APOGEE INTEGRATION

tableau

Tableau Essentials I

APOGEE TRAINING SERIES

Your Training Team

Domenick Albanese
dalbanese@apogeeintegration.com

Jordan Wilson
jwilson@apogeeintegration.com

Kathleen Callaghan
kcallaghan@apogeeintegration.com

To learn more about Apogee's Data Science & Analysis offerings, check out our website!

Housekeeping

1. Please review & acknowledge APOGEE IT Policy and Sign-In
2. Asking a question
3. Restroom locations and access

Course Outline

Topics Covered

- Tableau User Interface
- Connecting to Data
  - Introduction to Joins
  - Introduction to Unions
- Data Parts & Concepts
  - Discrete and Continuous
  - Dimensions and Measures
  - Dates
  - Aggregation
- Intro to Visual Analytics
  - Easy Calculation
  - Easy Groups
  - Easy Hierarchies
  - Easy Sort
  - Formatting
  - Mark labels and Annotations
- Intermediate Visual Analytics
  - Sort
  - Filter
  - Sets
  - Pages
- Basic Mapping
  - Point Maps
  - Filled Maps
  - Path Maps
  - Heat Maps
- Basic Graphic Standards
  - Color
  - Data-ink Ratio
  - Preattentive Properties
- Intro to Dashboards & Stories

Welcome to tableau

- What is Tableau, Data Science & BI?
- Key Features
- Why Tableau?
- User Interface

What is Tableau, Data Science & BI?

- **Tableau** is a data visualization software that is used for data science and business intelligence
- **Data science** is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from data in various forms, both structured and unstructured
- **Business intelligence (BI)** is a technology-driven process for analyzing data and presenting actionable information to help executives, managers and other stake-holders make informed business decisions

Tableau is….

- a high quality, interactive data visualization tool
- a self-service analytics tool
- read only
- scalable
- a data fusion tool
- collaborative

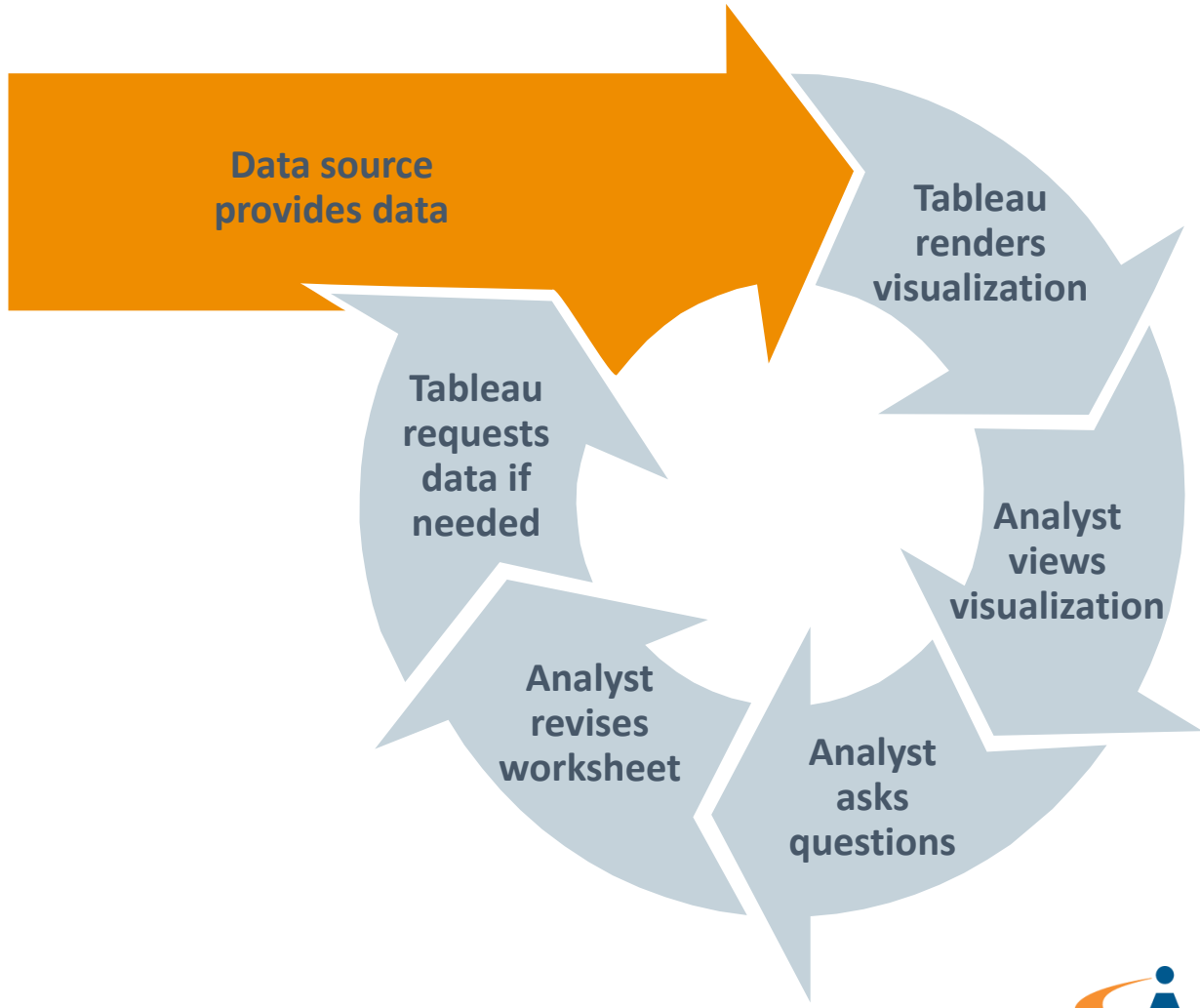
Tableau is not….

