

Lesson Change Log AN 4.1 - 4.2

Below are the specific changes made in this lesson.

Returning Instructor Directions:

1. Click on the hyperlinks below, and it will direct you to the specific slide.
2. Copy the entire slide, and paste in your existing curriculum deck.
 - [Connecting to Superstore Data](#): Updated verbiage. We now have the ability to see what products were returned. Added a Talking Point.
 - [Adding More Data](#): Updated to show relationship on order_id and product_id
 - [Adding Labels](#): Number of records is now orders.csv(Count). Updated how to add label
 - [Building a Stacked Bar Chart](#): Number of records is now orders.csv(Count).

Share how the lesson went through our Instructor Lesson Exit Ticket - the Curriculum Feedback form: <http://ga.co/curriculum-feedback>



Visualizing Data With Tableau

Overview

Import data into Tableau and navigate the interface to build graphs, calculations, dashboards, and stories — more specifically, SQL versus Tableau (VizQL) as well as CAST, aggregation types, and worksheets versus dashboards versus stories.

Learning Objectives

In this lesson, students will:

- Prepare data for import into Tableau.
- Navigate the Tableau interface to build visualizations.
- Aggregate measures and dimensions.
- Work with Marks Card and discrete versus continuous dates.
- Connect to the PostgreSQL server.
- Create calculated fields to analyze data.
- Apply filters to single or multiple sheets.
- Create data visualizations and simple dashboards in Tableau



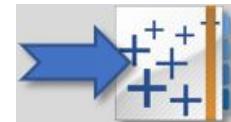
Pre-Class Materials and Preparation

Accessing Tableau: Copy and paste the product key provided by your IM on the “Before We Begin...” slide.

All Tableau workbooks for Superstore can be found in [this folder](#). For this lesson, we'll need:

- AN4.2_Tableau_Lesson_1_Activities.twbx (*to be used in the morning*)
- AN4.2_Tableau_Lesson_1_SOLUTIONS.twbx
- AN4.2_Tableau_Lesson_2_Activities.twbx (*to be used in the afternoon*)
- AN4.2_Tableau_Lesson_2_SOLUTIONS.twbx

Look for this icon to switch over from the slides to Tableau.
You'll often need to switch back and forth between the two.



Pre-Class Materials and Preparation

Optional Practice: At the end of each section, you'll find solo exercises titled “Optional Practice”. If time permits, students can work on them at the end of a section or these can be grouped together to be completed during **Open Workshop** or **Practice Application** hours. See the  [Lesson Pacing and Schedule](#) for more details.

Pre-Class Materials and Preparation

For Remote classrooms: Virtual breakout rooms and Slack may be needed to facilitate partner/group exercises and discussions. As you plan for your lesson:

- Consider the number and size of groups that would be appropriate for these exercises and discussions.
- Determine if exercise timing may need to be adjusted.
- Keep an eye out for the tag **For Remote Classrooms** in the speaker notes for helpful tips.
- Prepare screenshots and answers to exercises in advance so they can be easily shared in Slack during your lecture.



Suggested Agenda

Time	Activity
9:00–9:10	Welcome and Introduction
9:10–10:30	Getting Started with Tableau
10:30–10:45	Break
10:45–12:15	Marks Card + Working with Dates
12:15–1:15	Lunch + Open Workshop
1:15–2:45	Connecting to PostgreSQL
2:45–3:00	Break
3:00–4:30	Show Me + Filters + Starting a Dashboard
4:30–4:50	Practice Application
4:50–5:00	Wrap Up, Q&A, and Exit Ticket Completion



Data Analytics

Introduction to Tableau

Before We Begin...

- [Download Tableau Desktop here.](#)
- Select each product download link to get started.
 - When prompted, enter your email address for Business Email and enter General Assembly for Organization.
- Activate with your product key: **XXXX-XXXX-XXXX-XXXX-XXXX**



Our Learning Goals

In this lesson, we'll:

- Prepare data for import into Tableau.
- Navigate the Tableau interface to build visualizations.
- Aggregate measures and dimensions.
- Work with Marks Card and discrete versus continuous dates.



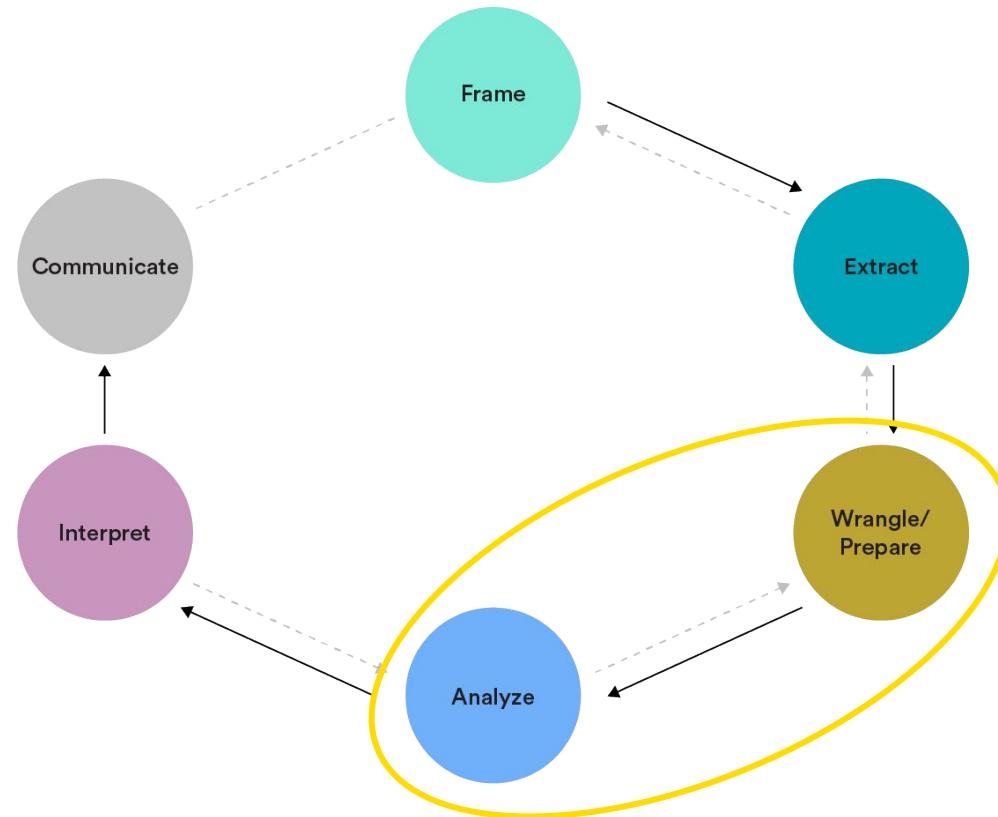
The DA Workflow

Wrangle/Prepare:

Clean and prepare relevant data.

Analyze:

Structure, comprehend, and visualize data.



Business Intelligence Tools

- Applications that transform data into business insights, fueling **data driven decisions!**
- Traditional Business Intelligence originally emerged in the 1960s as a system of sharing information across organizations.
- Modern BI solutions prioritize flexible self-service analysis, governed data on trusted platforms, empowered business users, and **accelerate time to insight.**





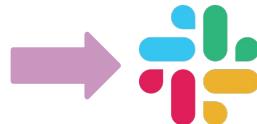
Discussion:

Consider This

In your data world, how is data analyzed?



Which reporting software do you and your team use for analysis and visualization?



Visualizing Data With Tableau

Getting Started With Tableau



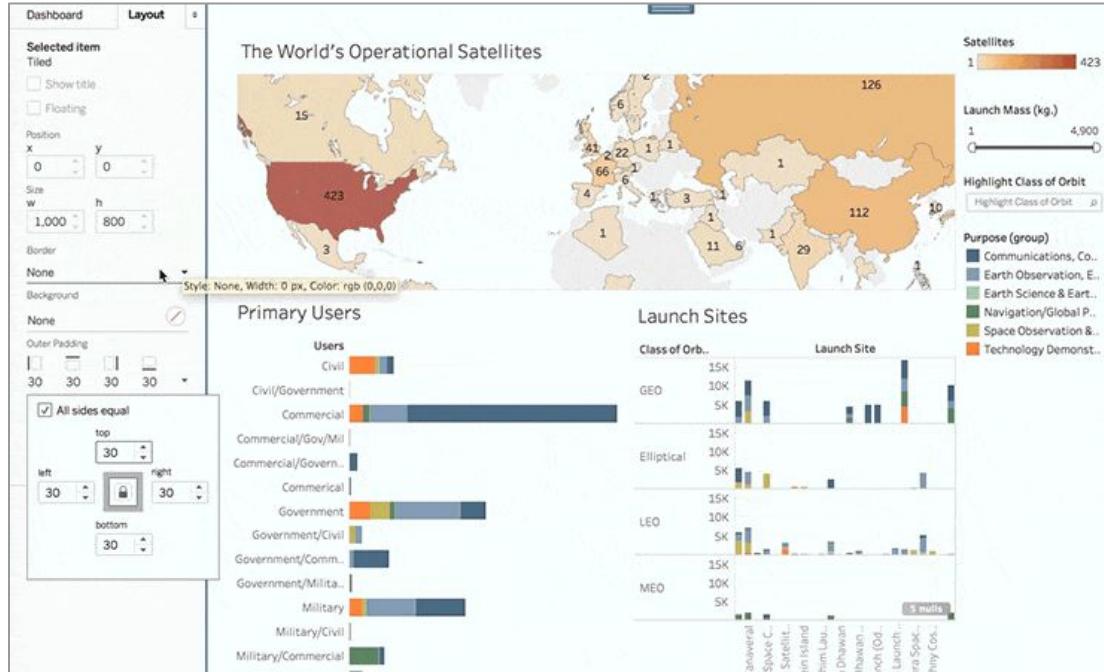
Meet Tableau

Tableau is a powerful visualization tool:

- Supports visual analytics and data discovery.
- Seamlessly creates background SQL queries called “VizQL.”
- Is designed to be used by a wide range of technical and non-technical audiences.



What Makes Tableau Stand Out?



Data

- Optimized performance for large data sets.

VizQL

- Perform visual SQL queries to combine data from multiple sources.

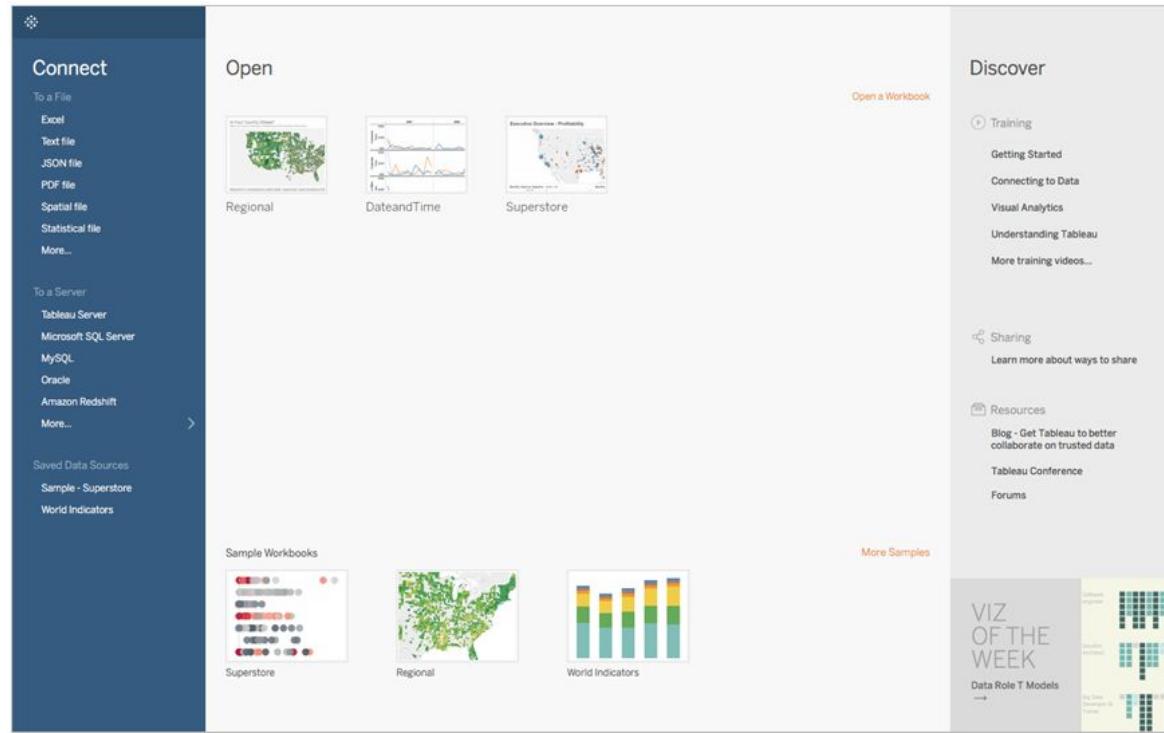
Visualization

- Interactive visual analytics.

Visualizing Data With Tableau

Navigating the Tableau Interface

Connecting to Data and Navigation





Group Exercise:

Anatomy of the Tableau Workspace

3 minutes



Every time we create a visualization in Tableau, we begin with a blank worksheet, which consists of the components shown opposite.

- a.) First name each component.**
- b.) Next define what the component is used for.**



Group Exercise:

Anatomy of the Tableau Workspace

1. Workbook name: A workbook contains sheets. A sheet can be a worksheet, a dashboard, or a story.

2. Cards and shelves: Drag fields to the cards and shelves in the workspace to add data to your **viz**.

3. Toolbar: Use the toolbar to access commands and analysis and navigation tools.

4. View: This is the canvas in the workspace where you create a visualization (also referred to as a "viz").

5. Start page: Where you can connect to data.





Group Exercise:

Anatomy of the Tableau Workspace

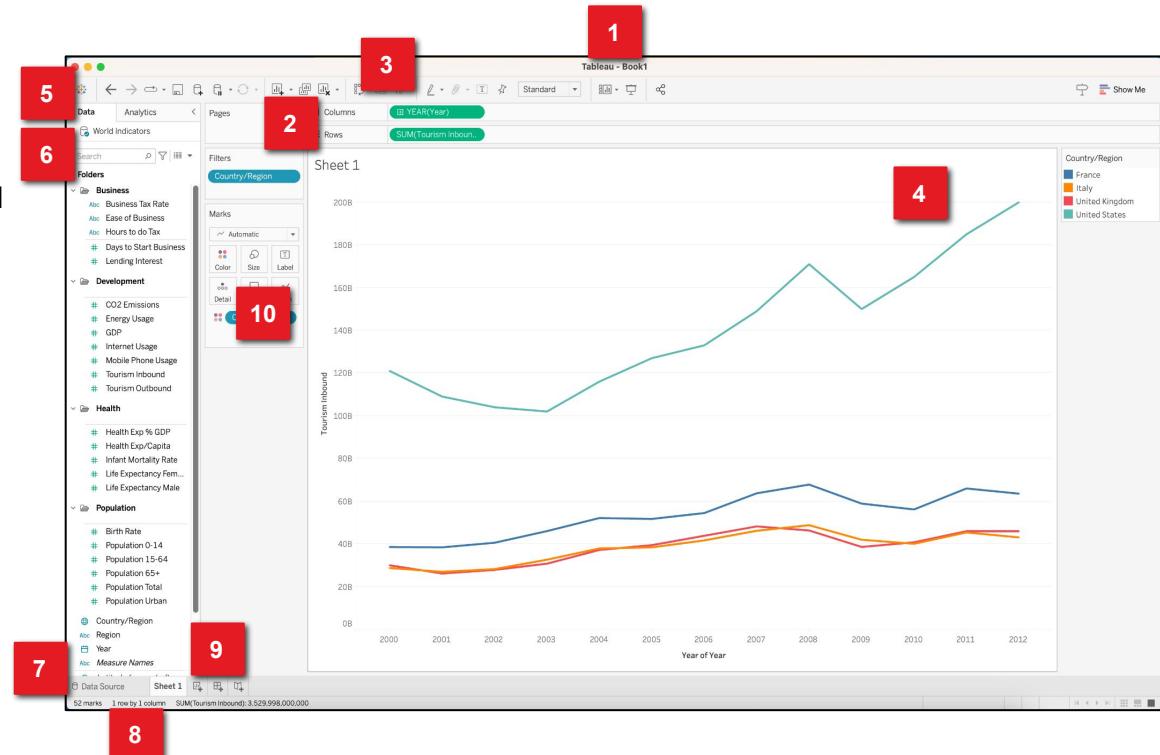
6. Side Bar: The side bar area contains the Data pane and the Analytics pane.

7. Data Source: Page where you can add and review data sources for the workbook, edit how connections are made, and view a sample of rows.

8. Status bar: Displays information about the current view.

9. Sheet tabs: Represent each sheet in your workbook. This can include worksheets, dashboards, and stories.

10. Marks card: Set the mark type to encode your data with color, size, shape, text, and detail.



Visualizing With Tableau

Worksheets	Dashboards	Story Points
<p>Enable you to:</p> <ul style="list-style-type: none">• Build a single visualization.• Use the Marks Card to add extra info to that visualization.	<p>Enable you to:</p> <ul style="list-style-type: none">• Combine multiple worksheets into a single view.• Add interactions between those visualizations.	<p>Enable you to:</p> <ul style="list-style-type: none">• Combine multiple dashboards or worksheets into an interactive full-screen presentation.

