BUSINESS INTELLIGENCE AND DATA ANALYTICS

22MCA254

MODULE 1

MODULE-1	CREATING VISUALIZATIONS AND DASHBOARD	22MCA254.1	8 Hours

Connections to data, Foundations for building visualizations, Visualizing data, Creating charts, Creating geographic visualizations.

Working with Data: Connecting to data, Managing data source metadata, Working with extracts instead of live connections, File types, Joins and blends, Filtering data

CREATING VISUALIZATIONS AND DASHBOARD

Introduction to Business Intelligence (BI)

Definition

- **Business Intelligence (BI)** is the technology-driven process of analyzing data and presenting actionable information to help executives, managers, and other corporate end users make informed business decisions.

Importance of BI

- Enhances decision-making capabilities by providing accurate and timely information.
- Improves efficiency by optimizing business processes.
- Increases competitive advantage through better market understanding.
- Helps in identifying business opportunities and trends.

Components of BI

- **Data Warehousing:** Central repository of integrated data from multiple sources.
- **Data Mining:** Extracting patterns and knowledge from large amounts of data.
- **OLAP (Online Analytical Processing):** Enables users to analyze multidimensional data interactively.
- **Reporting and Querying Software:** Tools to generate informative reports and queries.
- **Dashboards:** Visual interfaces displaying key performance indicators (KPIs) and metrics.

Data Analytics

- The term data analytics refers to the science of analyzing raw data to make conclusions about information.
- Many of the techniques and processes of data analytics have been automated into mechanical processes and algorithms that work over raw data for human consumption.
- Data analytics can be used by different entities, such as businesses, to optimize their performance and maximize their profits.
- This is done by using software and other tools to gather and analyze raw data.

Need for Data Anaytics

1. **Gain greater insight into target markets**

- Access digital footprints to learn customer preferences and needs.
- Understand customer browsing and purchasing behaviors.
- Analyze data to identify market trends and patterns.
- Customize products and services to meet specific customer needs.
- Increase customer loyalty by understanding their desires.
- Ensure customer satisfaction and happiness.
- Boost sales through targeted insights.
- Prevent losing customers to data-savvy competitors.

2. Enhance decision-making capabilities

- Make faster, better-informed business decisions.
- Avoid ineffective strategies and inefficient operations.
- Prevent misguided marketing campaigns and unproven concepts.
- Use a data-driven decision-making model for proactive opportunities.
- Rely on data validity rather than intuition or experience.
- Understand the potential inaccuracies in data analysis.

3. Create targeted strategies and marketing campaigns

- Businesses can also use data to inform their strategies and drive targeted marketing campaigns to help ensure promotions engage the right audiences.
- By analyzing customer trends, monitoring online shopping, and evaluating point-of-sale transactional data, marketers can create customized advertising to reach new or evolving consumer segments and increase the efficiency of overall marketing efforts.
- And by taking advantage of these insights on consumer behavior and desires in customer-oriented marketing, businesses can meet and exceed expectations, boost brand loyalty, and encourage growth.
- If you are interested in developing targeted marketing or advertising campaigns, it’s critical you understand the process by which quantitative social science and data science research is conducted.

4. Improve operational inefficiencies and minimize risk

- Another major benefit to data analytics is the ability to use insights to increase operational efficiencies. By collecting large amounts of customer data and feedback, businesses can deduce meaningful patterns to optimize their products and services.
- Data analytics can also help organizations identify opportunities to streamline operations, reduce costs, or maximize profits. Companies can use insights from data analytics to quickly determine which operations lead to the best results—and which areas are underperforming.
- This allows decision-makers to adjust their strategies accordingly and proactively anticipate problems, manage risks, and make improvements.
- Predictive modeling of data is one of the most sought-after skills in data science because it can help companies strategize future investments, nonprofits organize fundraising drives, or political candidates decide where to focus their canvassing efforts.

5. Identify new product and service opportunities

- When it comes to innovation, data analytics allows businesses to understand their current target audience, anticipate and identify product or service gaps, and develop new offerings to meet these needs.
- Not only can companies use data to track customer feedback and product performance in real-time, they can also track what rivals are doing so they can remain more competitive.
- Insights from data analytics can also allow organizations to update their existing products or services to reflect changing consumer demands, tweak marketing techniques, and optimize customer services.
- The enhanced adaptability afforded by big data can mean the difference between thriving or failing as a business.

Types of Data Analytics

Descriptive analytics: This describes what has happened over a given period of time. Have the number of views gone up? Are sales stronger this month than last?

Diagnostic analytics: This focuses more on why something happened. It involves more diverse data inputs and a bit of hypothesizing. Did the weather affect beer sales? Did that latest marketing campaign impact sales?

Predictive analytics: This moves to what is likely going to happen in the near term. What happened to sales the last time we had a hot summer? How many weather models predict a hot summer this year?

Prescriptive analytics: This suggests a course of action. For example, we should add an evening shift to the brewery and rent an additional tank to increase output if the likelihood of a hot summer is measured as an average of these five weather models and the average is above 58%,

Data Connection Tools:

BI Tools:

- **Tableau:** A powerful tool for creating interactive and shareable dashboards.
- **Power BI:** Microsoft's business analytics tool for data visualization and reporting.
- **QlikView:** Another popular BI tool for creating guided analytics applications.

Data Integration Tools:

- **Apache Nifi:** A software project designed to automate the flow of data between systems.
- **Talend:** An open-source data integration tool.

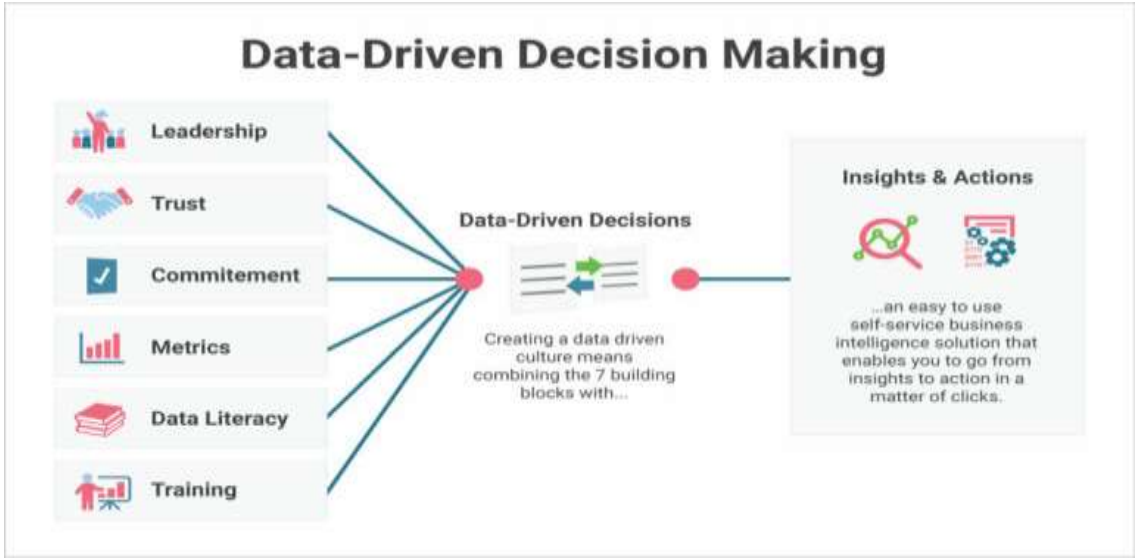
Difference Between BI and Data Analytics

Business Intelligence	Data Analytics
It is the process of using tools and technologies to gather, store, and analyze data to support business decision-making.	It is the science of analyzing raw data to make conclusions about that information.
It focuses on historical data and current data to drive business decisions.	It focuses on discovering patterns, predicting future trends, and making recommendations based on data.
The purpose is To provide a comprehensive overview of the business and assist in strategic decision-making.	The purpose is to perform deep analysis to understand data and derive insights that can lead to predictive modeling and decision-making.
The techniques used are Reporting, dashboards, scorecards, data warehousing.	The techniques used are Statistical analysis, machine learning, data mining, predictive modeling.
Some of the tools are Power BI, Tableau, SAP BusinessObjects, QlikView.The output will be in the form of Reports, visualizations,	Some of the tools are R, Python, Apache Spark, SAS, MATLAB.The output will be Insights, patterns, predictions, models.

Data-Driven Decision Making

Utilizing data analysis and interpretation to guide strategic business decisions.

Enhances objectivity and reduces reliance on intuition or "gut feeling."



Key Roles in BI and Data Analytics

Data Scientist

- Uses advanced statistical, analytical, and programming skills to collect, analyze, and interpret large datasets.

Data Analyst

- Examines data to generate actionable insights and reports, often using tools like SQL, Excel, and BI software.

BI Developer

- Designs and builds BI solutions, including data warehouses, dashboards, and reports.

Overview of Data-Driven Decision Making

Data Collection: Gathering data from various sources.

Data Processing: Cleaning and organizing the collected data.

Data Analysis: Applying statistical methods to interpret the data.

Data Interpretation: Making sense of the data analysis results.

Decision Making: Using insights from the data to inform business strategies and actions.

Key Takeaways

- **BI** focuses on providing historical insights through reports and dashboards.
- **Data Analytics** includes predictive and prescriptive analytics to forecast and optimize future outcomes.
- Both BI and Data Analytics are critical for informed decision-making in modern businesses.

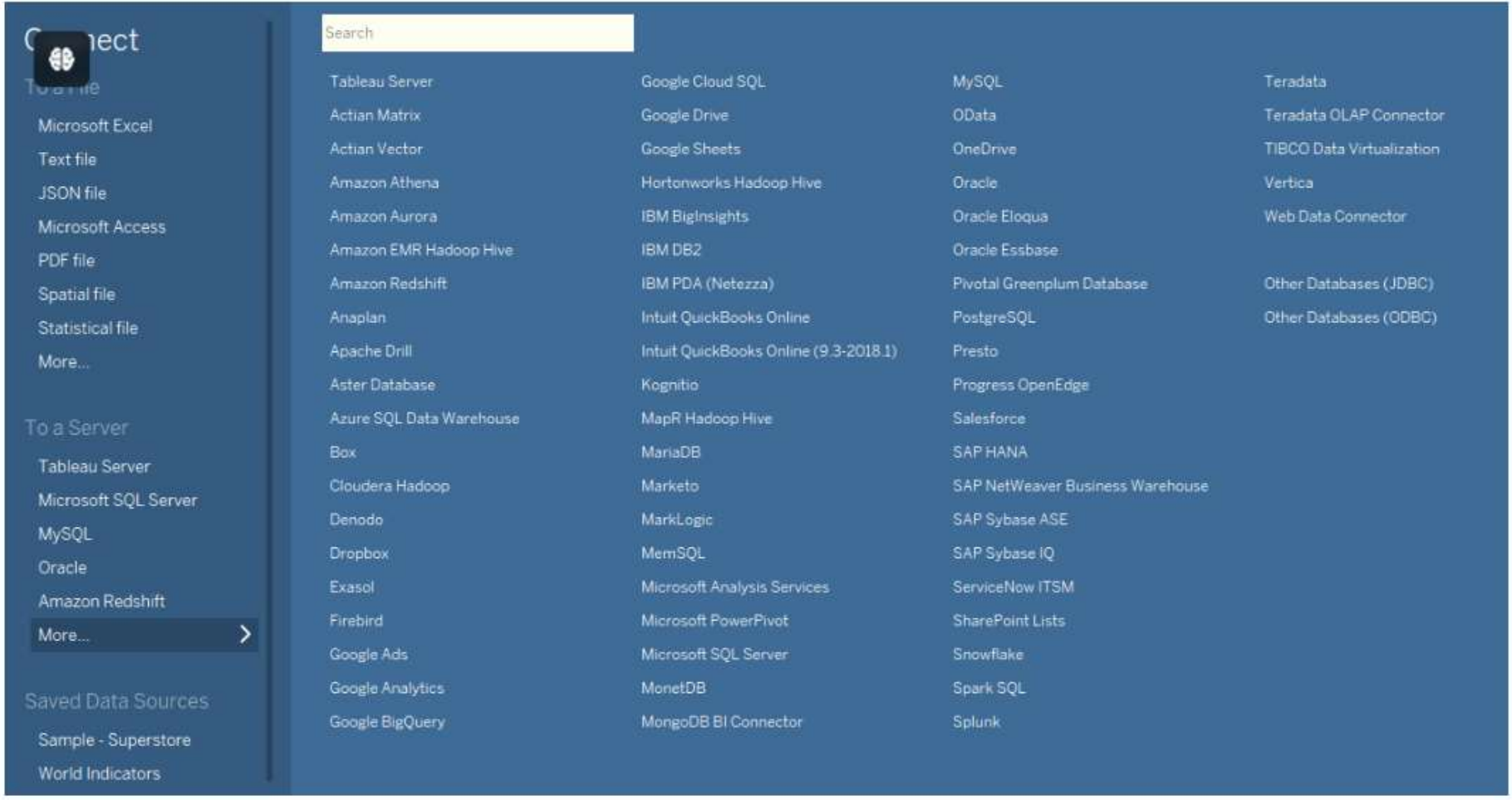
Data Connection with Data Sources

- Tableau can connect with all the accessible data sources which are broadly used. It can link to Excel files, PDF files, text files, etc. It can also connect to various databases using its ODBC connector. Tableau can connect to web connectors and servers.

  Tableaus native connectors can connect to the following types of data sources:

- **File Systems:** Such as Microsoft Excel, CSV, etc.
- **Cloud Systems:** Such as Google bigQuery, Windows Azure, etc.
- **Relational System:** Such as Microsoft SQL Server, Oracle, DB2, etc.
- **Other Sources:** It uses ODBC.

The given below picture shows all of the data sources available through Tableau's native data connectors.



Connect Live

- The Connect Live feature is used in real-time data analysis. In connect live case, Tableau connects with the ral-time data source, and it keeps reading the data.
- Thus, the result of the data analysis is up to the second, and the latest changes are reflected in this result.
- However, the drawback, it's the source system as it has to keep sending the data to Tableau.

In-Memory

- Tableau can also process the data in memory by caching them in memory, and it not being connected to the source anymore while analyzing the data.
- Of course, there will be a limit on the amount of data cached depending on the availability of the memory.

Combine Data Sources

- Tableau can connect with different data sources at the same time.
- **For example:** In a single workbook, you can connect to a relational source and a flat file by defined the multiple connections.
- This is also used in data blending, which is a unique feature in Tableau.

Understanding Data Types:

Numeric Data:

- **Integer:** Whole numbers (e.g., 1, 2, 100).
- **Float:** Decimal numbers (e.g., 1.5, 3.14).

Categorical Data:

- **String:** Text data (e.g., names, categories).
- **Boolean:** True/False values.

