVG([Credit Score])}

Pages

iii Columns Measure Names

Rows Portfolio Loan Type

Filters

Measure Names

Marks

Automatic

Color Size Text

Detail Tooltip

Measure Values

ATTR(Average Credit Score Excl..)

AVG(Credit Score)

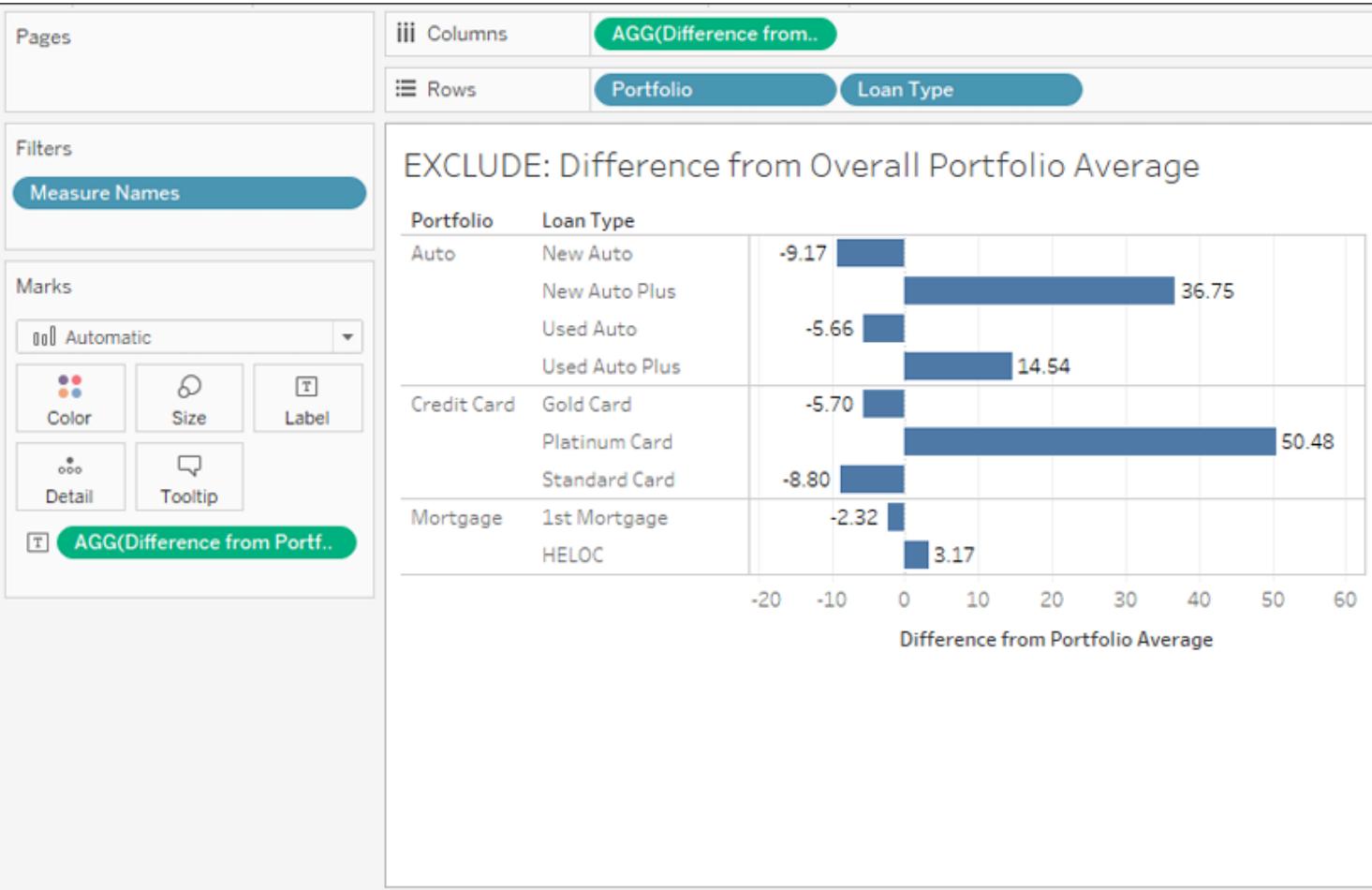
EXCLUDE: Avg Credit Scores

Portfolio	Loan Type	Average Credit Score Excluding Loan Type	Avg. Credit Score
Auto	New Auto	686.02	676.85
	New Auto Plus	686.02	722.77
	Used Auto	686.02	680.36
	Used Auto Plus	686.02	700.56
Credit Card	Gold Card	674.51	668.81
	Platinum Card	674.51	724.99
	Standard Card	674.51	665.71
Mortgage	1st Mortgage	706.85	704.53
	HELOC	706.85	710.02

Average credit score per loan type

- We can extend the calculation a bit to give us the difference between the overall portfolio average and the average of each loan type.
- The code would look like this:

AVG([Credit Score]) - AVG([Average Credit Score
Excluding Loan Type])

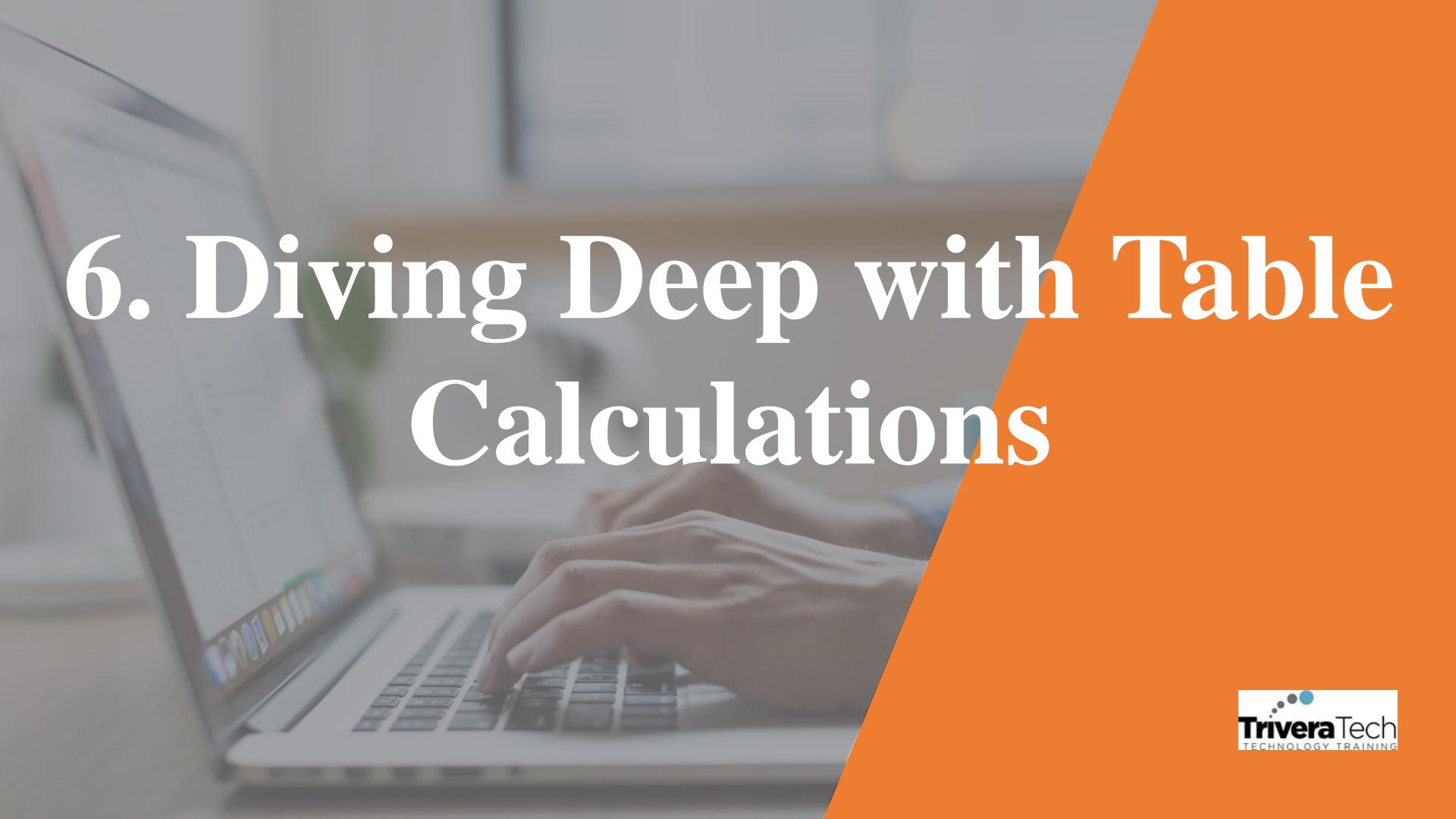


Summary

- Level of detail expressions greatly extend what you can accomplish with calculations.
- You now have a toolset for working with data at different levels of detail.
- With fixed level of detail calculations, you can identify the first or last event in a series or whether a condition is ever true across entire subsets of data.

COMPLETE LAB 5

6. Diving Deep with Table Calculations

A blurred background image of a person's hands typing on a laptop keyboard, positioned on the left side of the slide. The right side of the slide features a solid orange diagonal band.

Diving Deep with Table Calculations

- They make it possible to use data that isn't structured well and still get quick results without waiting for someone to fix the data at the source.
- They make it possible to compare and perform calculations on aggregate values across rows of the resulting table.
- They open incredible possibilities for analysis and creative approaches to solving problems, highlighting insights, or improving the user experience.

Diving Deep with Table Calculations

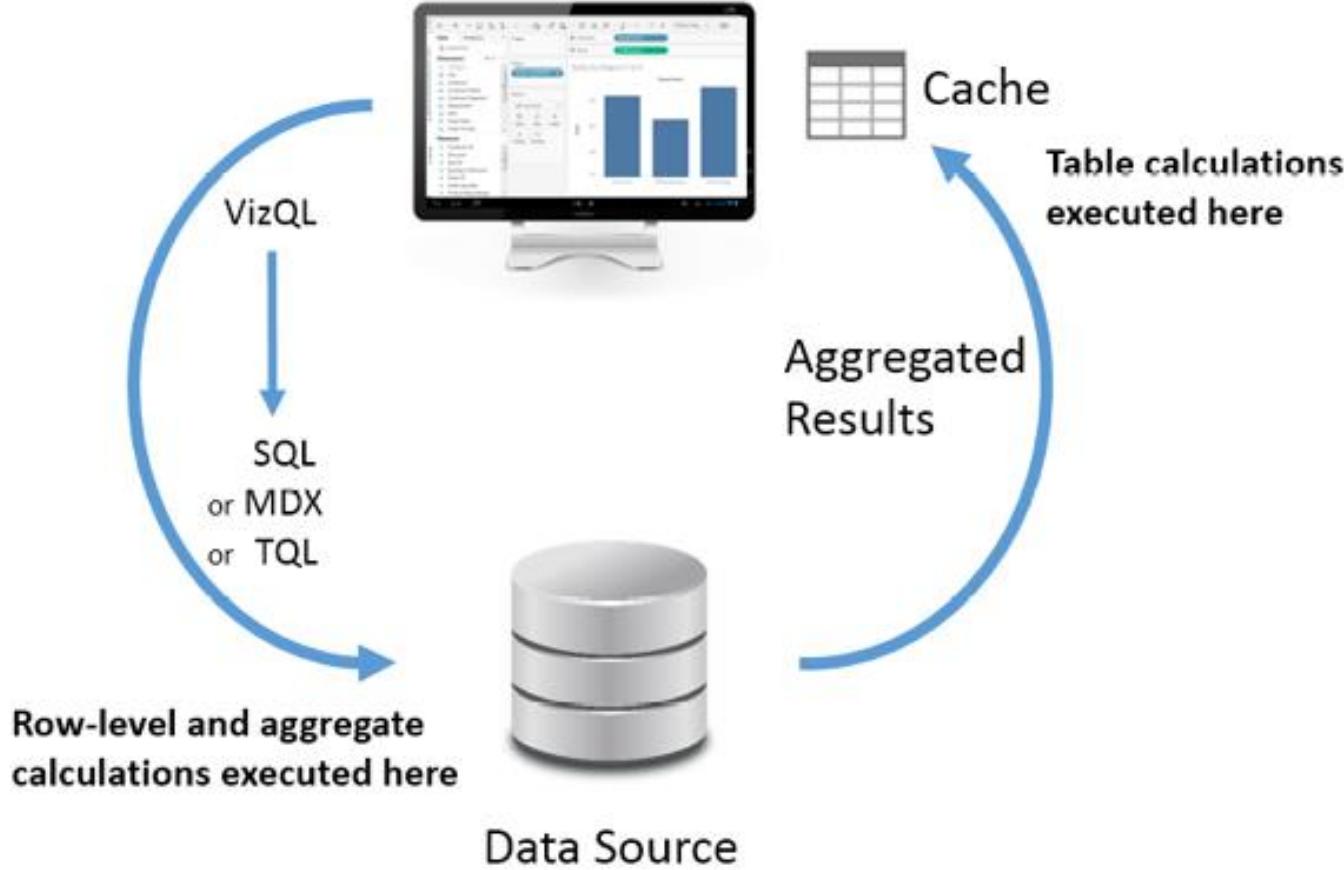
We'll consider these topics:

- An overview of table calculations
- Quick table calculations
- Scope and direction
- Addressing and partitioning
- Custom table calculations
- Practical examples

An overview of table calculations

- Table calculations are different from all other calculations in Tableau. Row-level, aggregate calculations, and LOD expressions, which we explored in the previous lessons, are performed as part of the query to the data source.
- If you were to examine the queries sent to the data source by Tableau, you'd find the code for your calculations translated into whatever implementation of SQL the data source used.

Tableau



Creating and editing table calculations

There are several ways to create table calculations in Tableau, including:

- Using the drop-down menu for any active field used as a numeric aggregate in the view, select Quick Table Calculation and then the desired calculation type.
- Using the drop-down menu for any active field that is used as a numeric aggregate in the view, select Add Table Calculation, then select the calculation type, and adjust any desired settings.
- Creating a calculated field and using one or more table calculation functions to write your own custom table calculations.

An overview of table calculations

Following is a snippet of an active field without a table calculation:



SUM(Sales)

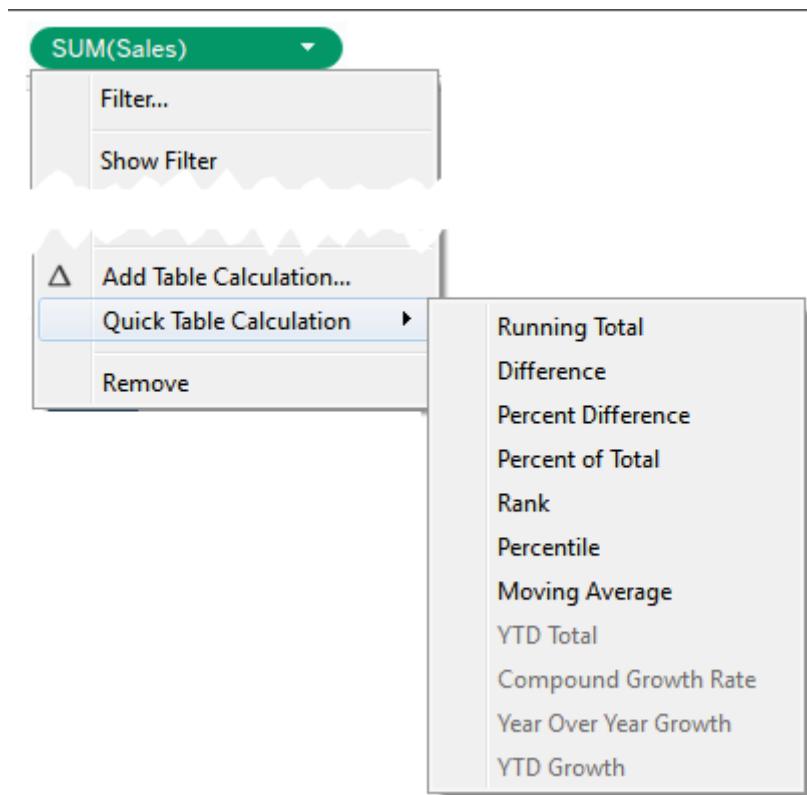
An active field without a table calculation applied

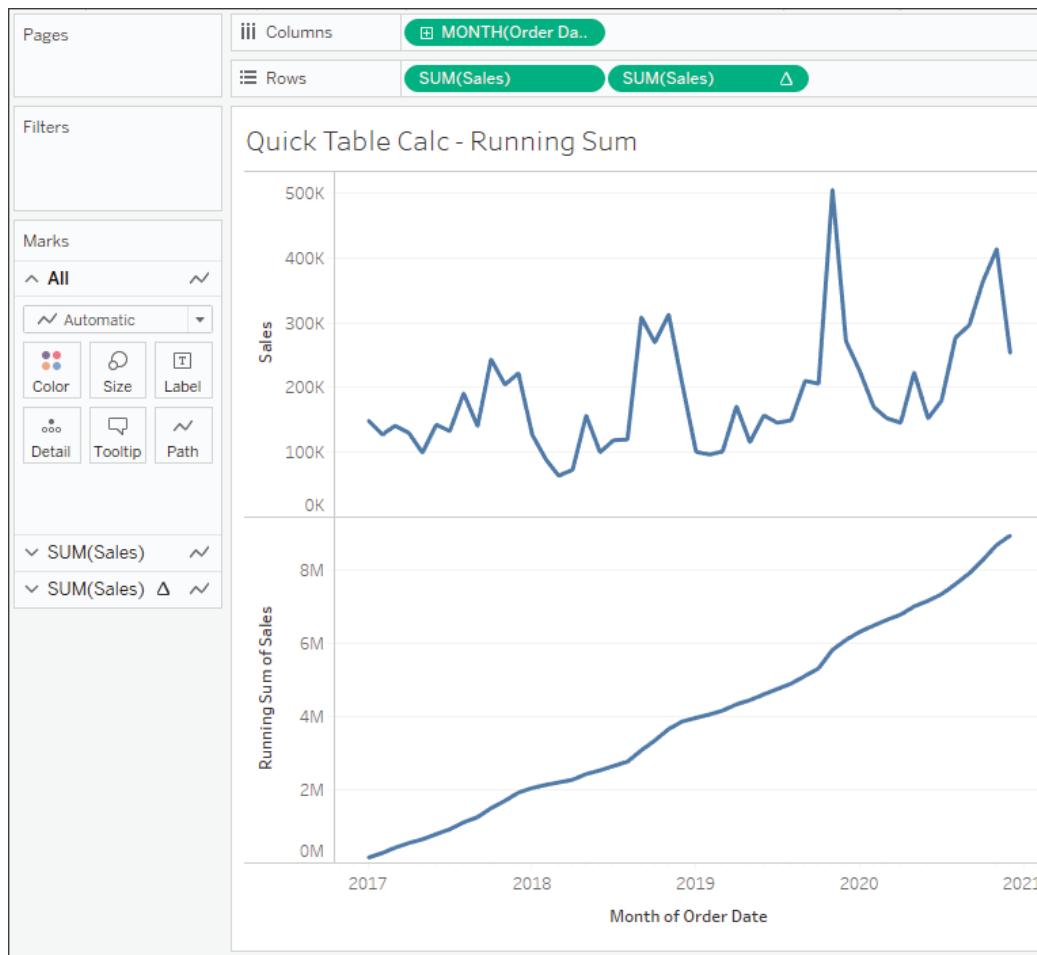
Following is the active field with a table calculation:



SUM(Sales) △

Quick table calculations





Some examples of Quick Table Calcs

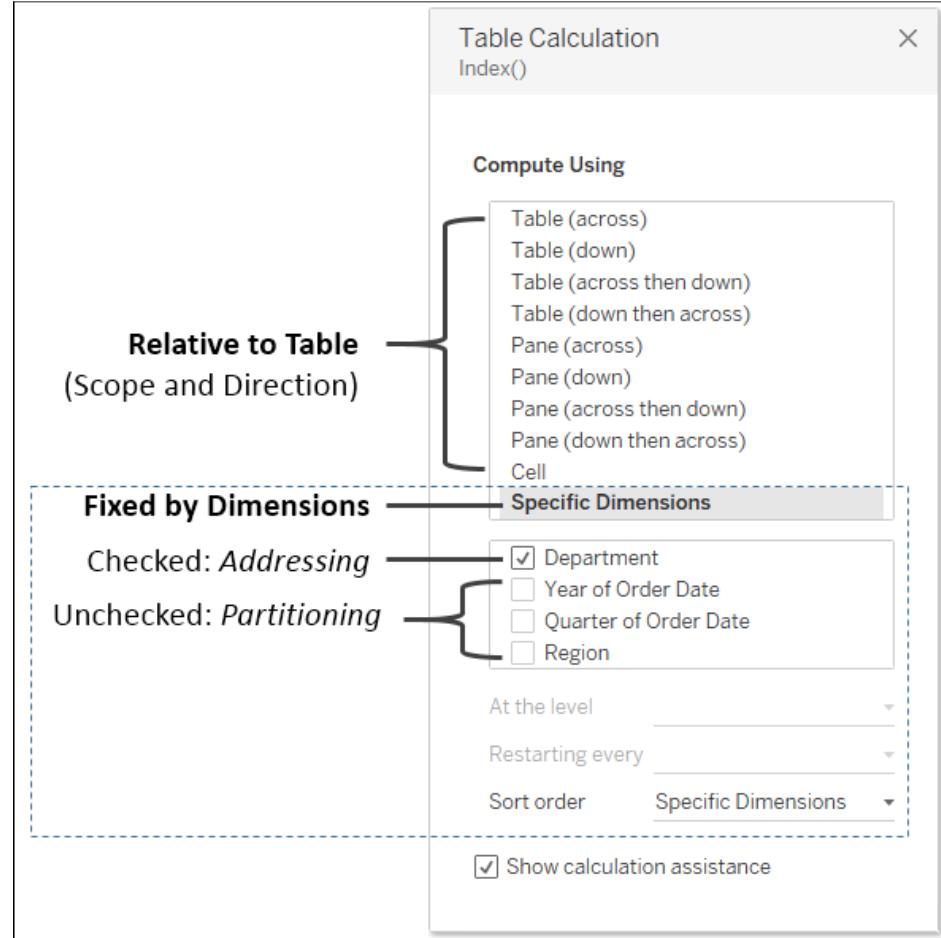
Year of Order Date	Quarter of Order Date	Sales	Running Sum of Sales along Table (Down)	Difference in Sales from the Previous along Table (Down)	Rank of Sales along Table (Down)
2017	Q1	417,555	417,555		12
	Q2	372,289	789,844	-45,266	13
	Q3	464,319	1,254,163	92,030	10
	Q4	670,182	1,924,345	205,863	5
2018	Q1	279,148	2,203,493	-391,034	16
	Q2	330,269	2,533,762	51,121	14
	Q3	546,875	3,080,637	216,606	7
	Q4	788,255	3,868,892	241,380	3
2019	Q1	298,848	4,167,740	-489,407	15
	Q2	443,764	4,611,504	144,916	11
	Q3	505,453	5,116,957	61,689	9
	Q4	982,675	6,099,632	477,222	2
2020	Q1	547,656	6,647,288	-435,019	6
	Q2	521,650	7,168,938	-26,006	8
	Q3	752,933	7,921,871	231,283	4
	Q4	1,030,156	8,952,027	277,223	1

Relative versus fixed

We'll look at the details shortly, but first it is important to understand that table calculations may be computed in one of the two following ways:

- Relative
- Fixed

- You can see these concepts in the user interface.
- The following is the Table Calculation editor that appears when you select Edit Table Calculation from the menu of a table calculation field:



Scope and direction

Scope and direction are terms that describe how a table calculation is computed relative to the table. Specifically, scope and direction refer to the following:

- **Scope:** The scope defines the boundaries within which a given table calculation can reference other values.
- **Direction:** The direction defines how the table calculation moves within the scope.

Scope and direction

The options for scope and direction relative to the table are as follows:

- Scope options: Table, pane, and cells
- Direction options: Down, across, down then across, across then down

Pages

iii Columns Region Department

Rows YEAR(Order Date) QUARTER(Order ...)

Filters

Region

YEAR(Order Date)

Marks

Automatic

Color Size Text

Detail Tooltip

SUM(Sales)

Tables, Panes, Cells

Region / Department

		East			West		
		Furniture	Office Supplies	Technology	Furniture	Office Supplies	Technology
2017	Q1	103,094	21,517	47,770	41,391	6,221	24,258
	Q2	50,254	77,176	35,663	44,256	5,275	30,549
	Q3	33,016	34,014	30,213	32,794	30,200	119,668
	Q4	72,409	124,765	108,417	43,313	35,524	26,782
2018	Q1	55,241	26,176	19,534	20,296	9,793	28,652
	Q2	47,499	29,320	41,988	22,534	19,487	31,268
	Q3	102,001	44,762	81,264	46,923	15,995	74,876
	Q4	63,724	75,442	110,800	41,742	39,343	101,173
2019	Q1	33,938	20,638	23,858	14,296	17,108	13,947
	Q2	47,778	51,273	73,519	37,173	36,998	35,092
	Q3	83,456	59,039	32,302	34,504	16,305	76,077
	Q4	124,363	65,704	86,405	171,909	51,372	62,808
2020	Q1	51,050	74,600	83,149	53,471	24,272	33,619
	Q2	57,299	83,749	52,066	22,994	19,818	42,868
	Q3	92,243	51,459	70,612	56,791	62,900	66,427
	Q4	132,884	98,473	157,206	42,129	45,894	71,268

Table

Pane

Cell

Working with scope and direction

- In order to see how scope and direction work together, let's work through a few examples.
- We'll start by creating our own custom table calculations. Create a new calculated field named Index with the code `Index()`.
- `Index()` is a table calculation function that starts with a value of 1 and increments by one as it moves in each direction and within a given scope.

Table (across):



		East			West		
		Furniture	Office Supplies	Technolo..	Furniture	Office Supplies	Technolo..
2015	Q1	1	2	3	4	5	6
	Q2	1	2	3	4	5	6
	Q3	1	2	3	4	5	6
	Q4	1	2	3	4	5	6
2016	Q1	1	2	3	4	5	6
	Q2	1	2	3	4	5	6
	Q3	1	2	3	4	5	6
	Q4	1	2	3	4	5	6

Table (down):

		East			West		
		Office			Office		
		Furniture	Supplies	Technolo..	Furniture	Supplies	Technolo..
2015	Q1	1	1	1	1	1	1
	Q2	2	2	2	2	2	2
	Q3	3	3	3	3	3	3
	Q4	4	4	4	4	4	4
2016	Q1	5	5	5	5	5	5
	Q2	6	6	6	6	6	6
	Q3	7	7	7	7	7	7
	Q4	8	8	8	8	8	8



Table (across then down):



		East			West		
		Furniture	Office Supplies	Technolo..	Furniture	Office Supplies	Technolo..
2015	Q1	1	2	3	4	5	6
	Q2	7	8	9	10	11	12
	Q3	13	14	15	16	17	18
	Q4	19	20	21	22	23	24
2016	Q1	25	26	27	28	29	30
	Q2	31	32	33	34	35	36
	Q3	37	38	39	40	41	42
	Q4	43	44	45	46	47	48

Pane (across):

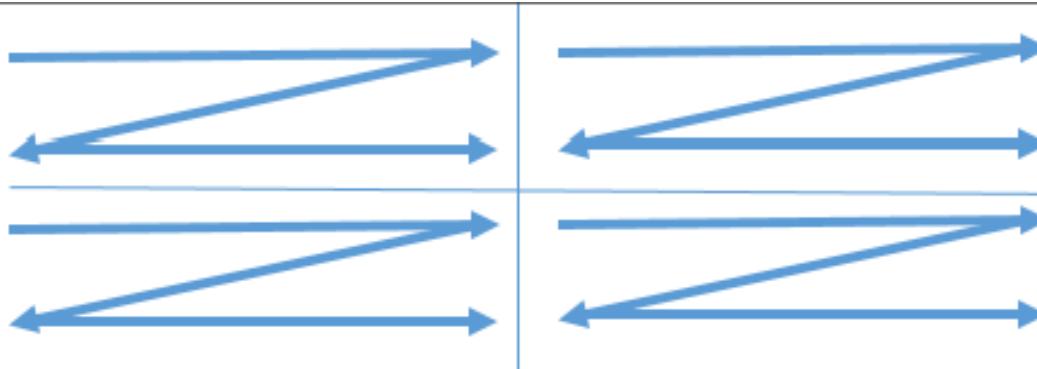


		East			West		
		Furniture	Supplies	Technolo..	Furniture	Supplies	Technolo..
2015	Q1	1	2	3	1	2	3
	Q2	1	2	3	1	2	3
	Q3	1	2	3	1	2	3
	Q4	1	2	3	1	2	3
2016	Q1	1	2	3	1	2	3
	Q2	1	2	3	1	2	3
	Q3	1	2	3	1	2	3
	Q4	1	2	3	1	2	3

Pane (down):

		East			West			
		Furniture	Office Supplies	Technolo..	Furniture	Office Supplies	Technolo..	
2015	Q1	1	1	1	1	1	1	1
	Q2	2	2	2	2	2	2	2
	Q3	3	3	3	3	3	3	3
	Q4	4	4	4	4	4	4	4
2016	Q1	1	1	1	1	1	1	1
	Q2	2	2	2	2	2	2	2
	Q3	3	3	3	3	3	3	3
	Q4	4	4	4	4	4	4	4





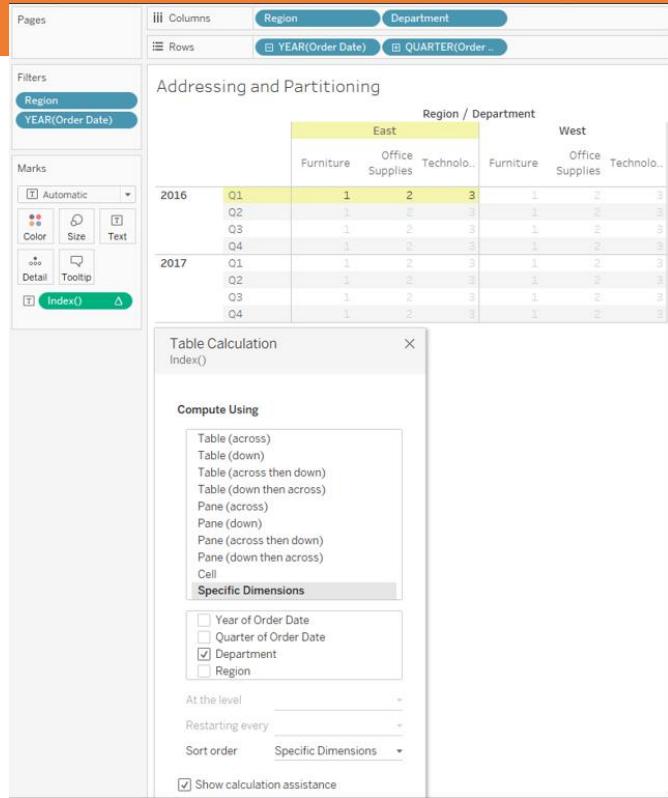
		East			West		
		Office			Office		
		Furniture	Supplies	Technolo..	Furniture	Supplies	Technolo..
2015	Q1	1	2	3	1	2	3
	Q2	4	5	6	4	5	6
	Q3	7	8	9	7	8	9
	Q4	10	11	12	10	11	12
2016	Q1	1	2	3	1	2	3
	Q2	4	5	6	4	5	6
	Q3	7	8	9	7	8	9
	Q4	10	11	12	10	11	12

Addressing and partitioning

- Addressing and partitioning are very similar to scope and direction but are most often used to describe how table calculations are computed with absolute reference to certain fields in the view.
- With addressing and partitioning, you define which dimensions in the view define the addressing (direction) and all others define the partitioning (scope).

Addressing and partitioning

- The result of selecting Department is as follows:



Pages YEAR(Order Date) QUARTER(Order ...)

Rows Region Department

Filters
 Region
 YEAR(Order Date)

Marks
 Automatic
Color Size Text
Detail Tooltip
 Index() △

Addressing and Partitioning

		Order Date				Order Date			
		2016				2017			
		Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
East	Furniture	1	1	1	1	1	1	1	1
	Office Supplies	2	2	2	2	2	2	2	2
	Technology	3	3	3	3	3	3	3	3
West	Furniture	1	1	1	1	1	1	1	1
	Office Supplies	2	2	2	2	2	2	2	2
	Technology	3	3	3	3	3	3	3	3

Table Calculation Index()

Compute Using

- Table (across)
- Table (down)
- Table (across then down)
- Table (down then across)
- Pane (across)
- Pane (down)
- Pane (across then down)
- Pane (down then across)
- Cell
- Specific Dimensions**

Year of Order Date
 Quarter of Order Date
 Department
 Region

At the level _____

Restarting every _____

Sort order Specific Dimensions

Show calculation assistance

Working with addressing and partitioning

The screenshot shows the Tableau Data Prep interface. On the left, there are sections for Pages, Filters, and Marks. The Marks section includes a dropdown for 'Automatic' and buttons for Color, Size, Text, Detail, Tooltip, and Index().

The main area displays a table with the following schema:

	Year	Quarter	Region / Department			Region / Department			
			East	Office Supplies	Technology	West	Office Supplies	Technology	
2016	Q1	Furniture	1	5	9	Furniture	1	5	9
		Office Supplies	2	6	10	Office Supplies	2	6	10
		Technology	3	7	11	Technology	3	7	11
			4	8	12		4	8	12
2017	Q1	Furniture	1	5	9	Furniture	1	5	9
		Office Supplies	2	6	10	Office Supplies	2	6	10
		Technology	3	7	11	Technology	3	7	11
			4	8	12		4	8	12

A 'Table Calculation' window is open, titled 'Index()'. It contains a 'Compute Using' section with options like 'Table (across)', 'Table (down)', etc., and a 'Specific Dimensions' section with checked boxes for 'Department' and 'Quarter of Order Date'. It also includes settings for 'At the level' (set to 'Deepest'), 'Restarting every' (set to 'None'), 'Sort order' (set to 'Specific Dimensions'), and a checkbox for 'Show calculation assistance'.

Pages Columns Region Department

Rows YEAR(Order Date) QUARTER(Order ..)

Addressing and Partitioning

		Region / Department			West		
		East	Office Supplies	Technology	Furniture	Office Supplies	Technology
2016	Q1	1	3	5	1	3	1
	Q2	1	3	5	1	3	1
	Q3	1	3	5	1	3	1
	Q4	1	3	5	1	3	1
2017	Q1	2	4	6	2	4	2
	Q2	2	4	6	2	4	2
	Q3	2	4	6	2	4	2
	Q4	2	4	6	2	4	2

Table Calculation X

Index()

Compute Using

- Table (across)
- Table (down)
- Table (across then down)
- Table (down then across)
- Pane (across)
- Pane (down)
- Pane (across then down)
- Pane (down then across)
- Cell
- Specific Dimensions**

Department
 Quarter of Order Date
 Region
 Year of Order Date

At the level **Deepest** ▼

Restarting every **None** ▼

Sort order **Specific Dimensions** ▼

Show calculation assistance

Custom table calculations

- Before we move on to some practical examples, let's briefly discuss how to write your own table calculations, instead of using quick table calculations.
- You can see a list of available table calculation functions by creating a new calculation and selecting Table Calculation from the drop-down list under Functions.

Meta table functions

- These are the functions that give you information about the partitioning and addressing.
- These functions also include Index, First, Last, and Size:

Meta table functions

Meta Table Calculations		Index along Category	First along Category	Last along Category	Size along Category
Department	Category				
Furniture	Bookcases	1	0	3	4
	Chairs & Chairmats	2	-1	2	4
	Office Furnishings	3	-2	1	4
	Tables	4	-3	0	4
Office Supplies	Appliances	1	0	8	9
	Binders and Binder Accessories	2	-1	7	9
	Envelopes	3	-2	6	9
	Labels	4	-3	5	9
	Paper	5	-4	4	9
	Pens & Art Supplies	6	-5	3	9
	Rubber Bands	7	-6	2	9
	Scissors, Rulers and Trimmers	8	-7	1	9
	Storage & Organization	9	-8	0	9
Technology	Computer Peripherals	1	0	3	4
	Copiers and Fax	2	-1	2	4
	Office Machines	3	-2	1	4
	Telephones and Communication	4	-3	0	4

Lookup and previous value

Lookup and Previous Value			
Department	Category	Lookup	Previous Value
Furniture	Bookcases	Null	,Bookcases
	Chairs & Chairmats	Bookcases	,Bookcases,Chairs & Chairmats
	Office Furnishings	Chairs & Chairmats	,Bookcases,Chairs & Chairmats,Office Furnishings
	Tables	Office Furnishings	,Bookcases,Chairs & Chairmats,Office Furnishings,Tables
Office Supplies	Appliances	Null	,Appliances
	Binders and Binder Accessor..	Appliances	,Appliances,Binders and Binder Accessories
	Envelopes	Binders and Binder Acc..	,Appliances,Binders and Binder Accessories,Envelopes
	Labels	Envelopes	,Appliances,Binders and Binder Accessories,Envelopes,Labels
Technology	Paper	Labels	,Appliances,Binders and Binder Accessories,Envelopes,Labels,Paper
	Pens & Art Supplies	Paper	,Appliances,Binders and Binder Accessories,Envelopes,Labels,Paper,Pens & Art Supplies
	Rubber Bands	Pens & Art Supplies	,Appliances,Binders and Binder Accessories,Envelopes,Labels,Paper,Pens & Art Supplies,Rubber Bands
	Scissors, Rulers and Trimmers	Rubber Bands	,Appliances,Binders and Binder Accessories,Envelopes,Labels,Paper,Pens & Art Supplies,Rubber Bands,Scissors,
Technology	Storage & Organization	Scissors, Rulers and ..	,Appliances,Binders and Binder Accessories,Envelopes,Labels,Paper,Pens & Art Supplies,Rubber Bands,Scissors,
	Computer Peripherals	Null	,Computer Peripherals
	Copiers and Fax	Computer Peripherals	,Computer Peripherals,Copiers and Fax
	Office Machines	Copiers and Fax	,Computer Peripherals,Copiers and Fax,Office Machines
Technology	Telephones and Communication	Office Machines	,Computer Peripherals,Copiers and Fax,Office Machines,Telephones and Communication

Lookup and previous value

- The first row in each partition gets a NULL result from the lookup (because there isn't a row before it).
- For Previous_Value, we used this code:

Previous_Value("") + "," + ATTR([Category])

Running Functions

Department	Category	Sales	Running Sum of Sales along Category	Running Min of Sales along Category
Furniture	Bookcases	507,496	507,496	507,496
	Chairs & Chairmats	1,164,586	1,672,082	507,496
	Office Furnishings	444,634	2,116,716	444,634
	Tables	1,061,922	3,178,638	444,634
Office Supplies	Appliances	456,736	456,736	456,736
	Binders and Binder ..	638,583	1,095,319	456,736
	Envelopes	147,915	1,243,234	147,915
	Labels	23,446	1,266,680	23,446
	Paper	253,620	1,520,300	23,446
	Pens & Art Supplies	103,265	1,623,565	23,446
	Rubber Bands	8,670	1,632,235	8,670
	Scissors, Rulers and ..	40,432	1,672,667	8,670
	Storage & Organizat..	585,717	2,258,384	8,670
Technology	Computer Periphera..	490,851	490,851	490,851
	Copiers and Fax	661,215	1,152,066	490,851
	Office Machines	1,218,655	2,370,721	490,851
	Telephones and Com..	1,144,284	3,515,005	490,851

Window functions

Window Functions

Department	Category	Sales	Window Sum along Category	Window Max along Category
Furniture	Bookcases	507,496	3,178,638	1,164,586
	Chairs & Chairmats	1,164,586	3,178,638	1,164,586
	Office Furnishings	444,634	3,178,638	1,164,586
	Tables	1,061,922	3,178,638	1,164,586
Office Supplies	Appliances	456,736	2,258,384	638,583
	Binders and Binder ..	638,583	2,258,384	638,583
	Envelopes	147,915	2,258,384	638,583
	Labels	23,446	2,258,384	638,583
	Paper	253,620	2,258,384	638,583
	Pens & Art Supplies	103,265	2,258,384	638,583
	Rubber Bands	8,670	2,258,384	638,583
	Scissors, Rulers and ..	40,432	2,258,384	638,583
	Storage & Organizat..	585,717	2,258,384	638,583
Technology	Computer Periphera..	490,851	3,515,005	1,218,655
	Copiers and Fax	661,215	3,515,005	1,218,655
	Office Machines	1,218,655	3,515,005	1,218,655
	Telephones and Com..	1,144,284	3,515,005	1,218,655

Rank functions

Rank Functions

Department	Category	Sales	Rank along Category
Furniture	Bookcases	507,496	3
	Chairs & Chairmats	1,164,586	1
	Office Furnishings	444,634	4
	Tables	1,061,922	2
Office Supplies	Appliances	456,736	3
	Binders and Binder ..	638,583	1
	Envelopes	147,915	5
	Labels	23,446	8
	Paper	253,620	4
	Pens & Art Supplies	103,265	6
	Rubber Bands	8,670	9
	Scissors, Rulers and ..	40,432	7
	Storage & Organizat..	585,717	2
Technology	Computer Periphera..	490,851	4
	Copiers and Fax	661,215	3
	Office Machines	1,218,655	1
	Telephones and Com..	1,144,284	2

Script functions

- These functions allow integration with the R analytics platform or Python, either of which can incorporate simple or complex scripts for everything from advanced statistics to predictive modeling.
- It's beyond the scope of this course to dive into all that is possible, but documentation and examples are readily available on Tableau's website and from various members of the Tableau community.

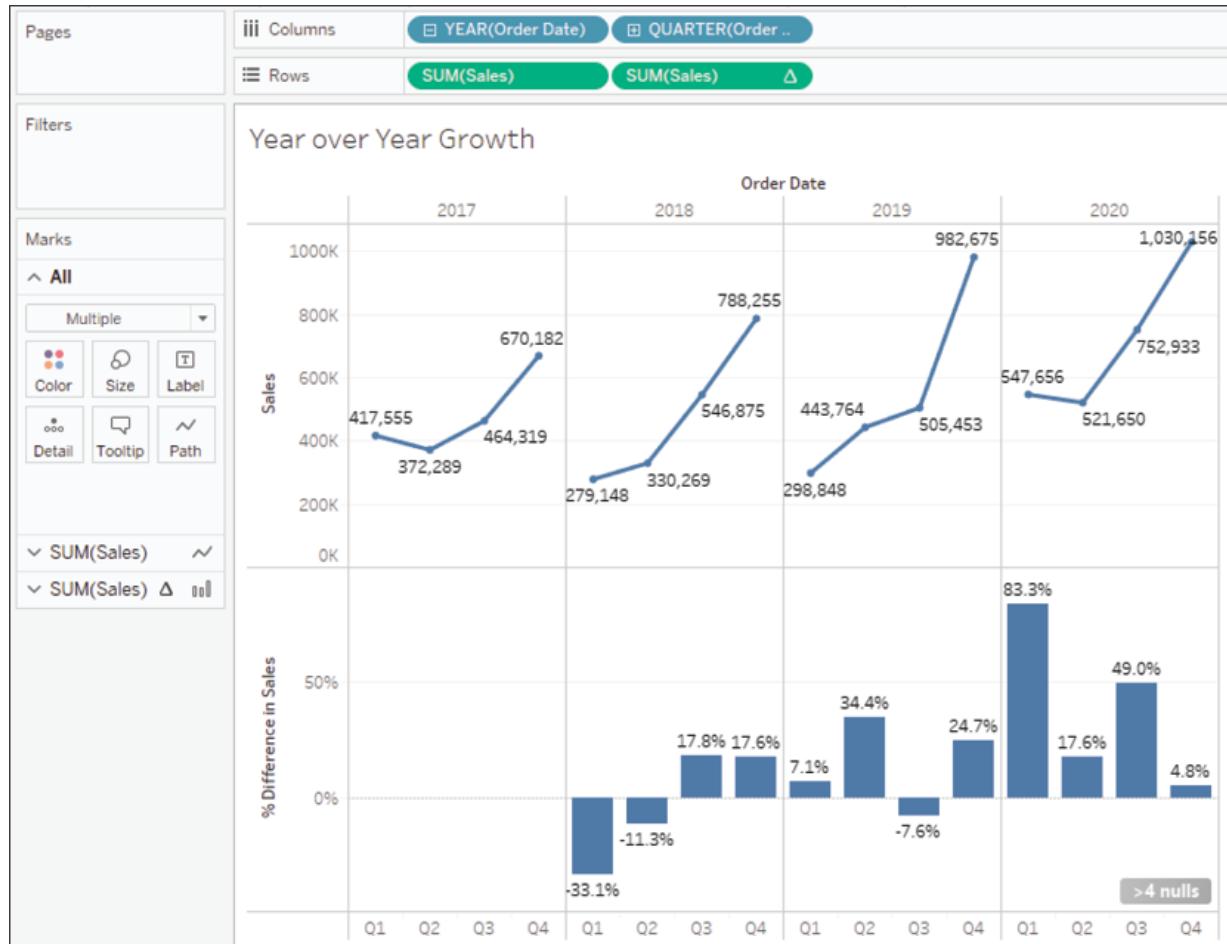
The Total function

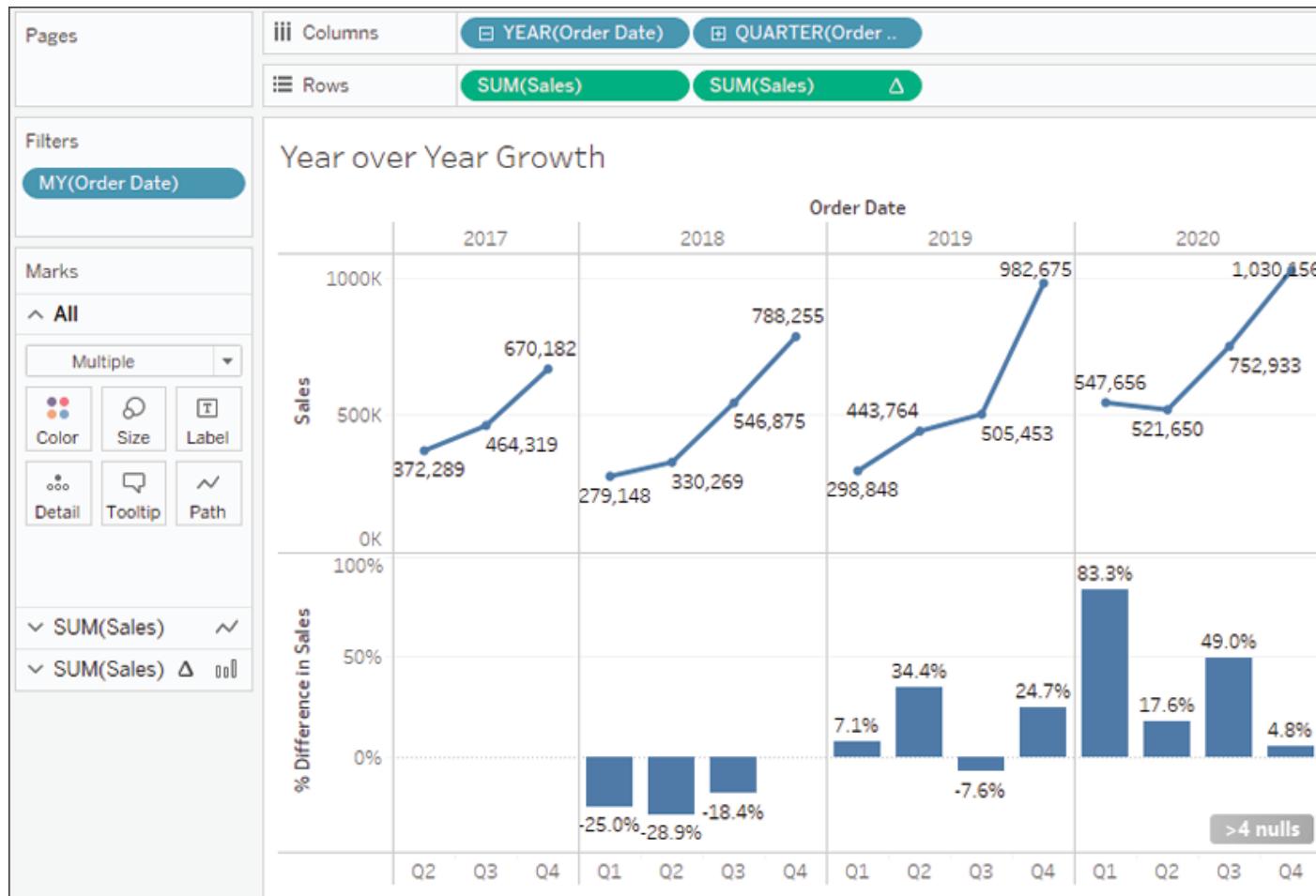
- The Total function deserves its own category because it functions a little differently from the others.
- Unlike the other functions that work on the aggregate table in the cache, Total will re-query the underlying source for all the source data rows that make up a given partition.
- In most cases, this will yield the same result as a window function.

Practical examples

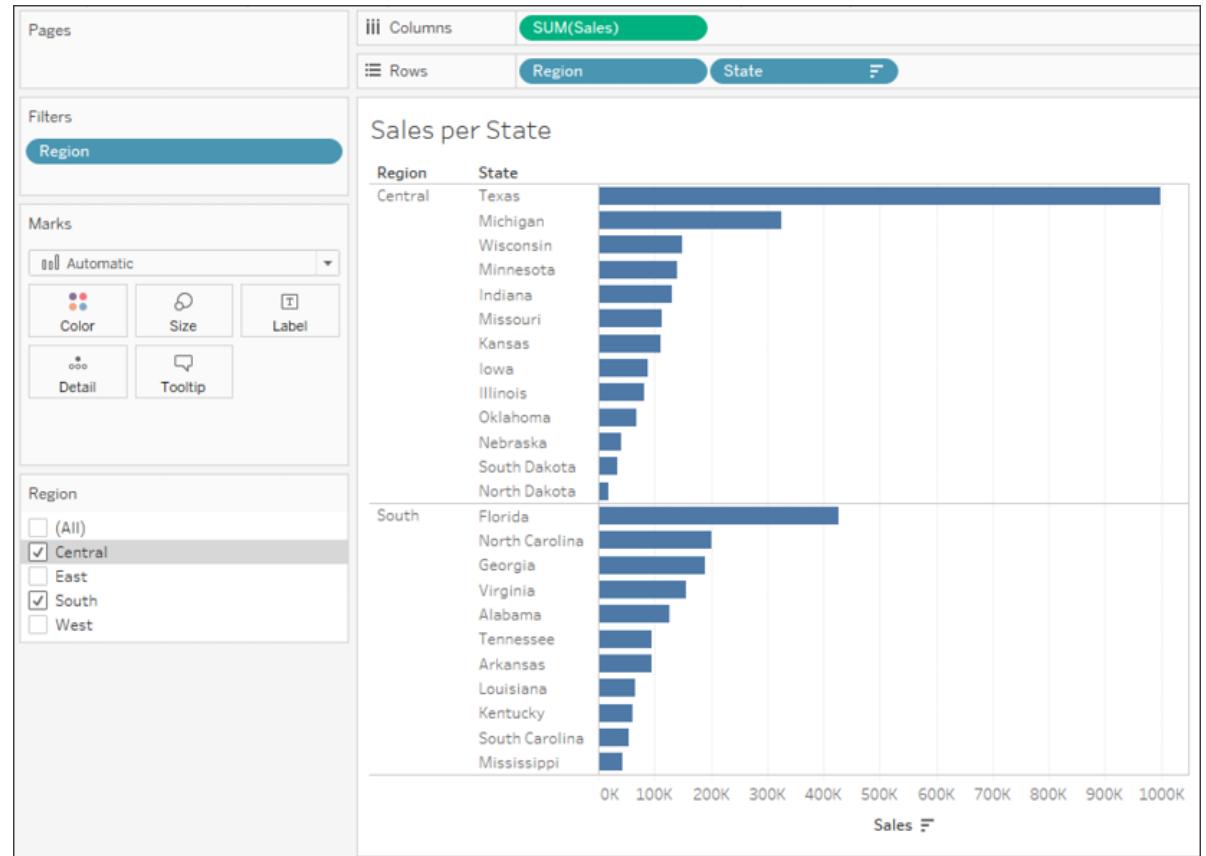
- Having looked at some of the essential concepts of table calculations, let's consider some practical examples.
- We'll look at several examples, although the practical use of table calculations is nearly endless.

Year over year growth





Dynamic titles with totals



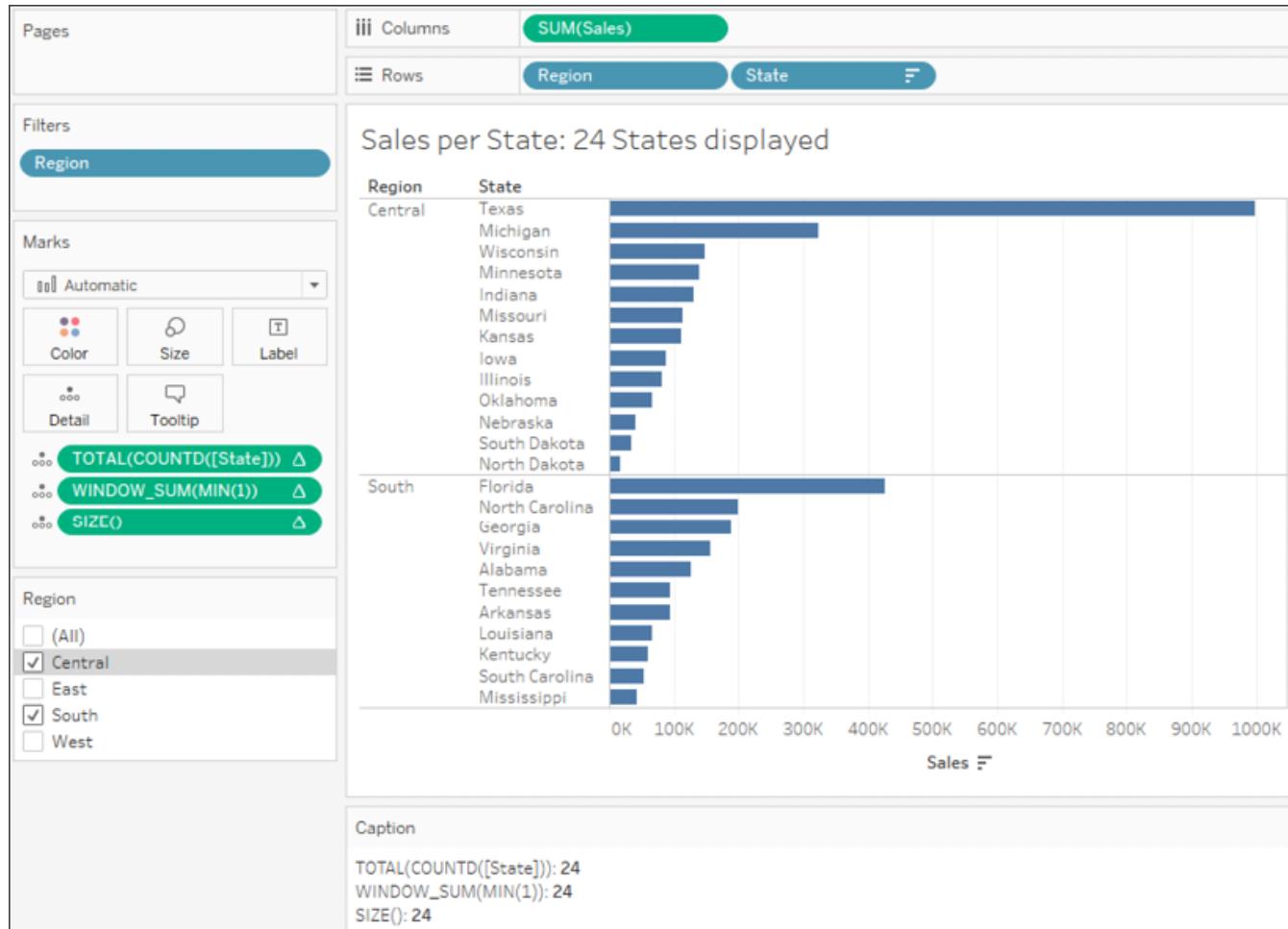
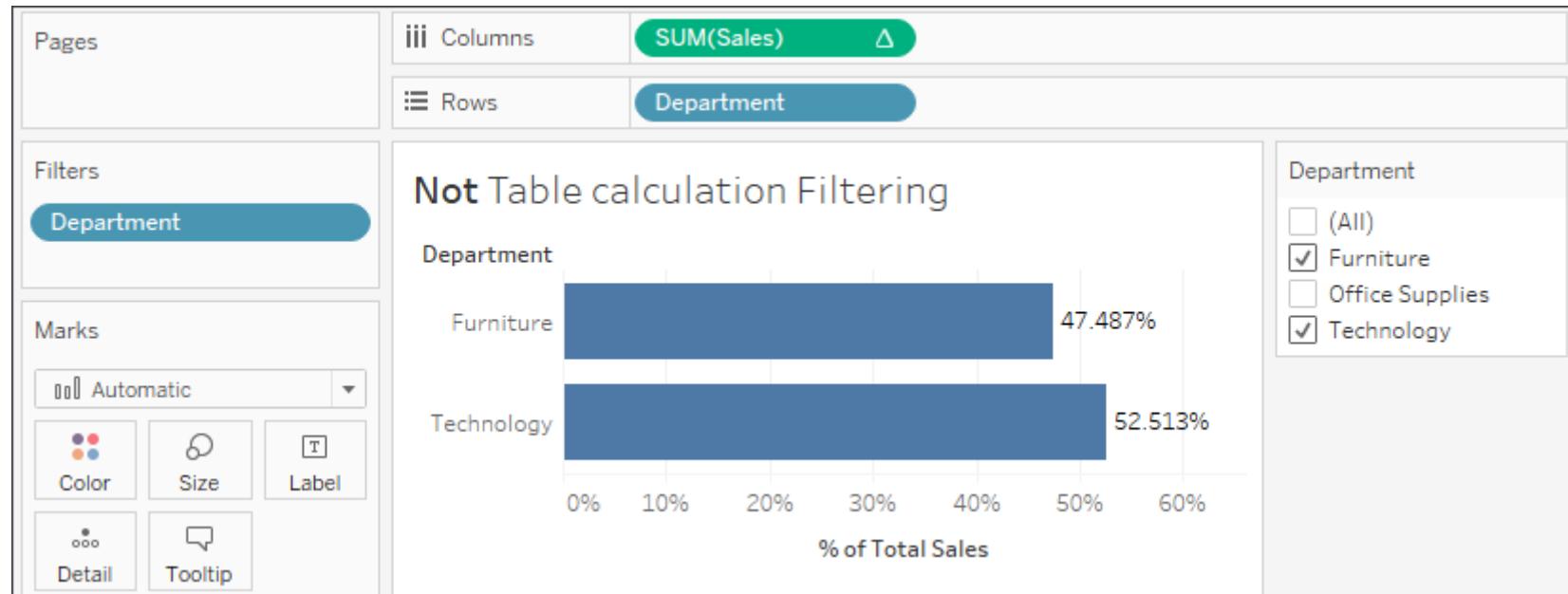
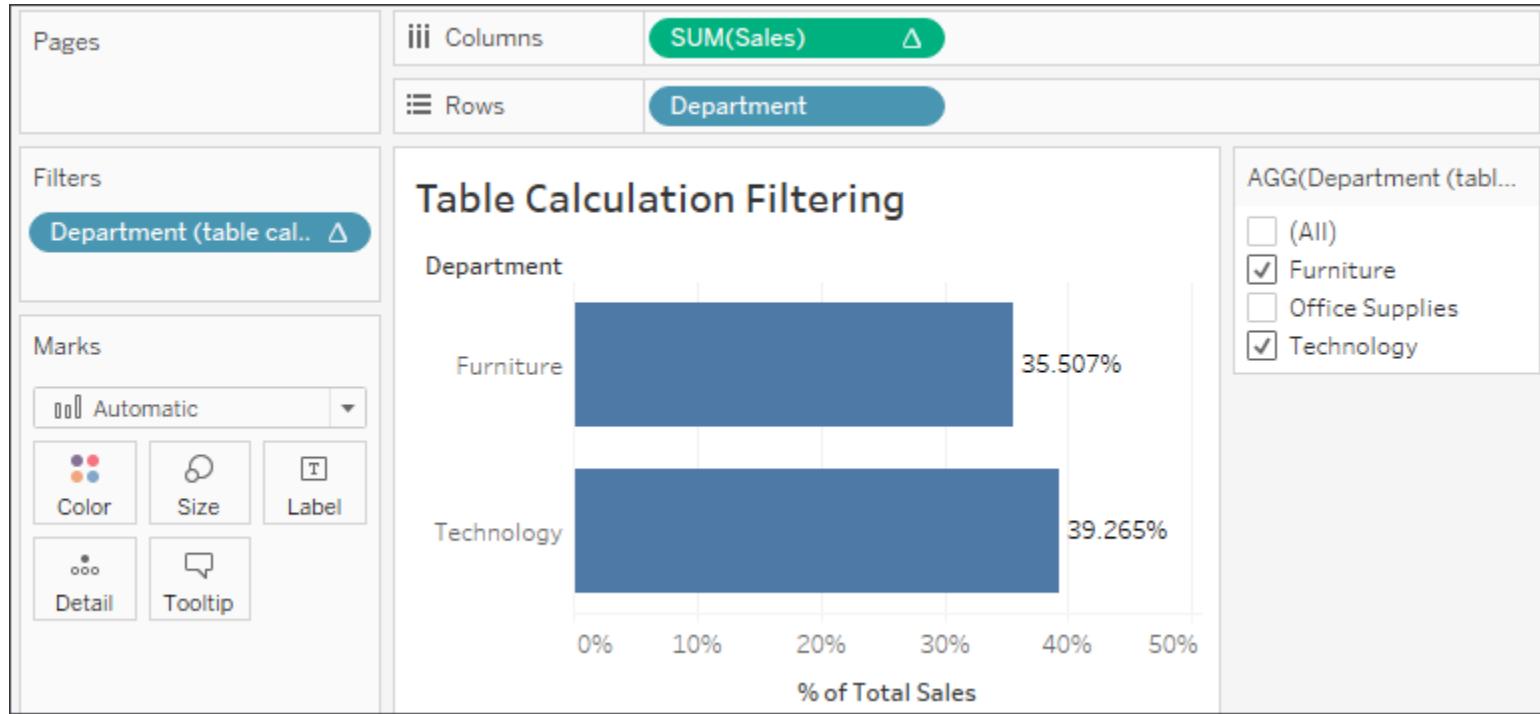


Table calculation filtering (late filtering)



Year over year growth



Summary

- We've covered a lot of concepts surrounding table calculations in this lesson.
- You now have a foundation for using the simplicity of quick table calculations and leveraging the power of advanced table calculations.
- We've looked at the concepts of scope and direction as they apply to table calculations that operate relative to the row and column layout of the view.

COMPLETE LAB 6

7. Making Visualizations That Look Great and Work Well

A blurred background image of a person's hands typing on a laptop keyboard, suggesting a workspace or data analysis environment.

Making Visualizations

This lesson will cover the following topics:

- Visualization considerations
- Leveraging formatting in Tableau
- Adding value to visualizations

Visualization considerations

- Tableau employs good practices for formatting and visualization from the time you start dropping fields on shelves.
- You'll find that the discrete palettes use colors that are easy to distinguish, the fonts are pleasant, the grid lines are faint where appropriate, and numbers and dates follow the default format settings defined in the metadata.
- The default formatting is more than adequate for discovery and analysis.

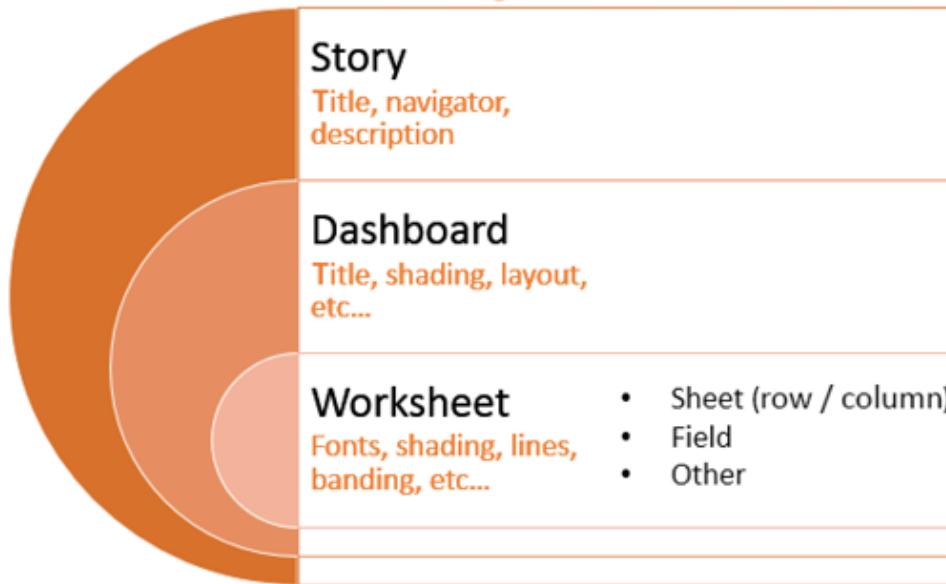
Data Source

Default metadata field
formats, colors, shapes



Workbook

Global Formatting



workbook-level formatting

Format Workbook X

Fonts

All
Tableau Book

Worksheets
Tableau Book 9 ▾

Tooltips
Tableau Book 10 ▾

Worksheet Titles
Tableau Light 15 ▾

Dashboard Titles
Tableau Book 18 ▾

Story Titles
Tableau Regular 18 ▾

Less ^

Lines

Grid Lines

Zero Lines

Trend Lines

Reference Lines

Drop Lines

Axis Rulers

Axis Ticks

Reset to Defaults

Worksheet-level formatting

Formatting: Parts of the View

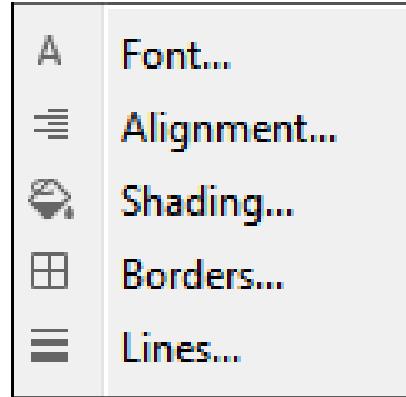
Diagram illustrating the parts of a worksheet view:

- 1**: Department
- 2**: Customer Segment
- 3**: Category
- 4**: Consumer
- 5**: Corporate
- 6**: Home Office
- 7**: Small Business
- a**: Grand Total

Department	Category	Customer Segment				Grand Total
		Consumer	Corporate	Home Office	Small Business	
Furniture	Bookcases	32,626	262,085	79,404	73,381	507,496
	Chairs & Chairmats	305,381	407,724	212,830	238,651	1,164,586
	Office Furnishings	69,528	115,506	197,188	62,412	444,634
	Tables	228,934	363,979	287,507	181,502	1,061,922
	(a) Total	696,469	1,149,294	776,929	555,946	3,178,638
Office Supplies	Appliances	63,813	167,941	124,757	100,225	456,736
	Binders and Binder Accessor..	103,625	225,160	148,472	161,326	638,583
	Envelopes	37,643	44,462	22,577	43,233	147,915
	Labels	3,713	7,929	5,411	6,393	23,446
	Paper	53,004	89,312	61,123	50,181	253,620
	Pens & Art Supplies	24,027	36,004	21,765	21,469	103,265
	Rubber Bands	1,710	2,197	2,294	2,469	8,670
	Scissors, Rulers and Trimme..	14,628	9,625	12,947	3,232	40,432
	Storage & Organization	121,719	154,918	179,151	129,929	585,717
Technology	Total	423,882	737,548	578,497	518,457	2,258,384
	Computer Peripherals	80,805	224,142	110,840	75,064	490,851
	Copiers and Fax	148,504	205,639	174,718	132,354	661,215
	Office Machines	260,011	516,513	245,019	197,112	1,218,655
	Telephones and Communicat..	225,571	436,295	282,962	199,456	1,144,284
(b)	Total	714,891	1,382,589	813,539	603,986	3,515,005
	Grand Total	1,835,242	3,269,431	2,168,965	1,678,389	8,952,027

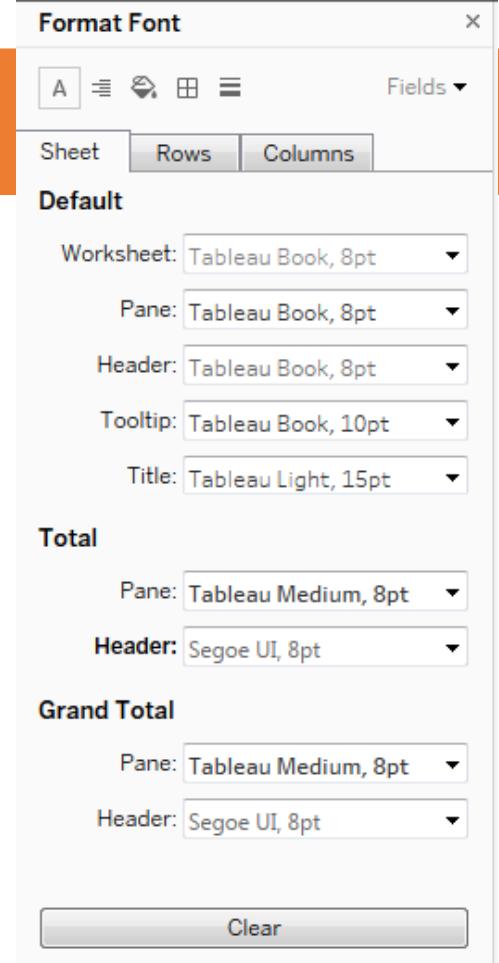
Worksheet-level formatting

- To view the format window, select Format from the menu and then Font..., Alignment..., Shading..., Borders..., or Lines....:



Worksheet-level formatting

- You should now see the format window on the left, in this case, Format Font. It will look like this:



Field-level formatting

- For example, is the window as it appears for the SUM(Sales) field:

Format SUM(Sales)

A ⌂ ⌂ Fields ▾

Axis Pane

Default

Font: Tableau Book, 8pt

Alignment: Automatic

Numbers: 123,456

Totals

Font: Tableau Medium, 8pt

Alignment: Automatic

Numbers: 123,456

Grand Totals

Font: Tableau Medium, 8pt

Alignment: Automatic

Numbers: 123,456

Special Values (eg. NULL)

Text: (Blank)

Marks: Show at Indicator

Custom number formatting

- When you alter the format of a number, you can select from several standard formats, as well as a custom format.
- The custom format allows you to enter a format string that Tableau will use to format the number.

Custom number formatting

- Some examples, assuming the positive number is 34,331.336 and the negative number is -8,156.7777:

Format String	Resulting Value
#;-#	34331 and -8157
,###.##; (#,###.##)	34,331.34 and (8,156.78)
,###.00000;-#,###.00000	34,331.33600 and -8,156.77770
"up "#,###;"down "#,###;"same"	up 34,331 and down 8,157
,###"▲"; #,###"▼"	34,331▲ and 8,157▼

Custom number formatting

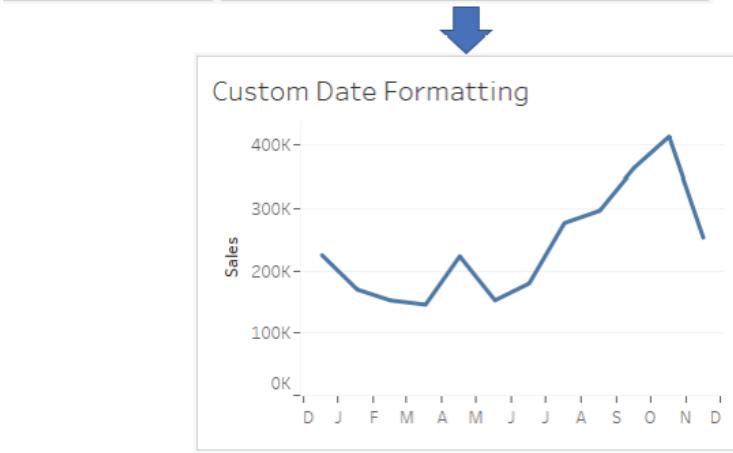
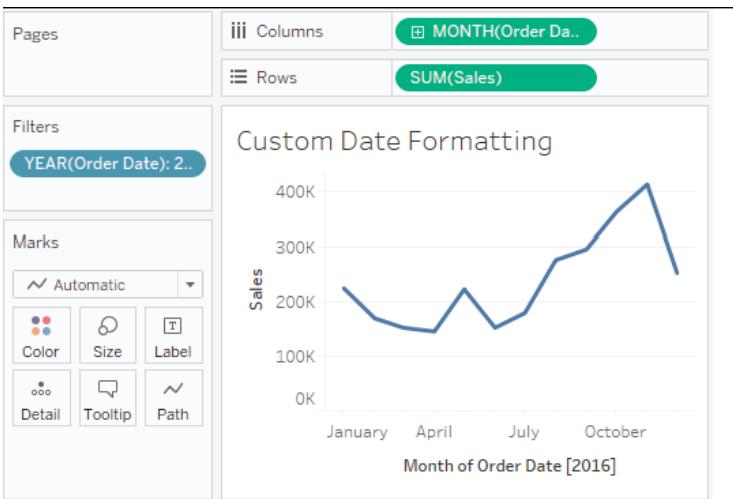
- You can replicate these examples and experiment with other format strings using the Custom Number Formatting view in the Starter or Complete workbooks:

The screenshot shows the Tableau 'Format' dialog box open over a data view. The dialog box is titled 'Format -8156.7777'. It has tabs for 'Axis' and 'Pane', with 'Axis' selected. Under the 'Default' tab, the 'Numbers' dropdown is set to '123,456 ▲'. A blue arrow points from this dropdown to a 'Custom' section where the format string '#,###"▲";#,###"▼"' is displayed. Below this, a list of format types is shown: Automatic, Number (Standard), Number (Custom), Currency (Standard), Currency (Custom), Scientific, Percentage, and Custom (which is selected). The main data view shows a 'Rows' shelf with the value '34331.336' and a 'Columns' shelf. The bottom right corner of the data view shows a green button with the value '-8156.7777'. On the left side of the dialog box, there are sections for 'Pages', 'Filters', and 'Marks' (with options for Color, Size, Text, Detail, and Tooltip).