Finding Worksheets, Dashboards, and Story Points

The screenshot shows the Tableau interface with the following elements:

- Data Source:** A red box highlights the "Data Source" tab at the bottom left of the interface.
- Sheet 1:** A red box highlights the "Sheet 1" tab at the bottom left of the interface.
- Dashboard:** A blue callout bubble labeled "Dashboards" points to a dashboard area containing three cards: "Data Source", "Sheet 1", and "Story Points".
- Story Points:** A green callout bubble labeled "Story Points" points to the "Story Points" card on the dashboard.
- Worksheets:** A yellow callout bubble labeled "Worksheets" points to the "Worksheets" card on the dashboard.

The main workspace displays a data source named "orders" from a CSV file. The data includes columns such as Order Id, Order Date, Ship Date, Ship Mode, and various item details like Item ID, Category, and Sub-Category. The interface also shows connection settings (Live or Extract), filters, and a note about extracting all data.

Order Id	Order Date	Ship Date	Ship Mode	Item ID	Category	Sub-Category	Quantity	Unit Price	Line Total	Profit Margin
CA-2018-365835	28/11/2018			IM-15055	TEC-CO-10003763	2,799.96	5	0.430		
CA-2018-3658380	07/12/2018			GW-14605	FUR-FU-10003773	318.08	4	0.080		
CA-2018-3658445	27/09/2018	02/10/2018	Standard Class	HL-15040	OFF-EN-10001099	15.65	2	0.230		
CA-2018-3658511	01/03/2018	05/03/2018	Standard Class	NP-18700	FUR-TA-10001095	209.15	2	0.330		
CA-2018-3658636	09/09/2018	12/09/2018	First Class	CP-208	FUR-BEV-10000388	315.75	1	0.230		
CA-2018-3658729	04/09/2018	04/09/2018	Standard Class	NM-18520	TEC-AC-10003870	1,649.95	5	0.230		
CA-2018-3659040	21/11/2018	21/11/2018	Standard Class	RF-19345	TEC-AC-10003441	101.70	6	0.030		
CA-2018-3659285	07/12/2018	12/12/2018	Second Class	ND-18370	OFF-PA-10003641	79.14	3	0.030		
CA-2018-3659373	20/11/2018	21/11/2018	Second Class	MS-17530	OFF-AR-10002766	8.34	3	0.330		
CA-2018-3659414	04/06/2018	06/06/2018	Same Day	LW-17215	OFF-BI-10004318	276.78	2	0.230		
CA-2018-3659580	24/10/2018	25/10/2018	Standard Class	JW-16075	OFF-AR-10001955	79.36	5	0.230		
CA-2018-3659777	17/04/2018	18/04/2018	Standard Class	MH-17440	OFF-BI-10000948	34.25	3	0.230		
CA-2018-3660075	23/11/2018	25/11/2018	Second Class	SB-20290	OFF-AP-10002968	103.97	6	0.430		
CA-2018-3660094	21/11/2018	21/11/2018	Standard Class	SA-20710	OFF-PA-10003640	29.90	5	0.230		

Guided Walk-Through: Tableau Welcome Screen



Let's get started!

Open up Tableau.
You should see
something similar
to this:

Connect

To a File

- Microsoft Excel
- Text file
- JSON file
- PDF file
- Spatial file
- Statistical file
- More...

To a Server

- Tableau Server
- Microsoft SQL Server
- MySQL
- Oracle
- Amazon Redshift
- More...

Saved Data Sources

- Sample - Superstore
- World Indicators

Open

Mannie_Brown_Fin...
Mary_Kay_Cummi...
FinalProject_Food...
GA Final_Brand Ou...
Final GA_10wk Pre...

Final Project-GA-C...
Project 3 - US Food...
Final Project-GA-C...
tori-final-project
Wheres-Will-Activit...

Lesson-18-Activiti...
Lesson-17-activiti...
Wheres-Will-Activit...
Dashboard_Action...
Lesson-18-Activiti...

Superstore
Regional
World Indicators

Discover

Open a Workbook

Training

- Getting Started
- Connecting to Data
- Visual Analytics
- Understanding Tableau
- More training videos...

Resources

- Get Tableau Prep
- Blog - What a year—explore the top vizies & authors of 2018 in our #VizInReview

Forums

VIZ OF THE WEEK

The Words of Christmas Movies →

Christmas with the Kranks

Update to 2018.3.2 Now





Guided Walk-Through: Tableau Connect Panel

The first time you start up a new Tableau Workbook, the Connect panel will be open by default.



You can open and close the Connect panel with this button:

The screenshot shows the Tableau Connect panel. At the top left, there's a 'Connect' button with a gear icon, which is highlighted with a red box and a red arrow from the previous slide. Below it, there's a search bar labeled 'Search'. The main area is divided into sections: 'To a File' (To a File, Microsoft Excel, Text file, JSON file, PDF file, Spatial file, Statistical file, More...), 'To a Server' (Tableau Server, Microsoft SQL Server, MySQL, Oracle, Amazon Redshift, More...), and 'Saved Data Sources' (Sample - Superstore, World Indicators). To the right, a large grid lists many data sources: Tableau Server, Amazon Athena, Amazon Aurora, Amazon EMR Hadoop Hive, Amazon Redshift, Anaplan, Apache Drill, Aster Database, Box, Cloudera Hadoop, Denodo, Dropbox, EXASOL, Firebird, Google Analytics, Google BigQuery, Google Cloud SQL, Google Sheets, Hortonworks Hadoop Hive, Intuit QuickBooks Online, Intuit QuickBooks Online (9.3-2018.1), Kognitio, MapR Hadoop Hive, Marketo, MemSQL, Microsoft SQL Server, MongoDB BI Connector, MySQL, OData, OneDrive, Oracle, Oracle Eloqua, Pivotal Greenplum Database, PostgreSQL, Presto, Salesforce, SAP HANA, ServiceNow ITSM, SharePoint Lists, Snowflake, Spark SQL, Teradata, Vertica, Web Data Connector, and Other Databases (ODBC).

Visualizing Data with Tableau

Let's Load Data and Get Started!





Guided Walk-Through:

Connecting to Data | Step by Step

Follow these steps to connect to the Superstore data set:

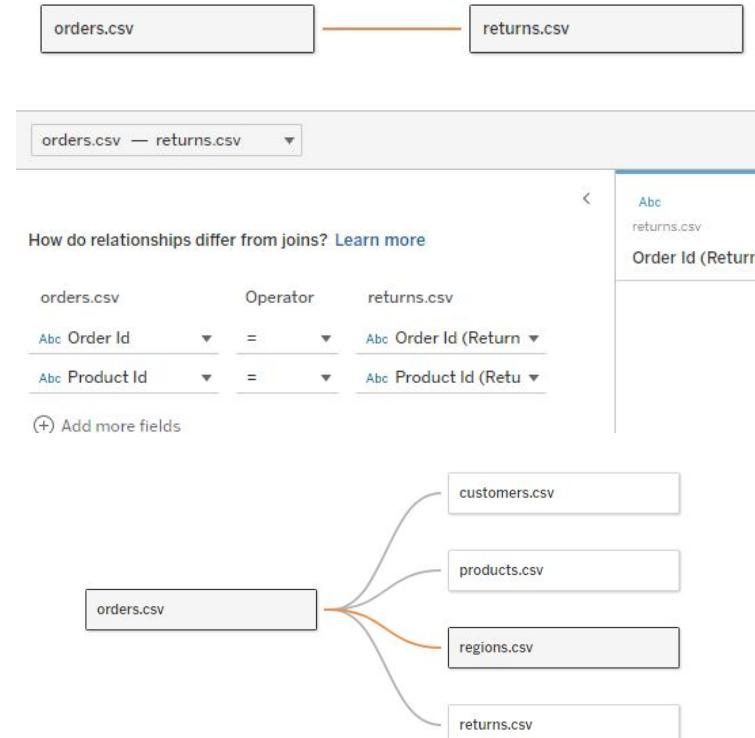
1. Open the ***Tableau_Lesson_1_Activities.twbx*** workbook.
2. Click on the Tableau Icon at the top left to bring you back to the start experience.
3. Navigate to your unzipped Superstore files from ***Superstore.zip***.
4. Choose the ***orders.csv*** file.
5. Tableau will load the file into the data pane, and the related data files will show up to the left.

The screenshot shows the Tableau Start Experience. On the left, there's a 'Connections' pane with a list of connections: 'orders' (Text file) is selected and highlighted in blue. Below it is a 'Files' section containing several CSV files: 'customers.csv', 'orders.csv', 'products.csv', 'regions.csv', and 'returns.csv'. To the right of the connections pane, there's a large 'orders' icon. Above the connections pane, a 'Connect' button is visible, along with options for 'To a File', 'Excel', 'Text file' (which is highlighted with a red box), 'Access', 'Statistical file', and 'More...'. At the bottom right of the interface, there's a small red ribbon icon.



Guided Walk-Through: Connect the Data

1. On the **Data Source** tab, notice **orders.csv** is on the data pane.
2. Find **returns.csv** in the left-hand pane, drag to data pane. The “noodle” will connect **orders.csv** to **returns.csv**.
3. Edit **Relationship** to connect using the **order_id** and **product_id** to create the relationship between the two files.
4. Add the remaining csv files for customers, products, and regions. Connect with the appropriate keys.

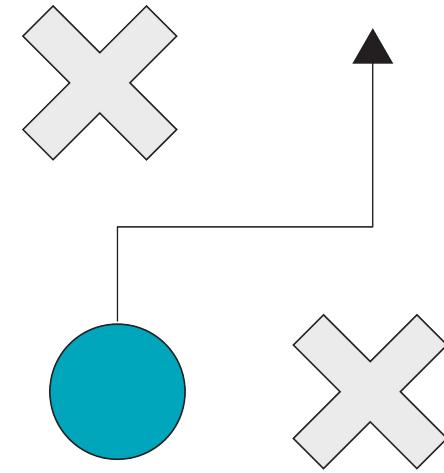


Connecting to Superstore Data

We'll continue using the Superstore data set and address stakeholder questions as we go.

Here are a few conditions to keep in mind:

- Every *row* of data is a purchase.
- Some purchases were returned.
- Every purchase has a category, sub-category, and pricing details.



Categorizing Data Types With CAST

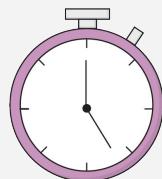
Tableau **automatically derives the data type** (e.g., string, numeric, date, etc.) from the data set upon import.

If a field's data type is *incorrectly* categorized, you can change the data type on the **Data Source** pane by clicking the data type icon above the field name.

Abc Orders Order ID	Abc Orders Order Date	Abc Orders Ship Date	Abc Orders Ship Mode	Abc Orders Segment	Abc Orders Customer Name
---------------------------	-----------------------------	----------------------------	----------------------------	--------------------------	--------------------------------

Extracts vs. Live

Extracts are snapshots of the connected data at the time of the extract that bring the live data into a Tableau columnar database.



You can schedule an Extract refresh in Tableau's server.

Live connections constantly run queries against the data connection and retrieve data in real or near real time.



LCs are best for relatively small or very fast data sources.

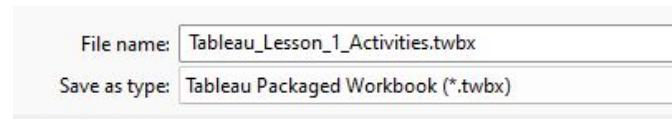
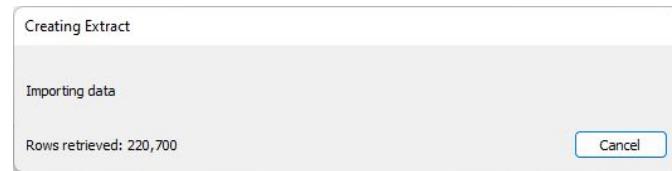


Guided Walk-Through:

Connecting to Data | Superstore Extract

Change the data connection from Live to Extract to improve workbook performance.

1. Click on the “Extract” button under “Connection.”
2. Select the **Guided Practice_01** sheet.
3. If prompted, save the extract file in your files.
 - Tableau will take up to one minute to build the Extract.
 - Saving Tableau as a “Packaged Workbook” will include data from the extract.



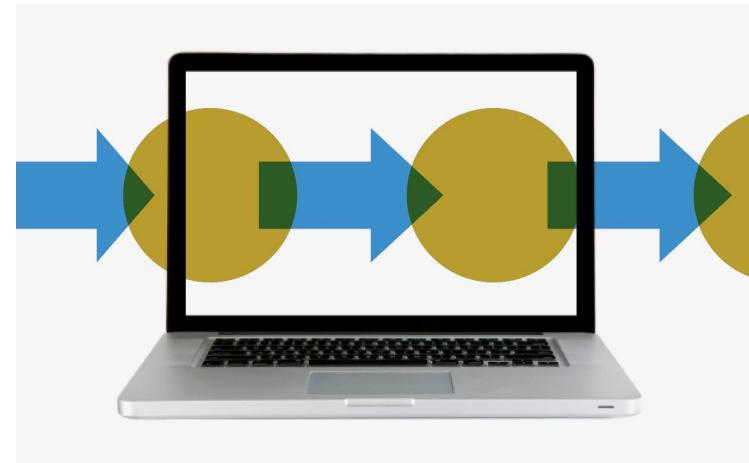
Transitioning From SQL to Tableau

From SQL to Tableau

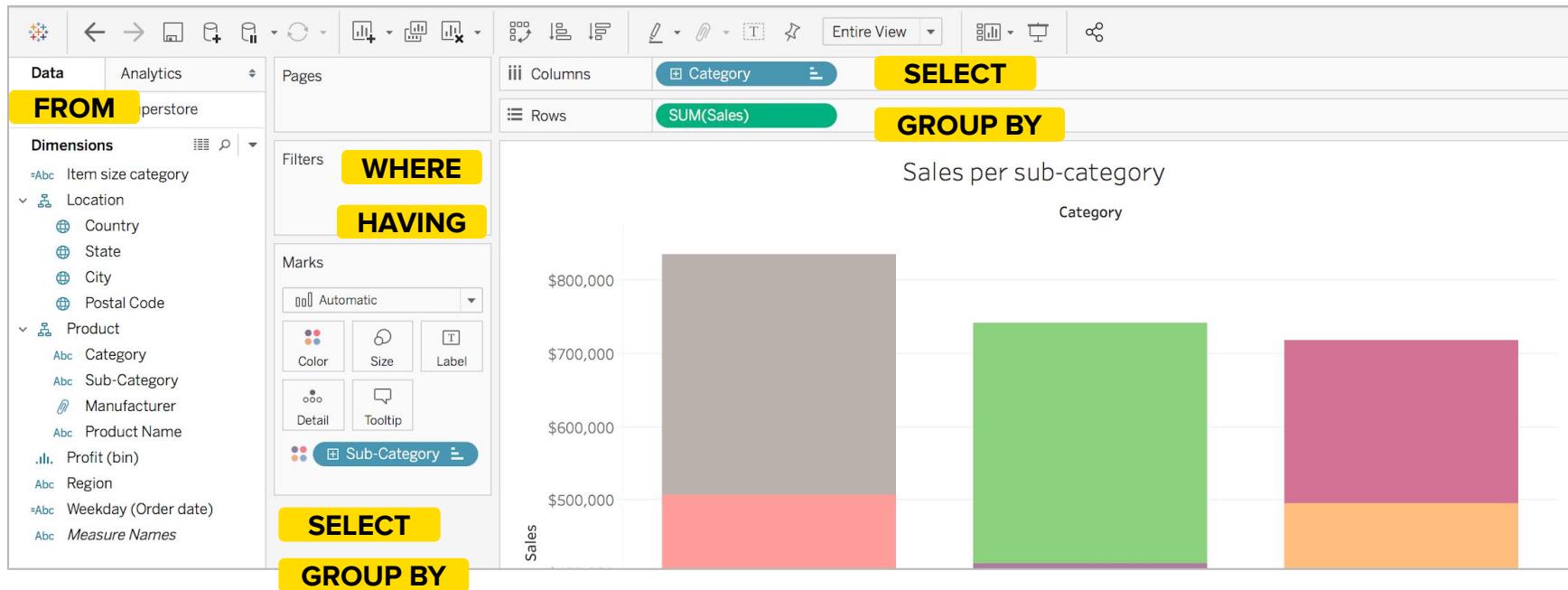
Tableau is a visual way to generate SQL statements (“VizQL”).