Date/Time Data:

- Timestamps representing dates and times.

Data Preparation:

Data Cleaning:

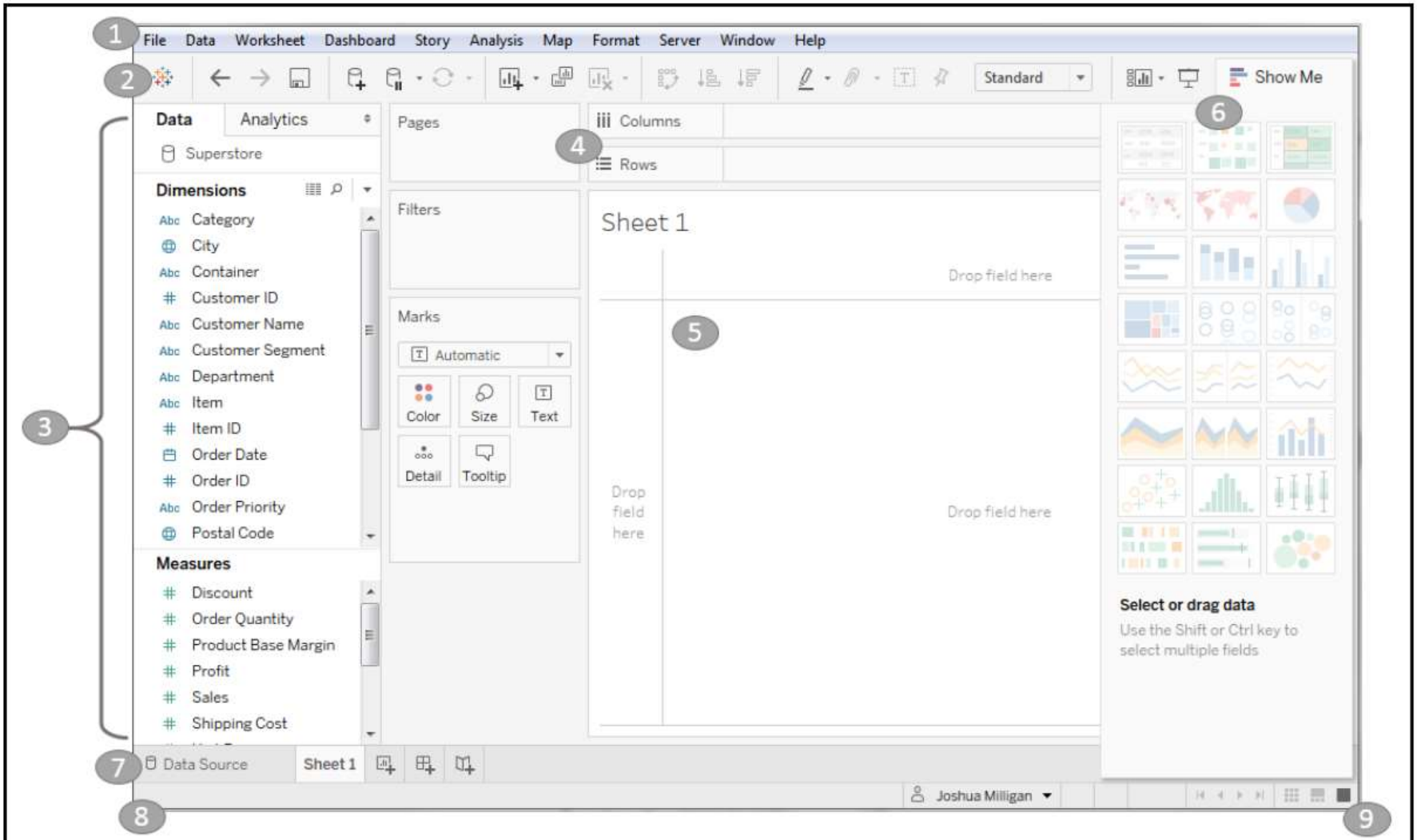
- Removing duplicates.
- Handling missing values (e.g., filling, deleting).
- Correcting errors (e.g., typos, incorrect values).

Data Structuring:

- **Normalization:** Organizing data to reduce redundancy (e.g., splitting data into related tables).
- **Denormalization:** Combining tables to improve read performance (e.g., flattening data for faster querying).

Connecting to the Data

1. Connecting to Text file
2. Connecting to Excel File


1
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
2
Standard
Show Me
Data
Analytics
Superstore
Dimensions
Abc Category
City
Abc Container
Customer ID
Abc Customer Name
Abc Customer Segment
Abc Department
Abc Item
Item ID
Order Date
Order ID
Abc Order Priority
Postal Code
Measures
Discount
Order Quantity
Product Base Margin
Profit
Sales
Shipping Cost
3
Pages
Filters
Marks
Automatic
Color
Size
Text
Detail
Tooltip
4
Columns
Rows
Sheet 1
Drop field here
5
Drop field here
Drop field here
6
Select or drag data
Use the Shift or Ctrl key to select multiple fields
7
Data Source
Sheet 1
8
Joshua Milligan
9

1. The Menu contains various menu items for performing a wide range of functions.

2. The Toolbar allows for common functions such as undo, redo, save, add a data source, and so on.

3. The Side Bar contains tabs for Data and Analytics. When the Data tab is active, we'll refer to the side bar as the data pane. When the Analytics tab is active, we'll refer to the side bar as the analytics pane.

4. Various shelves such as Columns, Rows, Pages, and Filters serve as areas to drag and drop fields from the data pane. The Marks card contains additional shelves such as Color, Size, Text, Detail, and Tooltip. Tableau will visualize data based on the fields you drop on to the shelves.

5. The canvas or view is where Tableau will draw the data visualization. You may also drop fields directly on to the view. You'll find the seamless title at the top of the canvas. By default, it will display the name of the sheet, but it can be edited or even hidden.

6. Show Me is a feature that allows you to quickly iterate through various types of visualizations based on data fields of interest.

7. The tabs at the bottom of the window give you options for editing the data source, as well as navigating between and adding any number of sheets, dashboards, or stories. Many times, any tab (whether it is a sheet, a dashboard, or a story) is referred to generically as a sheet.

8. As you work, the status bar will display important information and details about the view, selections, and the user.

9. Various controls allow you to navigate between sheets, dashboards, and stories, as well as view the tabs with Show Filmstrip or switch to a sheet sorter showing an interactive thumbnail of all sheets in the workbook. Now that you have connected to the data in the text file, we'll explore some examples that lay the foundation for data visualization and then move on to building some foundational visualization types.

Important Terms

- A Tableau **workbook** is a collection of data sources, sheets, dashboards, and stories. All of this is saved as a single Tableau workbook file (.twb or .twbx).
- A workbook is organized into a collection of tabs of various types: A sheet is a single data visualization, such as a bar chart or a line graph. Since Sheet is also a generic term for any tab, we'll often refer to a sheet as a view because it is a single view of the data.
- A dashboard is a presentation of any number of related views and other elements (such as text or images) arranged together as a cohesive whole to communicate a message to an audience. Dashboards are often designed to be interactive.
- A story is a collection of dashboards or single views arranged to communicate a narrative from the data. Stories may also be interactive

Foundations for Building Visualisations

- 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 on the left Side Bar.
- Fields can be dragged from the data pane onto the canvas area or onto various shelves such as Rows, Columns, Color, or Size.

- Tableau identifies each field as a dimension or measure in the **Data** pane, depending on the type of data the field contains. You use these fields to build your analysis.
- Data fields are made from the columns in your data source. Each field is automatically assigned a data type such as integer, string, or date, and a role: a discrete dimension or continuous measure (or less commonly, a continuous dimension or discrete measure).
- *Dimensions* contain qualitative values (such as names, dates, or geographical data).
- You can use dimensions to categorize, segment, and reveal the details in your data. Dimensions affect the level of detail in the view.
- *Measures* contain numeric, quantitative values that you can measure. Measures are aggregated by default. When you drag a measure into the view, Tableau applies an aggregation on the pill.

Blue and Green Fields

Tableau represents data differently in the view depending on whether the field is discrete or continuous). *Continuous* and *discrete* are mathematical terms.

- *Continuous* means "forming an unbroken whole, without interruption". These fields are colored green. When a continuous field is put on the Rows or Columns shelf, an axis is created in the view.
- *Discrete* means "individually separate and distinct." These fields are colored blue. When a discrete field is put on the Rows or Columns shelf, a header is created in the view.

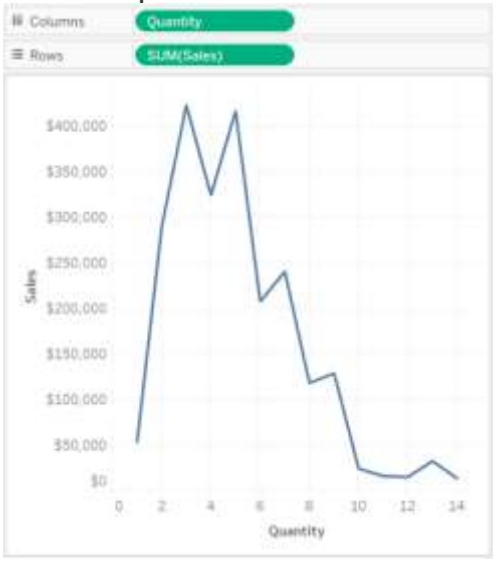
Possible combinations of fields in Tableau

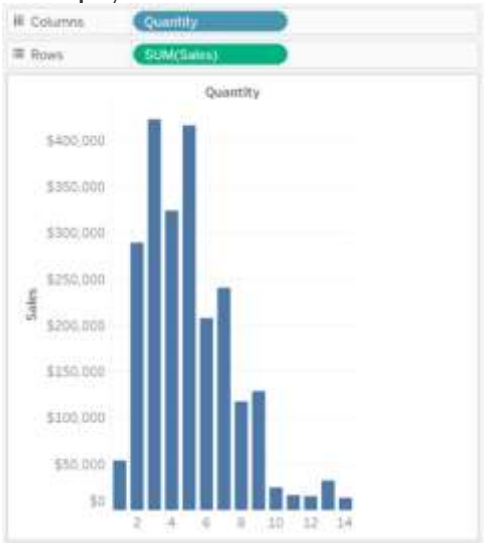
- This table shows examples of what the different fields look like in the view. An instance of a field in the view (on the Rows shelf, for example) is sometimes referred to as a "pill".

Discrete Dimensions	Product Name
Continuous Dimensions (dimensions with a data type of String or Boolean can't be continuous)	YEAR(Order Date)
Discrete Measures	SUM(Profit)
Continuous Measures	SUM(Profit)

Examples of Continuous and discrete fields used in view

- In the example on the left, because the **Quantity** field is Continuous, it creates a horizontal axis along the bottom of the view. The green pill and the axis help you to see that it's a continuous field.
- In the example on the right, the Quantity field is **Discrete**. It creates a horizontal headers instead of an axis. The blue pill and the horizontal headers help you to see that it's discrete.





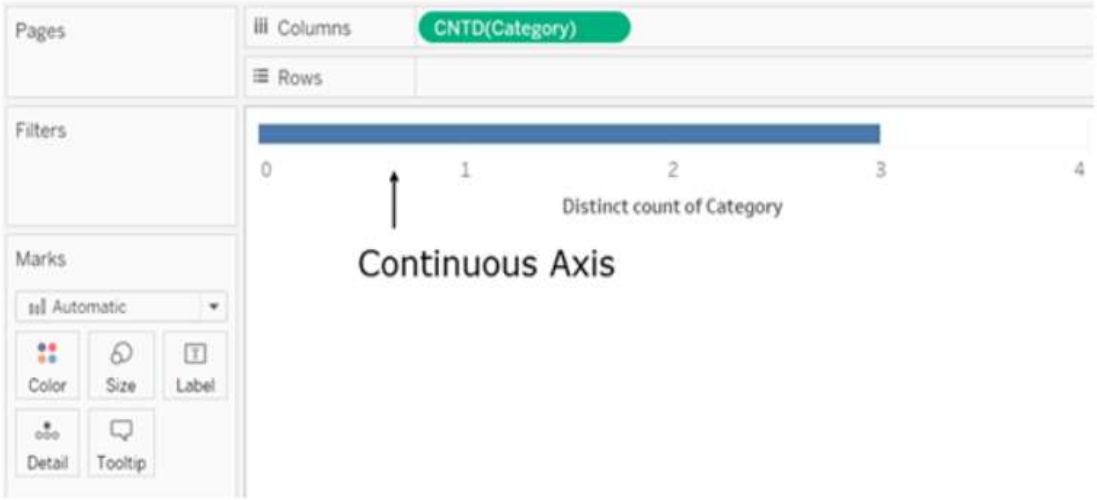
- In both examples, the Sales field is set to Continuous. It creates a vertical axis because it's continuous and on the Rows shelf. If it was on the Columns shelf the axis would be horizontal. The SUM aggregation indicates that it's a measure.
- The absence of an aggregation function in the Quantity field name indicates that it's a dimension.

Dimension Fields in the view

When you drag a discrete dimension field to **Rows** or **Columns**, Tableau creates column or row headers.

Often, fields from the **Dimension** area are discrete when you add them to a view, with a blue background. Date dimensions and numeric dimensions can be discrete or continuous, and all measures can be discrete or continuous.

After you drag a dimension to **Rows** or **Columns**, you can change the field to a measure by clicking the field and choosing **Measure**. Now the view contains a continuous axis instead of column or row headers, and the pill is green:



Date dimensions can be discrete or continuous. Dimensions containing strings or

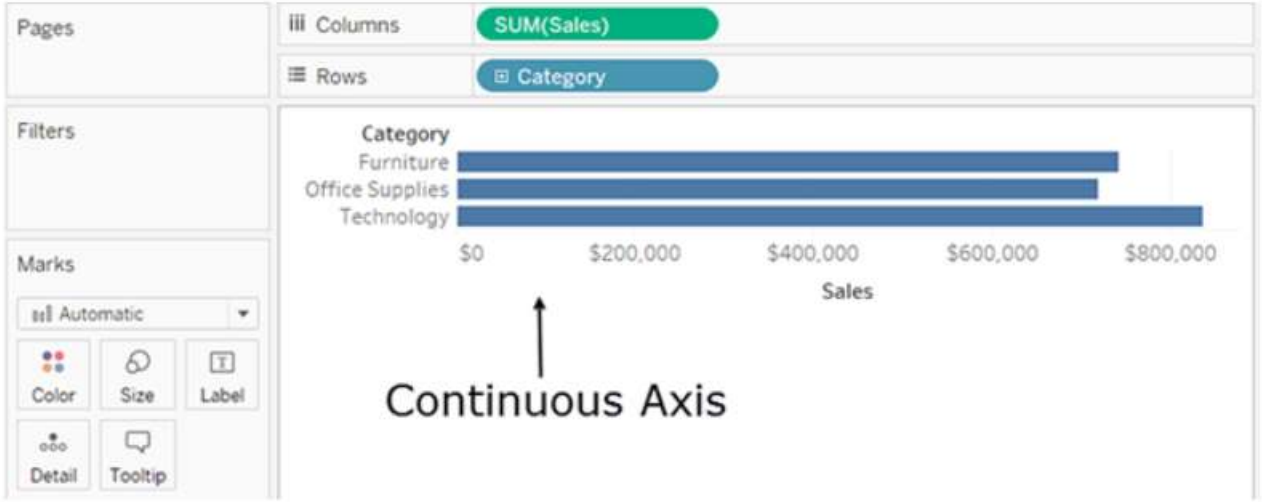
Boolean values can't be continuous.
Tableau doesn't aggregate dimensions.