Custom date formatting

Format String	Resulting Value
m/d/yyyy	11/8/2018
dd/mm/yyyy	08/11/2018
"The date is" m/d/yyyy	The date is 11/8/2018
mmmm d, yyyy	Nov 8, 2018
mmmm dd yyyy	November 08 2018
mm/dd/yyyy h:mm AM/PM	11/08/2018 1:30PM
tttt	1:30:28 PM
dddd, mmmm d, HH:MM:ss	Thursday, November 8, 13:30:28
ddd	Thu

- Notice how applying some custom date formatting improves the readability of the axis for a small timeline in this example:



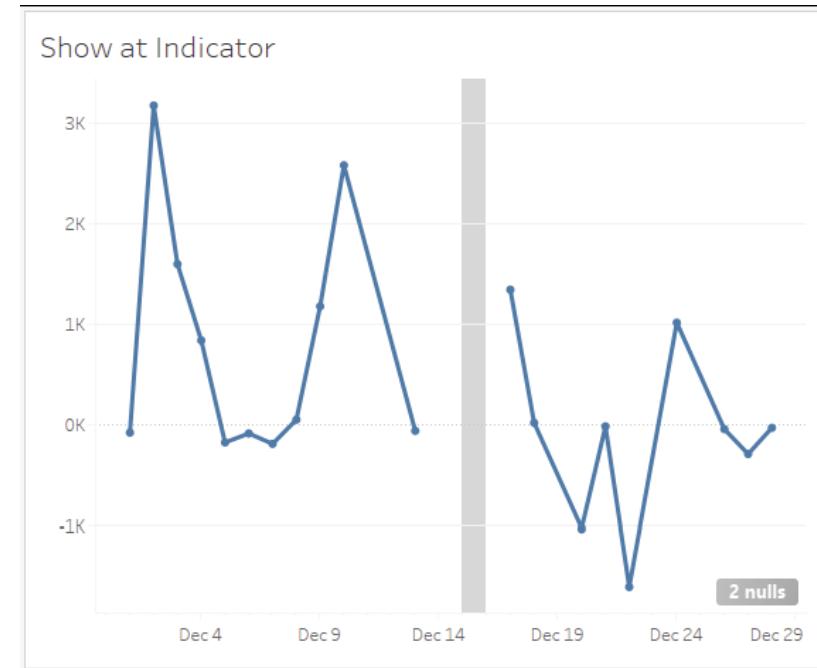
Null formatting

- An additional aspect of formatting a field is specially formatting Null values.
- When formatting a field, select the Pane tab and locate the Special Values section, as shown in the following screenshot:



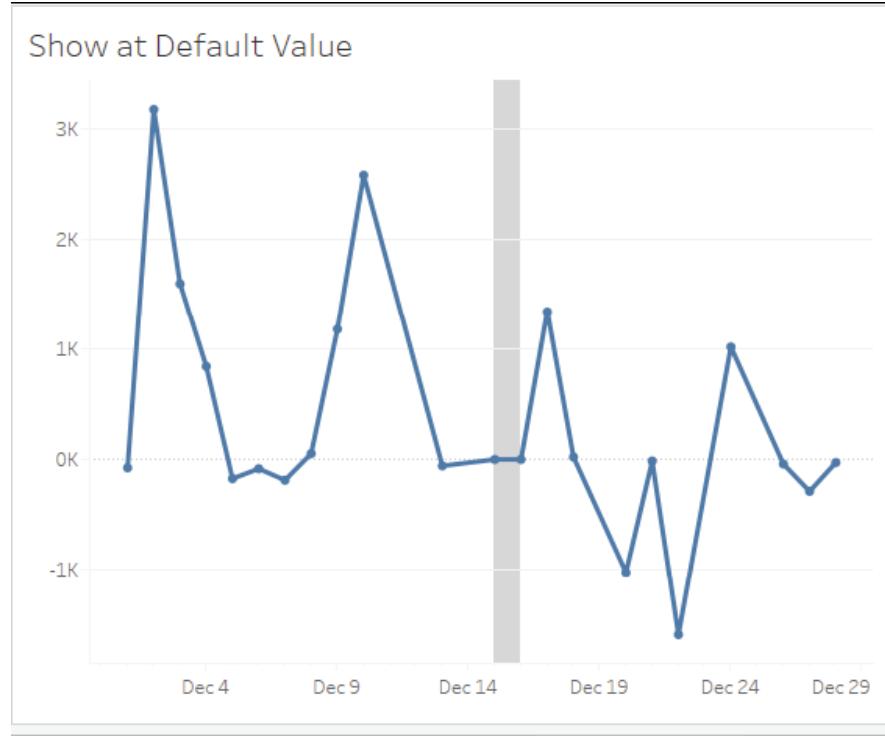
Null formatting

- Show at Indicator reveals no marks in the gray band with the number of null values indicated in the lower-right corner:



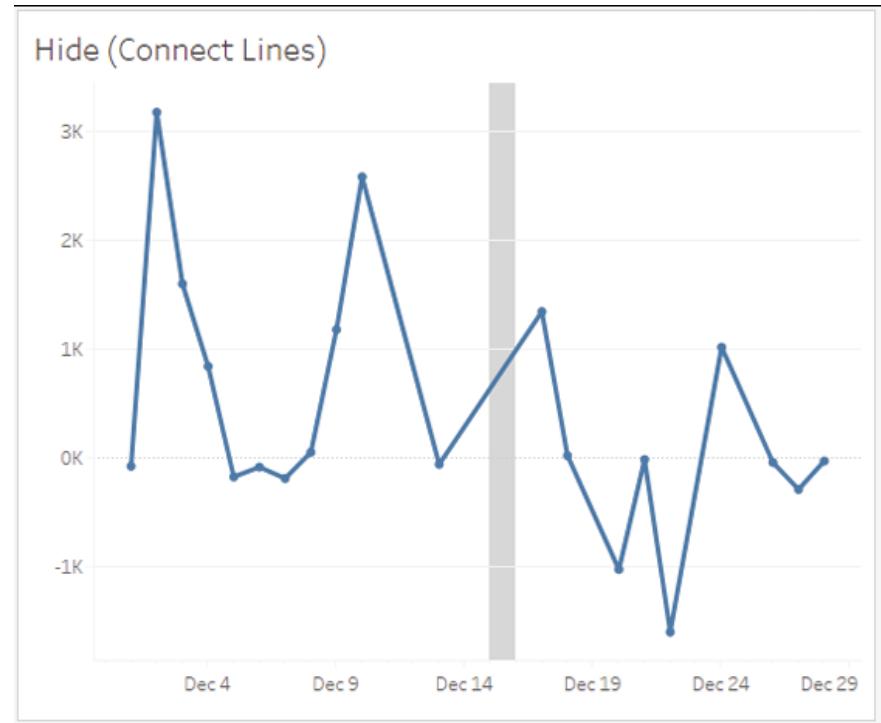
Null formatting

- Show at Default Value places marks at 0 and connects the lines:



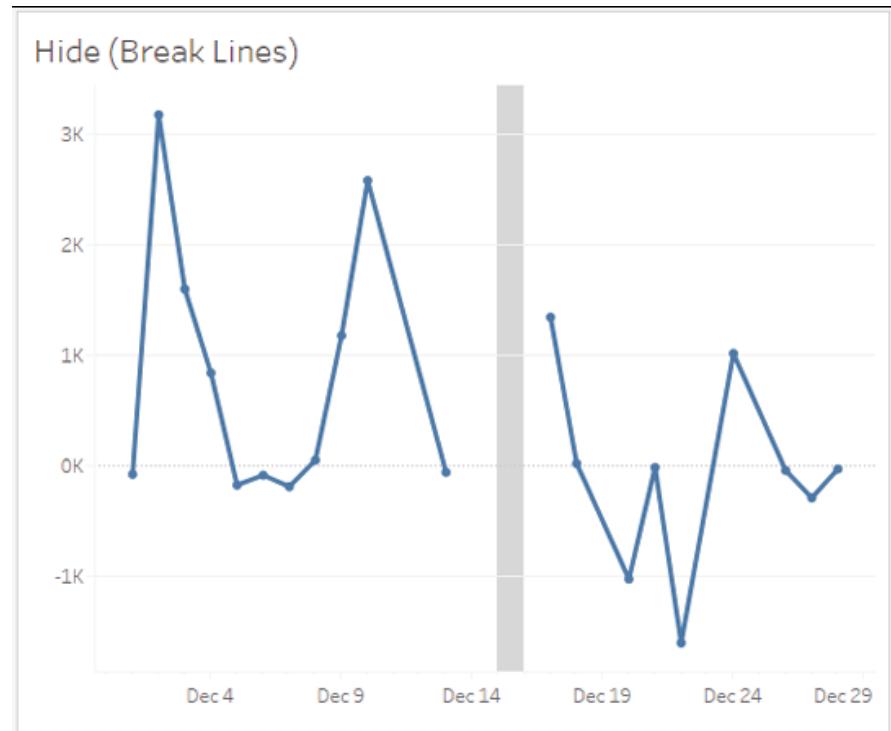
Null formatting

- Hide (Connect Lines) removes marks for the missing values, but does connect the existing marks:



Null formatting

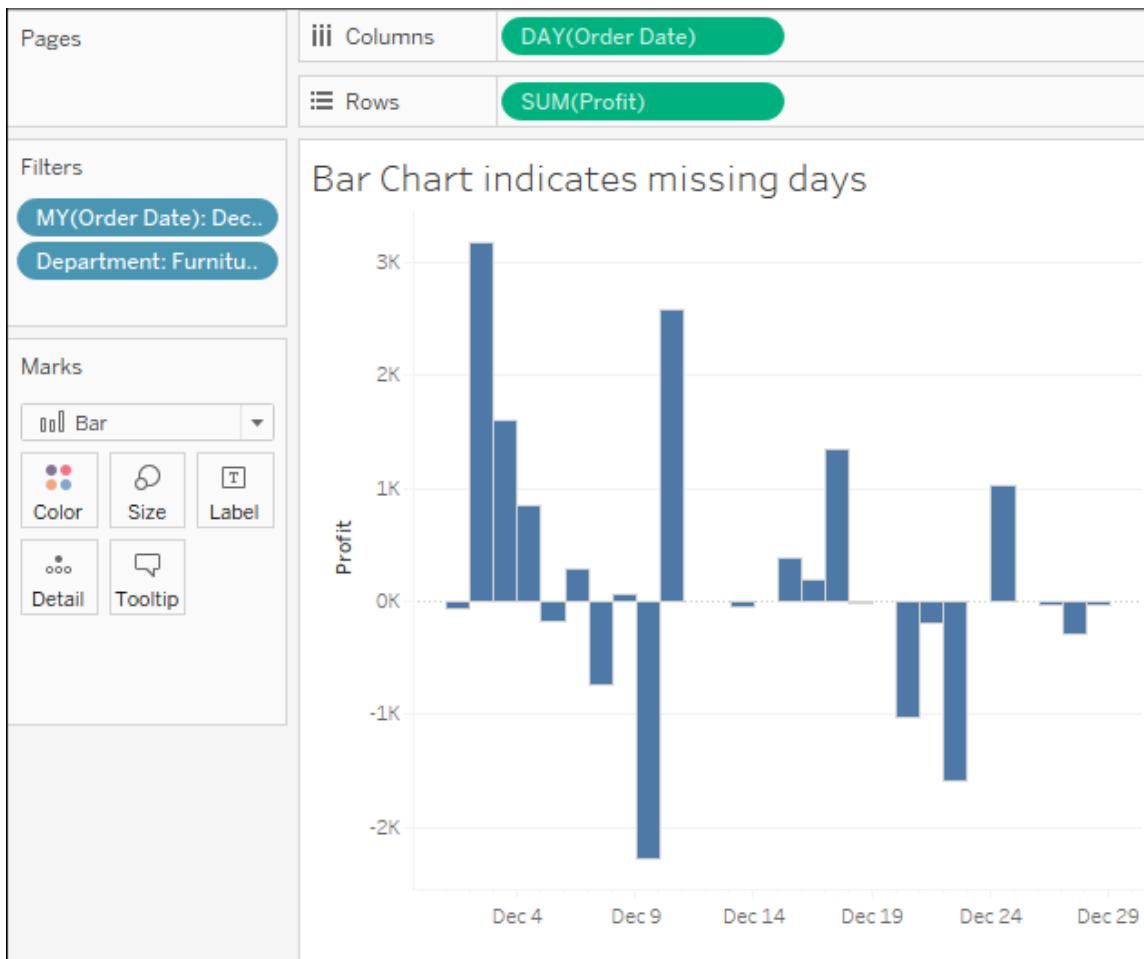
- Hide (Break Lines) removes the marks for the missing values and does not connect the existing marks:



Null formatting

- Some mark types support additional options such as border and/or halo, as shown here:



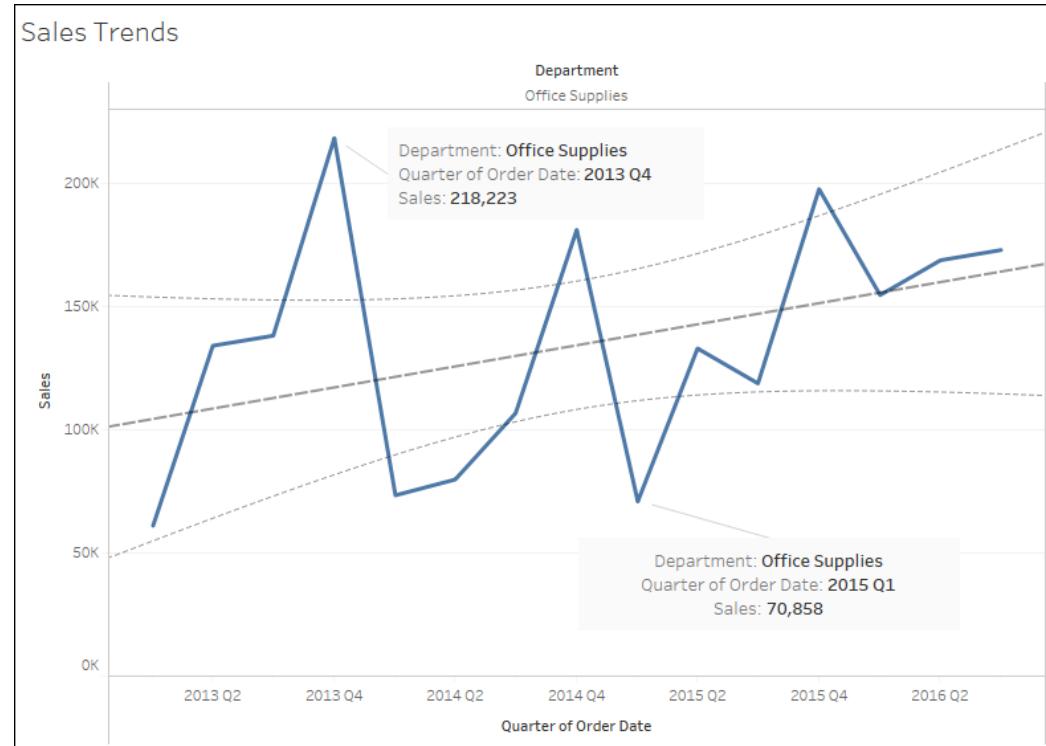


Adding value to visualizations

- Now that we've considered how formatting works in Tableau, let's look at some ways in which formatting can add value to a visualization.
- When you apply custom formatting, always ask yourself what the formatting adds to the understanding of the data.

Adding value to visualizations

- Consider the following visualization, all using default formatting:



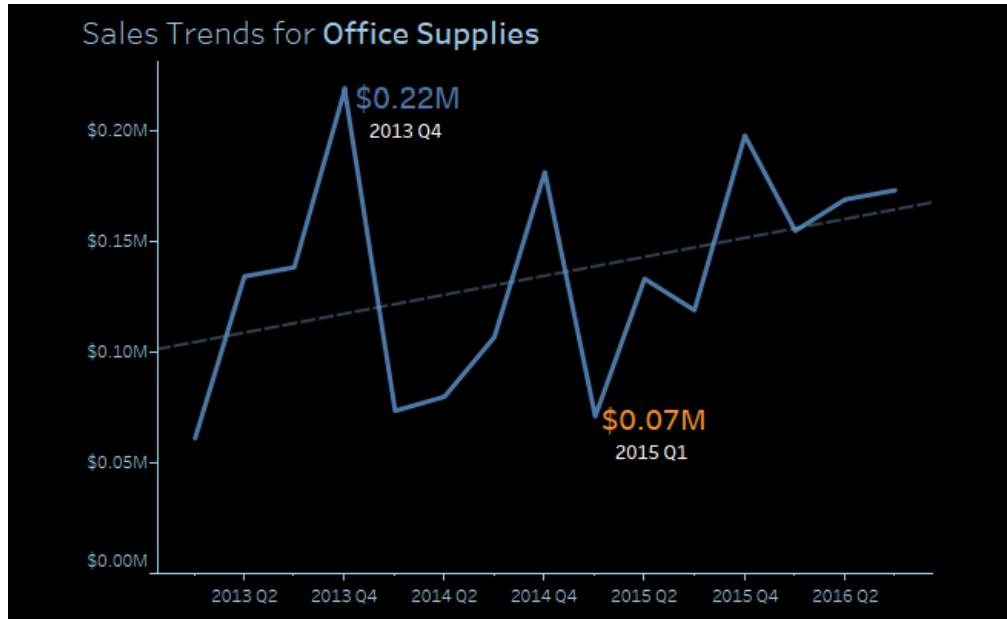
Adding value to visualizations

- The default format works fairly well, but compare that to this visualization:



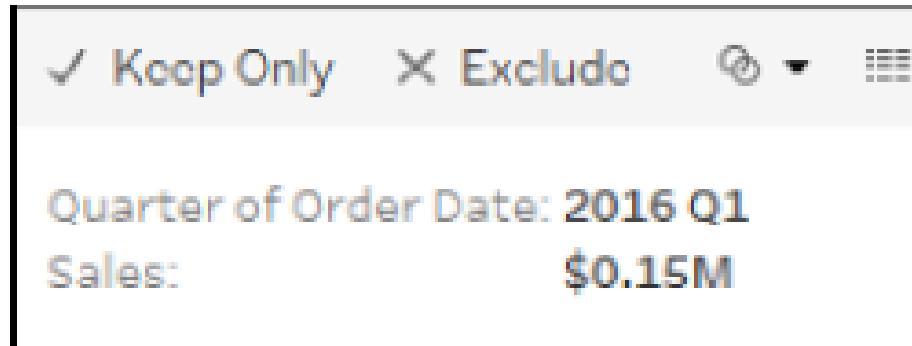
Adding value to visualizations

- Formatting can also be used to dramatically alter the appearance of a visualization.
- Consider the following chart:



Tooltips

- Consider the following default tooltip that displays when the end user hovers over one of the marks shown in the preceding screenshot:

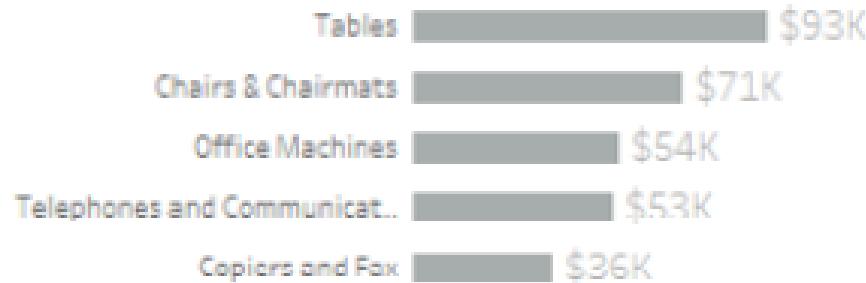


Tooltips

- Compare it to this tooltip:

Office Supplies

Sales of \$0.12M in 2015 Q3



Edit Tooltip

Tableau Book 10 B I U Insert X

<ATTR(Department)>
Sales of <SUM(Sales)> in <QUARTER(Order Date)>

<Sheet name="Categories" maxwidth="300" maxheight="300"
filter="

Show tooltips Responsive - Show tooltips instantly ▾
Include command buttons
Allow selection by category

OK Cancel

Sheets Data Source Name Data Update Time Sheet Name Workbook Name Page Count Page Name Page Number Default Caption Default Title Full Name User Name ATTR(Department) Department QUARTER(Order Date) SUM(Sales) All Fields

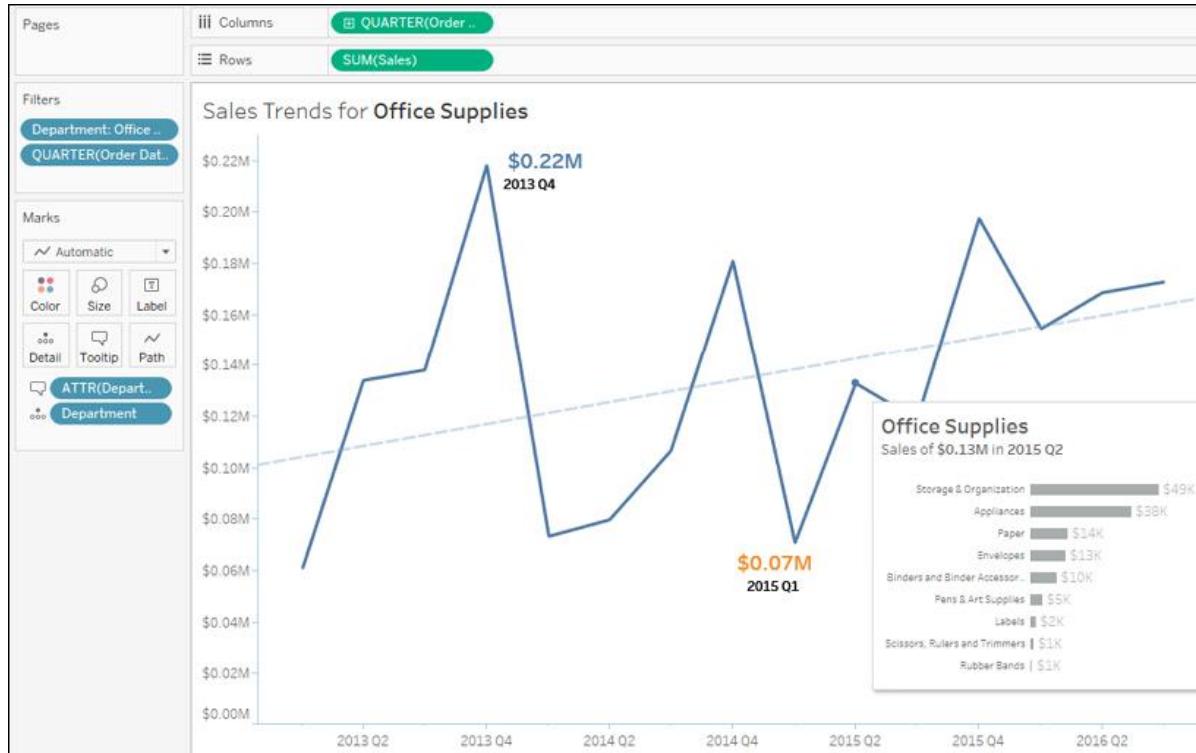
Bar Chart indicates missing days Categories Data Densification Dramatic Formatting Formatting: Parts of the View Hide (Break Lines) Hide (Connect Lines) Sales Trends Show at Default Value Show at Indicator Some formatting

Viz in Tooltip

- In the preceding screenshot, the following tag was added to the tooltip by selecting Insert | Sheets | Categories:

```
<Sheet name="Categories" maxwidth="300"  
maxheight="300" filter="<All Fields>">
```

Notice the final view with the tooltip:



Summary

- The goal of formatting is to increase effective communication of the data at hand.
- Always consider the audience, setting, mode, mood, and consistency as you work through the iterative process of formatting.
- Look for formatting that adds value to your visualization and avoid useless clutter.

COMPLETE LAB 7

8. Telling a Data Story with Dashboards

Telling a Data Story with Dashboards

This lesson will cover the following topics:

- Introduction to dashboards
- Designing dashboards in Tableau
- Designing for different displays and devices
- Interactivity with actions
- Stories

Introduction to dashboards

Dashboard definition

- From a Tableau perspective, a dashboard is an arrangement of individual visualizations, along with other components, such as legends, filters, parameters, text, containers, images, extensions, buttons, and web objects, that are arranged on a single canvas. Ideally, the visualizations and components should work together to tell a complete and compelling data story.
- Dashboards are usually (but not always) interactive.

Dashboard objectives

- The primary objective of a dashboard is to communicate data to a certain audience with an intended result.
- Often, we'll talk about telling the data story, That is, there is a narrative (or multiple narratives) contained within the data that can be communicated to others.
- While you can tell a data story with a single visualization or even a series of complex dashboards, a single Tableau dashboard is the most common way to communicate a single story.

Dashboard objectives

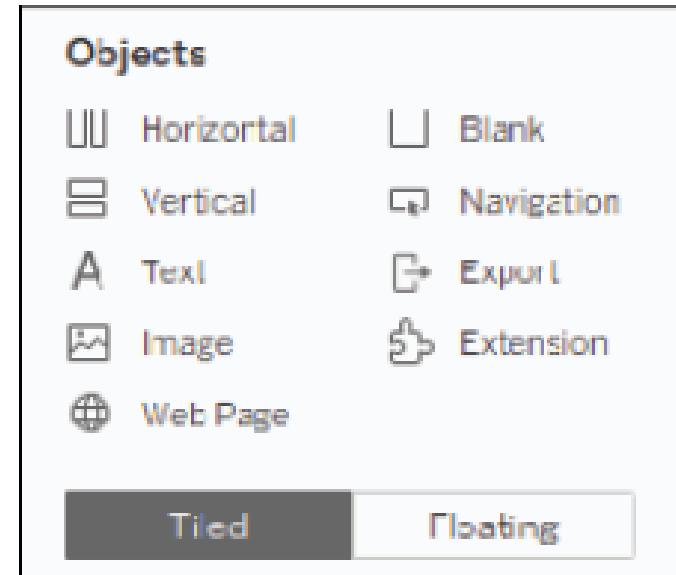
There are numerous possible approaches to building dashboards based on your objectives. The following is by no means a comprehensive list:

- Guided analysis
- Exploratory
- Scorecard/Status snapshot
- Narrative

Designing dashboards in Tableau

Objects

- Dashboards are made up of objects that are arranged on a canvas.
- You'll see a list of objects that can be added to a dashboard in the left-hand pane of a dashboard:

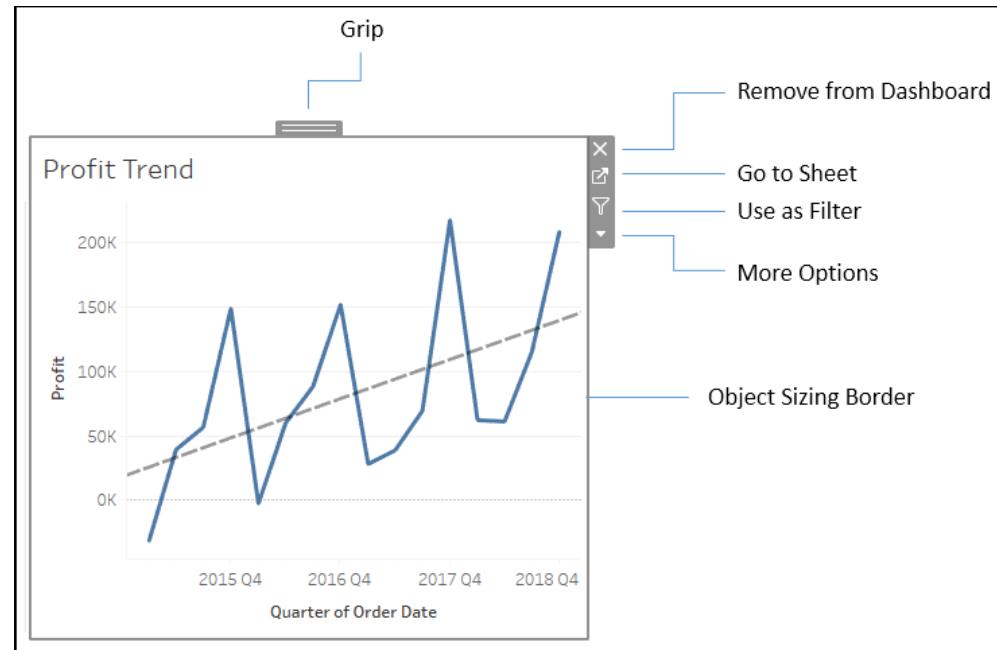


Tiled versus floating

- An object is either tiled or floating. If it is a tiled object, it will snap into the dashboard or layout container where you drop it.
- If it is a floating object, it will float over the dashboard in layers.
- You can change the order of the layers for a floating object.
- You'll notice the Tiled or Floating buttons directly beneath the Objects pallet in the preceding image.

Manipulating objects on the dashboard

- You may wish to manipulate an object once it is part of a dashboard.
- Every object has certain controls that become visible when you select it:



A dashboard to understand profitability

Having covered some conceptual topics as well as practical matters related to dashboard design, we'll dive into an example.

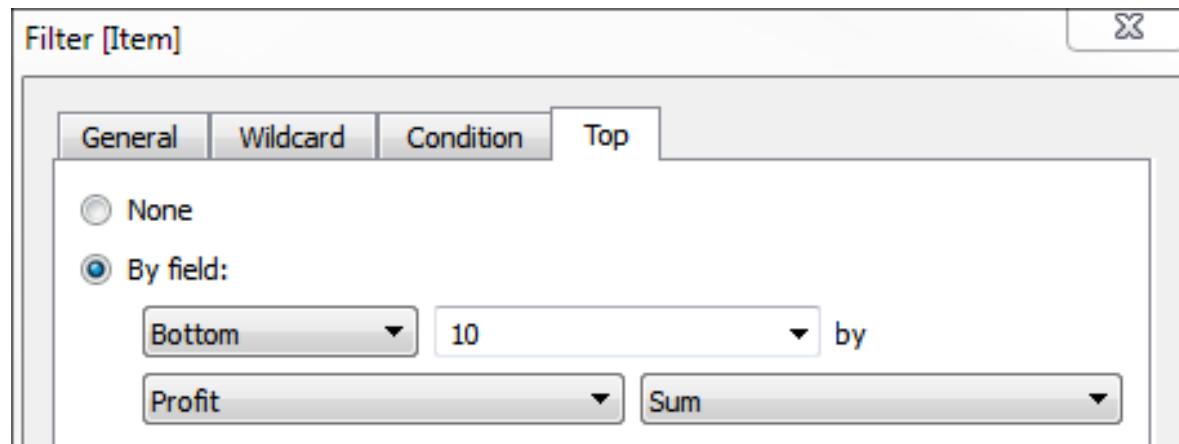
- Let's say you've been tasked with helping management find which items are the least profitable.
- Management feels that most of the least profitable items should be eliminated from their inventory.

Building the views



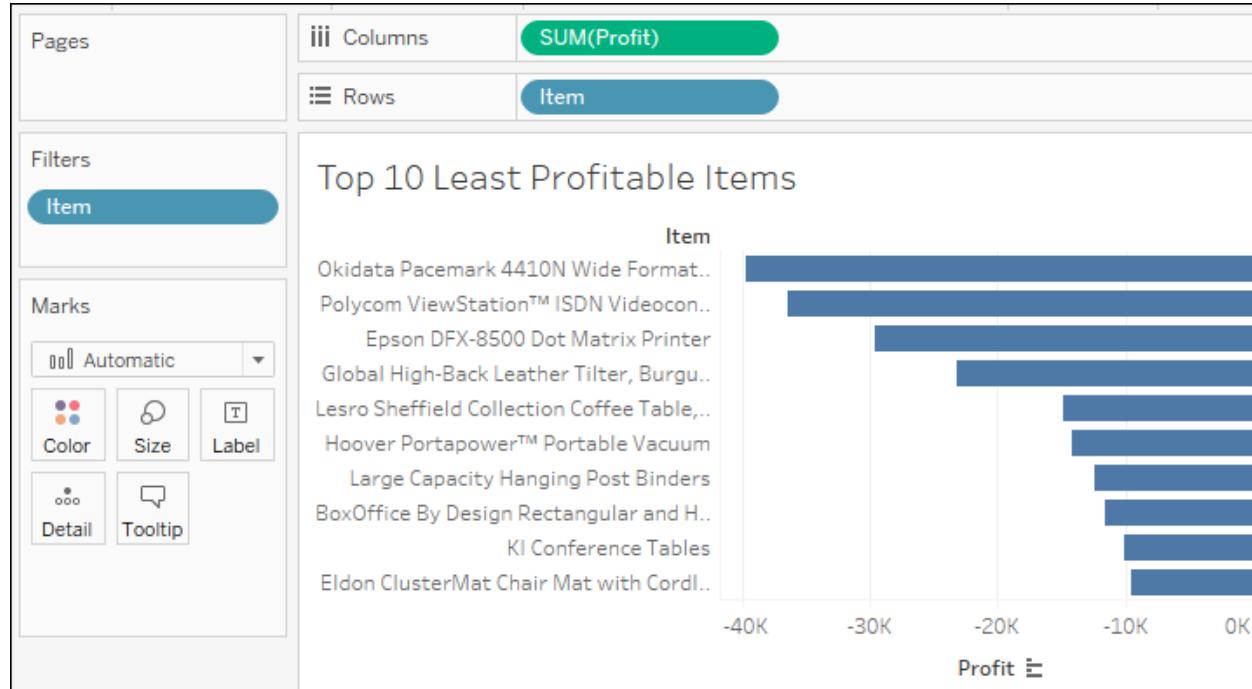
Building the views

- Select the Top tab, and adjust the settings to filter by field.
- Specify Bottom 10 by Sum(Profit):



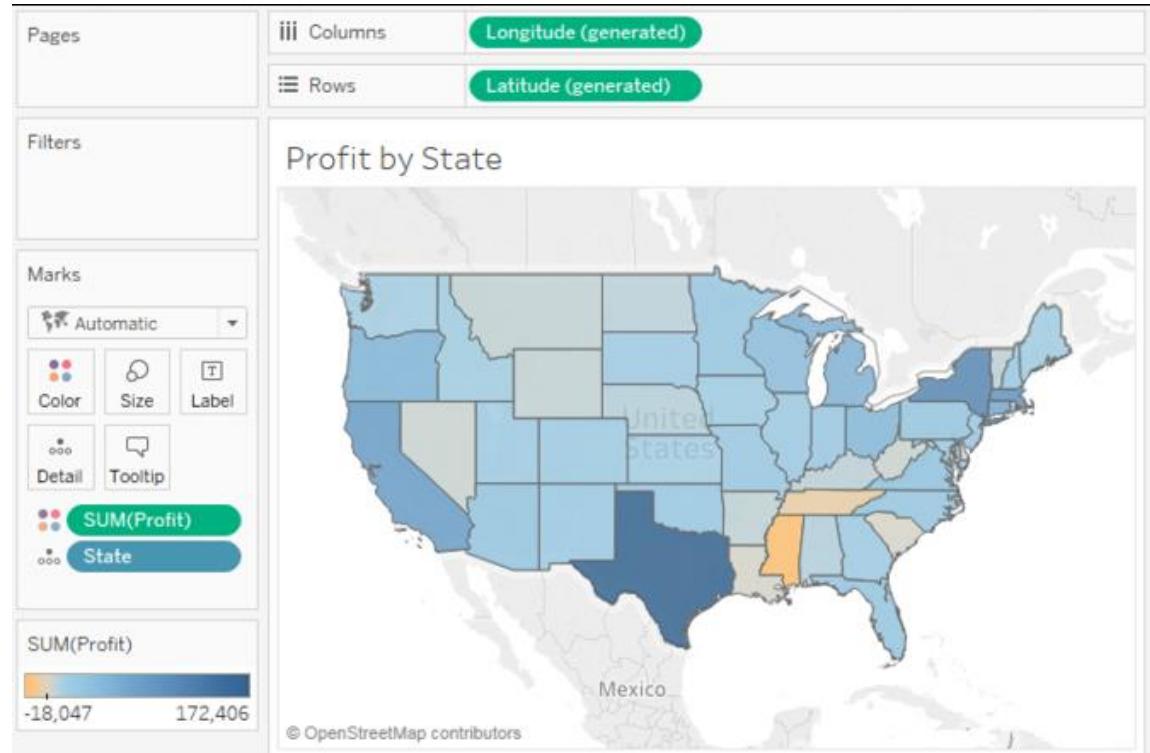
Building the views

- Rename the sheet Top 10 Least Profitable Items:



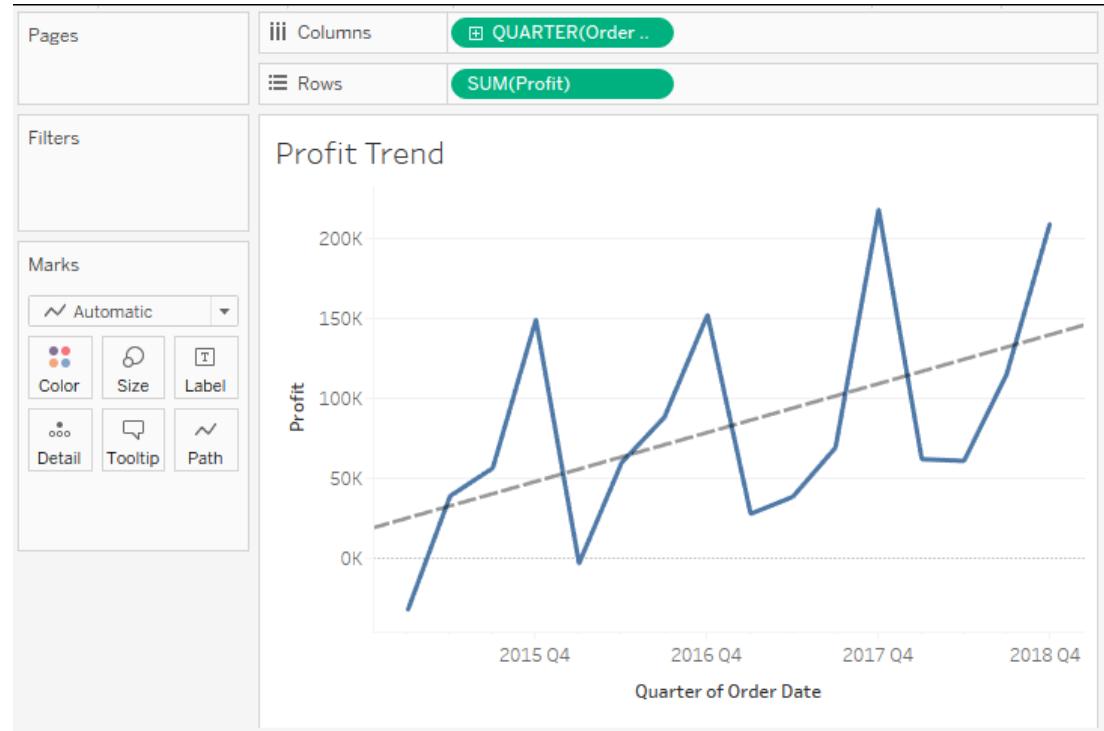
Building the views

- Rename the sheet to Profit by State



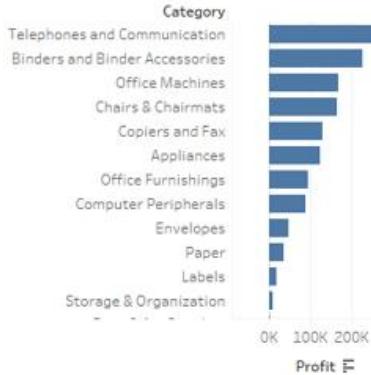
Building the views

- Rename the sheet to Profit Trend:

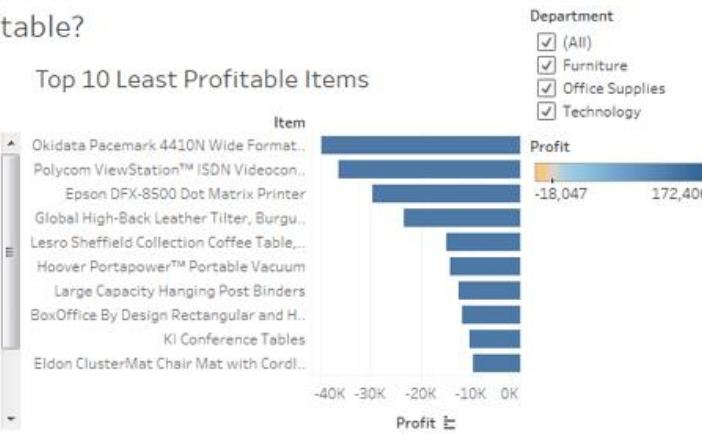


Is Least Profitable Always Unprofitable?

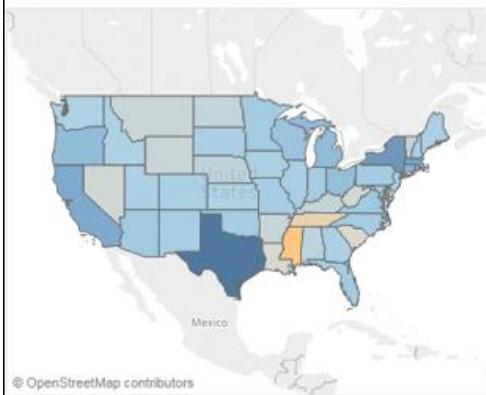
Overall Profit by Category



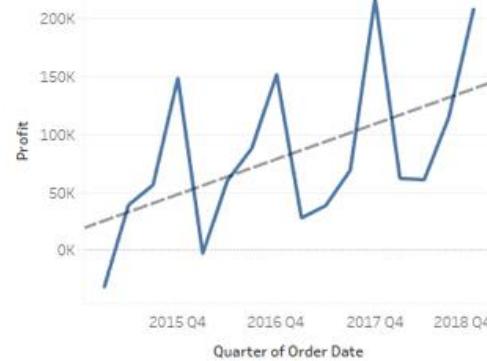
Top 10 Least Profitable Items



Profit by State



Profit Trend



Is Least Profitable Always Unprofitable?

1. Select a Department from the dropdown
2. Select a category below
3. Select an item below

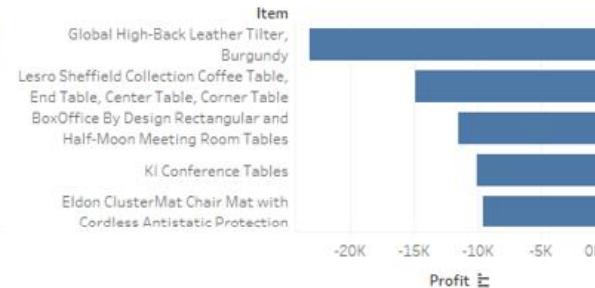
Department

Furniture

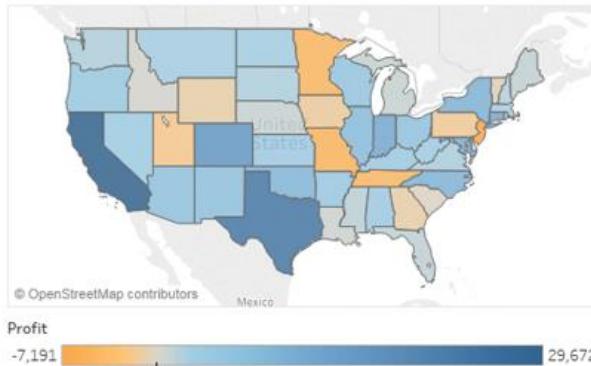
Overall Profit by Category



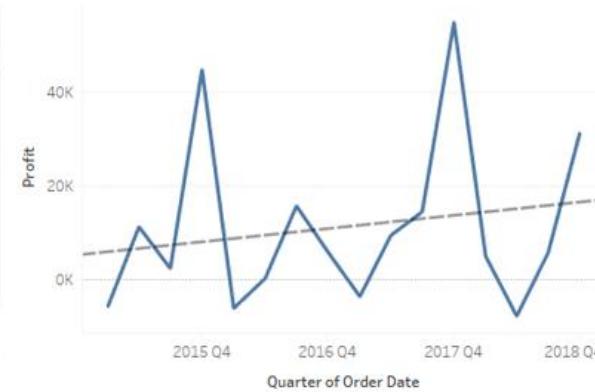
Top 10 Least Profitable Items



Profit by State

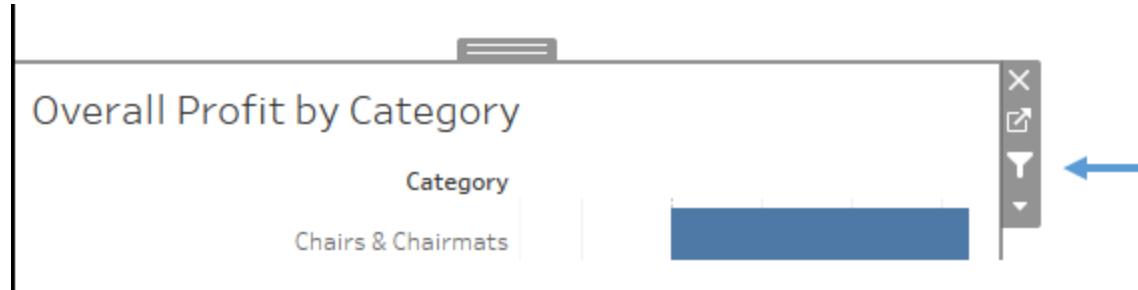


Profit Trend

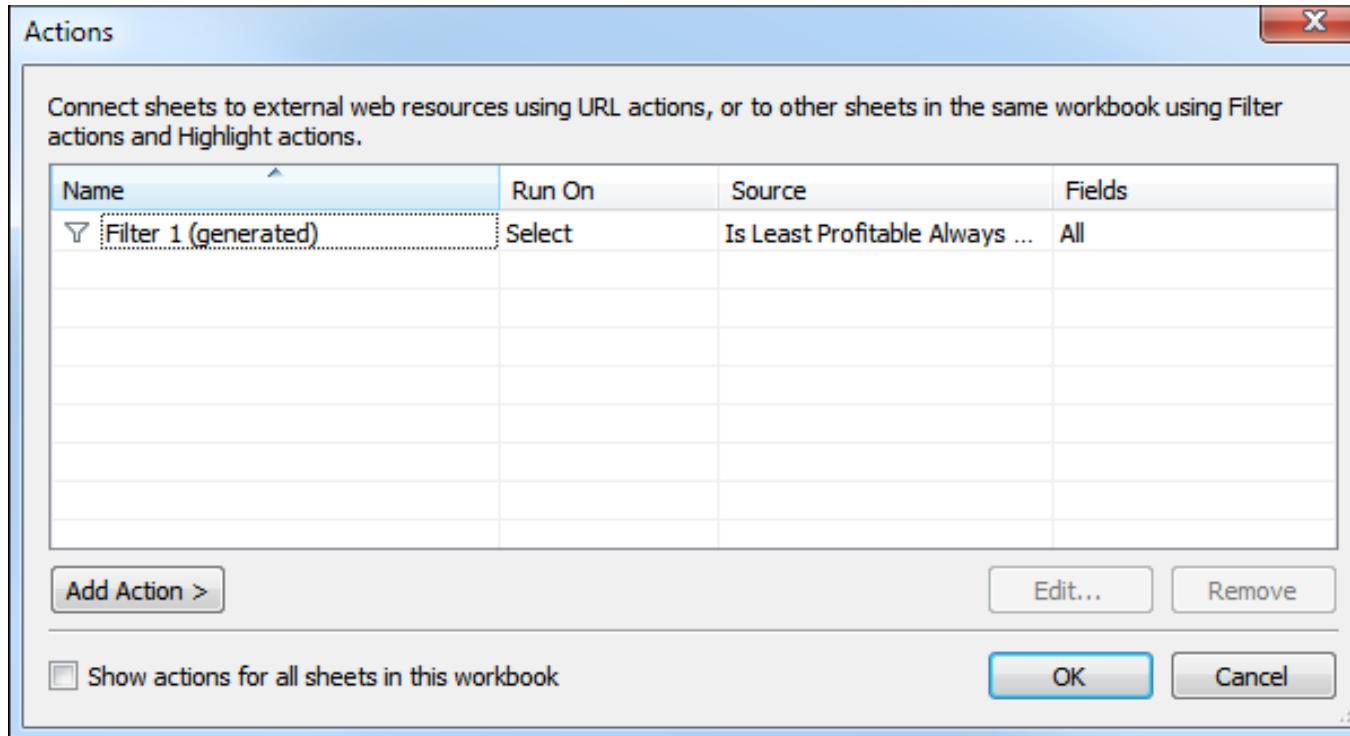


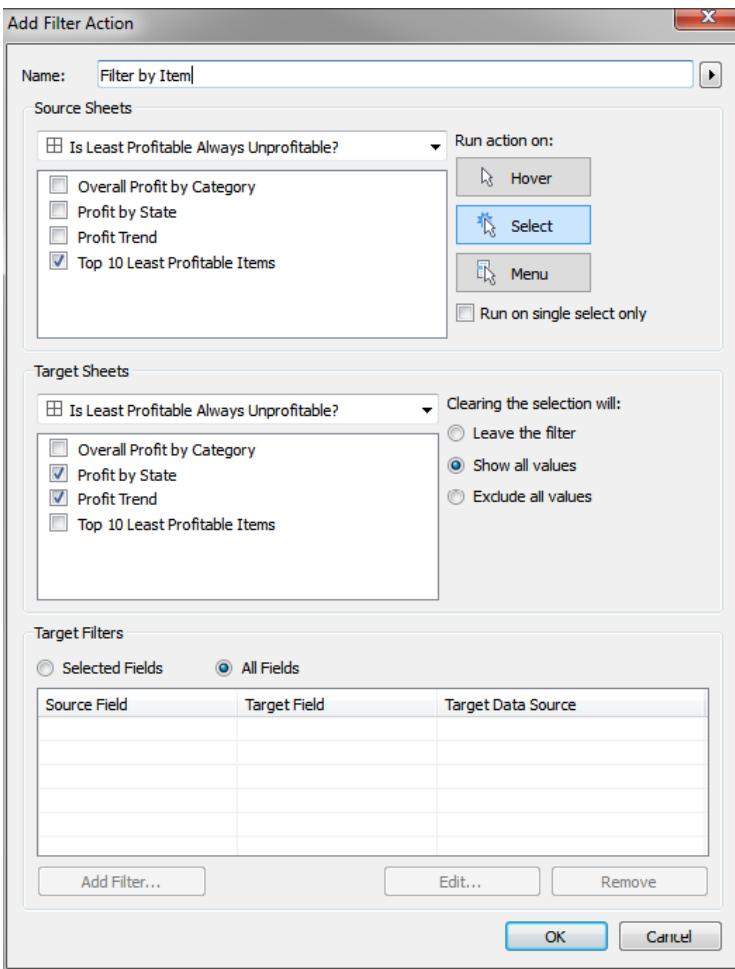
Implementing actions to guide the story

- Click the Use as Filter button on the Overall Profit by Category view.
- This will cause the view to be used as an interactive filter for the entire dashboard.
- That is, when the user selects a bar, all other views will be filtered based on the selection:



Implementing actions to guide the story





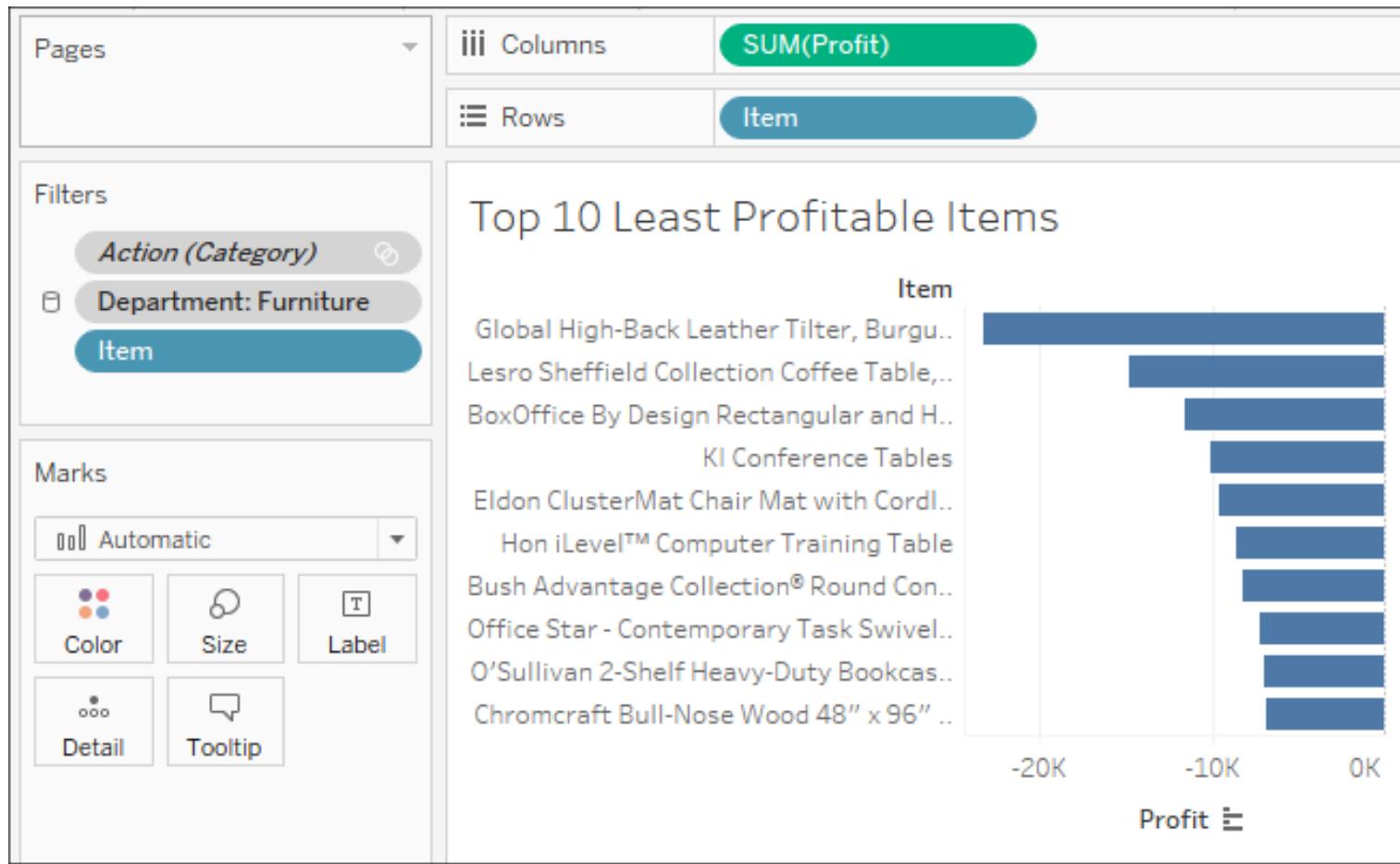
Implementing actions to guide the story

You now have three filters (two are actions) that drive the dashboard:

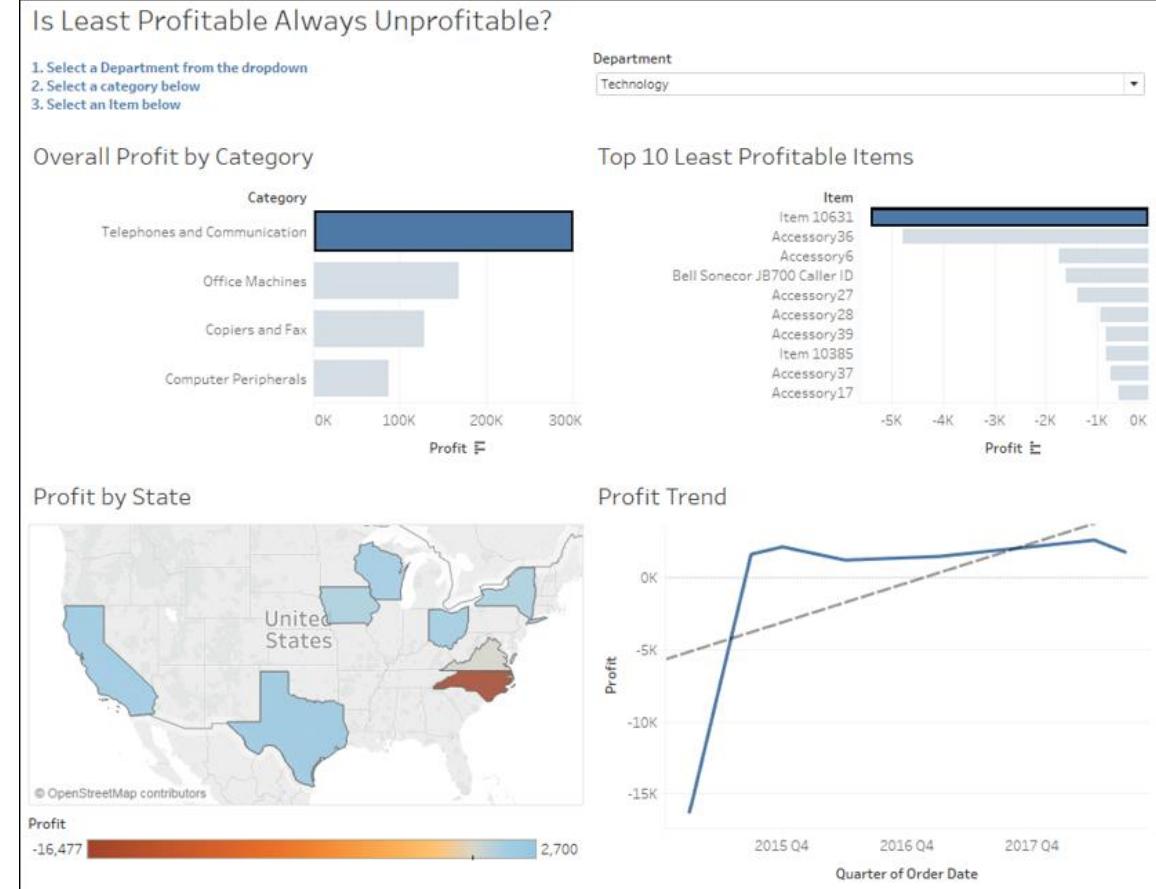
- Selecting a department from the drop-down will filter the entire dashboard (and all views in the workbook as you set it to filter every view using the data source)
- Selecting a category (clicking a bar or header) will filter the entire dashboard to that selection
- Selecting an item (clicking a bar or header) will filter the Profit by State and Profit Trend dashboards

Context filtering

- You may have noticed that when you use the drop-down filter to select a single department or select a single category, you have fewer than 10 items in the Top 10 view.
- For example, selecting Furniture from the Department filter and clicking on the bar for Tables results in only three items being shown.

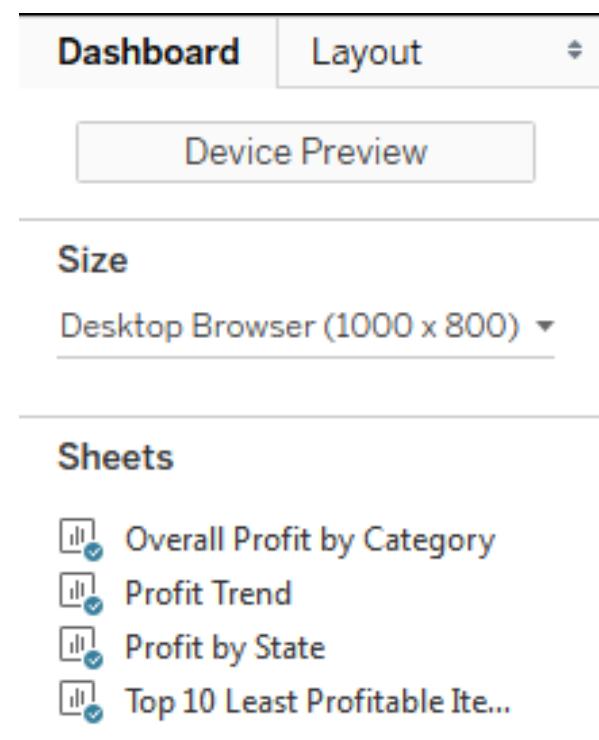


Implementing actions to guide the story



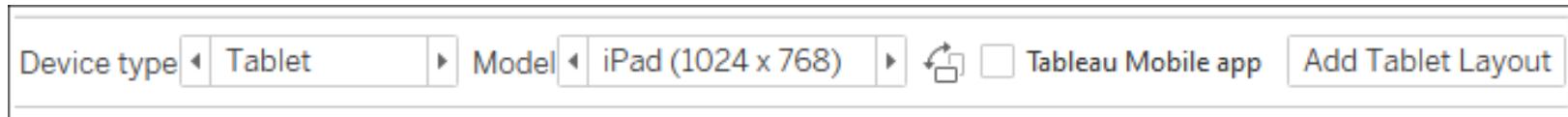
Designing for different displays and devices

- The top of the Dashboard tab on the left sidebar reveals a button to preview the dashboard on various devices, as well as a drop-down for Size options:



Designing for different displays and devices

- Clicking the Device Preview button not only allows you to see how your dashboard will look with various device types (and even specific models) but also allows you to add a layout for each device type, which you can customize:



Designing for different displays and devices

- Clicking the Add Layout button (that is, the Add Tablet Layout button in the preceding screenshot) will add a layout under the Dashboard tab on the left sidebar:

The screenshot shows a user interface for managing dashboard layouts. At the top, there are tabs for 'Dashboard' and 'Layout'. Below the tabs, there's a dropdown menu with options: 'Default', 'Tablet' (which is selected and highlighted in grey), and 'Phone'. A large dark grey button labeled 'Device Preview' is positioned below the device selection. Under the 'Tablet' section, there are three radio buttons: 'Default' (unchecked), 'Fit all' (checked), and 'Fit width' (unchecked). To the right of these buttons is a 'Height' input field with up and down arrows. Below this section is another titled 'Layout - Tablet' with two buttons: 'Default' (selected and highlighted in grey) and 'Custom'. A descriptive note below the buttons states: 'The content and layout of the dashboard on all devices will match the Default dashboard.' At the bottom, another note says: 'Click "Custom" to change how the dashboard will look for the devices.'

Dashboard **Layout**

Default **Phone**

Device Preview

Size - Phone

- Default
- Fit all
- Fit width

Height

Layout - Phone

- Default **Custom**

- Overall Profit by Category
- Department
- Profit Trend
- Profit by State
- Profit
- Top 10 Least Profitable Items
- 1. Select a Department from the dropdown
- Is Least Profitable Always Unprofitable?

Device Preview **Device type**

Is Least Profitable Always Unprofitable?

1. Select a Department from the dropdown
2. Select a category below
3. Select an item below

Department: Technology

Overall Profit by Category

Category	Profit
Telephones and Comm...	~250K
Office Machines	~150K
Copiers and Fax	~100K
Computer Peripherals	~50K

Top 10 Least Profitable Items

Item	Profit
Item 10631	~-3K
Accessory36	~-2K
Accessory6	~-1.5K
Bell Sonoco JB700 Caller ID	~-1K
Accessory27	~-800
Accessory28	~-700
Accessory39	~-600
Item 10385	~-500
Accessory37	~-400
Accessory17	~-300

Profit by State

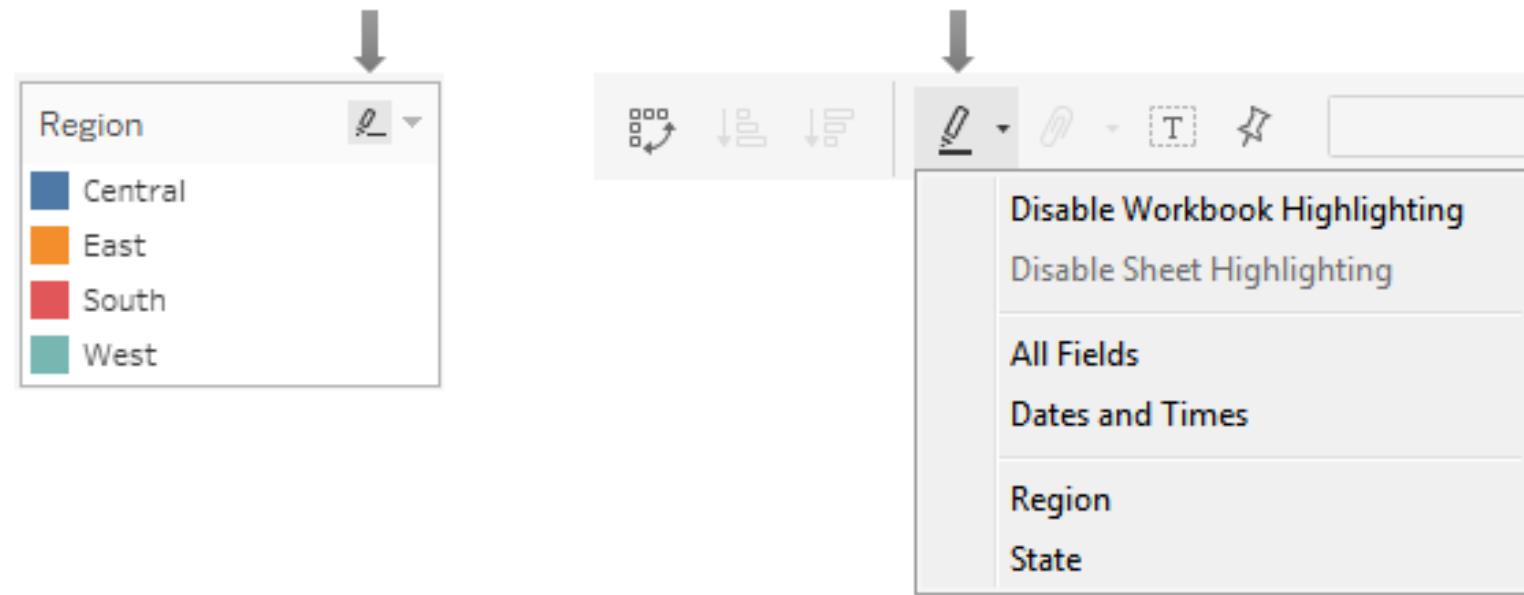
The map shows state-level profit data across the United States. States colored in blue include California, Texas, New York, Massachusetts, Connecticut, Rhode Island, New Jersey, Pennsylvania, Ohio, Michigan, Indiana, Illinois, Wisconsin, Minnesota, Iowa, Missouri, Kansas, Nebraska, Oklahoma, New Mexico, Colorado, Wyoming, Montana, North Dakota, South Dakota, and Alaska. States colored in red include Virginia, North Carolina, South Carolina, Georgia, and Florida.

Interactivity with actions

Tableau supports six kinds of actions:

- Filter actions
- Highlight actions
- URL actions
- Go to Sheet
- Parameter actions
- Set actions

Interactivity with actions

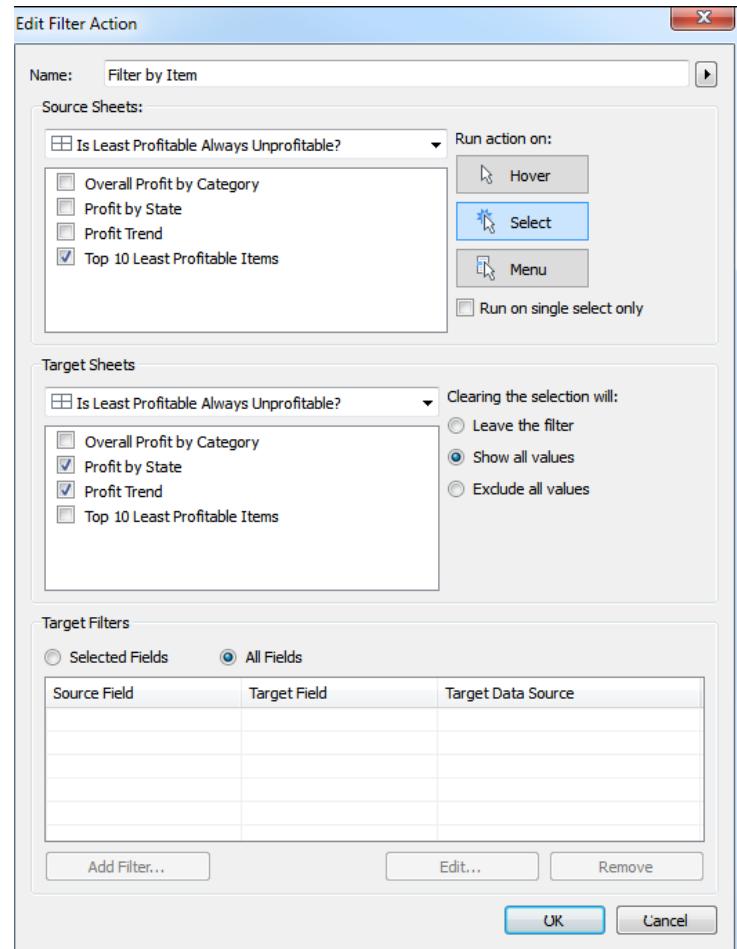


Filter actions

- Filter actions are defined by one or more source sheets that pass one or more dimensional values as filters to target sheets upon an action.
- Remember that every mark on a sheet is defined by a unique intersection of dimensional values.

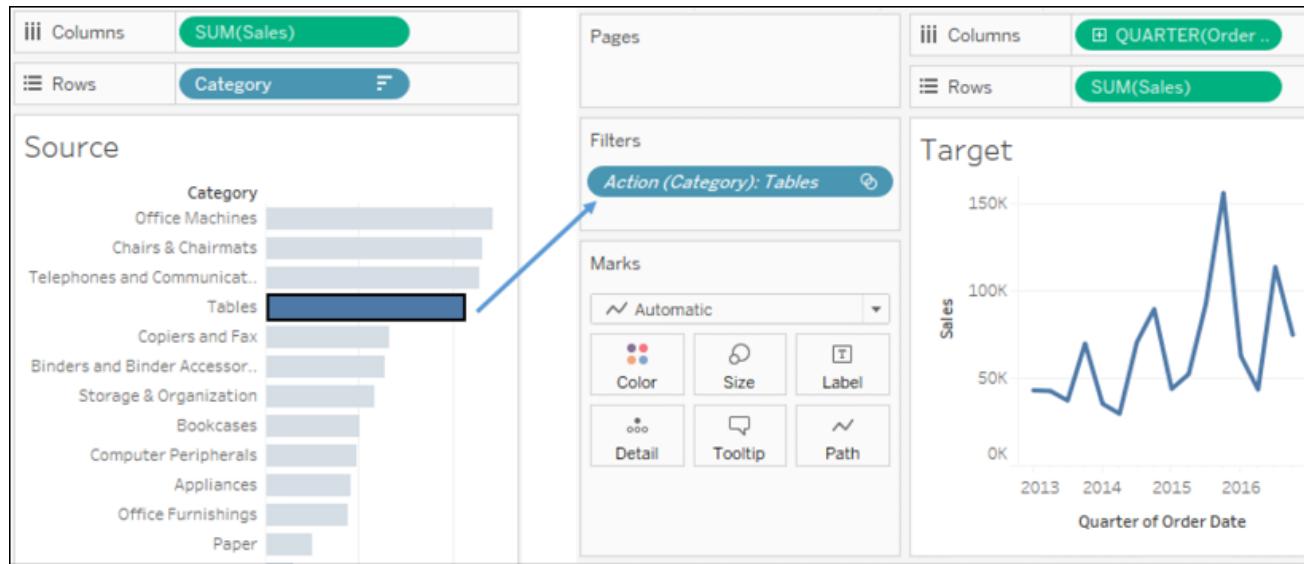
Filter actions

- When you create or edit a filter action, you will see options like these:



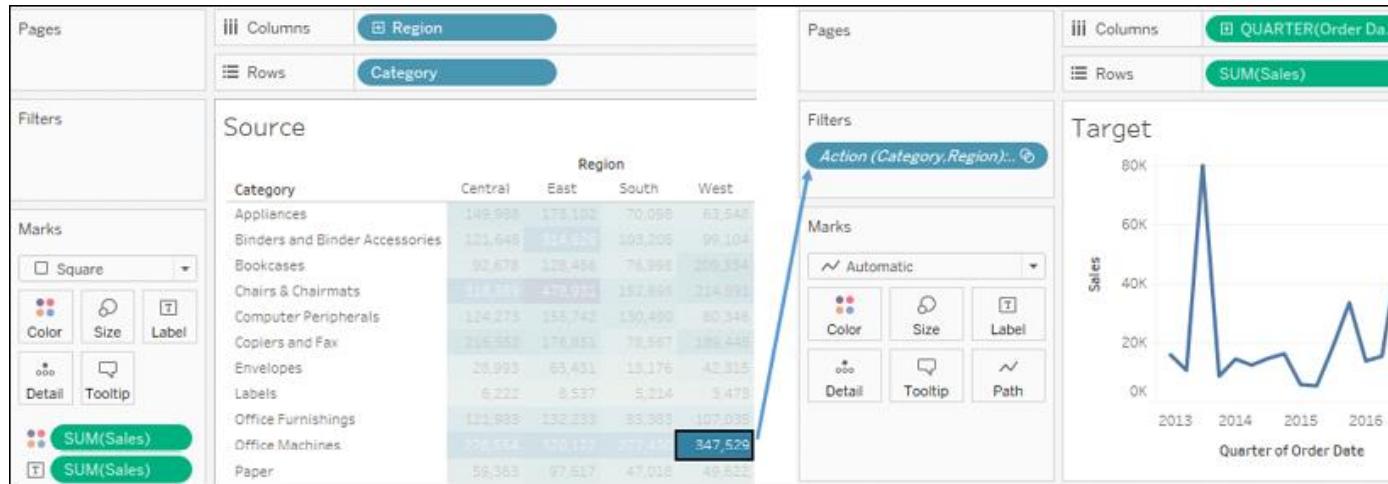
Filter actions

- Consider the following example of a filter action that's triggered when a bar is selected in the source:



Filter actions

- In this example, the filter contains Office Machines and West, matching the dimensions that define the selected square:



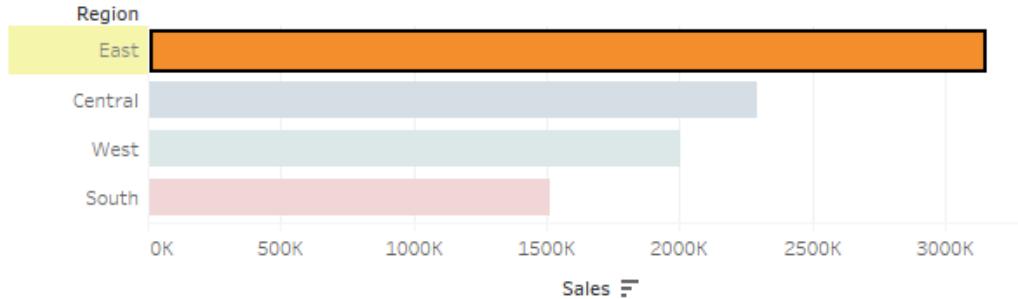
Highlight actions

- This type of action does not filter target sheets. Instead, highlight actions cause marks that are defined, at least in part, by the selected dimensional value(s) to be highlighted in the target sheets.
- The options for highlight actions are very similar to filter actions, with the same configuration options for source and target sheets, and which events are able to trigger the action.