Let’s consider and discuss these statements as we explore how SQL queries are constructed in Tableau:

- *How do parts of the Tableau interface relate to SQL commands?*
- *How are SQL and Tableau similar and different?*



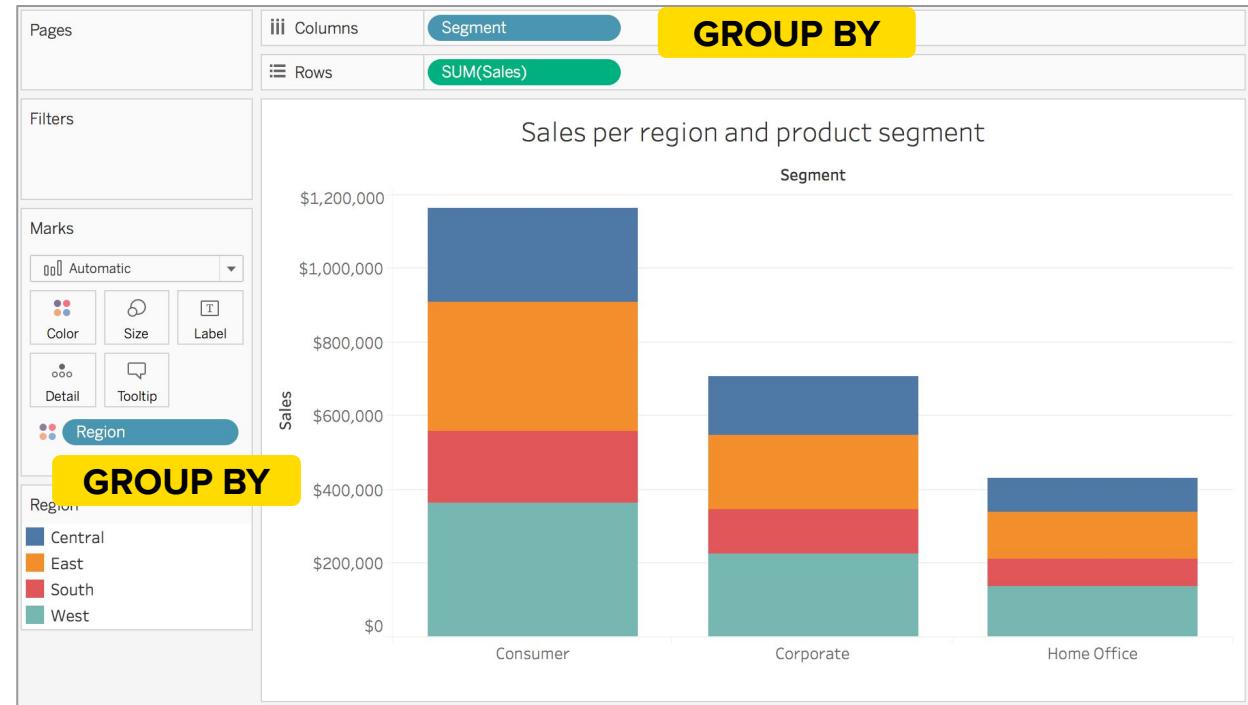
VizQL | SQL Analogs



Dimensions

Fields that contain discrete qualitative data. They are added into GROUP BY and are *not* aggregated.

Examples: dates, customer names, and segments.

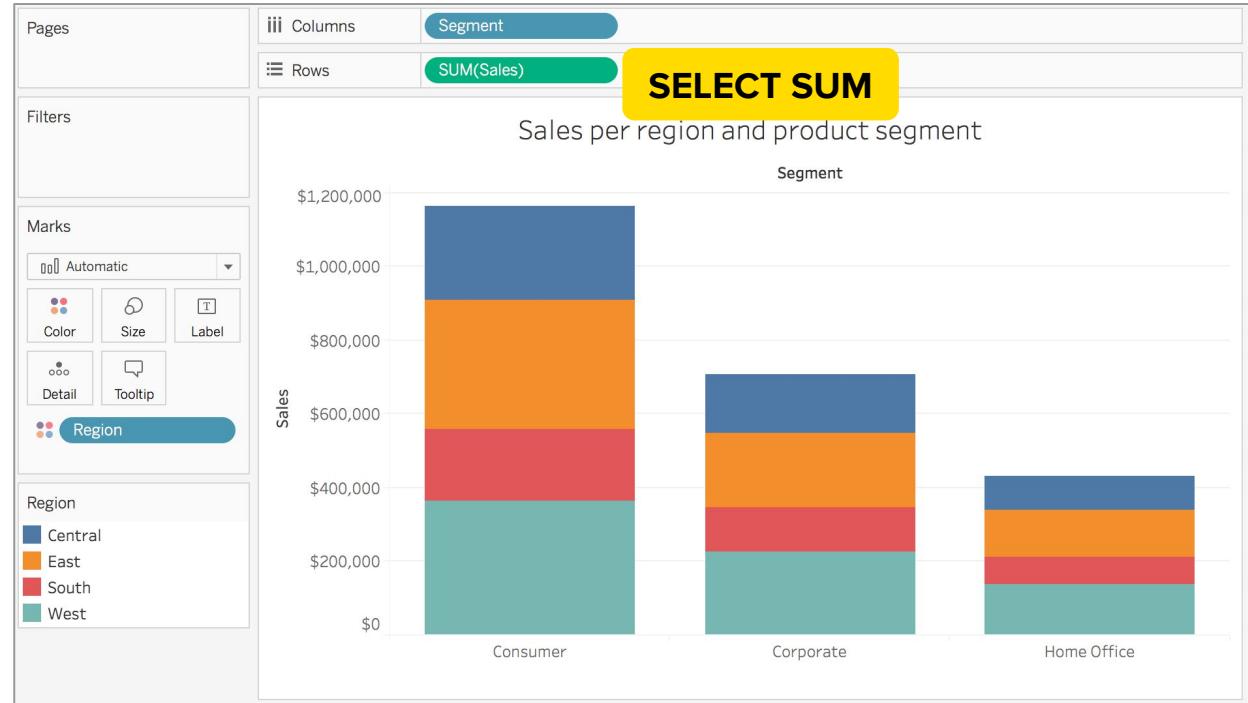


Measures

Fields that contain continuous numerical data that are aggregated via SELECT.

Measures aggregate on the pill, like SUM() or AVG().

Examples: SUM(Sales),
AVG(Profit),
COUNT(Employees).

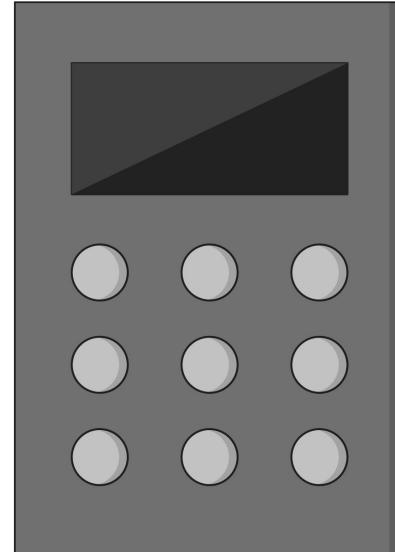


Aggregating Measures

When connecting to a data source, each measure gets a default aggregation, which you can view or change.

The aggregation for measures include:

- SUM
- AVG
- Median
- Count/Count (Distinct)
- Min/Max
- Percentile
- Std. Dev/Std. Dev (Pop.)
- Variance/Variance (Pop.)

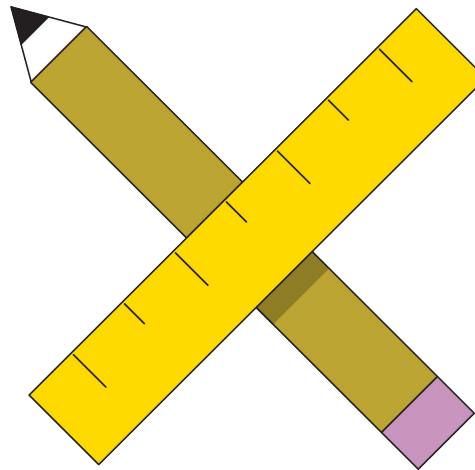


Aggregating Dimensions

While this is less common, you can also aggregate dimensions (such as dates) in the view as:

- Count/Count (Distinct)
- Min/Max

When aggregating a dimension, you create a temporary *measure* column.

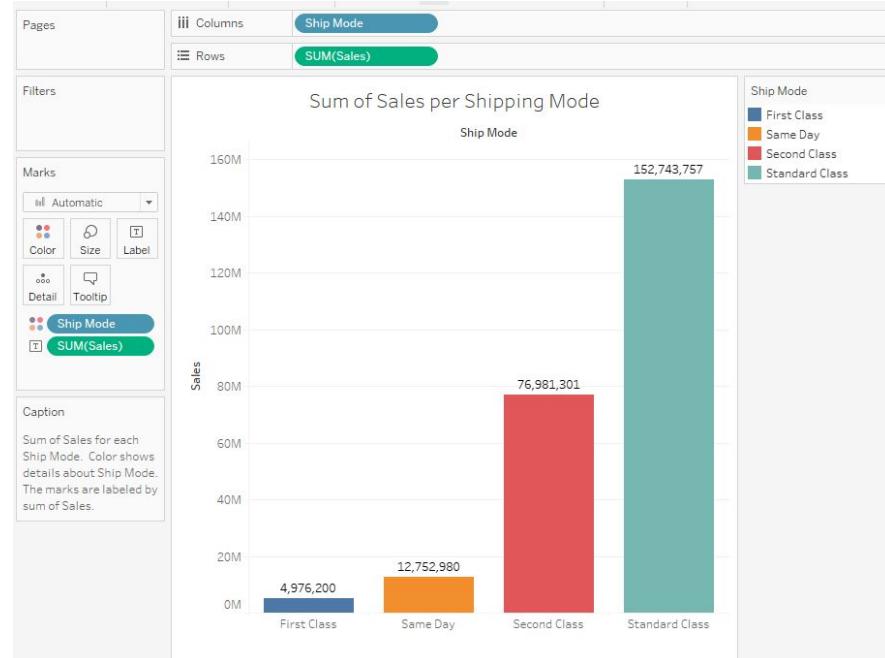




Guided Walk-Through: Visualization for Columns

On the worksheet **Guided Practice_01**:

1. Drag **Ship Mode** to **Columns** (creates x axis).
2. Drag **Sales** to **Rows** (creates y axis).
3. Drag **Ship Mode** from the Data pane to **Color**.
4. Drag **Sales** to **Label**.
5. Double-click on the Title, change to **Sum of Sales per Shipping Mode**.



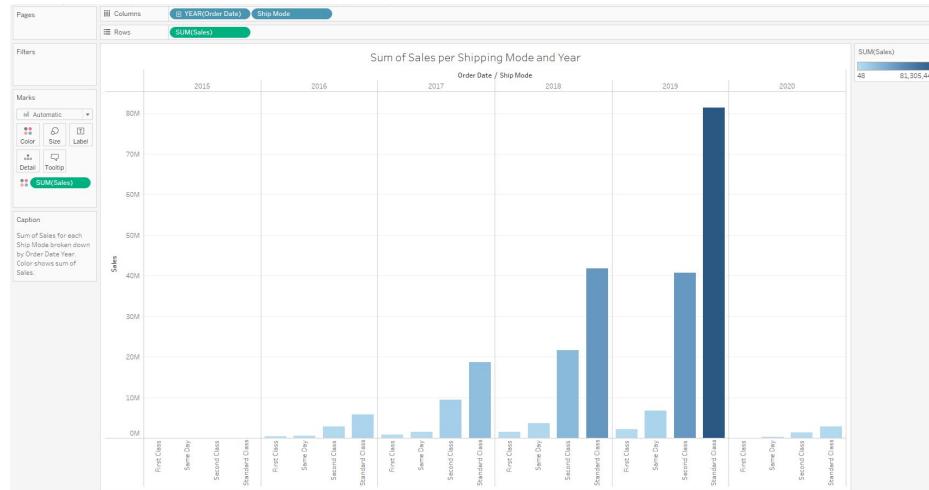
Save your work! Tableau Desktop does not automatically save!



Guided Walk-Through: Visualization for Columns

Select **Guided Practice_02**:

1. Drag **Order Date** (defaults to Year) and **Ship Mode** to **Columns** (creates a secondary x axis).
2. Drag **Sales** to **Rows** (creates y axis).
3. Drag **Sales** from the Data pane to **Color**.
4. Double-click on the Title, change to **Sum of Sales per Shipping Mode and Year**.



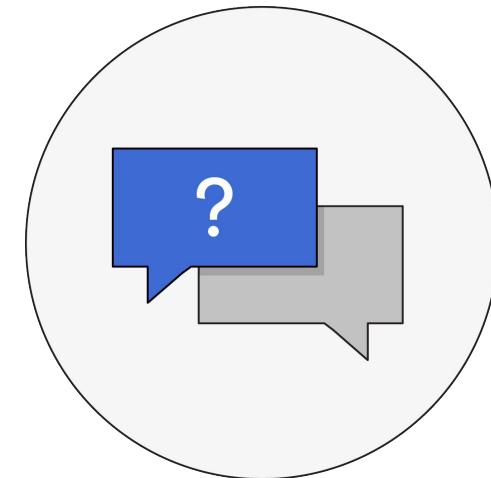


Interpreting and Refining Your Bar Chart

Let's review and interpret our chart from **Guided Practice_02** to add visual appeal for your audience.

Consider the following:

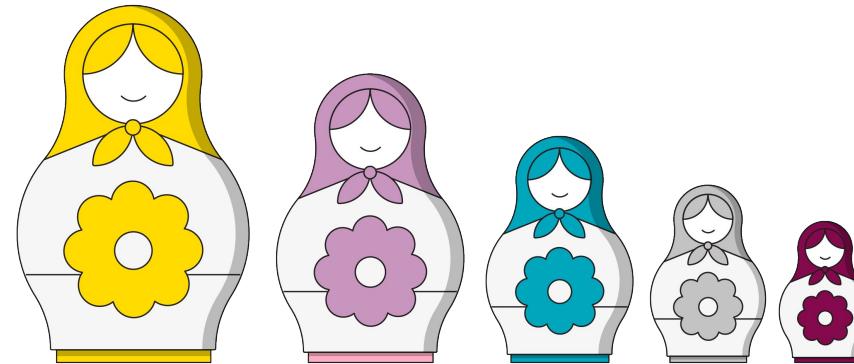
- What does **Sales** represent?
- How can we use **color**?
- How can we use **labels**?
- Can we add any **sorting**?



Formatting Tips | Sorting

Like SQL, we can add **ORDER BY** to our results to more clearly show a conclusion.

Example: We can sort the categories by their total number of sales from largest to smallest.



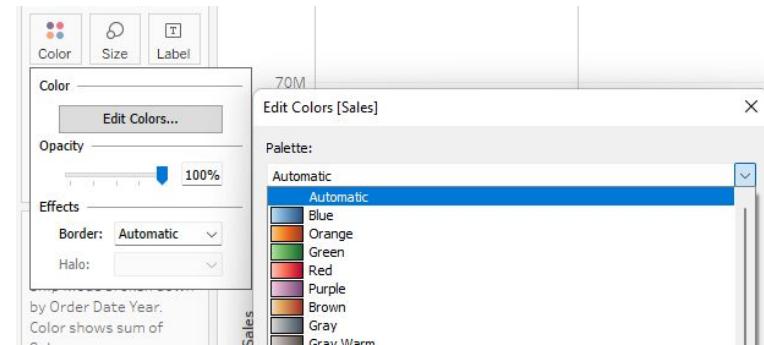
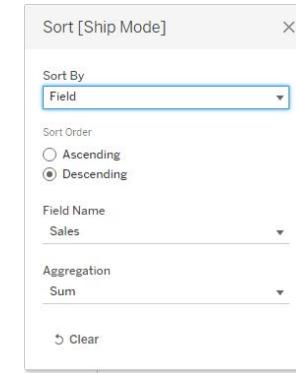


Guided Walk-Through:

Formatting Tips | Sorting Categories

Continue on **Guided Practice_02** to sort each category by the total number of sales.

1. Click the dropdown on **Ship Mode** in **Columns**.
2. Click **Sort**, choose **Sort By** dropdown and choose **Field**.
3. Change **Sort Order** to **Descending** using **Field Name** is **Sales** and **Aggregation** is **SUM**.
4. Click on **Color** > **Edit Colors** and choose a new color palette to see a change.



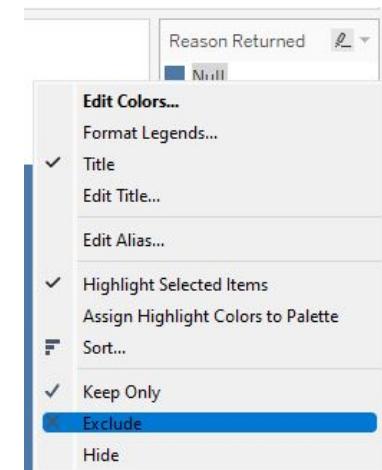


Guided Walk-Through:

Building a Stacked Bar Chart

Select **Guided Practice_03** to build a stacked bar chart that shows the Number of Sales (y axis) within each Ship Mode (x axis) by each reason returned.

1. Drag **Ship Mode** to **Columns** and **orders.csv(Count)** to **Rows**.
2. Drag **Reason Returned** to the **Color**.
3. Reason Returned of **Null** are orders with no returns.
Right-click on the **Null** in the legend and select **Exclude** to remove the nulls.
4. Drag **orders.csv(Count)** to **Label**.
5. Double-click to change the title to **Number of Returned Orders by Ship Mode**.





Group Exercise:

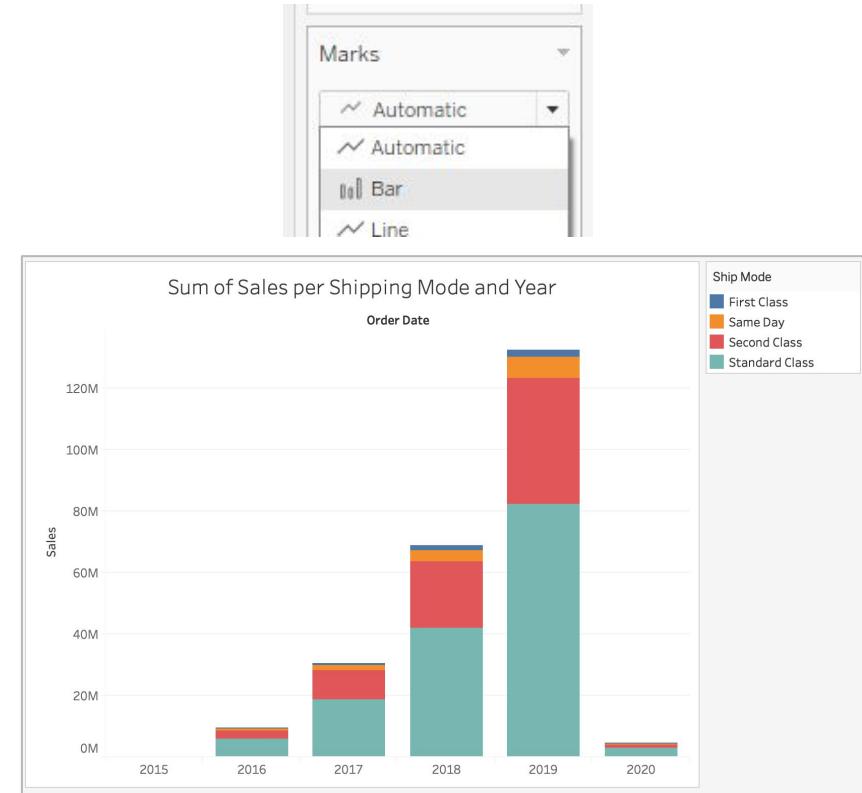
Visualization Practice

5 minutes



Use color to create a stacked bar chart in the **Group_Exercise_01** tab:

1. Add **Order Date** (as Year) to **Columns** and **Sales** to **Rows**.
2. Add **Ship Mode** to the **Color** mark.
3. Change the mark type to **Bar** on the Marks Card.
4. Change the title to **Sum of Sales per Shipping Mode and Year**.