In Tableau queries, dimensions in the view are expressed in SQL as "Group By" clauses.

List of Predefined Aggregations in Tableau.

Measures Field of the View

- When you drag a measure to the view, it’s aggregated by default.
- The type of aggregation varies depending on the type of view.
- Always check the aggregation and change it if necessary.
- For details, see "Change the default aggregation" in Edit Default Settings for Fields. For more details about aggregation, see Data Aggregation in Tableau.
- When you drag a continuous field from the **Data** pane to **Rows** or **Columns**, Tableau creates a continuous axis for that field.



...contd

- If you click the field and change it to **Discrete**, the values become column headers. ,

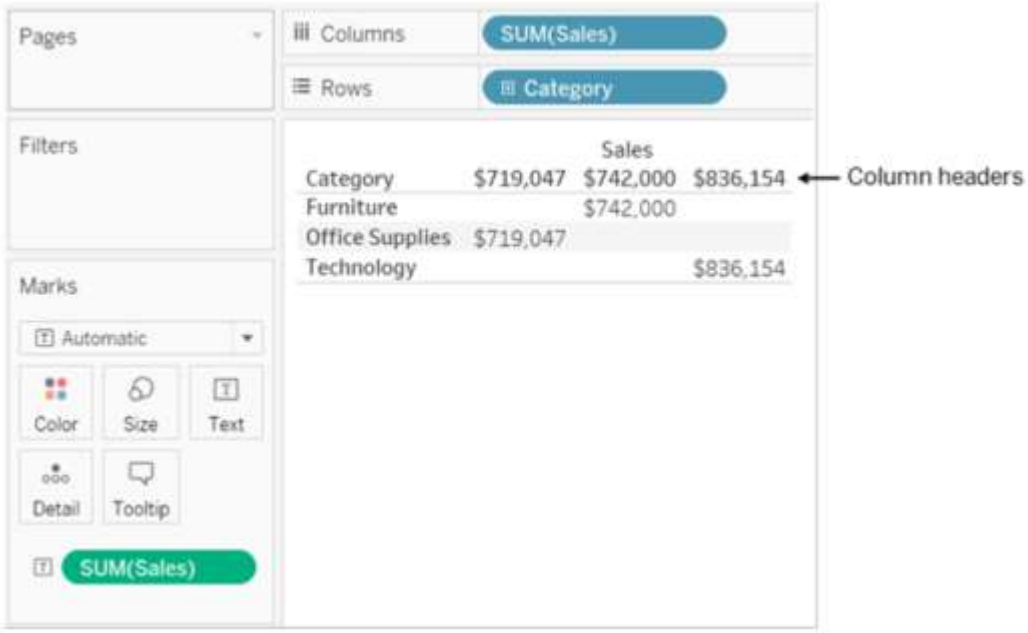
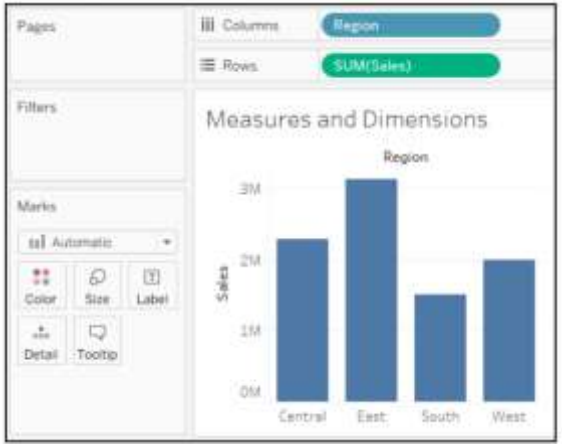


Tableau continues to aggregate values for the field, because even though the field is now discrete, it's still a measure, and Tableau aggregates measures by default.

In cases where Tableau has misclassified a field as a dimension or a measure possibly because of the data type, you can convert it and change its role. If a measure doesn't need to be aggregated (such as numeric IDs), convert the field to a dimension.

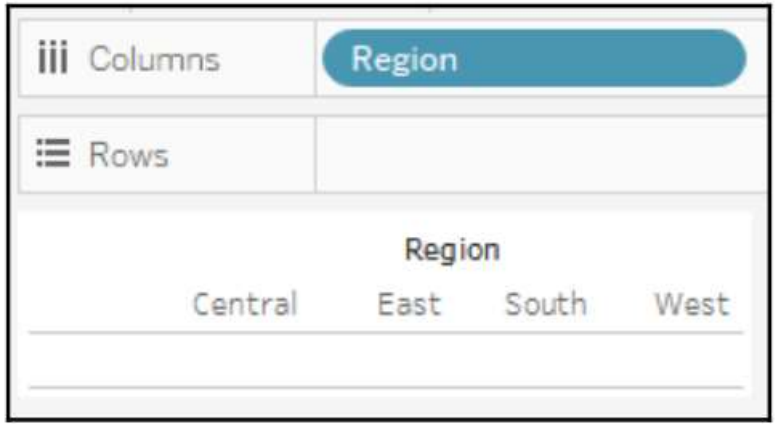
Discrete and continuous fields

- 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.
- Tableau will give a visual indication of the default for a field (the color of the icon in the data pane) and how it is being used in the view (the color of the active field on a shelf).
- Discrete fields, such as the Region in the previous example, are blue. Continuous fields, such as Sales, are green.



Discrete Fields

- Discrete (blue) fields have values that are shown as distinct and separate from one another. Discrete values can be reordered and still make sense. For example, you could easily rearrange the values of Region to be East, South, West, and Central, instead of the default order in the preceding screenshot. When a discrete field is used on the Rows or Columns shelves, the field defines headers. Here, the discrete field Region defines column headers.



Continuous Fields

- Continuous fields Continuous (green) fields have values that flow from first to last as a continuum. Numeric and date fields are often (though not always) used as continuous fields in the view. The values of these fields have an order that it would make little sense to change. When use
- d on Rows or Columns, a continuous field defines an axis:

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

It is very important to note that continuous and discrete are different concepts from Measure and Dimension.
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.

- To change the default of a field, right-click the field in the data pane and select Convert to Discrete or Convert to Continuous.
- To change how a field is used in the view, right-click the field in the view and select Discrete or Continuous.
- Alternatively, you can drag and drop the fields between Dimensions and Measures in the data pane.
- In general, you can think of the differences between the types of fields as follows: Choosing between 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.

Visualising Data

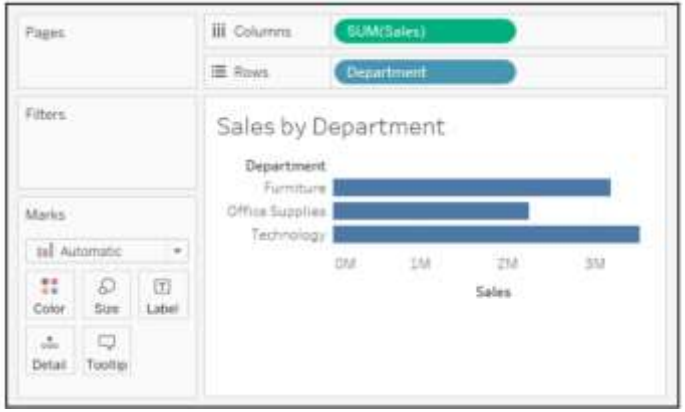
- At times, you may come to the data with very well-defined questions and a strong sense of what you expect to find.
- Other times, you will come to the data with general questions and very little idea of what you will find.
- The visual analytics capabilities of Tableau empower you to rapidly and iteratively explore the data, ask new questions, and make discoveries.
- The following visualization examples cover a few of the most foundational visualization types.
    - Bar charts
    - Line Charts

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.
- Creating bar charts in Tableau is very easy.
- Simply drag and drop the measure you want to see onto either the Rows or Columns shelf and the dimension that defines the categories onto the opposing Rows or Columns shelf.
- As an analyst for Superstore, you are ready to begin a discovery process focused on sales (especially the dollar value of sales).

Example

1. Click on the Sales by Department tab to view that sheet.
2. Drag and drop the Sales field from Measures in the data pane onto the Columns shelf. You now have a bar chart with a single bar representing the sum of sales for all the data in the data source.
3. Drag and drop the Department field from Dimensions in the data pane to the Rows shelf. This slices the data to give you three bars, each having a length that corresponds to the sum of sales for each department



- You now have a horizontal bar chart. This makes comparing the sales between the departments easy.
- The mark type drop-down menu on the Marks card is set to Automatic and shows that Tableau has determined that bars are the best visualization given the fields you have placed in the view.
- As a dimension, the Department slices the data.
- Being discrete, it defines row headers for each department in the data. As a measure, the Sales field is aggregated.
- Being continuous, it represents an axis.
- The mark type of bar causes individual bars for each department to be drawn from 0 to the value of the sum of sales for that department

Iterations of bar charts for deeper analysis

- Using the preceding bar chart, you can easily see that the technology department has more total sales than either the furniture or office supplies departments.
- What if you want to further understand sales amounts for departments across various regions?
- Follow these two steps:
    1. Navigate to the Bar Chart (two levels) sheet, where you will find an initial view 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

Stacked Bar Chart

1. Navigate to the Bar Chart (stacked) sheet, where you will find a view identical to the original bar chart.
2. Drag the Region field
   from the Rows shelf and drop it on to the Color shelf

- Instead of a side-by-side bar chart, you now have a stacked bar chart.
- Each segment of the bar is color-coded by the Region field.
- Additionally, a color legend has been added to the workspace.
- You haven't changed the level of detail in the view, so sales are still summed for every combination of region and department.

3. Navigate to the Bar Chart (experimentation) sheet.

4. Try dragging the Region field from Color to the other various shelves on the Marks card, such as Size, Label, and Detail. Observe that in each case the bars remain stacked but are redrawn based on the visual encoding defined by the Region field

5. Use the Swap button on the Toolbar to swap fields on Rows and Columns. This allows you to very easily change from a horizontal bar chart to a vertical bar chart (and vice versa).

6. Drag and drop Sales from the Measures section of the data pane on top of the Region field on the Marks card to replace it. Drag the Sales field to Color if necessary, and notice how the color legend is a gradient for the continuous field.

7. Experiment further by dragging and dropping other fields onto various shelves. Note the behavior of Tableau for each action you take.

8. From the File menu, select Save

Line Charts

- Line charts connect related marks in a visualization to show movement or 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.
- The most common kind of line chart is a Time Series.
- A time series shows the movement of values over time.
- Creating one in Tableau requires only a date and a measure.

    1. Navigate to the Sales over time sheet.
    2. Drag the Sales field from Measures to Rows. This gives you a single, vertical bar representing the sum of all sales in the data source
    3. To turn this into a time series, you must introduce a date. Drag the Order Date field from Dimensions in the data pane on the left and drop it into Columns. Tableau has a built-in date hierarchy, and the default level of Year has given you a line chart connecting four years. Notice that you can clearly see an increase in sales year after year

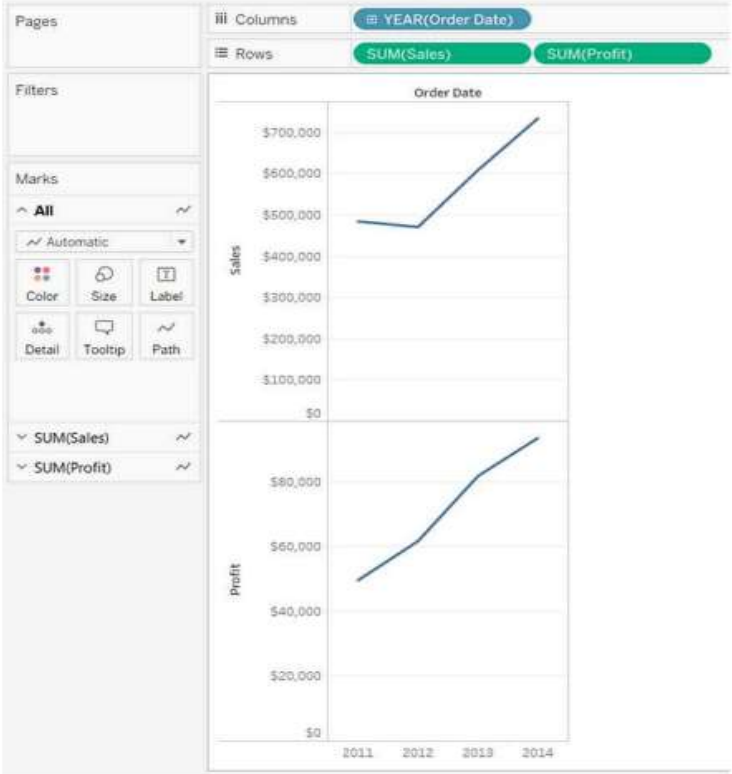
To create a view that displays the sum of sales and the sum of profit for all years, and then uses forecasting to determine a trend, follow these steps:

1. Connect to the **Sample - Superstore** data source.
2. Drag the **Order Date** dimension to **Columns**.

   Tableau aggregates the date by year, and creates column headers.
3. Drag the **Sales** measure to **Rows**.

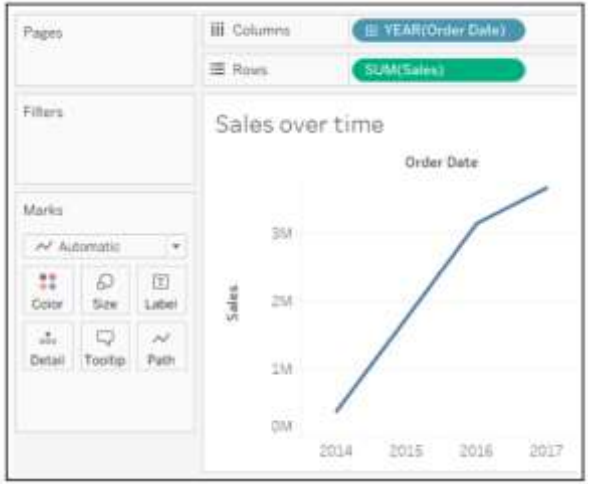
   Tableau aggregates **Sales** as SUM and displays a simple line chart.
4. Drag the **Profit** measure to **Rows** and drop it to the right of the **Sales** measure.