Highlight actions

- Both the map and scatter plot have Region on the Detail part of the Marks card:

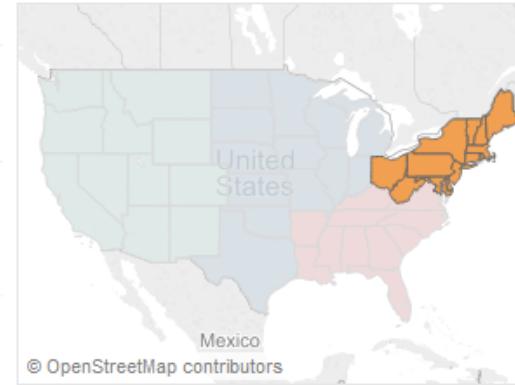
Sales by Region



Sales and Profit by State



Regions

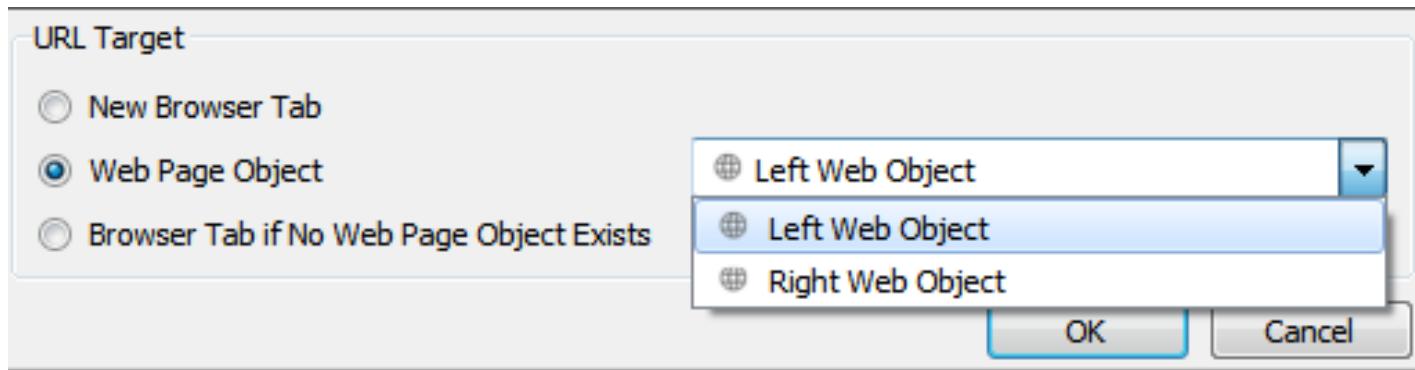


URL actions

- URL actions allow you to dynamically generate a URL based on an action and open it within a web object in the dashboard or in a new browser window or tab.
- URL actions can be triggered by the same hover, menu, and select events as other actions.
- The name of the URL action differentiates it and will appear as the link when used as a menu.

URL actions

- You may specify a target for the URL action when you create or edit the URL action:



URL actions

Item hierarchy

Dashboard

▼ Tiled

▼ Horizontal

Right Web Object

Left Web Object

Edit URL...

Floating

Fix Width

Edit Width...

Select Layout Container

Deselect

Remove from Dashboard

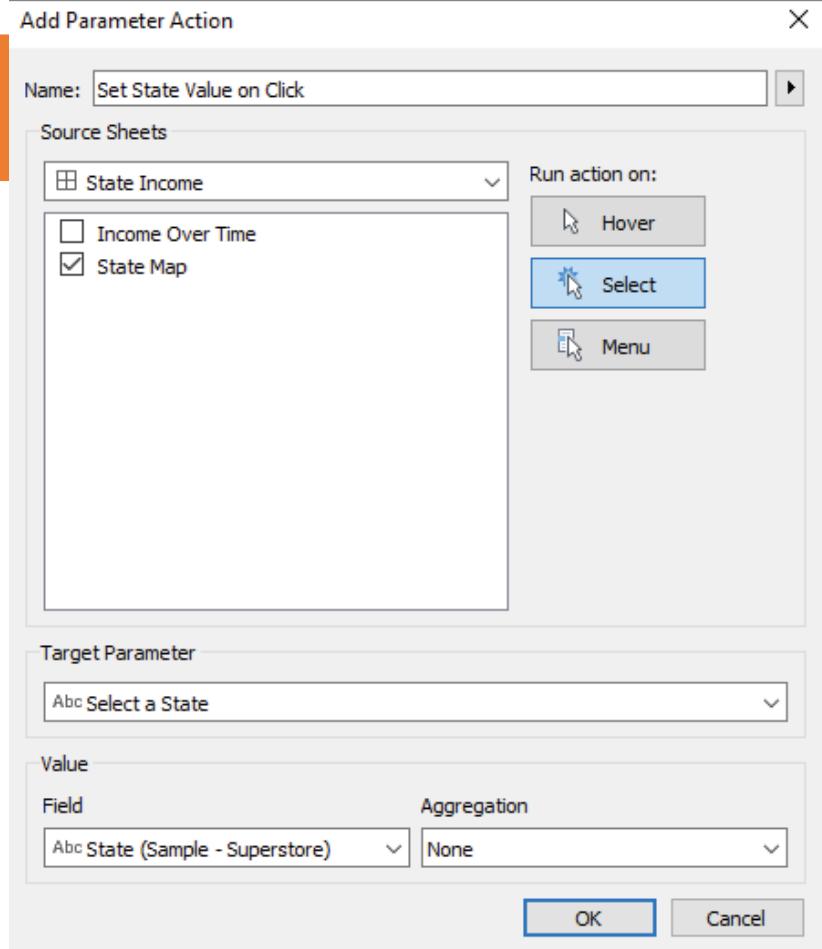
Rename Dashboard Item...

Go to Sheet actions

- Go to Sheet actions (also called navigation actions) allow you to specify navigation to another dashboard or sheet in the workbook based on user interaction with one or more views.
- For example, clicking a bar in a bar chart might navigate to another dashboard.
- These are similar to filter actions that define another sheet as a target but Go to Sheet actions do not set any filters.

Parameter actions

- Parameter actions allow you to set the value of a parameter based on a user action.
- When you create a parameter action, you'll set options using a screen like this:



Set actions

- Set actions allow you to populate a set with one or more values for one or more dimensions.
- This is very powerful because sets can be used on any shelf to encode any visual element, can be leveraged in calculations, and can be used as filters.
- They can be used in all of these ways—and in different ways—in different views.
- This gives you incredible flexibility in what can be accomplished with set actions.

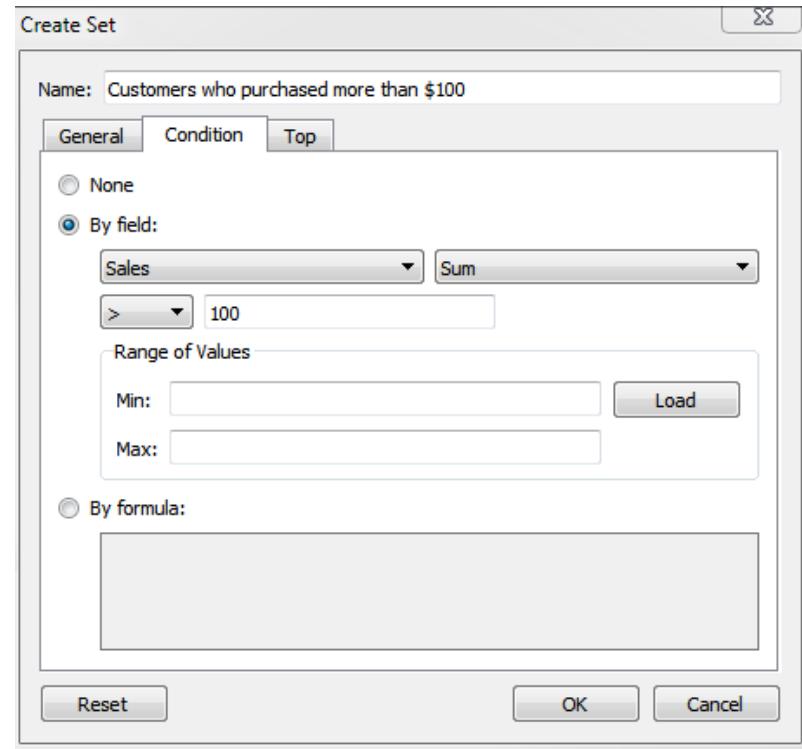
Sets

A set in Tableau defines a collection of records from the data source. At a row-level, each record is either in or out of the set. There are two types of sets:

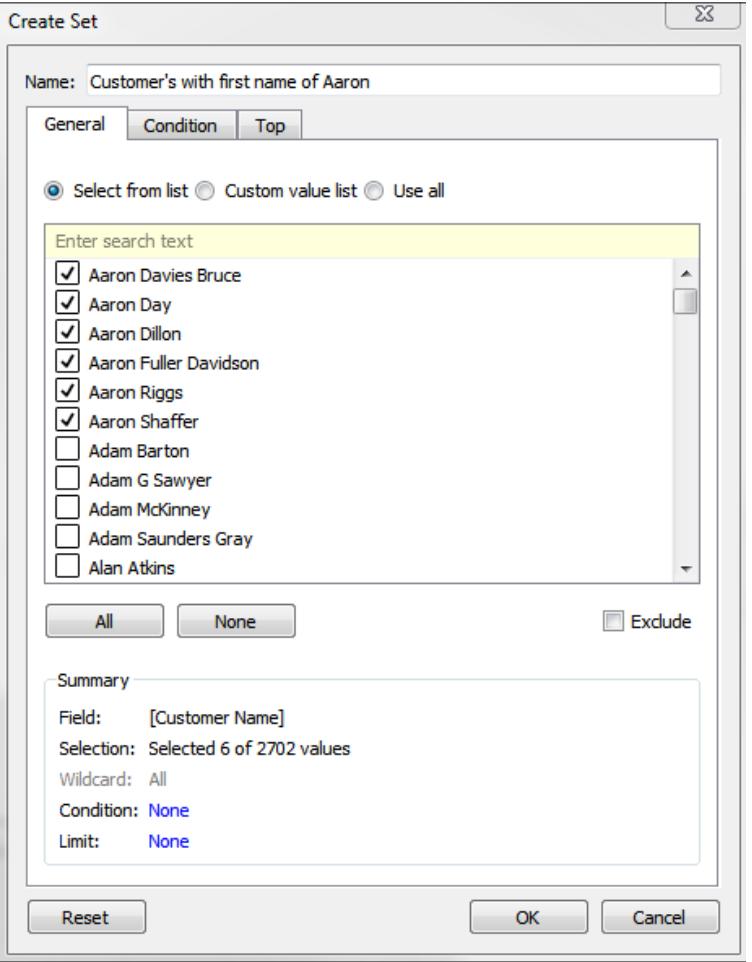
- Dynamic sets (sometimes called computed or calculated sets)
- Fixed sets

Sets

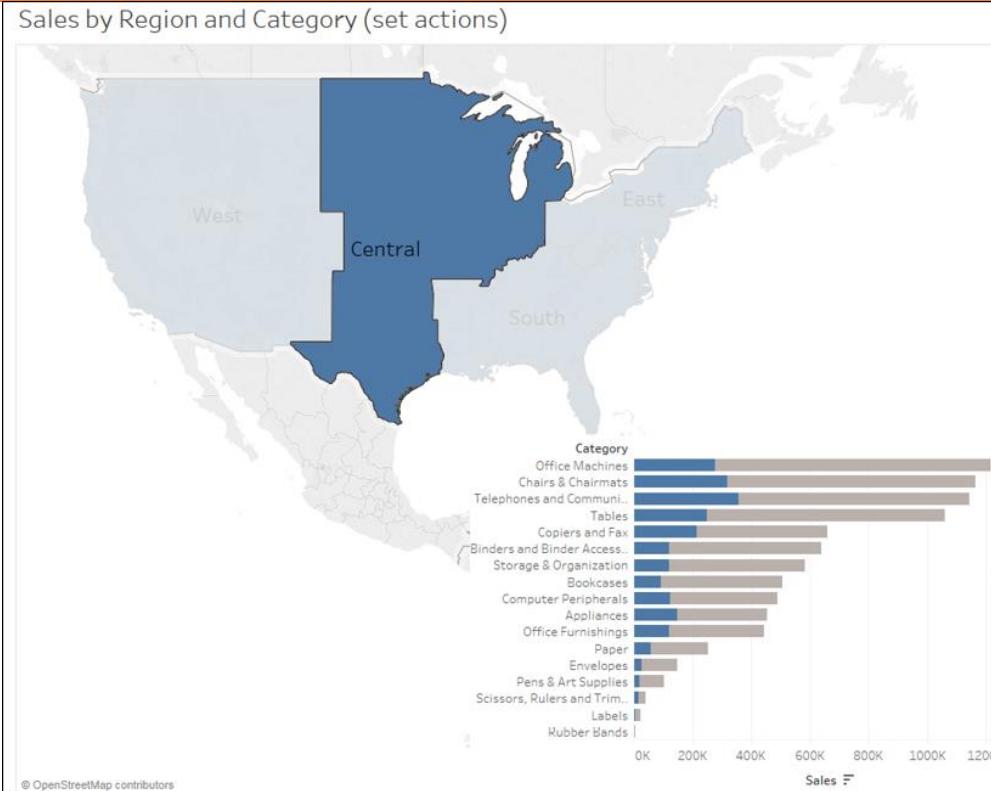
- For example, if you were to use the drop-down menu on Customer Name in the data pane and select Create | Set, then you could stipulate a condition that defines which records belong to the set:



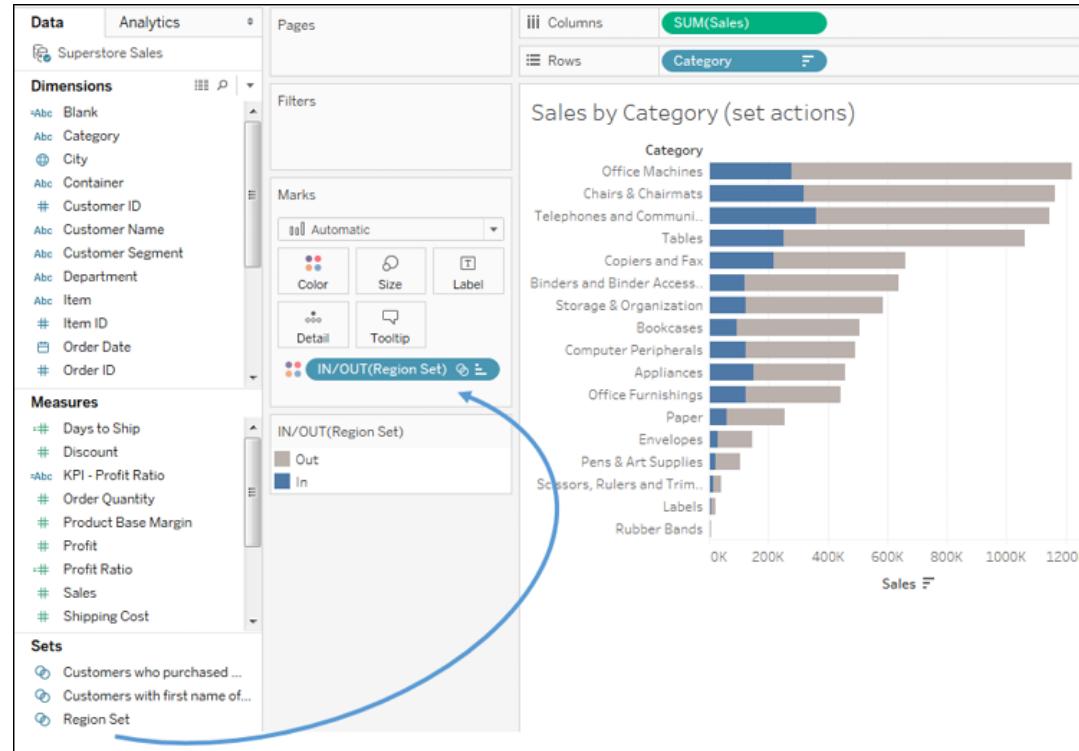
Sets



A set action example

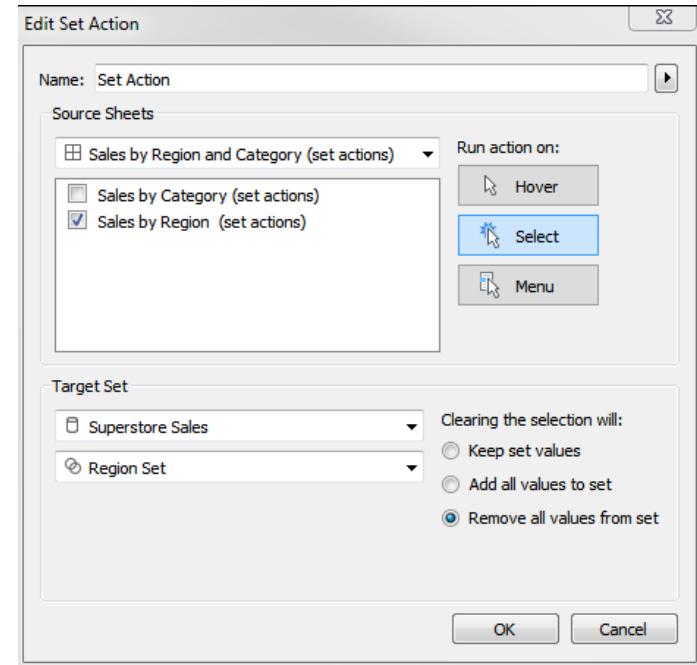


A set action example



A set action example

- Add a set action by selecting Dashboard | Actions from the menu and then Add Action | Change Set Values... in the resulting dialog.
- The resulting dialog has many similar features to other action types:

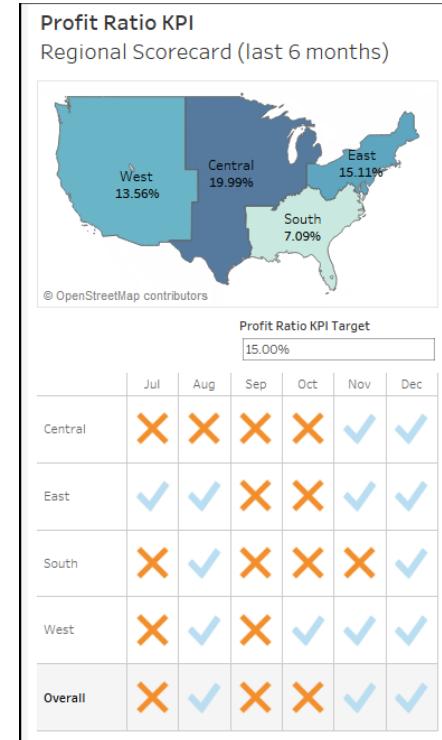


A regional scorecard dashboard

- Now, we'll consider another example dashboard that demonstrates slightly different objectives. Let's say everyone in the organization has agreed upon a KPI of profit ratio.
- Furthermore, there is consensus that the cut-off point between an acceptable and poor profit ratio is 15%, but management would like to have the option of adjusting the value dynamically to see whether other targets might be better.

A regional scorecard dashboard

- Consider the following dashboard:



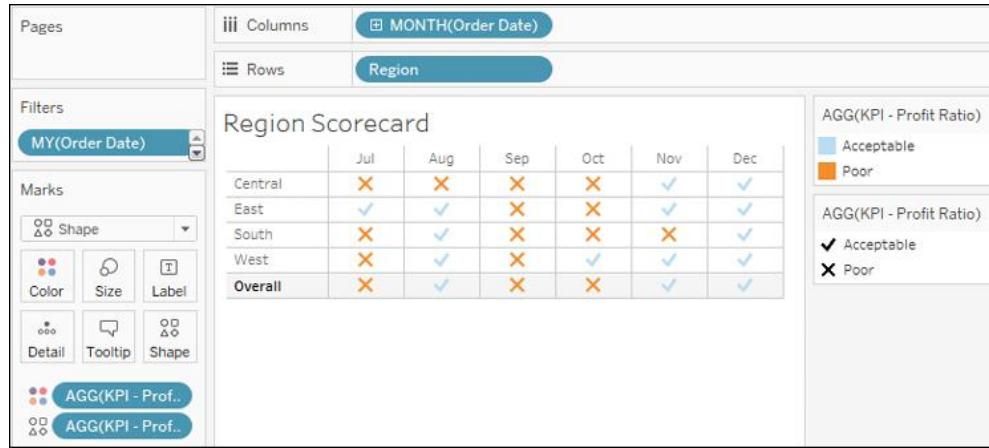
A regional scorecard dashboard

- Create a second calculation named KPI - Profit Ratio with the following code:

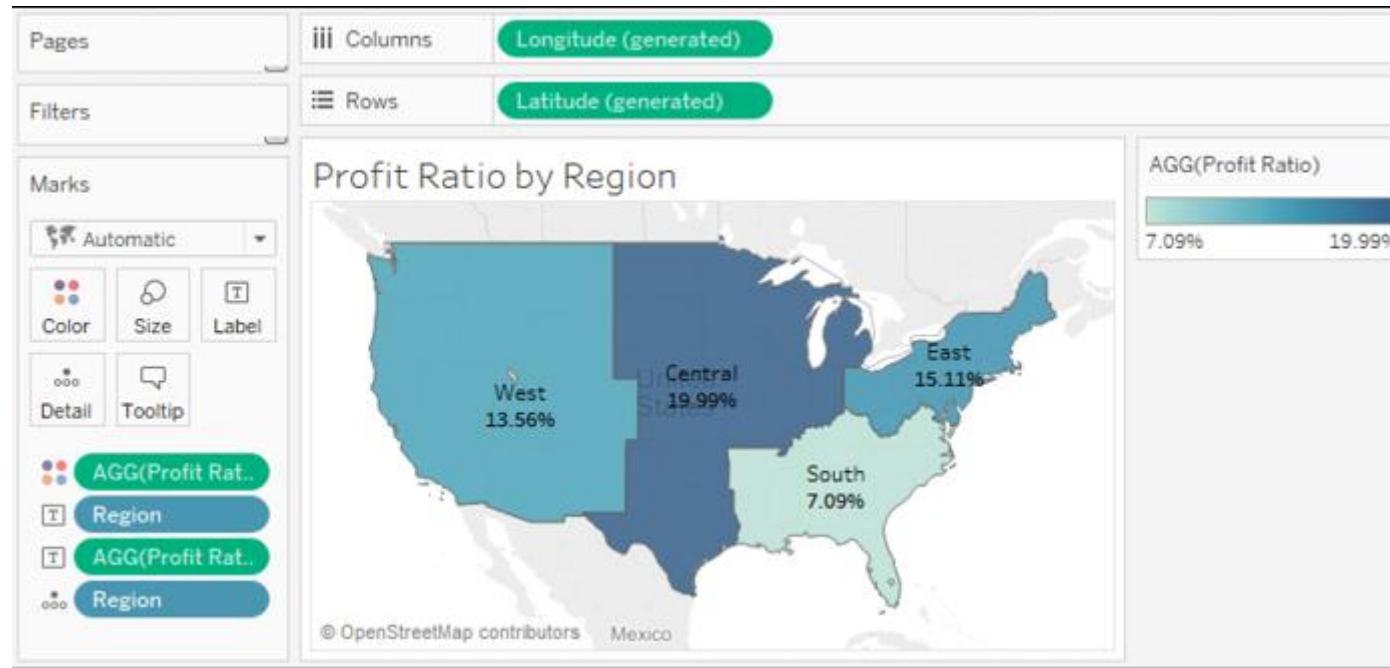
```
IF [Profit Ratio] >= [Profit Ratio KPI Target]  
THEN "Acceptable"  
ELSE "Poor"  
END
```

A regional scorecard dashboard

- Add Order Date as a filter and set it to the top 6 by field (Order Date as Min).
- This will dynamically filter the view to the last six months:

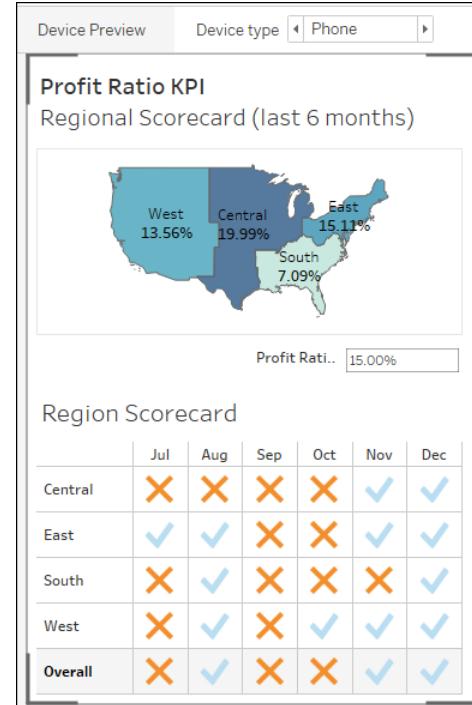


A regional scorecard dashboard



A regional scorecard dashboard

- Once both views have been created, you can arrange the views as a dashboard.
- The example in the lesson 08 Complete workbook has a phone layout applied to it as seen here:

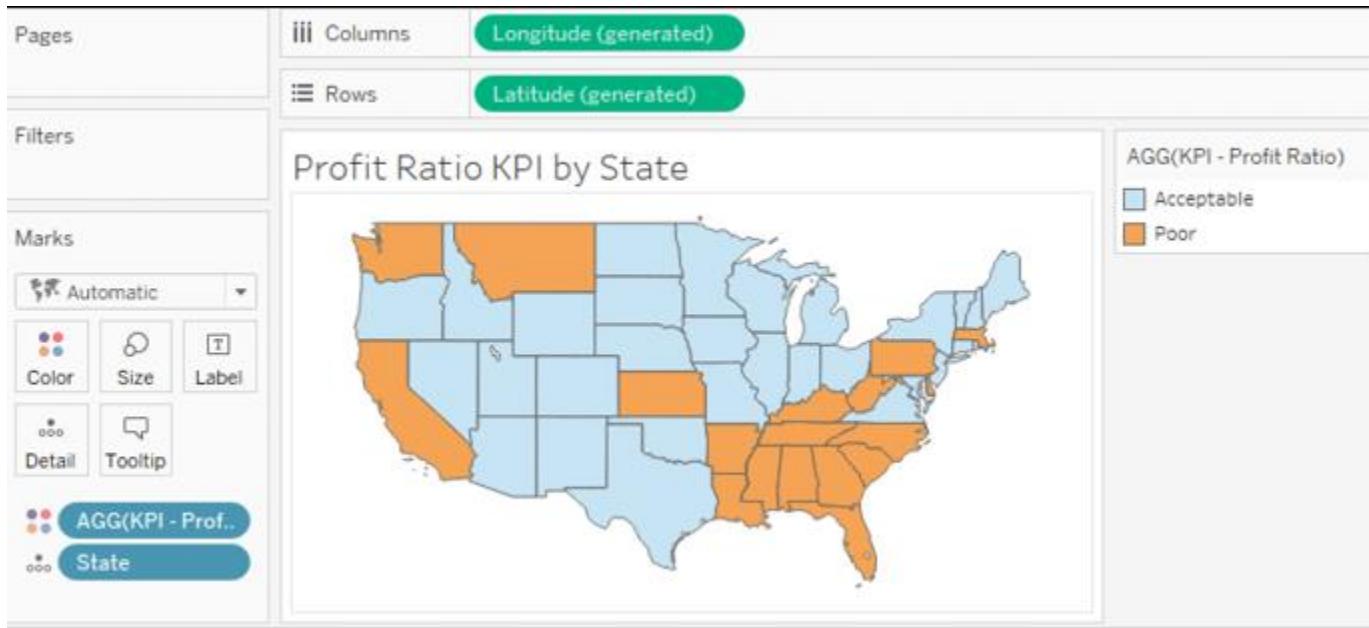


Stories

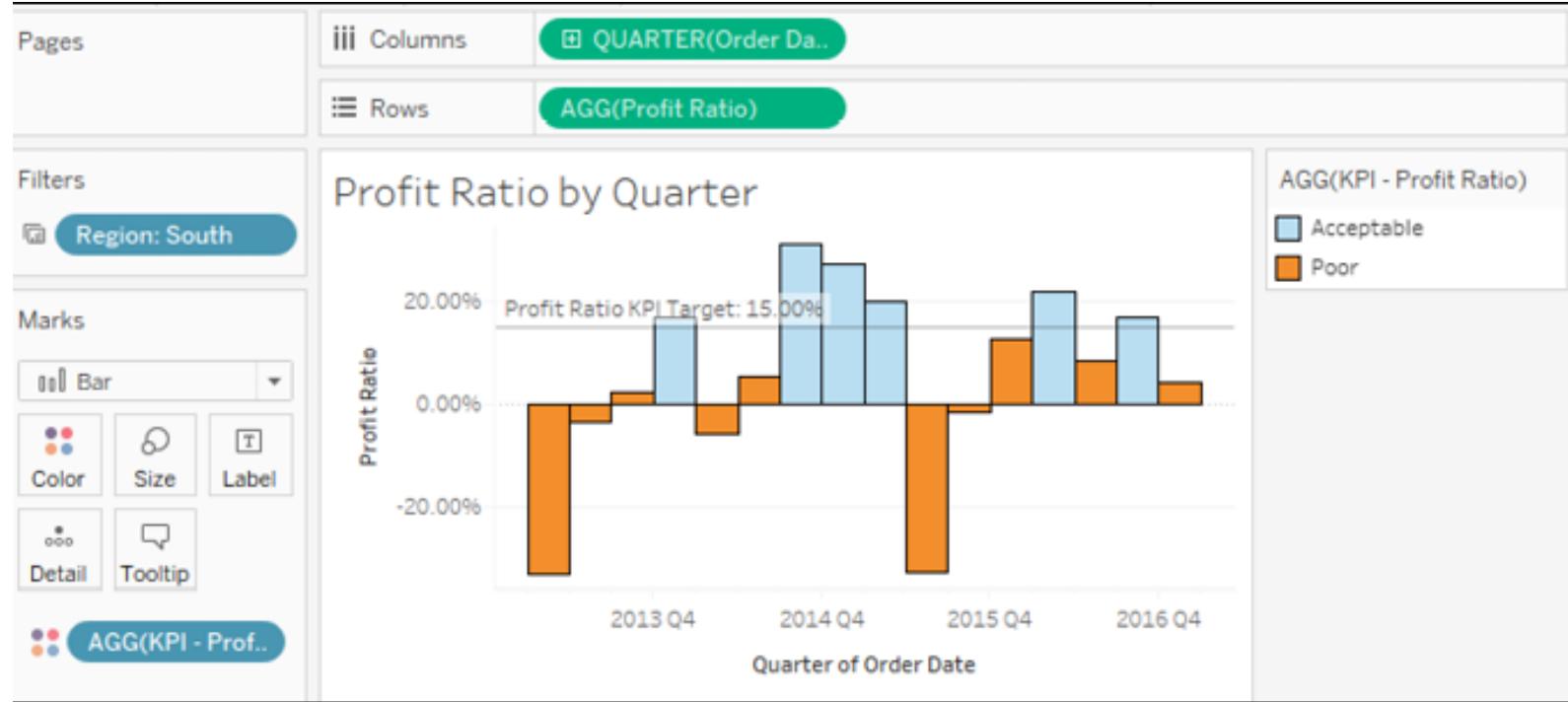
- The stories feature allows you to tell a story using interactive snapshots of dashboards and views, The snapshots become points in a story.
- This allows you to construct a guided narrative or even an entire presentation.
- Let's consider an example in which story points might be useful.

Stories

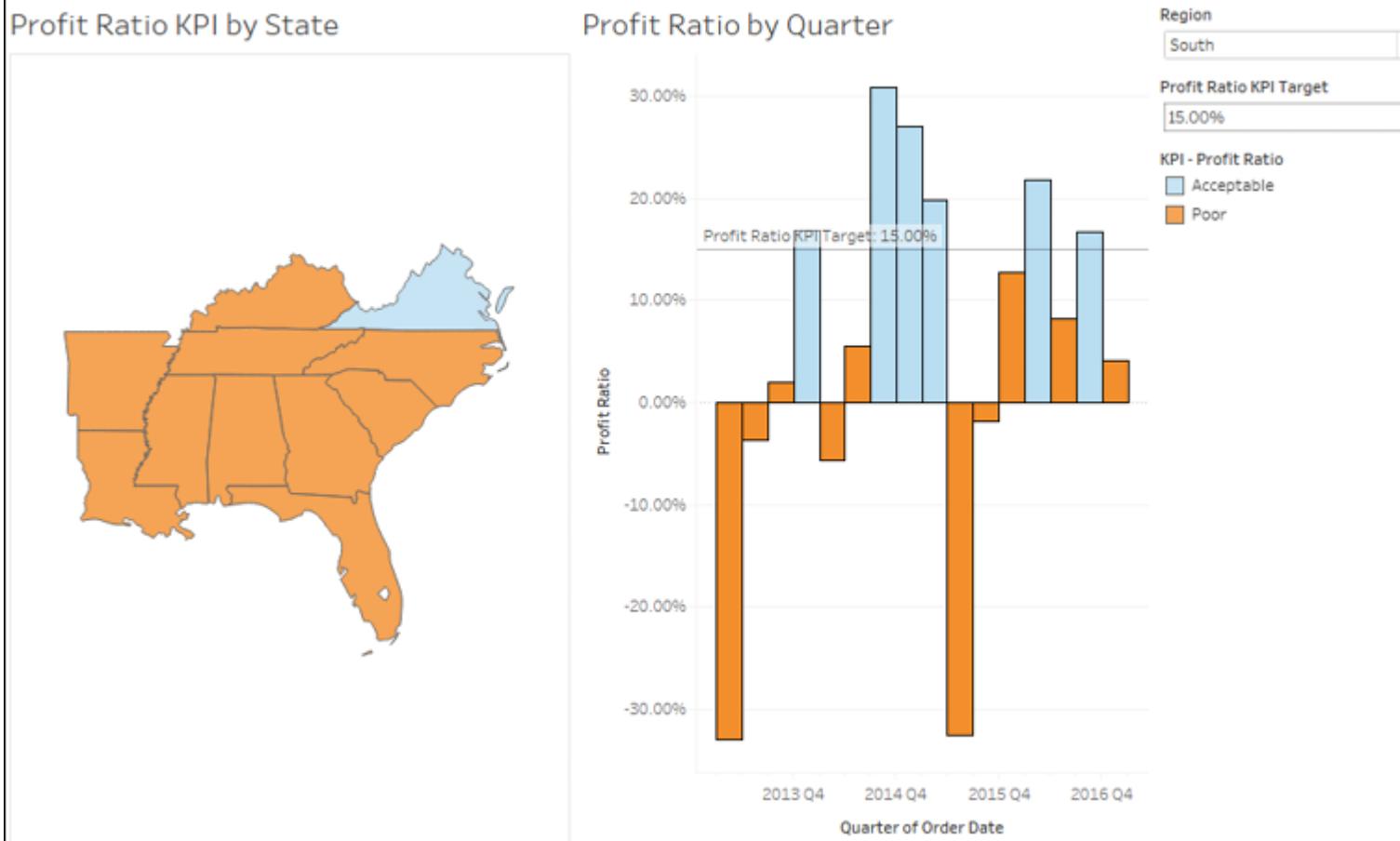
- Add Profit Ratio to the Detail part of the Marks card so that it is available for later use:



Stories



Stories



Stories

Story Title

Update X

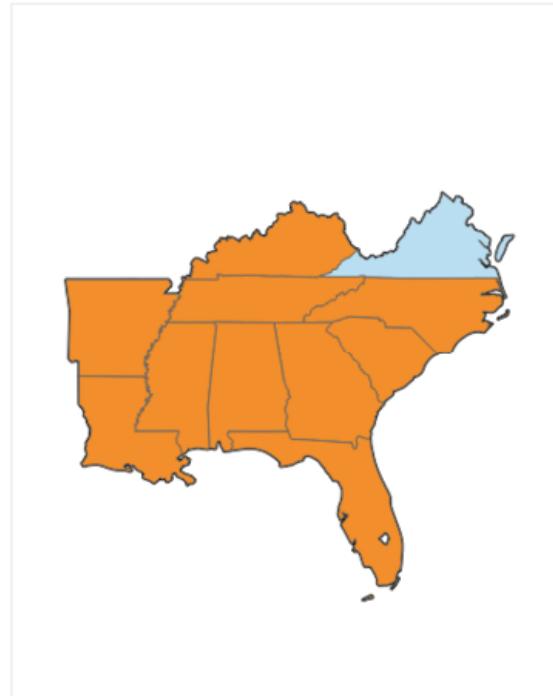
Story Point 1 Story Point 2

< >

New Blank Point

Save as a New Point

Profit Ratio KPI by State



Profit Ratio by Quarter



Stories

- Add the Regional Scorecard dashboard as the first story point. Select the South region in the map.
- Give the story point the following text: The South Region has not performed well the last 6 months:

South Region Analysis

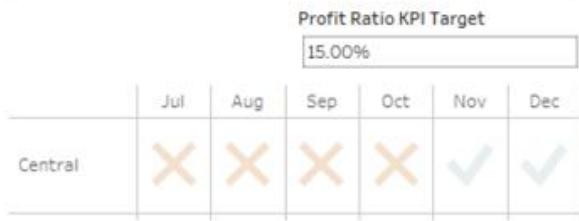
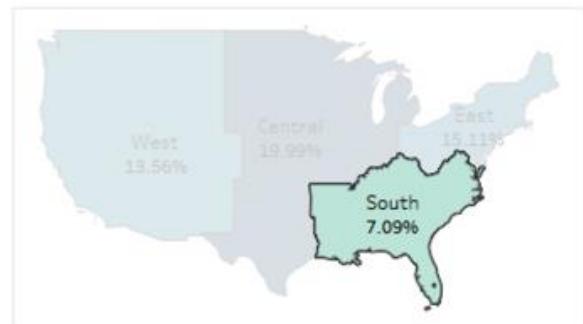
The South Region has not performed well the last 6 months.

New Blank Point

Duplicate

Profit Ratio KPI

Regional Scorecard (last 6 months)



South Region Analysis



The South Region has not performed well the last 6 months.

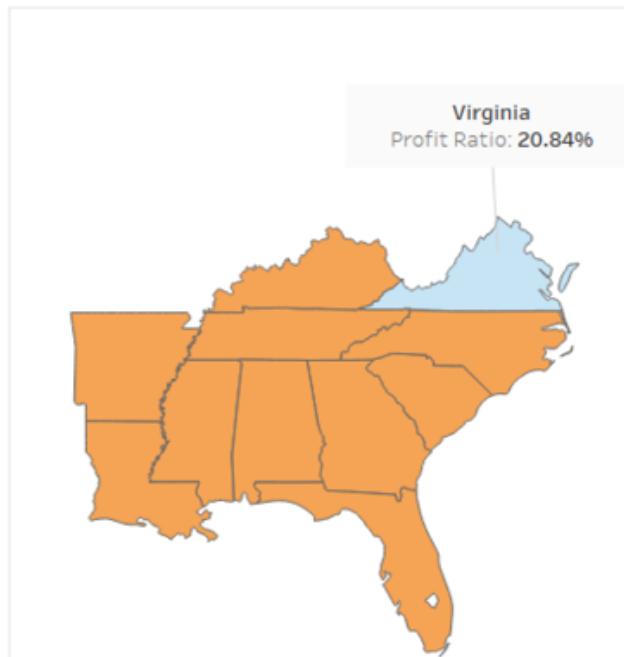


Only one state has met the 15% target overall

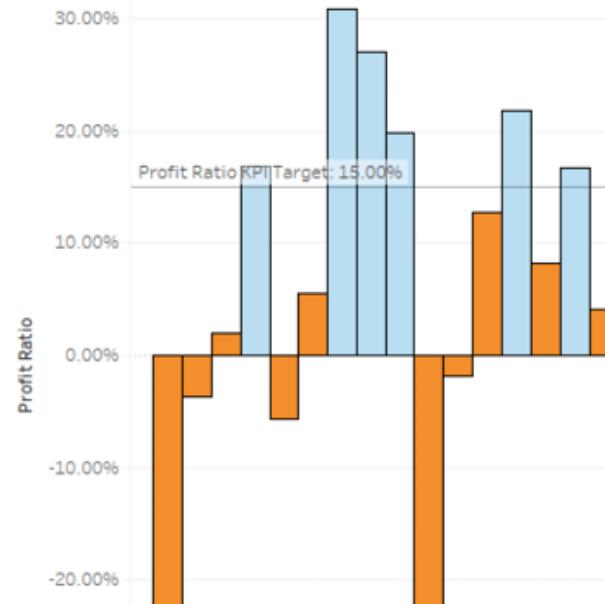
New Blank Point

Duplicate

Profit Ratio KPI by State



Profit Ratio by Quarter



South Region Analysis

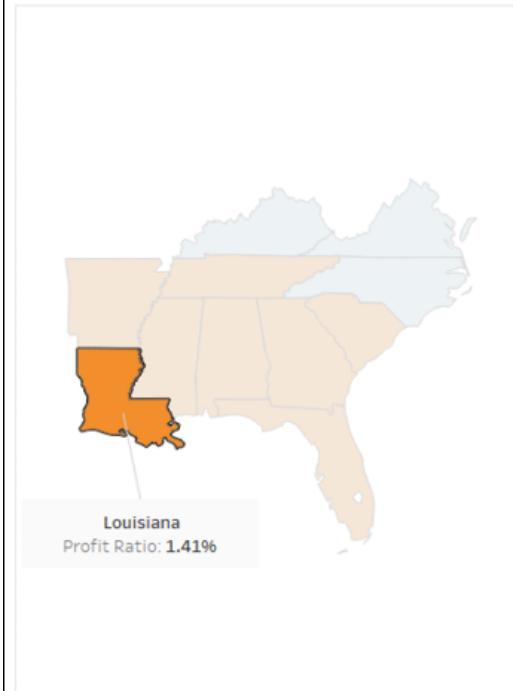
The South Region has not performed well the last 6 months.

Only one state has met the 15% target overall

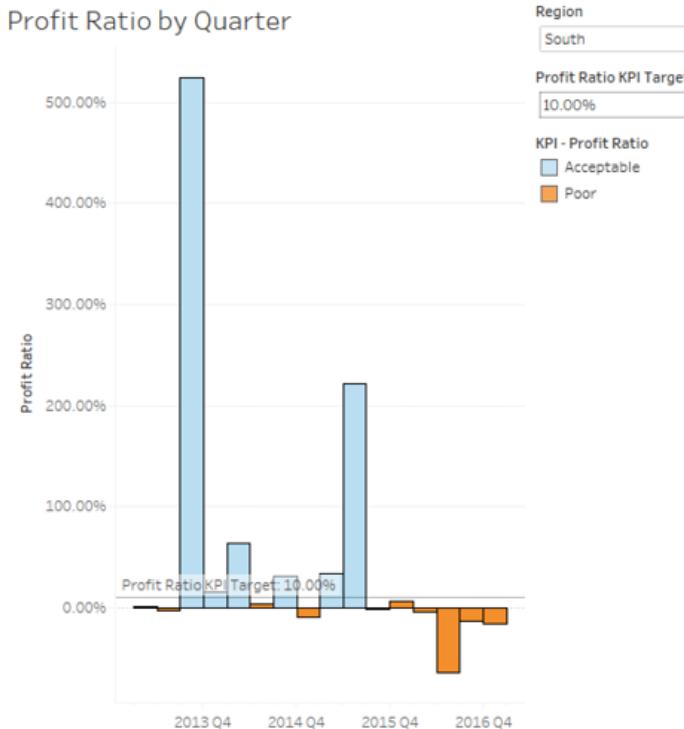
3 states would meet a goal of 10%

Certain states have performed well historically

Profit Ratio KPI by State



Profit Ratio by Quarter



Summary

- When you are ready to share your discovery and analysis, you will likely use dashboards to relate the story to your audience.
- The way in which you tell the story will depend on your objectives, as well as your audience and the mode of delivery.
- Using a combination of views, objects, parameters, filters, and legends, you can create an incredible framework for telling a data story.

COMPLETE LAB 8

9. Visual Analytics – Trends, Clustering, Distributions, and Forecasting

A blurred background image of a person's hands typing on a laptop keyboard, positioned on the left side of the slide. The right side of the slide features a solid orange diagonal band.

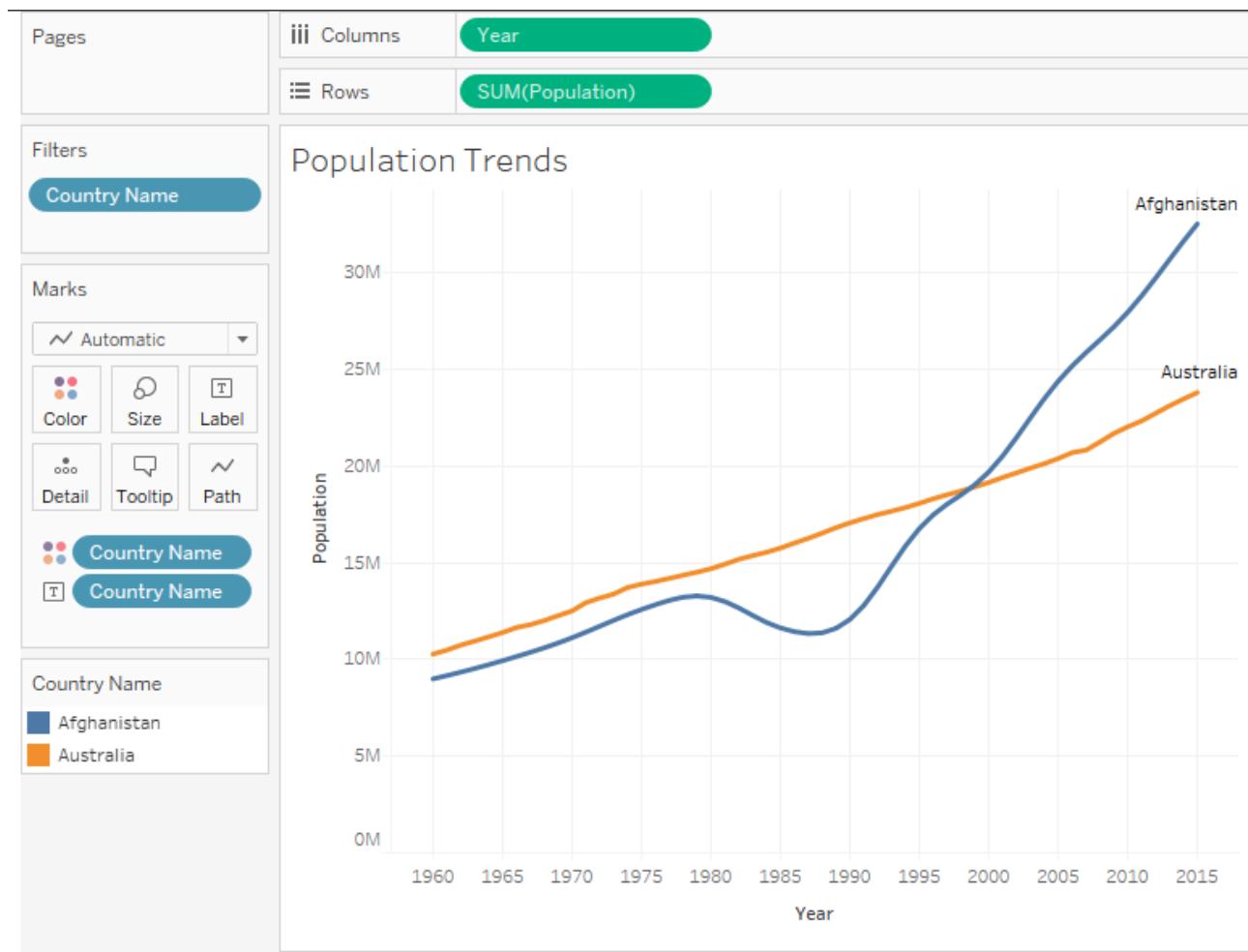
Visual Analytics – Trends, Clustering, Distributions, and Forecasting

This lesson will cover the built-in statistical models and analysis, including the following topics:

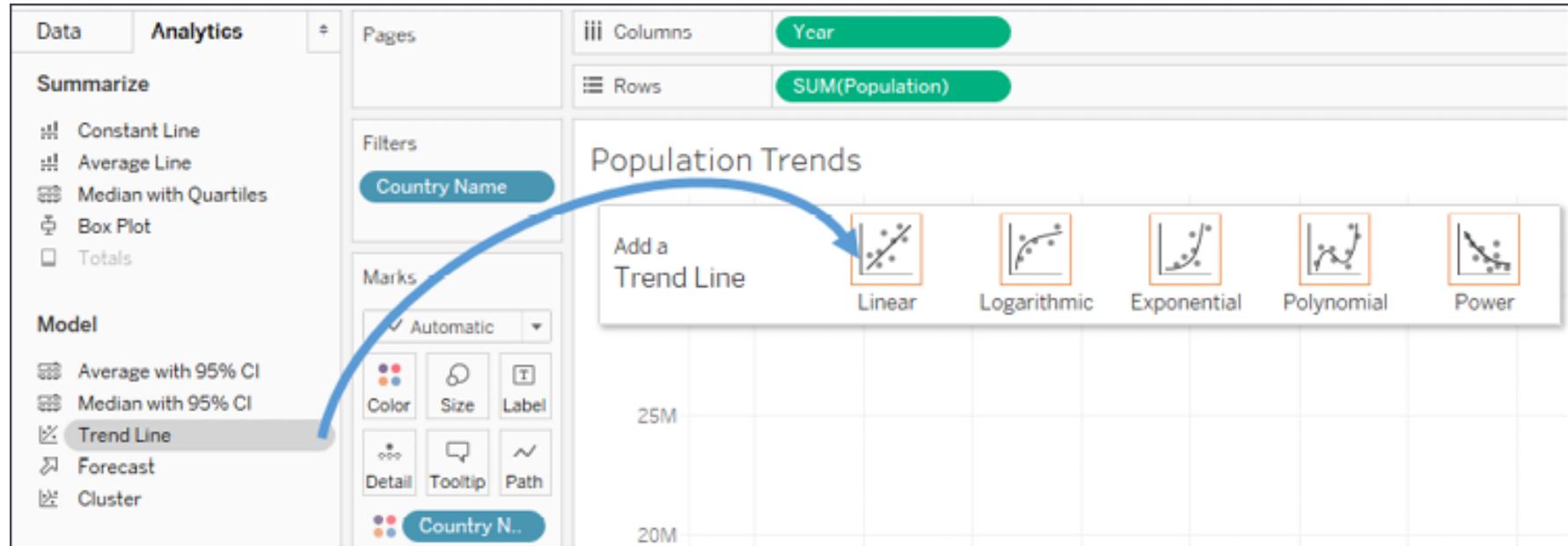
- Trends
- Clustering
- Distributions
- Forecasting

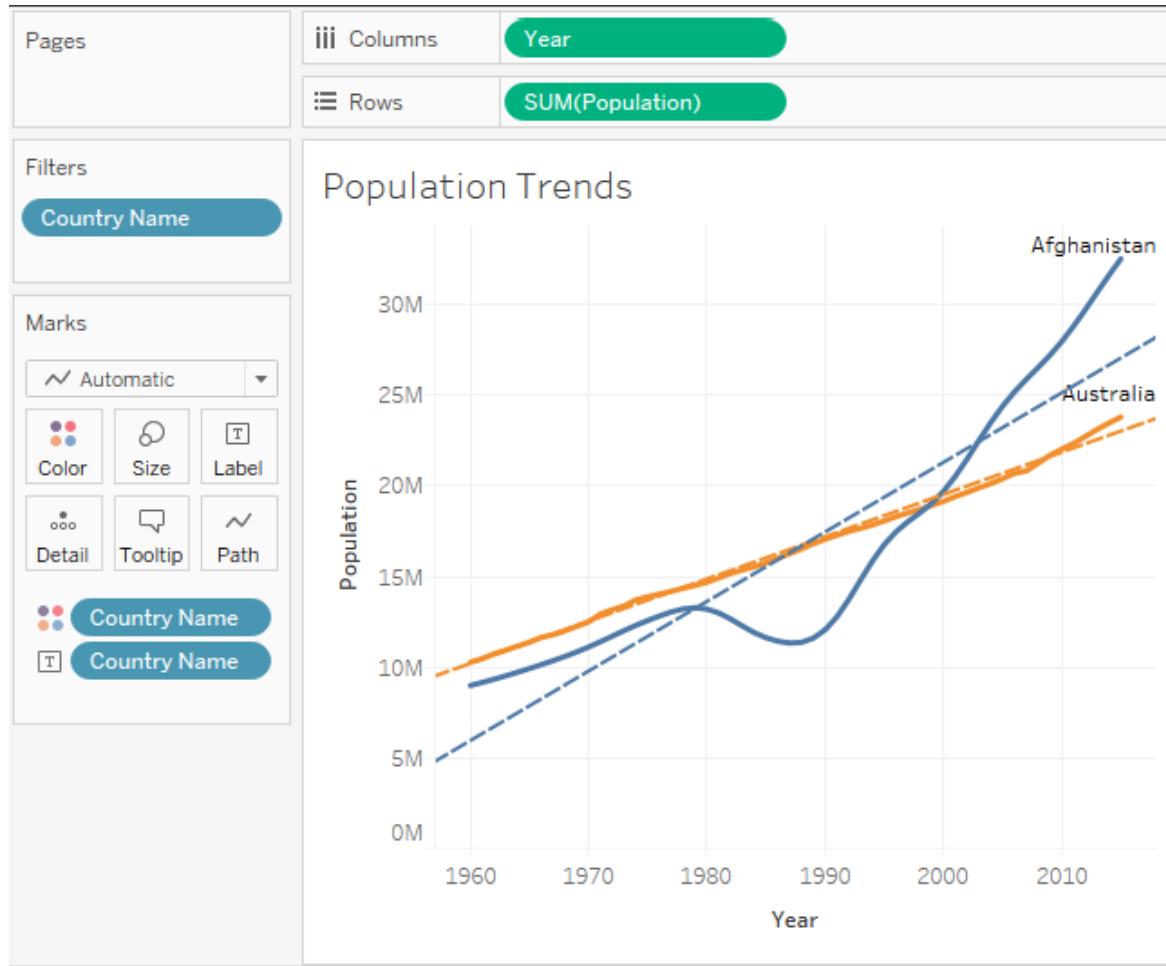
Trends

- World Population.xlsx is included in the lesson 09 directory.
- It contains one record for each country for each year from 1960 to 2015, measuring population. Using this dataset, let's take a look at the historical trends of various countries.
- Create a view similar to the one shown in the following screenshot, which shows the change in population over time for Afghanistan and Australia.

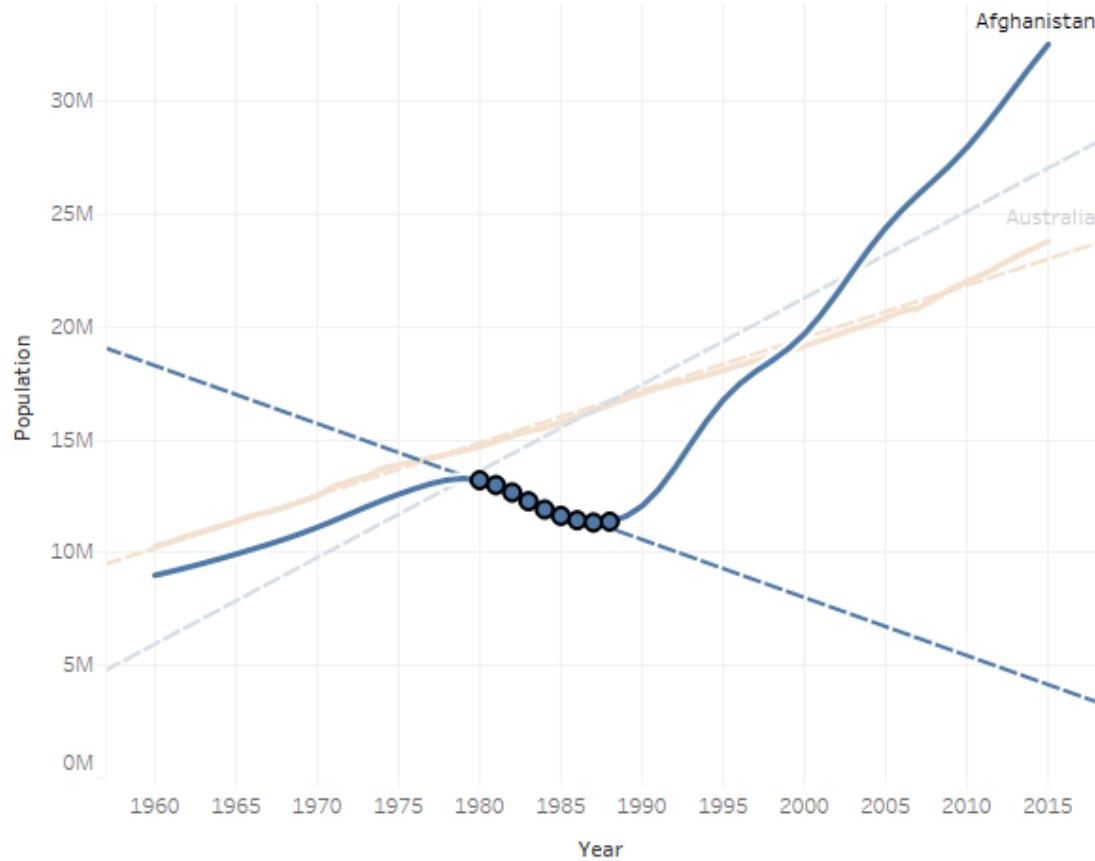


Trends





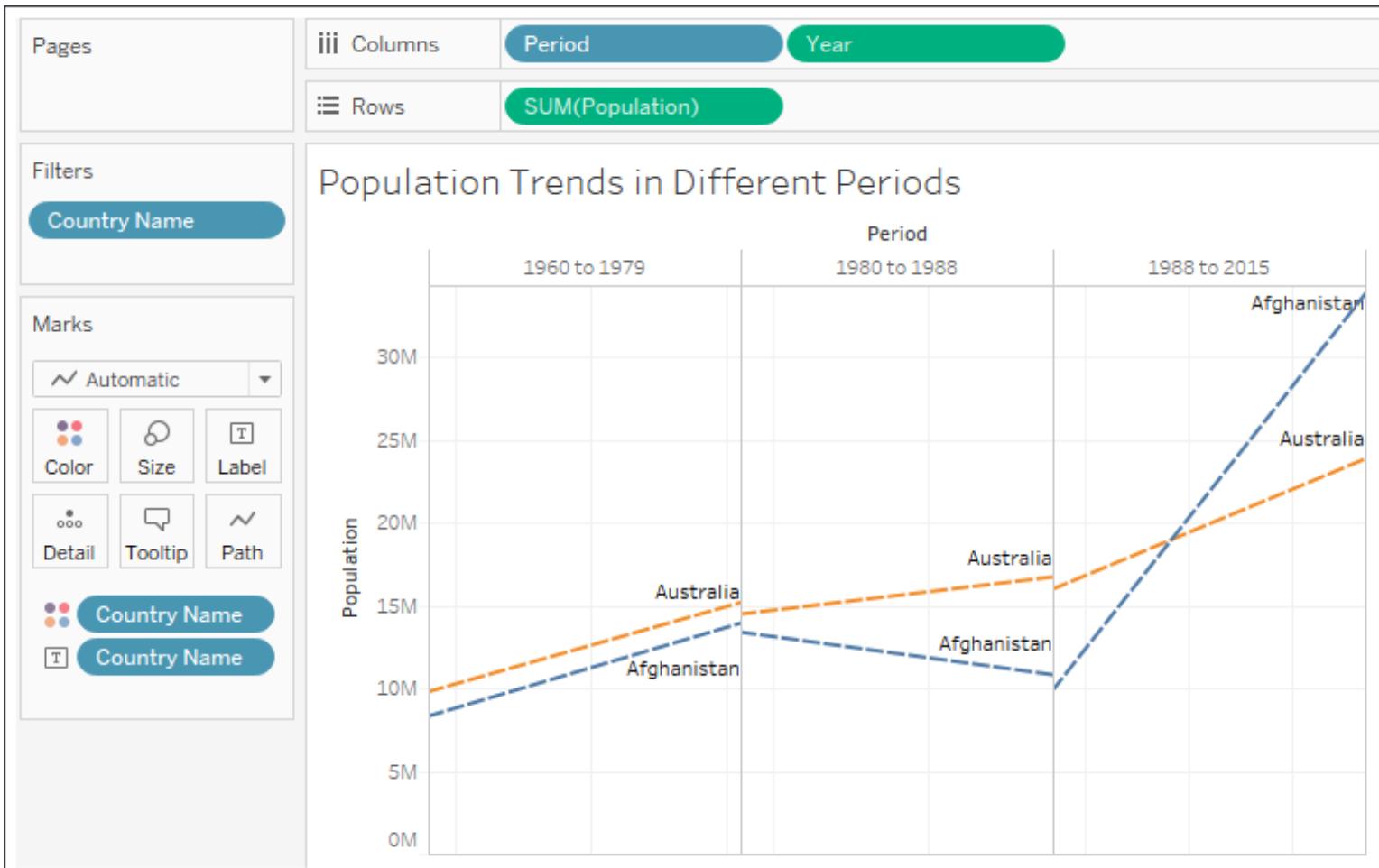
Population Trends



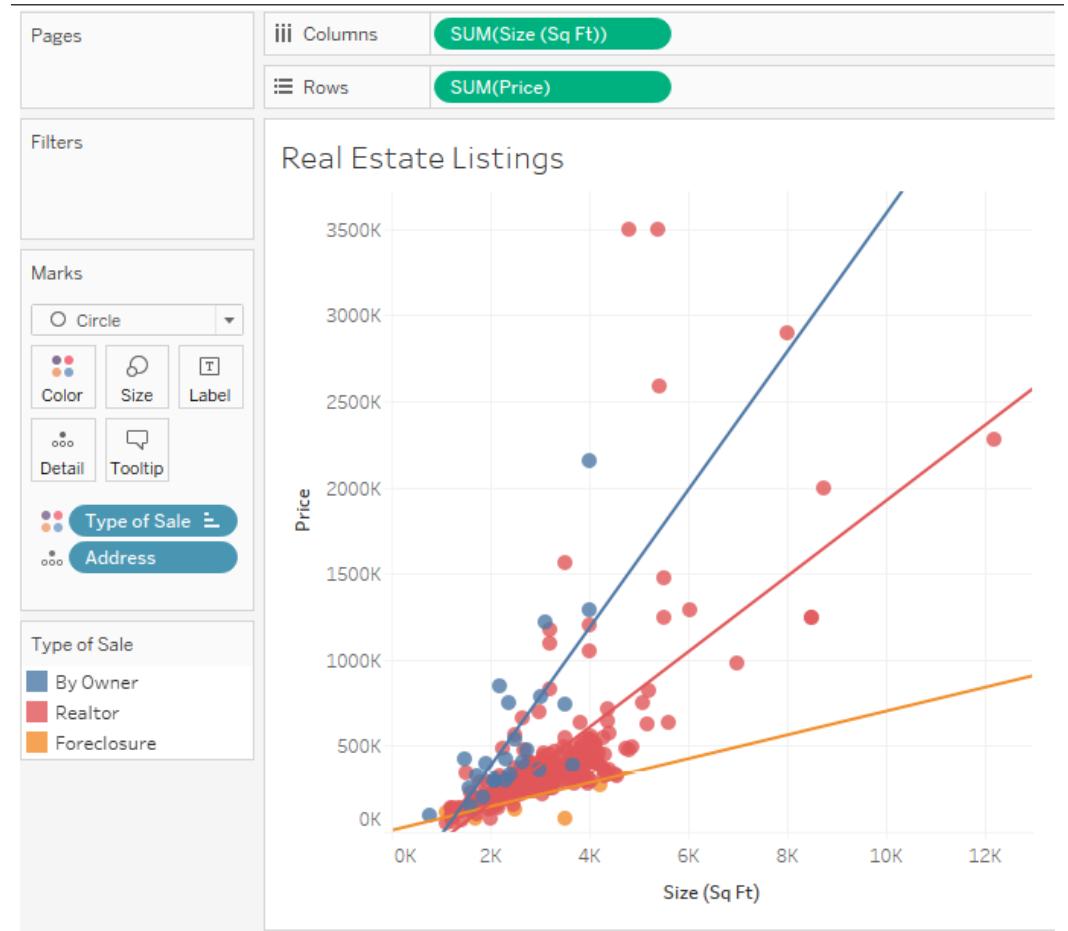
Trends

- Go ahead and create a calculated field called Period that defines discrete values for the different historical periods using code like this:

```
IF [Year] <= 1979  
    THEN "1960 to 1979"  
ELSEIF [Year] <= 1988  
    THEN "1980 to 1988"  
ELSE "1988 to 2015"  
END
```



Customizing trend lines



Trend Lines Options

X

Model type

- Linear
- Logarithmic
- Exponential
- Power
- Polynomial, Degree:

Options

Include the following fields as factors:

Type of Sale

Allow a trend line per color

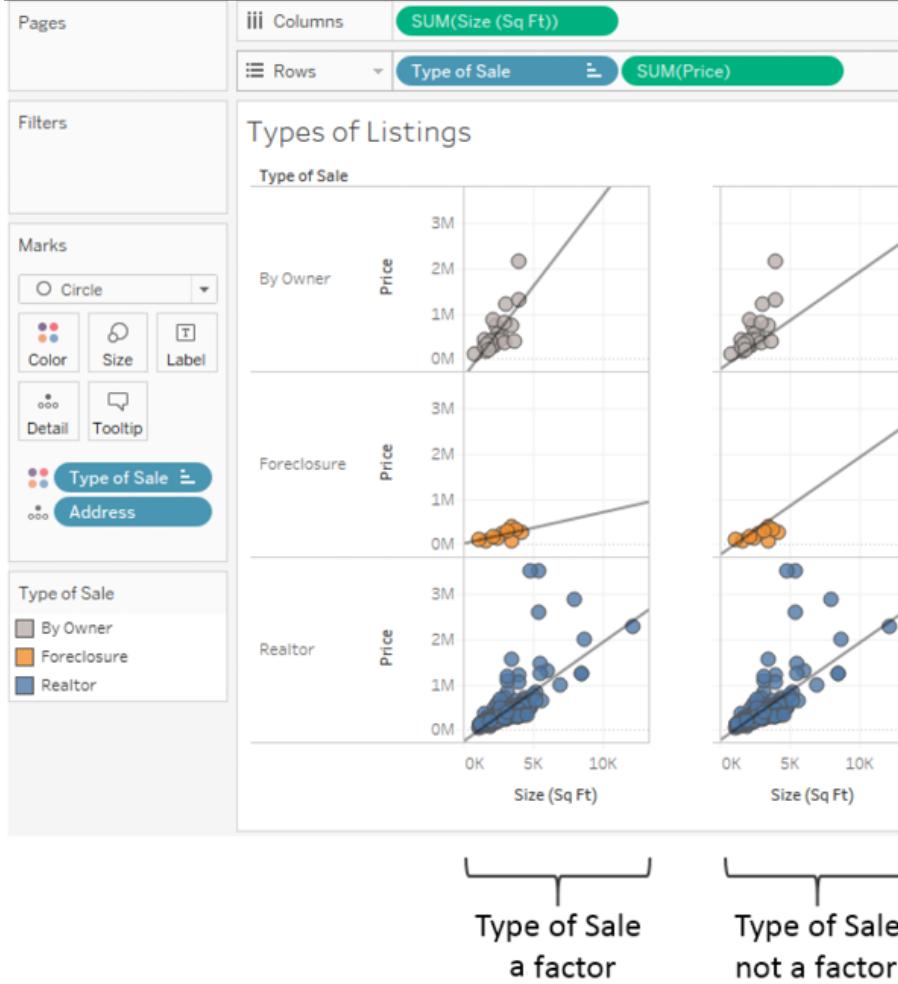
Show Confidence Bands

Force y-intercept to zero

Show recalculated line for highlighted or selected data points

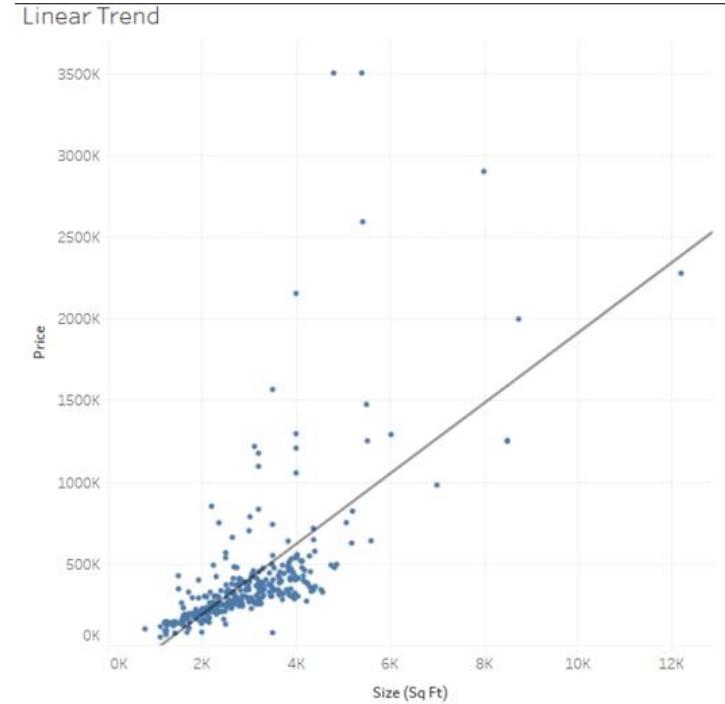
Show tooltips

OK



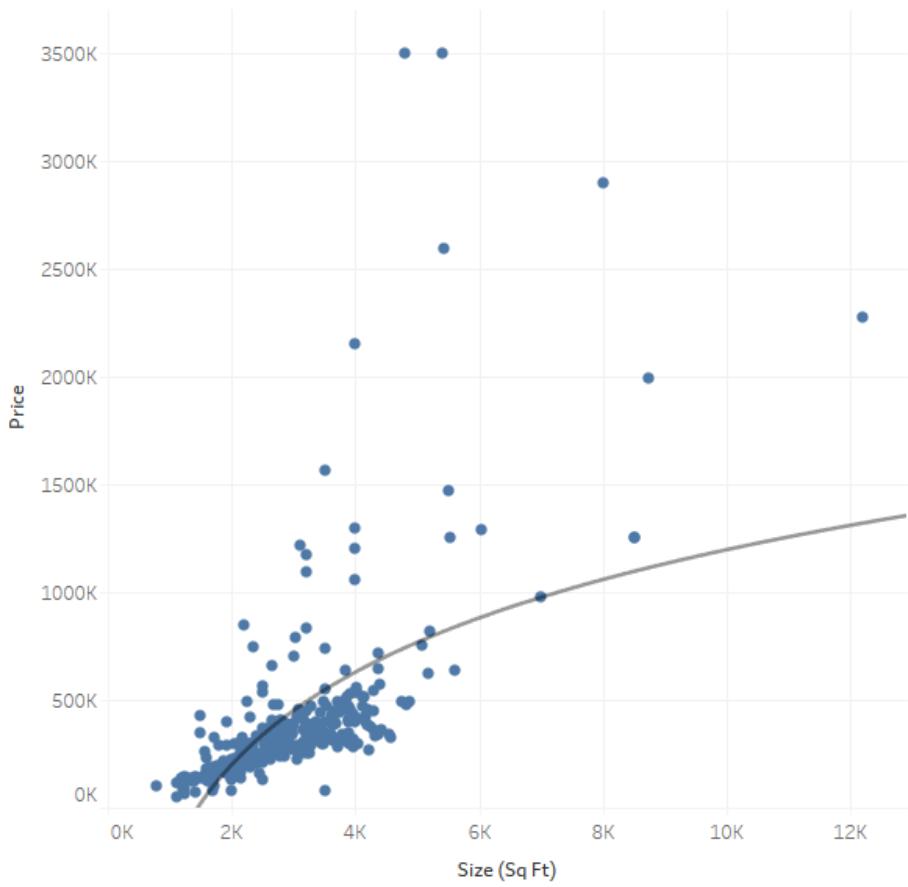
Trend models(Linear)

- We'd use a linear model if we assumed that, as Size increases, Price will increase at a constant rate.
- No matter how much Size increased, we'd expect Price to increase so that new data points fall close to the straight line:

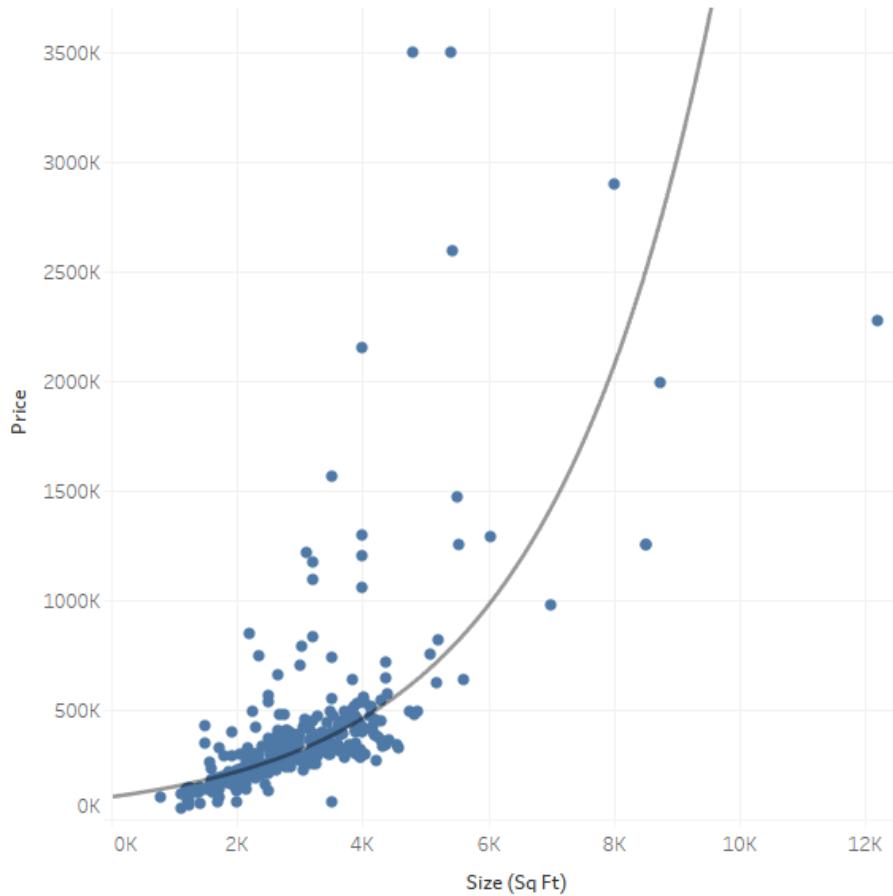


Logarithmic

Logarithmic Trend



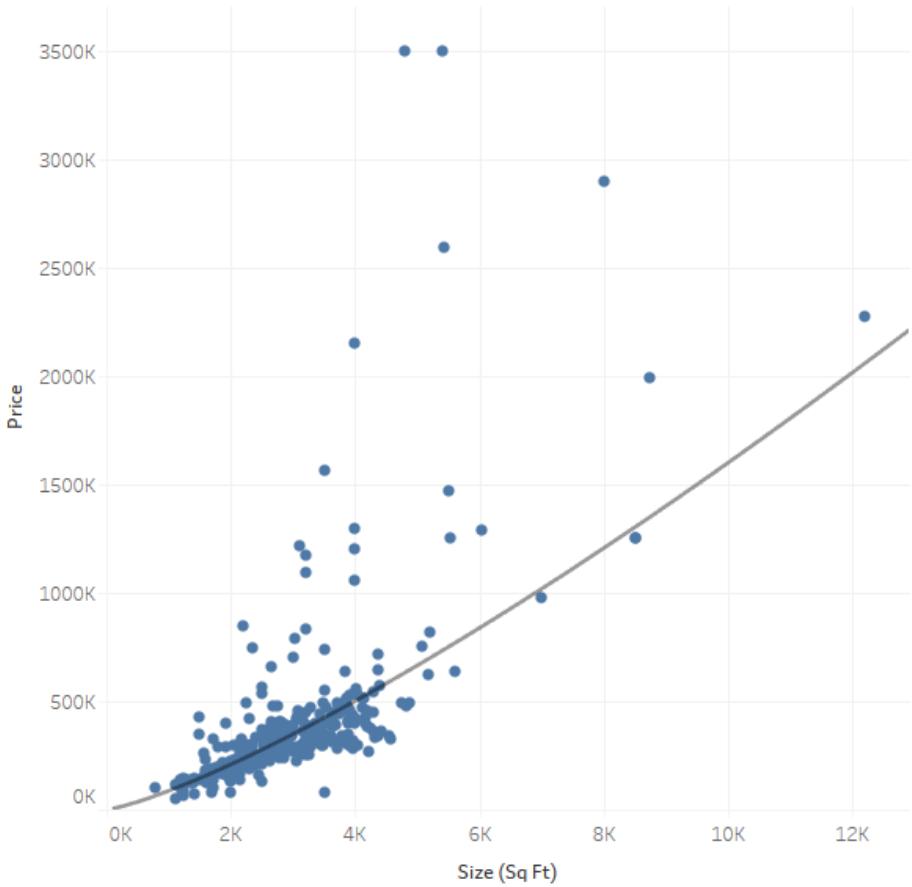
Exponential Trend



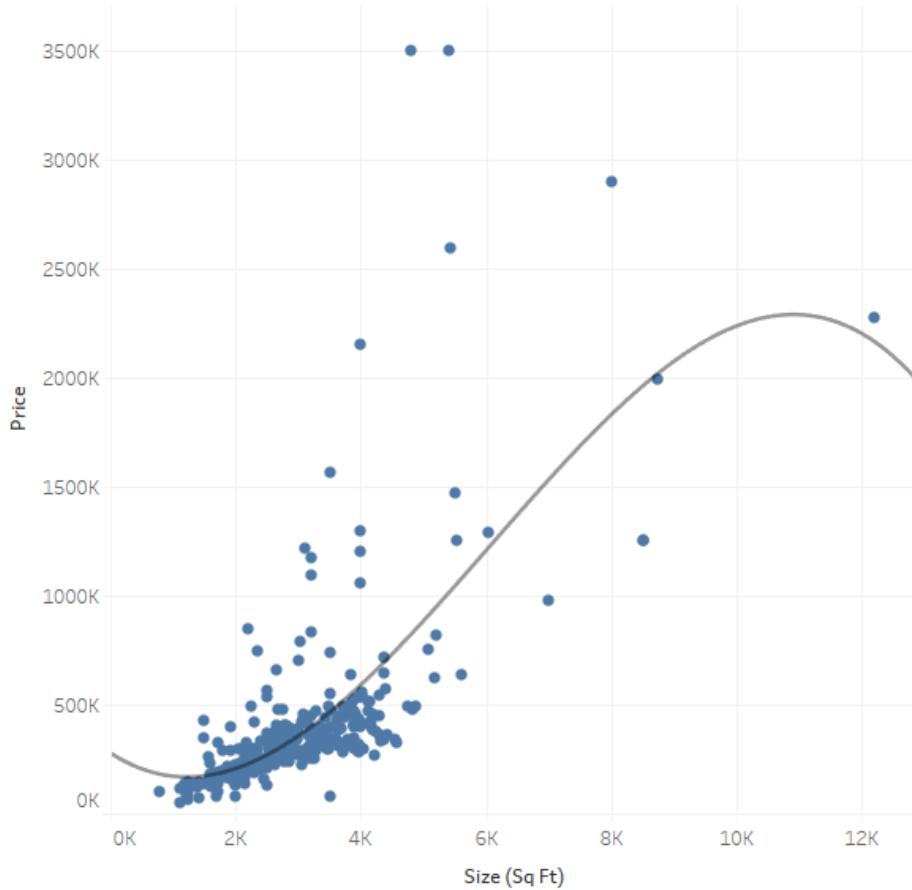
Exponential

Power

Power Trend

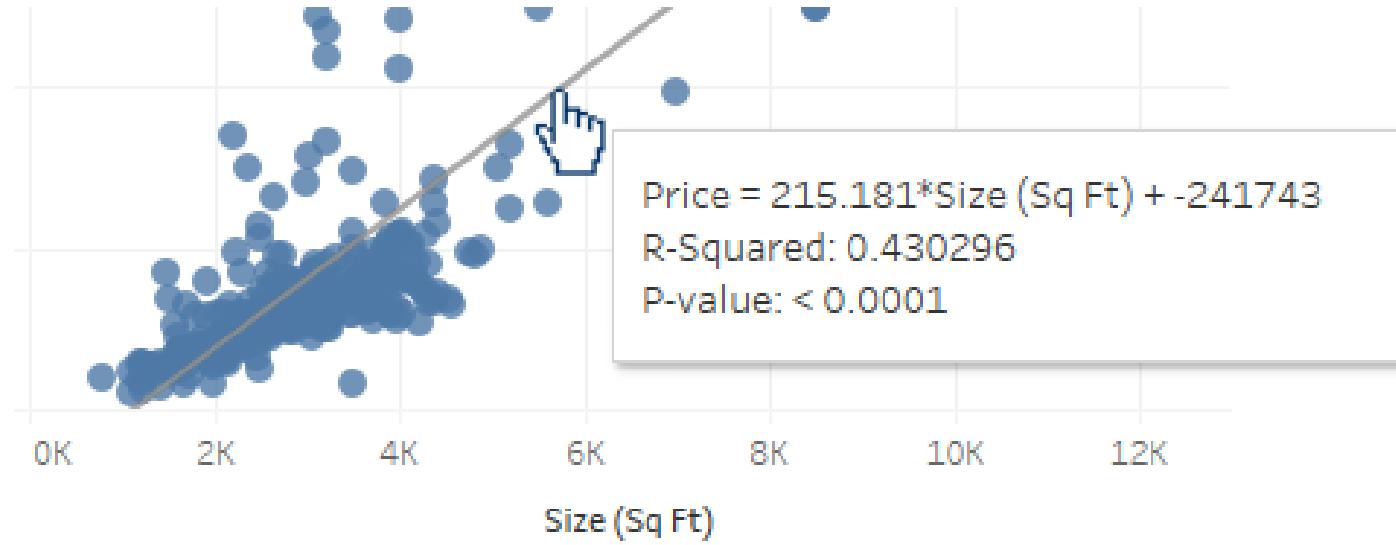


3rd Degree Polynomial Trend



Polynomial

Analyzing trend models



**Trend Lines Model**

A linear trend model is computed for sum of Price given sum of Size (Sq Ft). The model may be significant at $p \leq 0.05$. The factor Type of Sale may be significant at $p \leq 0.05$.