Advanced Analytics

Working With Dates



Tableau: Discrete versus Continuous

Objective: Explore different ways that dates and times can be used in Tableau. Use **Order Date** and **Sales** to generate each visualization.

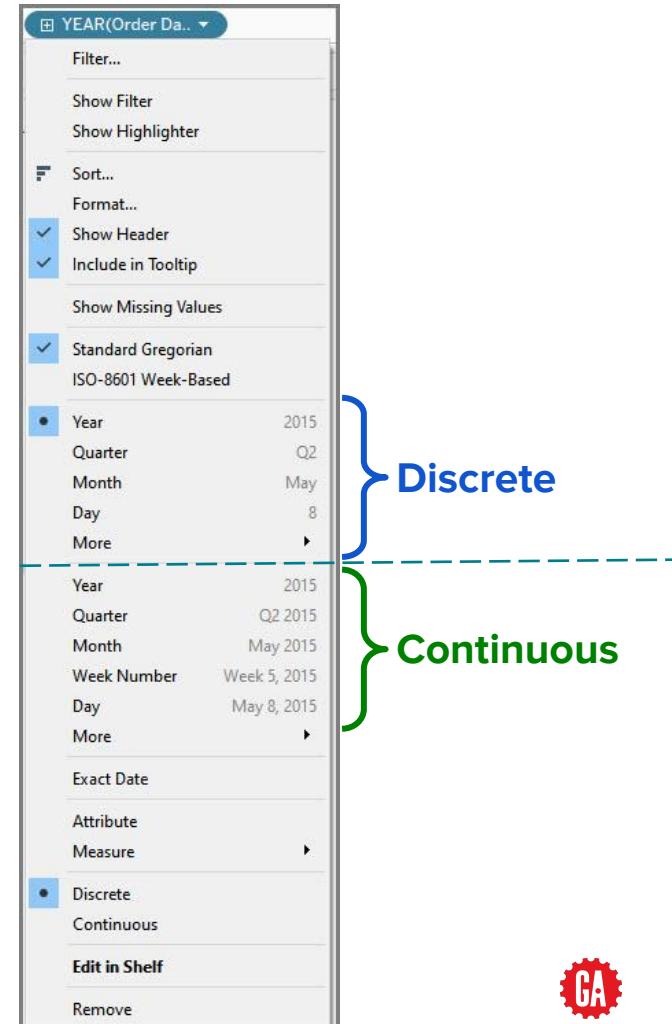
How it Works:

By default, Tableau begins with Year Discrete.

The top half of selections are **Discrete**.

The lower half of selections are **Continuous**.

Use examples on the right to confirm your selection and see what happens to the axis.



SUM(Profit)

YEAR(Order Date)

“

"forming an unbroken whole, without
interruption"

Continuous

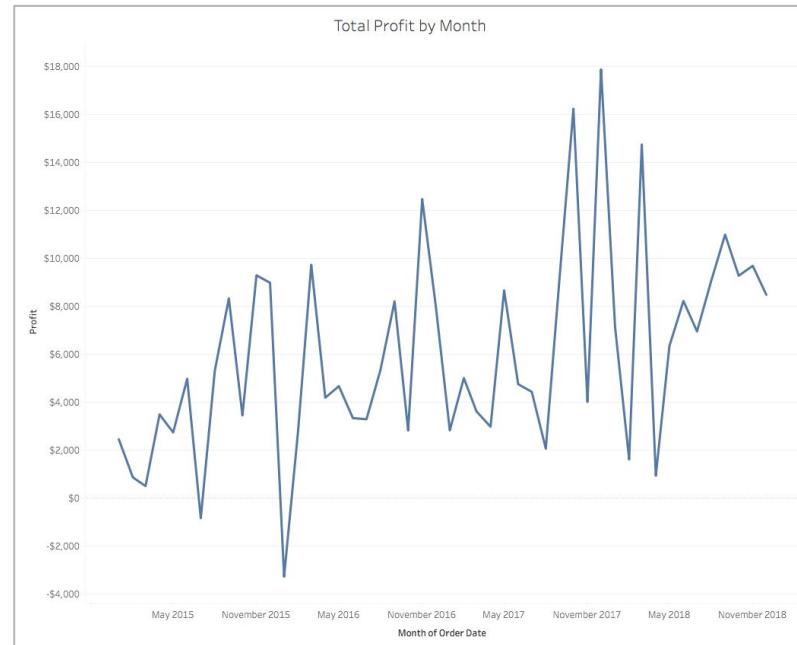


Continuous Dates

Create axes. Think "Time Series."

If month is used as a continuous date,
then **an axis that separates each
month by year will be visualized.**

You cannot sort this axis.

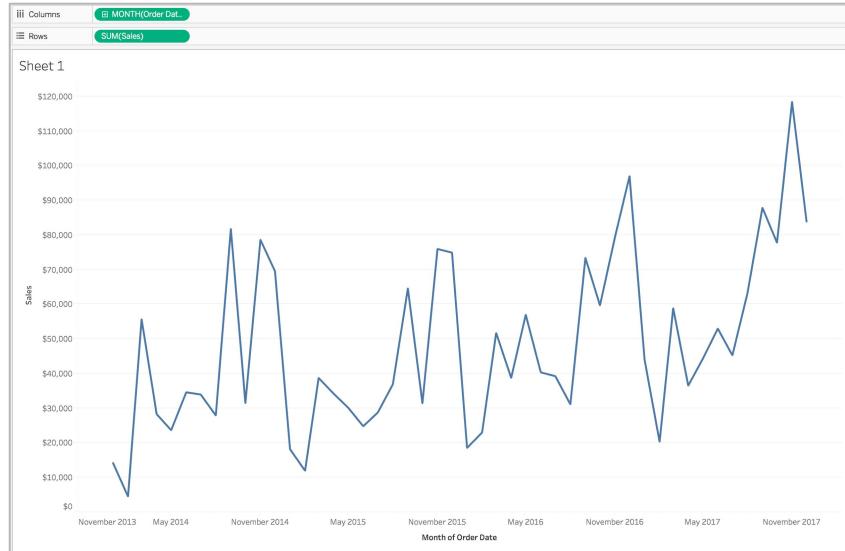


Continuous Date Aggregation

This graph is slightly different, showing
sales by month and year.

It's a **standard time series** type of chart.

These dates *cannot* be sorted.



Continuous Date Aggregation

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

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

A screenshot of a date aggregation dropdown menu. The top bar shows 'MONTH(Order Dat.. ▾)'. Below it are several options: 'Filter...', 'Show Filter', 'Format...', and two checked checkboxes: 'Show Header' and 'Include in Tooltip'. A section titled 'Show Missing Values' follows. The main list contains five items: 'Year' (2015), 'Quarter' (Q2), 'Month' (May) which is checked and highlighted in blue, 'Day' (8), and 'More' (with a right-pointing arrow). Below this is another list starting with 'Year' (2015), 'Quarter' (Q2 2015), and 'Month' (May 2015) which is also highlighted in blue. This second list also includes 'Week Number' (Week 5, 2015), 'Day' (May 8, 2015), and 'More'. Further down are sections for 'Exact Date', 'Attribute Measure', and 'Discrete', with the last item being 'Continuous' which is also checked.

Product Name

YEAR(Order Date)

“

"individually separate and distinct."

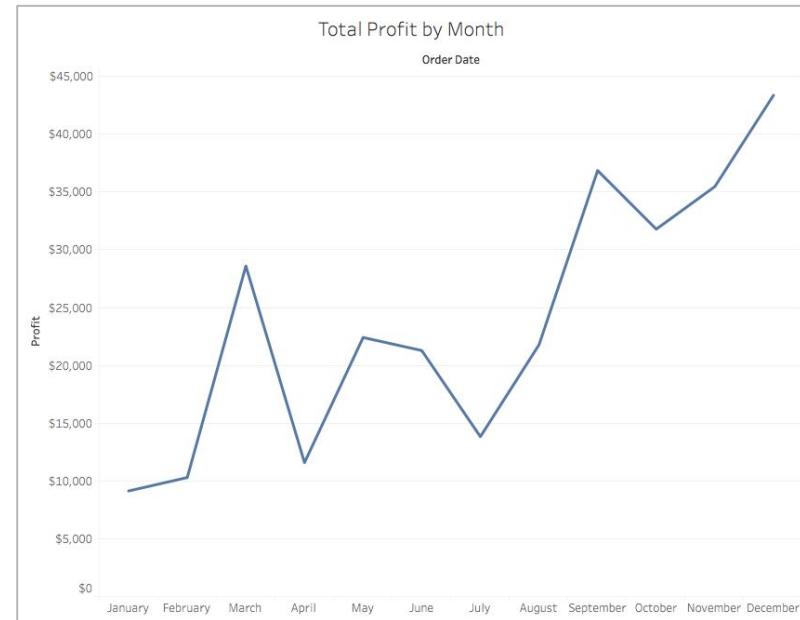
Discrete



Discrete Dates

Create bins.

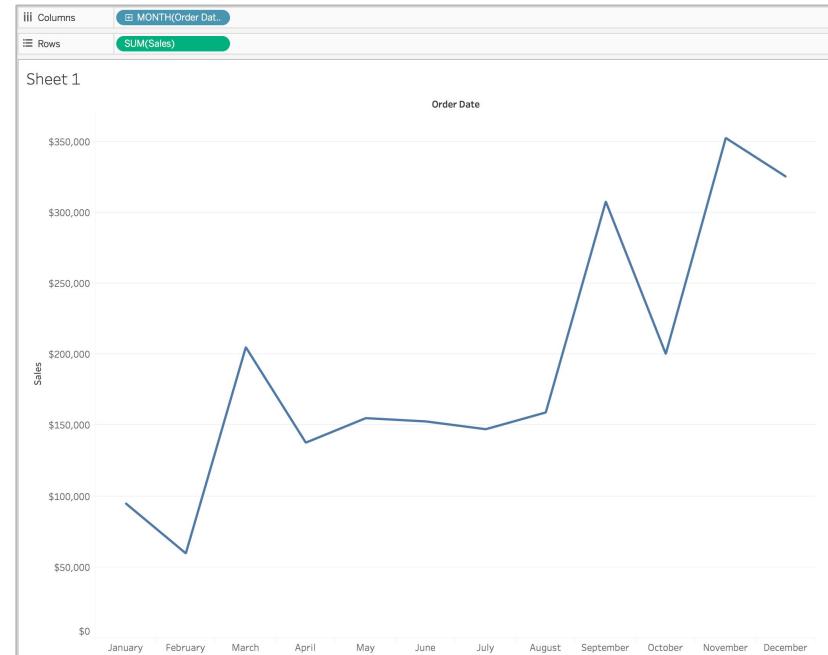
If you visualize each month as a discrete date, **it will not distinguish between years**. However, you can sort the time axis.



Discrete Date Aggregation

This graph shows **the sum of sales by month**, regardless of the year in which the sale occurred.

The dates are stored in bins and may be sorted.



Discrete Date Aggregation

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

MONTH(Order ..	▼
Filter...	
Show Filter	
Format...	
✓ Show Header	
✓ Include in Tooltip	
Show Missing Values	
Year	2015
Quarter	Q2
Month	May
Day	8
More	►
Year	2015
Quarter	Q2 2015
✓ Month	May 2015
Week Number	Week 5, 2015
Day	May 8, 2015
More	►
Exact Date	
Attribute	
Measure	►
Discrete	
✓ Continuous	

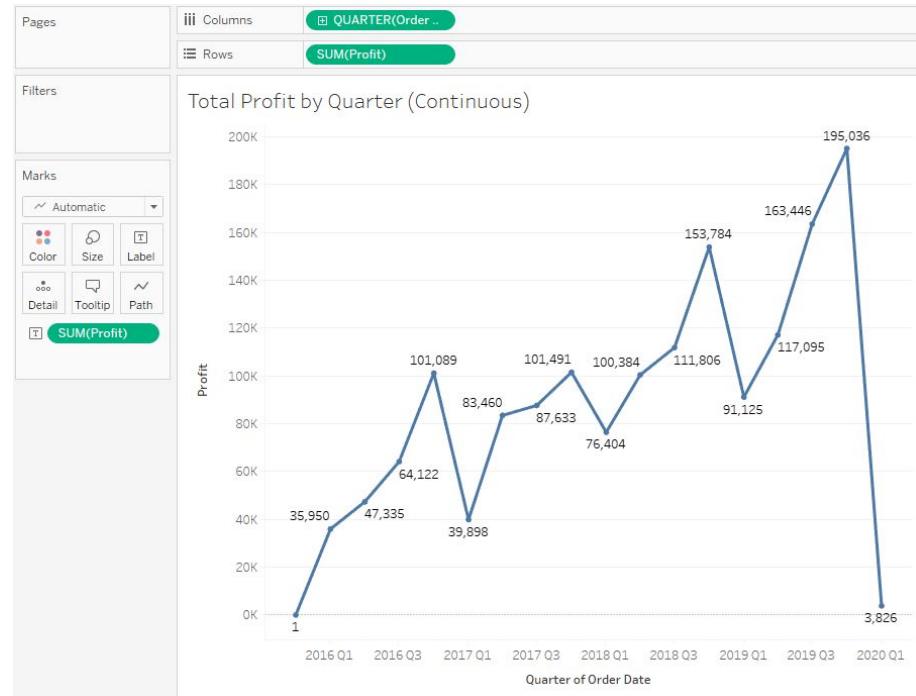


Guided Walk-Through:

Creating a Line Chart

Select **Guided Practice_04**. Superstore wants to have a bird's-eye view over the Total Profit by Year. Create a line chart that shows the **Total Profit by Quarter (Continuous)**.

1. Drag **Order Date** to **Columns**.
2. Click the dropdown on **YEAR(Order Date)** to change **Order Date to Continuous Quarter**.
3. Drag **Profit** to **Rows and Label**.
4. Change title to **Total Profit by Quarter (Continuous)**.

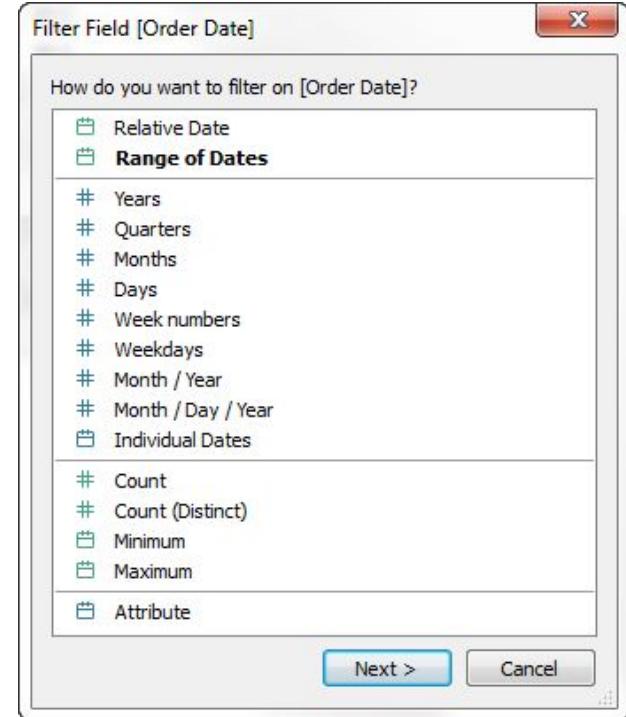


Filtering Dates

When dragging a **Date** field from the **Data** pane to the **Filters** shelf in Tableau Desktop, you'll see a dialog box like this one.

Then, decide whether you want to filter:

- On a **relative date**.
- Between a **range of dates**.
- Or select **discrete or individual dates** to filter from the view.





Guided Walk-Through:

Creating a Line Chart

Continue on **Guided Practice_04**, update the **Total Profit by Quarter (Continuous)** to add a date filter.

1. Drag **Order Date** to **Filters**.
2. Choose **Range of Dates**, click **Next**.
3. Adjust range **1 Jan 2016 - 31 Dec 2019** to remove partial years. Click **OK**.

Filter Field [Order Date] X

How do you want to filter on [Order Date]?

- Relative Date
- Range of Dates
- ...