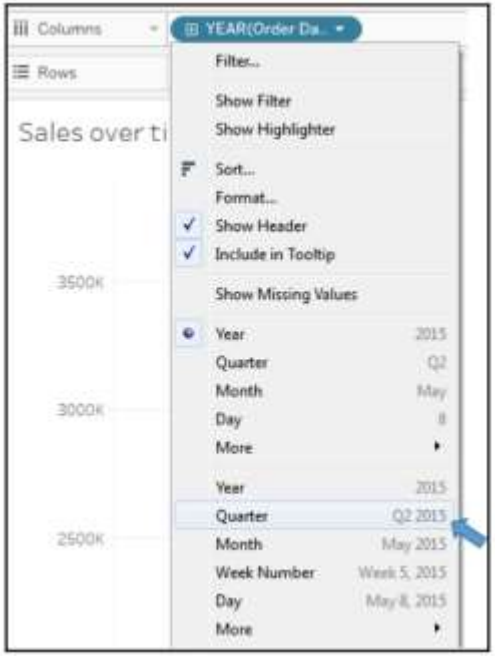
Tableau creates separate axes along the left margin for **Sales** and **Profit**.



Notice that the scale of the two axes is different—the **Sales** axis scales from $0 to $700,000, whereas the **Profit** axis scales from $0 to $100,000. This can make it hard to see that sales values are much greater than profit values.

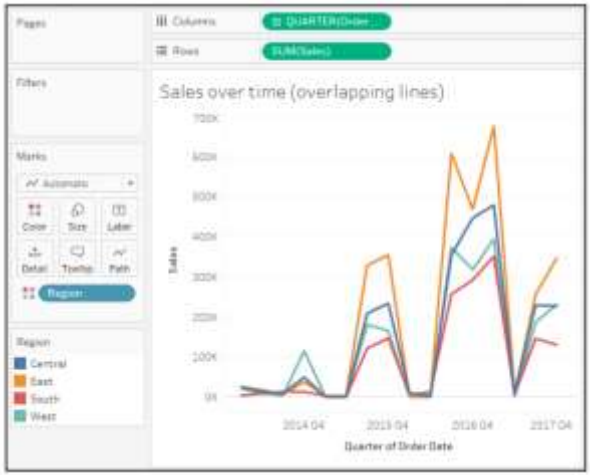
4. Use the drop-down menu on the YEAR(Order Date) field on Columns (or right-click the field) and switch the date field to use Quarter. You may notice that Quarter is listed twice in the drop-down menu.





Iterations of line charts for deeper analysis

- 1. Navigate to the Sales over time (overlapping lines) sheet, where you will find a view identical to the one you just created.
- 2. Drag the Region field from Dimensions to Color. Now you have a line per region, with each line a different color, and a legend indicating which color is used for which region. As with the bars, adding a dimension to color splits the marks. However, unlike the bars, where the segments were stacked, the lines are not stacked. Instead, the lines are drawn at the exact value for the sum of sales for each region and quarter. This allows for easy and accurate comparison. It is interesting to note that the cyclical pattern can be observed for each region



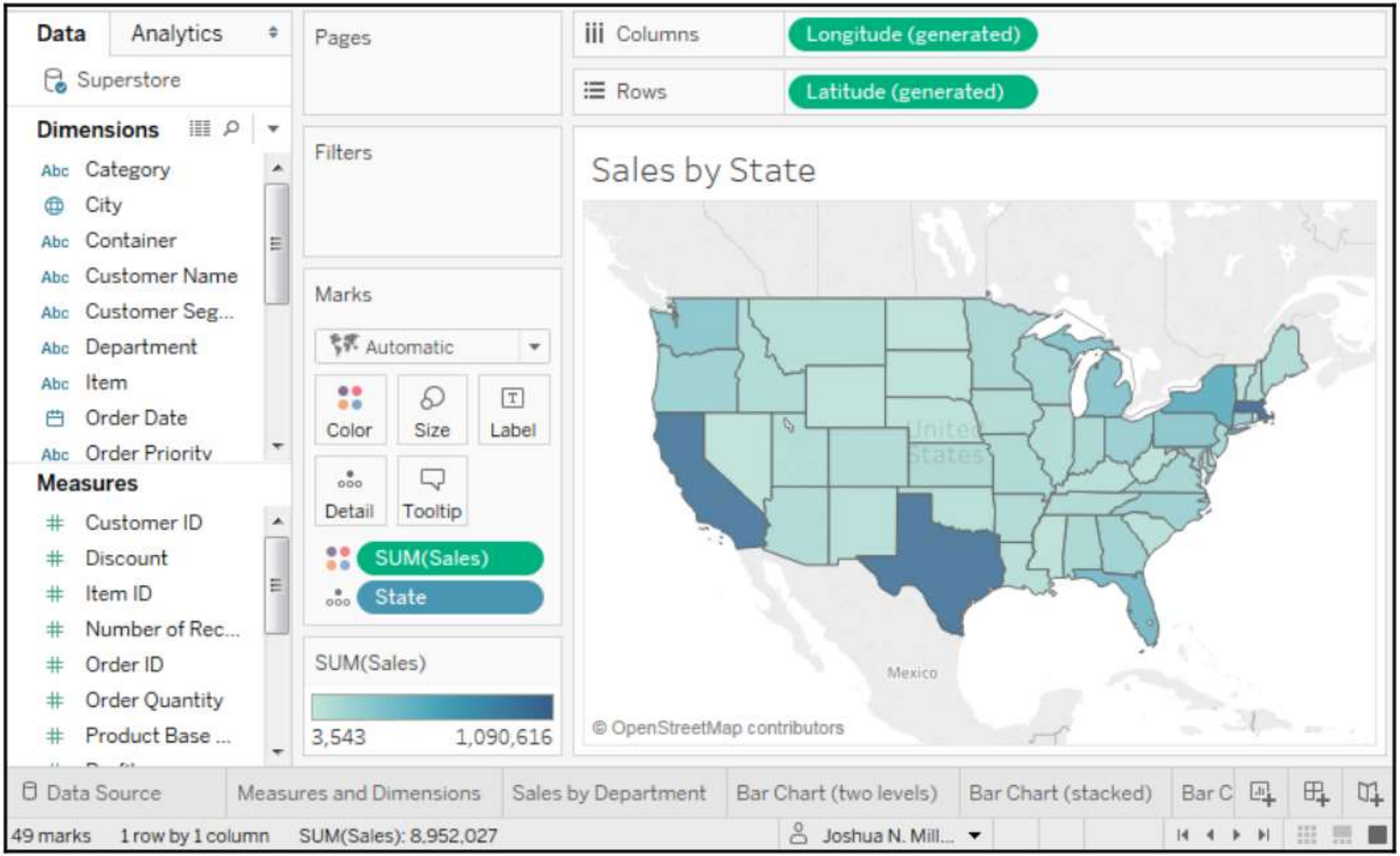
Geographic Visualisations

- 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.
- If your data does contain latitude and longitude fields, you may use those instead of the generated values.
- Tableau will automatically assign geographic roles to some fields based on a field name and a sampling of values in the data.
- You can assign or reassign geographic roles to any field by right-clicking the field in the data pane and using the Geographic Role option.
- This is also a good way to see what built-in geographic roles are available.

- 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 three main types of geographic visualization:
  - Filled maps (simply referred to as maps in the Tableau interface)
  - Symbol maps
  - Density maps

Filled Maps

- Filled maps fill areas such as countries, states, counties, or ZIP codes to show a location.
- The color that fills the area can be used to encode values, most often of aggregated measures but sometimes also dimensions.
- These maps are also called choropleth maps.
- Let's say you want to understand sales for Superstore and see whether there are any patterns geographically.
- You might take an approach similar to the following:
    - 1. Navigate to the Sales by State sheet.
    - 2. Double-click the State field in the data pane. Tableau automatically creates a geographic visualization using the Latitude (generated), Longitude (generated), and State fields.
    - 3. Drag the Sales field from the data pane and drop it on the Color shelf on the Marks card. Based on the fields and shelves you've used, Tableau has switched the automatic mark type to Map

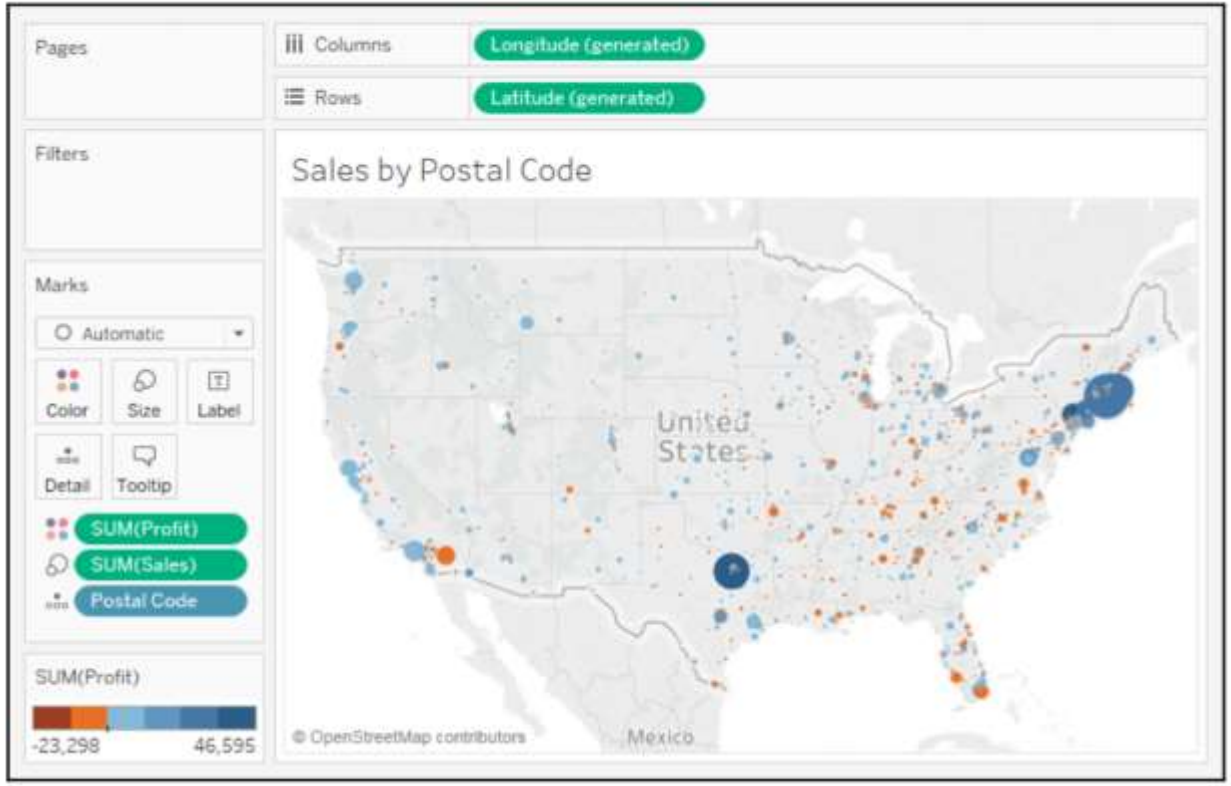

Data
Analytics
Superstore
Dimensions
Abc Category
City
Abc Container
Abc Customer Name
Abc Customer Seg...
Abc Department
Abc Item
Order Date
Abc Order Priority
Measures
Customer ID
Discount
Item ID
Number of Rec...
Order ID
Order Quantity
Product Base ...
Pages
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
SUM(Sales)
State
SUM(Sales)
3,543
1,090,616
Columns
Longitude (generated)
Rows
Latitude (generated)
Sales by State
Mexico
© OpenStreetMap contributors
Data Source
Measures and Dimensions
Sales by Department
Bar Chart (two levels)
Bar Chart (stacked)
Bar C
49 marks
1 row by 1 column
SUM(Sales): 8,952,027
Joshua N. Mill...

- Filled maps can work well in interactive dashboards and have quite a bit of aesthetic value.
- However, certain kinds of analyses are very difficult with filled maps.
- Unlike other visualization types, where size can be used to communicate facets of the data, the size of a filled geographic region only relates to the geographic size and can make comparisons difficult.
- For example: which state has the highest sales?
- You might be tempted to say Texas or California because they appear larger, but would you have guessed Massachusetts?
- Some locations may be small enough that they won't even show up compared to larger areas.
- Use filled maps with caution and consider pairing them with other visualizations on dashboards for clear communication

Symbol Maps

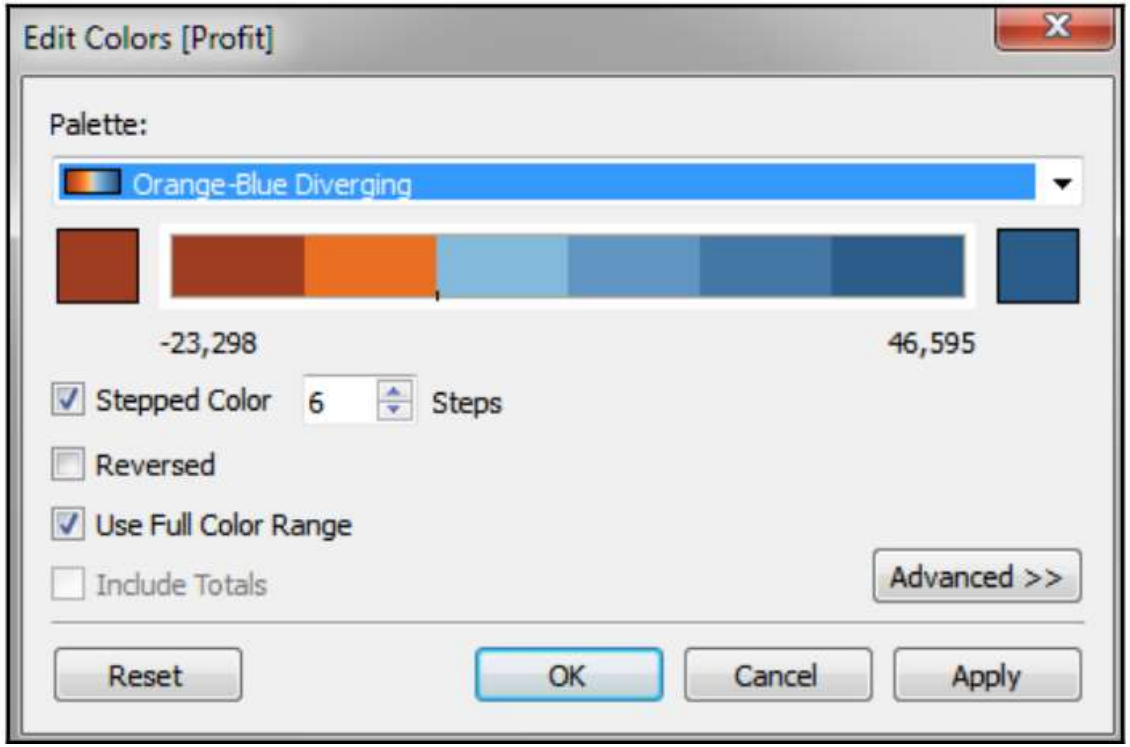
- With symbol maps, marks on the map are not drawn as filled regions; rather, marks are shapes or symbols placed at specific geographic locations.
- The size, color, and shape may also be used to encode additional dimensions and measures. Continue your analysis of Superstore sales by following these steps:
  - Navigate to the Sales by Postal Code sheet.
  - Double-click Postal Code under Dimensions. Tableau automatically adds Postal Code to the Detail of the Marks card and Longitude (generated) and Latitude (generated) to Columns and Rows. The mark type is set to a circle by default, and a single circle is drawn for each postal code at the correct latitude and longitude.
  - Drag Sales from Measures to the Size shelf on the Marks card. This causes each circle to be sized according to the sum of sales for that postal code.
  - Drag Profit from Measures to the Color shelf on the Marks card. This encodes the mark color to correspond to the sum of profit.
- You can now see the geographic location of profit and sales at the same time. This is useful because you will see some locations with high sales and low profit, which may require some action