Model formula: Type of Sale*(Size (Sq Ft) + intercept)
Number of modeled observations: 336
Number of filtered observations: 0
Model degrees of freedom: 6
Residual degrees of freedom (DF): 330
SSE (sum squared error): 2.99019e+13
MSE (mean squared error): 9.06118e+10
R-Squared: 0.480548
Standard error: 301018
p-value (significance): < 0.0001

Analysis of Variance:

Field	DF	SSE	MSE	F	p-value
Type of Sale	4	2.8926758e+12	7.23169e+11	7.98095	< 0.0001

Individual trend lines:

Panes	Color	Line	Coefficients						
Row	Column	Type of Sale	p-value	DF	Term	Value	StdErr	t-value	p-value
Price	Size (Sq Ft)	Realtor	< 0.0001	300	Size (Sq Ft)	219.042	13.5019	16.2231	< 0.0001
					intercept	-269860	44427	-6.07422	< 0.0001
Price	Size (Sq Ft)	Foresclosure	0.0455421	8	Size (Sq Ft)	69.1147	29.2138	2.36583	0.0455421
					intercept	9557.19	87520	0.1092	0.915733
Price	Size (Sq Ft)	By Owner	< 0.0001	22	Size (Sq Ft)	400.666	82.8705	4.83484	< 0.0001
					intercept	-421058	213628	-1.97099	0.0614346

Copy**Close**

Analyzing trend models

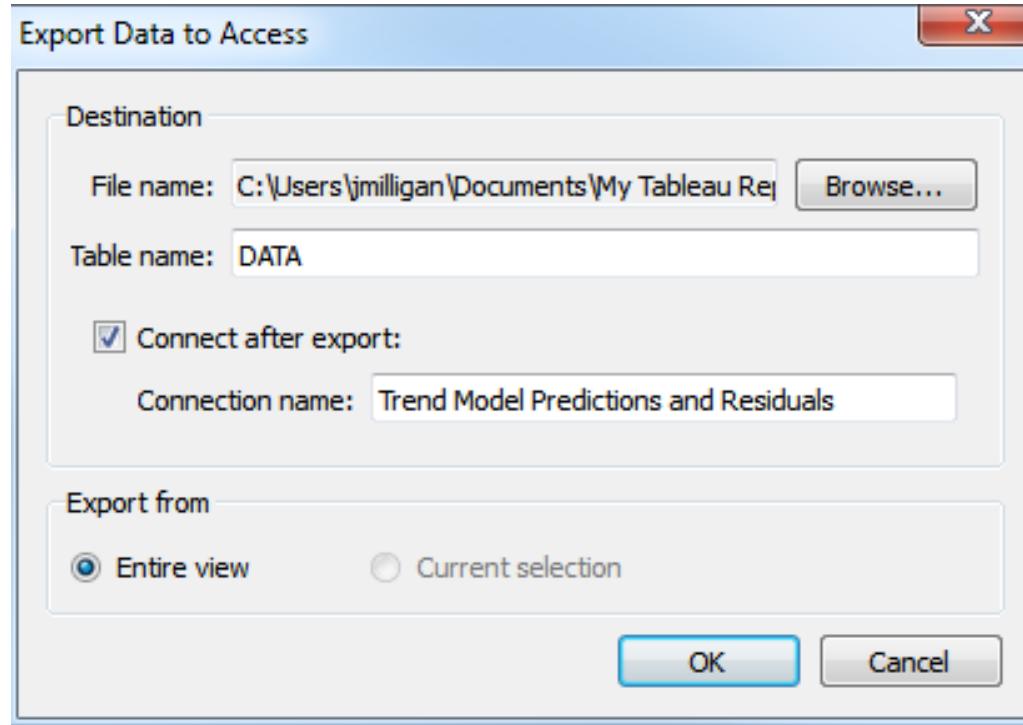
- Additional summary statistical information can be displayed in Tableau Desktop for a given view by showing a summary.
- From the menu, select Worksheet | Show Summary.
- The information displayed in the summary can be expanded using the drop-down menu on the Summary card:

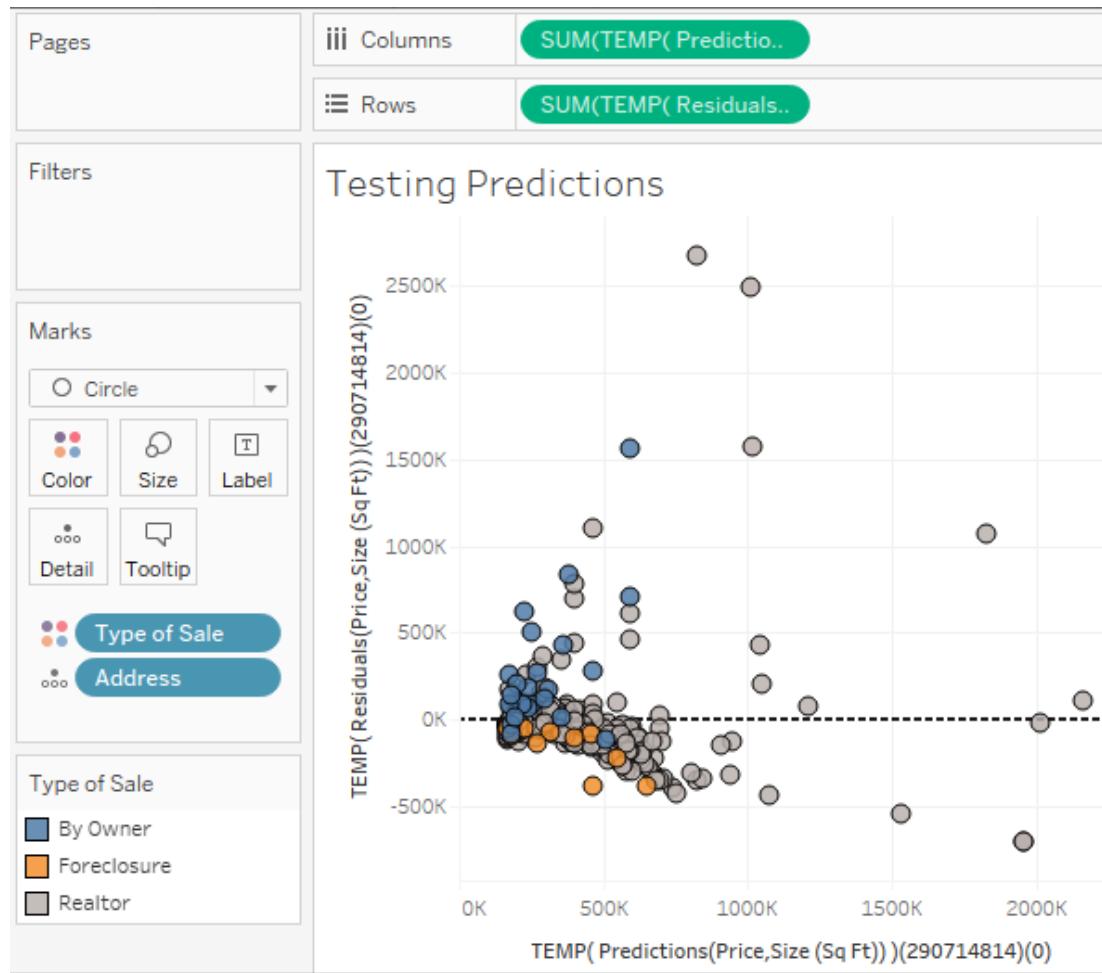
Summary		
Count:	336	
SUM(Price)		
Sum:	134,239,384	
Average:	399,521.98	
Minimum:	50,000	
Maximum:	3,500,000	
Median:	300,000.00	
Standard deviation:	414,528	
First quartile:	211,500.00	
Third quartile:	410,000.00	
Skewness:	4.59	
Excess Kurtosis:	25.74	
SUM(Size (Sq Ft))		
Sum:	1,001,318	
Average:	2,980.11	
Minimum:	784	
Maximum:	12,200	
Median:	2,800.50	
Standard deviation:	1,264	
First quartile:	2,108.50	
Third quartile:	3,627.00	
Skewness:	2.27	
Excess Kurtosis:	10.95	

Exporting statistical model details

- Tableau also gives you the ability to export data, including data related to trend models.
- This allows you to, more deeply—and even visually, analyze the trend model itself.
- Let's analyze the third-degree polynomial trend line of the real estate price and size scatterplot without any factors.
- To export data related to the current view, use the menu to select Worksheet | Export | Data.

Exporting statistical model details



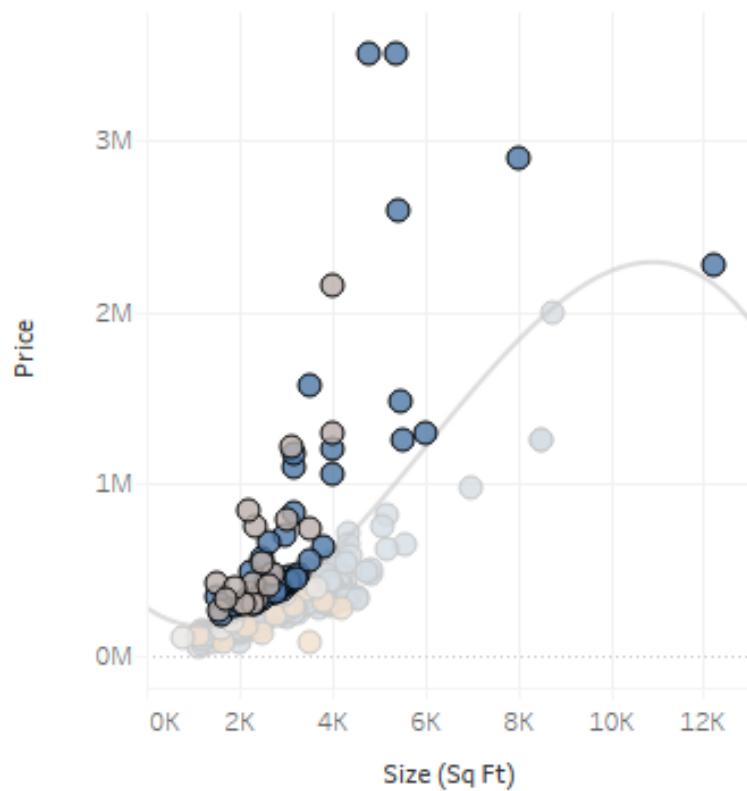


Exporting statistical model details

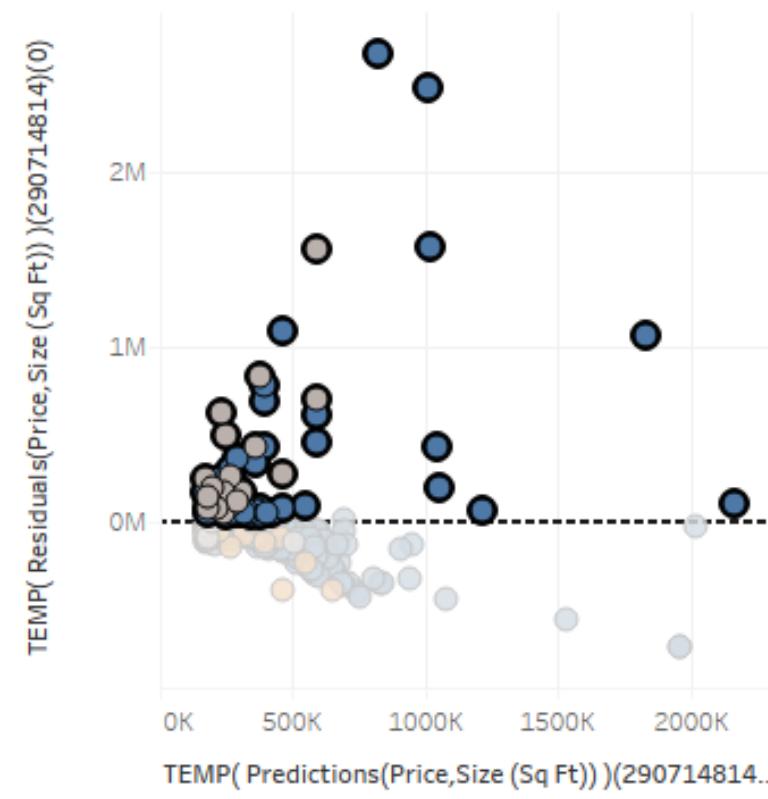
- You can include this new view along with the original on a dashboard to explore the trend model visually.
- Use the highlight button on the toolbar to highlight the Address field:



Trend Model to Test



Testing Predictions



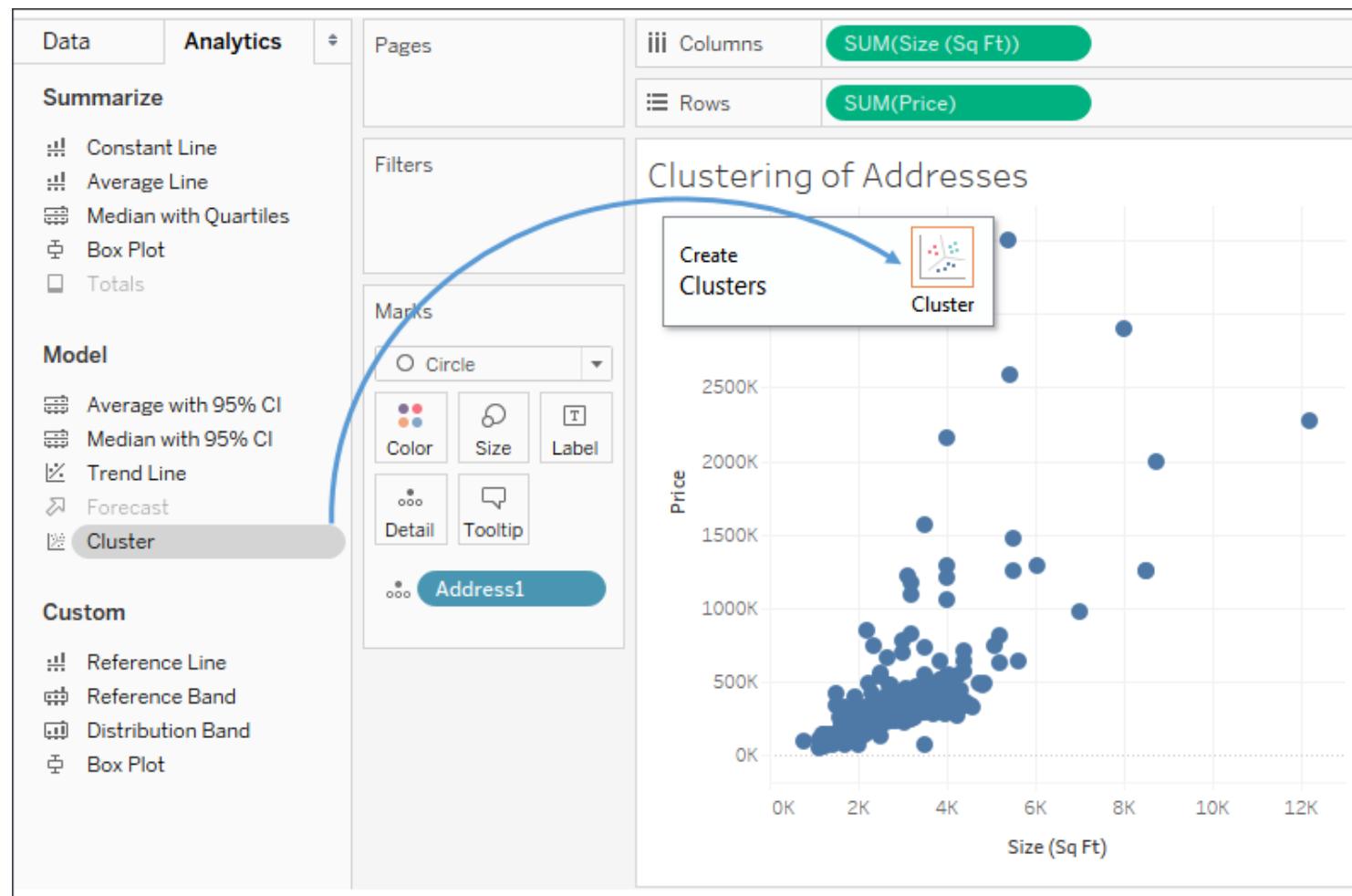
Advanced statistics and more with R and Python

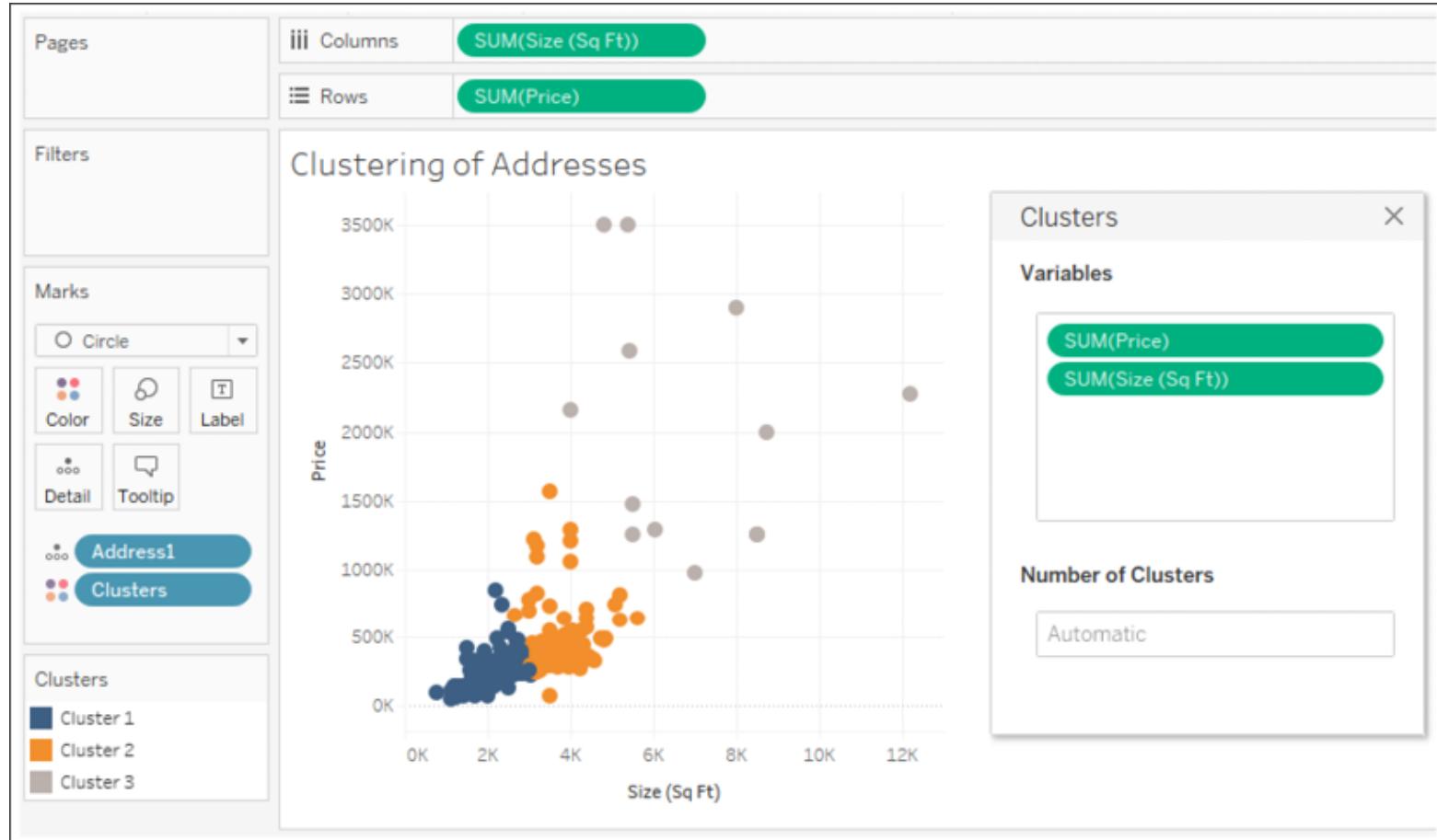
- You can achieve even more sophisticated statistical analysis leveraging Tableau's ability to integrate with R or Python.
- R is an open source statistical analysis platform and programming language with which you can define advanced statistical models.
- Python is a high-level programming language that has quickly gained a wide following among data analysts and data scientists for its ease of use.

Clustering

Tableau gives you the ability to quickly perform clustering analysis in your visualizations. This allows you to find groups, or clusters, of individual data points that are similar based on any number of your choosing. This can be useful in many different industries and fields of study, as in the following example:

- Marketing may find it useful to determine groups of customers related to each other based on spending amounts, frequency of purchases, or times and days of orders.





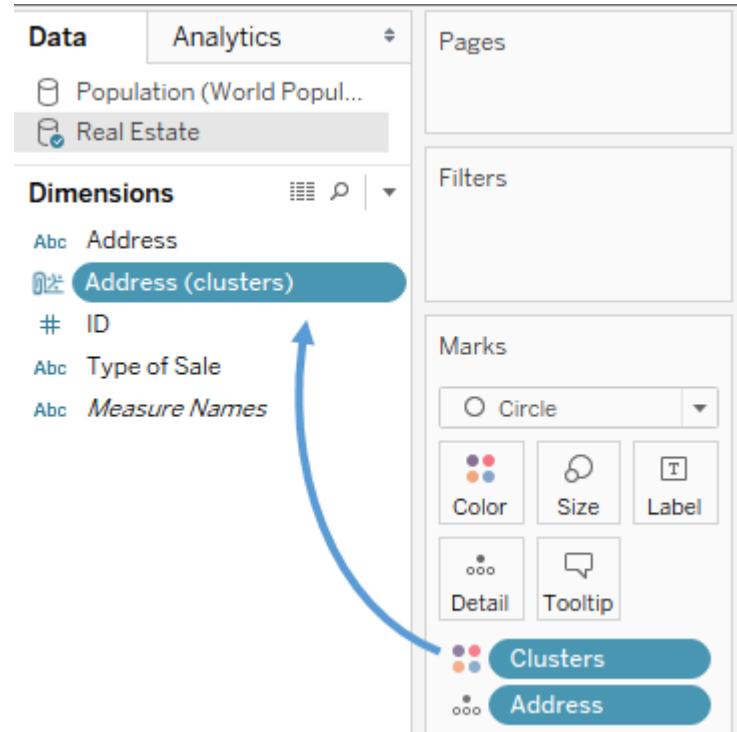
Clustering

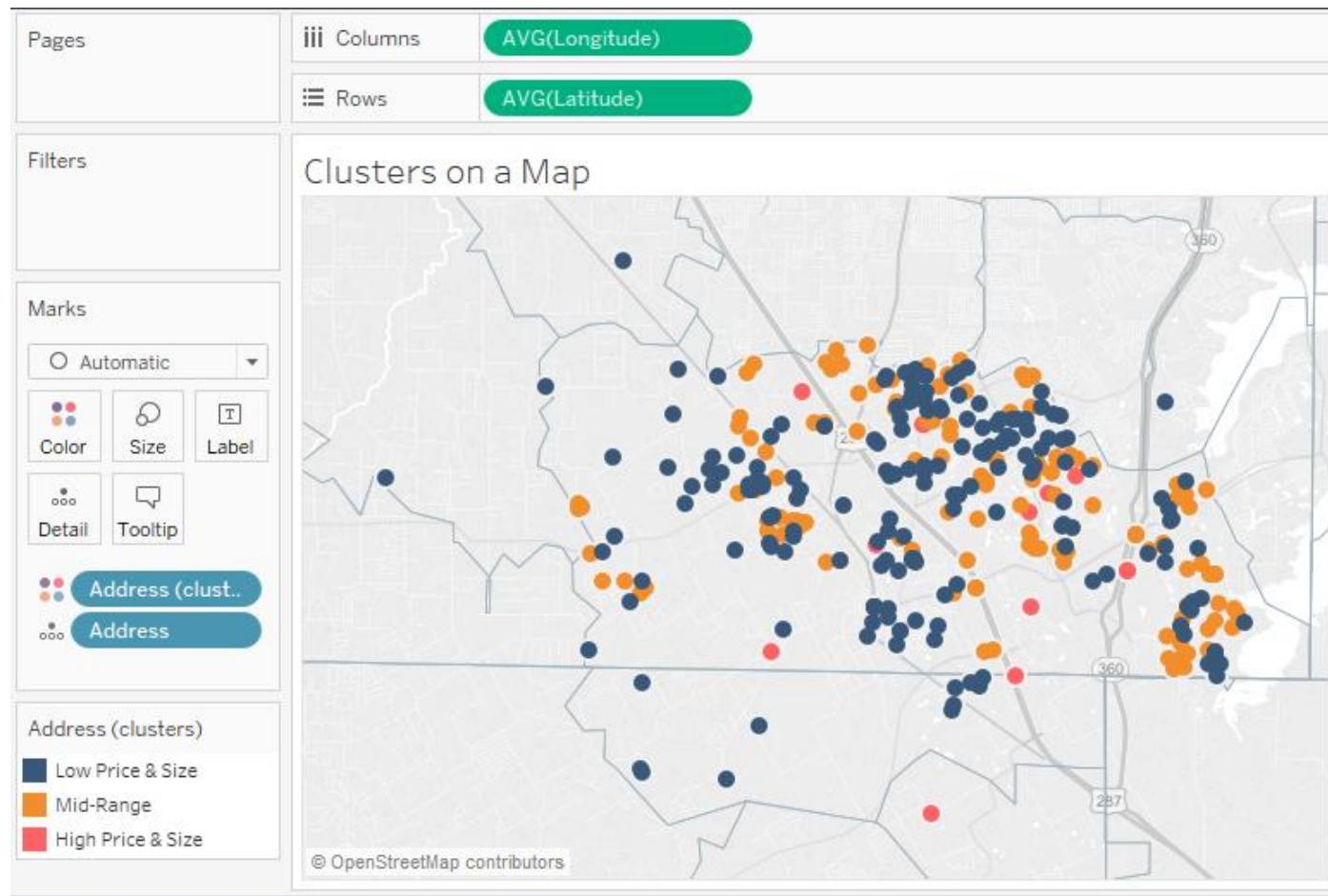
Variables are all the factors that the clustering algorithm uses to determine related data points. Number of Clusters determines into how many groups the data is partitioned. In the preceding view, you'll observe three clusters of houses:

- Those with a low price and a smaller size (blue)
- Those with an average price and size (orange)
- Those with a high price and a large size (gray)

Clustering

- Once you have meaningful clusters, you can materialize the clusters as groups in the data source by dragging them from the view and dropping them into the Data pane:



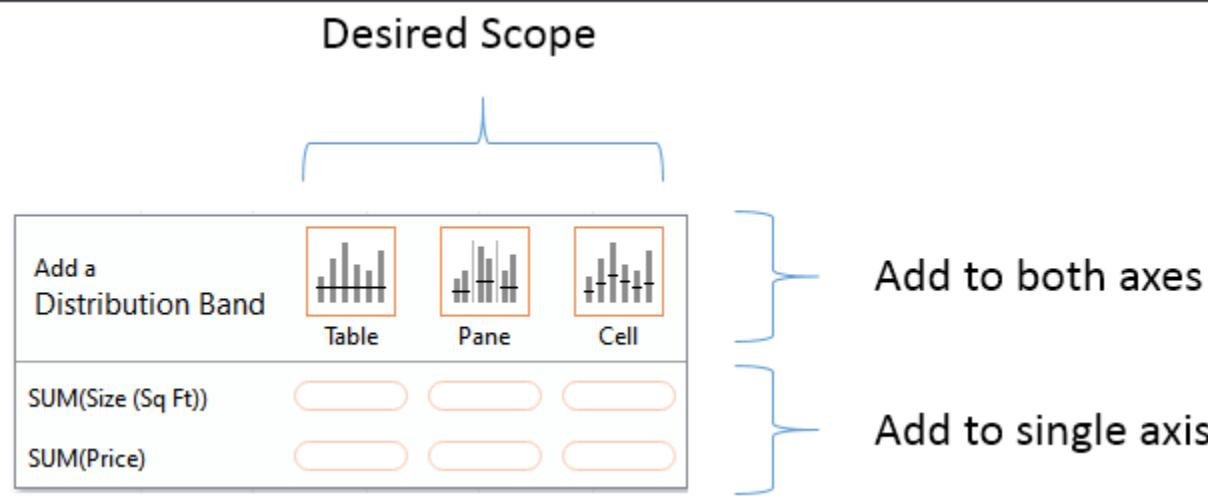


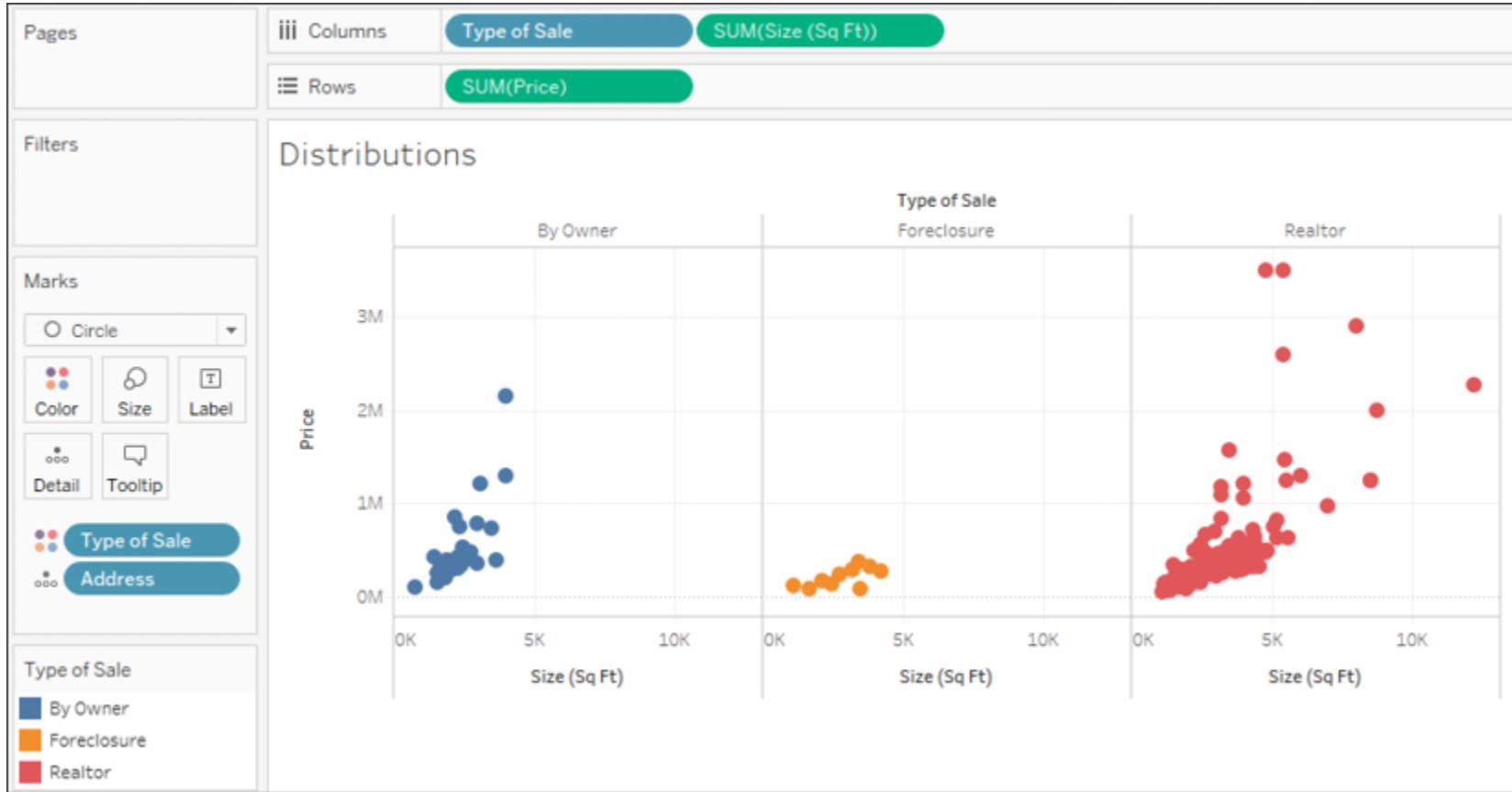
Distributions

- Analyzing distributions can be quite useful.
- We've already seen that certain calculations are available for determining statistical information such as averages, percentiles, and standard deviations.
- Tableau also makes it easy to quickly visualize various distributions, including confidence intervals, percentages, percentiles, quantiles, and standard deviations.

Distributions

- In the following example, we've dragged and dropped Distribution Band from the Analytics pane onto the scope of Pane for the axis defined by Sum(Price):





Edit Reference Line, Band, or Box

X



Line



Band



Distribution



Box Plot

Scope

Entire Table Per Pane Per Cell

Computation

Value: -2,-1,1,2 Standard Deviation

Label: None

Formatting

Line:

Fill Above

Fill: [Color Swatch]

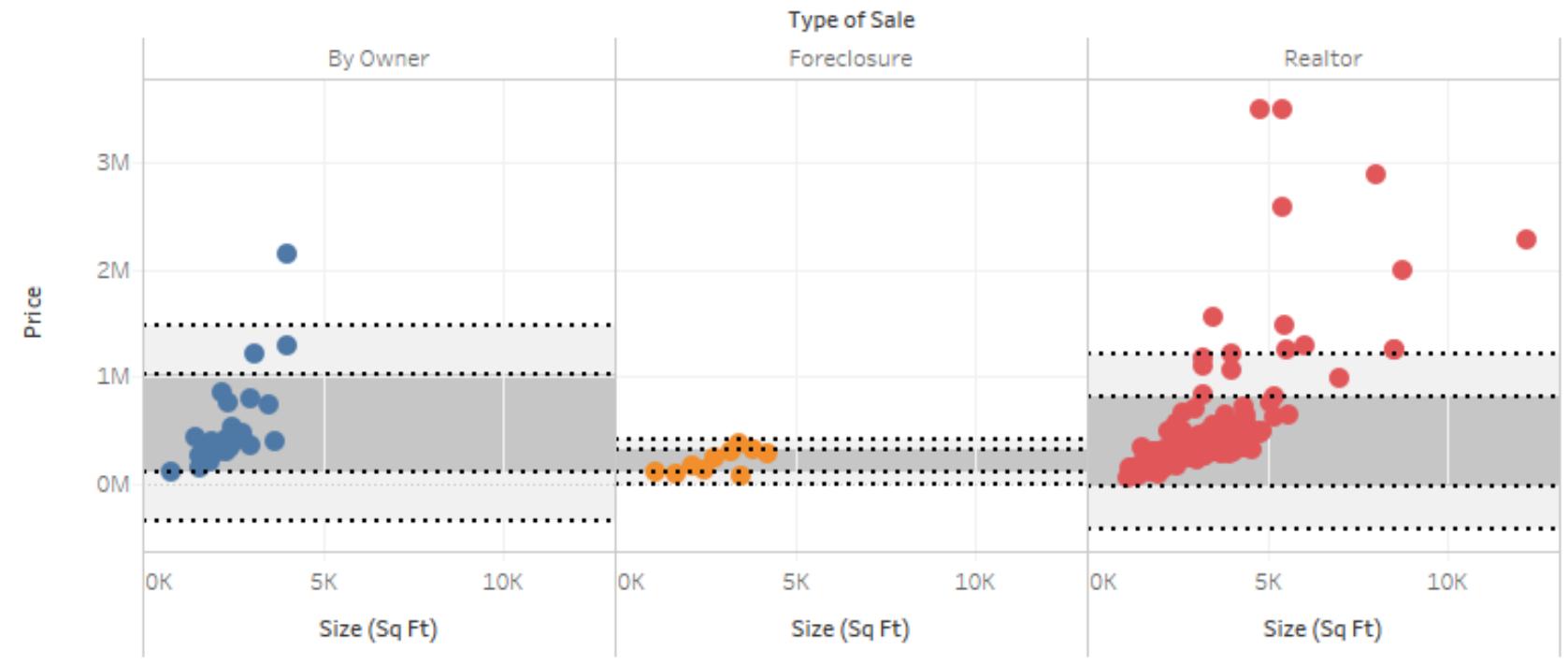
Fill Below

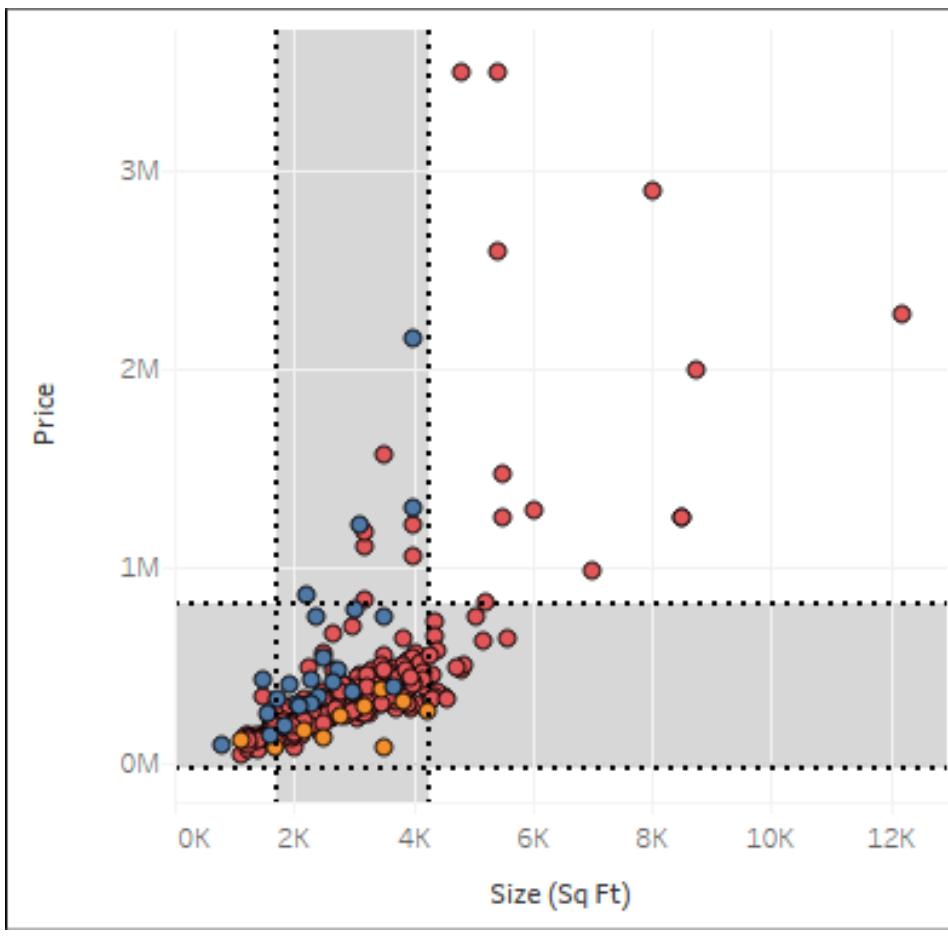
Symmetric

Show recalculated band for highlighted or selected data points

OK

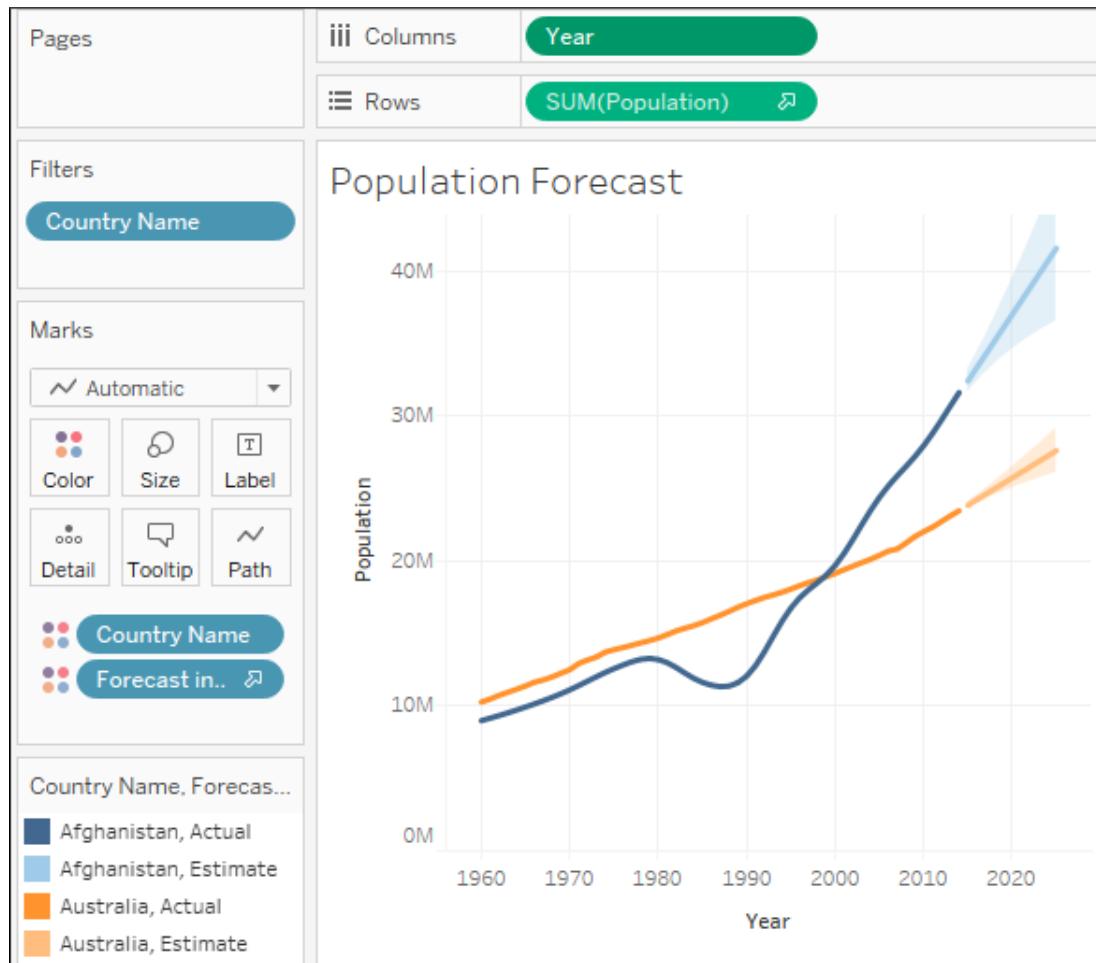
Distributions: Standard Deviation





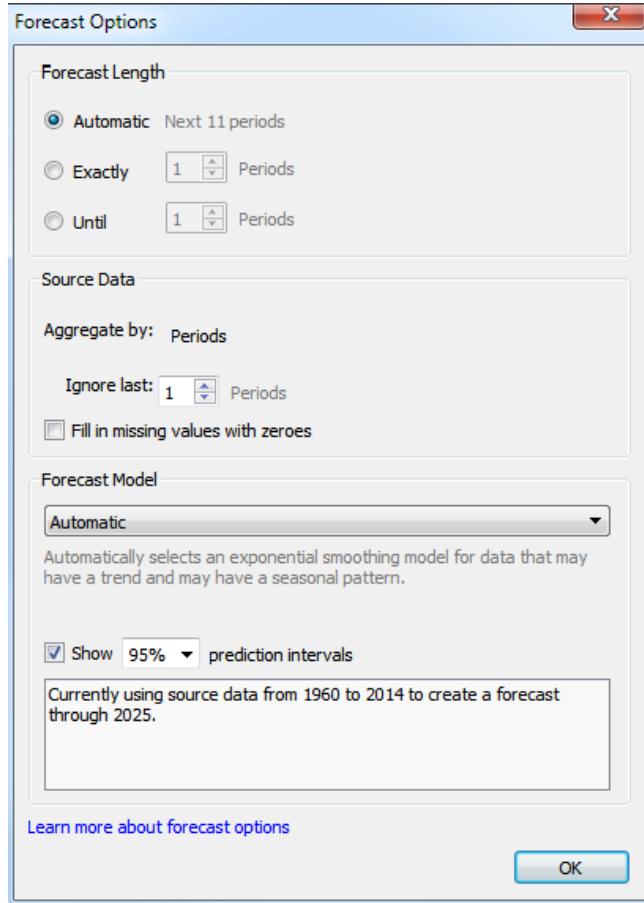
Forecasting

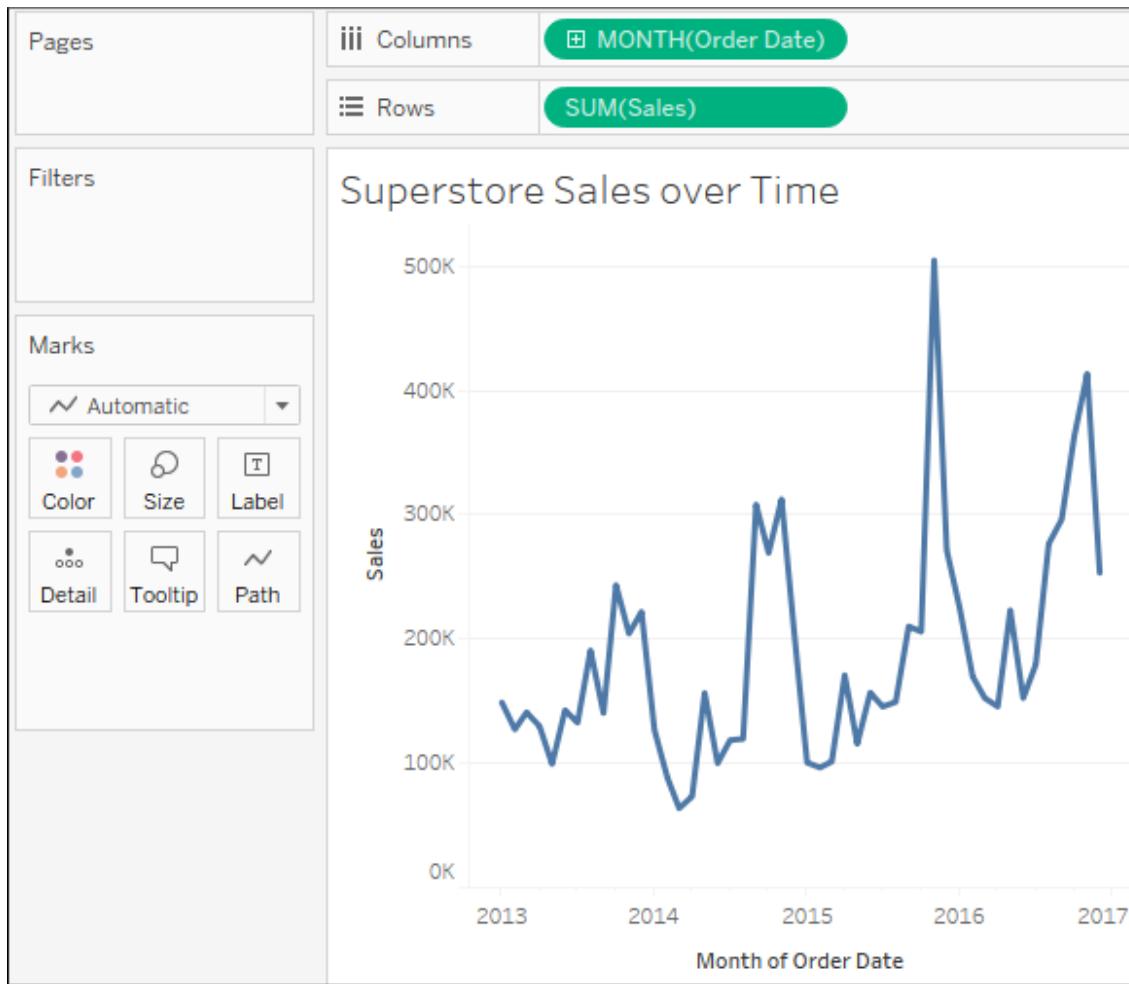
- To use forecasting, you'll need a view that includes a date field or enough date parts for Tableau to reconstruct a date (for example, a Year and a Month field).
- Tableau also allows forecasting based on integers instead of dates.
- You may drag and drop a forecast from the Analytics pane, select Analytics | Forecast | Show Forecast from the menu, or right-click on the view's pane and select the option from the context menu.



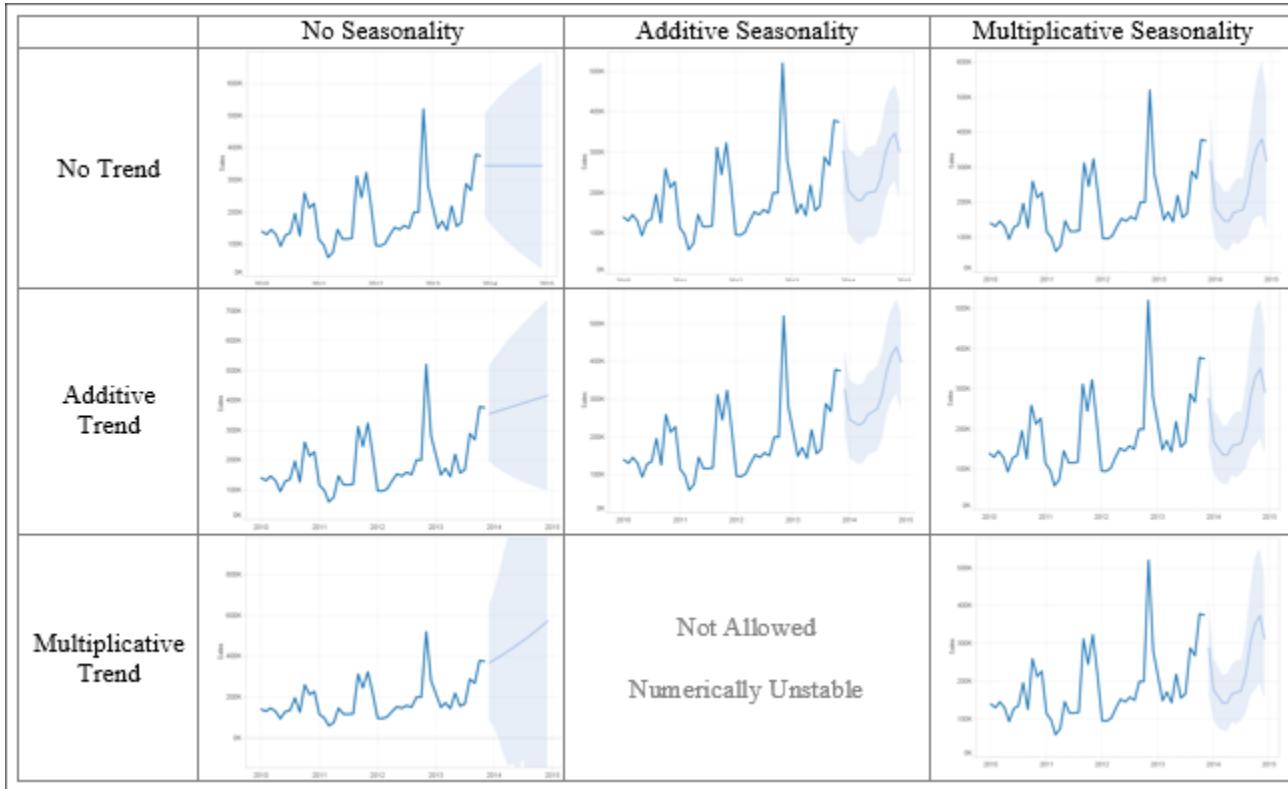
Forecasting

- When you edit the forecast by selecting Analytics | Forecast | Forecast Options... from the menu or use the right-click context menu on the view, you'll be presented with various options for customizing the trend model, like this:





Forecasting



Describe Forecast



Summary

Models

Options Used to Create Forecasts

Time series: Month of Order Date

Measures: Sum of Sales

Forecast forward: 13 months (Dec 2016 – Dec 2017)

Forecast based on: Jan 2013 – Nov 2016

Ignore last: 1 month (Dec 2016)

Seasonal pattern: 12 month cycle

Sum of Sales

Initial Dec 2016	Change From Initial Dec 2016 – Dec 2017	Seasonal Effect		Contribution		Quality
		High	Low	Trend	Season	
275,593 ± 128,837	23,738	Nov 2017	2	Apr 2017	1	100.0% 0.0% Ok

Show values as percentages

[Copy to Clipboard](#)

[Learn more about the forecast summary](#)

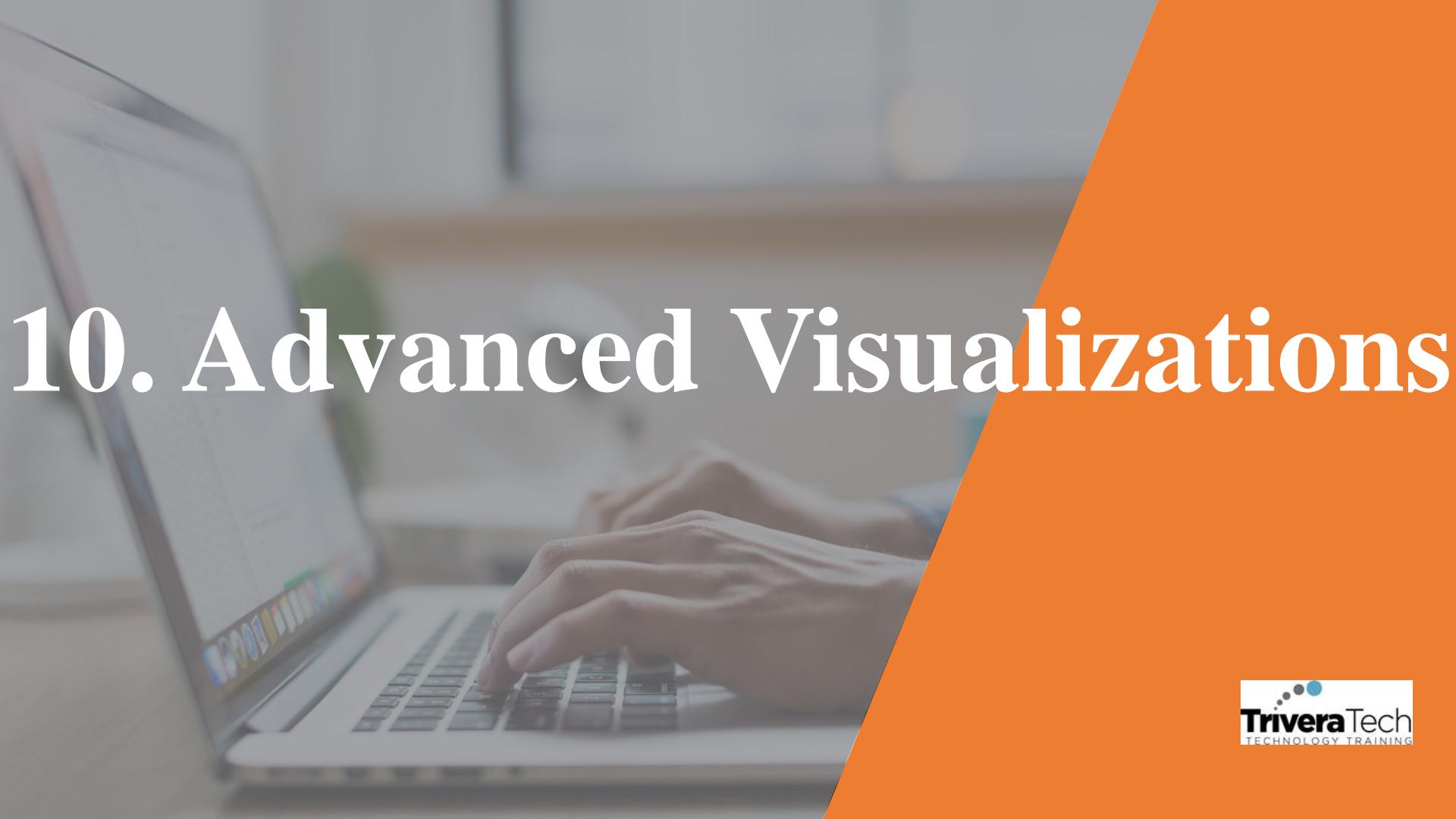
[Close](#)

Summary

- Tableau provides an extensive set of features for adding value to your analysis.
- Trend lines allow you to more precisely identify outliers, determine which values fall within the predictions of certain models, and even make predictions of where measurements are expected.
- Tableau gives extensive visibility into the trend models and even allows you to export data containing trend model predictions and residuals.

COMPLETE LAB 9

10. Advanced Visualizations



Advanced Visualizations

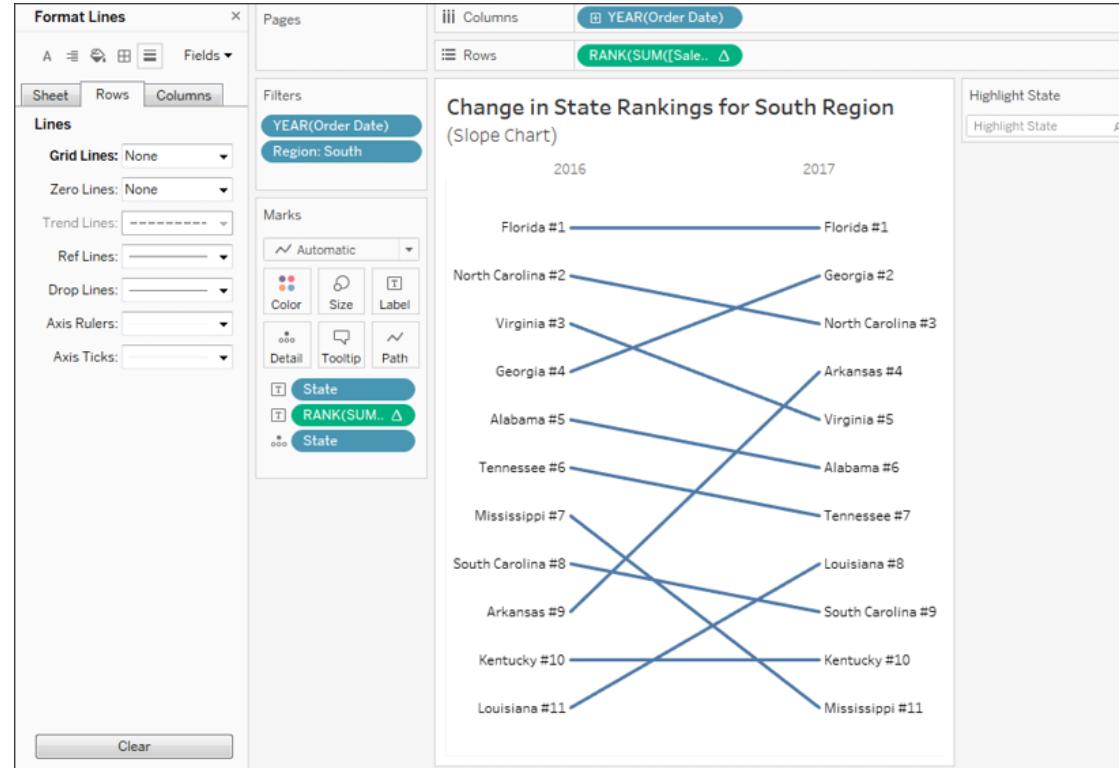
We'll consider the following topics:

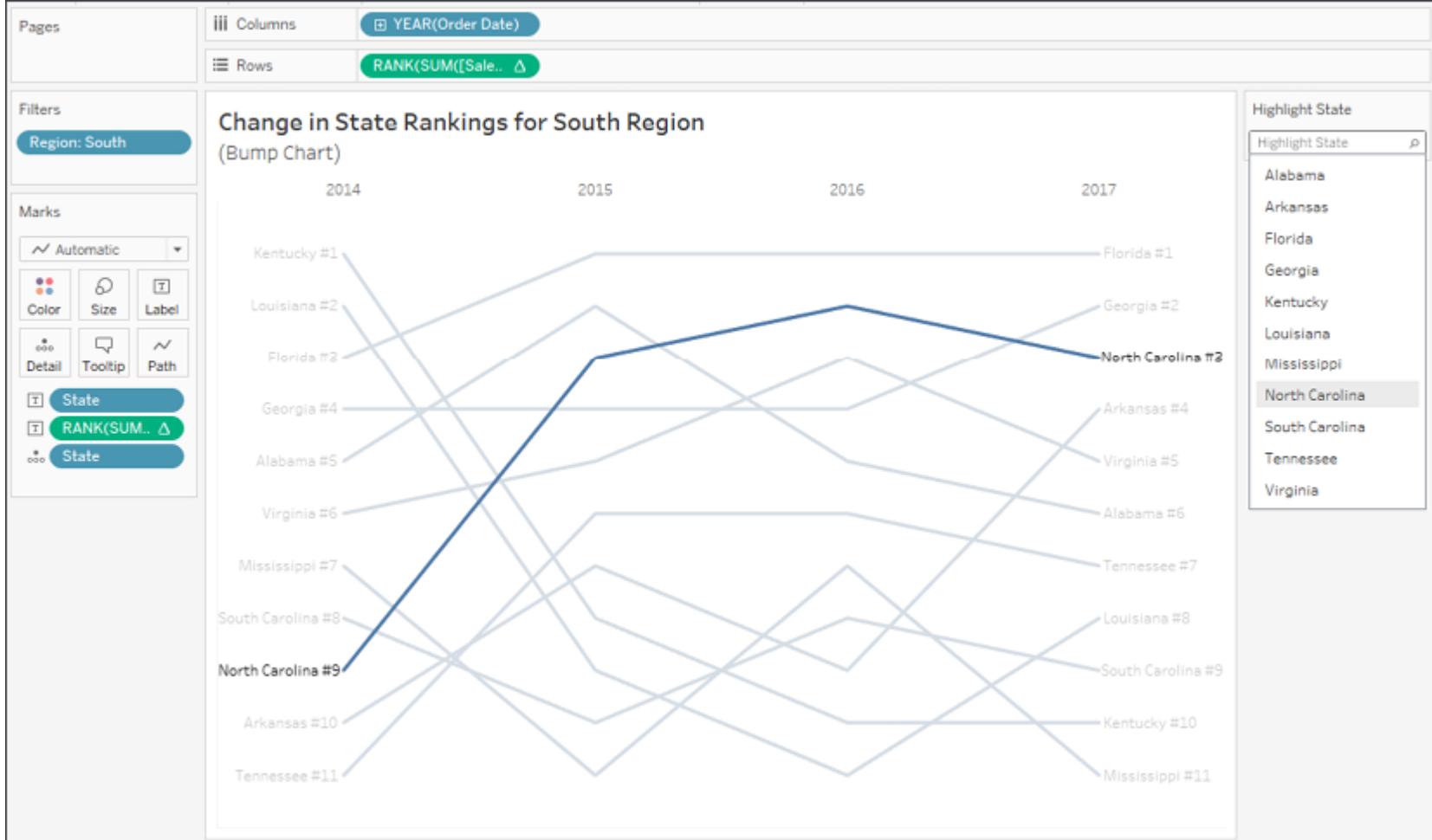
- Advanced visualizations – when and why to use them
- Slope charts and bump charts
- Waterfall charts
- Step lines and jump lines
- Spark lines
- Dumbbell charts
- Unit/symbol charts
- Marimekko charts
- Animated visualizations

Advanced visualizations – when and why to use them

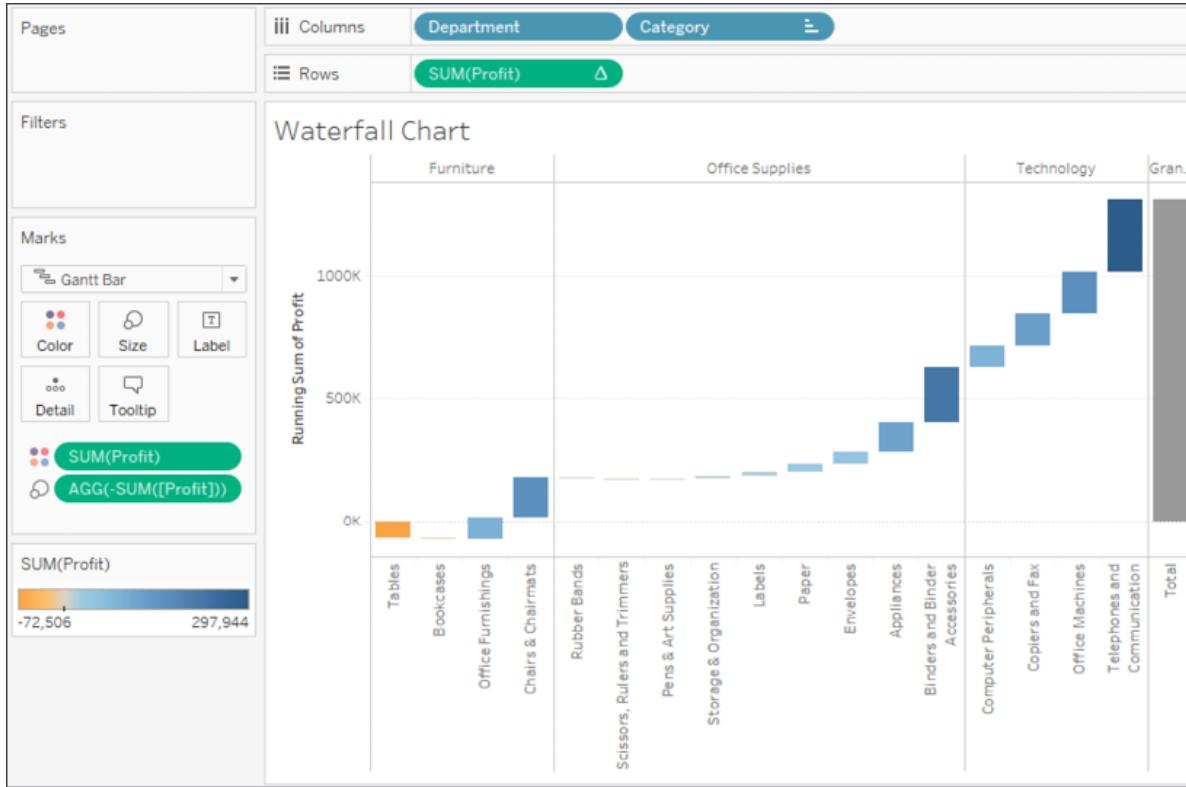
- The visualization types we've seen up to this point will answer many, if not most, of the questions you have about your data.
- If you are asking questions of when?, then a time series is the most likely solution.
- If you are asking how much?, a bar chart gives a good, quick result.
- But there are times when you'll ask questions that are better answered with a different type of visualization.