Filter [Order Date]

Relative dates Range of dates Starting date Ending date Special

Range of dates

1/1/2016 12/31/2019

12/18/2015 1/30/2020

Show: Only Relevant Values Include Null Values

Reset OK Cancel Apply

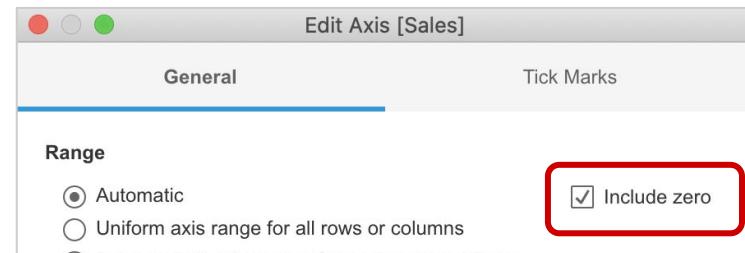


Guided Walk-Through:

Formatting Tips | Editing Axis

Line Charts may not always have to start at zero. Allowing the data to find its lowest point can show the peaks and valleys of the data more clearly. Continue on **Guided Practice_04**, to adjust the y-axis.

1. Right click on the y-axis and click **Edit Axis**.
2. Uncheck **Include Zero**



Bar Charts must ALWAYS start at zero; this is unique for our Line Chart example!

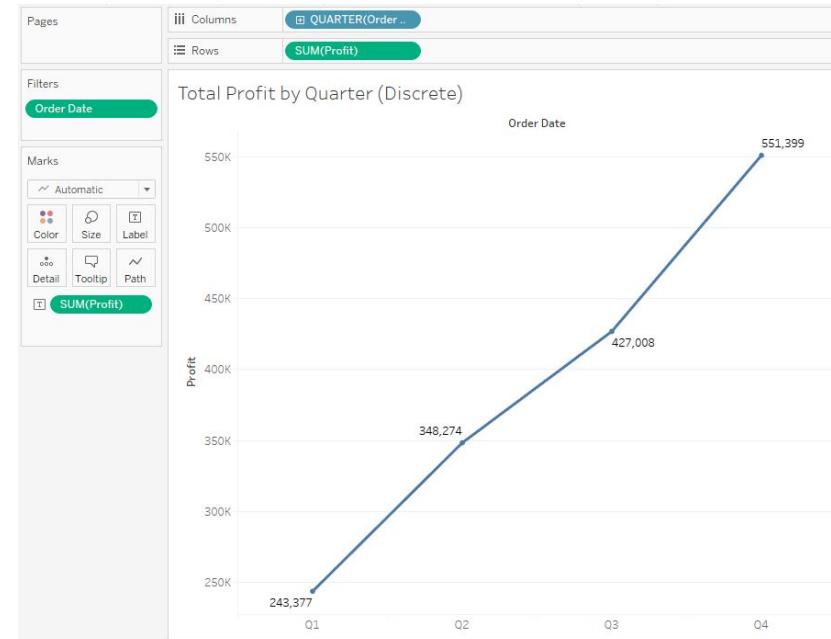


Guided Walk-Through:

Creating a Line Chart

Superstore wants to see seasonality by month. Duplicate the **Total Profit by Quarter (Continuous)** chart to create a line chart that shows the **Total Profit by Quarter (Discrete)**.

1. Right click on the **Guided Practice_04** tab.
2. Select **Duplicate**.
3. Rename the new sheet to **Guided Practice_05**.
4. Change Order Date in Columns to a **Discrete Quarter** and change the chart title to match.





Group Exercise:

Create a Multi-Line Chart

5 minutes



It's time to up your game. Superstore asked for a visual that offers insights into how each ship mode is doing by year. Create a multi-line chart that shows the **Total Profit by Year and Ship Mode**. Be prepared to share your visualization with the class!

1. Click into **Group Exercise_02**.
2. Drag **Order Date** to **Columns** and **Profit** to **Rows**.
3. Drag **Ship Mode** to **Color**.
4. **Exclude the zero** in the y-axis.
5. Add **Labels** for **Profit**.
6. Filter to include only **2016 - 2019**.
 - Hint: Choose Years instead of Range of Dates in the filter.
7. Update title to **Total Profit by Year and Ship Mode**.

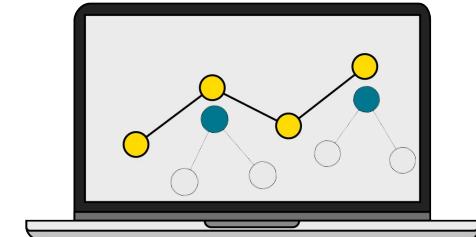


Tableau: Independent Practice



Solo Exercise:

Tableau Independent Practice

30 min



Let's bring it all together and apply the various concepts we learned today!

In the sheet tabs at the bottom of the Tableau workbook, you will see **SoloExercise_01** and **SoloExercise_02**. In those tabs, work to accomplish the following:

1. **SoloExercise_01:** Create a visualization that shows sales for each customer segment in each region.
2. **SoloExercise_02:** Create a visualization that shows sales by product category over time by quarter.



If at any point you feel stuck, go ahead and post in Slack and a classmate or instructor will be happy to assist!

Optional Guided Walkthroughs

Creating Groups and Sets



Guided Walk-Through: Creating a Set

Select **Guided Practice_06** to create a set for weekend/weekdays

1. Drag **Order Date** to **Rows**. Change the Year to **Discrete WEEKDAY**.
2. Holding your **Ctrl** or **Cmd** key and click on **Saturday** and **Sunday**. Click the venn diagram icon, select **Create Set**.
3. Name the set, **Weekend/Weekday**. Click **OK**.
4. From the Data Pane, drag the set to Rows, replacing the Order Date.
5. Right click on **In**, edit the Alias to **Weekend**. Change the Alias for **Out** to **Weekday**.

Weekday of..	
Sunday	Abc
Monday	Abc
Tuesday	Abc
Wednesday	Abc
Thursday	Abc
Friday	Abc
Saturday	Abc

✓ Keep Only ⚙ Exclude 2 items selected Create Set...

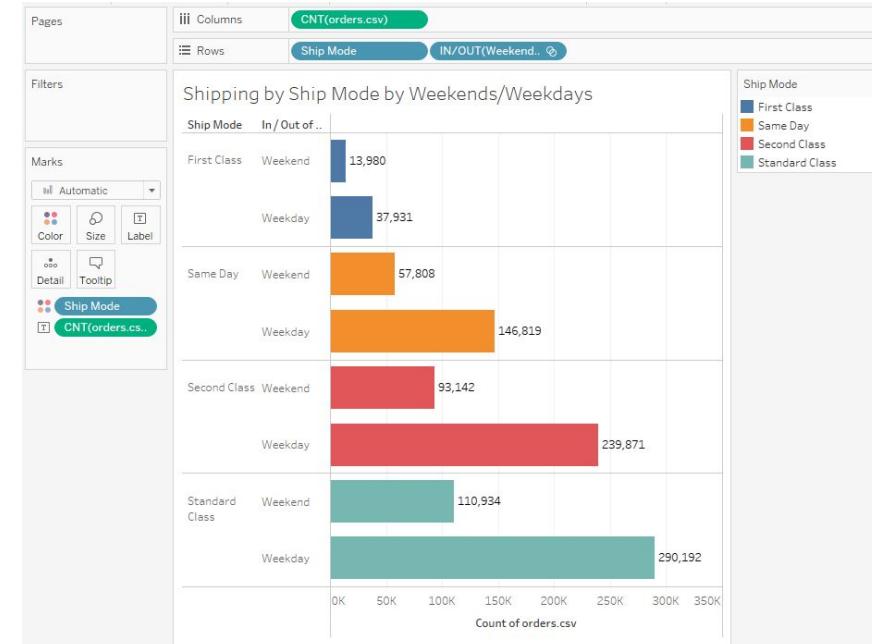
Saturday



Guided Walk-Through: Using a Set

Continue on **Guided Practice_06** to use the set for weekend/weekdays to see the number of orders shipped by ship mode.

6. Drag **Ship Mode** to **Rows** to the left of the **In/Out(Weekend..)** and to **Color**.
7. Drag **orders.csv(Count)** to **Columns** and **Label**.
8. Edit the **Title** to **Shipping by Ship Mode by Weekends/Weekdays**

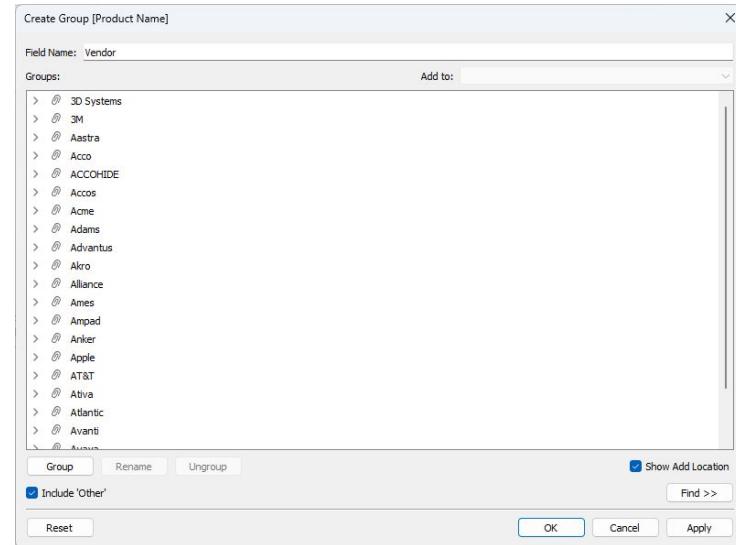




Guided Walk-Through: Creating Groups

Select **Guided Practice_07** to create groups for product vendors.

1. Click the drop down on **Product Name**. Select **Create > Group**.
2. Scroll down to the first product name for **3D Systems**. Click on all of the 3D Systems products. Click **Group**. Name it **Vendor**.
3. A line with a paperclip is added above the products. Edit to change the name to **3D Systems**.
4. Repeat to create a group for **3M, Aastral, Acco**, etc. Click the checkbox **Include 'Other'** to group the remaining products under 'Other'.

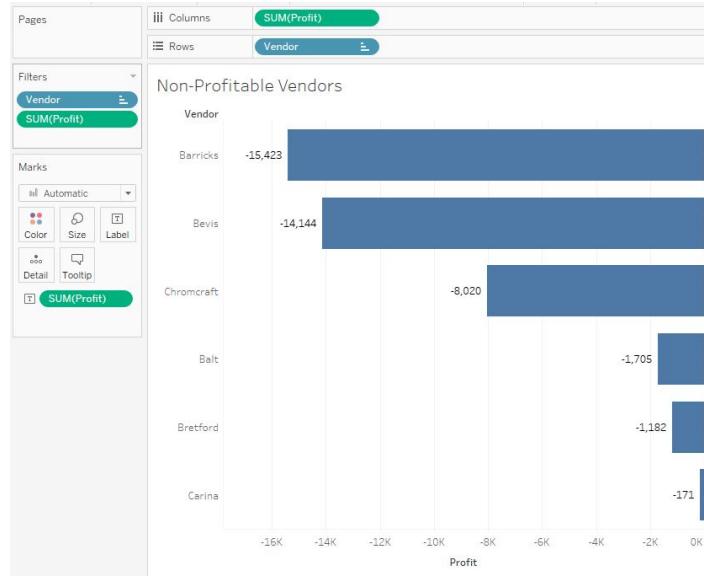




Guided Walk-Through: Creating Groups

Continue on **Guided Practice_07** to use the product vendor groups.

1. Drag the **Vendor** group to **Rows** and **Filters**.
2. Exclude **Other** in the filter.
3. Drag **Profit** to **Columns, Label** and **Filters**.
4. Filter **Profit** to **At Most 0**. This will filter any non-profitable vendors.
5. Sort **Vendor Ascending**.
6. Edit **Title** to be **Non-Profitable Vendors**
 - o Depending on how the vendors are grouped, the results may vary.





Solo Exercise:

Optional Homework

Use the Homework tab in your workbook to accomplish the following. You may need to add more data via JOINs to achieve the visualization!

Homework_03: Create a visualization that shows profit by product category to see seasonality by month.

Lunch Time!

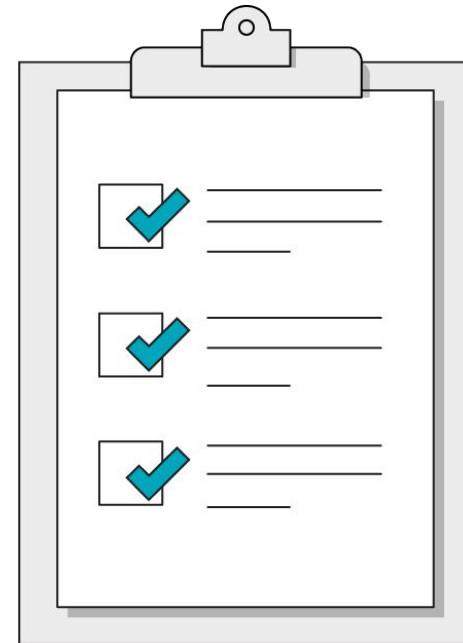


Data Analytics

Data Manipulation in Tableau

Our Learning Goals

- Connect to the PostgreSQL server.
- Create calculated fields to analyze data.
- Apply filters to single or multiple sheets.
- Create data visualizations and simple dashboards in Tableau





Discussion:

Making Data Accessible

Superstore needs a dashboard that shows **sales performance across regions and by salespersons**.

?

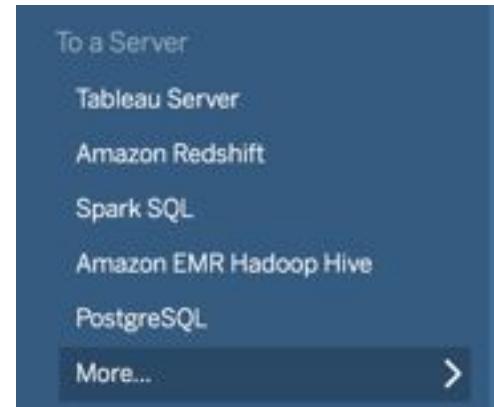
As a data analyst, what can you do to share data-driven insights AND empower the teams at Superstore to access, review, and understand their own data?

Data Manipulation in Tableau

Connecting to PostgreSQL

Connecting to PostgreSQL

1. Open up Tableau_Lesson_2_Activities.twbx in Tableau.
2. In the **Data Source** panel, choose **PostgreSQL** from the list of options under **Connect to a Server**. (You may have to click **More**.)
3. Enter the login details (see next slide).



Enter the login details for our GA server:

Server: analyticsga-psql.generalassembly.ly

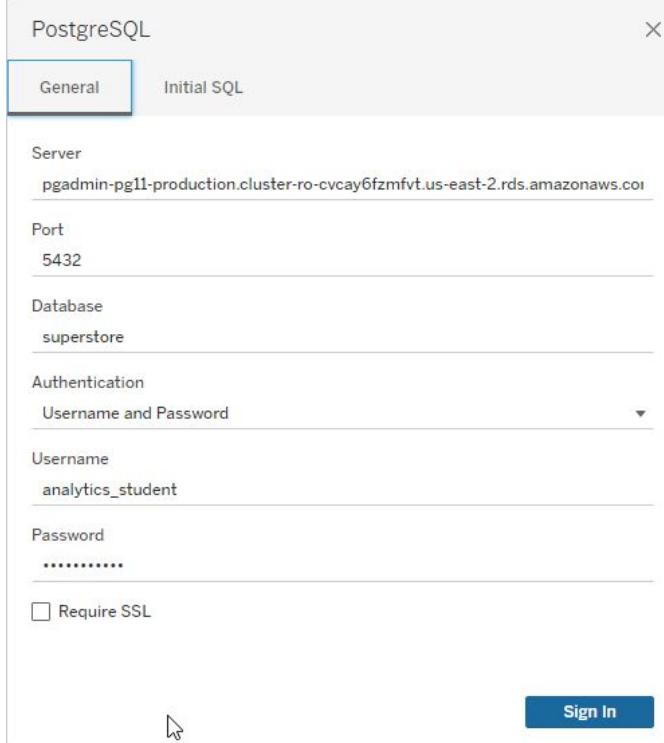
Port: 5432

Database: superstore

Authentication: Username and password

Username: analytics_student

Password: analyticsga



The screenshot shows a configuration dialog for a PostgreSQL connection. The title bar says "PostgreSQL". There are two tabs: "General" (which is selected) and "Initial SQL".

General Tab Fields:

- Server: pgadmin-pg11-production.cluster-ro-cvcay6fzmfvt.us-east-2.rds.amazonaws.com
- Port: 5432
- Database: superstore
- Authentication: Username and Password
- Username: analytics_student
- Password: (redacted)
- Require SSL

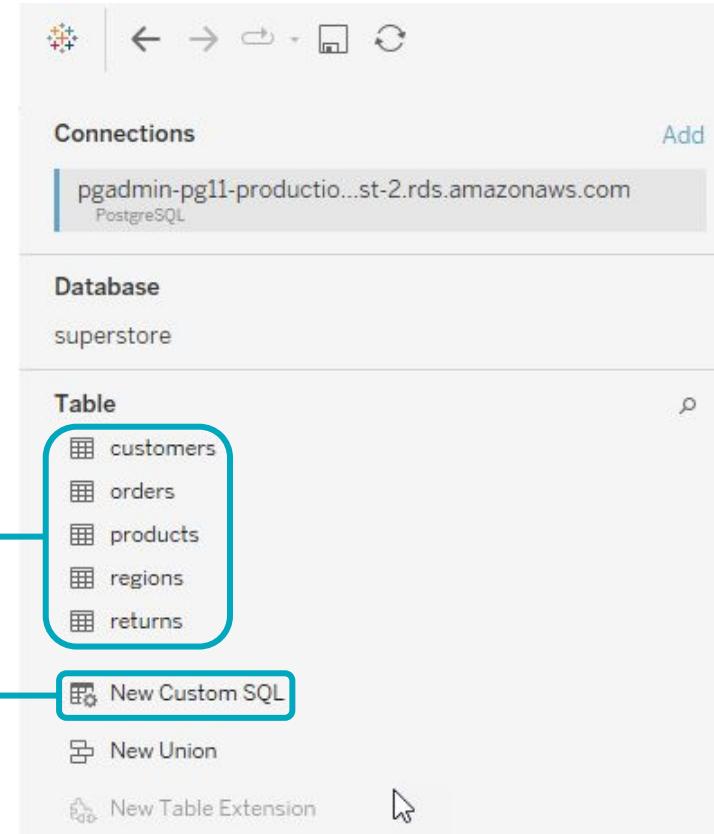
Sign In button is located at the bottom right.