The final view should look like this, after making some fine-tuned adjustments to the size and color



- Sometimes, you'll want to adjust the marks on symbol map to make them more visible.
- Some options include the following:
    - If the marks are overlapping, click the Color shelf and set the transparency to somewhere between 50% and 75%.
    - Additionally, add a dark border. This makes the marks stand out, and you can often better discern any overlapping marks.
    - marks are too small, click on the Size shelf and adjust the slider.
    - You may also double-click the size legend and edit the details of how Tableau assigns size.
    - If the marks are too faint, double-click the color legend and edit the details of how Tableau assigns color.
    - This is especially useful when you are using a continuous field that defines a color gradient

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



- Unlike filled maps, symbol maps allow you to use size to visually encode aspects of the data. Symbol maps also allow for greater precision.
- In fact, if you have latitude and longitude in your data, you can very precisely plot marks at a street address-level of detail.
- This type of visualization also allows you to map locations that do not have clearly defined boundaries.
- Sometimes, when you manually select Map in the Marks card drop-down menu, you will get an error message indicating that filled maps are not supported at the level of detail in the view.
- In those cases, Tableau is rendering a geographic location that does not have built-in shapes.
- Other than cases where filled maps are not possible, you will need to decide which type best meets your needs.

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 intensity in areas with a high concentration.
- You can control color, size, and intensity.
- Let's say you want to understand the geographic concentration of orders.
- You might create a density map using the following steps:
  - 1. Navigate to the Density of Orders sheet.
  - 2. Double-click the Postal Code field in the data pane. Just as before, Tableau automatically creates a symbol map geographic visualization using the Latitude (generated), Longitude (generated), and State fields.
  - 3. Using the drop-down menu on the Marks card, change the mark type to Density.
- The individual circles now blend together showing concentrations

Pages

Filters

Marks

Density

Color | Size | Label

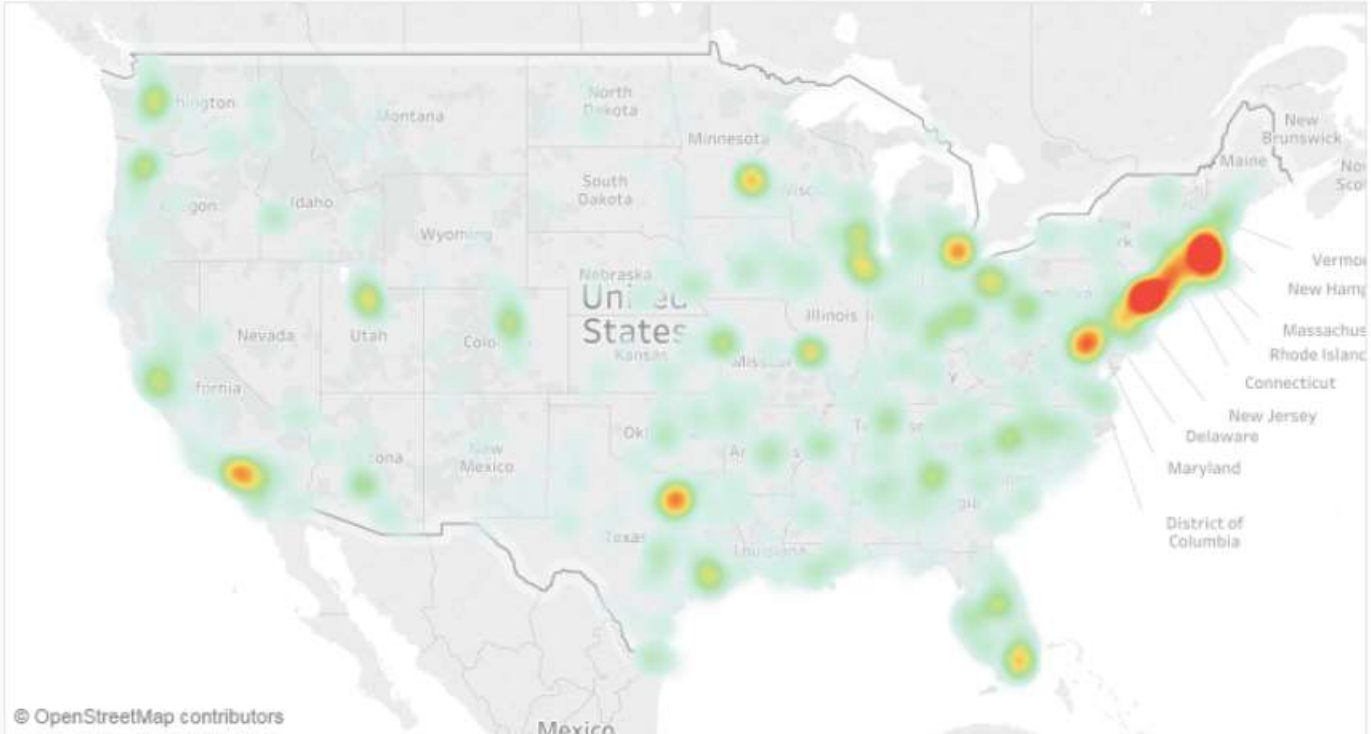
Detail | Tooltip

Postal Code

Columns: Longitude (generated)

Rows: Latitude (generated)

Density of Orders



Working with Data

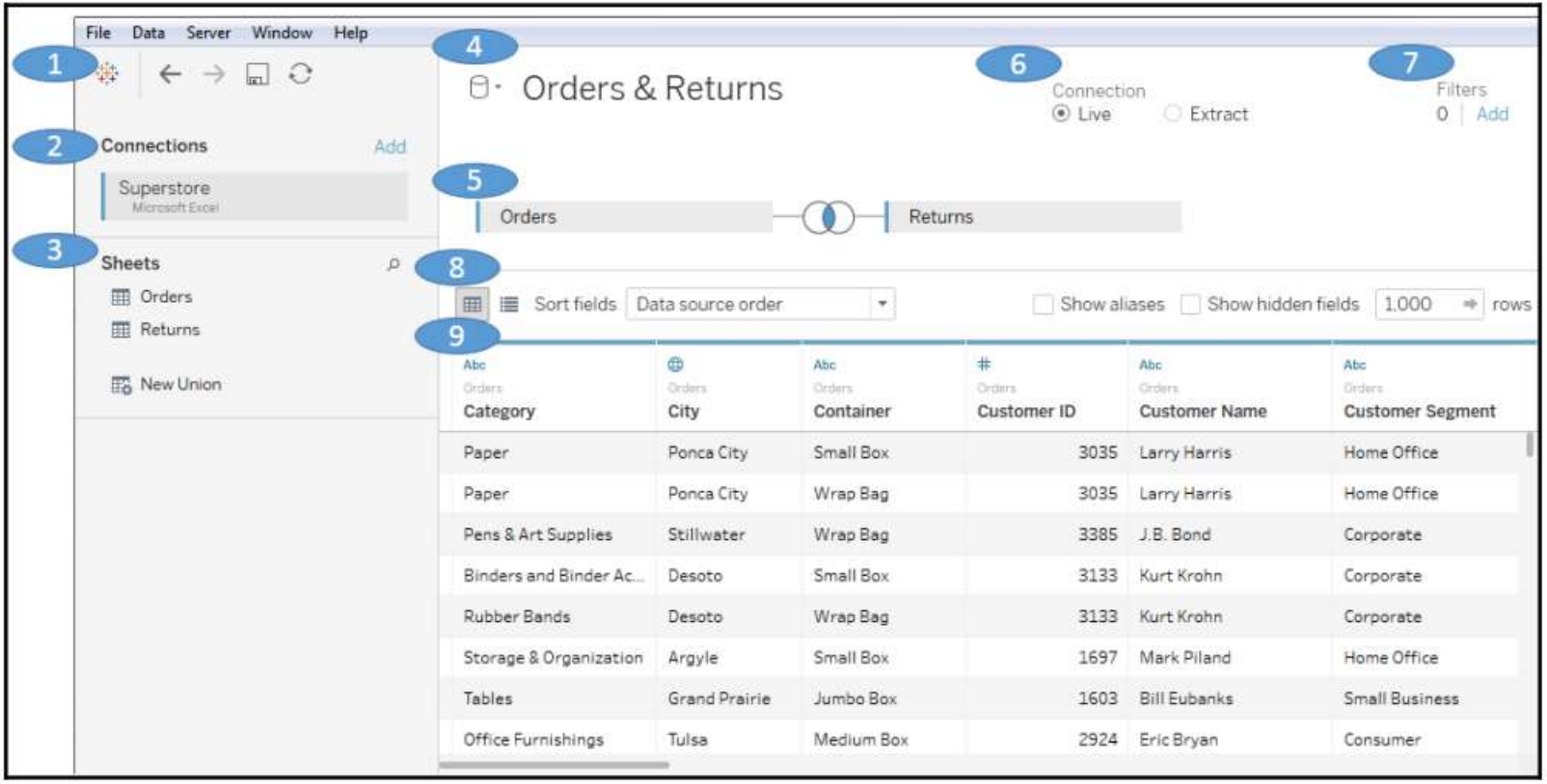
- Connecting to data in a file File-based data includes all sources of data where the data is stored in a file.
- File-based data sources include the following:
- Extracts: A .hyper or .tde file containing data that was extracted from an original source.
- Microsoft Access: An .mdb or .accdb database 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.
- Text file: A delimited text file, most commonly .txt, .csv, or .tab.
- Multiple text files in a single directory may be joined or unioned together in a single connection.
- Local cube file: A .cub file that contains multi-dimensional data. These files are typically exported from OLAP databases.

- Adobe PDF: A .pdf file that may contain tables of data that can be parsed by Tableau.
- Spatial file: A .kml, .shp, .tab, .mif, or .geojson file that contains spatial objects that can be rendered by Tableau.
- Statistical file: An .sav, .sas7bdat, .rda, or .rdata file generated by statistical tools, such as SAS or R.
- JSON file: A .json file that contais data in JSON format.

In addition to those mentioned previously, you can connect to Tableau files to import connections that you have saved in another Tableau workbook (.twb or .twbx).
The connection will be imported and changes will only affect the current workbook

Connection to an Excel file:

1. Navigate to the Connect to Excel sheet in the Chapter 02 Starter.twbx workbook.
2. From the menu, select Data | Create new data source and select Excel from the list of possible connections.
3. In the open dialogue, open the Superstore.xlsx file from the \Learning Tableau\Chapter 02 directory. Tableau will open the Data Source screen. You should see the two sheets of the Excel document listed on the left.
4. Double-click the Orders sheet and then the Returns sheet. Your data source screen should look similar to the following screenshot


File
Data
Server
Window
Help
1
2
Connections
Add
Superstore
Microsoft Excel
3
Sheets
Orders
Returns
New Union
4
Orders & 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
Abc
Orders
Category
Orders
City
Abc
Orders
Container
Orders
Customer ID
Abc
Orders
Customer Name
Abc
Orders
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
Office Furnishings
Tulsa
Medium Box
2924
Eric Bryan
Consumer

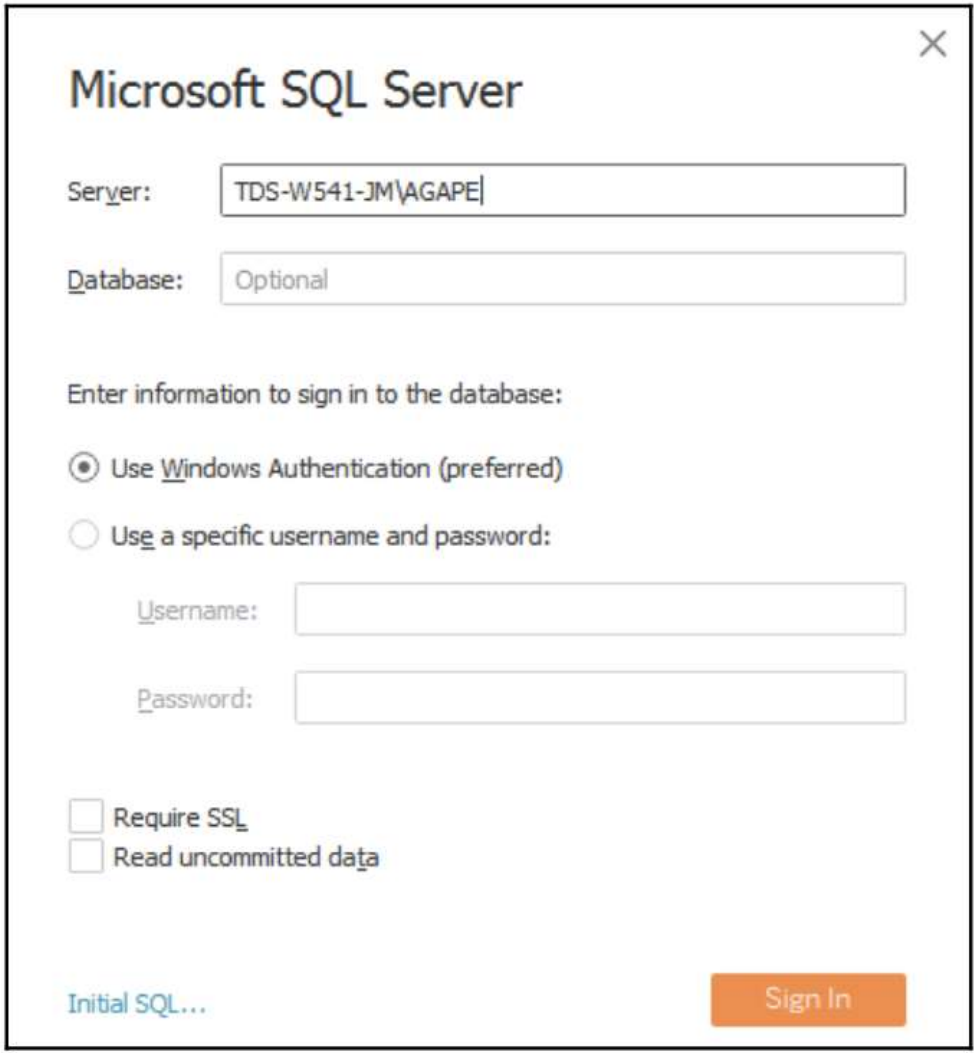
Data Source screen interface, which has the following features (numbered in the preceding screenshot):

- 1. Toolbar: The toolbar has a few of the familiar controls, including undo, redo, and save. It also includes the option to refresh the current data source.
- 2. Connections: All the connections in the current data source. Click Add to add a new connection to the current data source. This allows you to join data across different connection types. Each connection will be color-coded so that you can distinguish what data is coming from which connection.
- 3. Sheets (or Tables): This lists all the tables of data available for a given connection. This includes sheets, sub-tables, and named ranges for Excel; tables, views, and stored procedures for relational databases; and other connection dependent options, such as New Union or Custom SQL.
- 4. Data Source Name: This is the name of the currently selected data source. You may select a different data source using the drop-down arrow next to the database icon. You may click the name of the data source to edit it.
- 5. Connection Editor: Drop sheets and tables from the left into this area to make them part of the connection. For many connections, you may add multiple tables that will be joined or unioned together.