Slope charts and bump charts



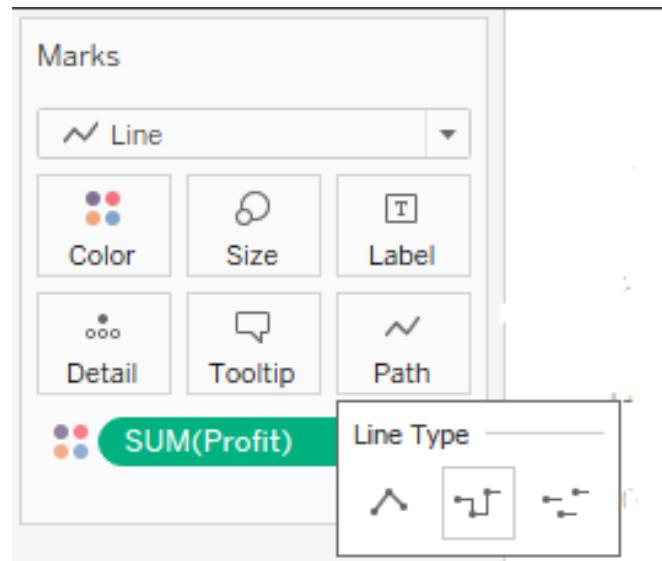


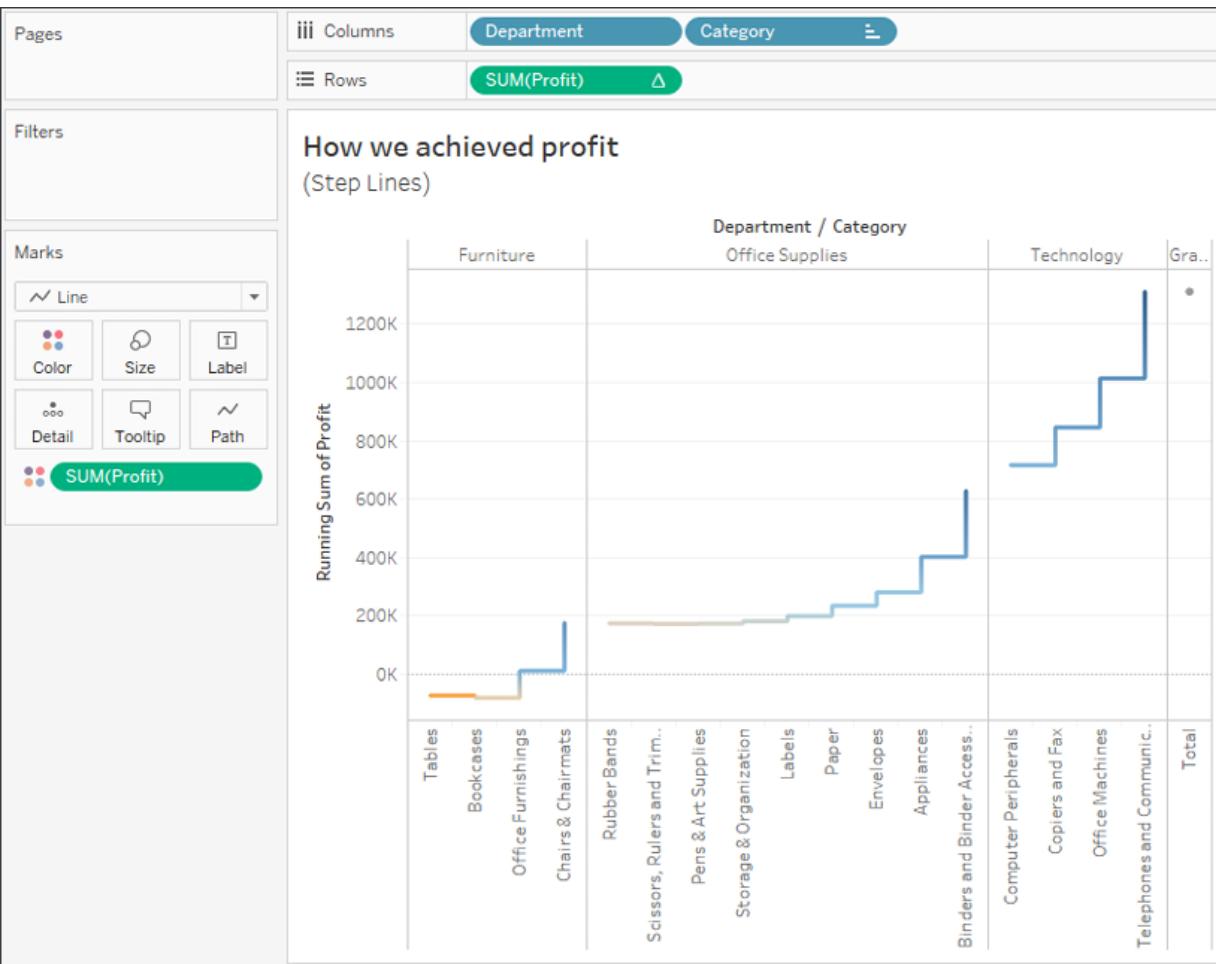
Waterfall charts



Step lines and jump lines

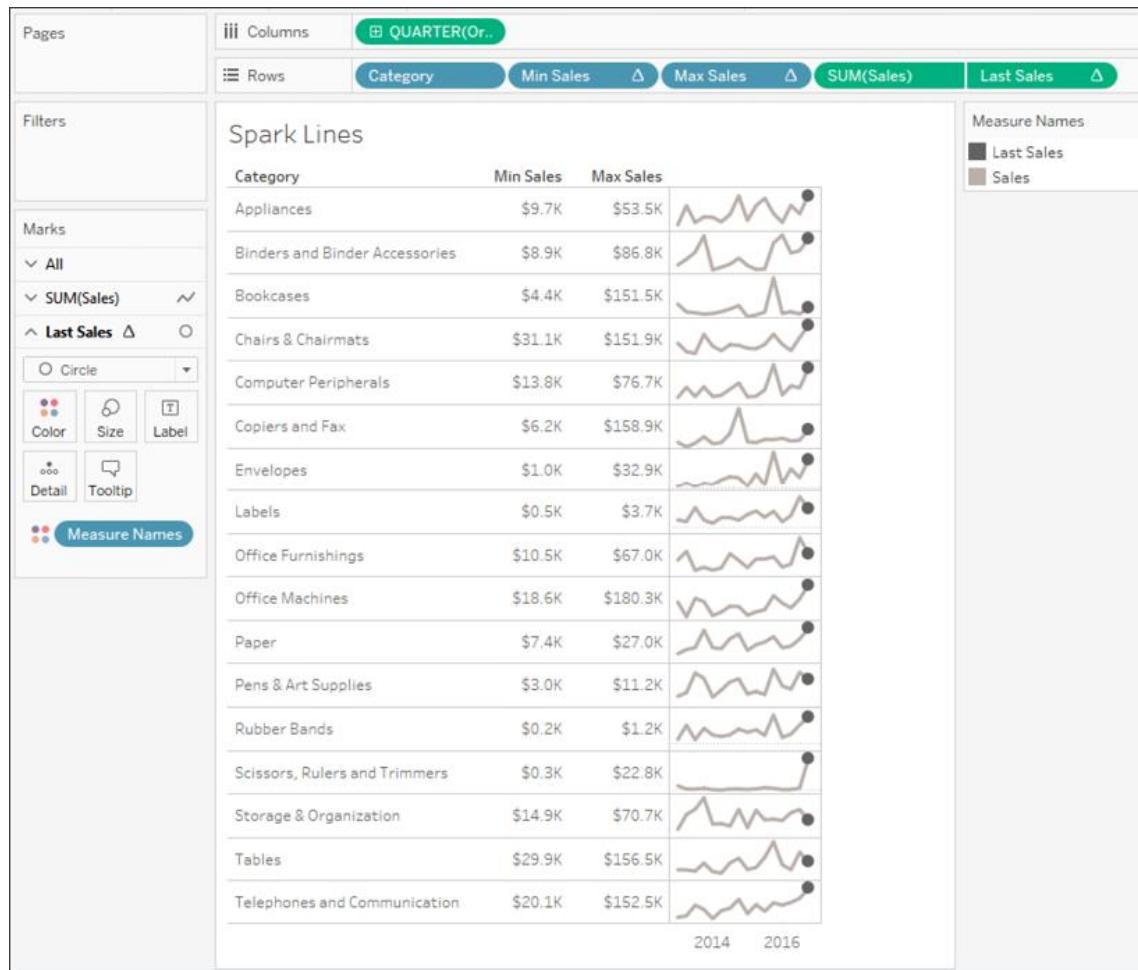
- With a mark type of Line, click the Path shelf and you'll see three options for Line Type:



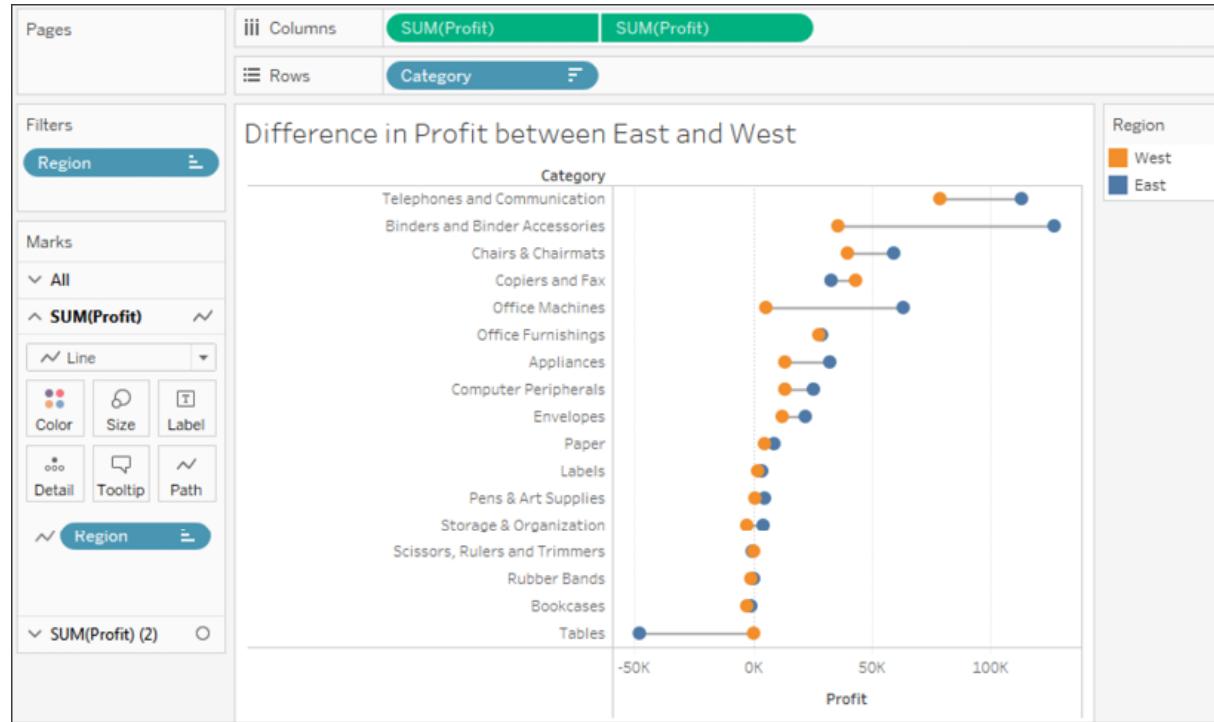


Sparklines

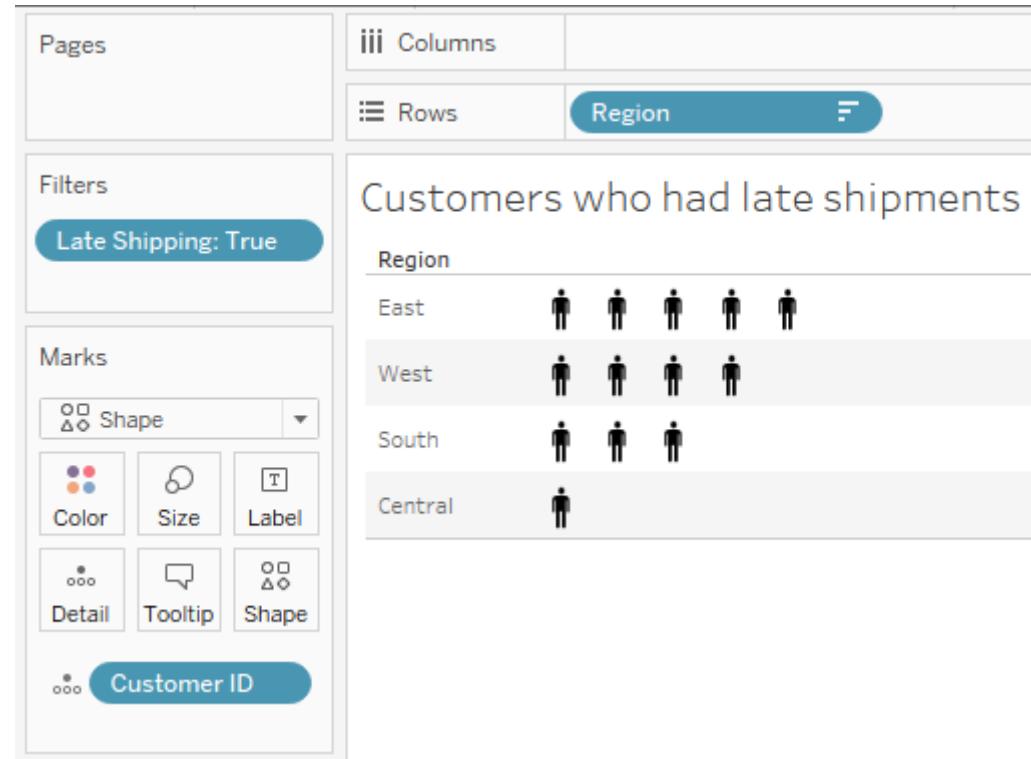
- Sparklines are visualizations that use multiple small line graphs that are designed to be read and compared quickly.
- The goal of sparklines is to give a visualization that can be understood at a glance.
- You aren't trying to communicate exact values, but rather give the audience the ability to quickly understand trends, movements, and patterns.



Dumbbell charts



Unit/symbol charts

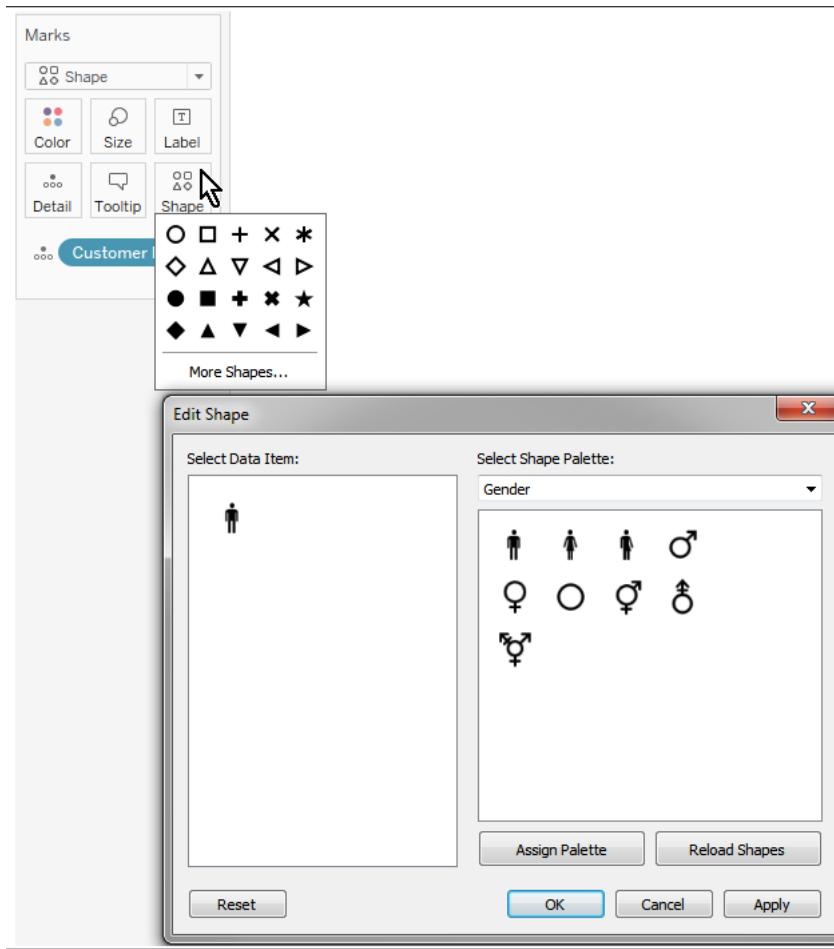


Unit/symbol charts

The view was created with the following techniques:

- The view is filtered where Late Shipping is True. Late Shipping is a calculated field that determines if it took more than 14 days to ship an order. The code is as follows:

DATEDIFF('day', [Order Date], [Ship Date]) > 14

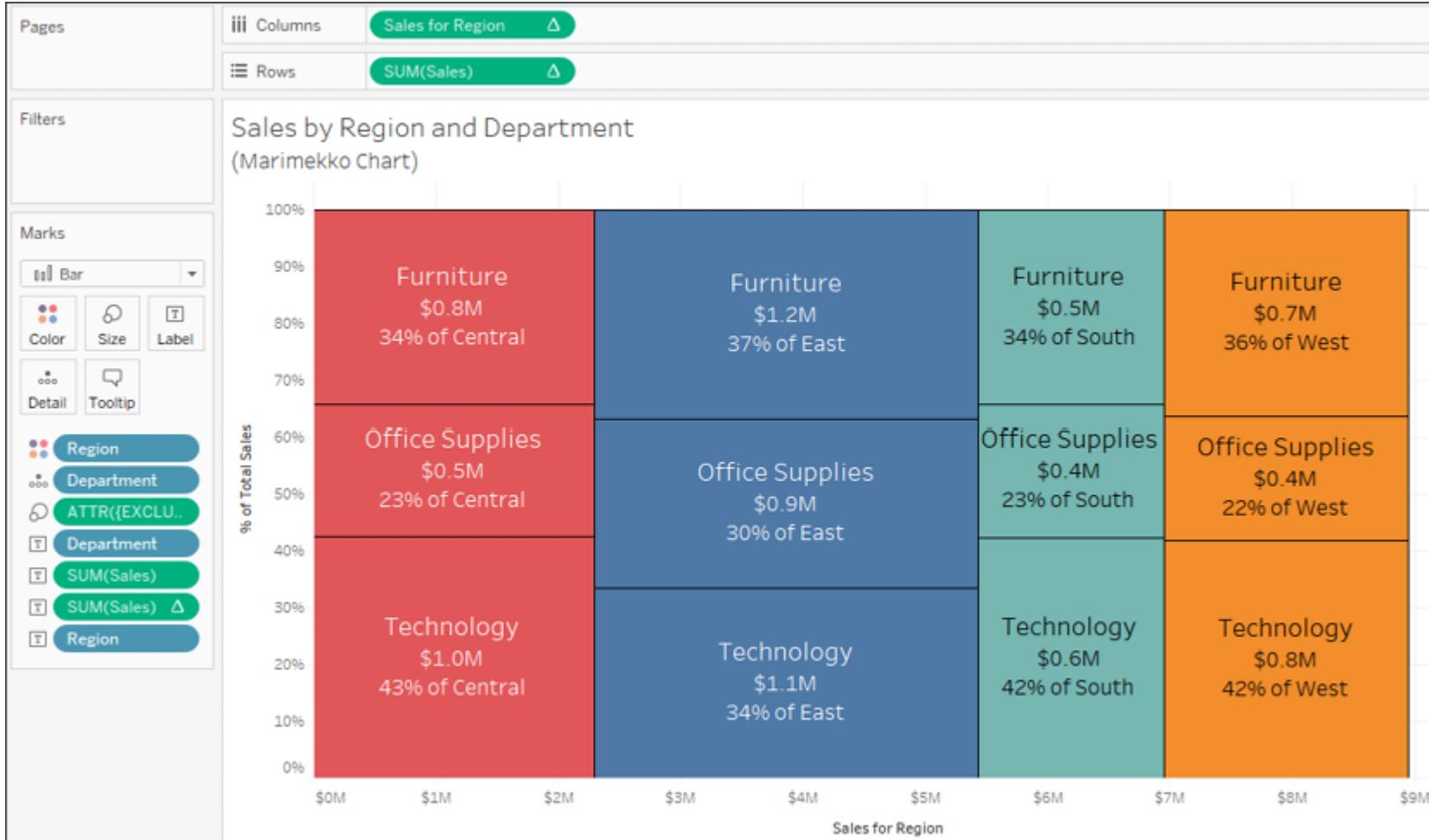


Unit/symbol charts



Marimekko charts

- A Marimekko chart (sometimes alternately called a Mekko chart) is similar to a vertically stacked bar chart, but additionally uses varying widths of bars to communicate additional information about the data.
- Here, for example, is a Marimekko chart showing the breakdown of sales by region and department.



Marimekko charts

- The calculated field Sales for Region calculates the x axis location for the right-side position of each bar.
- The code is as follows:

```
IF FIRST() = 0
    THEN MIN({EXCLUDE [Department] : SUM(Sales)})
ELSEIF LOOKUP(MIN([Region]), -1) <> MIN([Region])
    THEN PREVIOUS_VALUE(0) + MIN({EXCLUDE
[Department] : SUM(Sales)})
ELSE
    PREVIOUS_VALUE(0)
END
```

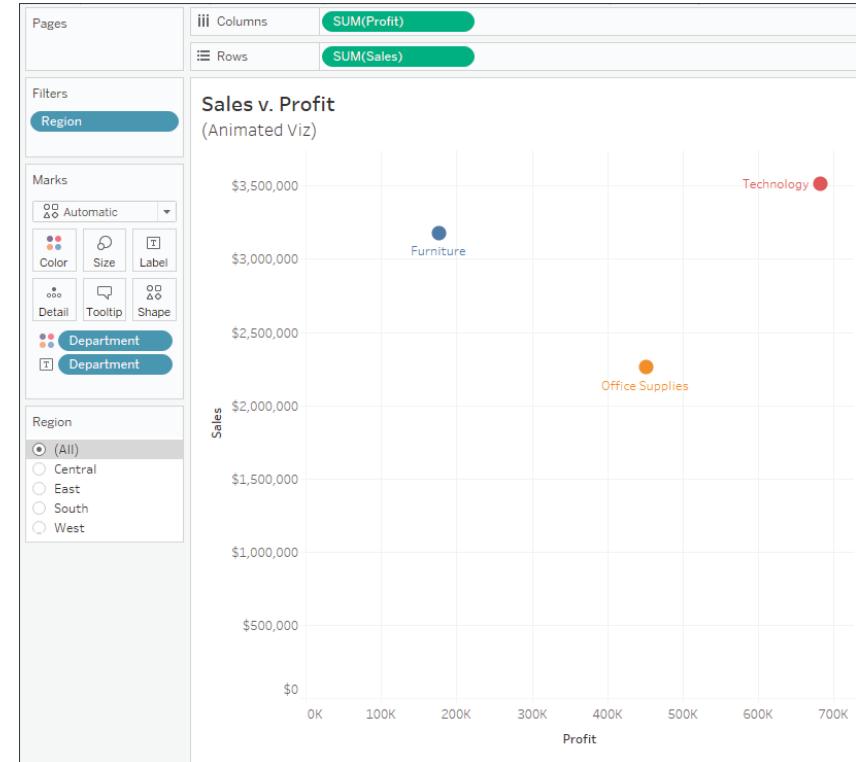
Animated visualizations

Consider leveraging animation to extend your analytical potential in a couple of ways:

- Turn it on while exploring and analyzing your data. This allows you to gain analytical insights you might otherwise miss, such as seeing how far and in which direction marks in a scatterplot move as a filter changes.
- Use it strategically to enhance the data story. Animation can be used to capture interest, draw attention to important elements, or build suspense toward a conclusion.

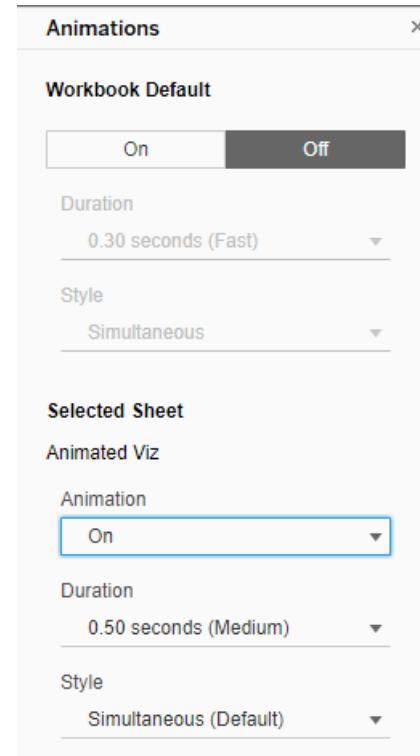
Enhancing analysis with animation

- Consider the following bar chart, which shows the correlation of Sales and Profit for each Department:



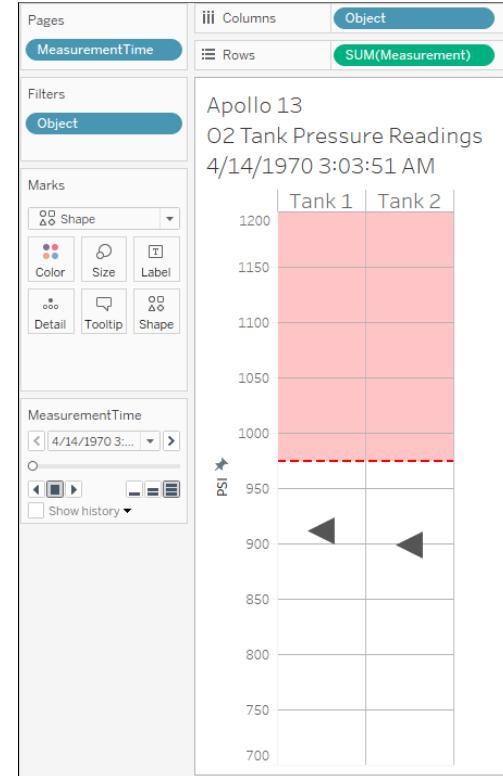
Enhancing analysis with animation

- Now, turn on animations for the view. To do this, use the menu to select Format | Animations... The Animations format pane will show on the left.
- Use it to turn Animation on for the Selected Sheet:



Enhancing data storytelling with animation

- Beyond providing analytical insight as you perform your data discovery and analysis, you can also leverage animation to more effectively drive interest and highlight decision points, opportunities, or risks in your data stories.
- As an example, consider this view in the lesson 10 workbook:



Summary

- We've covered a wide variety of advanced visualization types in this lesson!
- We've considered slope and bump charts that show changes in rank or value, step and jump lines that show discretely changing values, and unit charts that help materialize abstract concepts.

COMPLETE LAB 10

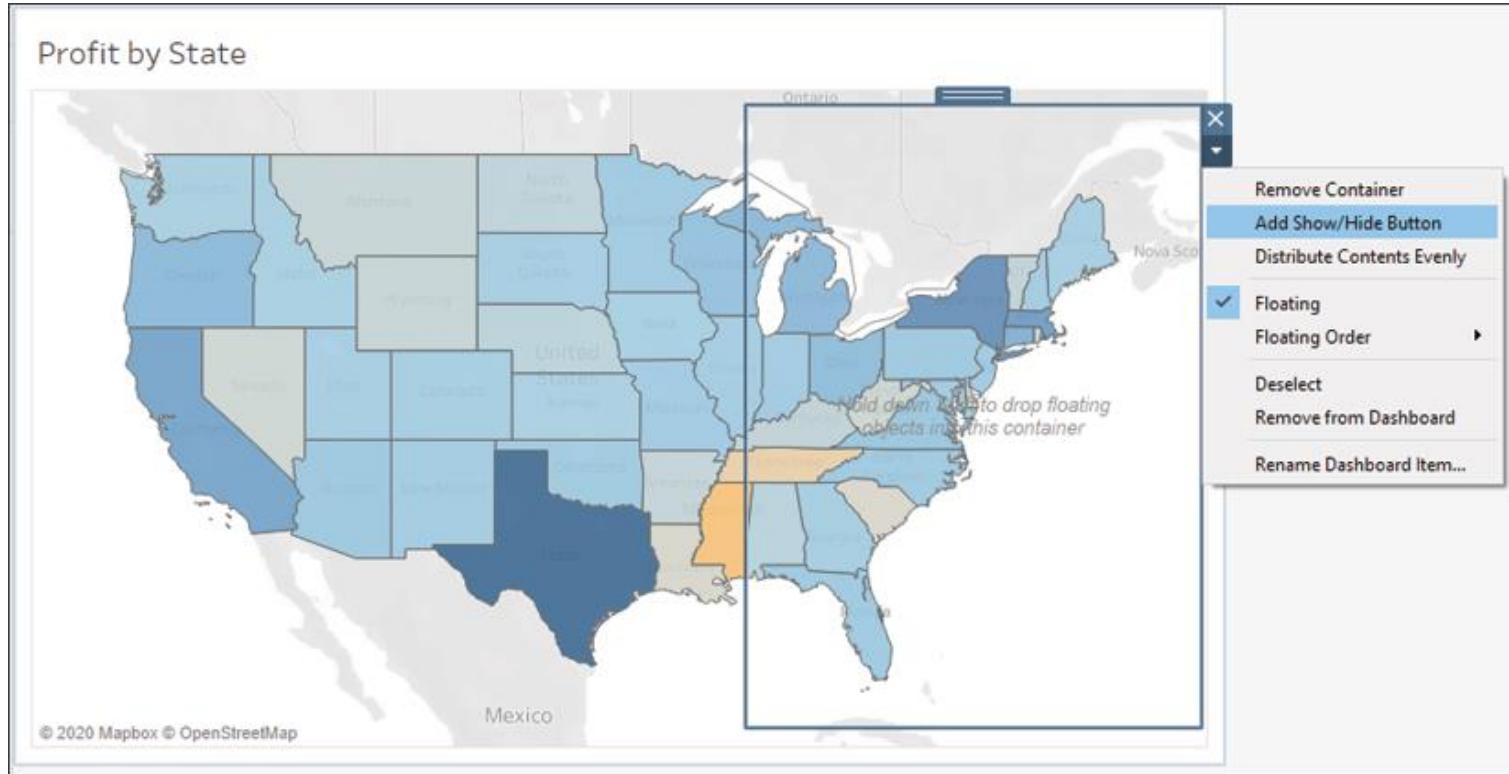
11. Dynamic Dashboards

Dynamic Dashboards

We'll consider the following types of dynamic interactivity:

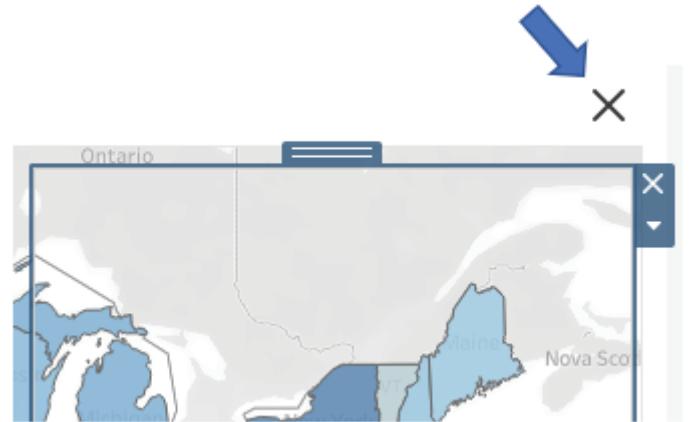
- Show/Hide buttons
- Sheet swapping
- Automatically showing and hiding controls

Show/Hide buttons



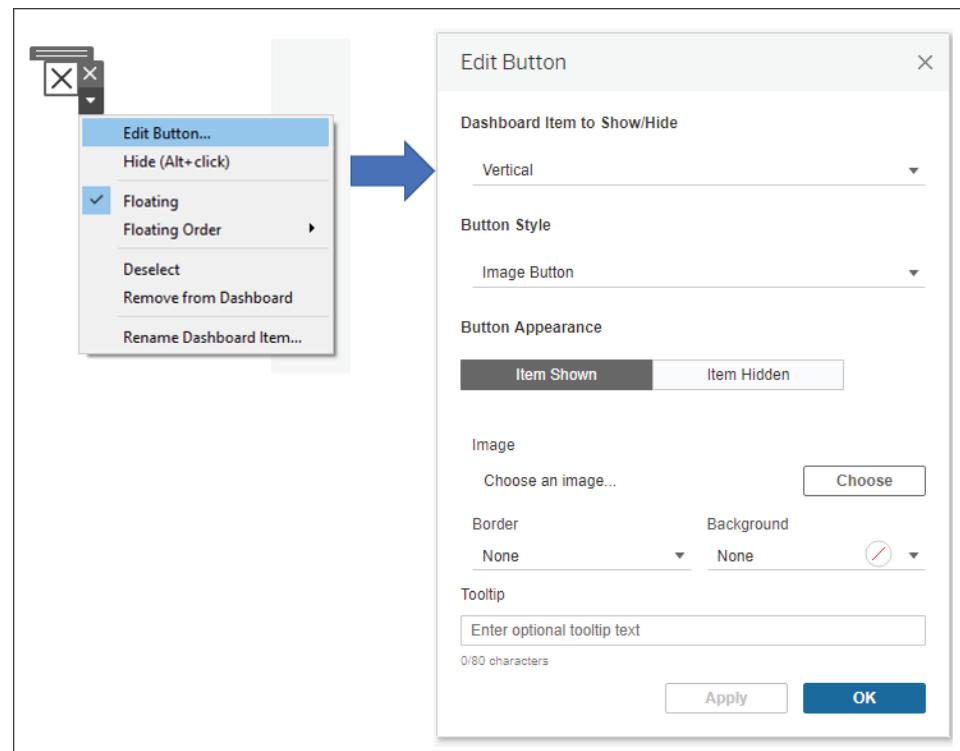
Show/Hide buttons

- On this dashboard, we've added a vertical layout container floating over the map.
- When we select Add Show/Hide Button, Tableau adds a small button to the dashboard:



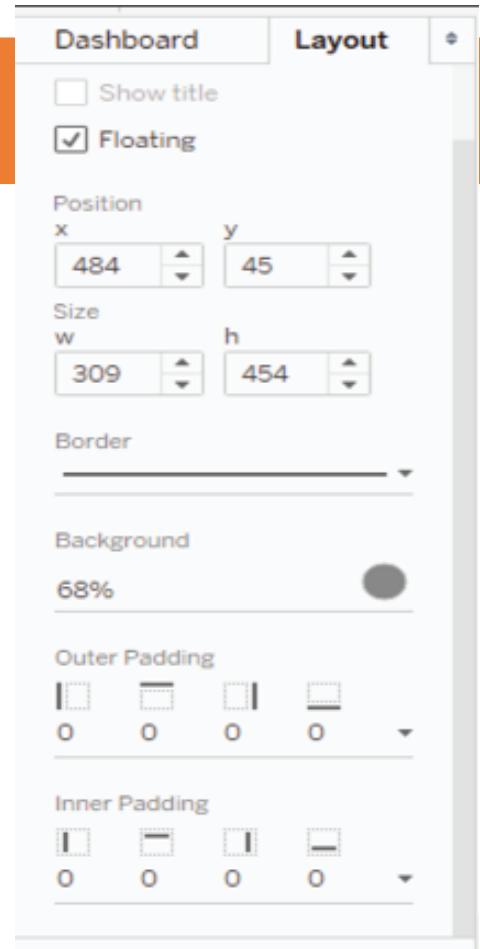
Show/Hide buttons

- You may edit the button's appearance and actions using the button's dropdown to select Edit Button....:



Show/Hide buttons

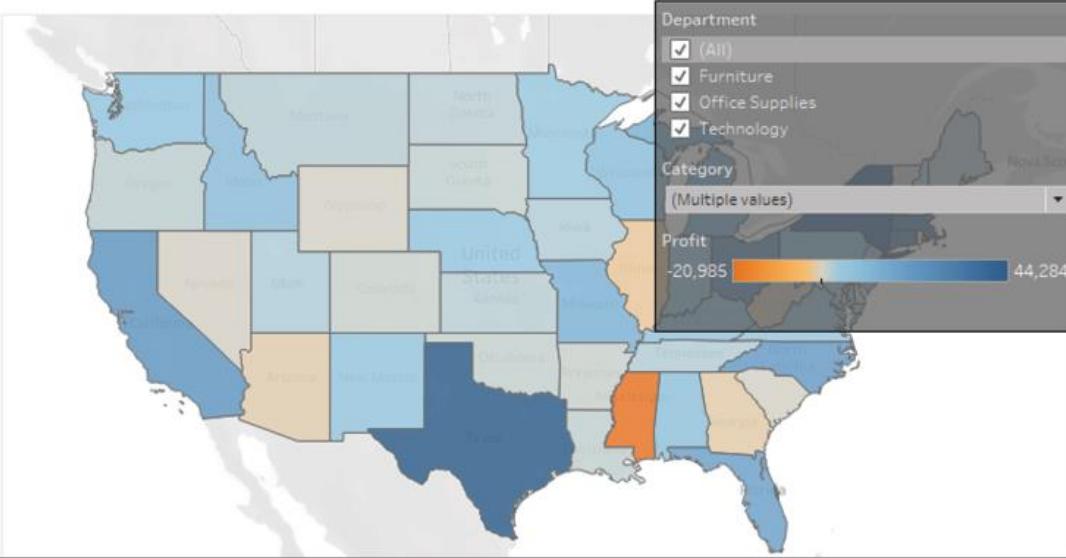
- Additionally, the container has been given a border and a semi-transparent background so that the map is slightly visible beneath the container.
- This was accomplished by selecting the container and then using the Layout tab of the Dashboard pane, as shown here:



Profit by State

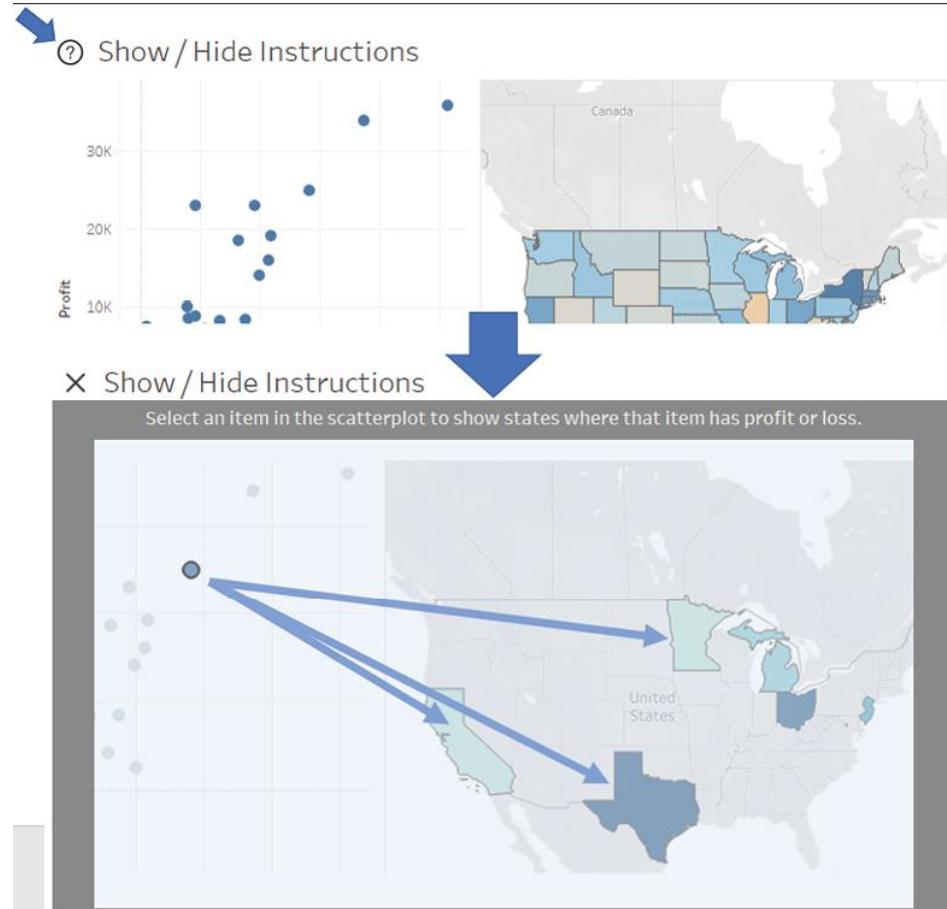


Profit by State



Show/Hide buttons

- Consider another example in the lesson 11 Complete workbook that illustrates the possibility of showing and hiding instructions:



Sheet swapping

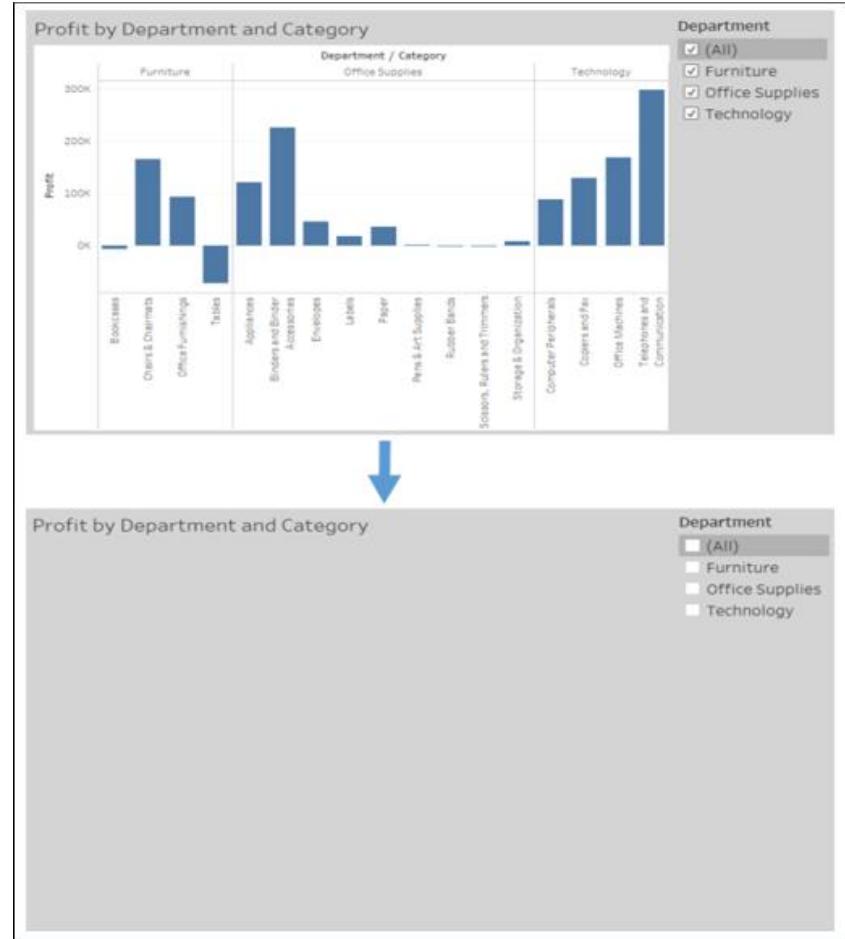
- Sheet swapping, sometimes also called sheet selection, is a technique in which views are dynamically shown and hidden on a dashboard, often with the appearance of swapping one view for another.
- The dynamic hiding and showing of views on a dashboard has an even broader application.

Basic principles of sheet swapping

The basic principles are relatively simple. A view on a dashboard collapses when the following occur:

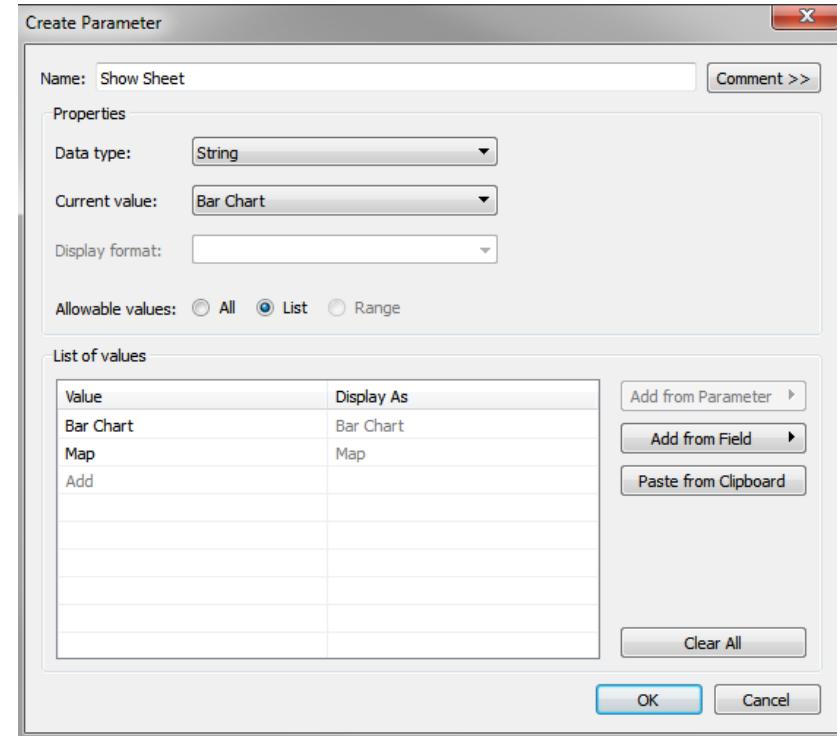
- At least one field is on rows or columns
- It is in a horizontal or vertical layout container
- It does not have a fixed height or width and the layout container is not set to distribute items equally
- A combination of filters or hiding prevents any marks from being rendered

Using sheet swapping to change views on a dashboard



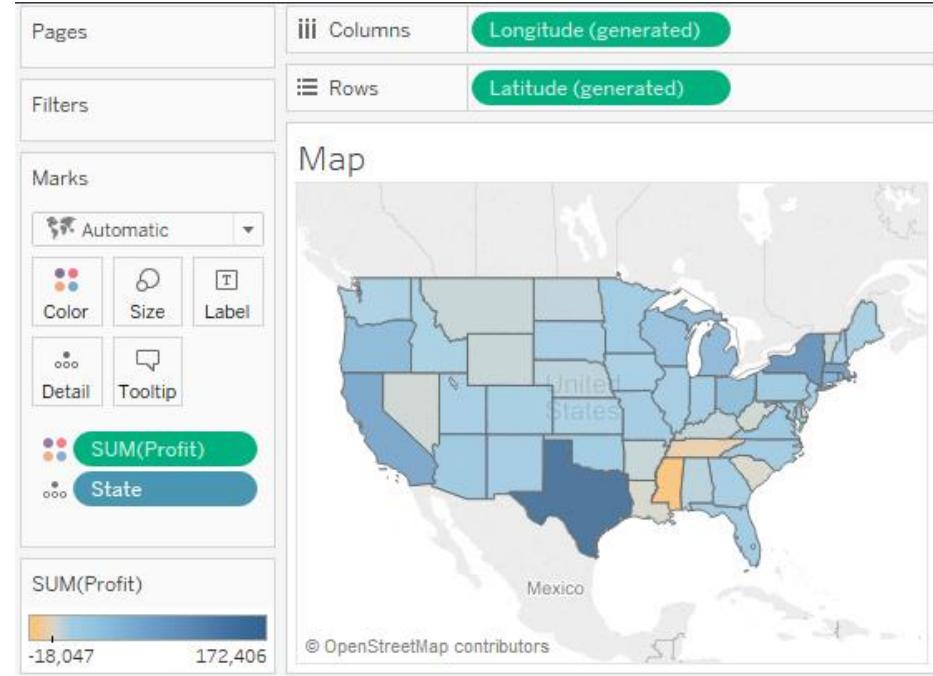
Using sheet swapping to change views on a dashboard

- Create an integer parameter named Show Sheet with a list of String values set to Bar Chart and Map:



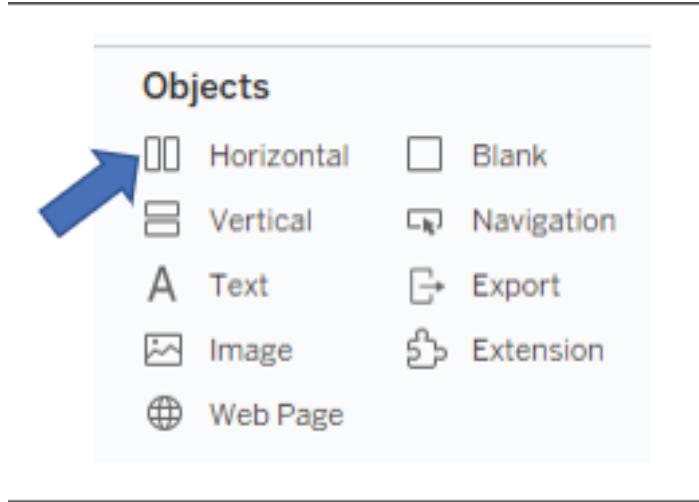
Using sheet swapping to change views on a dashboard

- Create another sheet named Map that shows a filled map of states by profit:

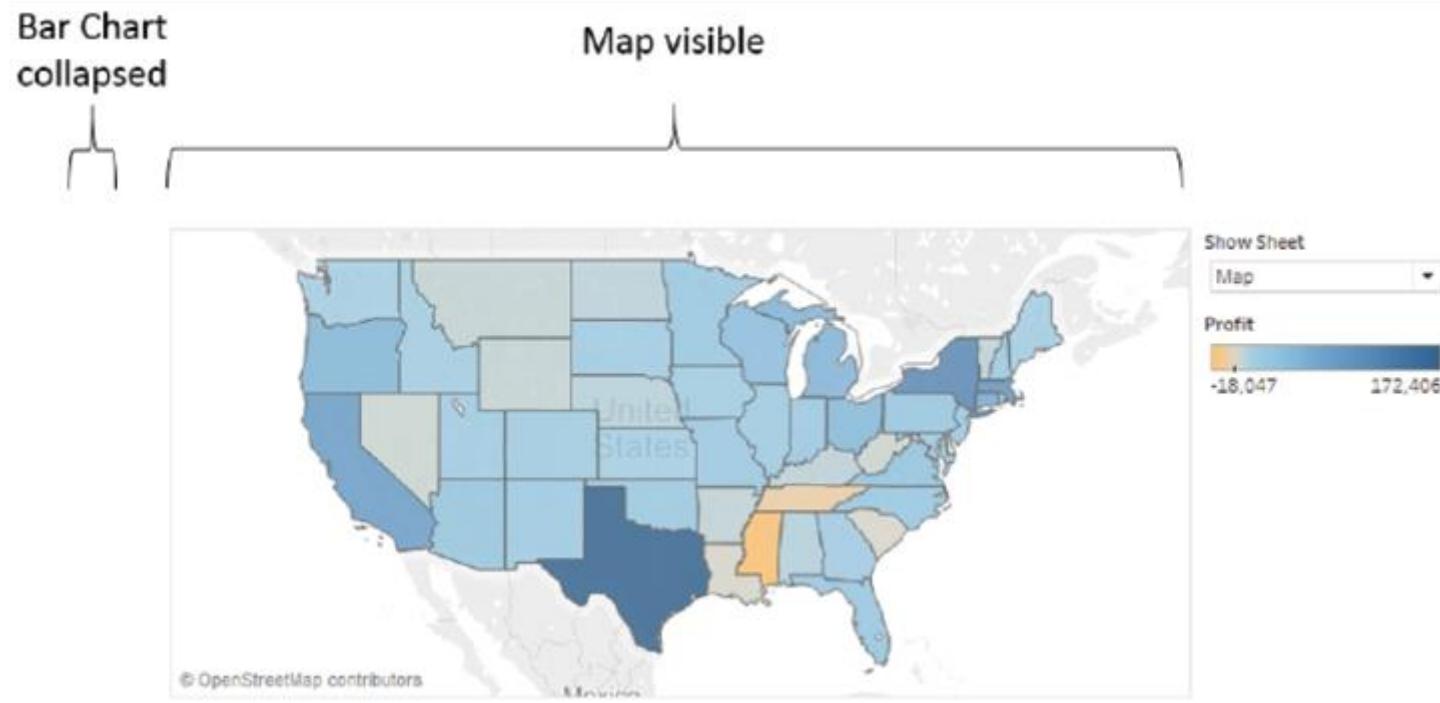


Using sheet swapping to change views on a dashboard

- Add a Horizontal layout container to the dashboard from the objects in the left window:

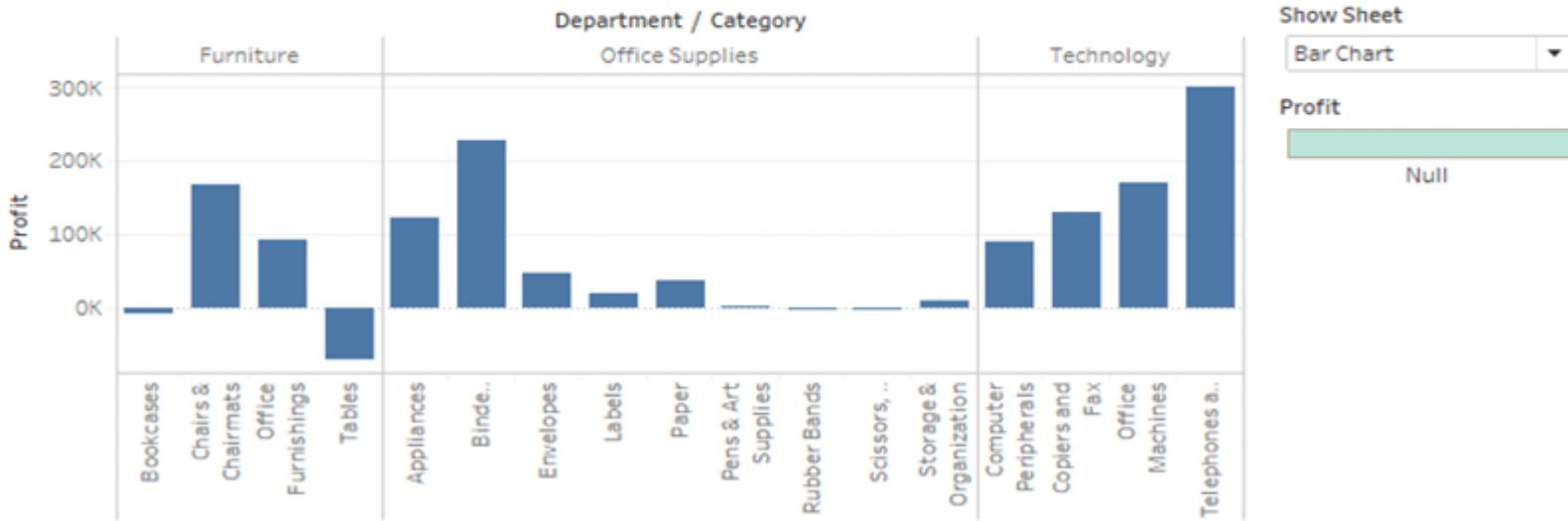


Using sheet swapping to change views on a dashboard



Bar Chart visible

Map collapsed



Automatically showing and hiding other controls

- Views will collapse when all data is filtered out.
- However, other controls, such as quick filters, parameters, images, legends, and textboxes, will not collapse.
- You could use a Show/Hide button as we previously discussed, but often you'll want to show or hide these controls automatically as filters change, without forcing the user to take additional action.

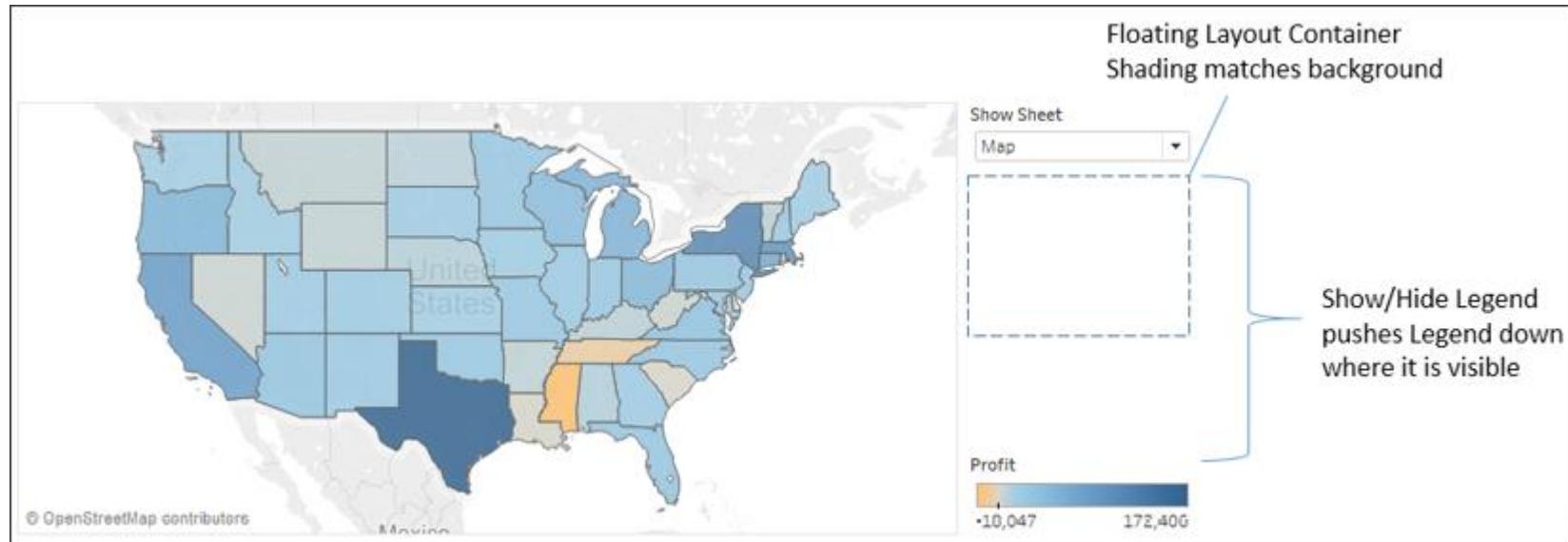
Automatically showing and hiding other controls

The screenshot shows a data visualization interface with several panels:

- Pages:** Contains "Columns" and "Rows" sections, each with a button labeled "AGG(MIN(1))".
- Filters:** Contains a "Show Sheet Filter: M.." button.
- Marks:** Contains a dropdown menu set to "Automatic", and buttons for "Color", "Size", "Label", "Detail", "Tooltip", and "Shape".
- Central Area:** Labeled "Show/Hide Legend".
- Show Sheet:** A dropdown menu currently set to "Map".

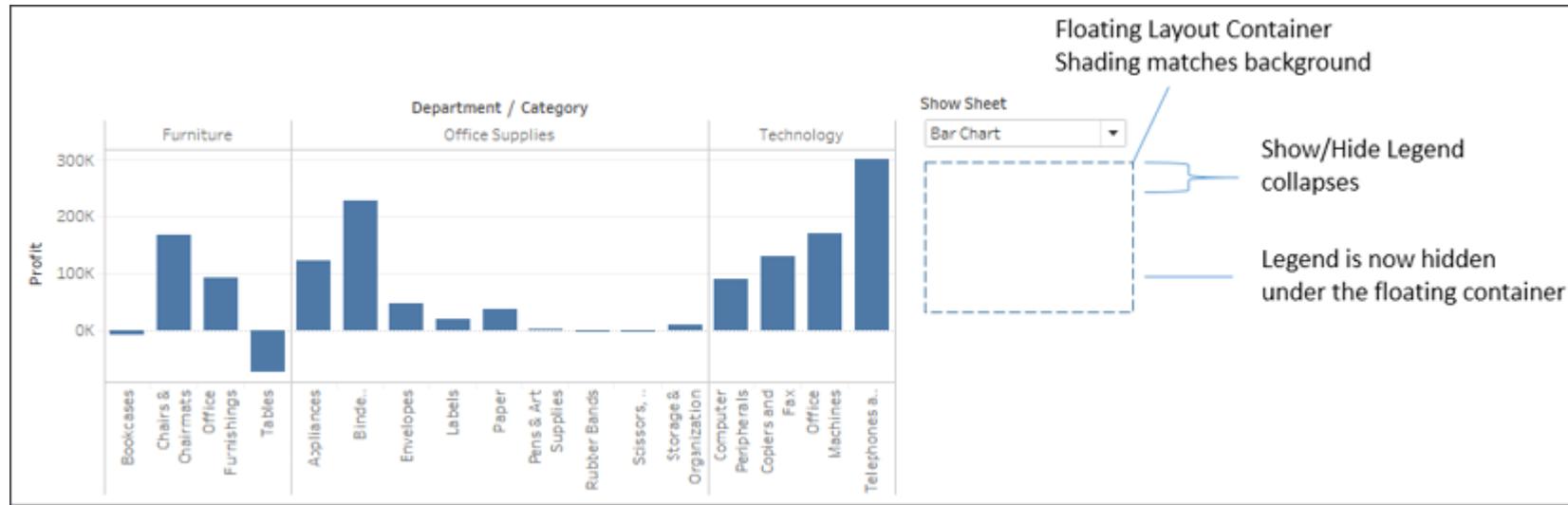
The "Rows" section under "Pages" and the "Show Sheet" dropdown are highlighted with green rounded rectangles, indicating they are the controls being demonstrated for automatic visibility.

Automatically showing and hiding other controls



Automatically showing and hiding other controls

- When Bar Chart is selected, the Show/Hide Legend sheet collapses and the legend, which is no longer applicable to the view, falls under/hides behind the Floating Layout Container:



Summary

- Creating truly dynamic dashboards give you incredible flexibility in your dashboard design. Show/Hide buttons give the end user the option to show or hide content on the dashboard.
- Sheet-swapping techniques allow you to swap out which views are shown and also automatically show or hide controls or other content.

COMPLETE LAB 11

12. Exploring Mapping and Advanced Geospatial Features

Exploring Mapping and Advanced Geospatial Features

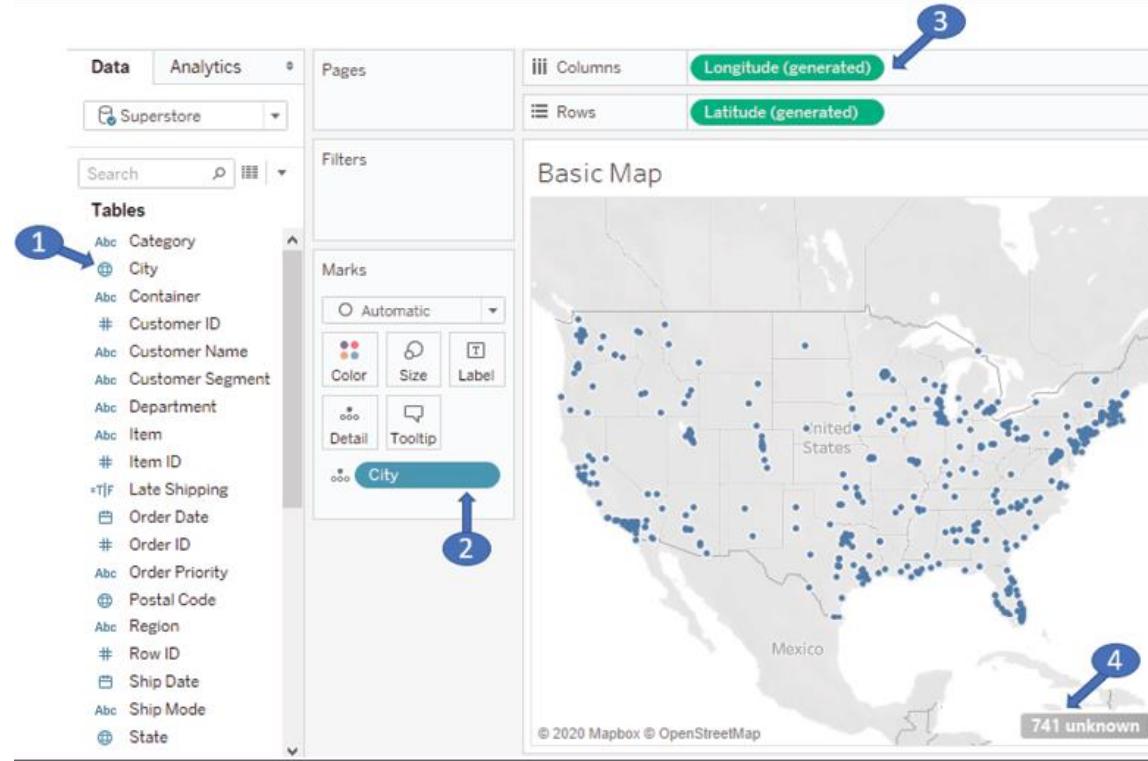
In this lesson, we'll cover the following topics:

- Overview of Tableau maps
- Rendering maps with Tableau
- Using geospatial data
- Leveraging spatial functions
- Creating custom territories
- Tableau mapping: tips and tricks
- Plotting data on background images

Overview of Tableau maps

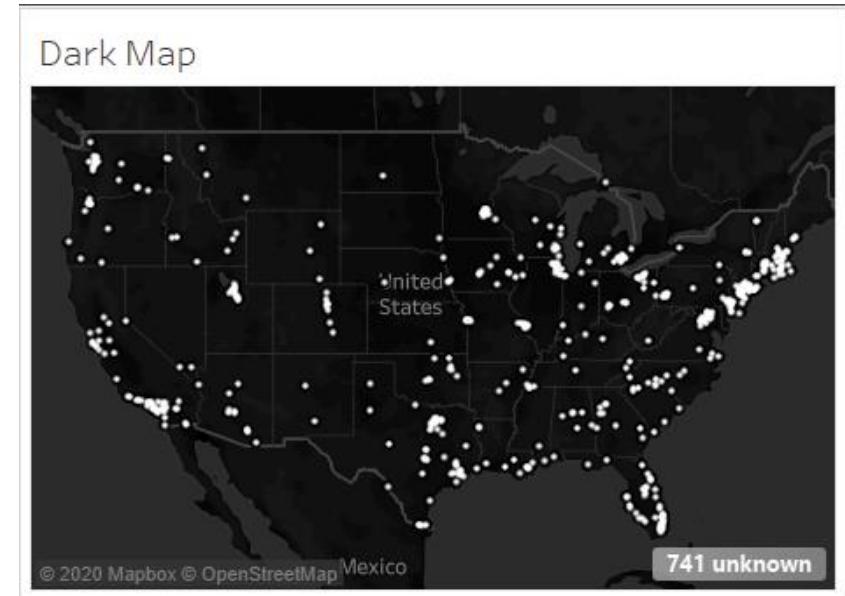
- Tableau contains an internal geographic database that allows it to recognize common geographic elements and render a mark at a specific latitude and longitude on a map.
- In many cases, such as with a country or state, Tableau also contains internal shapefiles that allow it to render the mark as a complex vector shape in the proper location.

Rendering maps with Tableau



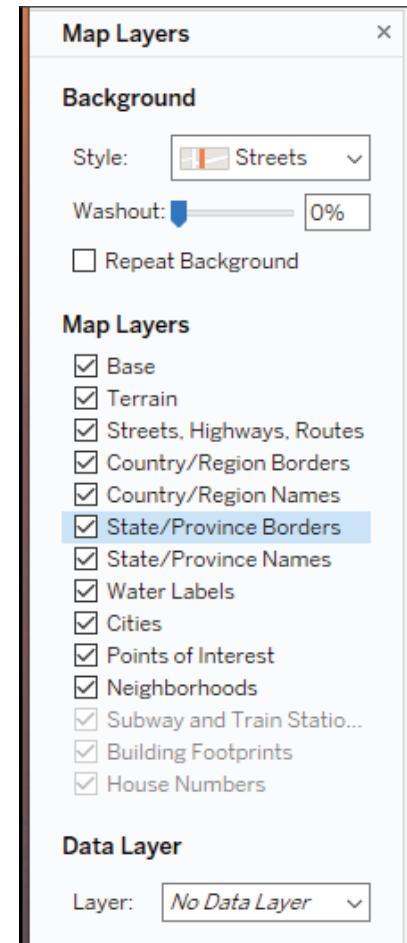
Customizing map layers

- Use the menu to explore various options by selecting Map | Background Maps. Here, for example, is a Dark Map:



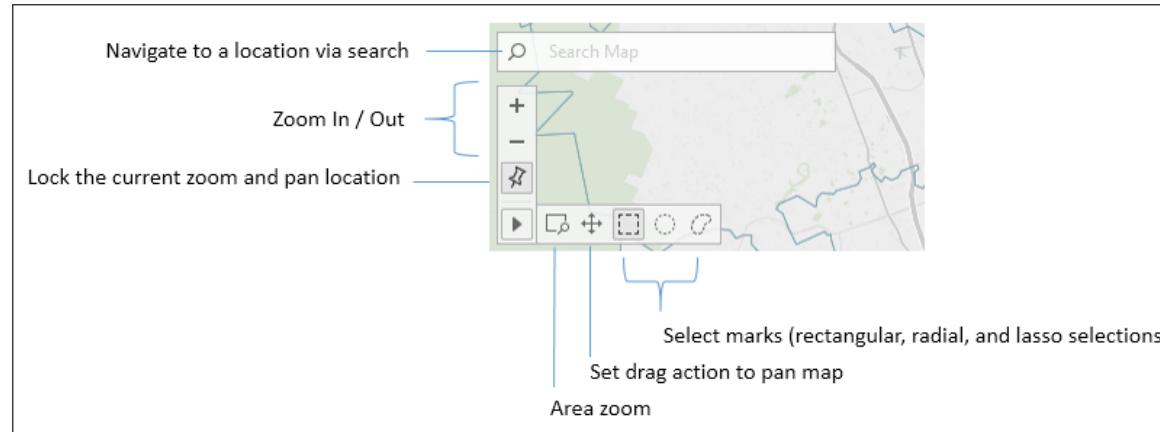
Customizing map layers

- Additional layer options can be found by selecting Map | Map Layers from the menu.
- This opens a Map Layers pane that looks like this:



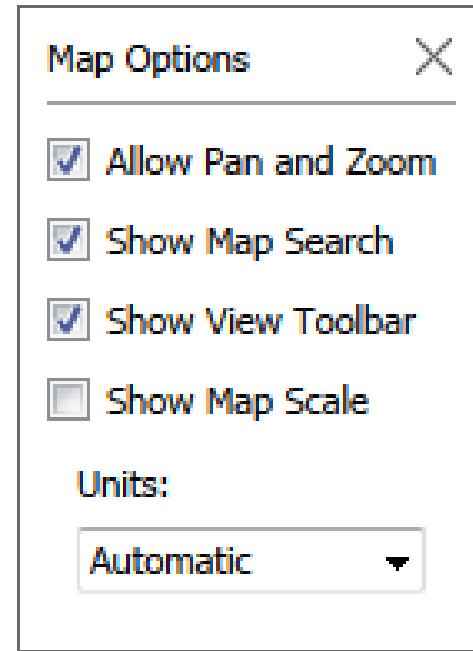
Customizing map options

- You can customize the map options available to end users.
- Notice the controls that appear when you hover over the map:



Customizing map options

- Additional options will appear when you select Map | Map Options from the top menu:



Using geospatial data

We'll consider ways to use data that enable geospatial visualizations, including the following:

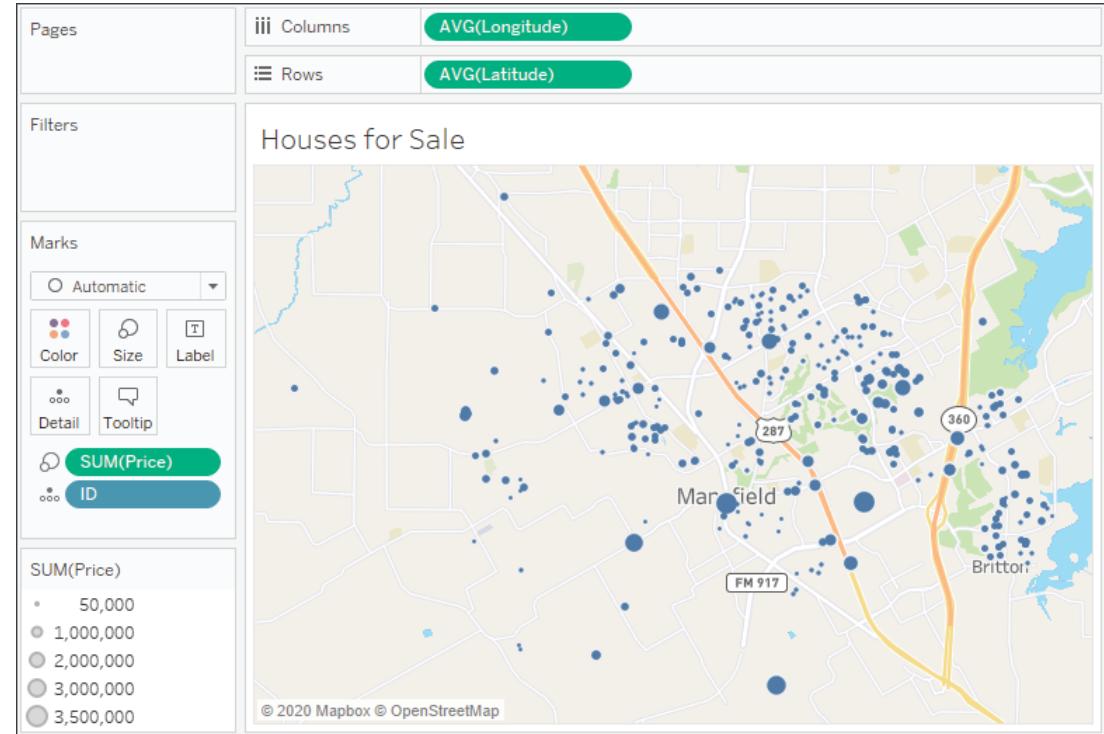
- Including Latitude and Longitude as values in your data.
- Importing a .csv file containing definitions of Latitude and Longitude into Tableau's database.
- Leveraging Tableau's ability to connect to various spatial files or databases that natively support spatial objects.

Including latitude and longitude in your data

- Including latitude and longitude in your data gives you a lot of flexibility in your visualizations (and calculations).
- For example, while Tableau has built-in geocoding for countries, states, and zip codes, it does not provide geocoding at an address level.

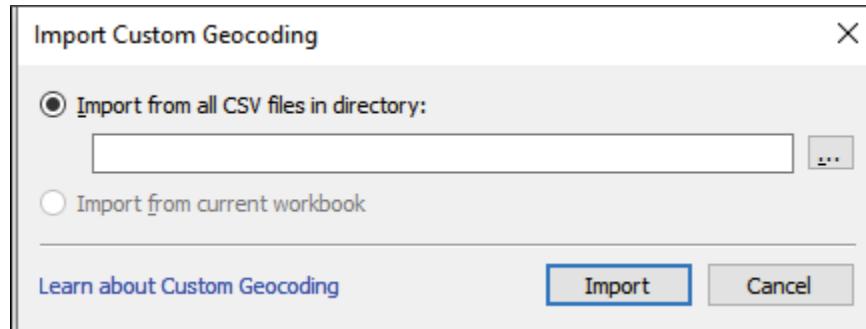
Including latitude and longitude in your data

- You'll find the following example in the lesson 12 workbook using the Real Estate data source:



Importing definitions into Tableau's geographic database

- In order to import from the menu, select Map | Geocoding | Import Custom Geocoding.... The import dialog contains a link to documentation describing the option in further detail:

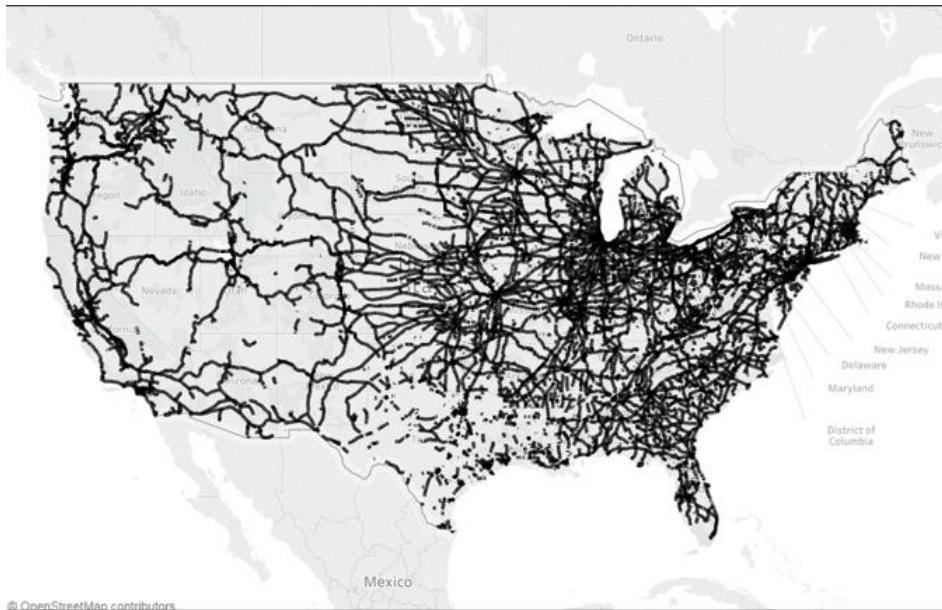


Leveraging spatial objects

- Spatial objects define geographic areas that can be as simple as a point and as complex as multi-sided polygons.
- This allows you to render everything from custom trade areas to rivers, roads, and historic boundaries of counties and countries.
- Spatial objects can be stored in spatial files and are supported by some relational databases as well.