Connecting to PostgreSQL (Cont.)

You should now see our tables in the list of tables on the left-hand side.

- We can choose a table to move over to the **Data** pane (click and drag)
- We can also write “**New Custom SQL**” to manipulate the data prior to moving it into the **Data** pane





Discussion:

Reviewing the Data Source

Now that we're connected to the Superstore database, let's review the data source and see how it aligns with our business question.

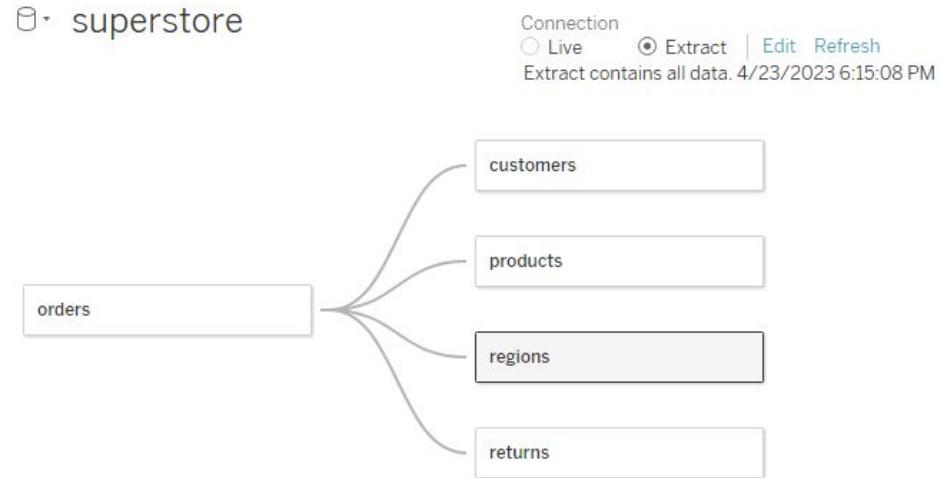
- 1. Which tables do we need to accomplish our goal of **seeing sales by region and salesperson?**
- 2. How should we connect those tables?





Guided Walk-Through: JOINs in Tableau

1. First, drag **orders** into the **Data** pane.
2. Then, drag **returns** into the **Data** pane and specify the **Relationship** on `order_id` and `product_id`.
3. Repeat with **regions**, **products** and **customers**.
4. Click on **Extract**.
5. Click on any tab to generate the extract. This may take a few minutes.
6. Go back to the Data Source page, click on the Data Source name, **orders+ (superstore)**, and change it to **superstore**.



What's in Your Data?



Now, work with your partner to identify the data points in your new table.



1. Given our business problem, what should we use in our analysis?
2. Are there data points that we can hide?
3. Are there data points we still need to create?

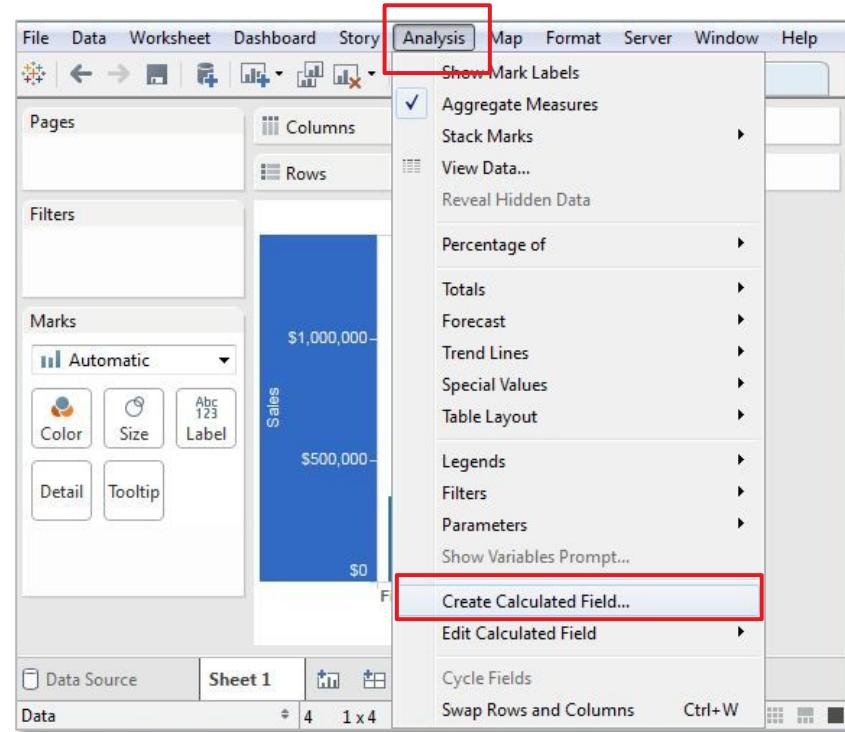
Be prepared to share your answers.

Calculated Fields

What Are Calculated Fields?

Calculated fields are **new fields you create from your original data source** — the values of which are determined by the calculation that you control.

You can create calculated fields in Tableau from the **Analysis** tab.



Use Cases for Calculated Fields

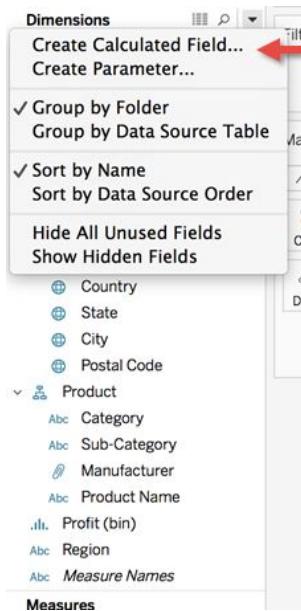
You can use calculated fields to do a lot of different things, including:

- Segmenting your data.
- Converting the data type of a field (e.g., a string to a date).
- Aggregating your data.
- Filtering the results.
- Calculating ratios.

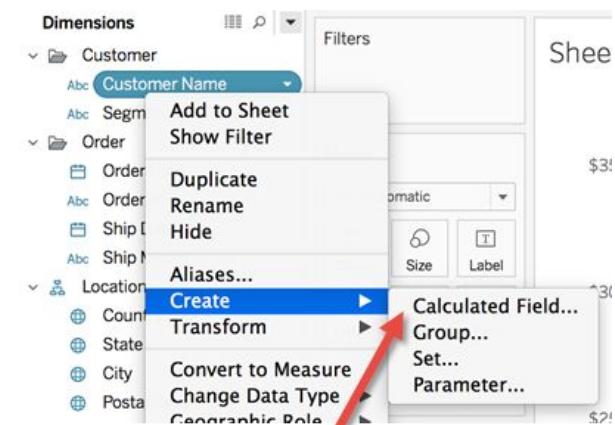


Do you have specific examples for any of the above? What are some other use cases? Let's discuss!

Finding Calculated Fields in Dimensions



You can also create calculations by clicking on the triangle next to **Dimensions** and selecting **Create Calculated Field...**



If you know the specific dimension or measure, you can right click on that pill and select **Create > Calculated Field...**

Data Manipulation and Calculation

Like SQL, Tableau uses different formulas to manipulate measures and dimensions. The three main types of calculations are:

- **Text manipulation:** Allows us to combine dimensions, create new dimensions, and organize our data.
- **Date manipulation:** Allows us to create new dimensions or measures in our data.
- **Logical manipulation:** We can create logical calculations using functions such as CASE, IF, CONTAINS, and ISNULL, which use Boolean logic (which returns a TRUE or FALSE) to produce different results depending on criteria being met.



Discussion:

Calculations in Tableau

Let's look at how Ship Mode affects Shipping Performance.

Drawing from what you know about both Excel and SQL functions, which of the following would you use to find shipping time?

- DATEDIFF
- LEFT
- FIND
- CONCAT
- CONTAINS
- IF

The screenshot shows the Tableau calculated field editor. A search bar at the top contains the text "date". Below it is a list of functions. The "DATEDIFF" function is highlighted with a gray background and has a tooltip displayed to its right. The tooltip contains the function signature "DATEDIFF(date_part, start_date, end_date, [start_of_week])" and a detailed description: "Returns the difference between two dates where start_date is subtracted from end_date. The difference is expressed in units of date_part. If start_of_week is omitted, the week start day is determined by the start day configured for the data source." At the bottom of the editor are two buttons: "Apply" and "OK".



Guided Walk-Through:

Performing Basic Calculations

To understand how shipping varies across Ship Modes, we'll start by looking at shipping times. First, we need a calculated field for Time to Ship.

1. Open the **Guided Practice_01** worksheet.
2. Create a new calculated field.
3. Name it “**Time to Ship**.”
4. Use the calculation:

DATEDIFF(“day”, [Order Date], [Ship Date]).

The screenshot shows a dialog box titled "Time to Ship". Inside, the formula `DATEDIFF("day", [Order Date], [Ship Date])` is displayed. Below the formula, a message says "The calculation is valid." At the bottom right are two buttons: "Apply" and "OK".

Tip: Make sure that **Order Date** and **Ship Date** are casted as dates in your **Orders** data set before building this calculation.



Performing Basic Calculations

Let's do some analysis on our shipping times. Create a line chart that shows the average shipping time over the past four years. Has any ship mode increased?

1. Continue on the **Guided Practice_01** worksheet.
2. Drag **Order Date** to columns and leave the aggregation as **YEAR**.
3. Drag **Order Date** to **Filters**, Choose **Range of Dates** for **1/1/2016 - 12/31/2019**.
4. Right click on Time to Ship and change the **Default Properties > Aggregation** to **Average**. Drag **Time to Ship** to rows.
5. Drag **Ship Mode** to **Color**.
6. Click **Label** on the Marks Card. Select **Line Ends** and **Show mark labels**.
7. Right click the y-axis, select **Edit Axis**. Uncheck **Include zero**.
 - a. **Note:** Line Charts do not have to start at zero. It can be easier to see trends if the y-axis is closer to the lowest data point. Bar charts MUST always start at zero!

Filter Multiple Sheets

Adding Filters to Multiple Sheets

You can add filters to worksheets that use...

A related primary data source.

On the **Filters** shelf, right click the field and select:

Apply to Worksheets > All Using Related Data Sources

Filters that use this option are *global* across the workbook.

Or

The current primary data source.

On the **Filters** shelf, right click the field and select:

Apply to Worksheets > All Using This Data Source

Filters that use this option are *global* across the workbook.





Applying Filters Globally

1. In **Guided Practice_01**, click the dropdown on **Order Date** in the **Filters** shelf.
2. Hover on **Apply to Worksheets** to see options.
 - a. Choose **All Using this Data Source**. This will apply the filter to all of our sheets and any new sheets we create automatically.
 - b. Choose **Selected Worksheets** and select only the Guided Practice worksheets. This option is visible after more than 1 visualization/tab has been created.
3. Choose **All Using this Data Source** for this workbook.
 - a. As you build new visualizations your **Order Date** filter has been added automatically.

Show Me

Show Me

You can use **Show Me** to create a view based on the fields already used in the view as well as any fields you've selected in the **Data pane**.

1. Select fields in the **Data pane** that you want to analyze.
2. Click **Show Me** on the toolbar and then select the type of view you want to create.
3. View the result!





Histogram

Histograms show the **spread or distribution** of values for a single variable.

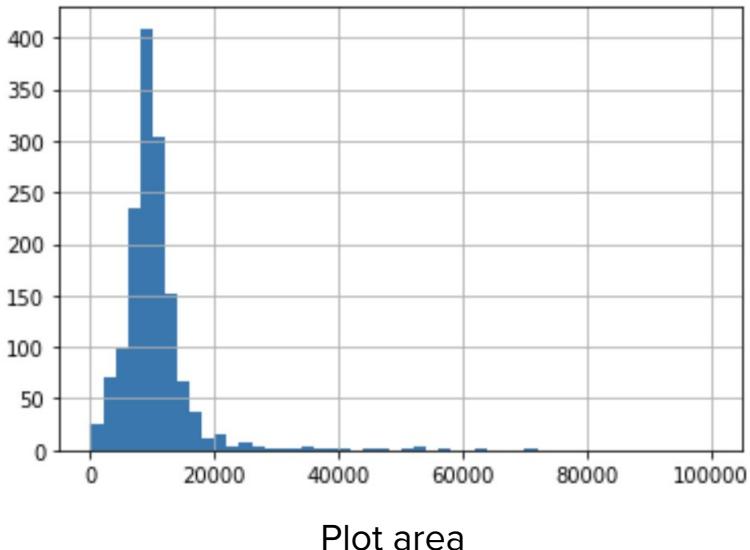
Histograms can help us identify:

- Whether a variable is **normally distributed** (a bell curve).
- Whether a variable contains lots of outliers.

This can help us to decide whether:

- A variable is suitable for being included in an analysis. (Some analyses rely on features being normally distributed, with few or no outliers.)
- Outliers need to be handled or removed.

Count





Guided Walk-Through:

Performing Basic Calculations

What is the distribution of the Time to Ship?

1. Open the **Guided Practice_02** worksheet.
2. Drag **Time to Ship** to Columns.
 - a. Creates “bins” to see the frequency of how often the time to ship occurs.
3. Go to **Show Me** and select **Histogram**.



Guided Walk-Through:

Performing Basic Calculations

What is the distribution of the Time to Ship?

1. Open the **Guided Practice_02** worksheet.
2. Drag **Time to Ship** to Columns.
 - a. Creates “bins” to see the frequency of how often the time to ship occurs.
3. Go to **Show Me** and select **Histogram**.
4. Click the down arrow on **Time to Ship (bin)** and select **Format**.
5. Under **Pane > Default Numbers > Numbers Custom** and set to zero decimals.
6. Holding the CTRL (or CMD) key, drag **Time to Ship (bin)** to Label.
 - a. The most frequent time to ship is 4 Days.
7. Drag **Ship Mode** to **Filter**, select **All**. Click the dropdown to show the filter to the right of the chart.
 - a. This will allow us to view each Ship Mode’s distribution around the median and the “shape” of the data (where does most of the data fall in the distribution).



Box Plot

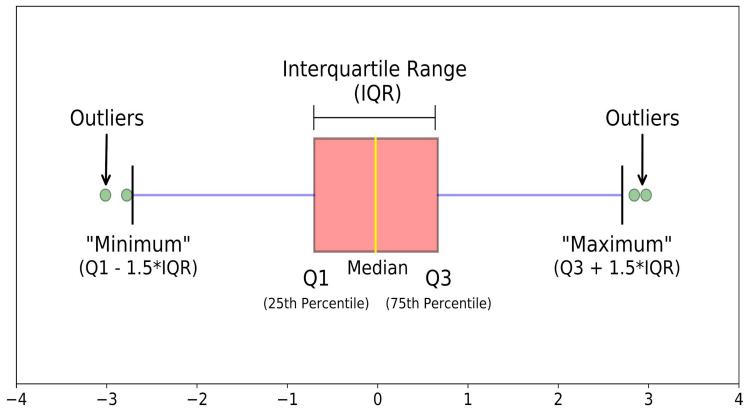
Box plots help us visualize a **five number summary** for a single variable:

1. **Min** = minimum value
2. **25%** = first quartile (Q1) = median of the lower half of the data
3. **50%** = second quartile (Q2) = median of the data
4. **75%** = third quartile (Q3) = median of the upper half of the data
5. **Max** = maximum value

Interquartile Range (IQR) = Q3 - Q1.

Outliers are:

- Below $Q1 - 1.5 * IQR$
- Above $Q3 + 1.5 * IQR$





Performing Basic Calculations

Are there any outliers in average Time to Ship by Weekday?

1. Open a new worksheet.
2. Drag **Ship Mode** to columns.
3. Drag **Time to Ship** to rows. Change the **SUM** aggregation to **AVG**.
4. Drag **Order Date** to **Detail**, change to a **Discrete Weekday**.
5. Go to **Show Me**, select **Box and Whisker Plots**.
6. Swap Rows and Columns.
7. Right click the y-axis, select **Edit Axis**. Uncheck **Include zero**.
8. Adjust the circles using **Size** on the **Marks Card**.
9. Add **Ship Mode** to **Color**.
10. Add an **Annotation** to the outlier in Standard Class.



Tree Map

Tree Maps is used to show the composition of a whole typically represented as rectangles.

- The size of the rectangle represents the size each variable is in relation to the whole.
- Great alternative to a Pie Chart when we have more than 3-5 part of a whole to display visually.
- They are only valid/effective if the data being represented are part of the same whole and does not display negative values.





Guided Walk-Through:

Applying Show Me to a Visualization | Treemaps

First, let's create a visualization that shows all sales per salesperson, then show the share of all sales by salesperson using a Treemap.

1. Click into the **Guided Practice_03** tab.
2. Drag **Salesperson** to **Columns**.
3. Drag **Sales** to **Rows**.
4. Right click on the **Sum(Sales)**, use **Quick Table Calculation**, select **Percent of Total**.
5. Click on **Show Me** and Select **Treemap**.
 - The Dimensions and Measures move to the Marks Card.
6. Hold the **CTRL** (CMD if using a Mac) key and drag one of the **Sum(Sales)** with the table calculation on the Marks Card to **Label**.