- 6. Live or Extract Options: For many data sources, you may choose whether you would like to have a live connection or an extracted connection.
- 7. Data Source Filters: You may add filters to the data source. These will be applied at the data-source level, and thus to all views of the data using this data source in the workbook.
- 8. Preview Pane Options: These options allow you to specify whether you'd like to see a preview of the data or a list of metadata, and how you would like to preview the data (examples include alias values, hidden fields shown, and how many rows you'd like to preview).
- 9. Preview Pane/Metadata View: Depending on your selection in the options, this space either displays a preview of data or a list of all fields with additional metadata. Notice that these views give you a wide array of options, such as changing data types, hiding or renaming fields, and applying various data transformation functions.

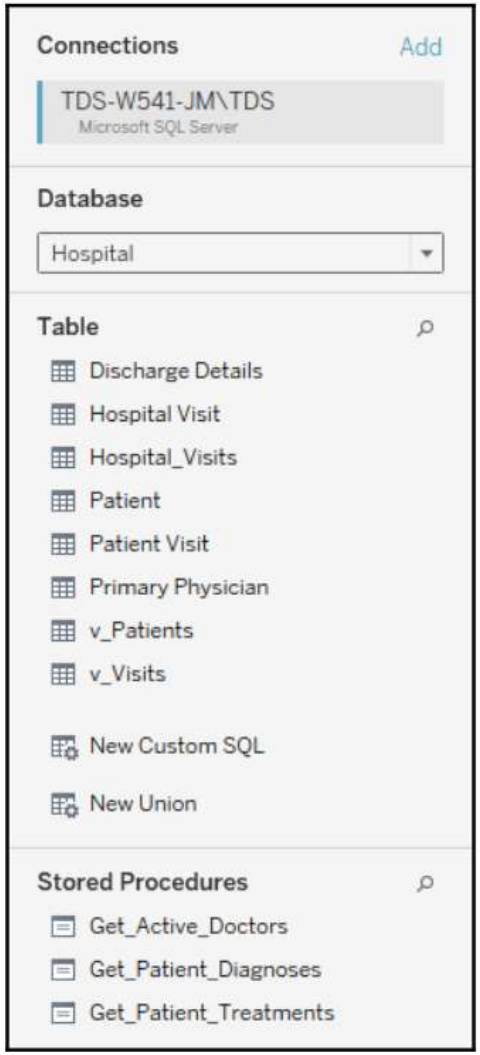
Connecting to data on a server

- Database servers, such as SQL Server, MySQL, 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.
- Alternately, data can be extracted from these sources and stored in an extract (.hyper or .tde).
- As an example of connecting to a server data source, we'll demonstrate connecting to SQL Server.
- If you have access to a server-based data source, you may wish to create a new data source and explore the details.
- As soon as the Microsoft SQL Server connection is selected, the interface displays options for some initial configuration as follows:



- A connection to SQL Server requires the Server name, as well as authentication information
- A database administrator can configure SQL Server to use Windows Authentication or a SQL Server username and password.
- With SQL Server, you can also optionally allow for reading uncommitted data.
- This can potentially improve performance, but may also lead to unpredictable results if data is being inserted, updated, or deleted at the same time Tableau is querying.
- Additionally, you may specify SQL to be run at connect time using the Initial SQL... link in the lower-left corner.

- In order to maintain high standards of security, Tableau will not save a password as part of a data source connection.
- This means that if you share a workbook using a live connection with someone else, they will need to have credentials to access the data.
- This also means that when you first open the workbook, you will need to re-enter your password for any connections requiring a password.
- 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



- Once you've selected a database, you will see the following:
- Table: This shows any data tables or views contained in the selected database.
- New Custom SQL: You may write your own custom SQL scripts and add them as tables. You may join these as you would any other table or view.
- New Union: You may union together tables in the database. Tableau will match fields based on name and data type, and you may additionally merge fields as needed.
- Stored Procedures: You may use a stored procedure that returns a table of data. You will be given the option of setting values for stored procedure parameters, or using or creating a Tableau parameter to pass values.

Once you have finished configuring the connection, click a tab for any sheet to begin visualizing the data.

Connecting Data to the Cloud

- Certain data connections are made to data that is hosted in the cloud.
- These include Amazon Redshift, Google Analytics, Google Sheets, Salesforce, and many others.
- It is beyond the scope of this book to cover each connection in depth, but as an example of a cloud data source, we'll consider connecting to Google Sheets.
- Google Sheets allows users to create and maintain spreadsheets of data online. Sheets may be shared and collaborated on by many different users.
- Here, we'll walk through an example of connecting to a sheet that is shared via link.

To follow the example, you'll need a free Google account. With your credentials, follow these steps:

1. Click the **Add new data source** button on the toolbar, as shown here:

2. Select **Google Sheets** from the list of possible data sources. You may use the search box to quickly narrow the list.
3. On the next screens, sign into your Google Account and allow Tableau Desktop the appropriate permissions. You will then be presented with a list of all your Google Sheets, along with preview and search capabilities, as shown in the following screenshot:

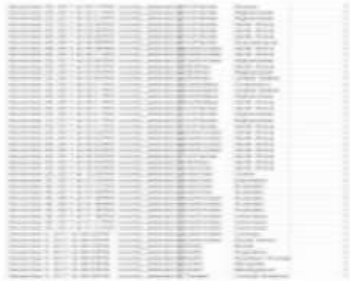
Select Your Google Sheet

Signed in as milligan. Sign Out

Search by sheet name or enter URL

Search

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

Cancel

Connect

4. Enter this URL
5. Select the resulting Superstore sheet in the list and then click the Connect button. You should now see the Data Source screen.
6. Rename the data source to Superstore (Google Sheets).
7. 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.
8. Navigate to the Connect to Google Sheets sheet. The data should be extracted within a few seconds.
9. Create a filled map of Profit by State, with Profit defining the Color and the Label.

Managing data source metadata

- Data sources in Tableau store information about the connection(s).
- In addition to the connection itself (example, database server name, database, and/or file names), 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.
- Right-clicking a field in the data pane reveals a menu of metadata options.
- These are some of the options available via right-clicking:
    - Renaming the field Hiding the field
    - Changing aliases for values of dimension (other than date fields)
    - Creating calculated fields, groups, sets, bins, or parameters Splitting the field

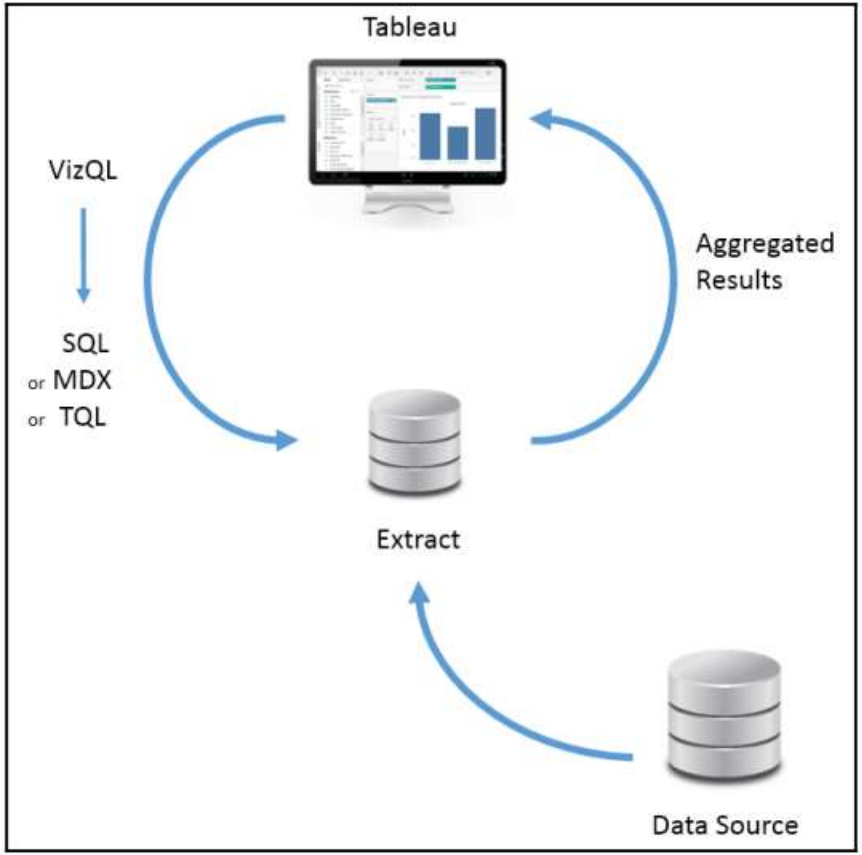
- Changing the default use of a date or numeric field to either discrete or continuous
- Redefining the field as a dimension or a measure
- Changing the data type of the field Assigning a geographic role of the field
- Changing defaults for how a field is displayed in a visualization, such as the default colors and shapes, number or date format, sort order (for dimensions), or type of aggregation (for measures)
- Adding a default comment for a field (which will be shown as a tooltip when hovering over a field in the data pane, or shown as part of the description when Describe... is selected from the menu).
- Adding or removing the field from a hierarchy
- Metadata options that relate to the visual display of the field, such as default sort order or default number format, define the overall default for a field. However, you can override the defaults in any individual view by right-clicking the active field on the shelf and selecting the desired options.

- To see how this works, use the filled map view of Profit by State that you created in the Connect to Google Sheets view.
- If you did not create this view, you may use the Orders and Returns data source, though the resulting view will be slightly different.
- With the filled map in front of you, follow these steps:
  - Right-click the Profit field in the data pane and select Default Properties | Number Format.... The resulting dialog gives you many options for numeric format.
  - Set the number format to Currency (Custom) with 0 Decimal places. After clicking OK, you should notice that the labels on the map have updated to include currency notation.
  - Right-click the Profit field again and select Default properties | Color....
  - The resulting dialog gives you an option to select and customize the default color encoding of the Profit field. Experiment with various palettes and settings.
  - Notice that every time you click the Apply button, the visualization updates.

Working with extracts instead of live connections

- Most data sources allow the option of either connecting live or extracting the data. However, some cloud-based data sources require an extract.
- Conversely, OLAP data sources cannot be extracted and require live connections.
- When using a live connection, Tableau issues queries directly to the data source (or uses data in the cache, if possible).
- When you extract the data, Tableau pulls some or all of the data from the original source and stores it in an extract file.
- Prior to version 10.5, Tableau used a Tableau Data Extract (.tde) file. Starting with version 10.5, Tableau uses Hyper extracts (.hyper) and will convert .tde files to .hyper as you update older workbooks.

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

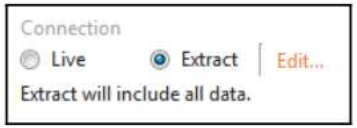


- The fundamental paradigm of how Tableau works with data does not change, but you'll notice that Tableau is now querying and getting results from the extract.
- Data can be retrieved from the source again to refresh the extract.
- Thus, each extract is a snapshot of the data source at the time of the latest refresh.
- Extracts offer the benefit of being portable and extremely efficient.