Leveraging spatial objects

- For example, is a map of US railroads:



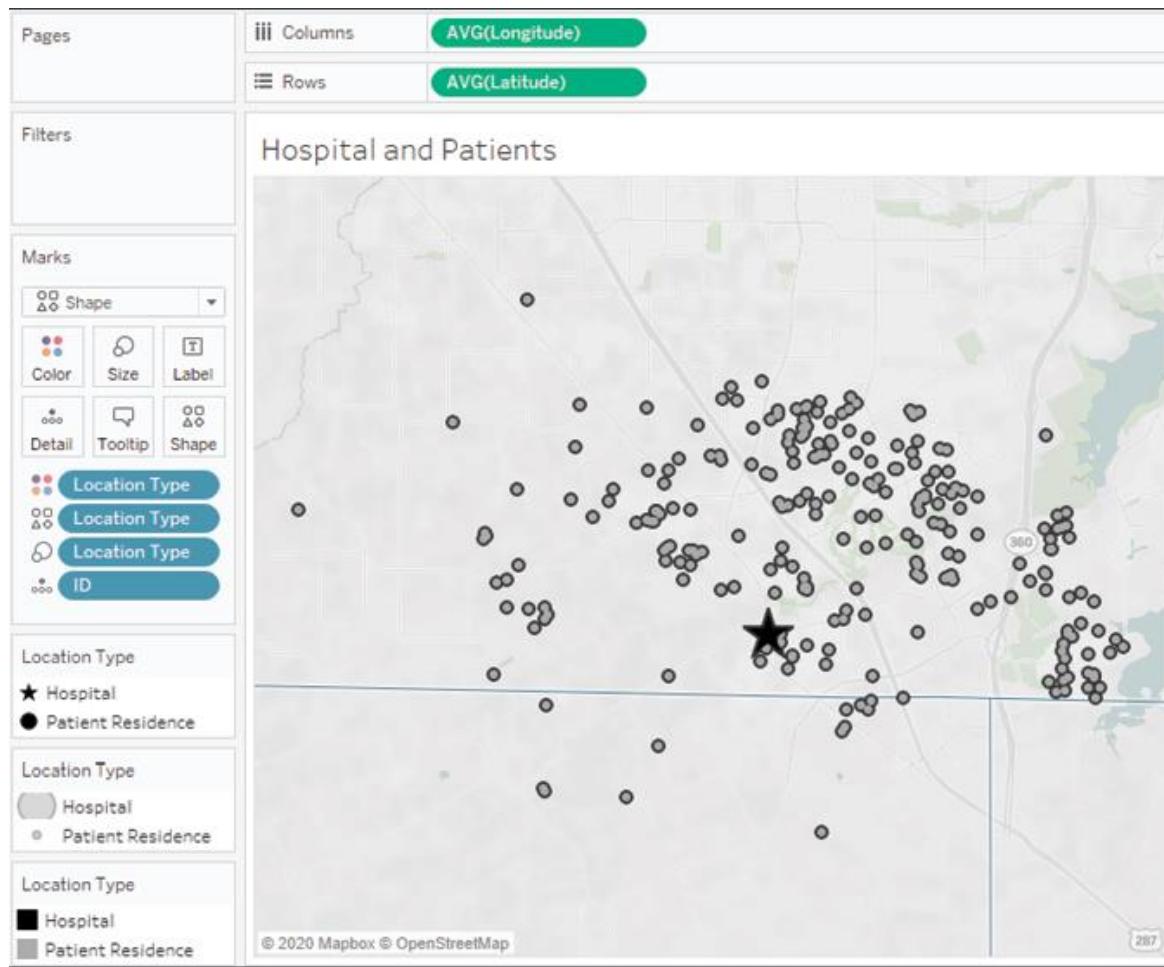
Leveraging spatial objects

The screenshot shows a spatial object viewer interface. At the top, there are navigation icons (back, forward, search, refresh) and a 'Connections' section with a single entry: 'tl_2015_us_rails Spatial file'. Below this is a 'Files' section containing 'tl_2015_us_rails.shp'. The main area displays a table titled 'tl_2015_us_rails.shp' with the following data:

Abc Linearid	Abc Fullname	Abc Mtfc...	Abc Geometry
110430696185	Union Pacific RR	R1011	LINESTRING
110430696262	Prescott and Northwestern RR	R1011	LINESTRING
110430744914	Norfolk Southern Rlwy	R1011	LINESTRING
110430744915	Norfolk Southern Rlwy	R1011	LINESTRING
110430744916	Norfolk Southern Rlwy	R1011	LINESTRING
110430744917	Norfolk Southern Rlwy	R1011	LINESTRING

Leveraging spatial functions

- Makeline() returns a line spatial object given two points.
- Makepoint() returns a point spatial object given two coordinates.
- Distance() returns the distance between two points in the desired units of measurement.
- Buffer() creates a circle around a point with a radius of the given distance. You may specify the units of measurement.



Leveraging spatial functions

There are numerous analytical questions we might ask.
Let's focus on these:

- How far is each patient from the hospital?
- How many patients fall within a given radius?
- Which patients are outside the radius?

Leveraging spatial functions

- We'll create a calculation called Hospital Latitude with the following code:

```
{FIXED : MIN(IF [Location Type] == "Hospital" THEN  
[Latitude] END)}
```

- And a corresponding calculation called Hospital Longitude with the following code:

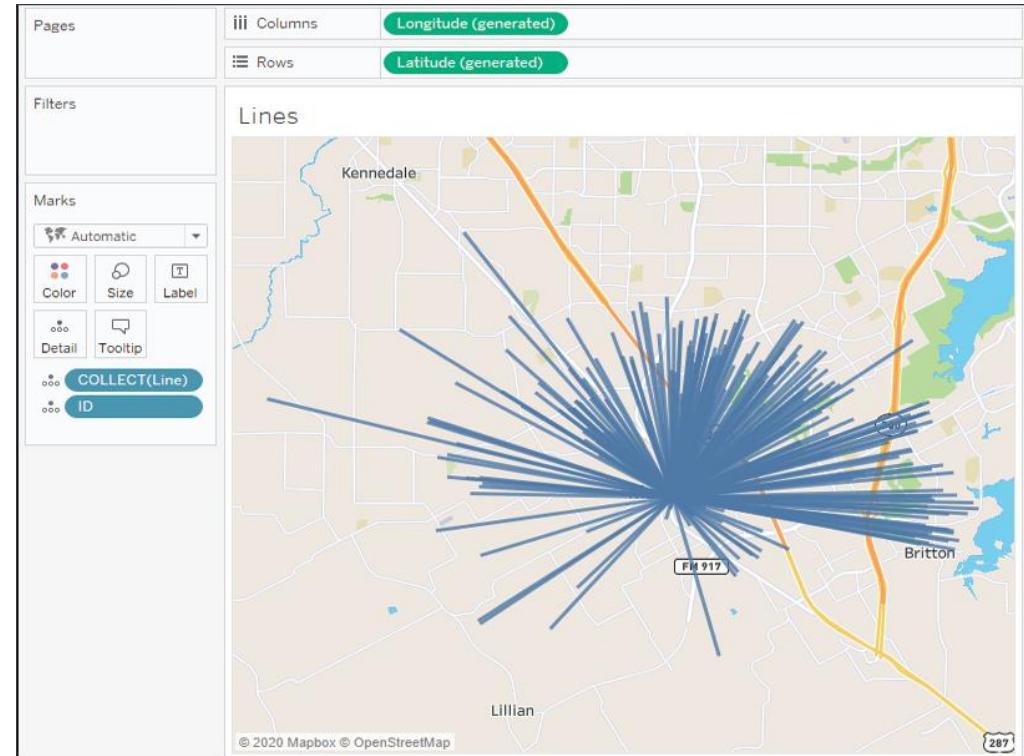
```
{FIXED : MIN(IF [Location Type] == "Hospital" THEN  
[Longitude] END)}
```

MAKELINE() and MAKEPOINT()

- As we consider these two functions, we'll create a calculated field to draw a line between the hospital and each patient.
- We'll name our calculation Line and write this code:
`MAKELINE(
 MAKEPOINT([Hospital Latitude], [Hospital
Longitude]),
 MAKEPOINT([Latitude], [Longitude])
)`

MAKELINE() and MAKEPOINT()

- In this case, each ID defines a separate line, so adding it to Detail on the Marks card splits the geospatial object into separate lines:



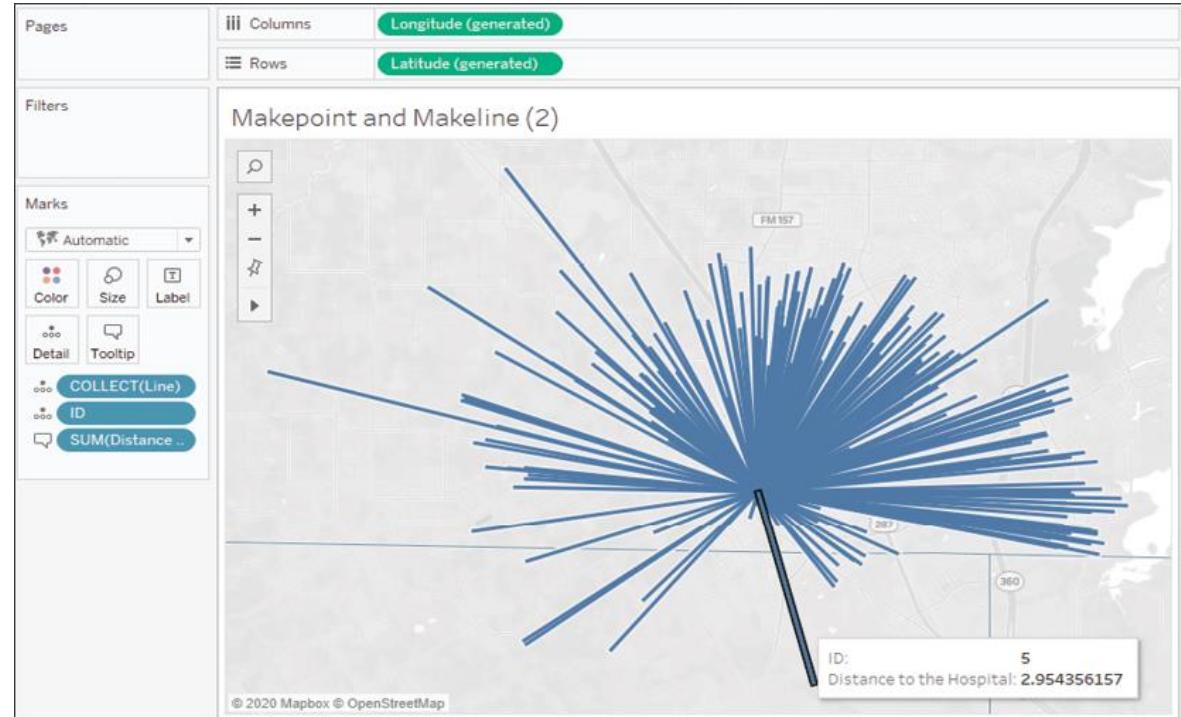
DISTANCE()

- The calculation itself is very similar to MAKELINE() and we might create a calculated field named Distance to the Hospital with the following code:

```
DISTANCE(  
    MAKEPOINT([Hospital Latitude], [Hospital  
Longitude]),  
    MAKEPOINT([Latitude], [Longitude]),  
    'mi'  
)
```

DISTANCE()

- We can place this calculation on Tooltip to see the distance covered by each line:



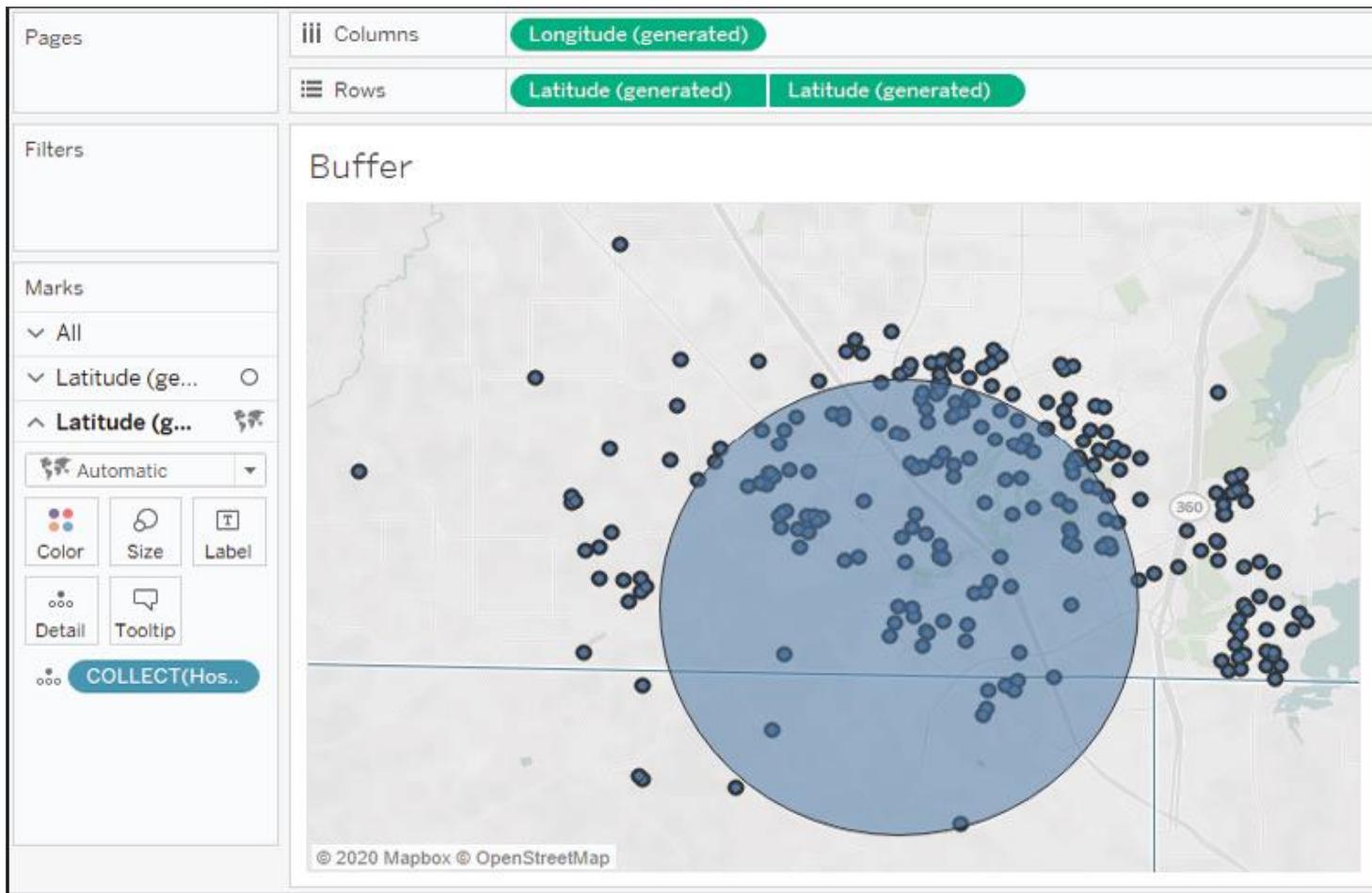
BUFFER()

- Buffer is similar to DISTANCE(), but the reverse.
- Rather than calculating a distance between two points, the BUFFER() function allows you to specify a point, a distance, and a unit of measurement to draw a circle with a radius of the specified distance around the point.

BUFFER()

- For example, you might want to visualize which patients fall within a 3-mile radius of the hospital.
- To do that, we'll create a calculated field named Hospital Radius, with the following code:

```
IF [Location Type] == "Hospital"  
THEN BUFFER(MAKEPOINT([Latitude], [Longitude]), 3, 'mi')  
END
```



Creating custom territories

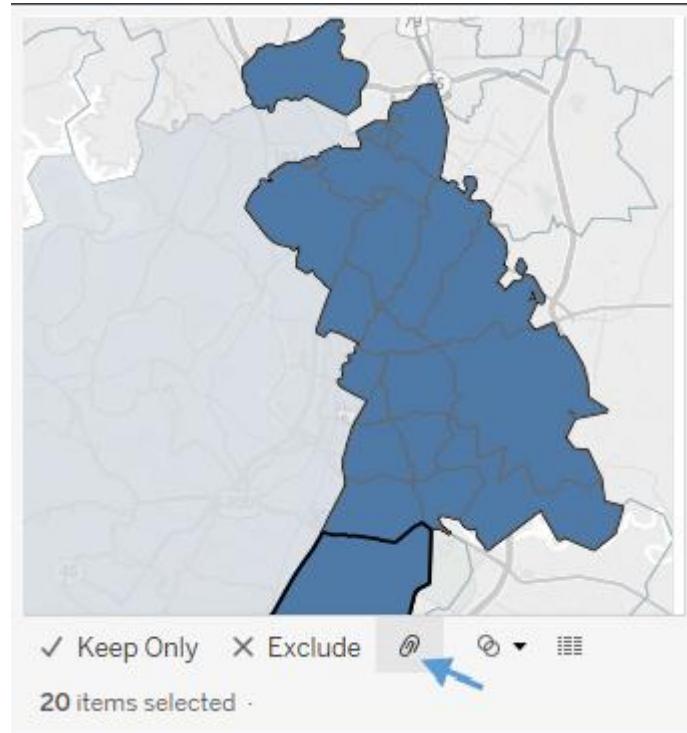
- Custom territories are geographic areas or regions that you create (or that the data defines) as opposed to those that are built in (such as country or area code).
- Tableau gives you two options for creating custom territories: ad hoc custom territories and field-defined custom territories. We'll explore these next.

Ad hoc custom territories

- You can create custom territories in an ad hoc way by selecting and grouping marks on a map.
- Simply select one or more marks, hover over one, and then use the Group icon.
- Alternately, right-click one of the marks to find the option.
- You can create custom territories by grouping by any dimension if you have latitude and longitude in the data or any geographic dimension if you are using Tableau's generated latitude and longitude.

Ad hoc custom territories

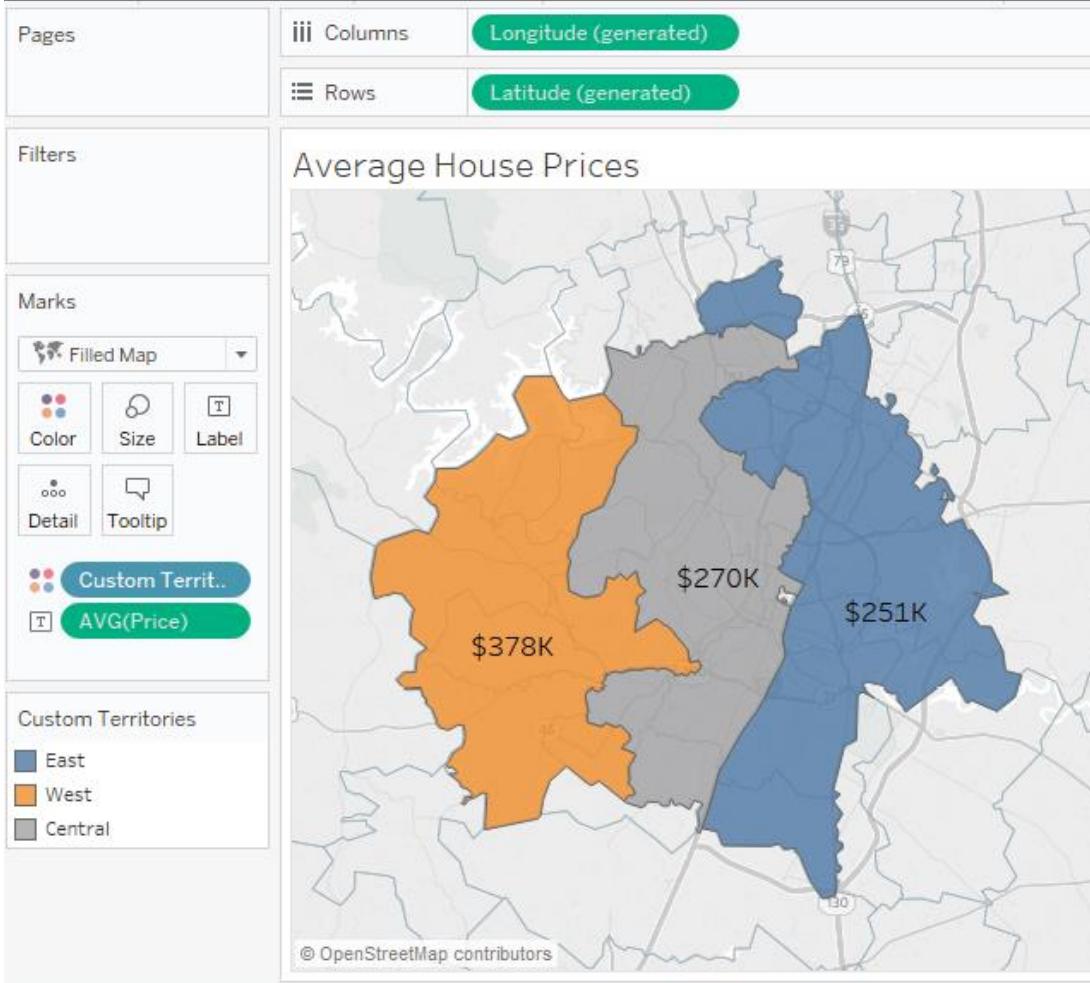
- Here, we'll consider an example using zip code:



Ad hoc custom territories

- You'll notice that Tableau creates a new field, Zip Code (group), in this example.
- The new field has a paperclip and globe icon in the data pane, indicating it is a group and a geographic field:





Field-defined custom territories

- Sometimes your data includes the definition of custom territories.
- For example, let's say your data had a field named Region that already grouped zip codes into various regions.
- That is, every zip code was contained in only one region.
- You might not want to take the time to select marks and group them manually.

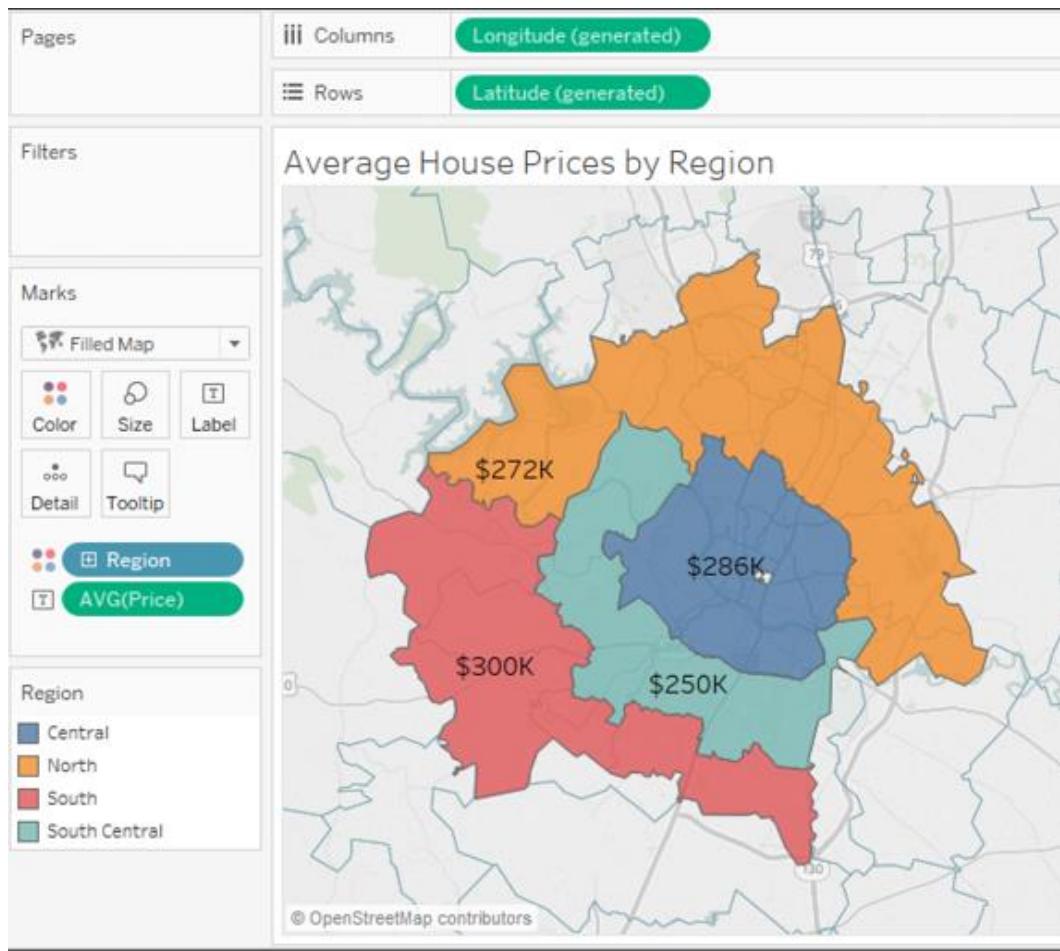
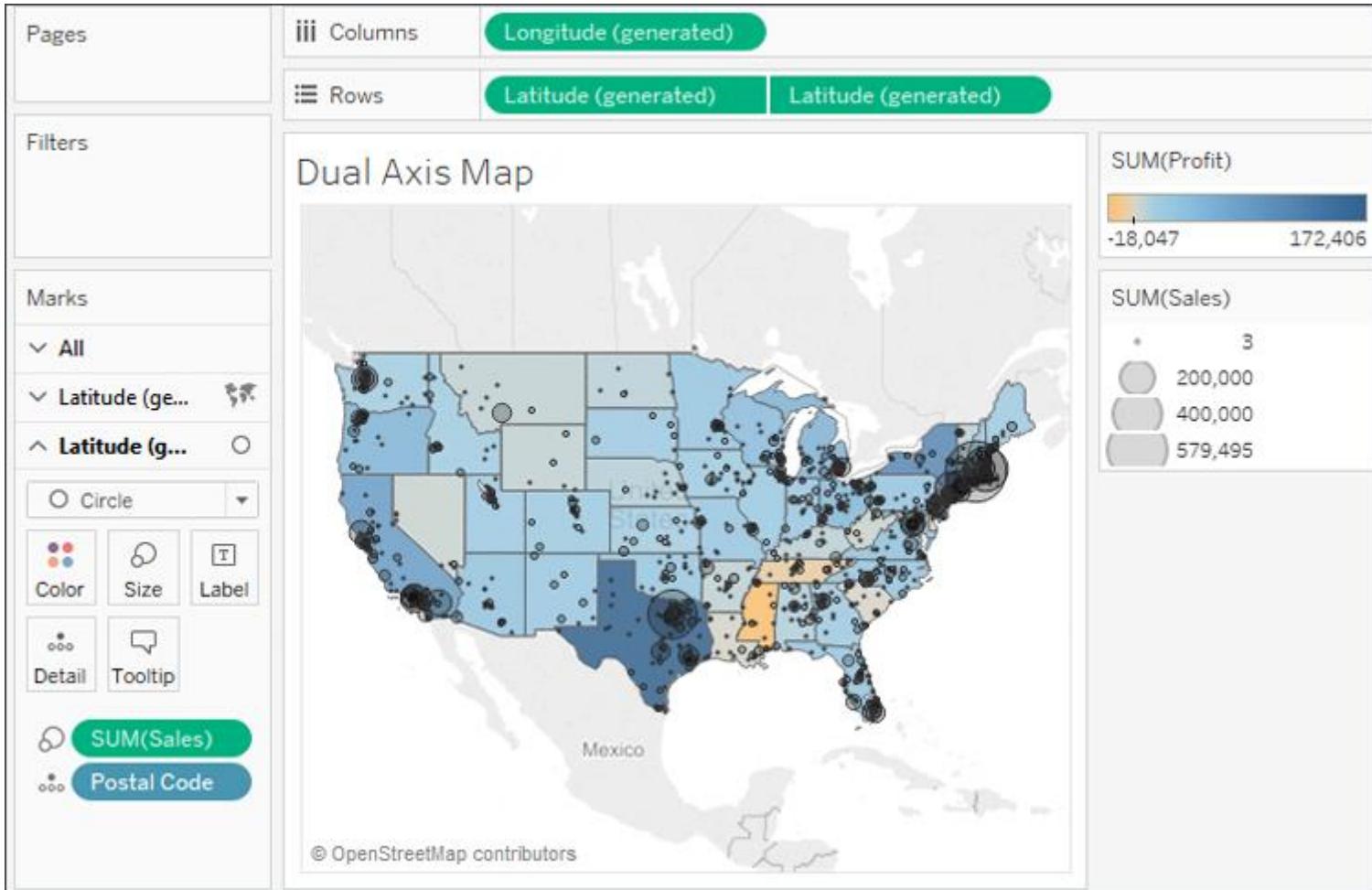
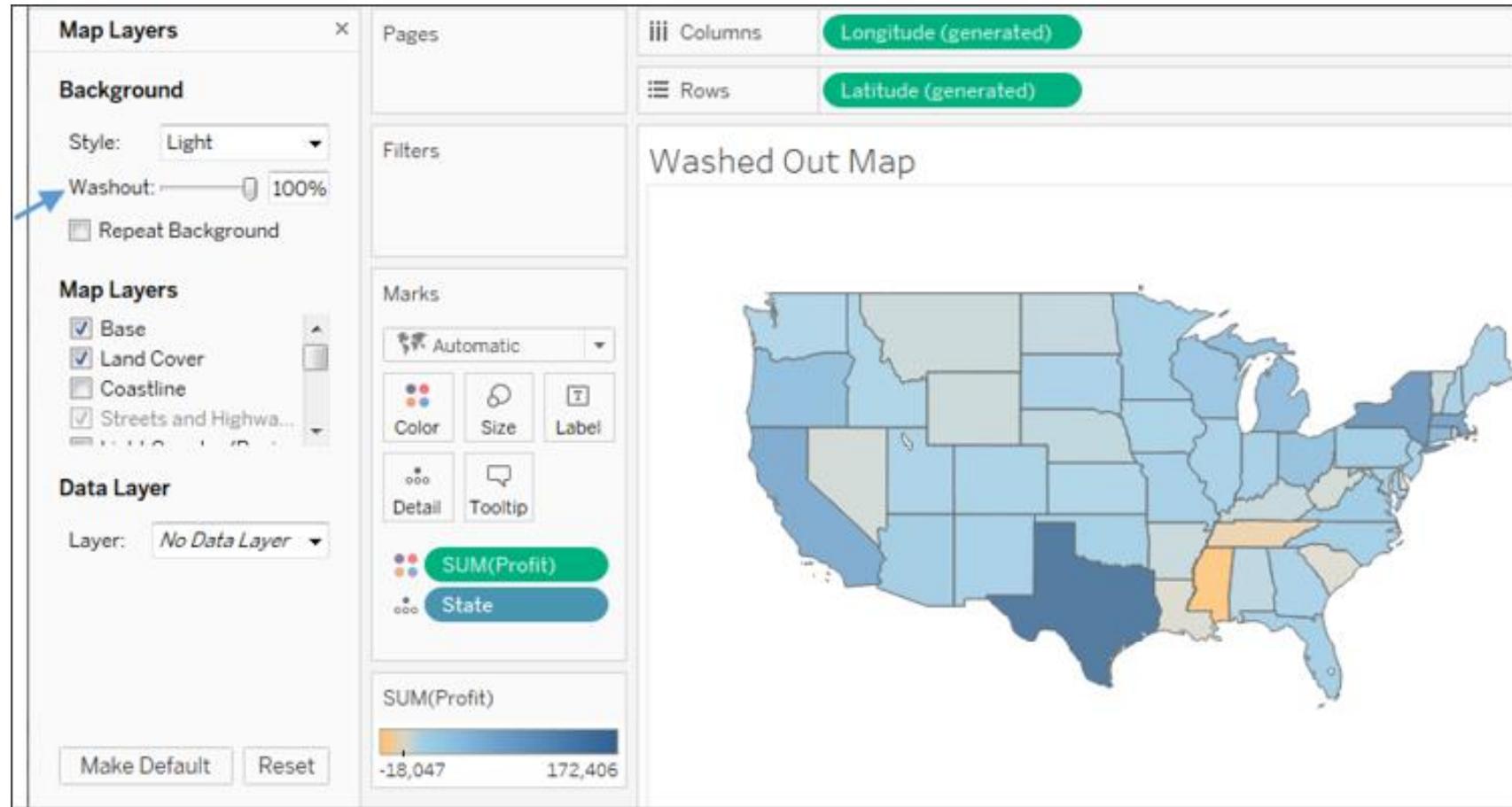


Tableau mapping – tips and tricks

There are a few other tips to consider when working with geographic visualizations:

- Use the top menu to select Map | Map Layers for numerous options for what layers of background to show as part of the map.
- Other options for zooming include using the mouse wheel, double-clicking, Shift + Alt + click, and Shift + Alt + Ctrl + click.
- You can click and hold for a few seconds to switch to pan mode.



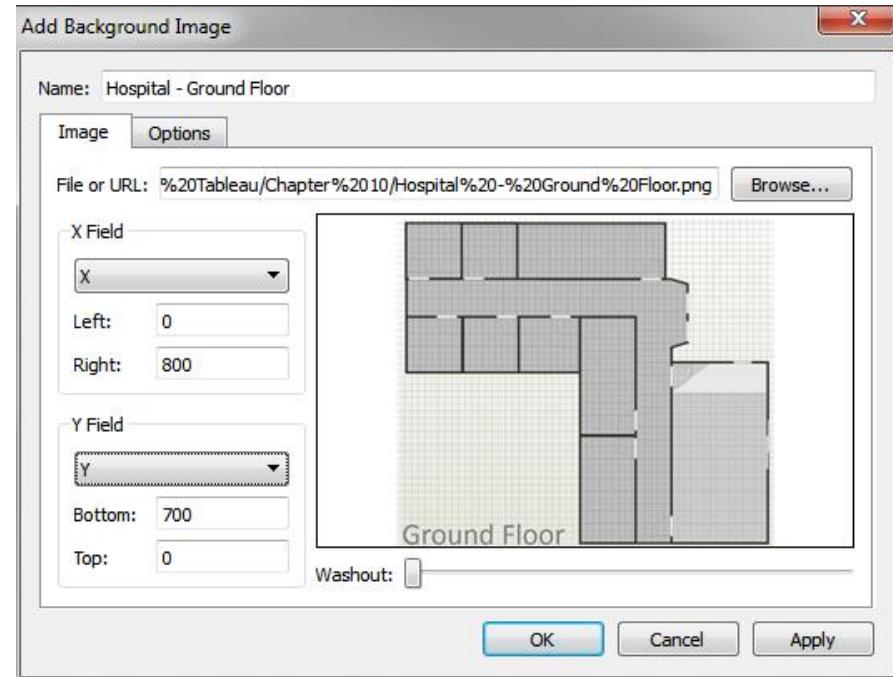


Plotting data on background images

- Background images allow you to plot data on top of any image.
- Consider the possibilities! You could plot ticket sales by seat on an image of a stadium, room occupancy on the floor plan of an office building, the number of errors per piece of equipment on a network diagram, or meteor impacts on the surface of the moon.

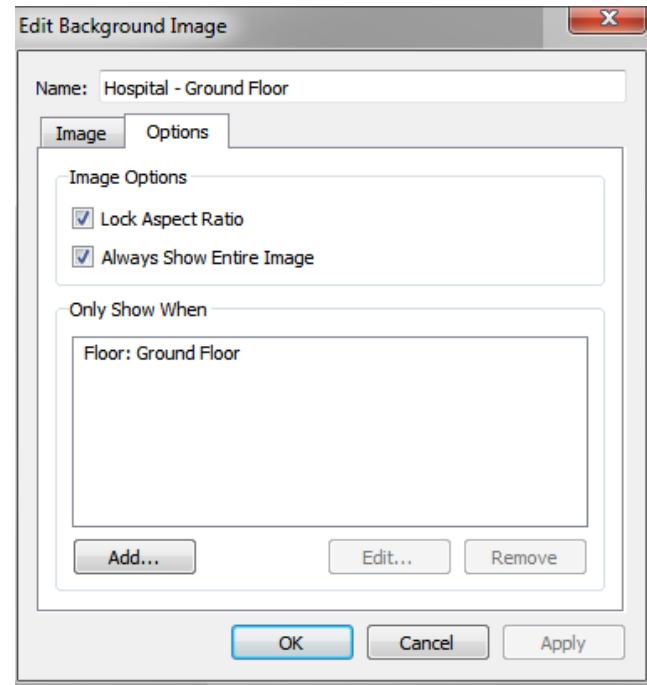
Plotting data on background images

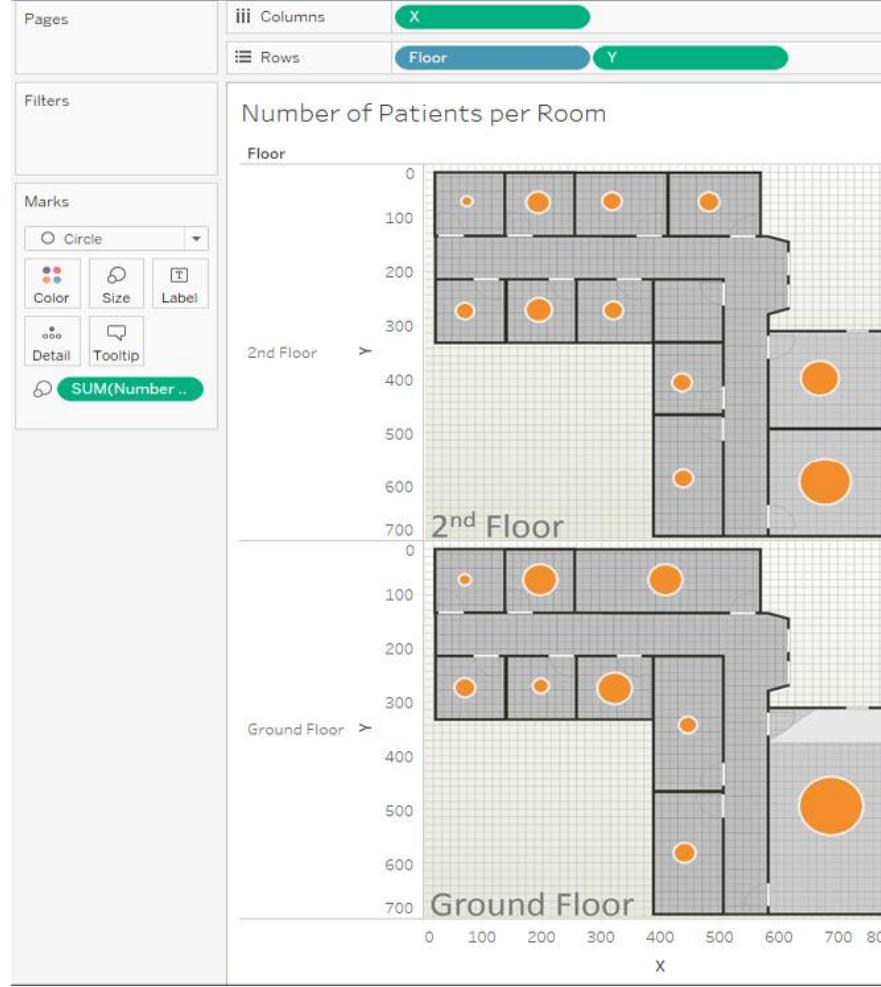
- We'll start with Hospital - Ground Floor.png, located in the lesson 12 directory:



Plotting data on background images

- We'll only want to show this blueprint for the ground floor, so switching to the Options tab, we'll ensure that the condition is set based on the data.
- We'll also make sure to check Always Show Entire Image:





Plotting data on background images

We've plotted circles with the size based on the number of patients in each room. We could clean up and modify the visualization in various ways:

- Hide the x and y axes (right-click the axis and uncheck Show Header)
- Hide the header for Floor, as the image already includes the label
- Add Floor to the Filter shelf so that the end user can choose to see one floor at a time

Summary

- We've covered a lot of ground in this lesson! The basics of visualizing maps are straightforward, but there is a lot of power and possibility behind the scenes.
- From using your own geospatial data to leveraging geospatial objects and functions, you have a lot of analytical options.
- Creating custom territories and plotting data on background images expand your possibilities even further.

COMPLETE LAB 12

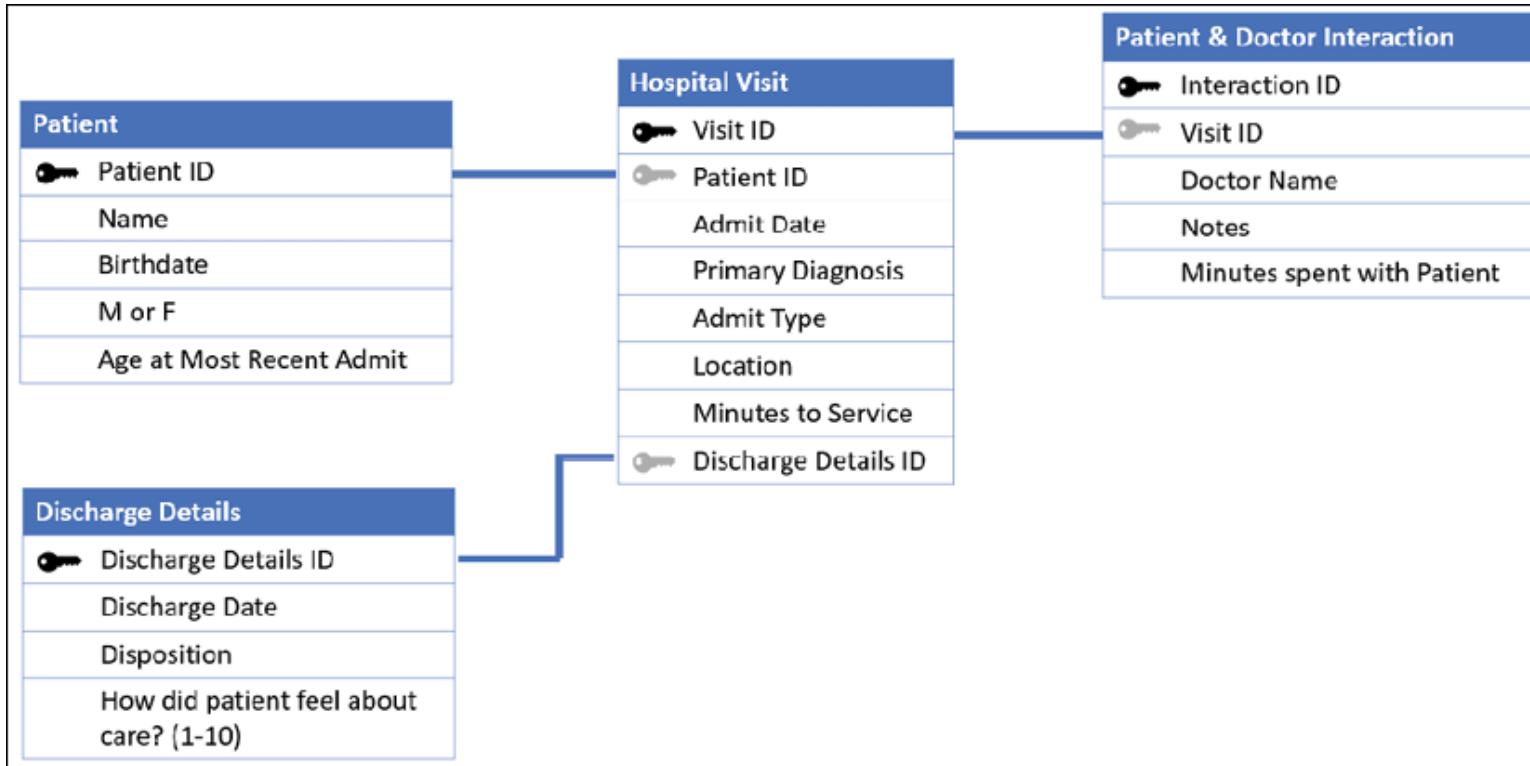
13. Understanding the Tableau Data Model, Joins, and Blends

Tableau Data Model, Joins, and Blends

In this lesson, we'll cover the following topics:

- Explanation of the sample data used in this lesson
- Exploring the Tableau data model
- Using joins
- Using blends
- When to use a data model, joins, or blends

Explanation of the sample data used in this lesson



Exploring the Tableau data model

- You'll find the data model as a new feature in Tableau 2020.2 and later.
- Every data source will use the data model.
- Data sources created in previous versions will be updated to the data model but will be contained in a single object, so, functionally, they will work in the same way as the previous version.

Creating a data model

The screenshot shows the Microsoft Power BI Data Model view. On the left, there's a sidebar with a 'Connections' section containing a single item: 'Hospital Visits' (Microsoft Excel). Below it is a 'Sheets' section listing several tables: 'Discharge Details', 'Hospital Visit', 'Patient', 'Patient&Doctor Interaction', and 'New Union'. A checkbox labeled 'Use Data Interpreter' is present, with a note explaining that it might clean the Excel workbook. The main workspace is titled 'Hospital Visits' and features a large grid icon with the text 'Drag tables here' below it. At the bottom, there are buttons for 'Sort fields', 'Data source order', and checkboxes for 'Show aliases' and 'Show hidden fields', along with a 'rows' button.

Hospital Visit+ (Hos...)

Connection



Live

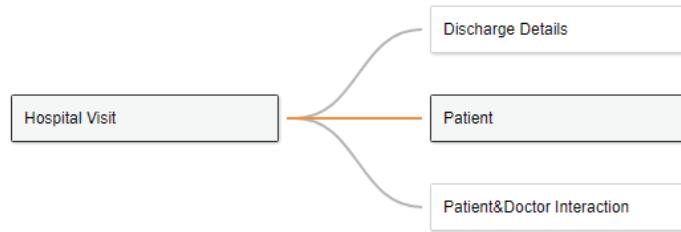


Extract

Filters

0

| Add



Edit Relationship

How do relationships differ from joins? [Learn more](#)

Hospital Visit	Patient
# Hospital Visit Visit ID	# Patient ID # Patient ID (Patient)

+ Add more fields

Performance Options

	101	2/20/2020	Pneumonia	Emergency	ER
5	104	2/20/2020	Fracture of Tibia	Emergency	ER
6	102	2/21/2020	Pneumonia	Emergency	ER
7	105	2/28/2020	Normal Delivery	Newborn	Matern
8	106	3/1/2020	Laceration	Emergency	ER
9	103	3/15/2020	Anemia	Elective	Lab
10	101	3/20/2020	CHE	Urgent	Triage

Creating a data model

- Notice the Performance Options drop-down menu in the relationship editor, as shown here:

Edit Relationship

How do relationships differ from joins? [Learn more](#)

Hospital Visit	Patient
# Patient ID	# Patient ID (Patient)
+ Add more fields	
^ Performance Options	
These settings help Tableau optimize queries during analysis. The default settings are recommended, if you aren't sure what to choose. Learn more	
Cardinality	Many
Referential Integrity	Some records match

[Revert to Default](#)

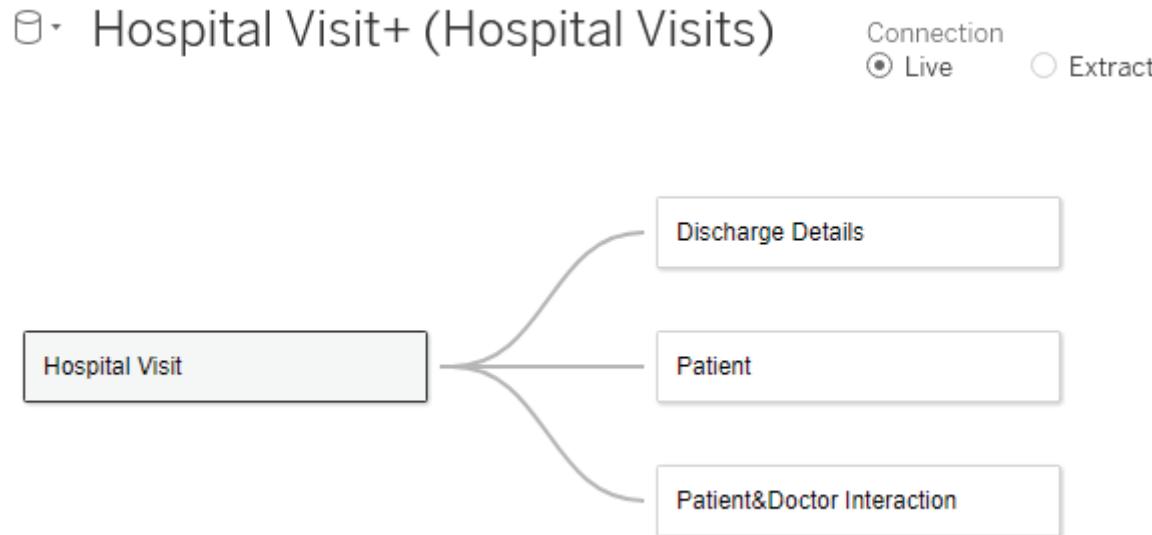
Layers of the data model

A data model consists of two layers:

- The logical layer: A semantic layer made up of logical tables or objects that are related. Each logical table might be made up of one or more physical tables.
- The physical layer: A layer made up of the physical tables that come from the underlying data source. These tables may be joined or unioned together with conventional joins or unions or created from custom SQL statements.

Layers of the data model

- Consider the following screenshot of a canvas containing our four tables:



Layers of the data model

- Double-click on the Hospital Visit table on the canvas, and you'll see another layer beneath the logical layer:



Using the data model

- For the most part, working with the data model will be relatively intuitive.
- If you've worked with previous Tableau versions, you'll notice some slight interface changes, and there are a few data model behaviors you should learn to expect.
- Once you are comfortable with them, your analysis will exceed expectations!

Data Analytics

Hospital Visit+ (Hospital Visits)

Search

Tables

- Discharge Details
 - Discharge Date
 - Discharge Details ID (Discharge Details)
 - Disposition
 - How did patient feel about care? (1-10)
 - Discharge Details (Count)
- Hospital Visit
 - Admit Date
 - Admit Type
 - Discharge Details ID
 - Location
 - Patient ID
 - Primary Diagnosis
 - Visit ID
 - Minutes to Service
 - Hospital Visit (Count)
- Patient
 - Birthdate
 - M or F
 - Name
 - Patient ID (Patient)
 - Age at Most Recent Admit
 - Patient (Count)
- Patient&Doctor Interaction
 - Doctor
 - Interaction ID
 - Notes
 - Visit ID (Patient&Doctor Interaction)
 - Minutes spent with Patient
 - Patient&Doctor Interaction (Count)
- Measure Names
- Measure Values

The new data pane interface

- One thing you may notice is the difference in the Data pane, which will look something like this:

Data model behaviors

- In the Analysis tab of the Starter workbook, experiment with creating different visualizations.
- Especially note dimensions, what values are shown, and how measures are aggregated.
- We'll walk through a few examples to illustrate (which you can replicate in the Starter workbook or examine in the Complete workbook).

Data model behaviors

The screenshot shows the Tableau Data Model Behaviors interface. On the left, there are three panels: 'Pages' (empty), 'Columns' (empty), and 'Rows'. The 'Rows' panel has a selected tab labeled 'Name'. In the center, there is a title 'How Many Visits for each Patient?' followed by a table with two columns: 'Name' and a numerical count.

Name	Count
Corinna	2
Darrel	0
Hillary	3
John	1
Leonard	2
Mary	1
Saeeda	0
Shantea	0
Takila	0
Toria	1

The 'Marks' panel on the left contains settings for 'Automatic' marks, with options for Color, Size, Text, Detail, and Tooltip. The 'Color' and 'Text' buttons are highlighted with green rounded rectangles, while 'Detail' and 'Tooltip' are highlighted with blue rounded rectangles. The 'CNT(Hospital Visit)' button is also highlighted with a green rounded rectangle.

Data model behaviors

The screenshot shows the Tableau Data Model Behaviors interface. On the left, there are three panels: 'Pages' (empty), 'Filters' (empty), and 'Marks'. The 'Marks' panel has dropdown menus for 'Automatic' (selected), 'Color', 'Size', 'Text', 'Detail', and 'Tooltip'. A green button at the bottom of the Marks panel is labeled 'CNT(Hospital Visit)'. On the right, the main workspace displays a table titled 'How many visits? Patients & Diagnosis'. The table has two columns: 'Name' and 'Primary Diagnos..'. The data is as follows:

Name	Primary Diagnos..
Corinna	Anemia
	Iron Deficiency
Hillary	CHF
	Pneumonia
John	Fracture of Tibia
Leonard	Fever
	Pneumonia
Mary	Normal Delivery
Toria	Laceration

Pages

iii Columns Measure Names

Rows Name Primary Diagnosis

Filters

Measure Names

Marks

Automatic

Name	Primary Diagnosis	Age at Most Recent Admit	Count of Hospital Visit
Corinna	Anemia	41.00	1.00
	Iron Deficiency	41.00	1.00
Darrel	Null		0.00
Hillary	CHF	72.00	2.00
	Pneumonia	72.00	1.00
John	Fracture of Tibia	18.00	1.00
Leonard	Fever	22.00	1.00
	Pneumonia	22.00	1.00
Mary	Normal Delivery	30.00	1.00
Saeeda	Null		0.00
Shantea	Null		0.00
Takila	Null		0.00
Toria	Laceration	26.00	1.00

Measure Values

SUM(Age at Most Recent ..)

CNT(Hospital Visit)

Pages

iii Columns Measure Names

Rows Name Primary Diagnosis

Filters

Measure Names

Marks

Automatic

Color Size Text

Detail Tooltip

Measure Values

Measure Values

SUM(Age at Most Recent ..)

CNT(Hospital Visit)

How old and how many visis? with subtotals

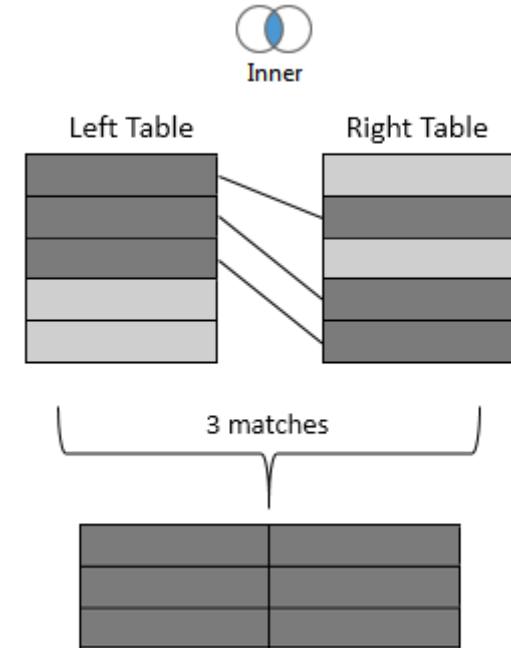
Name	Primary Diagnosis	Age at Most Recent Admit	Count of Hospital Visit
Corinna	Anemia	41.00	1.00
	Iron Deficiency	41.00	1.00
	Total	41.00	2.00
Darrel	Null	0.00	
	Total		0.00
Hillary	CHF	72.00	2.00
	Pneumonia	72.00	1.00
	Total	72.00	3.00
John	Fracture of Tibia	18.00	1.00
	Total	18.00	1.00
Leonard	Fever	22.00	1.00
	Pneumonia	22.00	1.00
	Total	22.00	2.00
Mary	Normal Delivery	30.00	1.00
	Total	30.00	1.00

Using joins

- A join at the physical level is a row-by-row matching of the data between tables.
- We'll look at some different types of joins and then consider how to leverage them in the physical layer of a data model.

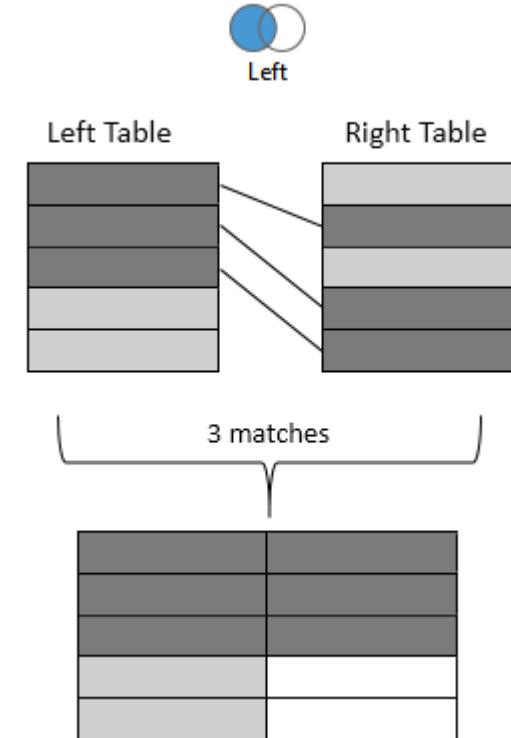
Types of joins

- Inner: Only records that match the join condition from both the table on the left and the table on the right will be kept.
- In the following example, only three matching rows are kept in the results:



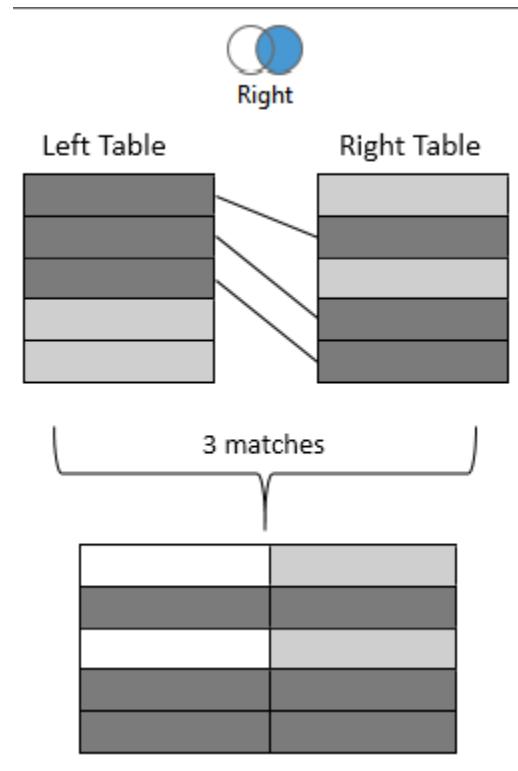
Using joins

- In the following example, the five rows from the left table are kept, with NULL results for any values in the right table that were not matched:



Using joins

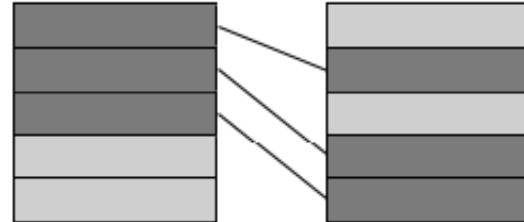
- In the following example, the five rows from the right table are kept, with NULL results for any values from the left table that were not matched:



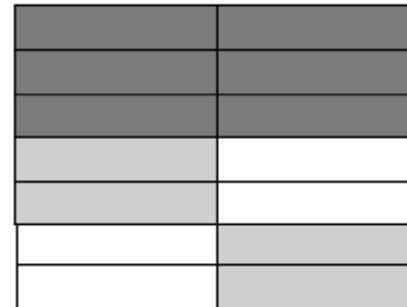


Full Outer

Left Table Right Table



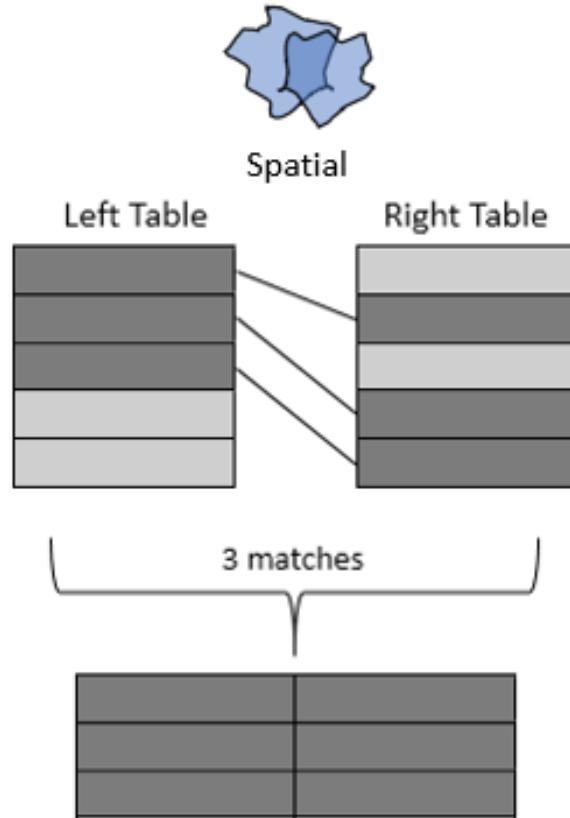
3 matches



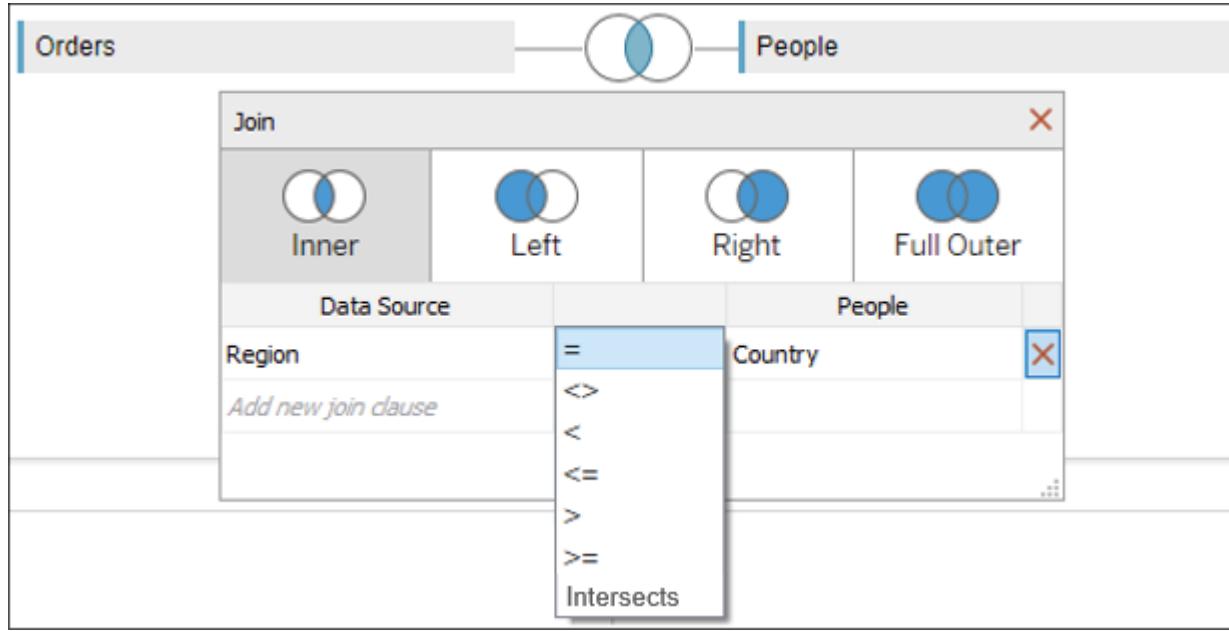
- In the following example, all rows are kept from both sides with NULL values where matches were not found:

Using joins

- Records will be kept for any records where the spatial object in one table overlaps with the spatial object specified for the other:



Using joins



Patient	
🔑	Patient ID
	Name
	Birthdate
	Age at Most Recent Admit

Hospital Visit	
🔑	Visit ID
🔑	Patient ID
	Admit Date
	Primary Diagnosis
	Admit Type
	Location
	Minutes to Service
🔑	Discharge Details ID

Discharge Details	
🔑	Discharge Details ID
	Discharge Date
	Disposition
	How did patient feel about care? (1-10)

Joining tables

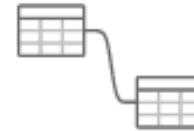
• Hospital Visit (Hospital Visits (Joins))

Connection

Live

Extract

Hospital Visit



Need more data?

Drag tables here to relate them. [Learn more](#)

Joining tables

⌚ Hospital Visit (Hospital Visits (Joins))

Connection

Live

Extract

Hospital Visit is made of 1 table. ⓘ

X

Hospital Visit

Joining tables

Hospital Visit is made of 2 tables. ⓘ



Join

Join	Discharge Detail
Inner	Discharge Detail
Left	Discharge Detail
Right	Discharge Detail
Full Outer	Discharge Detail

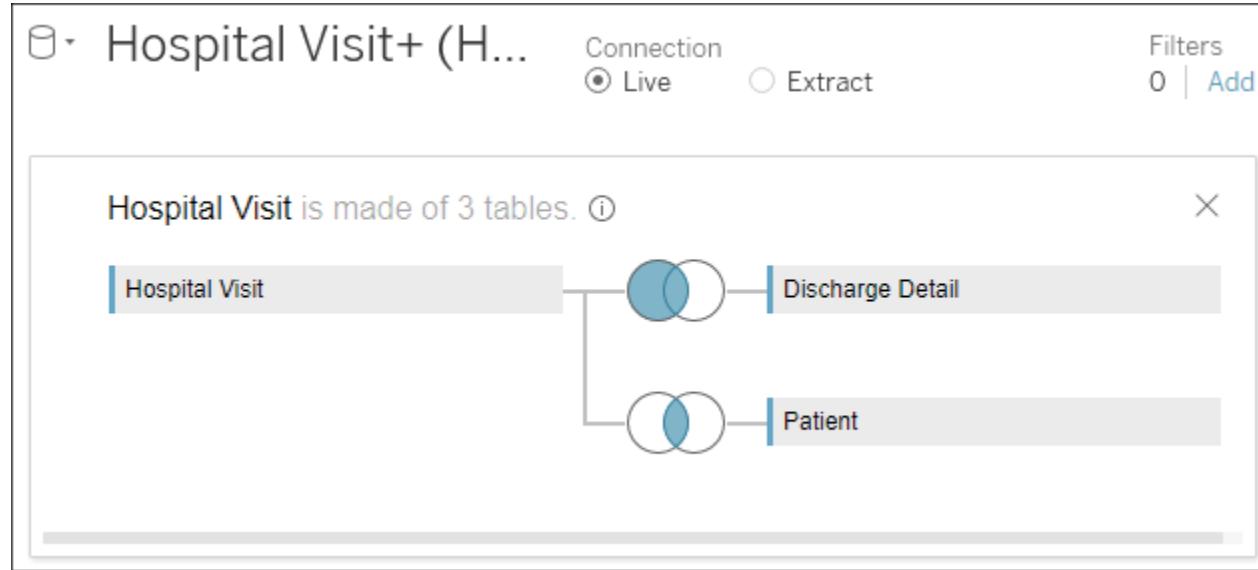
Data Source

Discharge Details ID	=	Discharge Details ID (...)
----------------------	---	----------------------------

Add new join clause

Joining tables

- Ultimately, the physical layer for Hospital Visit will look like this:



Joining tables

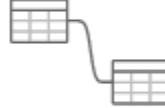
- It remains a single object in the logical layer of the data model and looks like this:

⊖ Hospital Visit+ (H...)

Connection
 Live Extract

Filters
0 | [Add](#)

 Hospital Visit



Need more data?

Drag tables here to relate them. [Learn more](#)

Other join considerations

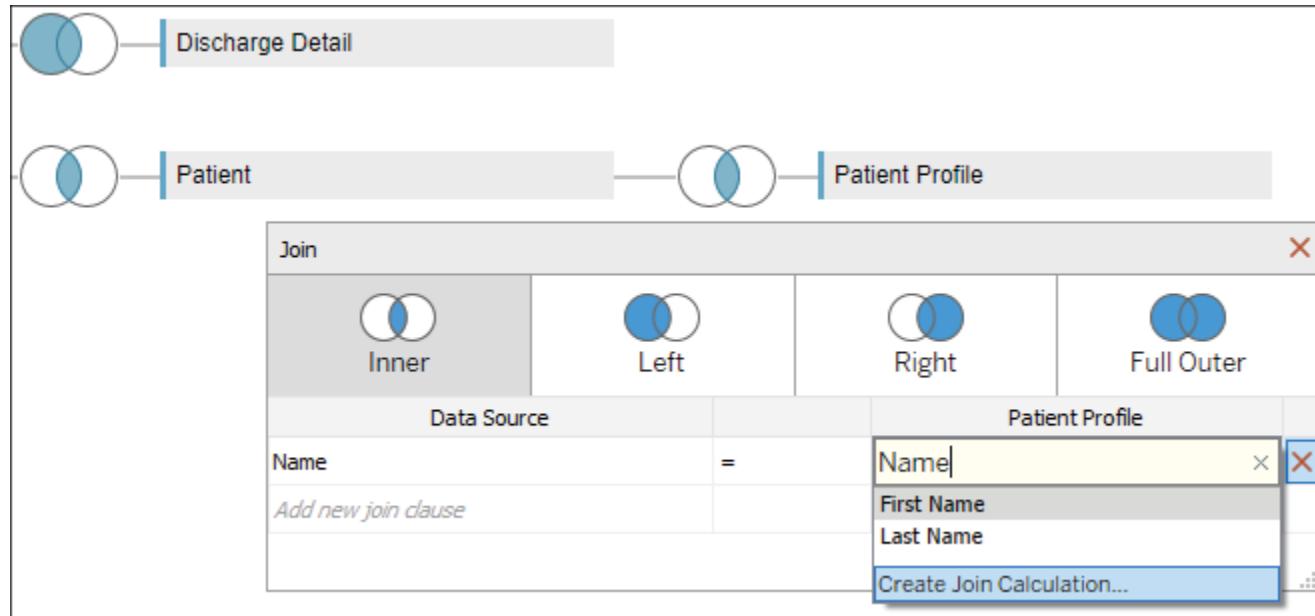
- We conclude this section with some further possibilities to leverage joins, as well as a caution regarding a potential problem that can arise from their use.

Join calculations

- In the previous example, we noted that Tableau joins row-by-row based on fields in the data.
- You may come across cases where you need to join based on values that are not present in the data but can be derived from the existing data.
- For example, imagine that there is a Patient Profile table that would add significant value to your dataset.
- However, it lacks a Patient ID and only has First Name and Last Name fields.

Join calculations

- Create Join Calculation:

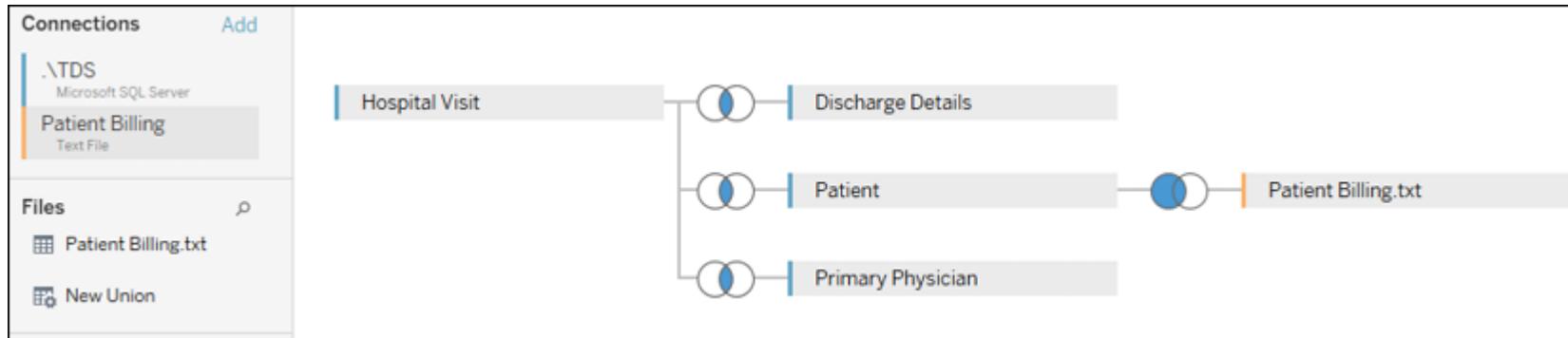


Cross-database joins

- With Tableau, you have the ability to join (at the row level) across multiple different data connections. Joining across different data connections is referred to as a cross-database join.
- For example, you can join SQL Server tables with text files or Excel files, or join tables in one database with tables in another, even if they are on a different server.
- This opens up all kinds of possibilities for supplementing your data or analyzing data from disparate sources.