Scatterplot

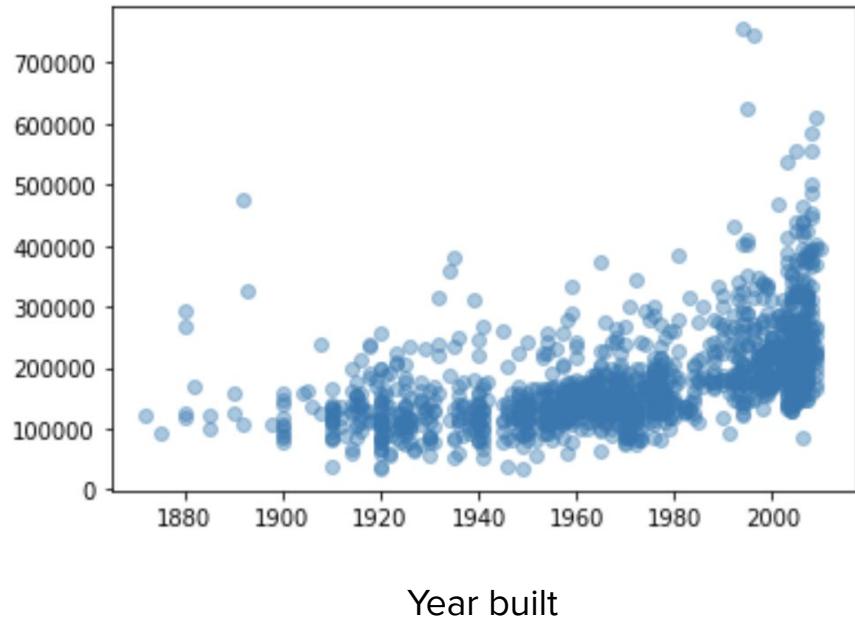
Scatterplots show the relationship between two or more quantitative variables as a general illustration.

Each dot on a Scatterplot represents one data point at the intersection of the dependent and independent variables.

Often when we're exploring a large data set, we'll want to answer questions like:

- What kind of relationship exists between a pair of variables?
- Does the dependent variable seem to have an effect on the independent variable?
- Are there outliers that can potentially skew our analysis?

House sale price



Correlation

Correlation (or the correlation coefficient or R² value) tells us whether there's an **association** between two variables. It can only take values from -1 to 1.

A **strongly positive** correlation between two variables **X** and **Y** means:

When **X** is high, **Y** is high

When **X** is low, **Y** is low

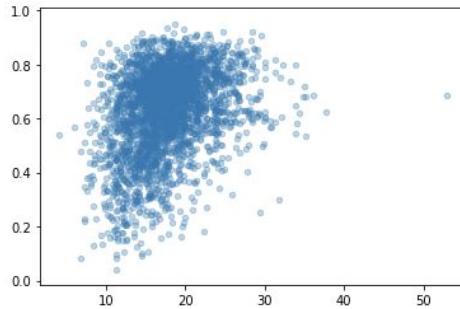
A **strongly negative** correlation between two variables **X** and **Y** means:

When **X** is high, **Y** is low

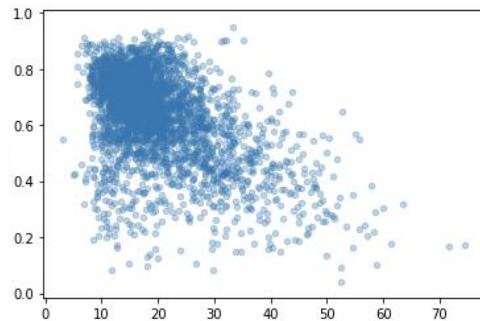
When **X** is low, **Y** is high

A correlation **close to zero** between two variables **X** and **Y** means there's no association between them, and both variables are just doing their own thing.

correlation = 0.4



correlation = -0.3



Reminder: Correlation does not equal causation!



Applying Show Me to a Visualization | Scatterplots

Let's create a Scatterplot to show the effect of discounts on sales for Subcategory.

1. Click into the **Guided Practice_04** tab.
2. Drag **Sales** to **Columns**.
3. Drag **Discount** to **Rows**. Change the aggregation to **Average**.
4. Use **Show Me** and select **Scatterplot**
5. Drag **Subcategory** to **Color**.
6. Right click on the Scatterplot, select **Trendlines > Show Trendlines**.
7. Hover over the **Trendline** to see information including R-squared (fit of regression line) and P-value (statistical significance).
8. Right click on the dot for **Appliances**. Choose **Annotate > Mark**.
9. Repeat for Phones to add an annotation.

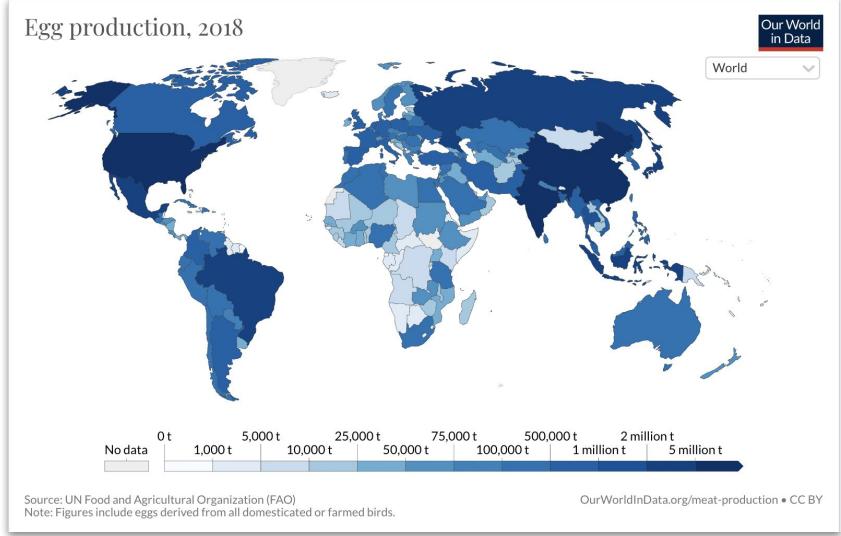
Filter Using Relevant Data



Filled Map Chart (Basic Choropleth)

Filled map charts are the most common way of representing data on a map. They are used when showing the location of data is more important than anything else. For example, showing population density, election results, or the prevalence of a language being spoken across the world.

The filled map or (basic choropleth) splits a map into areas, and color codes those areas based on the proportion of the measure(s) being shown.





Guided Walk-Through:

Applying Show Me to a Visualization | Filled Maps

Now, we want to show the total Sales per Country. Go to the **Guided Practice_05** tab.

1. Drag **Country** to **Rows**.
2. Drag **Sales** to **Columns**.
3. Click **Show Me** to use a **Filled Map**.
 - a. Country and Sales Move to the Marks Card.
4. Drag **Sales** to Label.

The screenshot shows the 'Format' dialog box for the 'SUM(Sales)' field. The 'Default' tab is selected. In the 'Numbers' section, the 'Font' is set to 'Tableau Boo...' and the 'Numbers' value is '0.1M'. The 'Totals' section shows various options like 'Automatic', 'Number (Standard)', and 'Number (Custom)', with 'Number (Custom)' currently selected. In the 'Special Values' section, the 'Thousands separator' checkbox is checked. To the right of the dialog, a 'Show Me' card is visible, showing filters for 'Order Date', 'Region', 'Sub Region', and 'Salesperson'. Below the card is a map visualization.



Guided Walk-Through:

Applying Show Me to a Visualization | Filled Maps

Now, we want to show the total Sales per Country. Go to the **Guided Practice_05** tab.

1. Drag **Country** to **Rows**.
2. Drag **Sales** to **Columns**.
3. Click **Show Me** to use a **Filled Map**.
 - a. Country and Sales Move to the Marks Card.
4. Drag **Sales** to Label.
5. Right click on any sales number on the map.
Choose **Format**.
6. In the Format pane, choose **Fields > Sum(Sales) > Pane**.
7. Set the **Default Numbers** to **Number (Custom)** to report in **Millions(M)** and **1 decimal**.

The screenshot shows the 'Format SUM(Sales)' pane in Tableau. The 'Fields' section is selected. The 'Default' tab is active, showing 'Font: Tableau Boo...' and 'Alignment: Automatic'. Under 'Numbers:', '0.1M' is selected. A dropdown menu for 'Number (Custom)' is open, with 'Number (Standard)' highlighted. Other options include 'Automatic', 'Currency (Standard)', 'Scientific', 'Percentage', and 'Custom'. To the right, the 'Number (Custom)' settings are displayed: 'Decimal places:' set to 1, 'Negative values:' set to '-1234', 'Display Units:' set to 'Millions (M)', and 'Prefix / Suffix:' fields empty. A checkbox for 'Include thousands separators' is checked. The 'Special Values' and 'Step' tabs are also visible.



Applying Show Me to a Visualization | Filled Maps

Using Filters we can tie relevant values together to create a “cascading filter”.

1. Return to the **Guided Practice_05** tab.
2. Add **Region**, **Sub Region**, **Salesperson** to Filters (select All when prompted)
3. Show Filters as you add each one.
4. On the Filters at the right, Click the dropdown and set **Only Relevant Values**, for Region, Sub Region, and Salesperson.
5. Click the filters to see how they interact.

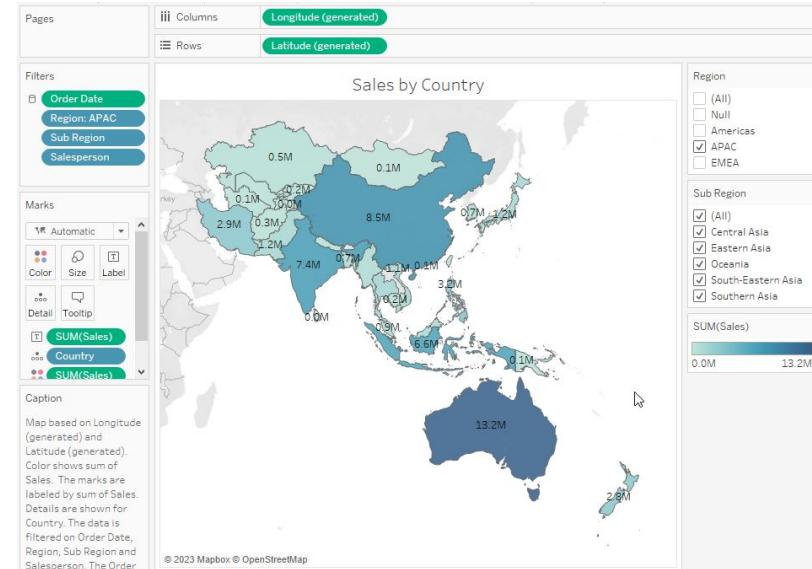


Table Calculations

Granularity and Aggregation

What is the difference between the following calculated field formulas, assuming we are trying to determine net profit?

- Profit ratio = $\text{SUM}([\text{PROFIT}]/[\text{SALES}])$
- Profit ratio = $\text{SUM}([\text{PROFIT}])/ \text{SUM}([\text{SALES}])$

Sales	Profit	Profit Ratio
\$100	15	15%
10	5	50%
50	20	40%
<u>100</u>	<u>30</u>	<u>30%</u>
		135%
\$260	\$ 70	27%

Granularity and Aggregation | Excel Calculations

Here's an illustration using Excel calculations:

The image shows two identical Excel spreadsheets side-by-side, each containing a table of data with four columns: Row ID, Profit, Sales, and Profit Ratio. The top spreadsheet is titled "INCORRECT - profit/sales" and contains a formula =SUM(D3:D7) in the cell for the Profit Ratio of the total row. The bottom spreadsheet is titled "CORRECT - SUM(profit) / SUM(sales)" and contains a formula =B8/C8 in the same cell.

A	B	C	D
INCORRECT - profit/sales			
1			
2	Row ID	Profit	Sales
3	89536	\$1.32	\$19.97
4	89547	\$51.35	\$132.10
5	89549	-\$1,685.04	\$180.29
6	89550	-\$415.70	\$3,422.18
7	89536	\$455.42	\$660.03
8	TOTAL	-\$1,592.64	\$4,414.57
$=\text{SUM}(\text{D3:D7})$			
SUM			
CORRECT - $\text{SUM}(\text{profit}) / \text{SUM}(\text{sales})$			
1			
2	Row ID	Profit	Sales
3	89536	\$1.32	\$19.97
4	89547	\$51.35	\$132.10
5	89549	-\$1,685.04	\$180.29
6	89550	-\$415.70	\$3,422.18
7	89536	\$455.42	\$660.03
8	TOTAL	-\$1,592.64	\$4,414.57
$=\text{B8}/\text{C8}$			

Table Calculations

Quick table calculations in Tableau are analogous to “Show data as” in Excel.

They allow us to **change how we display the data** and **perform some brief analysis** in our worksheets and dashboards.

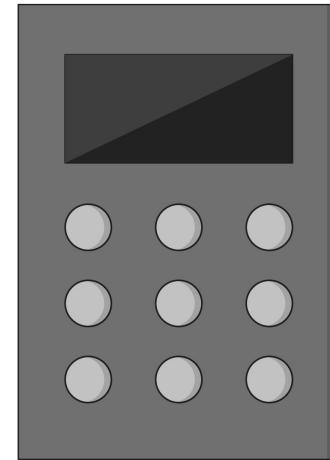
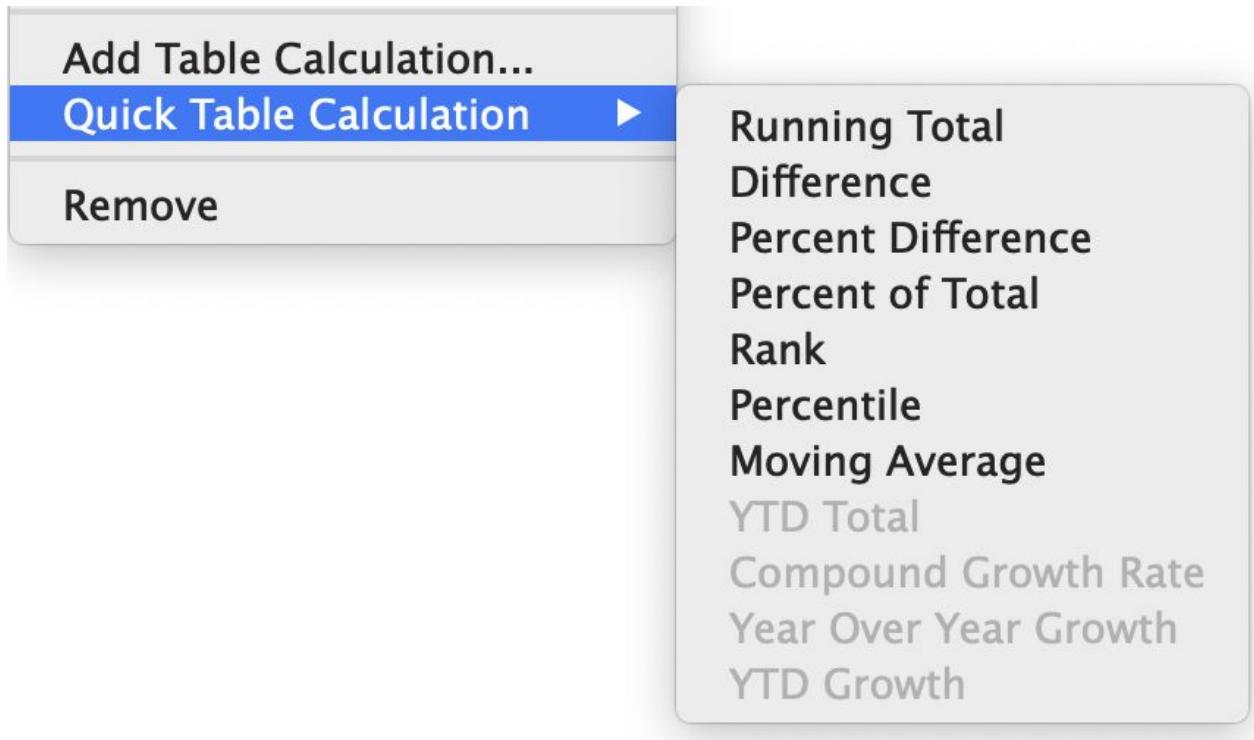


Table Calculations (Cont.)

A table calculation, such as **year-over-year profit growth**, can run across or down a measure.

Year of Order Date	Quarter of Order Date	Profit	Sales
2016	Q1	\$35,994	\$501,837
	Q2	\$49,861	\$1,261,058
	Q3	\$66,777	\$2,672,205
	Q4	\$103,819	\$5,184,743
2017	Q1	\$45,652	\$3,675,226
	Q2	\$86,874	\$5,816,826
	Q3	\$95,070	\$8,730,155
	Q4	\$106,843	\$12,625,317

Table Calculations (Cont.)





Discussion:

Performing Logical Calculations

?

Should we look at our Shipping Performance calculation to reflect each Subcategory versus on a per-order basis? Yes or no?





Performing Logical Calculations

In order to analyze shipping times by subcategory, we need to create a new calculated field using **IF/THEN** logic:

1. Create a new “**Shipping Performance**” field.
2. Use the following formula, then click **OK**.

```
IF [Time to Ship]< 1 THEN "Good"  
ELSEIF [Time to Ship]< 3 THEN "OK"  
ELSE "Bad"  
END
```

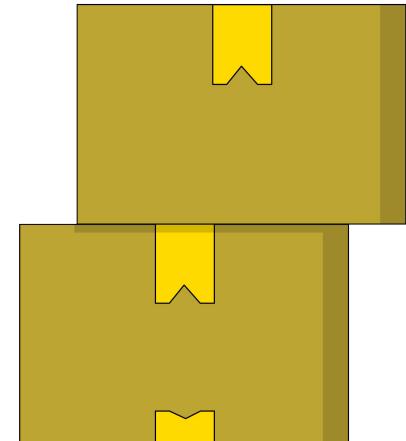


Guided Walk-Through:

Performing Logical Calculations (Cont.)