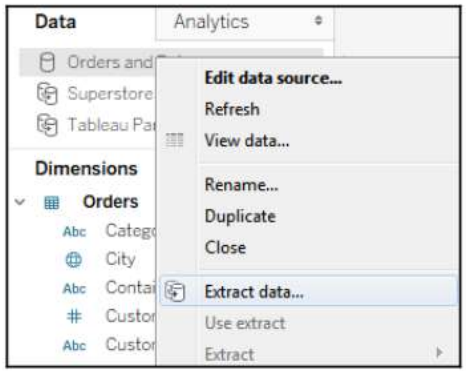
Creating extracts

Extracts can be created in multiple ways, as follows:

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

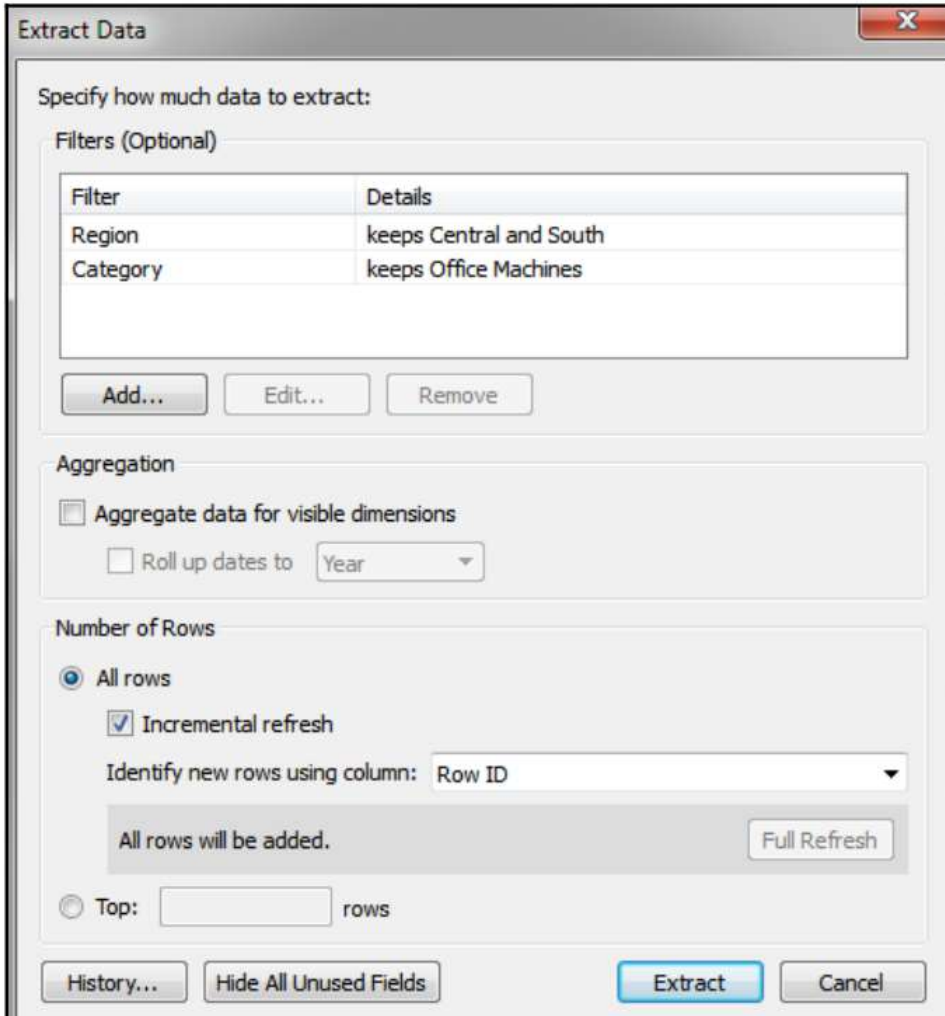


- Select the data source from the **Data** menu, or right-click the data source on the data pane and select **Extract data...**. You will be given a chance to set configuration options for the extract, as demonstrated in the following screenshot:



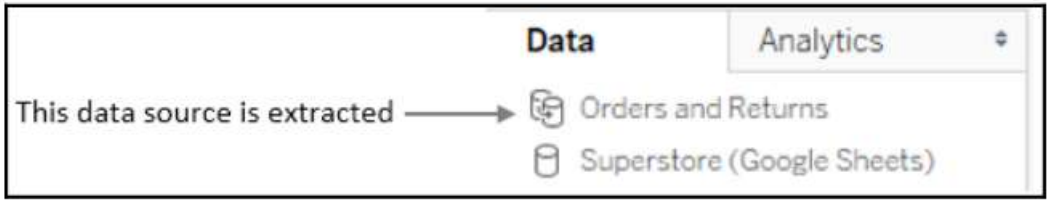
- Developers may create an extract using the Tableau Data Extract API. This API allows you to use Python or C/C++ to programmatically create an extract file. The details of this approach are beyond the scope of this book, but documentation is readily available on Tableau's website.
- Certain tools, such as Alteryx or Tableau Prep, are able to output Tableau extracts.

- When you first create or subsequently configure an extract, you will be prompted to select certain options., as shown here

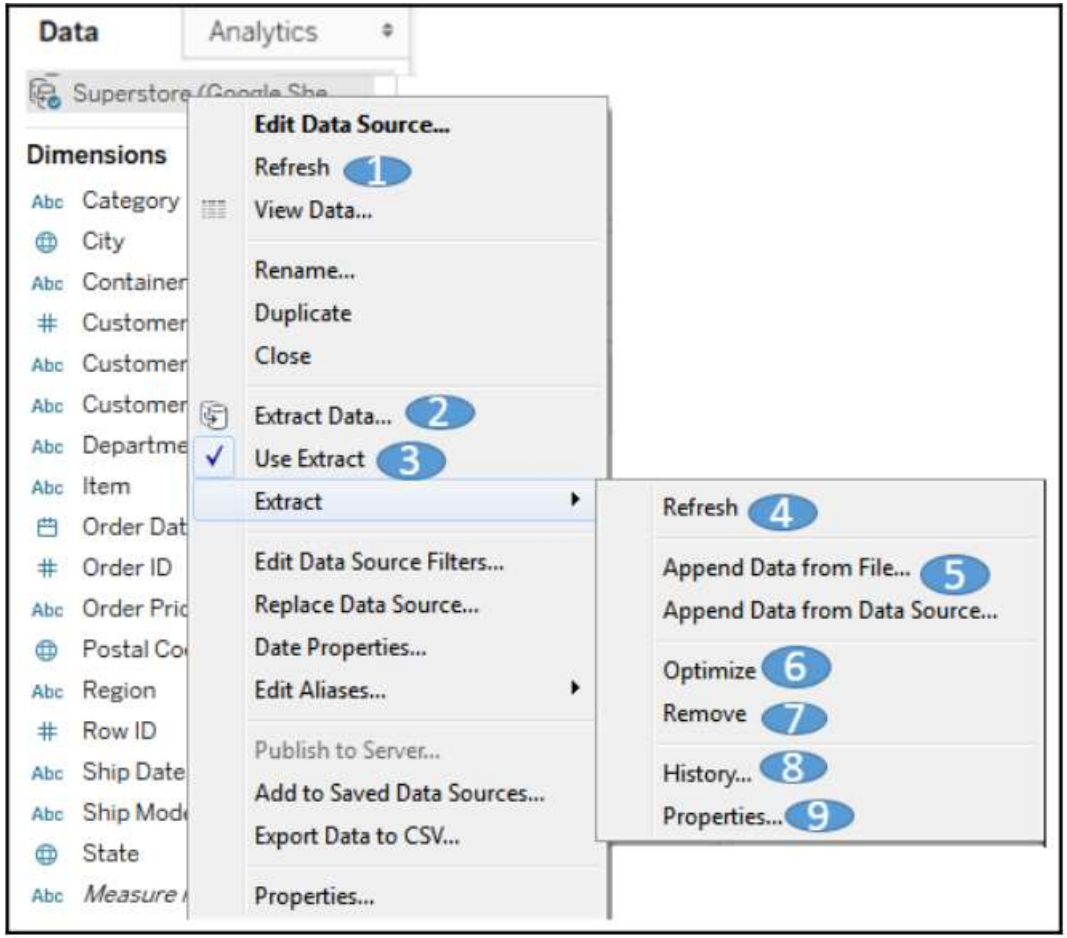


Using Extracts

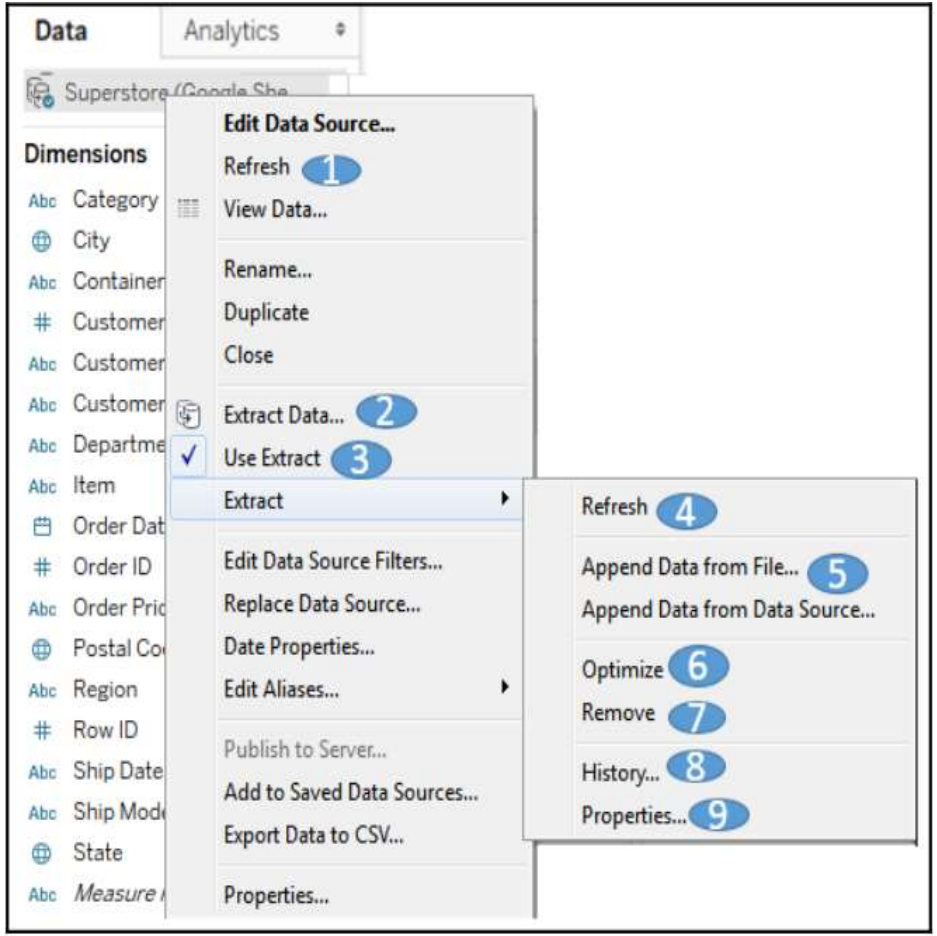
- Any data source that is using an extract will have a distinctive icon that indicates the data has been pulled from a source into an extract, as shown in the following screenshot:



The first data connection in the preceding data pane is extracted, while the second is not. After an extract has been created, you may choose to use the extract or not. When you right-click a data source (or Data from the menu and then the data source), you will see the following menu options, as demonstrated in this screenshot



1. Refresh: The Refresh option under the data source simply tells Tableau to refresh the local cache of data. With a live data source, this would require the underlying data. With an extracted source, the cache is cleared and the extract is required, but this Refresh option does not update the extract from the source.
2. Extract data...: This creates a new extract from the data source (replacing an existing extract if it exists).
3. Use Extract: This option is enabled if there is an extract for a given data source. Unchecking the option will tell Tableau to use a live connection instead of the extract. The extract will not be removed and may be used again by checking this option at any time. If the original data source is not available in this workbook, then Tableau will ask where to find it
4. Refresh: This Refresh option refreshes the extract with data from the source. It does not optimize the extract for some changes you make (such as hiding fields or creating new calculations).



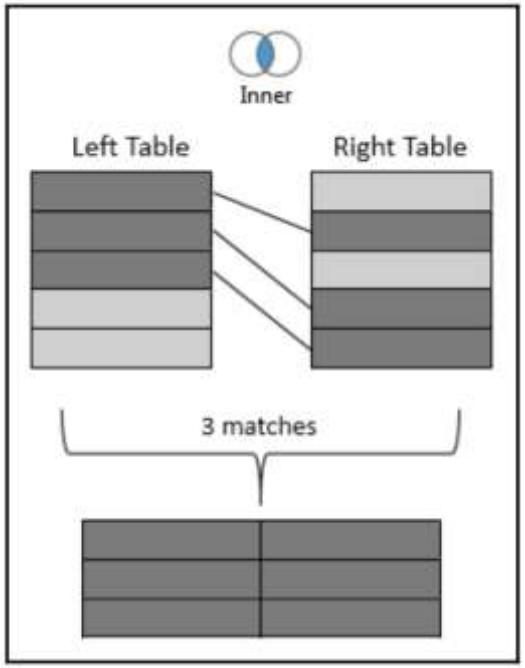
5. Append data from file...: This option allows you to append additional files to an existing extract, provided they have the same data structure as the source. This adds rows to your existing extract; it will not add new columns.
6. Optimize: This will restructure the extract, based on changes you've made since originally creating the extract, to make it as efficient as possible. For example, certain calculated fields may be materialized (that is, calculated once so that the resulting value can be stored), and newly hidden columns or deleted calculations will be removed from the extract.
7. Remove: This removes the definition of the extract, optionally deletes the extract file, and resumes a live connection to the original data source.
8. History: This allows you to view the history of the extract and refresh it.
9. Properties: This enables you to view the properties of the extract, such as the location, underlying source, filters, and row limits.

Joins

- Joins are performed to link tables.
- Types of Joins :
  - Inner Join
  - Left Join
  - Right Join
  - Full Outer Join

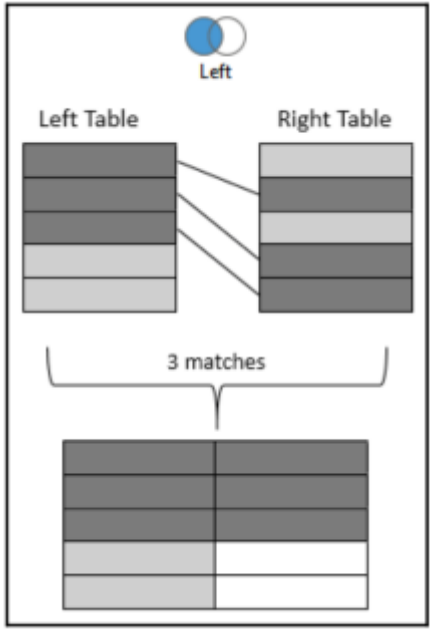
Inner Join

- 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 the three matching rows are kept in the results.