Cross-database joins

- You would be able to accomplish this by adding the text file as a data connection and then joining it to the existing tables, as follows:



The unintentional duplication of data

- Let's say you have a Visit table like this:

Visit ID	Patient Name	Doctor ID
1	Kirk	1
2	Picard	2
3	Sisko	3

The unintentional duplication of data

- And a Doctor table like this:

Doctor ID	Doctor Name
1	McCoy
2	Crusher
3	Bashir
2	Pulaski

The unintentional duplication of data

- Such a join would result in the following dataset:

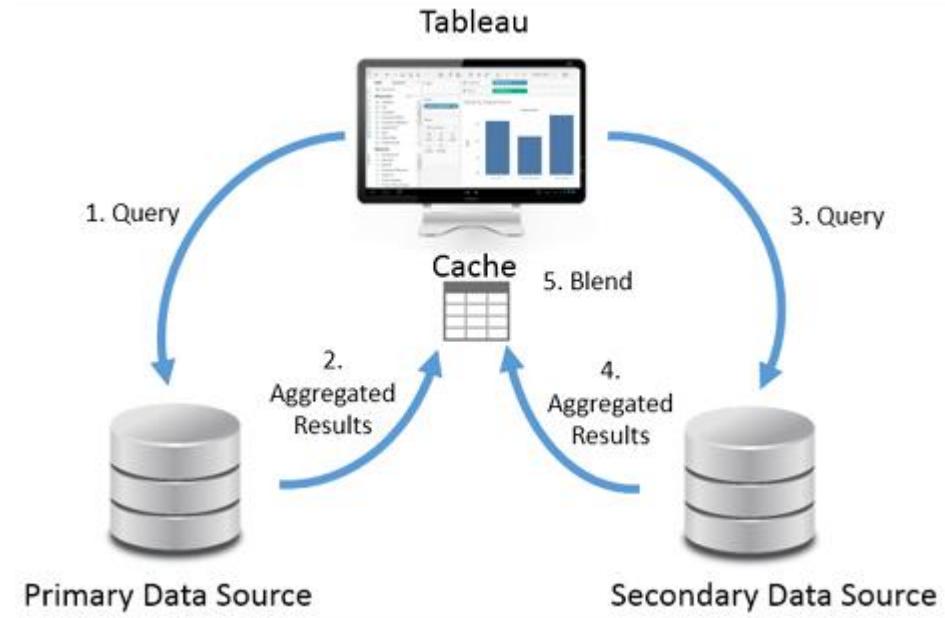
Visit ID	Patient Name	Doctor ID	Doctor Name
1	Kirk	1	McCoy
2	Picard	2	Crusher
3	Sisko	3	Bashir
2	Picard	2	Pulaski

Using blends

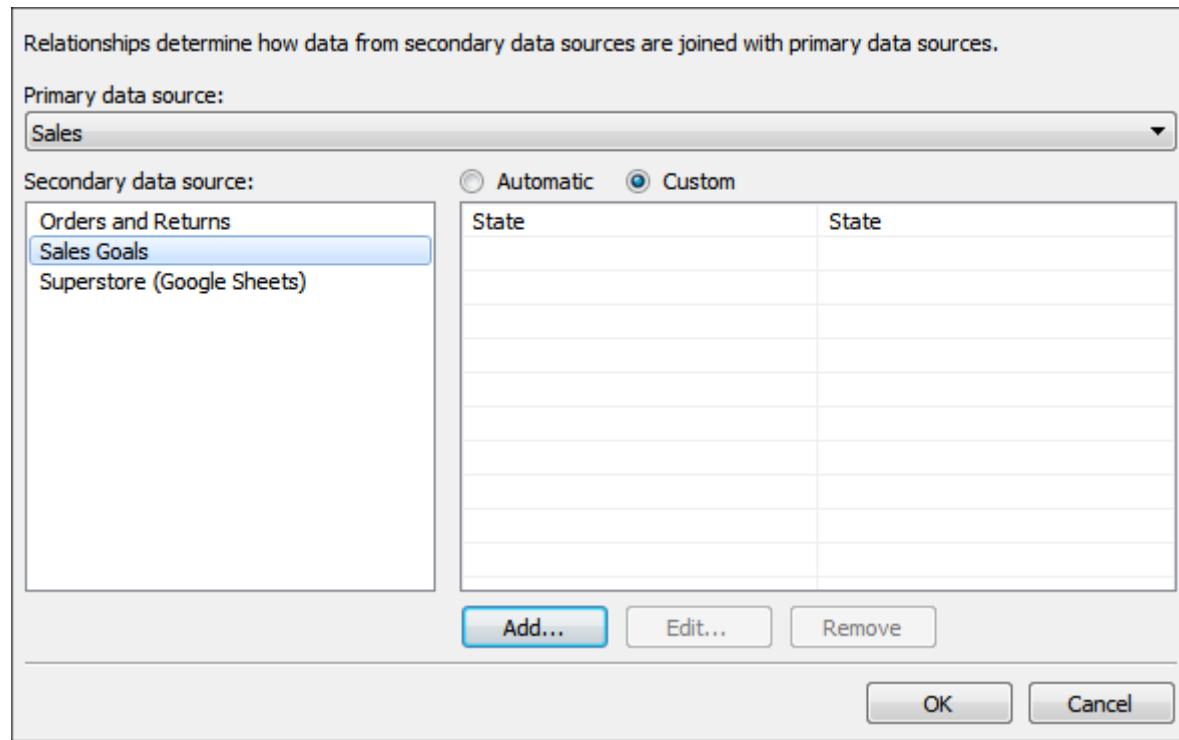
- Data blending allows you to use data from multiple data sources in the same view. Often, these sources may be of different types.
- For example, you can blend data from Oracle with data from Excel.
- You can blend Google Analytics data with a spatial file.
- Data blending also allows you to compare data at different levels of detail

Using blends

- A simple data blending process involves several steps, as shown in the following diagram:



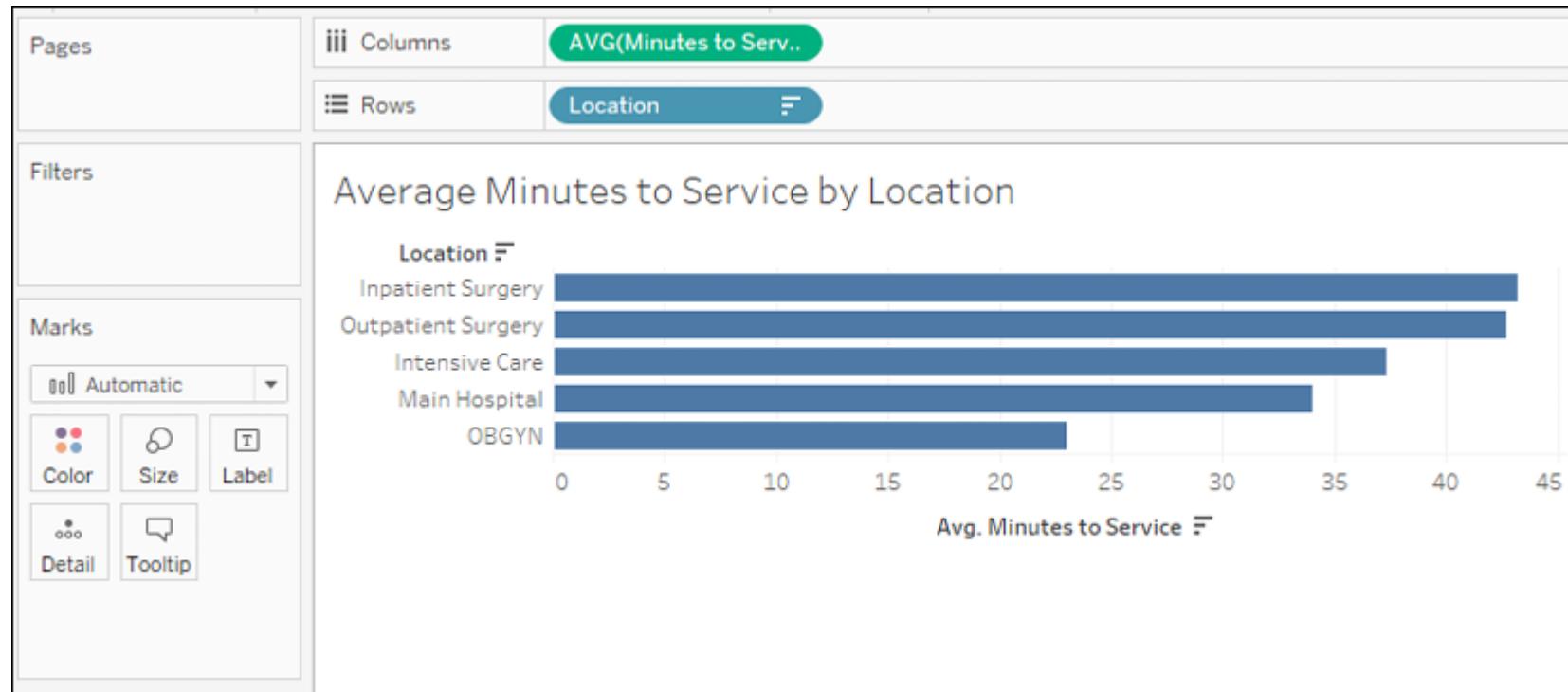
Using blends



A blending example

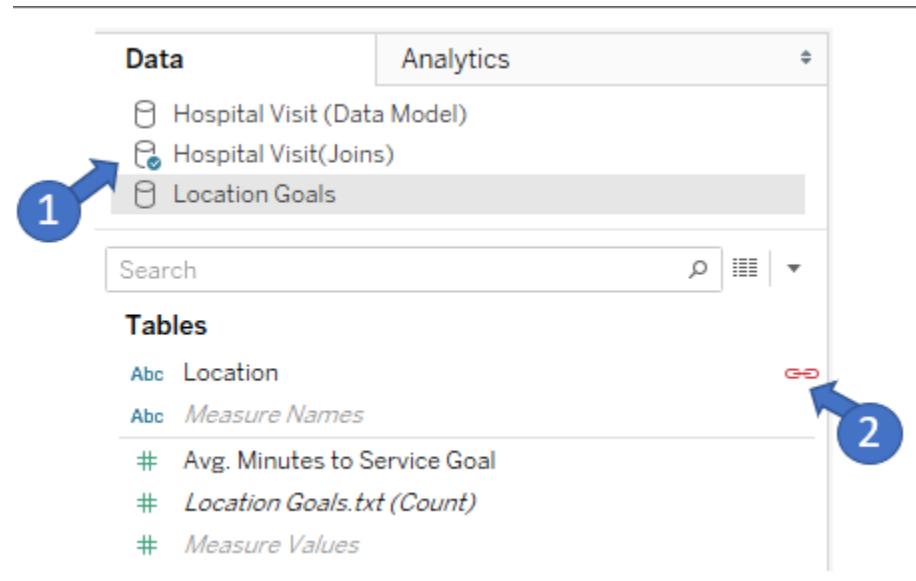
Location	Avg. Minutes to Service Goal
Inpatient Surgery	30
Outpatient Surgery	40
ICU	30
OBGYN	25
Lab	120

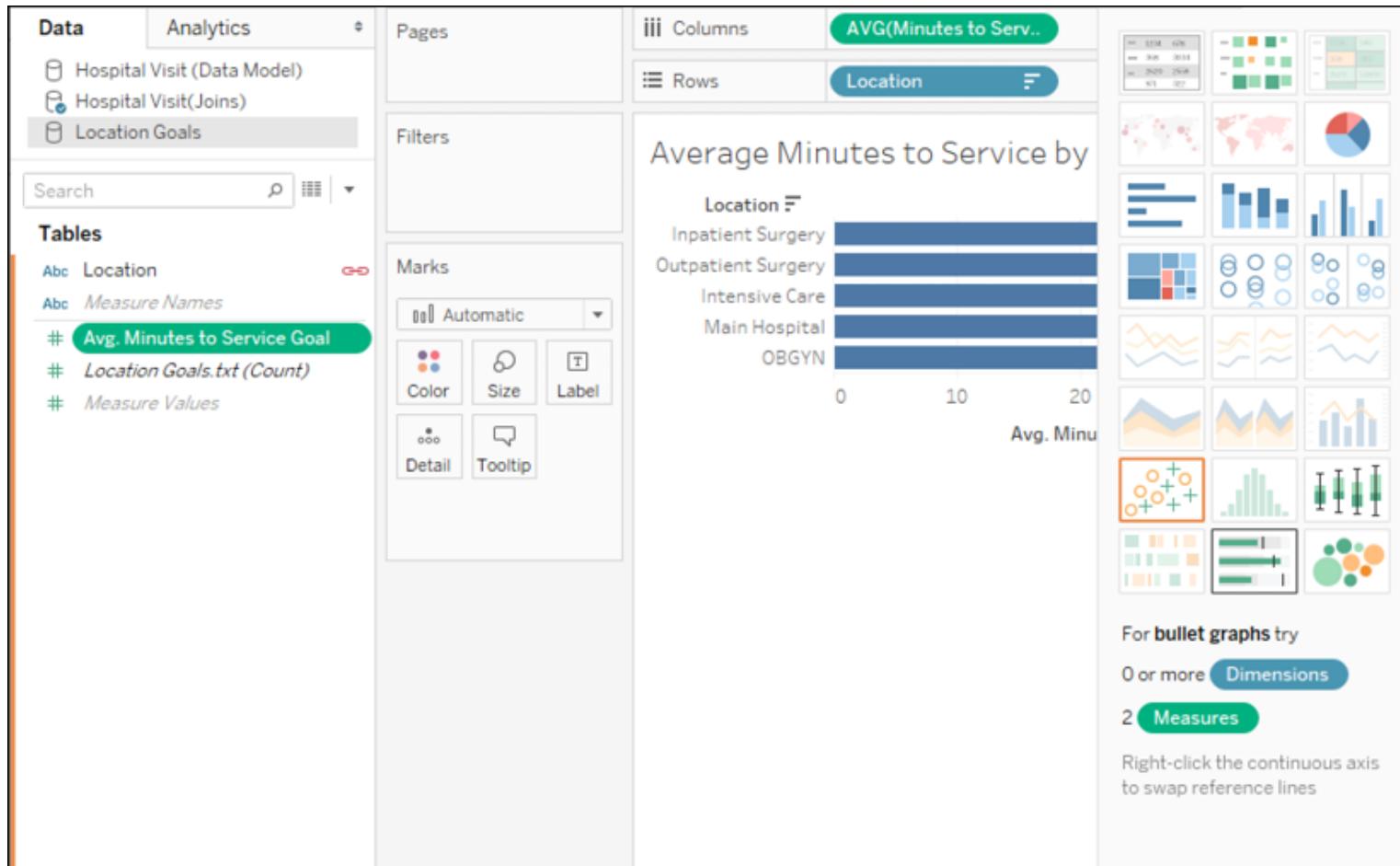
A blending example



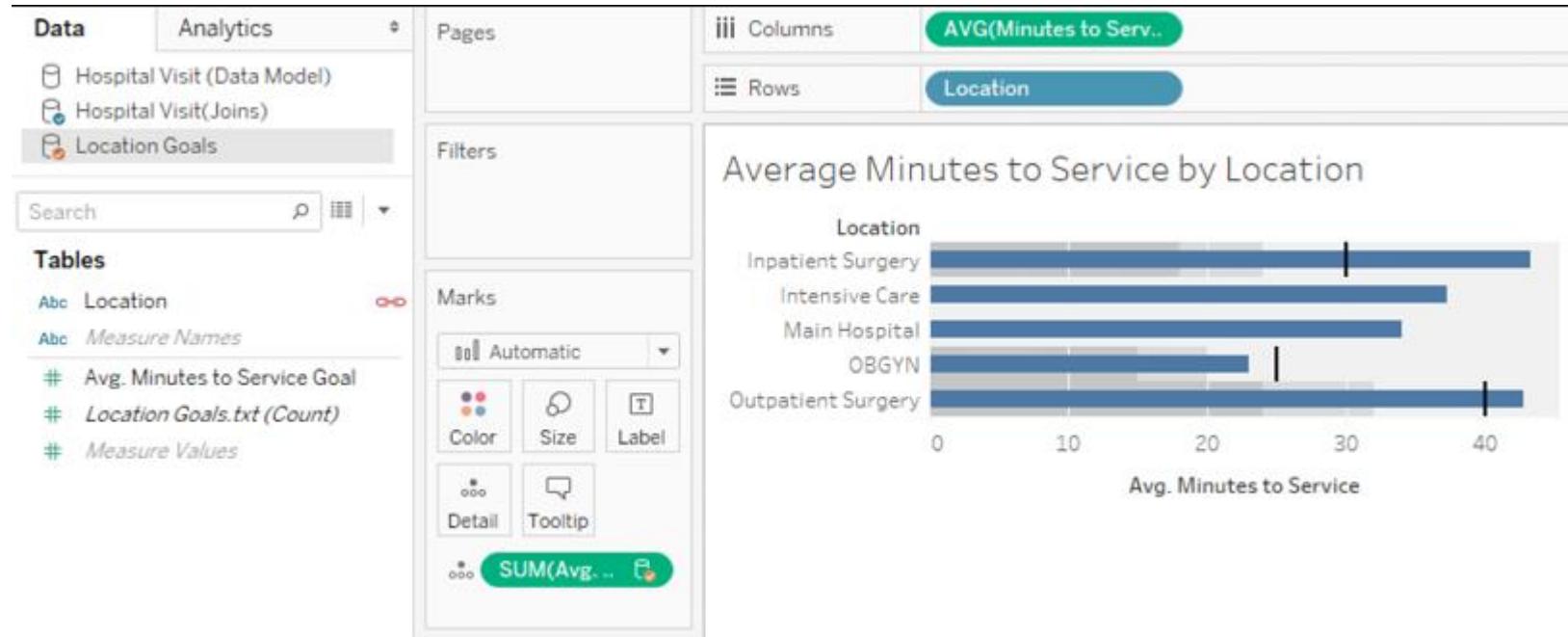
A blending example

- In the Data pane, we'll select the Location Goals data source. Observe the Data pane shown here:



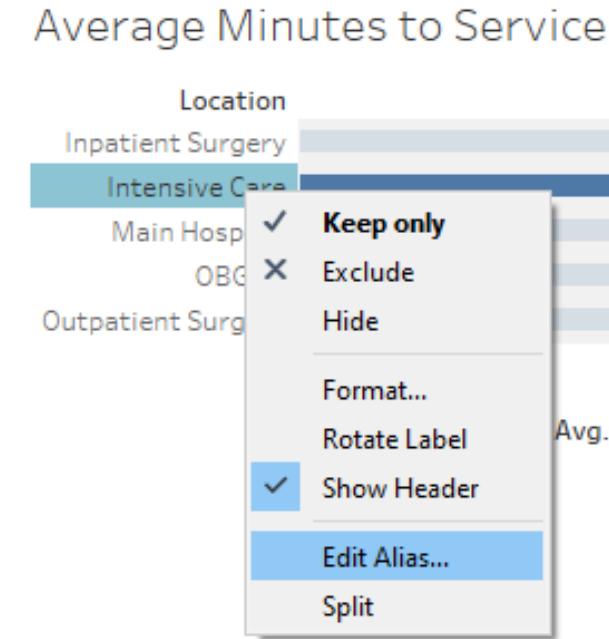


A blending example

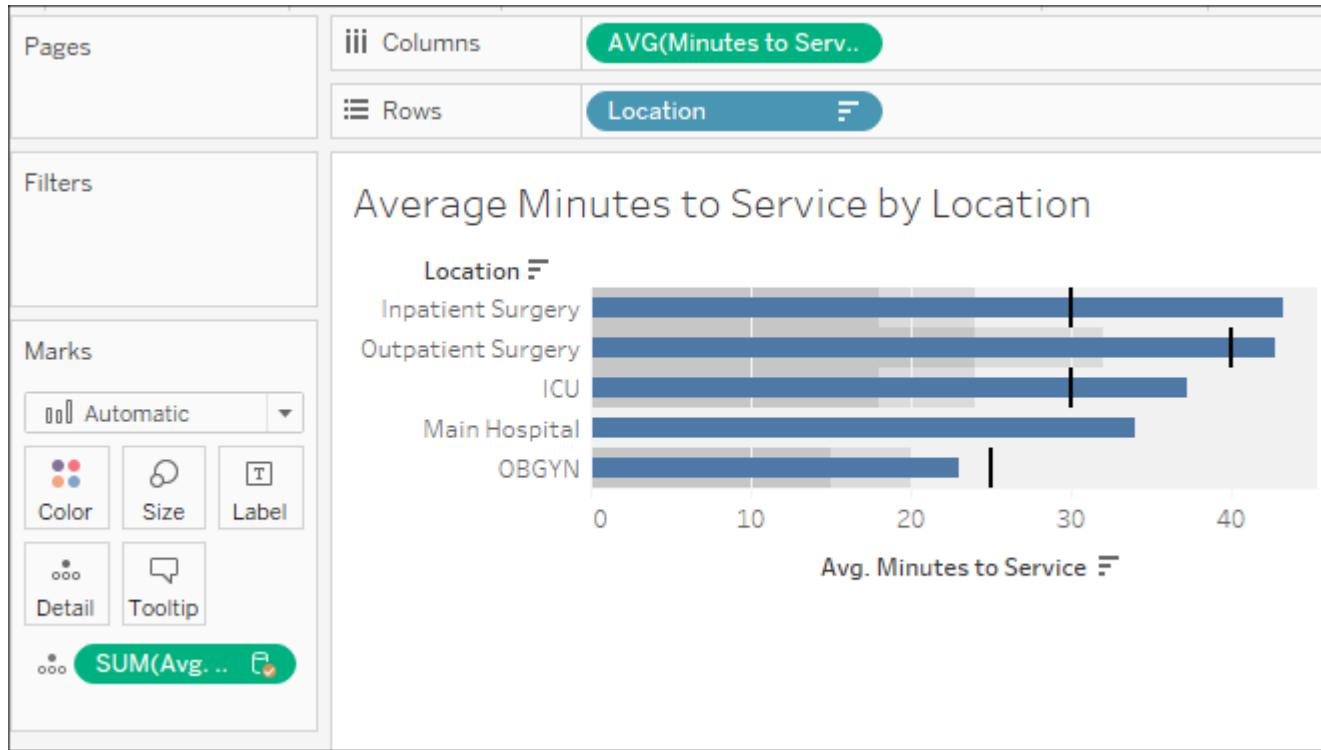


A blending example

- We can change the alias of a field by right-clicking on the row header in the view and using the Edit Alias... option, as shown here:



A blending example



When to use a data model, joins, or blends

In general, use a data model to relate tables:

- When joins would make correct aggregations impossible or require complex LOD expressions to get accurate results
- When joins would result in the duplication of data
- When you need flexibility in showing full domains of dimensions versus only values that match across relationships

When to use a data model, joins, or blends

In general, use joins at the physical level:

- When you must use an operator other than equality between fields
- When you want to do a spatial join
- When you want to specifically control the type of join used in your analysis
- When the performance of the data model is less efficient than it would be with the use of joins

When to use a data model, joins, or blends

In general, use blending when:

- You need to relate data sources that cannot be joined or related using a data model (such as OLAP cubes)
- You need flexibility to "fix" matching using aliases
- You need flexibility to adjust which fields define the relationship differently in different views

Summary

- You now have several techniques to turn to when you need to relate tables of data together.
- The data model, a new feature in the latest versions of Tableau, gives a new paradigm for relating logical tables of data together.
- It introduces a few new behaviors when it comes to showing the full and partial domains of dimensional values, but it also greatly simplifies aggregations by taking into account the natural level of detail for the aggregation.

COMPLETE LAB 13

14. Structuring Messy Data to Work Well in Tableau

Structuring Messy Data to Work Well in Tableau

In this lesson, we'll focus on some principles for structuring data to work well with Tableau, as well as some specific examples of how to address common data issues. This lesson will cover the following topics:

- Structuring data for Tableau
- The four basic data transformations
- Overview of advanced fixes for data problems

Structuring data for Tableau

There are two keys to ensure a good data structure that works well with Tableau:

- Every record of a source data connection should be at a meaningful level of detail
- Every measure contained in the source should match the level of detail of the data source or possibly be at a higher level of detail, but it should never be at a lower level of detail

Structuring data for Tableau

School	Classroom	Average GPA	Number of Students	Number of Students (School)
Pickaway Elementary	4 th Grade	78	153	1,038
Pickaway Elementary	5 th Grade	73	227	1,038
Pickaway Elementary	6 th Grade	84	227	1,038
McCord Elementary	4 th Grade	82	94	915
McCord Elementary	5 th Grade	77	89	915
McCord Elementary	6 th Grade	84	122	915

Well-structured data in Tableau

- The two keys to a good structure, which we mentioned in the previous section, should result in a data structure where a single measure is contained in a single column.
- You may have multiple different measures, but any single measure should almost never be divided into multiple columns.

Wide data

- Wide data is not typically a good structure for visual analysis in Tableau.
- Wide data describes a structure in which a measure in a single row is spread over multiple columns.
- This data is often more human-readable, Wide data often results in fewer rows with more columns.

Wide data

- Here is an example of what wide data looks like in a table of population numbers:

Country Name	1960	1961	1962	1963	1964
Afghanistan	8,774,440	8,953,544	9,141,783	9,339,507	9,547,131
Australia	10,276,477	10,483,000	10,742,000	10,950,000	11,167,000

Tall data

- Tall data is typically a good structure for visual analysis in Tableau.
- Tall data describes a structure in which each distinct measure in a row is contained in a single column.
- Tall data often results in more rows and fewer columns.

- Consider the following table, which represents the same data as earlier but in a tall structure:

Country Name	Year	Population
Afghanistan	1960	8,774,440
Afghanistan	1961	8,953,544
Afghanistan	1962	9,141,783
Afghanistan	1963	9,339,507
Afghanistan	1964	9,547,131
Australia	1960	10,276,477
Australia	1961	10,483,000
Australia	1962	10,742,000
Australia	1963	10,950,000
Australia	1964	11,167,000

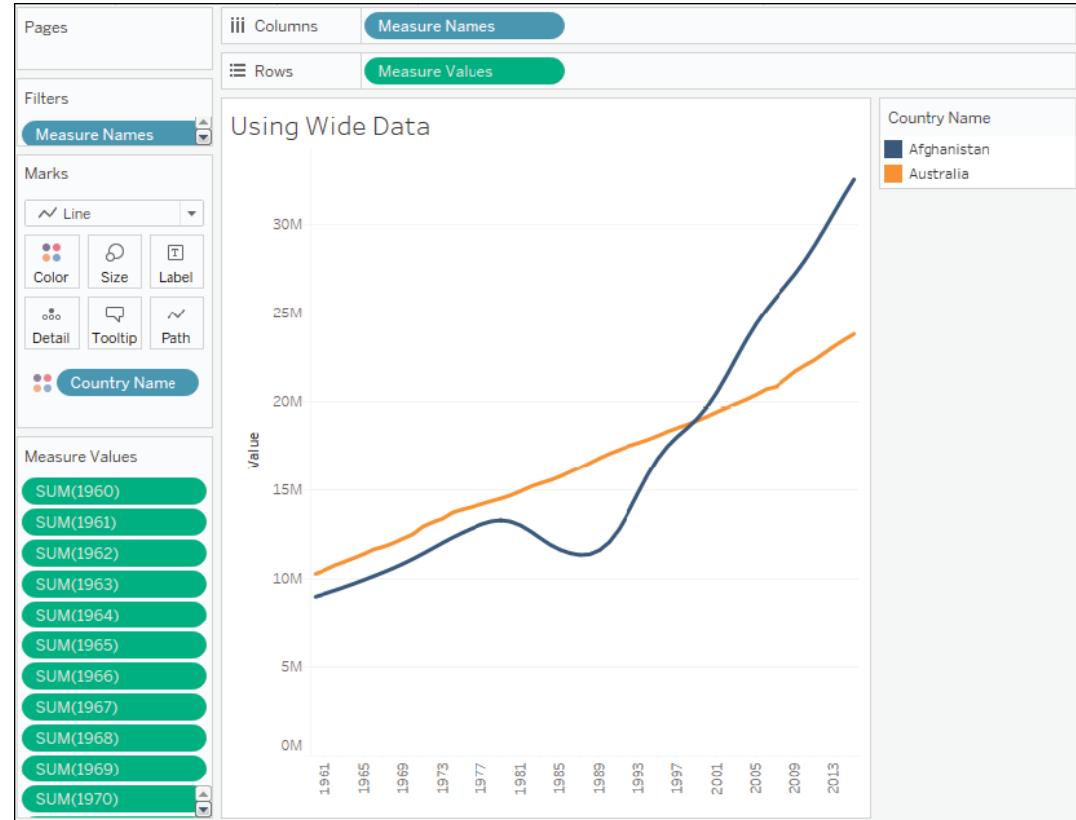
Wide versus tall data in Tableau

- You can easily see the difference between wide and tall data in Tableau.
- Here is what the wide data table looks like in the left Data window:

The screenshot shows the Tableau Data pane. At the top, there are tabs for 'Data' and 'Analytics'. Below the tabs, a dropdown menu shows 'Wide (Country)'. A search bar is followed by a 'Tables' section. Under 'Tables', there is a section for 'Country Name' and another for 'Measure Names'. The 'Measure Names' section lists the years from 1960 to 1965, each preceded by a green hash symbol (#). A vertical scrollbar is visible on the right side of the pane.

#	Measure Names
#	1960
#	1961
#	1962
#	1963
#	1964
#	1965

- For example, you might use Measure Names/Measure Values to plot all the Year measures in a single view, like this:

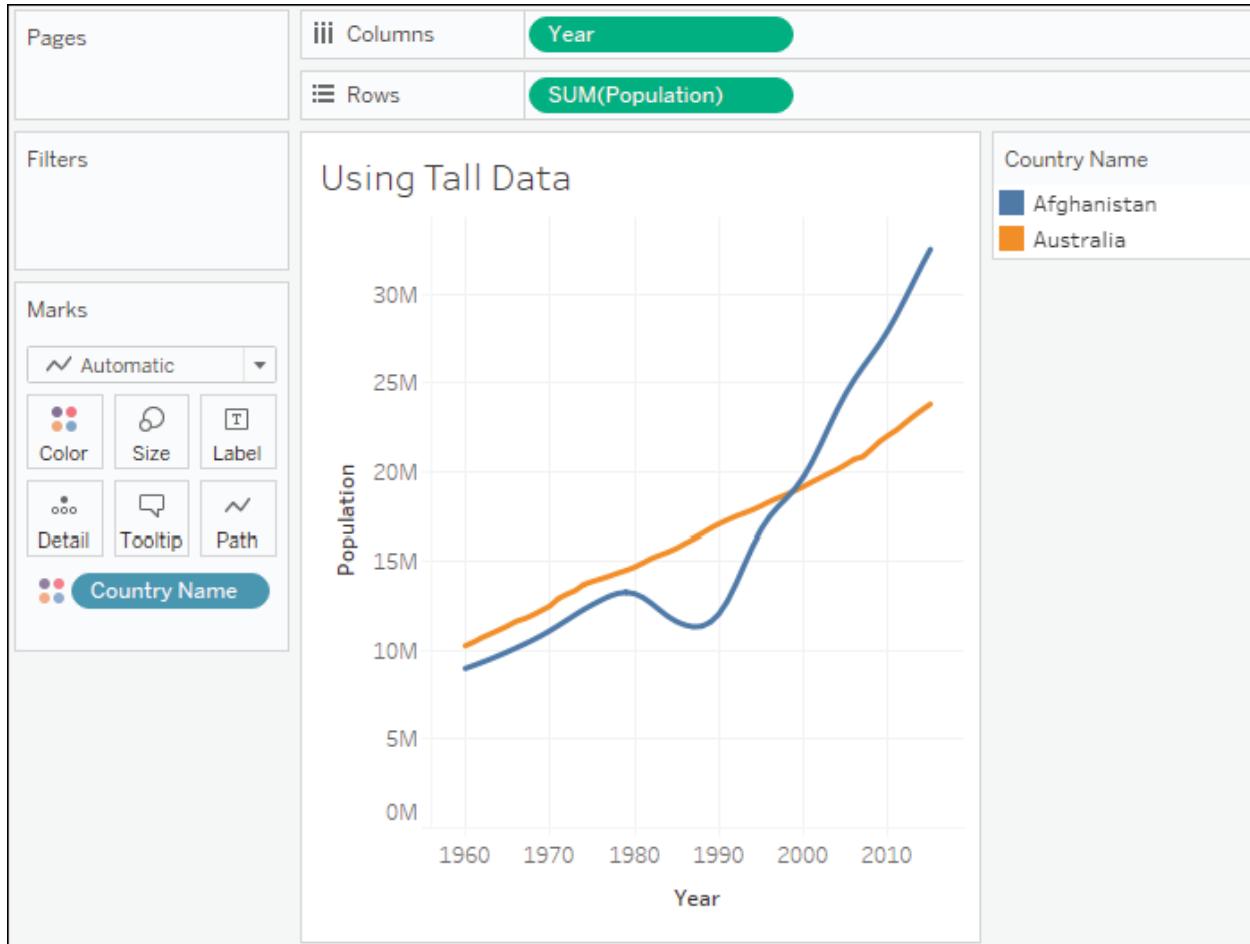


Wide versus tall data in Tableau

- In contrast, the tall data looks like this in the Data pane:

The screenshot shows the Tableau Data pane interface. At the top, there are tabs for "Data" and "Analytics". Below the tabs, a dropdown menu is open, showing "Tall (Country)" as the selected option. There is also a search bar and a refresh icon. The main area is titled "Tables" and lists various data items, each preceded by a small icon:

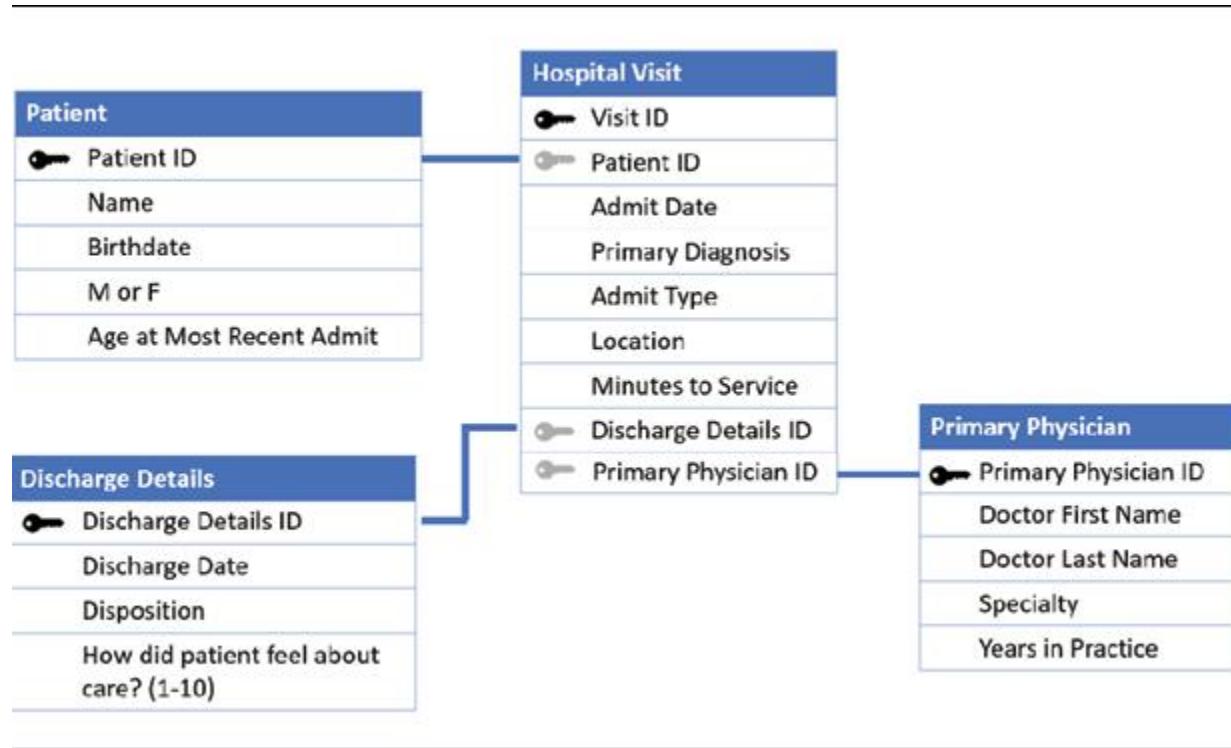
- 🌐 Country Name
- # Year
- Abc Measure Names
- # Population
- 🌐 Latitude (generated)
- 🌐 Longitude (generated)
- # Migrated Data (Count)
- =# Number of Records
- # Measure Values



Star schemas (Data mart/Data warehouse)

- Assuming they are well-designed, star schema data models work very well with Tableau because they have well-defined granularity, measures, and dimensions.
- Additionally, if they are implemented well, they can be extremely efficient to query.
- This allows for an ergonomic experience when using live connections in Tableau.

Star schemas (Data mart/Data warehouse)



The four basic data transformations

- In this section, we'll give you an overview of some basic transformations that can fundamentally change the structure of your data.
- We'll start with an overview and then look at some practical examples.

Overview of transformations

In Tableau (and Tableau Prep), there are four basic data transformations. The following definitions broadly apply to most databases and data transformation tools, but there are some details and terminology that are Tableau-specific:

- Pivots:
- Unions
- Joins
- Aggregations

Pivots (along with some simple data cleaning)

A	B	C	D	E	F	G	H	I	J	K	
1	World Population Data										
3	This is sample data only.										
4	Accuracy and completeness is not guaranteed.										
6	Country Name and Code	Indicator Name	Indicator Code	1960	1961	1962	1963	1964	1965	1966	1967
7	Aruba (ABW)	Population, total	SP.POP.TOTL	54208	55435	56226	56697	57029	57360	57712	58049
8	Andorra (AND)	Population, total	SP.POP.TOTL	13414	14376	15376	16410	17470	18551	19646	20755
9	Afghanistan (AFG)	Population, total	SP.POP.TOTL	8774440	8953544	9141783	9339507	9547131	9765015	9990125	10221902
10	Angola (AGO)	Population, total	SP.POP.TOTL	4965988	5056688	5150076	5245015	5339893	5433841	5526653	5619643
11	Albania (ALB)	Population, total	SP.POP.TOTL	1608800	1659800	1711319	1762621	1814135	1864791	1914573	1965598
12	United Arab Emirates (ARE)	Population, total	SP.POP.TOTL	89608	97727	108774	121574	134411	146341	156890	167360
13	Argentina (ARG)	Population, total	SP.POP.TOTL	20623998	20959241	21295290	21630854	21963952	22293817	22618887	22941477
14	Armenia (ARM)	Population, total	SP.POP.TOTL	1867396	1934239	2002170	2070427	2138133	2204650	2269475	2332624
15	American Samoa (ASM)	Population, total	SP.POP.TOTL	20012	20478	21118	21883	22701	23518	24320	25116
16	Antigua and Barbuda (ATG)	Population, total	SP.POP.TOTL	54681	55403	56311	57368	58500	59653	60818	62002
17	Australia (AUS)	Population, total	SP.POP.TOTL	10276477	10483000	10742000	10950000	11167000	11388000	11651000	11799000
18	Austria (AUT)	Population, total	SP.POP.TOTL	7047539	7086299	7129864	7175811	7223801	7270889	7322066	7376998
19	Azerbaijan (AZE)	Population, total	SP.POP.TOTL	3897889	4030130	4167558	4307315	4445653	4579759	4708485	4832098

Data (World Population Data)

Connection: Live Extract

Filters: 0 | Add

Connections Add
World Population Data Microsoft Excel

Sheets Use Data Interpreter
Data Interpreter might be able to clean your Microsoft Excel workbook.

Data

Data

Need more data?
Drag tables here to relate them. [Learn more](#)

Sort fields Data source order ▾ Show aliases Show hidden fields 239 rows

New Union

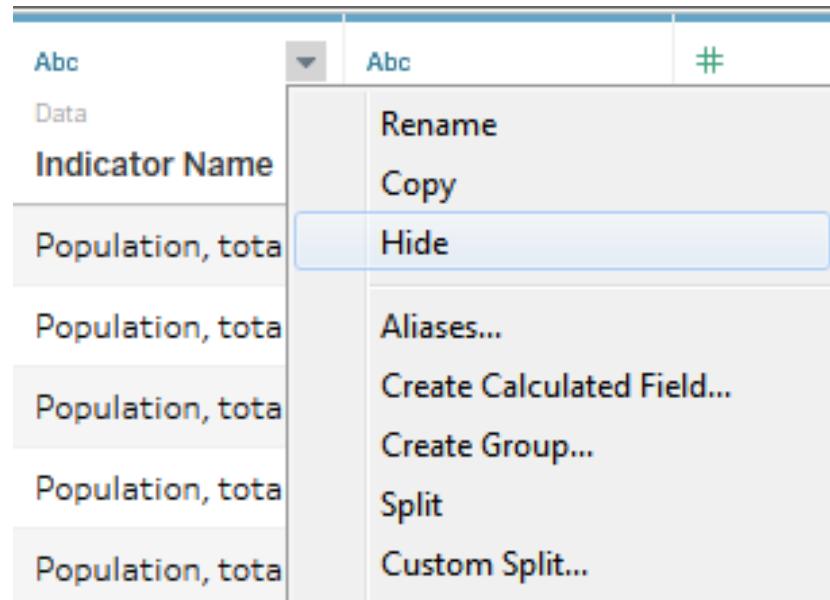
Abc Data F1	Abc Data F2	Abc Data F3	# Data F4	# Data F5	# Data F6	# Data F7	# Data F8
World Population Data	null	null	null	null	null	null	null
This is sample data o...	null	null	null	null	null	null	null
Country Name and Co...	Indicator Name	Indicator Code	1,960	1,961	1,962	1,963	1,96
Aruba (ABW)	Population, total	SP.POP.TOTL	54,208	55,435	56,226	56,697	57,02
Andorra (AND)	Population, total	SP.POP.TOTL	13,414	14,376	15,376	16,410	17,471
Afghanistan (AFG)	Population, total	SP.POP.TOTL	8,774,440	8,953,544	9,141,783	9,339,507	9,547,13
Angola (AGO)	Population, total	SP.POP.TOTL	4,965,988	5,056,688	5,150,076	5,245,015	5,339,89

Pivots (along with some simple data cleaning)

Country Name and ...	Indicator Name	Indicator Code	1960	1961	1962	1963	1964	1965
Aruba (ABW)	Population, total	SP.POP.TOTL	54,208	55,435	56,226	56,697	57,029	57,360
Andorra (AND)	Population, total	SP.POP.TOTL	13,414	14,376	15,376	16,410	17,470	18,551
Afghanistan (AFG)	Population, total	SP.POP.TOTL	8,774,440	8,953,544	9,141,783	9,339,507	9,547,131	9,765,015
Angola (AGO)	Population, total	SP.POP.TOTL	4,965,988	5,056,688	5,150,076	5,245,015	5,339,893	5,433,841
Albania (ALB)	Population, total	SP.POP.TOTL	1,608,800	1,659,800	1,711,319	1,762,621	1,814,135	1,864,791
United Arab Emirates (...)	Population, total	SP.POP.TOTL	89,608	97,727	108,774	121,574	134,411	146,341

Pivots (along with some simple data cleaning)

- Selecting Hide will remove the field from the connection and even prevent it from being stored in extracts:



Pivots (along with some simple data cleaning)

=@ Calculation Country Name	=@ Calculation Country Code	# Data 1961	# Data 1962	# Data 1963	# Data 1964	# Data 1965	# Data 1966
Aruba	ABW	55,435	56,226	56,697	57,029	57,360	57,712
Andorra	AND	14,376	15,376	16,410	17,470	18,551	19,646
Afghanistan	AFG	8,953,544	9,141,783	9,339,507	9,547,131	9,765,015	9,990,125
Angola	AGO	5,056,688	5,150,076	5,245,015	5,339,893	5,433,841	5,526,653
Albania	ALB	1,659,800	1,711,319	1,762,621	1,814,135	1,864,791	1,914,573
United Arab Emirates	ARE	97,727	108,774	121,574	134,411	146,341	156,890

Pivots (along with some simple data cleaning)

A screenshot of a data editing interface, likely from a spreadsheet or database application. On the left, there is a table with columns for Country Code and Name. The table contains the following data:

Country Code	Name
ABW	Aruba
AND	Andorra
AFG	Afghanistan
AGO	Angola
ALB	Albania

To the right of the table, a context menu is open, showing various data type options:

- Number (decimal)
- Number (whole)** (selected)
- Date & Time
- Date
- String
- Boolean
- Default
- Geographic Role ►

Pivots (along with some simple data cleaning)

- The final cleaned, pivoted dataset is far easier to work with in Tableau than the original:

=🌐 Calculation Country Name	=🌐 Calculation Country Code	# Pivot Year	# Pivot Population
Aruba	ABW	1961	55,435
Andorra	AND	1961	14,376
Afghanistan	AFG	1961	8,953,544
Angola	AGO	1961	5,056,688
Albania	ALB	1961	1,659,800
United Arab Emirates	ARE	1961	97,727
Argentina	ARG	1961	20,959,241
Armenia	ARM	1961	1,934,239
American Samoa	ASM	1961	20,478
Antigua and Barbuda	ATG	1961	55,403
Australia	AUS	1961	10,483,000

Unions

- A union is a concatenation of data tables that brings together rows of each table into a single table.
- For example, consider the following three tables of data.

Originals:

Name	Occupation	Bank account balance
Luke	Farmer	\$2,000
Leia	Princess	\$50,000
Han	Smuggler	-\$20,000

Prequels:

Name	Occupation	Bank account balance
Watto	Junk Dealer	\$9,000
Darth Maul	Face Painter	\$10,000
Jar Jar	Sith Lord	-\$100,000

Unions

Sequels:

Name	Occupation	Bank account balance
Rey	Scavenger	\$600
Poe	Pilot	\$30,000
Kylo	Unemployed	\$0

A union of these tables would give a single table containing the rows of each individual table:

Name	Occupation	Bank account balance
Luke	Farmer	\$2,000
Leia	Princess	\$50,000
Han	Smuggler	-\$20,000
Watto	Junk Dealer	\$9,000
Darth Maul	Face Painter	\$10,000
Jar Jar	Sith Lord	-\$100,000
Rey	Scavenger	\$600
Poe	Pilot	\$30,000
Kylo	Unemployed	\$0

Unions

The screenshot shows a user interface for managing data connections and files. On the left, under 'Connections', there is one entry: 'Originals' (Text File). Under 'Files', there are four entries: 'Originals.txt', 'Prequels.txt', 'Sequel.txt', and 'New Union'. A blue curved arrow originates from the 'Prequels.txt' entry in the 'Files' list and points towards the 'Originals' connection in the 'Connections' panel. An orange callout box with the text 'Drag table to union' is positioned over the 'Originals.txt' connection.

Connections

Add

Originals
Text File

Files

Originals.txt

Prequels.txt

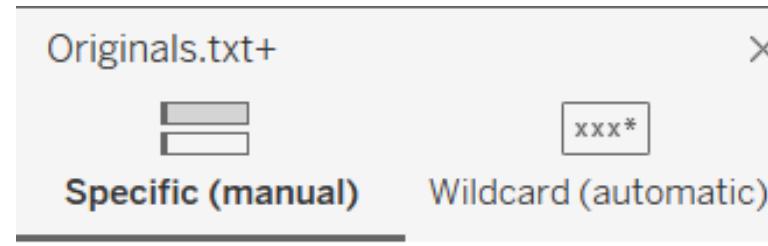
Sequel.txt

New Union

Originals.txt

Drag table to union

Unions



Connection: Originals

Originals.txt
Prequels.txt
Sequals.txt

Tables in union: 3

Apply

OK

Unions

- For example, if you had a directory of monthly data dump files, named 2020-01.txt, 2020-02.txt, 2020-03.txt, and so on, but no actual date field in the files, you could obtain the date using a calculated field with code such as the following:

`DATEPARSE('yyyy-MM', [Table Name])`

Abc Originals+	Abc Originals+	Abc Originals+	Abc	Abc
Name	Occupation	Job		
Luke	Farmer	null		
Leia	Princess	null		
Han	Smuggler	null		
Watto	null	Junk Dealer		
Darth Maul	null	Face Painter	Prequels	Prequels
Jar Jar	null	Sith Lord	Prequels	Prequels
Rey	Scavenger	null	Sequels	Sequels
Poe	Pilot	null	Sequels	Sequels
Kylo	Unemployed	null	Sequels	Sequels

- Rename
- Copy
- Hide
- Create Calculated Field...
- Pivot
- Merge Mismatched Fields

Joins

- You'll recall that the concept of joins and the types of joins were previously discussed in lesson 13, Understanding the Tableau Data Model, Joins, and Blends.
- While joins are quite useful in bringing together tables in the same database or even disparate data sources (data contained in different systems and formats), they can be used to solve other data issues too, such as reshaping data to make it easier to meet your objectives in Tableau.

Patient ID	Patient Name	Admit Date	Discharge Date
1	David	12/1/2018	12/20/2018
2	Solomon	12/3/2018	12/7/2018
3	Asa	12/5/2018	12/22/2018
4	Jehoshaphat	12/5/2018	12/6/2018
5	Joash	12/9/2018	12/16/2018
6	Amaziah	12/10/2018	12/14/2018
7	Uzziah	12/12/2018	12/24/2018
8	Jotham	12/16/2018	12/29/2018
9	Hezekiah	12/18/2018	12/22/2018
10	Josiah	12/22/2018	12/23/2018

Joins

- An alternative is to use a join to create the rows for all dates.
- In this case, we'll leverage a cross-database join to bring in another source of data altogether.
- You might quickly create an Excel sheet with a list of dates you want to see, like this:

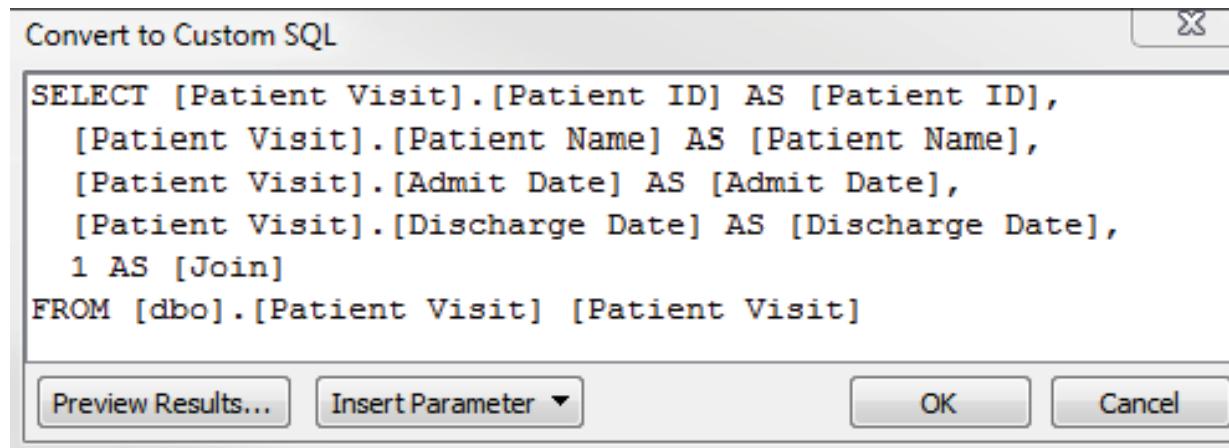
A
1 Date
2 12/1/2018
3 12/2/2018
4 12/3/2018
5 12/4/2018
6 12/5/2018
7 12/6/2018
8 12/7/2018
9 12/8/2018
10 12/9/2018
26 12/25/2018
27 12/26/2018
28 12/27/2018
29 12/28/2018
30 12/29/2018
31 12/30/2018
32 12/31/2018

- In this case, we'll select Create Join Calculation... for both tables and enter the single, hard-coded value, that is, 1, for both the left and right sides:

The screenshot shows the Power BI Data Flow interface. On the left, the 'Connections' pane lists 'Hospital Patients' (Text file) and 'Hospital Patients Dates' (Microsoft Excel). The 'Sheets' pane shows 'Dates' and 'New Union'. In the main area, under 'Hospital Patients Padded', it says 'Hospital Visit Padded is made of 2 tables.' A diagram shows 'Hospital Patients.txt' connected to 'Dates' via a join. A tooltip for the join operation is open, showing options: Inner, Left, Right, Full Outer. It also shows 'Data Source' (1) and 'Dates' (Enter search text, Date, Create Join Calculation...).

Joins

- Fields defined in Custom SQL can also be used in joins:



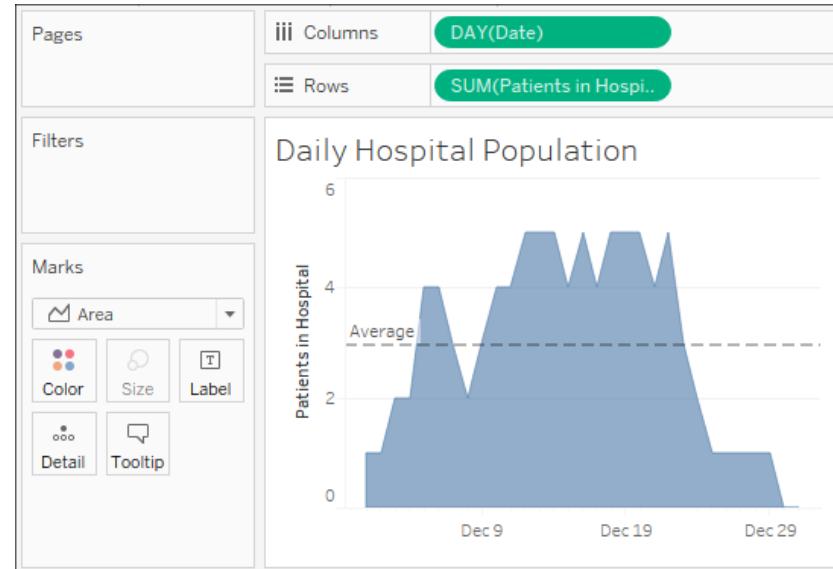
Joins

- The calculated field, named Patients in Hospital, has the following code:

```
IF [Admit Date] <= [Date] AND [Discharge Date] >= [Date]
THEN 1
ELSE 0
END
```

Joins

- Allows us to easily visualize the flow of patients, and even potentially perform advanced analytics based on averages, trends, and even forecasting:



Aggregation

Remember that the two keys to a good data structure are as follows:

- Having a level of detail that is meaningful
- Having measures that match the level of detail or that are possibly at higher levels of detail

Aggregation

- Consider, for example, the following data (included as Apartment Rent.xlsx in the lesson 14 directory), which gives us a single record each month per apartment:

Apartment	Month	Rent Collected	Square Feet
A	Jan	\$0	900
	Feb	\$0	900
	Mar	\$0	900
	Apr	\$0	900
	May	\$0	900
	Jun	\$1,500	900
	Jul	\$1,500	900
	Aug	\$1,500	900
	Sep	\$1,500	900
	Oct	\$1,500	900
	Nov	\$1,500	900
	Dec	\$1,500	900
B	Jan	\$1,200	750
	Feb	\$1,200	750
	Mar	\$1,200	750
	Apr	\$1,200	750
	May	\$1,200	750
	Jun	\$1,200	750
	Jul	\$0	750
	Aug	\$0	750
	Sep	\$0	750
	Oct	\$0	750
	Nov	\$0	750
	Dec	\$0	750

Aggregation

The two measures are really at different levels of detail:

- Rent Collected matches the level of detail of the data where there is a record of how much rent was collected for each apartment for each month.
- Square Feet, on the other hand, does not change month to month. Rather, it is at the higher level of detail, of Apartment only.

Pages ▾

Columns Measure Names ▾

Rows Apartment

Filters

Measure Names ▾

Marks

Automatic

Color Size Text

Detail Tooltip

Measure Values ▾

Measure Values

SUM(Rent Collected)

SUM(Square Feet)

AVG(Square Feet)

MIN(Square Feet)

MAX(Square Feet)

Measures per Apartment

Apartment	Rent Collected	Sum of Square Feet	Avg. Square Feet	Min. Square Feet	Max. Square Feet
A	\$10,500	10,800	900	900	900
B	\$7,200	9,000	750	750	750

Aggregation

- To understand why, consider what happens when we turn on the column grand totals (from the menu, select Analysis | Totals | Show Column Grand Totals, or drag and drop Totals from the Analytics tab):

Apartment	Rent Collected	Sum of Square Feet	Avg. Square Feet	Min. Square Feet	Max. Square Feet
A	\$10,500	10,800	900	900	900
B	\$7,200	9,000	750	750	750
Grand Total	\$17,700	19,800	825	750	900

Aggregation

- We'll use a fixed LOD calculation to keep the level of detail fixed at the apartment level.
- We'll create a calculated field, named Square Feet per Apartment, with the following code:

```
{ INCLUDE [Apartment] : MIN([Square Feet]) }
```

Aggregation

- The calculation returns the correct result in the view at the apartment level and at the grand total level, where Tableau includes Apartment to find 900 (the minimum for A) and 750 (the minimum for B) and then sums them to get 1,650:

Apartment	Rent Collected	Square Feet per Apartment
A	\$10,500	900
B	\$7,200	750
Grand Total	\$17,700	1,650

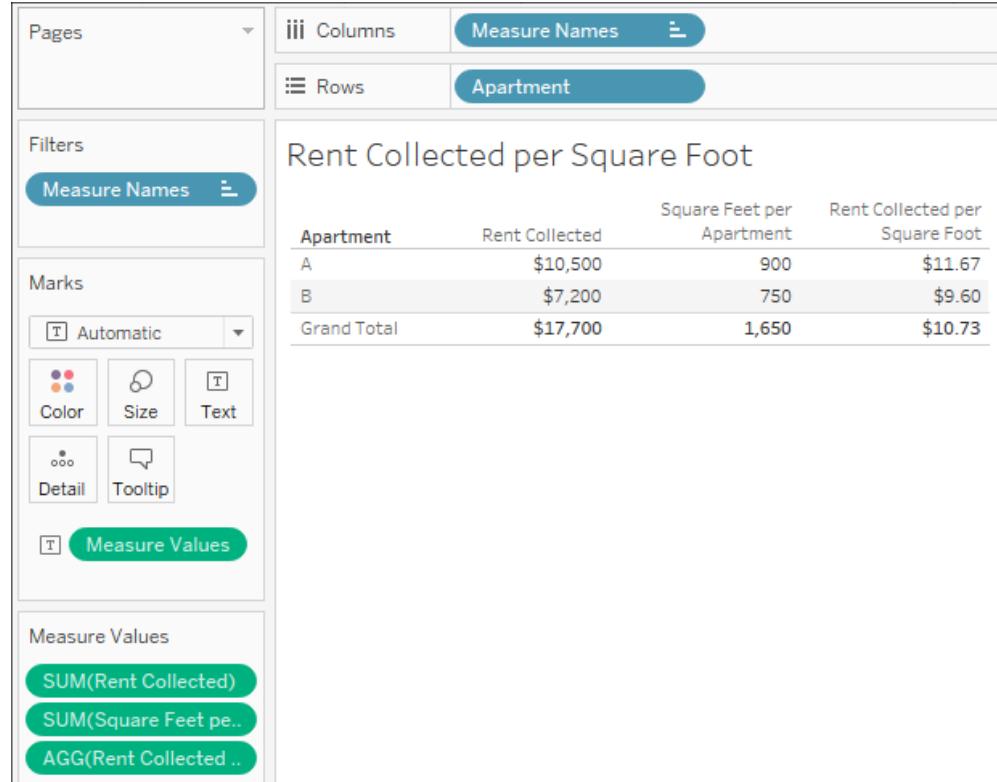
Aggregation

- Now, we can use the LOD calculated field in another calculation to determine the desired results.
- We'll create a calculated field, named Rent Collected per Square Foot, with the following code:

$\text{SUM}([\text{Rent Collected}])/\text{SUM}([\text{Square Feet per Apartment}])$

Aggregation

- When that field is added to the view and formatted to show decimals, the outcome is correct:



Overview of advanced fixes for data problems

- Consider a custom SQL script that takes the wide table of country populations we mentioned earlier in this lesson and restructures it into a tall table:

```
SELECT [Country Name],[1960] AS Population, 1960 AS Year  
FROM Countries
```

```
UNION ALL
```

```
SELECT [Country Name],[1961] AS Population, 1961 AS Year  
FROM Countries
```

```
UNION ALL
```

```
SELECT [Country Name],[1962] AS Population, 1962 AS Year  
FROM Countries
```

```
...
```

```
...
```

Summary

- Up until this lesson, we'd looked at data that was, for the most part, well-structured and easy to use.
- In this lesson, we considered what constitutes a good structure and ways to deal with poor data structures.
- A good structure consists of data that has a meaningful level of detail and that has measures that match that level of detail.

COMPLETE LAB 14

15. Taming Data with Tableau Prep

Taming Data with Tableau Prep

We'll cover quite a few topics in this lesson, including the following:

- Getting ready to explore Tableau Prep
- Understanding the Tableau Prep Builder interface
- Flowing with the fundamental paradigm
- Options for automating flows

Getting ready to explore Tableau Prep

- Tableau Prep Builder is available for Windows and Mac.
- If you do not currently have Tableau Prep Builder installed on your machine, please take a moment to download the application from
<https://www.tableau.com/products/prep/download>.
- Licenses for Tableau Prep Builder are included with Tableau Creator licensing.
- If you do not currently have a license, you may trial the application for 14 days.

Understanding the Tableau Prep Builder interface

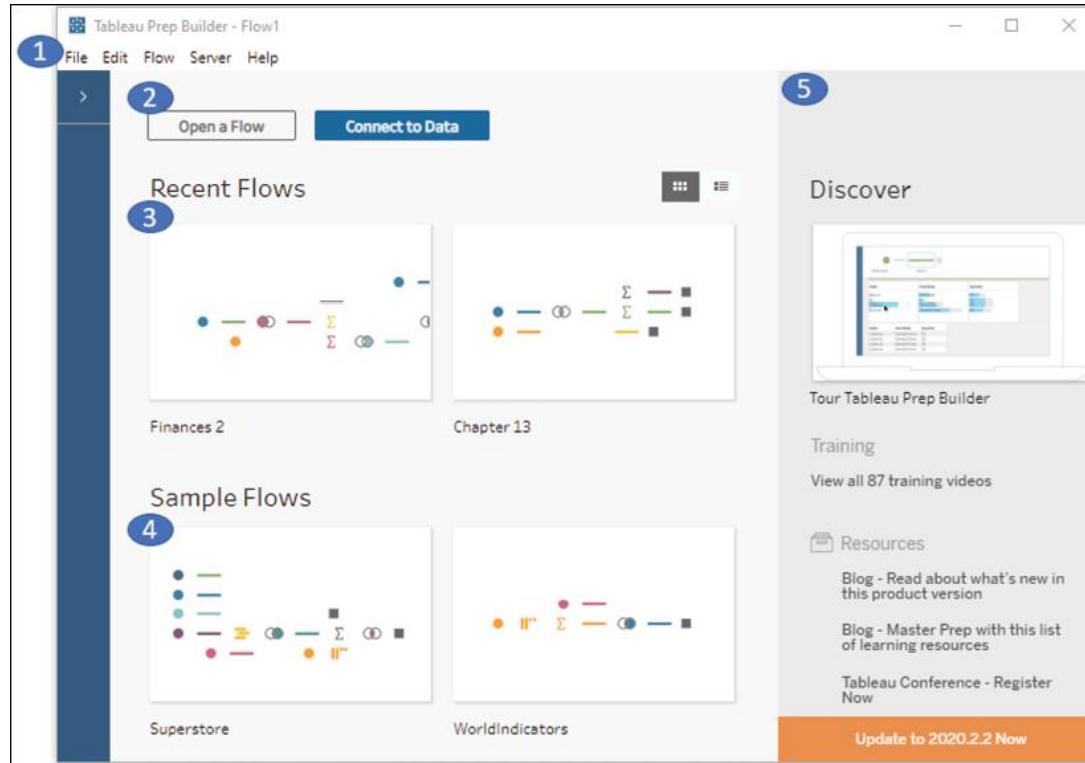


Tableau Prep Builder - Superstore*

File Edit Flow Server Help

Alerts (0)

1

Orders (Central) → Fix Dates → Orders + Returns → Roll Up Sales → Create 'Annual ...'

Returns (all) → Clean Notes/Ap... → Fix Dates

100%

Fix Dates 21 Fields 2K Rows Filter Values... 4 Recommendations Search

2 Changes (13)

Calculated Field [Region] "Central"

Calculated Field [Order Date] (STR([Order Day])+"/"+STR([Order Month])+"/"+STR([Order Year]))

Rename Field [Product Name] From [Product] to [Product Name]

Change Type [Ship Date] To Date type

Remove Field [Order Year]

Remove Field [Order Month]

3

Order Date 720
Region 1
Ship Date 746

01/01/2015
01/01/2019

Central

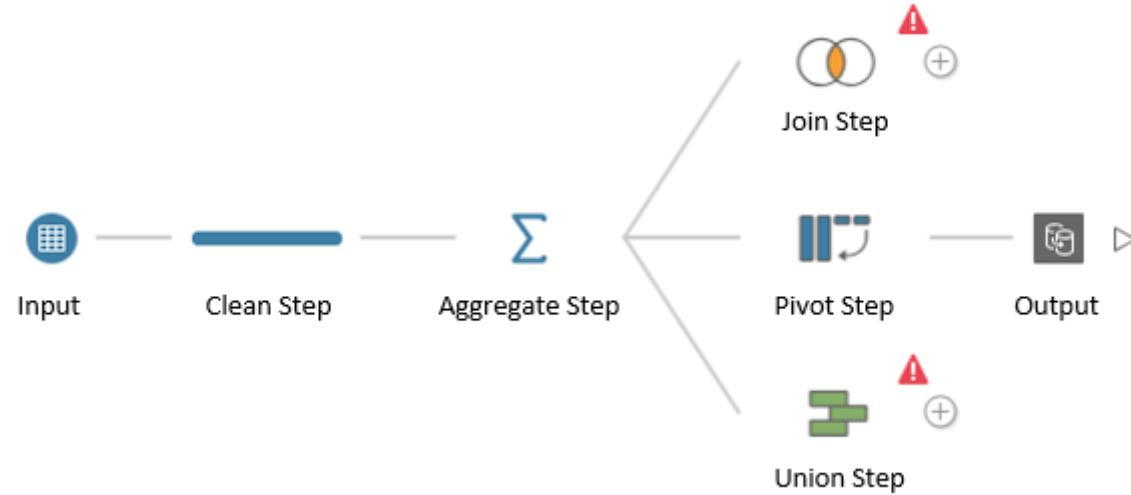
01/01/2015
01/01/2020

4

Order Date	Region	Ship Date	Row ID	Order ID	Ship Mode	Customer ID	Customer N...
11/22/2016	Central	11/26/2016	15	US-2016-118983	Standard Class	HP-14815	Harold Pa...
11/22/2016	Central	11/26/2016	16	US-2016-118983	Standard Class	HP-14815	Harold Pa...
11/11/2015	Central	11/18/2015	17	CA-2015-105893	Standard Class	PK-19075	Pete Kriz
12/09/2017	Central	12/13/2017	22	CA-2017-137330	Standard Class	KB-16585	Ken Black

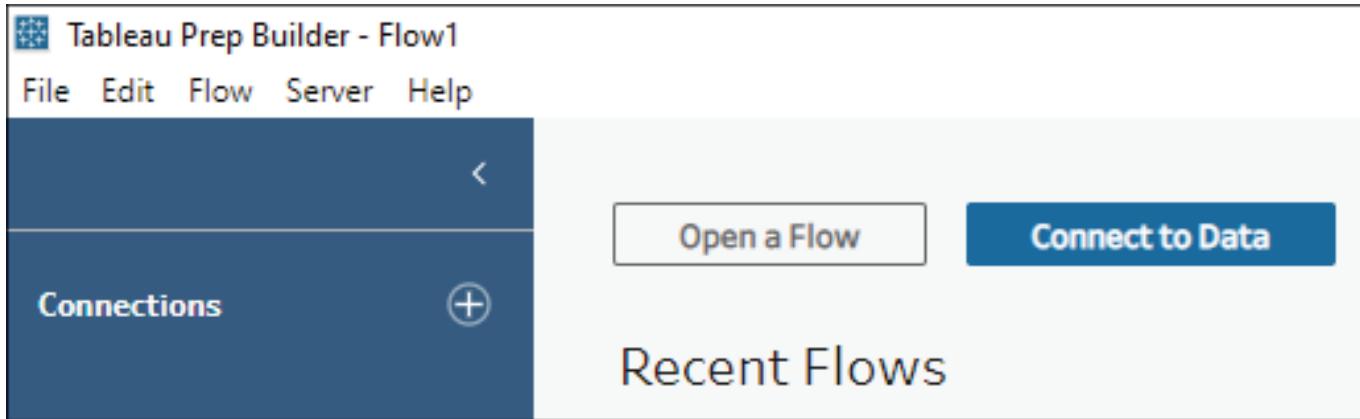
Flowing with the fundamental paradigm

- Example of what a flow looks like in the flow pane of Tableau Prep:



Connecting to data

- Connecting to data in Tableau Prep is very similar to connecting to data in Tableau Desktop.
- From the home screen, you may click either Connect to Data or the + button on the expanded Connections pane:



Employee Flights

Fields selected: 10 of 10

 Filter Values...

Select the fields to include in your flow, apply a filter, or change data types. To see and clean your data, add a cleaning step in the flow pane.

<input checked="" type="checkbox"/>	Type	Field Name	Original Field Name	Changes	Sample Values
<input checked="" type="checkbox"/>	#	Employee Airline...	Employee Airline Travel		null
<input checked="" type="checkbox"/>	Abc	F2	F2		null
<input checked="" type="checkbox"/>	Abc	F3	F3		null
<input checked="" type="checkbox"/>	Abc	F4	F4		null, 07/08/1905, 07/09/1905
<input checked="" type="checkbox"/>	Abc	F5	F5		Total Cost, \$100,287.00, \$108,788.00
<input checked="" type="checkbox"/>	Abc	F6	F6		null
<input checked="" type="checkbox"/>	#	F7	F7		null
<input checked="" type="checkbox"/>	Abc	F8	F8		null
<input checked="" type="checkbox"/>	Abc	F9	F9		null
<input checked="" type="checkbox"/>	Abc	F10	F10		null

Connections 

 Employee Flights.xlsx
Microsoft Excel

Search

Tables

Cleaned with Data Interpreter
To undo changes, clear the checkbox. If the Data Interpreter results were incorporated in the flow, this will result in flow errors.

 Employee Flights

 Employee Flights C3:E7

 Employee Flights Table1

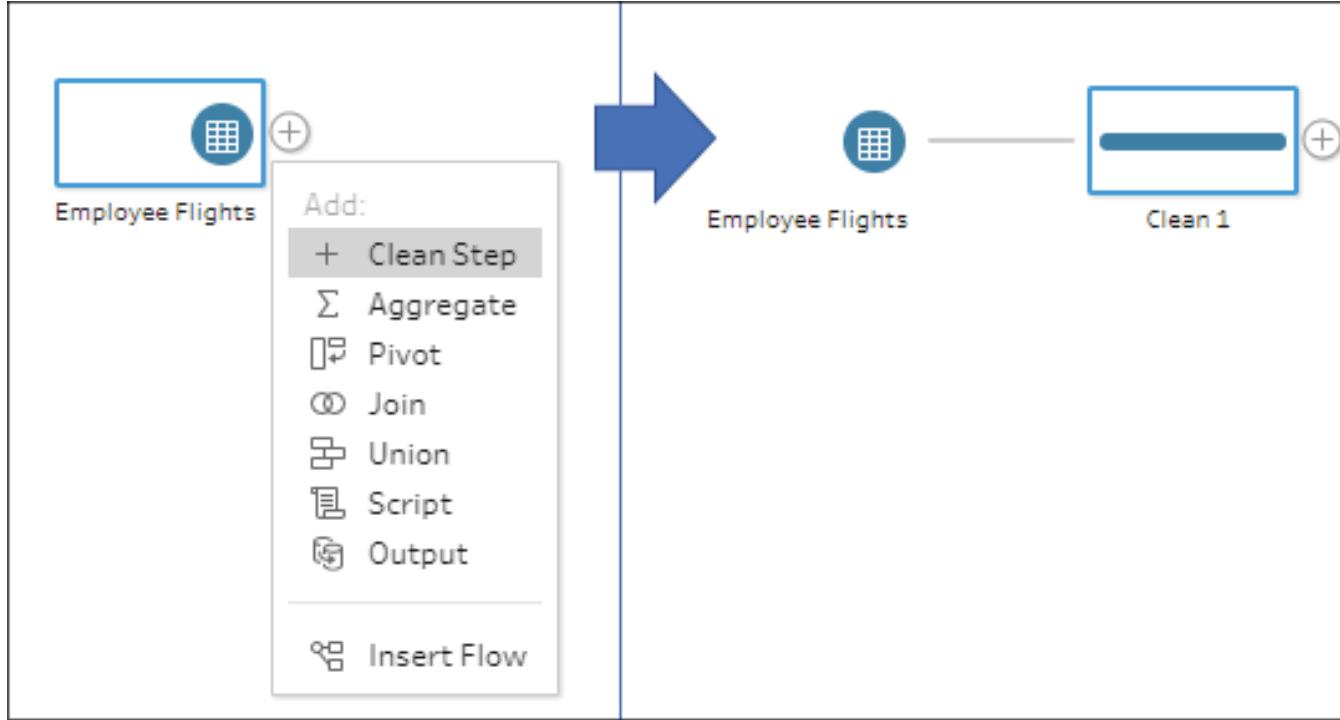


Employee Flights Fields selected: 10 of 10 

Select the fields to include in your flow, apply a filter, or change data types. To see and clean your data, add a cleaning step in the flow pane.

<input checked="" type="checkbox"/>	Type	Field Name	Original Field Name	Changes	Sample Values
<input checked="" type="checkbox"/>	#	Row ID	Row ID		4,144, 9,102, 2,194
<input checked="" type="checkbox"/>	Abc	Passenger Email	Passenger Email		ekiefer@vizpainter.com, abarton@viz
<input checked="" type="checkbox"/>	Icon	Purchase Date	Purchase Date		01/10/2019, 09/10/2017, 10/22/2018
<input checked="" type="checkbox"/>	Icon	Travel Date	Travel Date		02/06/2019, 11/12/2017, 12/17/2018
<input checked="" type="checkbox"/>	Abc	Passenger ID	Passenger ID		EH-14125, AB-10105, AP-10915
<input checked="" type="checkbox"/>	Abc	Airline	Airline		United, American Airlines
<input checked="" type="checkbox"/>	#	Ticket Price	Ticket Price		184, 214, 199

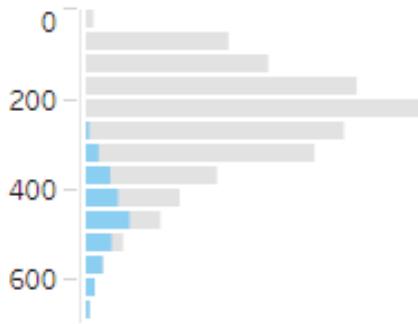
Connecting to data



>

#

Ticket Price 454



Abc

Route 17

- DFW-BUF
- DFW-BUR
- DFW-CHS
- DFW-CMH
- DFW-COS
- DFW-DSM
- DFW-ELP
- DFW-GRB
- DFW-HIO
- DFW-IAD
- DFW-JFK
- DFW-LNK

Abc

Ticket Type 3

- 1st Class
- Business
- Economy

Southwest 2019

Input

Text Settings Multiple Files Data Sample Changes (0)

Connection

Text file
Southwest 2019.csv Edit
Original Table Name: Southwest 2019

Text Options

First line contains header
 Generate field names automatically

Field Separator
Comma

Text Qualifier
Automatic

Southwest 2019 Fields selected: 9 of 9

	Type	Field Name
<input checked="" type="checkbox"/>	Abc	Passenger Email
<input checked="" type="checkbox"/>	T/F	Travel Insurance?
<input checked="" type="checkbox"/>	date	Purchase Date
<input checked="" type="checkbox"/>	date	Travel Date
<input checked="" type="checkbox"/>	Abc	Fare Type
<input checked="" type="checkbox"/>	#	Ticket Price
<input checked="" type="checkbox"/>	Abc	Route
<input checked="" type="checkbox"/>	#	Row_ID
<input checked="" type="checkbox"/>	Abc	Person

Input

Text Settings **Multiple Files** Data Sample Changes (0)

- Single table
 Wildcard union

Search in

Chapter 15



Include subfolders

Files

Include



Matching Pattern (xxx*)

Southwest*

Included files (4)

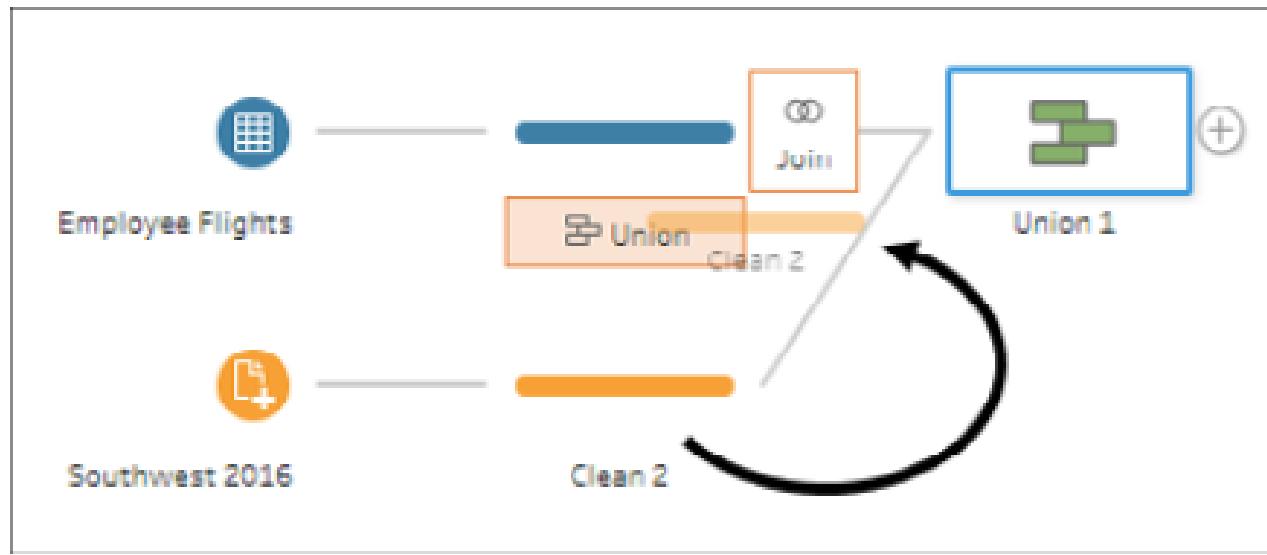
Southwest 2016.csv
Southwest 2017.csv
Southwest 2018.csv
Southwest 2019.csv

Apply

Cleaning the data

- The process of building out the flow is quite iterative, and you'll often make discoveries about the data that will aid you in cleaning and transforming it.
- We'll break this example into sections for the sake of reference, but don't let this detract from the idea that building a flow should be a flow of thought.
- The example is meant to be seamless!