Use **Guided Practice_06** to create a new visualization that shows each subcategory Shipping Performance. Each row should total 100%.

1. Drag **Subcategory** to **Rows**.
2. Drag **Shipping Performance** to **Columns**.
3. Drag **orders(count)** to **Columns** to the right of Shipping Performance.
4. Click the dropdown on **CNT(orders.csv)**. Choose **Quick Table Calculation** then **Percent of Total**.
5. Click the dropdown on **CNT(orders.csv)** again.
6. Hover on **Compute Using**, choose **Shipping Performance**.





Guided Walk-Through:

Performing Logical Calculations (Cont.)

Continue with **Guided Practice_06** to add color, labels, and sort the shipping performance manually.

1. Hold the **CTRL** (**CMD** if using a Mac) key and drag **CND(orders.csv)** with the table calculation to **Color** on the **Marks Card**.
Repeat but drag to **Label**.
2. Click the dropdown on **Shipping Performance**.
3. Select **Sort > Manual**. Use the arrows to sort by Good, OK, then Bad.



More on Filters

The Filters Order of Operations

Filtering plays an integral role in data analysis. Tableau has five types of filters, which are executed in this order:

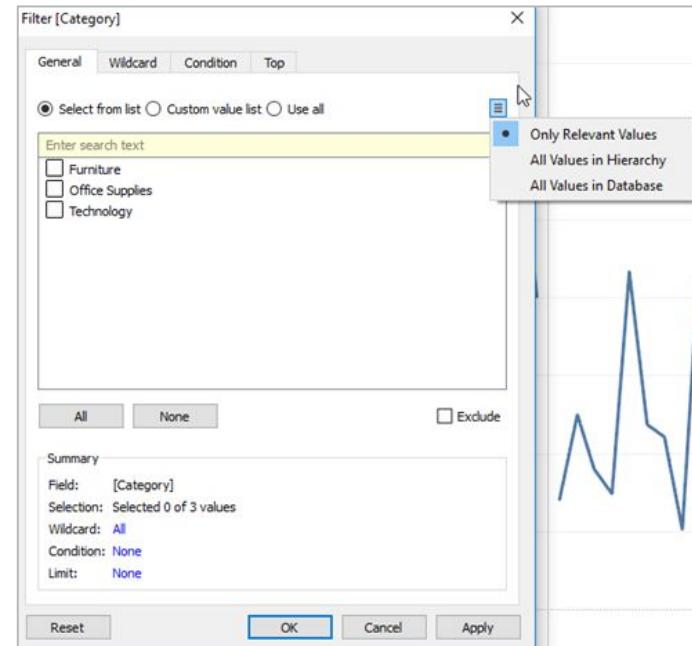
1. Extract filters.
2. Data source filters.
3. Context filters.
4. Filters on dimensions.
5. Filters on measures.



Filtering Dimensions

When dragging a dimension from the data pane to the **Filters** shelf in Tableau Desktop, you'll see a dialog box with these tabs:

- **General:** Select the values you want to include or exclude.
- **Wildcard:** Define a pattern to filter by.
- **Condition:** Define rules to filter by.
- **Top:** Define a formula that calculates the data to be included in the view.





Discussion:

How Many Filters Do You Need?

Let's take a moment to think about the following request from Superstore:



How can we find out if each salesperson has a sales category at which they excel?

How many filters will you apply to the data set? What might they be?

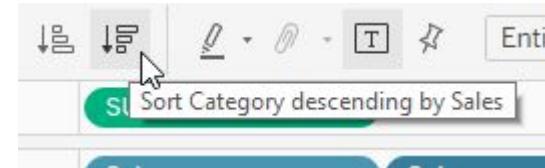


Guided Walk-Through:

Applying Multiple Filters

First, we need to create a visualization that shows Sales by Salesperson per Category. Does any Salesperson excel in any Category?

1. Click on the **Guided Practice_07** tab.
2. Drag **Salesperson** and **Category** to **Rows**.
3. Drag **Sales** to **Columns and Labels**
4. Drag **Category** to **Color**.
5. Click the **Sort Descending** button to sort by **Category** and **Sales**.
6. Click the dropdown on **Salesperson** in **Rows, Show Filters**.
7. Click the dropdown on **Category** in **Rows, Show Filters**.

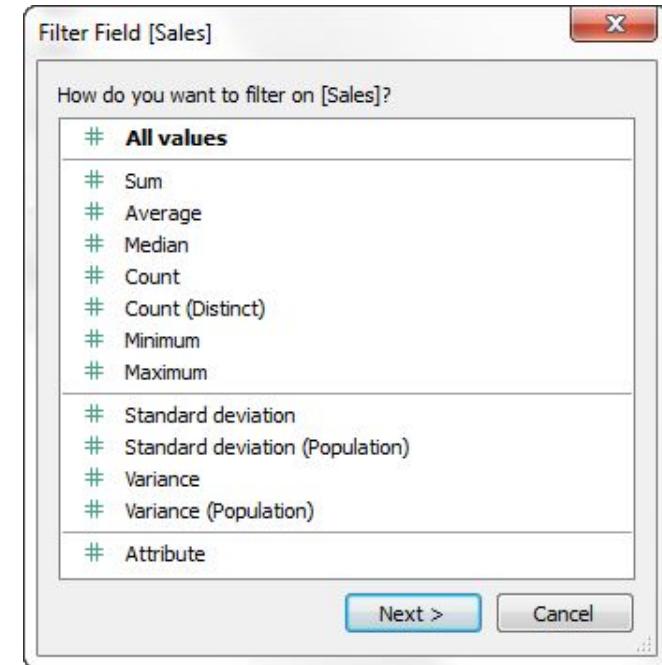


Filtering Measures

When dragging a measure from the **Data** pane to the **Filters** shelf in Tableau Desktop, you'll see a dialog box with a list of aggregations.

After clicking **Next**, you'll be asked to choose from these types of filters:

- Range of Values
- At Least
- At Most
- Special (Tableau Desktop only)





Guided Walk-Through: Filtering Measures

Let's return to **Guided Practice_07**, which shows the Sales per Salesperson and Category. We can filter on **SUM(Sales)** in order to only see salespeople and categories that made more than a certain amount in Sales and Profit. Are these high producing salesperson profitable at a glance?

1. Drag **Sales** to **Filters**.
2. Choose **SUM** as your aggregation.
3. Then, using the **At Least** option, set a minimum of **\$5,000,000** in Sales.
4. Add **Labels**. Notice only salespeople and categories with the minimum are displayed.



Group Exercise:

Applying a Single Filter

5 Minutes



Go to **Group Practice_01**, to create Returns per Salesperson and Category.

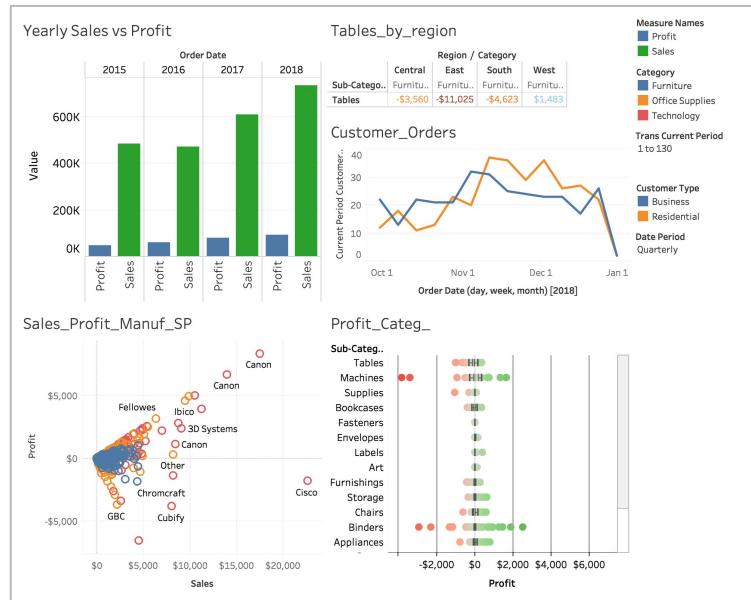
1. Drag **Reason Returned** and **Salesperson** to Rows.
2. Drag **Sales** to Columns.
3. Drag the **Reason Returned** and **Salesperson** to **Filters**.
 - a. Check everything except for “**Null**” in both lists. Show the filters.
4. Drag **Category** to **Color**.
5. Drag Sales to Filters. Using the **At Least** option, set a minimum of **\$200,000** in Sales.
6. Add **Labels**. Notice only salespeople and categories with the minimum are displayed.

Starting a Dashboard

What Is a Tableau Dashboard?

Dashboards are the main use of Tableau.

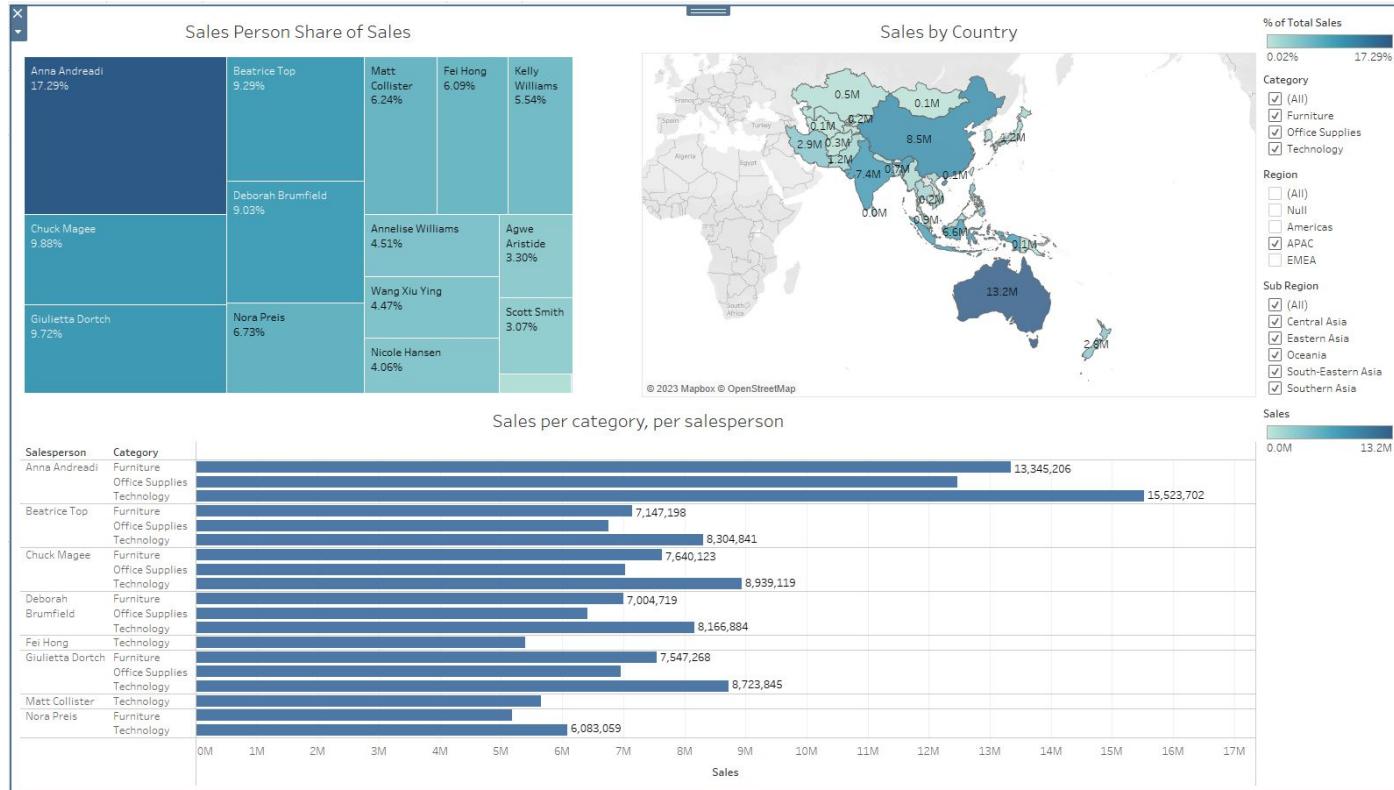
They allow casual viewers to **interact with data, manipulate views, and perform analysis** without having to go through the entire analytics workflow (i.e., Extract and Prepare).





Discussion:

What Goes Into a Dashboard?





Group Exercise:

From Worksheets to a Dashboard

Go to **Group Practice_02**. Add the following charts to the dashboard:

- **Guided Practice_03**
- **Guided Practice_05**
- **Guided Practice_07**

Manipulate the size of the dashboard using the controls on the dashboard panel on the left-hand side. Try changing the size of your charts, too!

In our next lesson, we will dive deeper into dashboards.

Data Manipulation in Tableau

Wrapping Up



Solo Exercise:

Optional Homework

Using your initial hypotheses, validate initial answers to the following questions. Add your visualizations to the Homework tabs in your workbook.

- **Homework 01:** Which product categories or subcategories are not doing well?
- **Homework 02:** Can a product be a problem even if it has high sales?
- **Homework 03:** How should we define underperforming products?
- **Homework 04:** If a product is not performing well but has a large market share (i.e., lots of customers), is it worth keeping?
 - Hint: Create a calculated field to do a distinct count for Total Customers: COUNTD([Customer Id]).

Use the caption boxes on the Tableau worksheets to share your insights.

Recap

Today, we...

- Prepared data for import into Tableau
- Navigated the Tableau interface to build visualizations
- Aggregated measures and dimensions
- Worked with Marks Card and discrete versus continuous dates
- Connected to the PostgreSQL server
- Created calculated fields to analyze data
- Applied filters to single or multiple sheets

Looking Ahead

Homework:

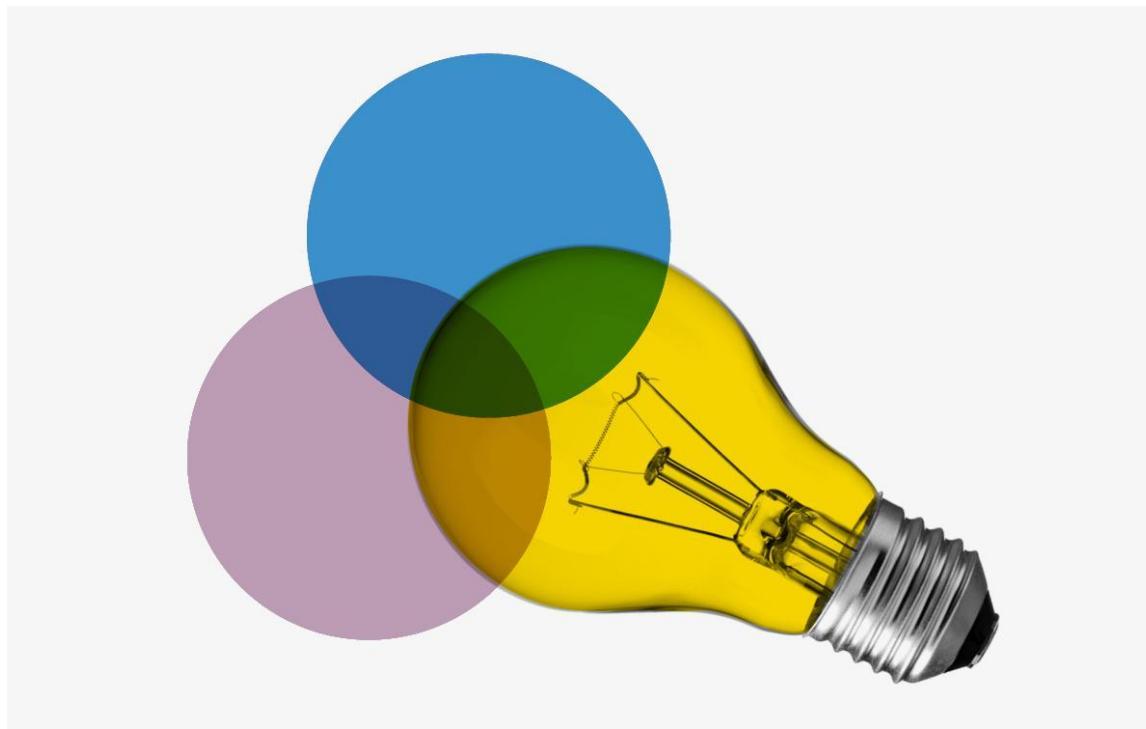
- Final project work
- Optional homework in the workbook

Up Next:

Dashboards and Data Narratives
in Tableau



Don't Forget: Exit Tickets!



Additional Resources

- [Visual Analysis Best Practices](#)
- [A Quick Guide to Better Visualizations \(Tableau\)](#)
- [Visualizing Time: Beyond the Line Chart](#)
- [Tableau “Starter Kits” \(role-based tutorials\)](#)

Additional Resources

- Additional Examples of Calculated Fields (Interworks):
<https://www.interworks.com/blog/ccapitula/2015/03/24/tableau-essentials-calculated-fields-introduction>
- Differences Between Calculated Fields and Table Calculations:
<https://www.thedataschool.co.uk/lena-caric/difference-table-calculations-calculated-fields/>
- Example: Descriptive and Inferential Analysis Dashboard Using Superstore Data Set:
<https://www.ryansleeper.com/super-sample-superstore/>