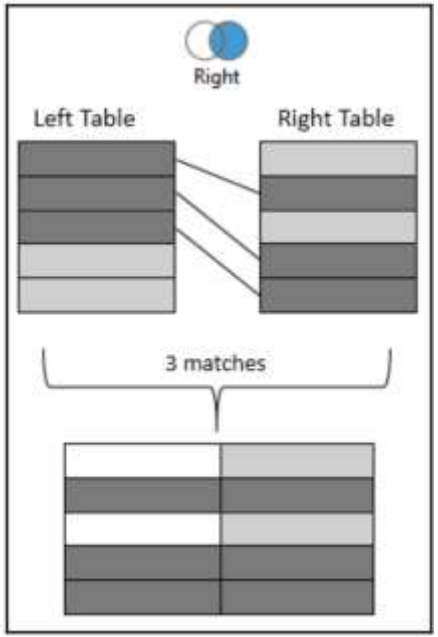
Left Join

- All records from the table on the left will be kept. Matching records from the table on the right will have values in the resulting table, while unmatched records will contain NULL values for all fields from the table on the right. In the following example, the five rows from the left table are kept with NULL results for right values that were not matched:



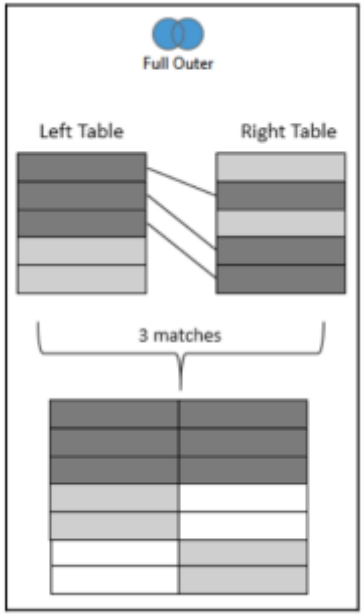
Right Join

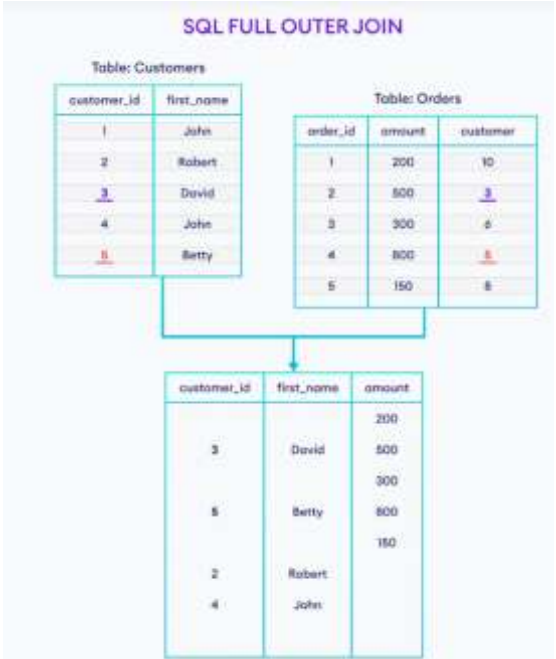
- All records from the table on the right will kept. Matching records from the table on the left will result in values, while unmatched records will contain NULL values for all fields from the table on the left. Not every data source supports a right join. If it is not supported, the option will be disabled. In the following example, the five rows from the right table are kept with NULL results for left values that were not matched:



Full Outer Join

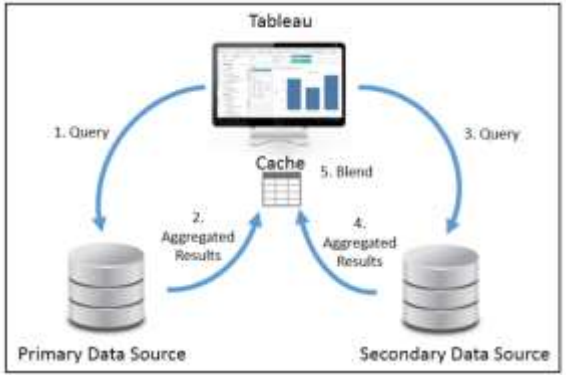
- All records from tables on both sides will be kept. Matching records will have values from the left and the right. Unmatched records will have NULL values where either the left or the right matching record was not found. Not every data source supports a full outer join. If it is not supported, the option will be disabled. In the following example, all rows are kept from both sides with NULL values where matches were not found:





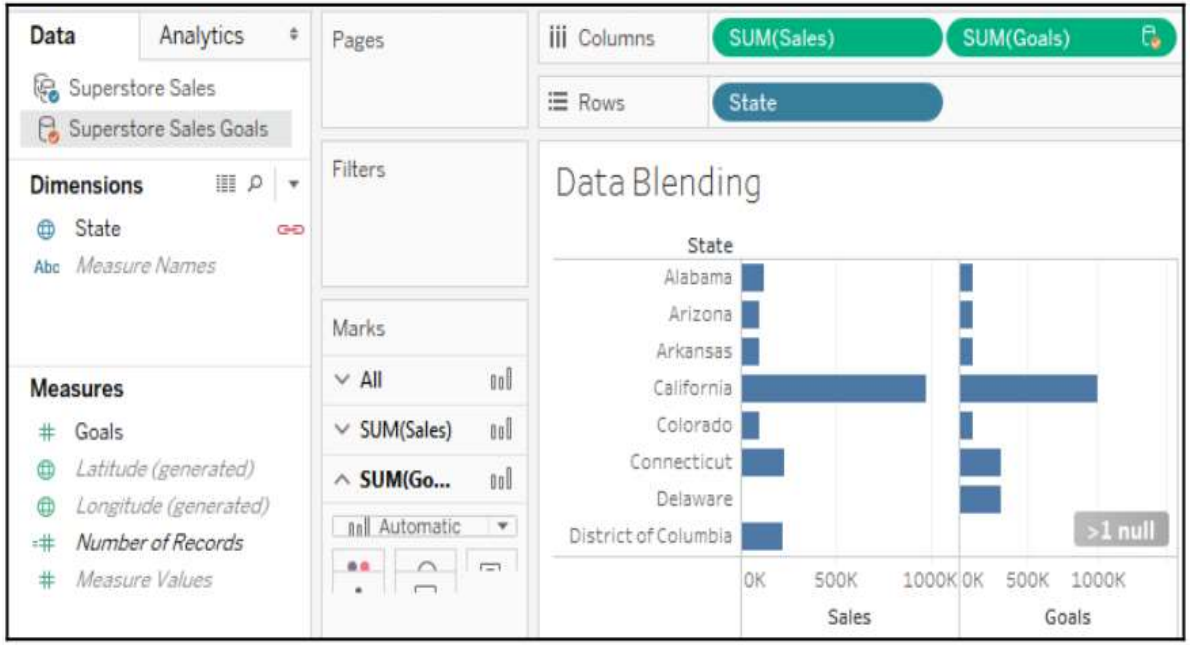
Blending

- Data blending is a powerful and innovative feature in Tableau.
- It 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.
- Data blending is done at an aggregate level and involves different queries sent to each data source, unlike joining, which is done at a row level and involves a single query to a single data source.



- Tableau issues a query to the primary data source.
- The underlying data engine returns aggregate results.
- Tableau issues another query to the secondary data source. This query is filtered based on the set of values returned from the primary data source for dimensions that link the two data sources.
- The underlying data engine returns aggregate results from the secondary data source.
- The aggregated results from the primary data source and the aggregated results from the secondary data source are blended together in the cache.

The following screenshot shows a simple example of data blending in action:



- There are two data source connections defined in this workbook, one for the Superstore data and the other for Superstore Sales Goals.
- The Superstore data source is the primary data source in this view (indicated by the blue checkmark) and Superstore Sales Goals is the secondary source (indicated by the orange checkmark).
- Active fields in the view that are from the secondary data source are also indicated with an orange checkmark icon.
- The Sales measure has been used from the primary source and the Goals from the secondary sources.
- In both cases, the value is aggregated. The State dimension is an active linking field, indicated by the orange chain link icon next to the field in the data pane.
- Both measures are aggregated at the level of State (sales by state in Superstore Sales, and goals by state in Superstore Sales Goals) and then matched by Tableau based on the value of the linking field State.

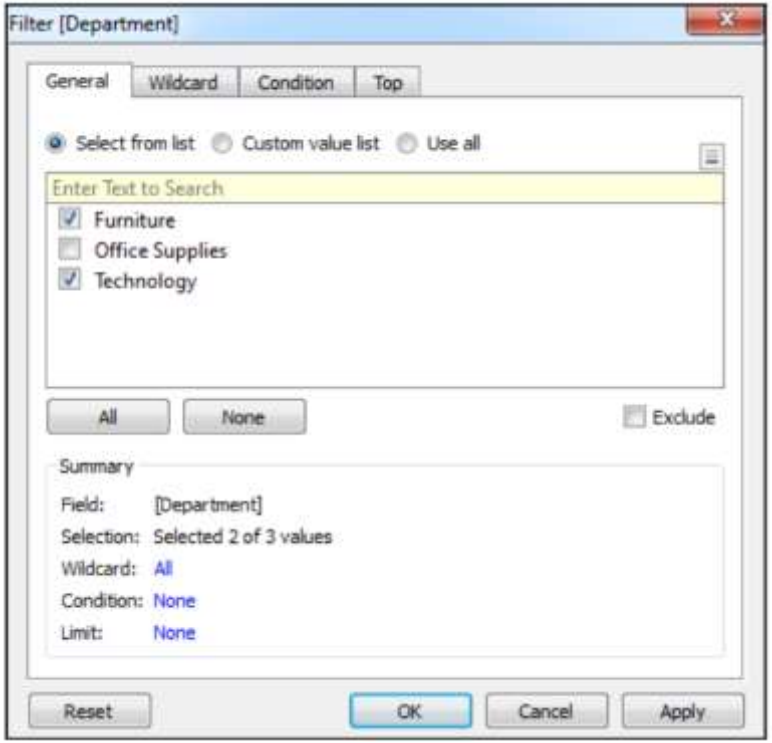
Filtering Data

- Often, you will want to filter data in Tableau to perform an analysis on a subset of data, narrow your focus, or drill into detail. Tableau offers multiple ways to filter data.
- Data Filtering Techniques
  - 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.
  - Custom SQL Filters can be accomplished using a live connection with custom SQL, which has a Tableau parameter in the WHERE clause.

- 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:
- 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:

- Right-click any field in the data pane or the view, and select Show Filter.
- The filter will be shown as a control (examples include a drop-down list and checkbox) to allow the end user of the view or dashboard the ability to change the filter.
- Use an action filter.

Filtering Discrete(Blue) Fields

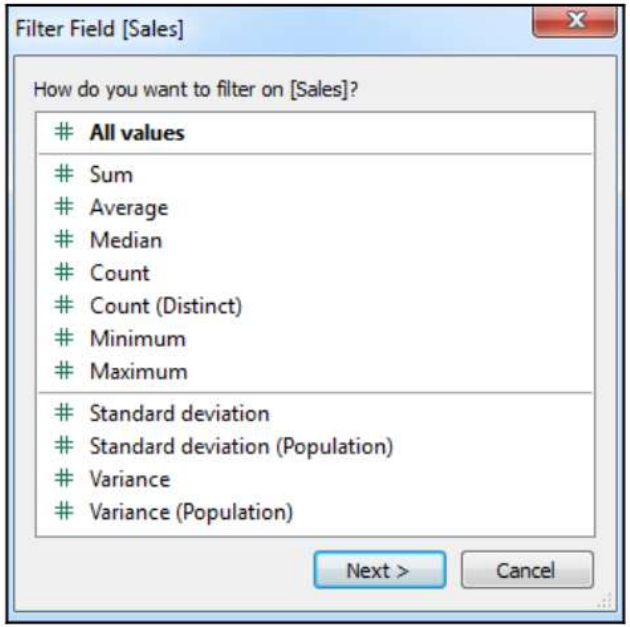


- When you filter using a discrete field, you will be given options for selecting individual values to keep or exclude.
- For example, when you drop the discrete Department dimension onto the Filters shelf, Tableau will give you the following options:

- The Filter options include General, Wildcard, Condition, and Top tabs. Your Filter can include options from each tab.
- The Summary section on the General tab will show all options selected:
    - The General tab allows you to select items from a list (you can use the custom list to add items manually if the dimension contains a large number of values that take a long time to load.)
    - You may use the Exclude option to exclude the selected items.
    - The Wildcard tab allows you to match string values that contain, start with, end with, or exactly match a given value
    - The Condition tab allows you to specify conditions based on aggregations of other fields that meet conditions (for example, a condition to keep any Department where the sum of sales was greater than $1,000,000). Additionally, you can write a custom calculation to form complex conditions
    - The Top tab allows you to limit the filter to only the top or bottom items. For example, you might decide to keep only the top five items by the sum of sales.
- Discrete measures (except for calculated fields using table calculations) cannot be added to the Filters shelf.
- If the field holds a date or numeric value, you can convert it to a continuous field before filtering.
- Other data types will require the creation of a calculated field to convert the values you wish to filter into continuous numeric values.

Filtering Continuous Fields

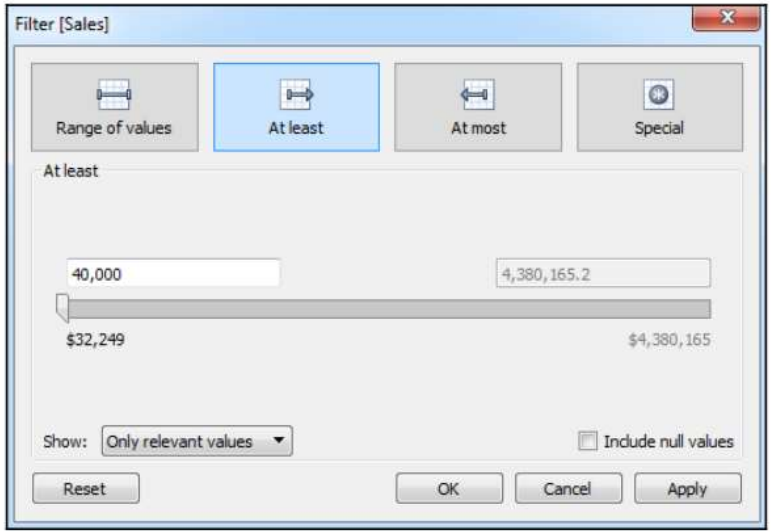
If you drop a continuous dimension onto the **Filters** shelf, you'll get a different set of options. Often, you will first be prompted as to how you want to filter the field, as follows:



The options here are divided into two major categories:

- **All Values**: The filter will be based on each individual value of the field. For example, an **All Values** filter keeping only sales above $100 will evaluate each record of underlying data and keep only individual sales above $100.
- **Aggregation**: The filter will be based on the aggregation specified (for example, **Sum**, **Average**, **Minimum**, **Maximum**, **Standard deviation**, and **Variance**) and the aggregation will be performed at the level of detail of the view. For example, a filter keeping only the sum of sales above $100,000 on a view at the level of category will keep only categories that had at least $100,000 in total sales.

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:



Here, you'll see options for filtering continuous values based on a range with a start, end, or both. The **Special** tab gives options for showing all values, `NULL` values, or non-NULL values.

Filtering Dates

We'll take a look at the special way Tableau handles dates in the *Visualizing Dates and Times* section of `Chapter 3`, *Venturing on to Advanced Visualizations*. For now, consider the options available when you drop an `Order Date` field onto the **Filters** shelf, as follows:

The options here include the following:

- **Relative date**: This option allows you to filter a date based on a specific date (for example, keeping the last three weeks from today, or the last six months from January 1).
- **Range of dates**: This option allows you to filter a date based on a range with a starting date, ending date, or both.
- **Date Part**: This option allows you to filter based on discrete parts of dates, such as **Years**, **Months**, **Days** or combinations of parts, such as **Month/Year.**
- **Individual dates**: This option allows you to filter based on each individual value of the date field in the data.
- **Count** or **Count (Distinct)**: This option allows you to filter based on the count, or distinct count, of date values in the data.