Unioning, merging mismatched fields, and removing unnecessary fields



Union 1 15 Fields 5K Rows

Merge Fields

Settings Changes (0)

Inputs

- Clean 1
- Clean 2

Resulting Fields

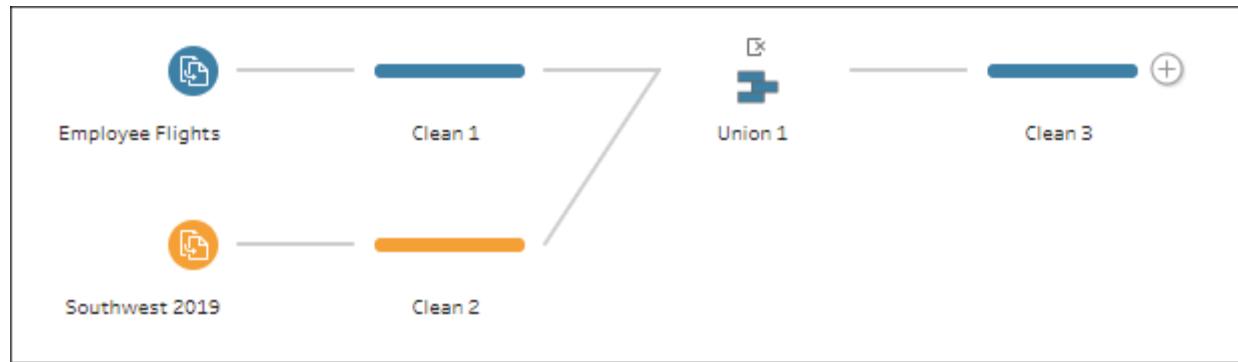
8 Mismatching fields from 15 resulting fields.

Mismatched Fields

Travel Insurance?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Fare Type	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Row_ID	<input type="checkbox"/>	<input checked="" type="checkbox"/>
File Paths	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Row ID	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Passenger ID	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Airline	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Ticket Type	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Unioning, merging mismatched fields, and removing unnecessary fields

- The single remaining mismatched field, Airline, is useful.
- Leave it for now and click the + icon on the Union 1 step in the flow pane and extend the flow by selecting Clean Step, At this point, your flow should look like this:



Grouping and cleaning

Table Names	2
Employee Flights.xlsx/Employee Flights	
Southwest 2016	

Abc

Q

Airline 7

Filter

Clean

Group Values

Split Values

View State

✓ Detail

Summary

Group Values by

Manual Selection

Pronunciation

Common Characters

Spelling

Ungroup All

Rename Field

Duplicate Field

Keep Only Field

Create Calculated Field

Publish as Data Role...

Remove

The screenshot shows a Microsoft Power BI interface. On the left, there's a list of airline names: AA, American, American (AA), American Airlines, Delta, Southwest, and United. The 'Southwest' item is currently selected. A context menu is open over the 'Southwest' item, with a blue arrow pointing to the ellipsis button in the top bar. The menu has several main sections: 'Filter', 'Clean', 'Group Values', 'Split Values', 'View State', and a checked 'Detail' option under 'Summary'. A sub-menu for 'Group Values by' is expanded, showing 'Pronunciation' as the selected option. Other options in this sub-menu include 'Manual Selection', 'Common Characters', and 'Spelling'. At the bottom of the main menu, there are options for 'Rename Field', 'Duplicate Field', 'Keep Only Field', 'Create Calculated Field', 'Publish as Data Role...', and 'Remove'.

Group Values by Pronunciation

Done

Airline 5

Grouping

- AA
- ② American Airlines
- Delta
- Southwest
- United

American Airlines 3 members

- American
- American (AA)
- American Airlines

- AA
- Delta
- Southwest
- United



Recommendations

Change Data Role

To: Email

Many values match this data role.
Apply it to identify invalid values.

Valid values (preview)

rbensley@vizpainter.com
npatel@vizpainter.com

Invalid values (preview)

!mhernandez
!kwillman

Apply

Passenger... 790

... 

Passenger Name	Email Address
null	
a	
aallen@vizpainter.com	
aandreadi@vizpainter....	
aarmold@vizpainter.c...	
aavila@vizpainter.com	

Grouping and cleaning

Recommendations X

Filter: Selected Values

Exclude rows with invalid values for the data role.

Values to exclude (preview)

- !mherandez
- !kwillman
- !asdf
- !pcoakley

Apply

Group Values

Replace invalid values for the data role with null.

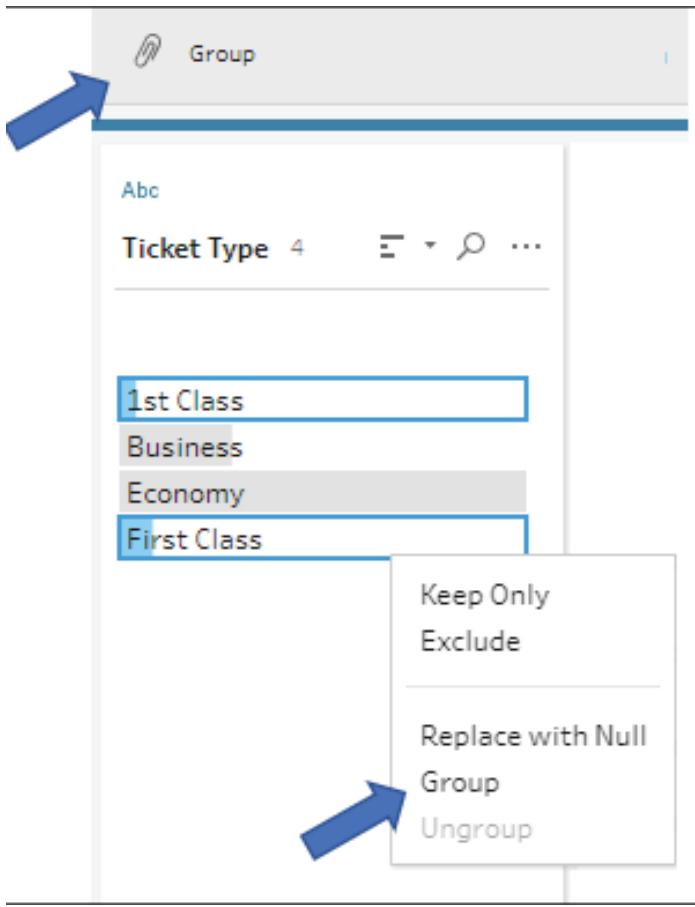
Values to replace (preview)

- !mherandez
- !kwillman
- !asdf

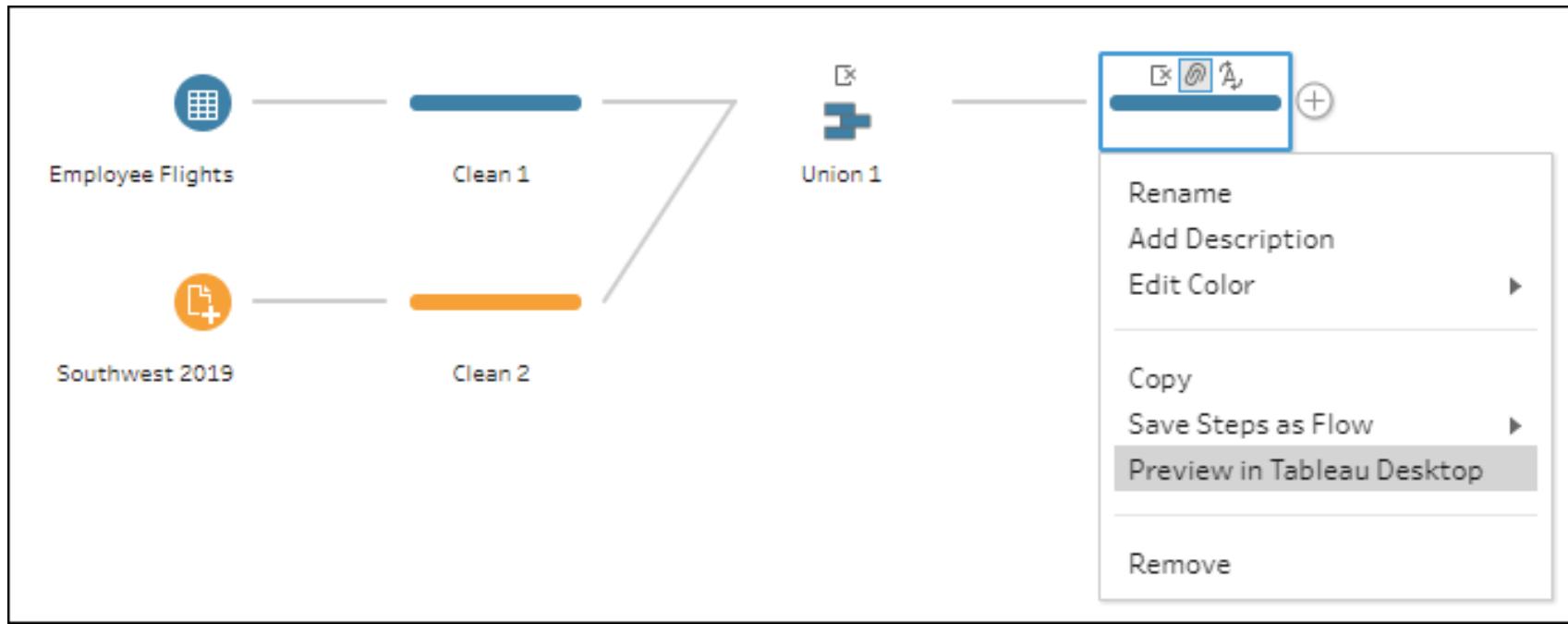
Replace with this value

- null

Apply



Grouping and cleaning

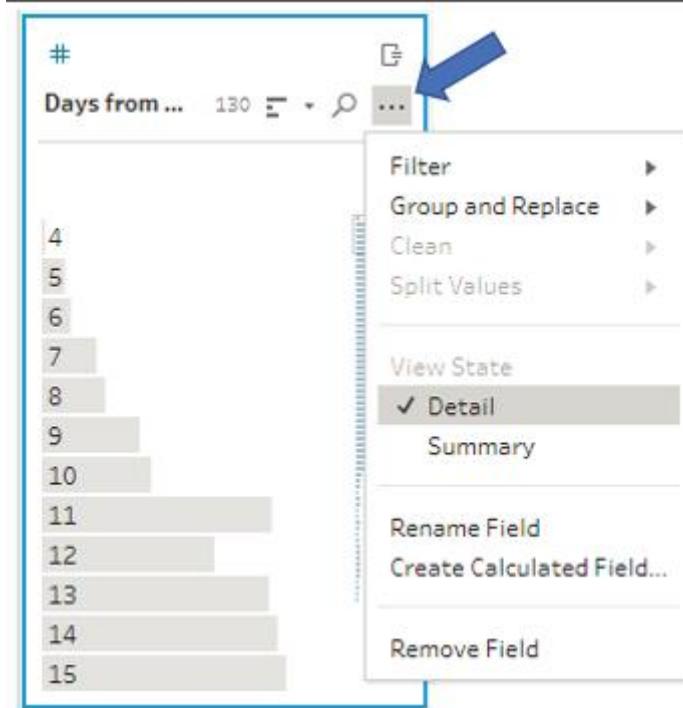


Calculations and aggregations in Tableau Prep

- Examine the results in the profile pane.
- The new field should look like this:



Row-level calculations



Level of detail calculations

- We'll accomplish this using a **FIXED** level of detail (LOD) expression.
- We could create the calculation from scratch, matching the syntax we learned for Tableau Desktop to write the calculation like this:

{**FIXED [Person]** : COUNTD([Row_ID])}

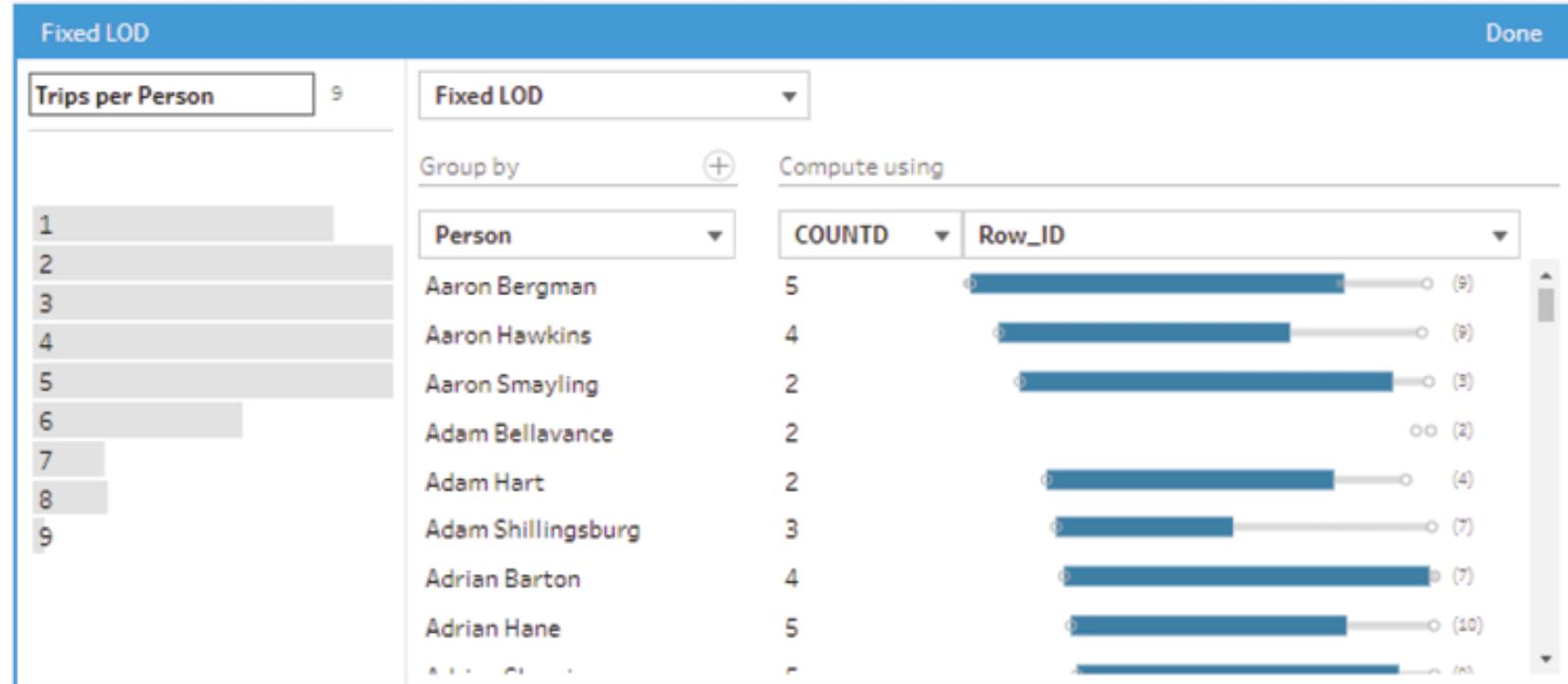
Abc

Person 765



- Filter
- Clean
- Group Values
- Split Values
-
- View State
- Detail
- Summary
-
- Rename Field
- Duplicate Field
- Keep Only Field
- Create Calculated Field ►
- Custom Calculation
- Fixed LOD
- Rank
-
- Remove

Level of detail calculations

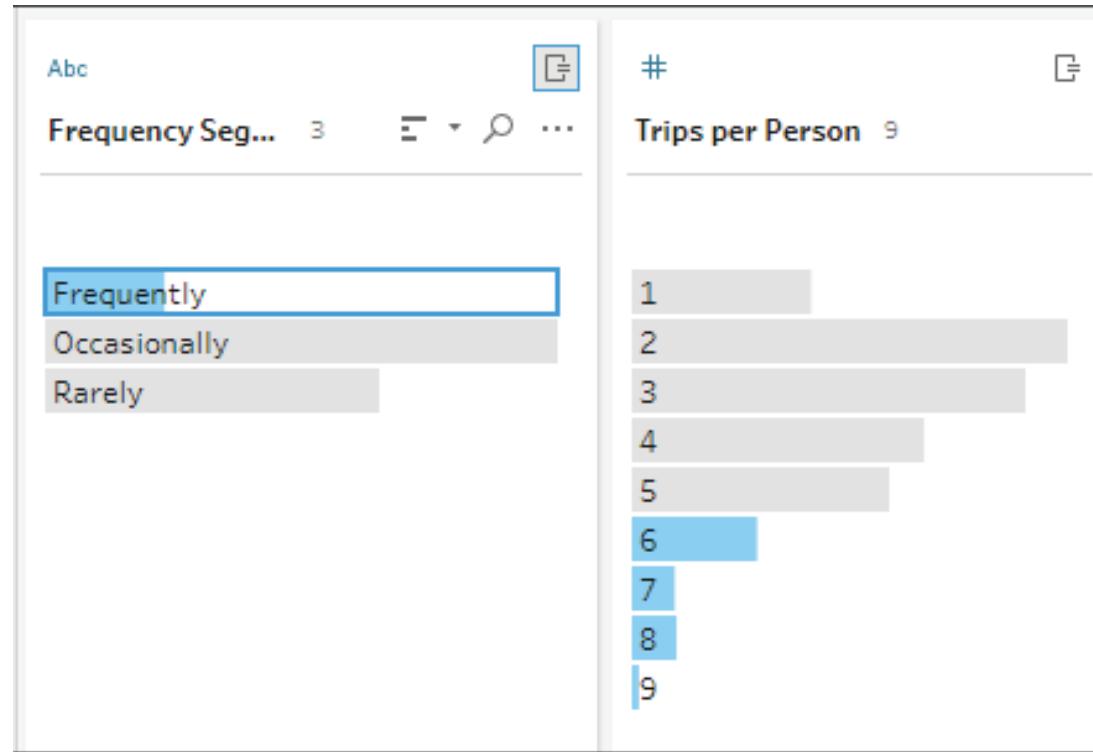


Level of detail calculations

- Name the field Frequency Segment and enter the following code:

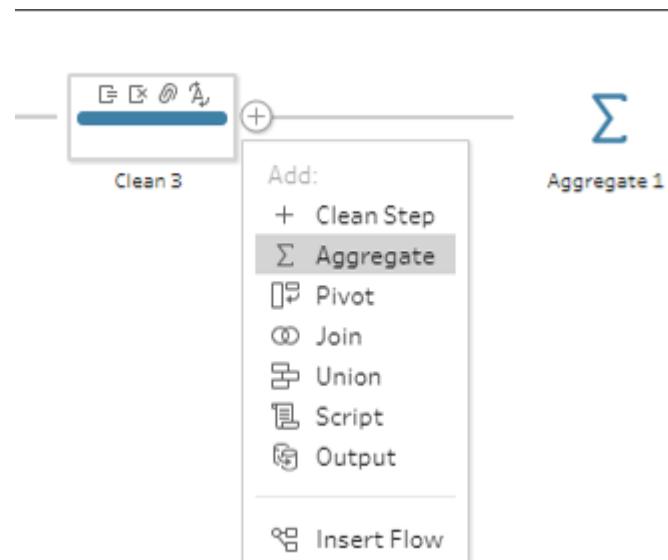
```
IF [Trips per Person] <= 2 THEN "Rarely"  
ELSEIF [Trips per Person] <= 5 THEN "Occasionally"  
ELSE "Frequently"  
END
```

Level of detail calculations



Aggregations

- Click the + symbol on Clean 3 and add an Aggregate step.
- The new step will be named Aggregate 1 by default:



Frequency Segment 2 Fields 3 Rows Filter Values... Search

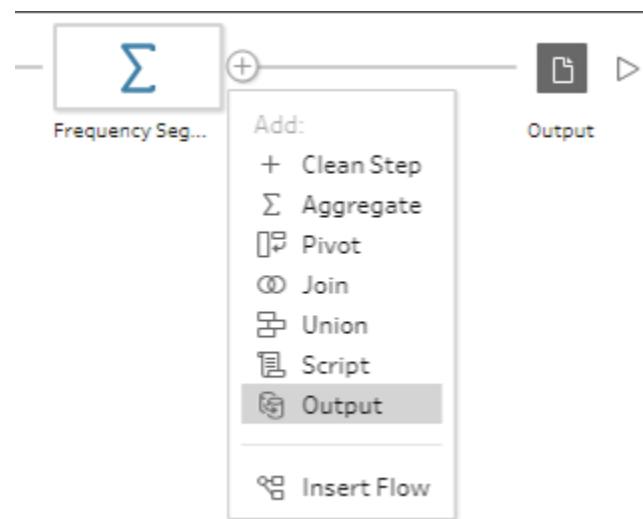
Settings		Changes (0)
Additional Fields Drag fields to aggregate or group them. <input type="text" value="Search"/>		
<input type="button" value="Add All"/> <input type="button" value="Remove All"/>		
Abc	GROUP	Airline
#	SUM	Days from Purchase to Travel
Abc	GROUP	File Paths
Abc	GROUP	Passenger Email
Abc	GROUP	Person
	GROUP	Purchase Date
Abc	GROUP	Route
#	SUM	Row_ID
Abc	GROUP	Ticket Type
	GROUP	Travel Date
#	SUM	Trips per Person
#		Number of Rows (Aggregated)

Grouped Fields	
Abc	GROUP
Frequency Segment 3	
Frequently	
Occasionally	
Rarely	

Aggregated Fields	
#	AVG
Ticket Price 3	
149.56439393939394	
173.5090534979424	
181.2991150442478	

Aggregations

- Conclude the example by clicking the + icon that appears when you hover over the Frequency Segment aggregate step and adding an Output step:



Output 2 Fields

Save output to



File

[Browse](#)

Name

Frequency Segment

Location

C:\..\Datasources

Output type

Comma Separated Values (.csv)

Write Options

Select an option to create or update your output table.

Full refresh

[Create table](#)

[Run Flow](#)

Save to Frequency Segment.csv

Frequency Segment	Ticket Price
Frequently	181.2991150442478
Rarely	149.56439393939394
Occasionally	173.5090534979424

Filtering data in Tableau Prep

There are a couple of ways to filter data in Tableau Prep:

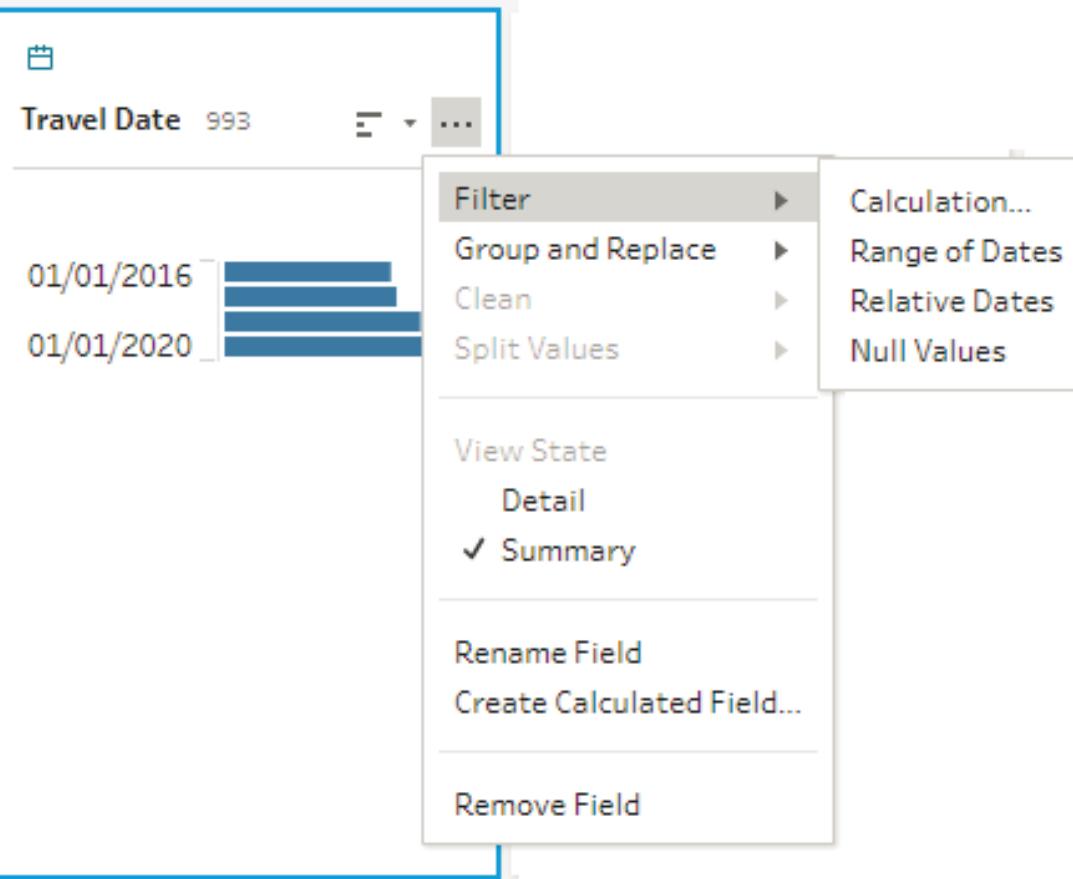
- Filter an input
- Filter within the flow

Filtering data in Tableau Prep

- To filter an input, select the input step and then click the Filter Values... button on the input pane:

The screenshot shows the Tableau Prep Input pane for an 'Employee Flights' flow. At the top, it displays 'Fields selected: 10 of 10' and a 'Filter Values...' button. Below this, a message says 'Select the fields to include in your flow. If you make changes to the data, the data source will be queried again.' A table lists four fields: Row ID, Passenger Email, Purchase Date, and Travel Date. Each row includes a checked checkbox, a type indicator (e.g., #, Abc), the field name, the original field name, changes made, and sample values.

<input checked="" type="checkbox"/>	Type	Field Name	Original Field Name	Changes	Sample Values
<input checked="" type="checkbox"/>	#	Row ID	Row ID		4,144, 9,102, 2,194
<input checked="" type="checkbox"/>	Abc	Passenger Email	Passenger Email		eklefer@vizpainter.com, abarton@vizpainter.com, achung@vizpainter.com
<input checked="" type="checkbox"/>	Alph	Purchase Date	Purchase Date		01/10/2019, 09/10/2017, 10/22/2018
<input checked="" type="checkbox"/>	Alph	Travel Date	Travel Date		02/06/2019, 11/12/2017, 12/17/2018



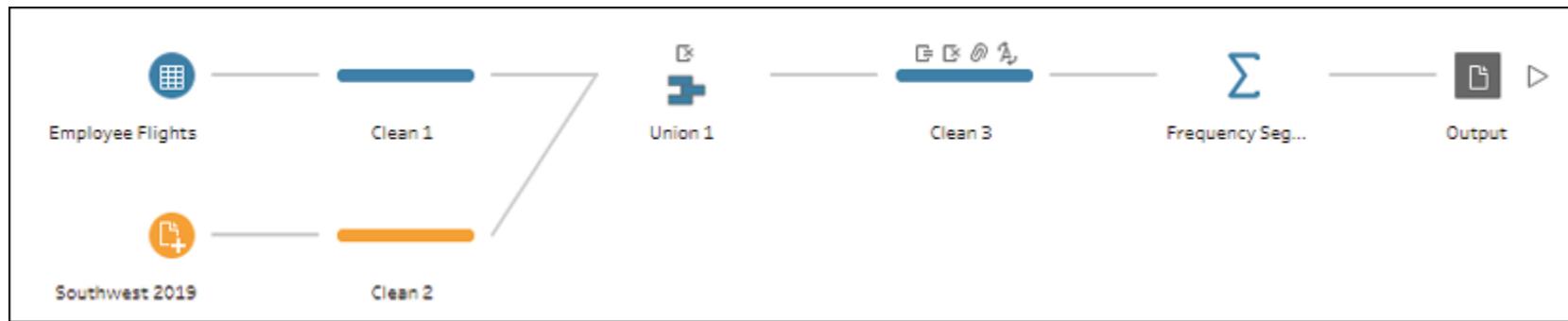
Filtering data in Tableau Prep

- Select a field and then Filter Values from the toolbar.
- Similar to the way filters work in the input pane, you will be prompted to write code that returns true for records you wish to keep.
- If, for example, you wanted to keep records for trips scheduled after January 1, 2016, you could write code such as the following:

[Travel Date] > MAKEDATE(2016, 1, 1)

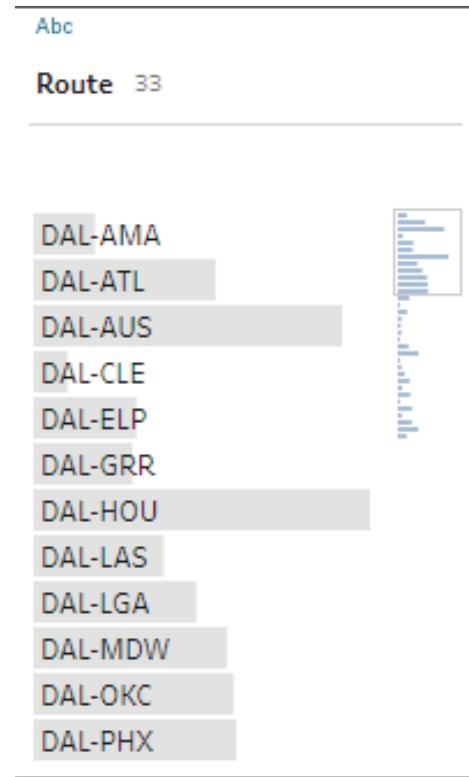
Filtering data in Tableau Prep

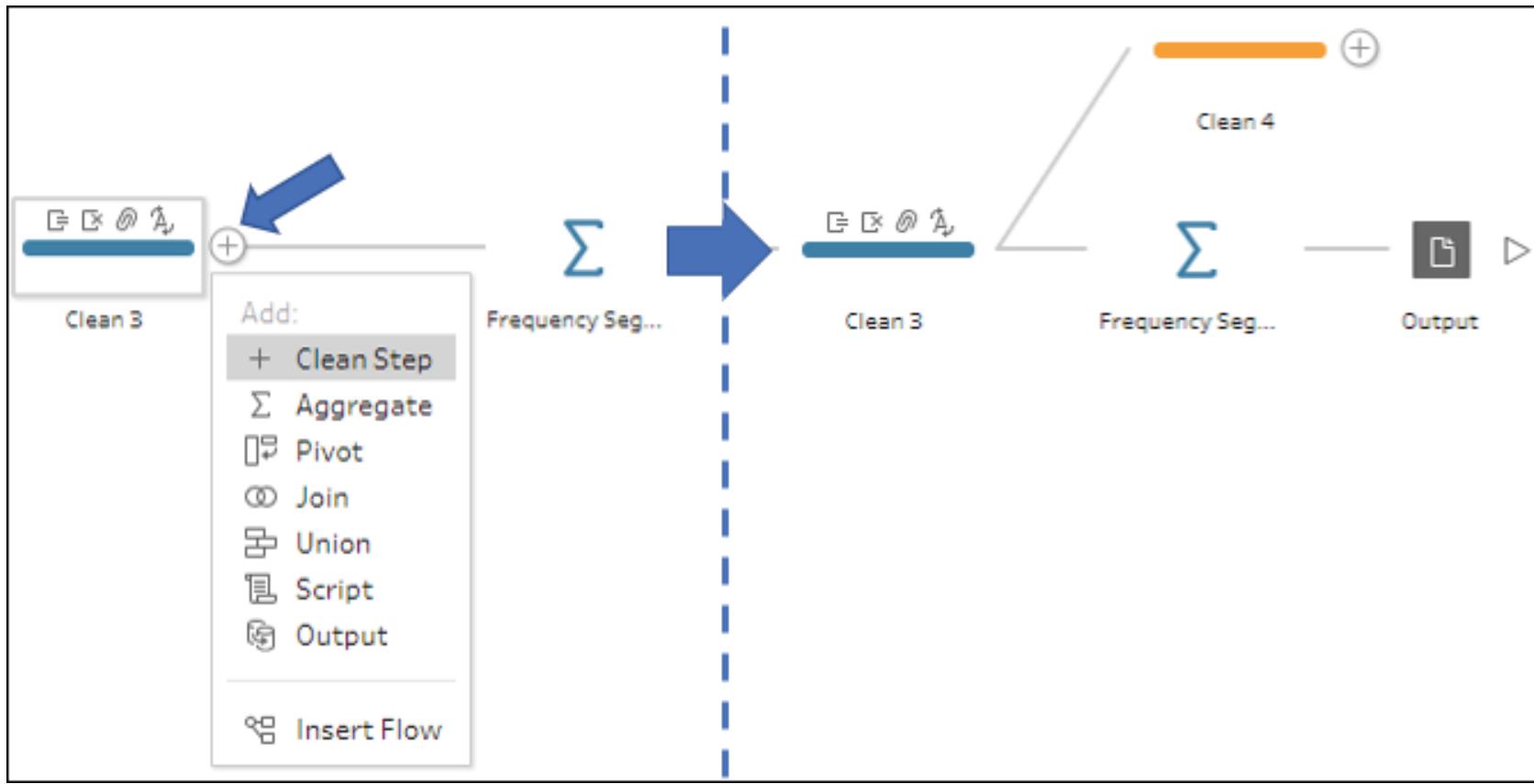
- At this point, your flow should look something like this:



Transforming the data for analysis

- Let's create a new branch in the flow to work once again with the detailed data.
- Click on the Clean 3 step and examine the preview pane.
- In particular, consider the Route field:





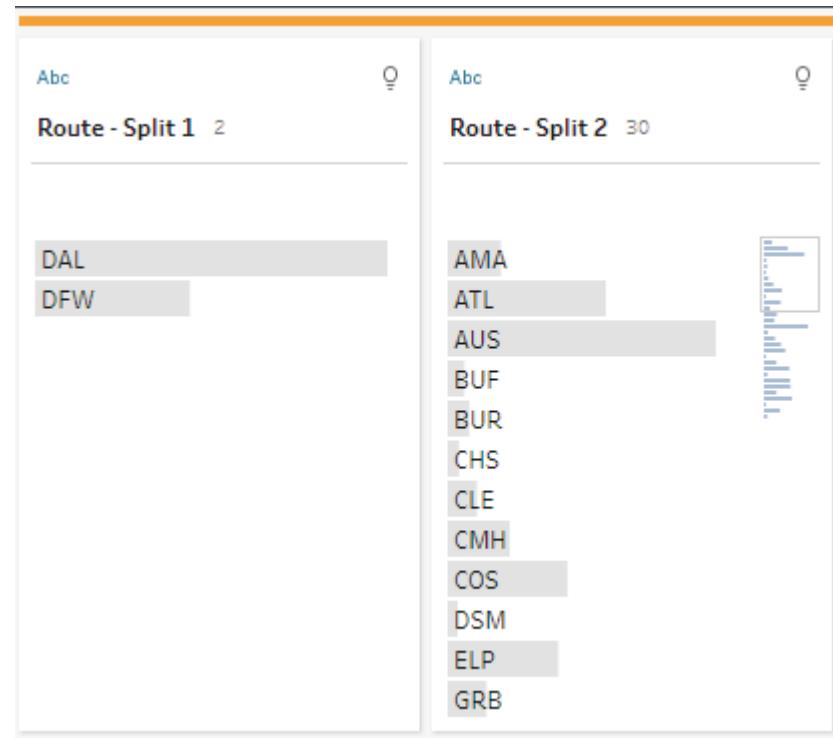
The screenshot shows a Microsoft Power BI interface. On the left, there is a list of cities with codes: DAL-AMA, DAL-ATL, DAL-AUS, DAL-CLE, DAL-ELP, DAL-GRR, DAL-HOU, DAL-LAS, DAL-LGA, DAL-MDW, DAL-OKC, and DAL-PHX. The 'DAL-HOU' entry is currently selected. At the top, there is a search bar with the placeholder 'Route 33' and a magnifying glass icon. To the right of the search bar is a three-dot ellipsis button, which has been clicked to open a context menu. The menu contains the following items:

- Filter
- Clean
- Group Values
- Split Values
 - Automatic Split
 - Custom Split...
- View State
 - Detail
 - Summary
- Rename Field
- Duplicate Field
- Keep Only Field
- Create Calculated Field
- Publish as Data Role...
- Remove

A blue arrow points from the top center towards the ellipsis button, indicating the action of opening the context menu.

Transforming the data for analysis

- You'll now see two new fields added to the step:

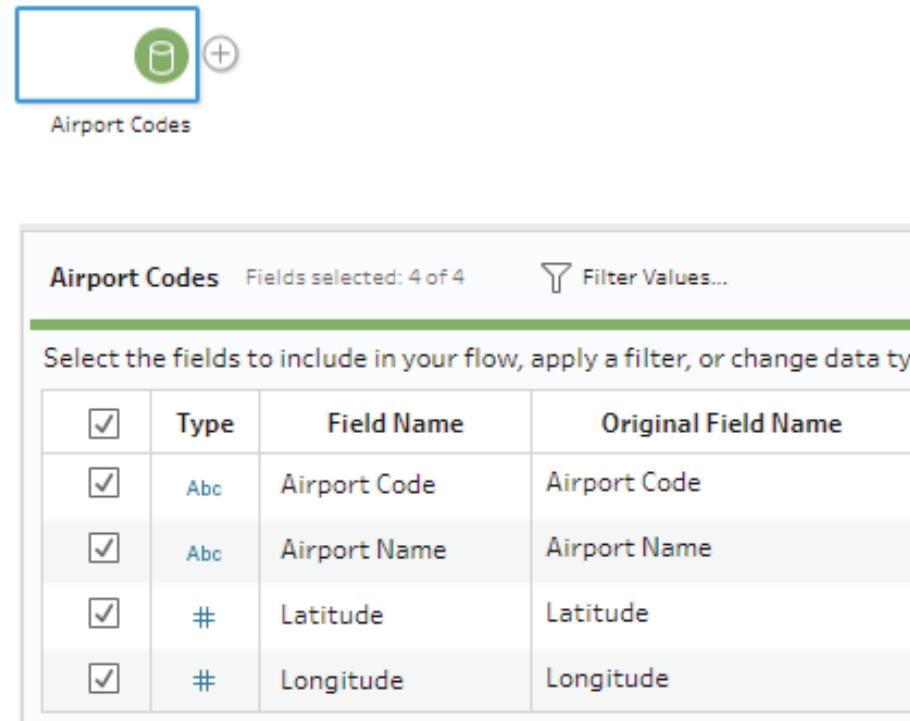


Transforming the data for analysis

- Locate the US Airports.hyper file in the lesson 15 directory.
- This file contains each Airport Code along with the Airport Name and Latitude and Longitude:

Airport Code	Airport Name	Latitude	Longitude
BTI	Barter Island LRRS Airport	70.1340	-143.582
LUR	Cape Lisburne LRRS Airport	68.8751	-166.110
PIZ	Point Lay LRRS Airport	69.7329	-163.005
ITO	Hilo International Airport	19.7214	-155.048
ORL	Orlando Executive Airport	28.5455	-81.333

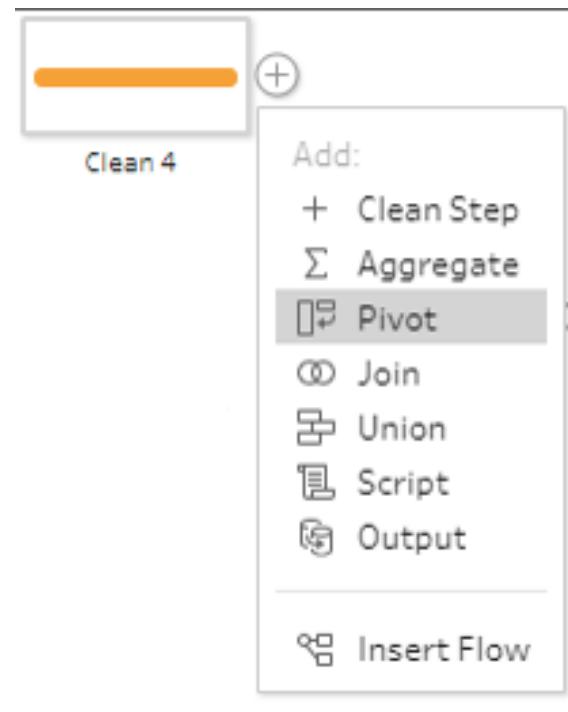
Transforming the data for analysis



The screenshot shows the 'Transform Data' interface in Microsoft Power BI. At the top, there's a green icon with a white cylinder and a plus sign, followed by the text 'Airport Codes'. Below this, the main title 'Airport Codes' is displayed along with the message 'Fields selected: 4 of 4' and a 'Filter Values...' button. A green bar below the title contains the instruction 'Select the fields to include in your flow, apply a filter, or change data type'. The data is presented in a table with four columns: 'Type', 'Field Name', and 'Original Field Name' (which is identical to the 'Field Name'). All five rows in the table have their 'Type' column set to 'Abc' and their 'Field Name' and 'Original Field Name' columns set to 'Airport Code', 'Airport Name', 'Latitude', and 'Longitude' respectively. Each row has a checked checkbox in the first column.

Type	Field Name	Original Field Name
Abc	Airport Code	Airport Code
Abc	Airport Name	Airport Name
#	Latitude	Latitude
#	Longitude	Longitude

Transforming the data for analysis



Keep the default of Columns to Rows

The screenshot shows the Power BI interface with two main sections: 'Pivoted Fields' on the left and 'Pivot Results' on the right.

Pivoted Fields: This section contains two columns: 'Pivot1 Names' and 'Pivot1 Values'. Under 'Pivot1 Names', there are two items: 'Origin' and 'Destination'. Under 'Pivot1 Values', there are two items: 'Abc' (under 'Origin') and 'Abc' (under 'Destination'). A blue arrow points from the text 'Keep the default of Columns to Rows' down to the 'Columns to Rows' button in the top center of the Pivoted Fields area.

Pivot Results: This section displays the transformed data. It includes a table with four columns: 'Pivot1 Values', 'Pivot1 Names', 'Frequency Segment', and 'Trips per Person'. Below this table is a list of cities: AMA, ATL, AUS, BUF, BUR, CHS, CLE, CMH, COS, DAL, DFW, and DSM. The 'DFW' entry in the list is highlighted with a gray bar.

Pivot1 Values	Pivot1 Names	Frequency Segment	Trips per Person
CMH	Destination	Occasionally	3
DFW	Origin	Occasionally	3
DFW	Origin	Occasionally	4
CMH	Destination	Occasionally	4

Transforming the data for analysis

The screenshot shows the Power BI Data Flow interface with two tables side-by-side.

Left Table (Origin):

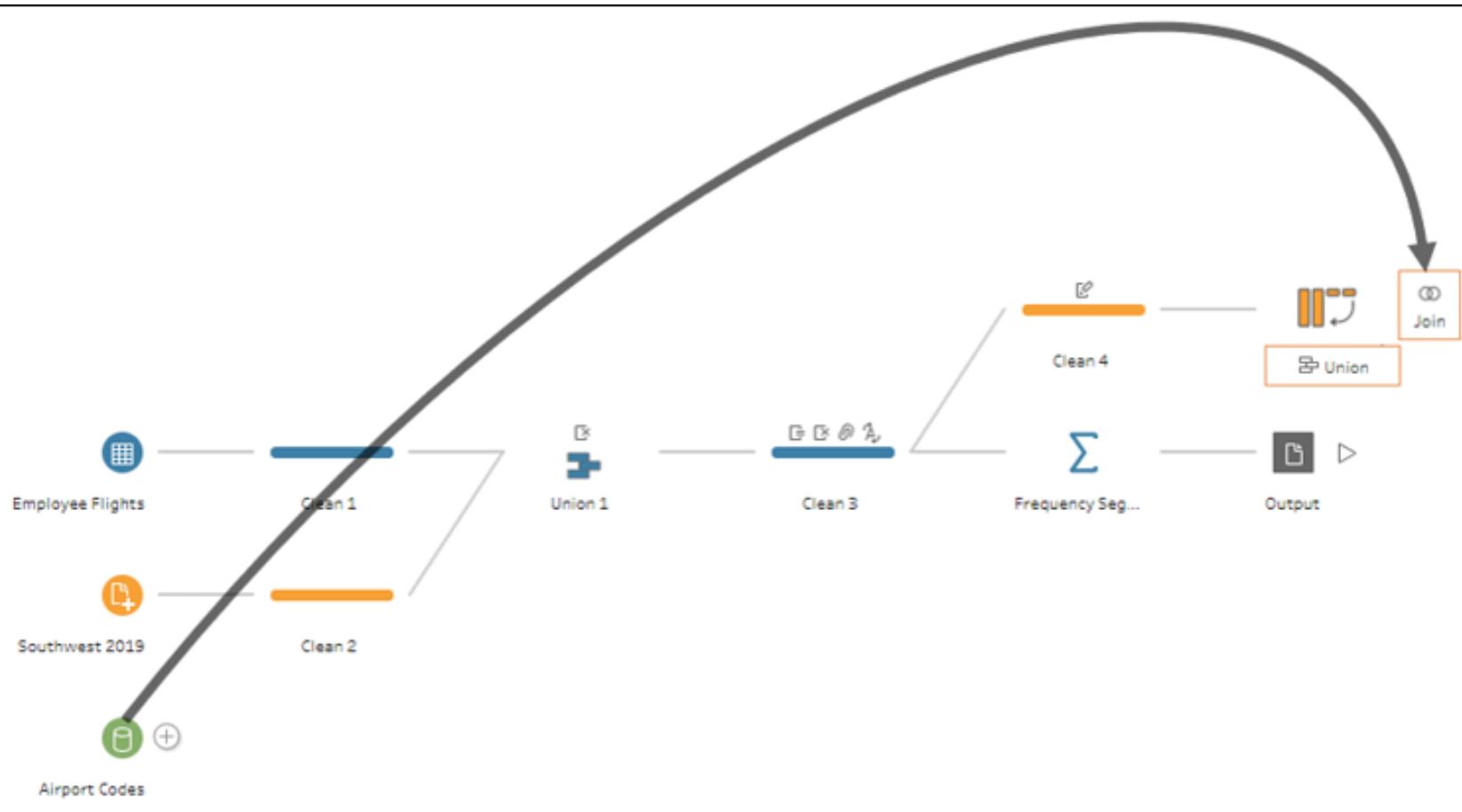
- Table name: Origin
- Rows: 2
- Columns: 1 (DAL, DFW)

Right Table (Destination):

- Table name: Destinati...
- Rows: 30
- Columns: 1 (AMA, ATL, AUS, BUF)

A context menu is open on the right table, showing the following options:

- Merge
- Keep Only
- Remove
- Pivot Columns to Rows



Join 1 19 Fields, 9K Rows

Filter Values...

Create Calculated Field...

Settings

Changes (0)

1 Applied Join Clauses

Pivot 1 Airport Codes

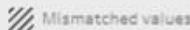
Airport Code = Airport Code

2 Join Type : inner

Click the graphic to change the join type.

**3 Summary of Join Results**

Click the bar segments to view the included and excluded values.



	Included	Excluded
Pivot 1	9,158	0
Airport C...	32	1,301
Join Result	9,158	

4 Join Clause Recommendations

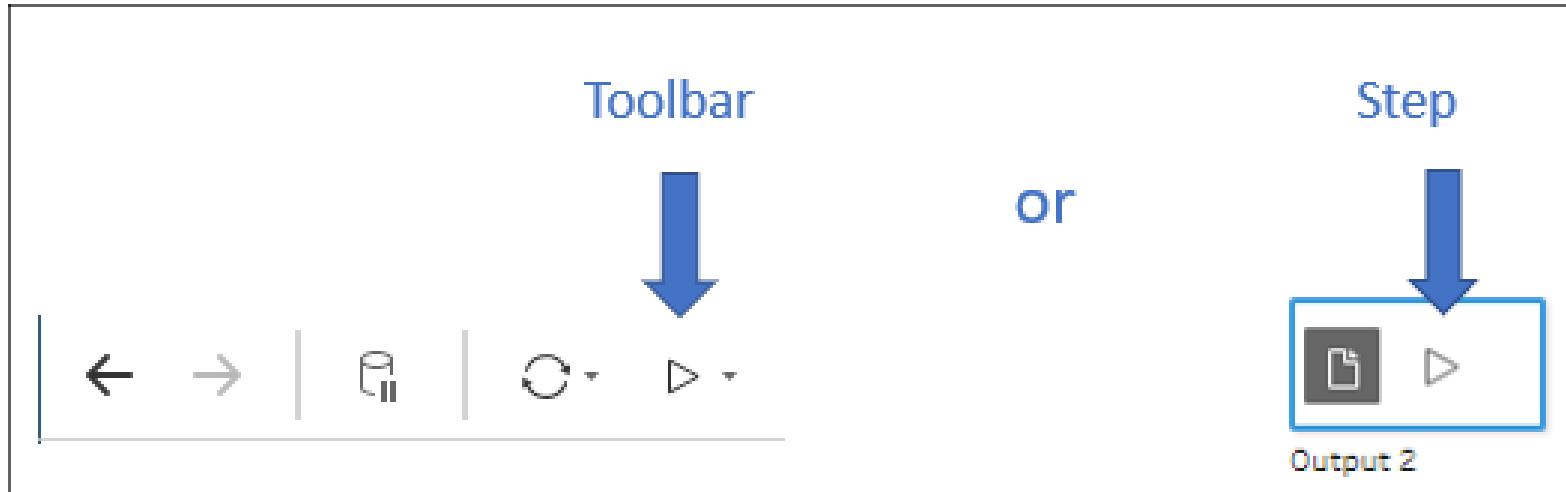
Airport Codes = Airport Name

5Join Clauses Show only mismatched values

Pivot 1	Airport Codes
↑ Airport Code	↑ Airport Code
AMA	0V4
ATL	19S
AUS	23M
BUF	2A5
BUR	2H0
CHS	2K7
CLE	3AU
CMH	4A7
COS	4U9
DAL	52A
DFW	57C
DSM	5B2
ELP	6J4
GRB	650
GRR	7W6
HIO	87K
HOU	A39
IAD	A50

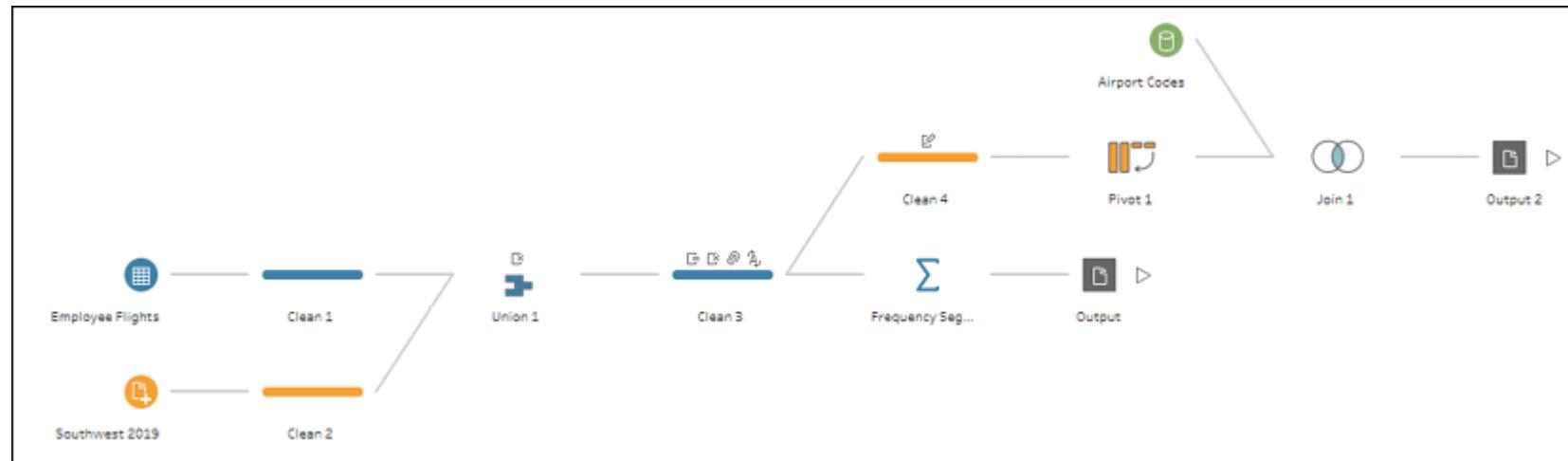
Transforming the data for analysis

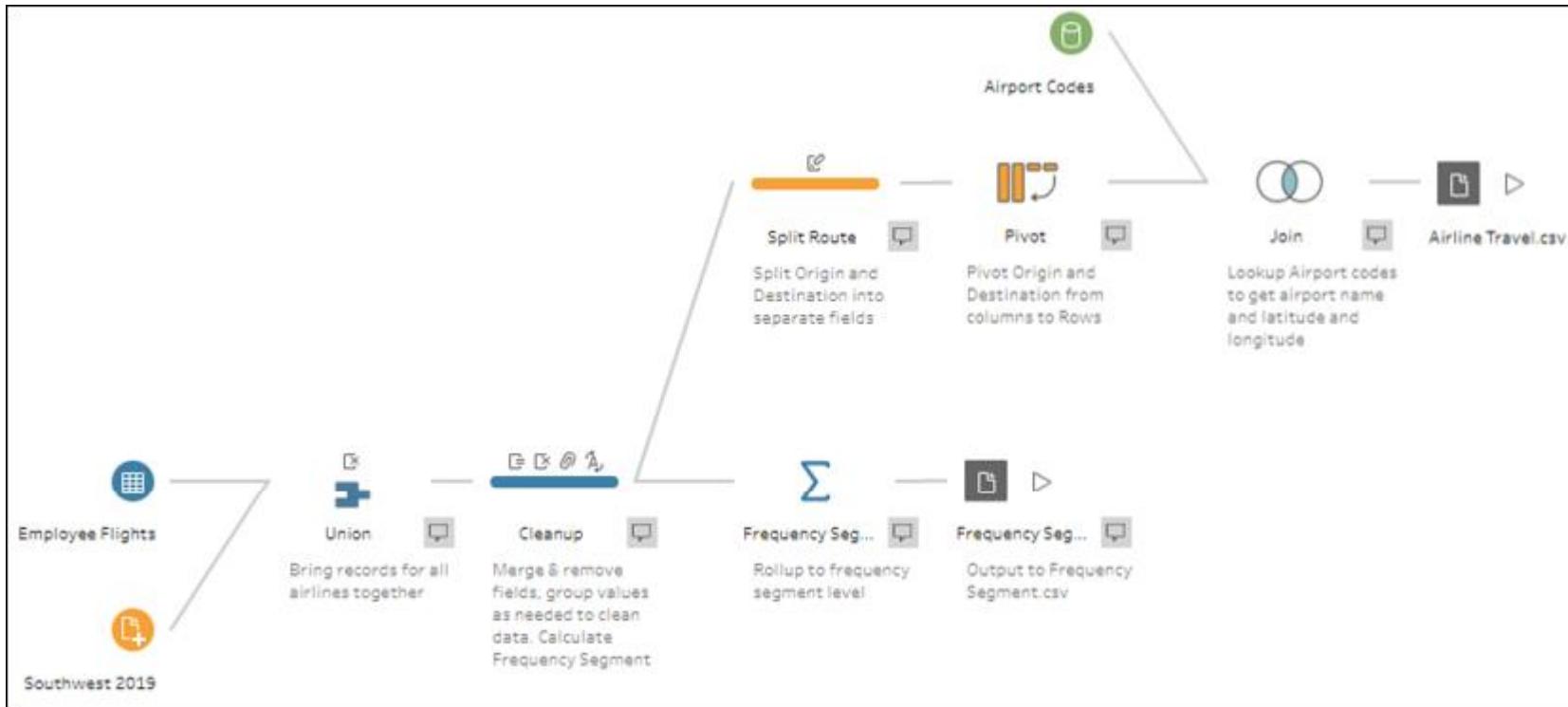
- Run the flow by using the run button at the top of the toolbar or by clicking the run button on the output step.

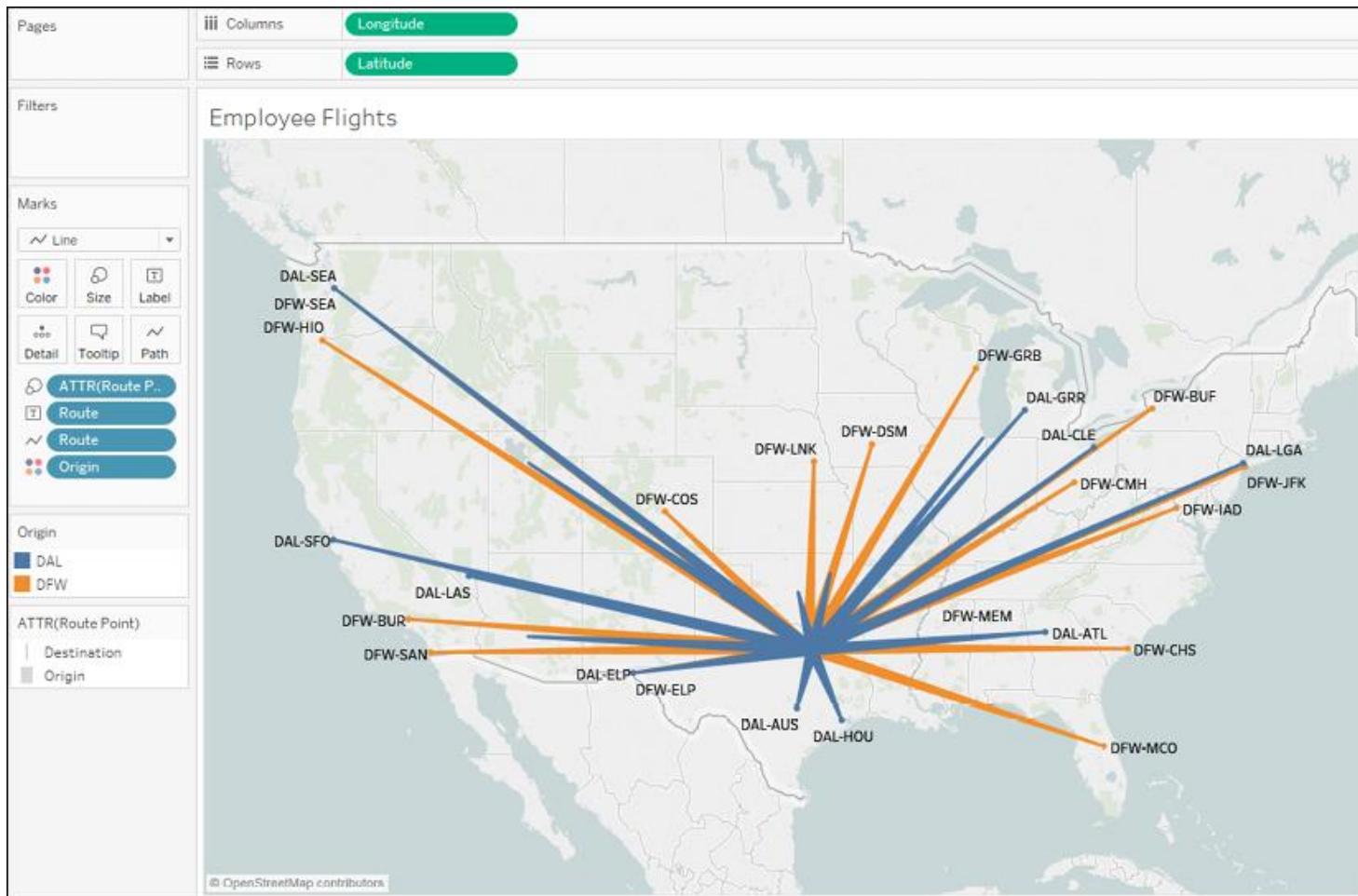


Transforming the data for analysis

- Your final flow will look something like this:







Options for automating flows

Tableau Prep Builder allows you to design and run flows using the application. Sometimes, data cleansing and prepping will be a one-time operation to support an ad hoc analysis. However, you will often want to run a flow subsequently to capture new or changed data and to cleanse and shape it according to the same pattern. In these cases, you'll want to consider some options for automating the flow:

- Tableau Prep Builder
- Tableau Prep Conductor

Summary

- Tableau Prep's innovative paradigm of hands-on data cleansing and shaping with instant feedback greatly extends the Tableau platform and gives you incredible control over your data.
- In this lesson, we considered the overall interface and how it allows you to iteratively and rapidly build out a logical flow to clean and shape data for the desired analysis or visualization.

COMPLETE LAB 15

16. Sharing Your Data Story

Sharing Your Data Story

We'll look at the following topics:

- Presenting, printing, and exporting
- Sharing with users of Tableau Desktop and Tableau Reader
- Sharing with users of Tableau Server, Tableau Online, and Tableau Public

Presenting, printing, and exporting

- Tableau is primarily designed to build richly interactive visualizations and dashboards for consumption on a screen.
- Often, you will expect users to interact with your dashboards and visualizations.

Presenting

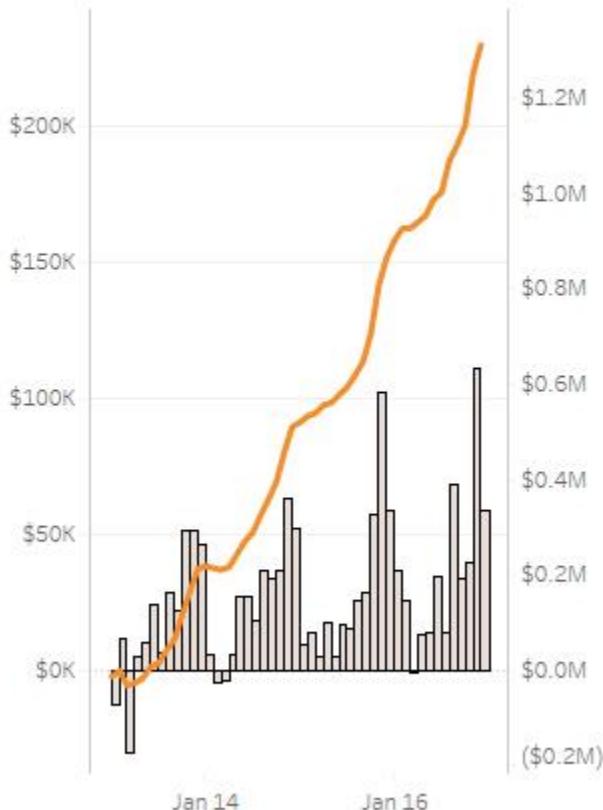
Tableau gives you multiple options for personally presenting your data story. You might walk your audience through a presentation of a single dashboard or view, or you might create an entire presentation. While there are multiple ways you might structure a presentation, consider the following options:

- Exporting to PowerPoint
- Presentation mode

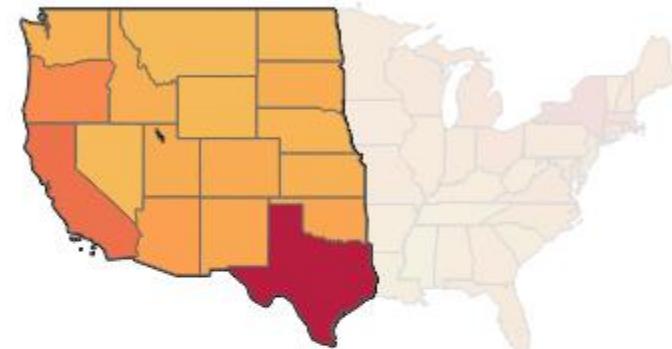
Profit Analysis

Printing

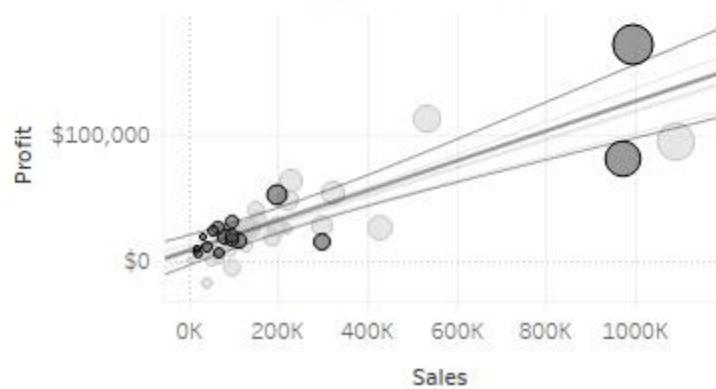
Profit over Time



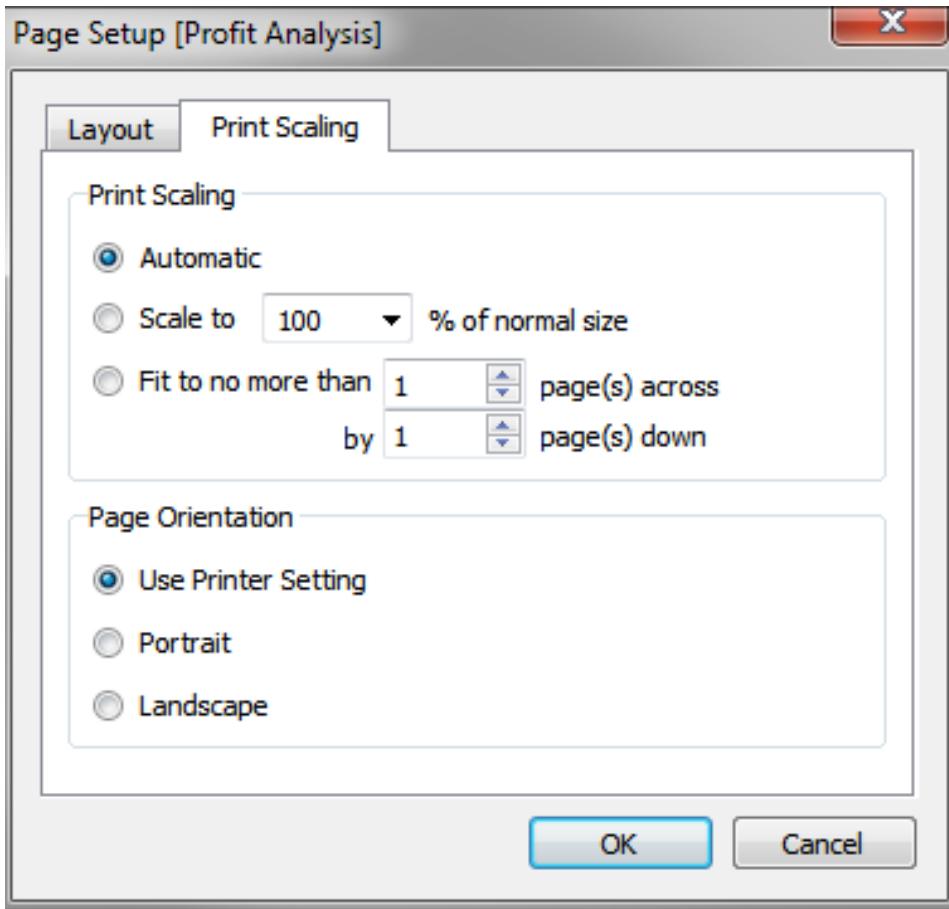
Profit by State

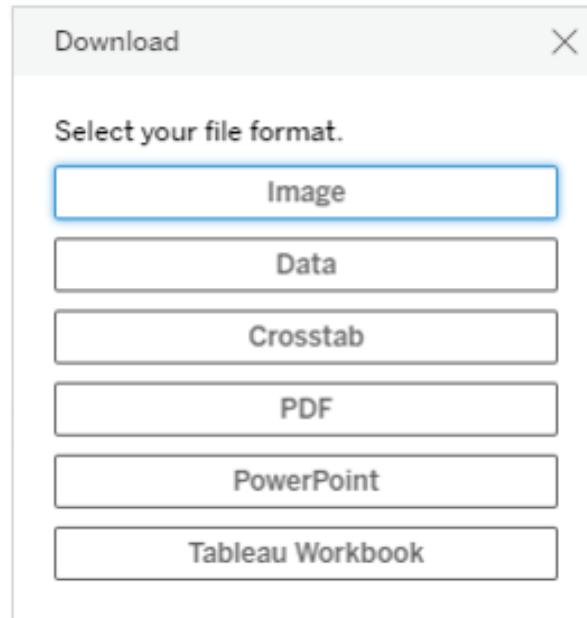


Profit and Sales Correlation (State)



Printing





Sharing with users of Tableau Desktop & Tableau Reader

Sharing with Tableau Desktop users

- Sharing a workbook with other Tableau Desktop users is fairly straightforward, but there are a few things to consider.
- One of the major considerations is whether you will be sharing a packaged workbook (.twbx) or an unpackaged workbook (.twb).

Sharing with Tableau Reader users

- Tableau Reader is a free application provided by Tableau Software that allows users to interact with visualizations, dashboards, and stories created in Tableau Desktop.
- Unlike Tableau Desktop, it does not allow for the authoring of visualizations or dashboards.
- However, all interactivity, such as filtering, drilldown, actions, and highlighting, is available to the end user.

Sharing with users of Tableau Server, Tableau Online, and Tableau Public

	Tableau Server	Tableau Online	Tableau Public
Description	A server application installed on one or more server machines that hosts views and dashboards created with Tableau Desktop.	A cloud-based service maintained by Tableau Software that hosts views and dashboards created with Tableau Desktop.	A cloud-based service maintained by Tableau Software that hosts views and dashboards created with Tableau Desktop or the free Tableau Public client.
Licensing cost	Yes	Yes	Free
Administration	Fully maintained, managed, and administered by the individual or organization that purchased the license.	Managed and maintained by Tableau Software with some options for project and user management by users.	Managed and maintained by Tableau Software.
Authoring and publishing	Users of Tableau Desktop may author and publish workbooks to Tableau Server. Web Authoring allows Tableau Server users the capability to edit and create visualizations and dashboards in a web browser.	Users of Tableau Desktop may author and publish workbooks to Tableau Online. Web Authoring allows Tableau Online users the capability to edit and create visualizations and dashboards in a web browser.	Users of Tableau Desktop or the free Tableau Public client can publish workbooks to Tableau Public. Future enhancements to allow online authoring have been announced.
Interaction	Licensed Tableau Server users may interact with hosted views. Views may also be embedded in intranet sites, SharePoint, and custom portals.	Licensed Tableau Online users may interact with hosted views. Views may also be embedded in intranet sites, SharePoint, and custom portals.	Everything is public-facing. Anyone may interact with hosted views. Views may be embedded in public websites and blogs.

	Tableau Server	Tableau Online	Tableau Public
Limitations	None.	Most data sources must be extracted before workbooks can be published. Most non-cloud-based data sources must have extracts refreshed using Tableau Desktop on a local machine or through the Tableau Online Sync Client .	All data must be extracted and each data source is limited to 15 million rows.
Security	The Tableau Server administrator may create sites, projects, and users and adjust permissions for each. Access to the underlying data can be restricted, and downloading of the workbook or data can be restricted.	The Tableau Server administrator may create projects and users and adjust permissions for each. Access to the underlying data can be restricted, and downloading of the workbook or data can be restricted.	By default, anyone may download and view data; however, access to these options may be restricted by the author.
Good uses	Internal dashboards and analytics and/or use across departments/divisions/clients through multi-tenant sites.	Internal dashboards and analytics, especially where most data sources are cloud-based. Sharing and collaboration with remote users.	Sharing visualizations and dashboards using embedded views on public-facing websites or blogs.

Publishing to Tableau Server and Tableau Online



Sign in to Tableau Online

jmilligan@teknionusa.com

.....

Remember me

Sign In

Publish Workbook to Tableau Online

X

Project

Default



Name

Engage



Description

Tags

Add

Sheets

All [Edit](#)

Permissions

Same as project (**Default**) [Edit](#)

Data Sources

1 embedded in workbook [Edit](#)

More Options

Show sheets as tabs

Show selections

Publish

Publishing to Tableau Server and Tableau Online

Permissions for Workbook "Borg Cube"

Sheet tabs shown: Permissions for views inherited from workbook

Permission Rules

Group/User	Template	Eye	Filter	Download	Upload	Print	Copy	Move	Share	Comment	Edit	Delete	Checkmark
All Users	Custom	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

+ Add Group/User Rule

Effective Permissions

Search for a user to view their effective permissions

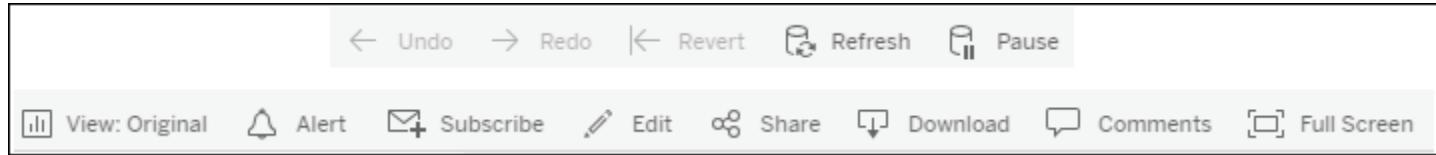
User	Site Role	Eye	Filter	Download	Upload	Print	Copy	Move	Share	Comment	Edit	Delete	Checkmark
Locutus	Creator	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✗
Hugh	Explorer (can pu...	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✗
ThreeOfFive	Explorer (can pu...	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✗
Dna	Explorer (can pu...	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✗
BQueen	Server Administ...	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
SevenOfNine	Creator	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✗

Interacting with Tableau Server

- After a workbook is published to Tableau Server, other users will be able to view and interact with the visualizations and dashboards using a web browser.
- Once logged in to Tableau Server, they will be able to browse content for which they have appropriate permissions.
- These users will be able to use any features built into the dashboards, such as quick filters, parameters, actions, or drilldowns.

Interacting with Tableau Server

- You'll find a toolbar that includes several features:



Summary

- Tableau is an amazing platform for exploring, prepping, and cleaning your data as you create useful and meaningful visualizations and dashboards to understand and communicate key insights.
- Throughout this course, we've considered how to connect to data—whether file-based, in an on-premises database, or in the cloud.

COMPLETE LAB 16