- a data editor
- a data cleanup/ETL tool
- a network graphing tool
- a database table creator
- a word processor
- a drawing package

Key Features

- Does not copy data
  - Queries the original data source
  - Sends computation to the data
  - Retrieves only the viz data needed to render
  - Never modifies your data
- Can access many types of data sources
  - To a file (Excel, Text, CSV, etc.)
  - To a server (Tableau, Microsoft SQL, etc.)
- Scales to large sizes
- Can query multiple sources
- Summarizes results for rapid analysis
- Publishes the visualization once
  - If live, it is always up to date

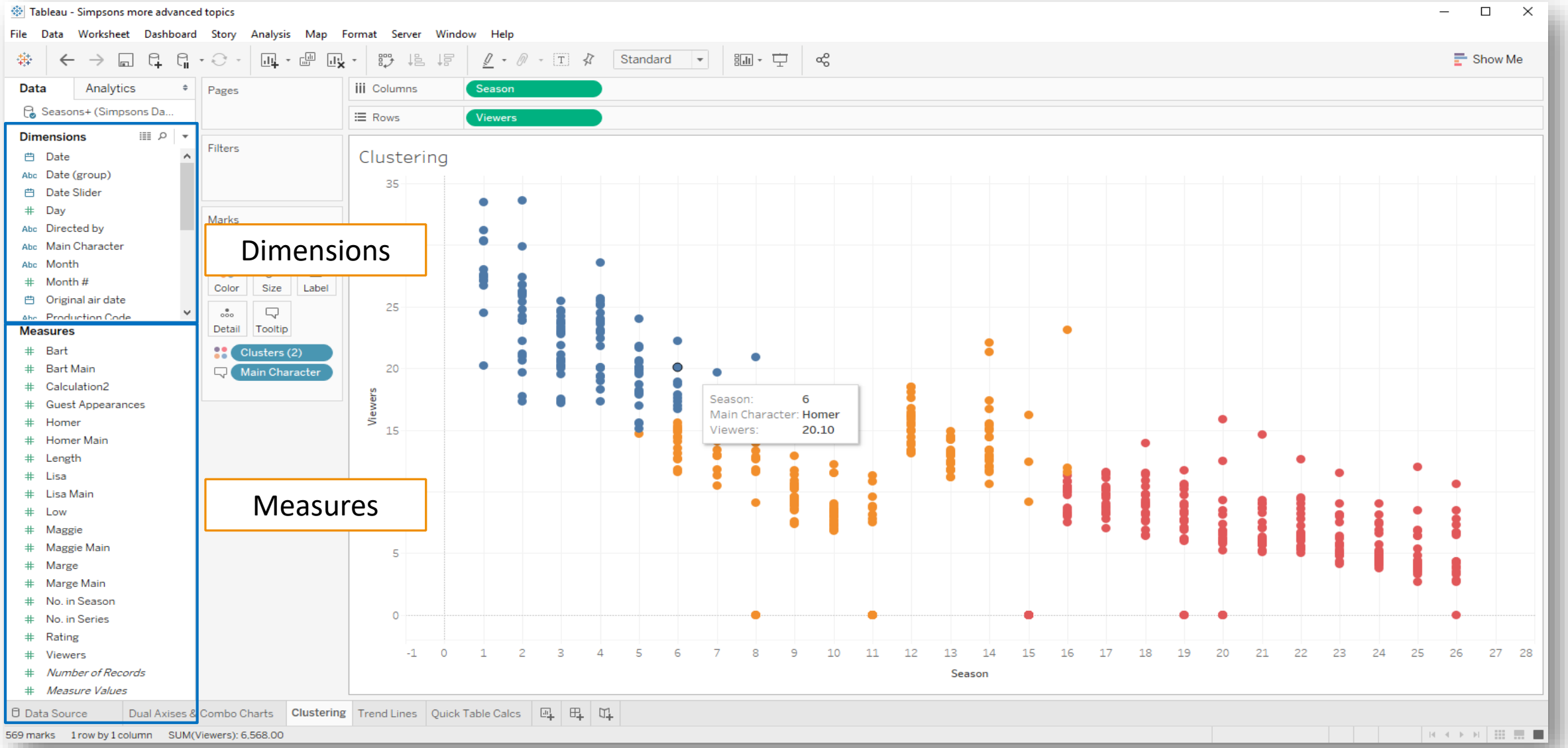
Rapid Analytical Workflow



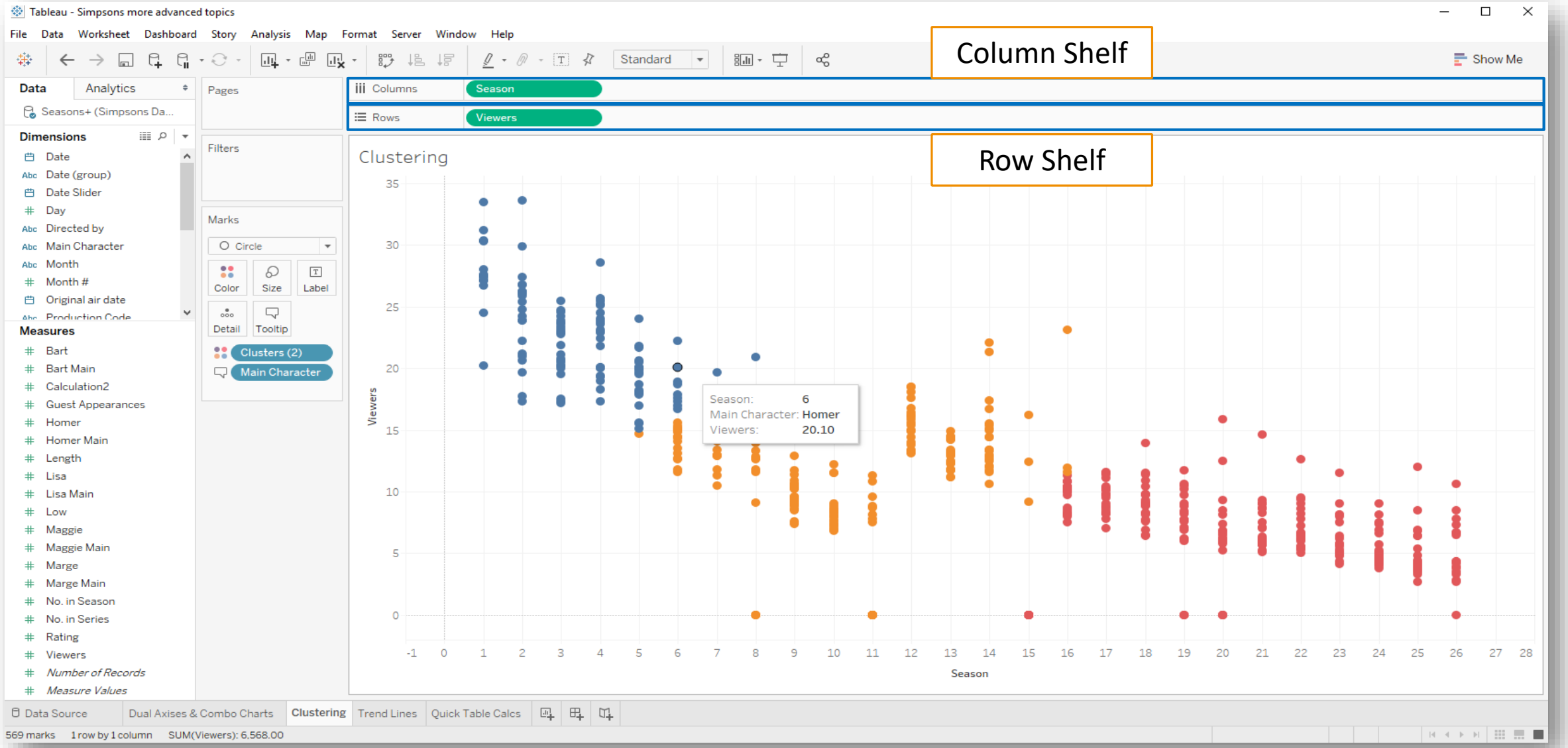
Why Tableau?

1. Tableau helps to see and understand the data
2. Tableau is flexible enough to suit different needs and different organizations
3. Tableau makes crafting narratives easy with dashboards, stories and visual analytics
4. You don't need to do any coding
5. Tableau leverages the power of the database
6. Tableau has an active professional network of users
7. Tableau is a leader within BI visualization

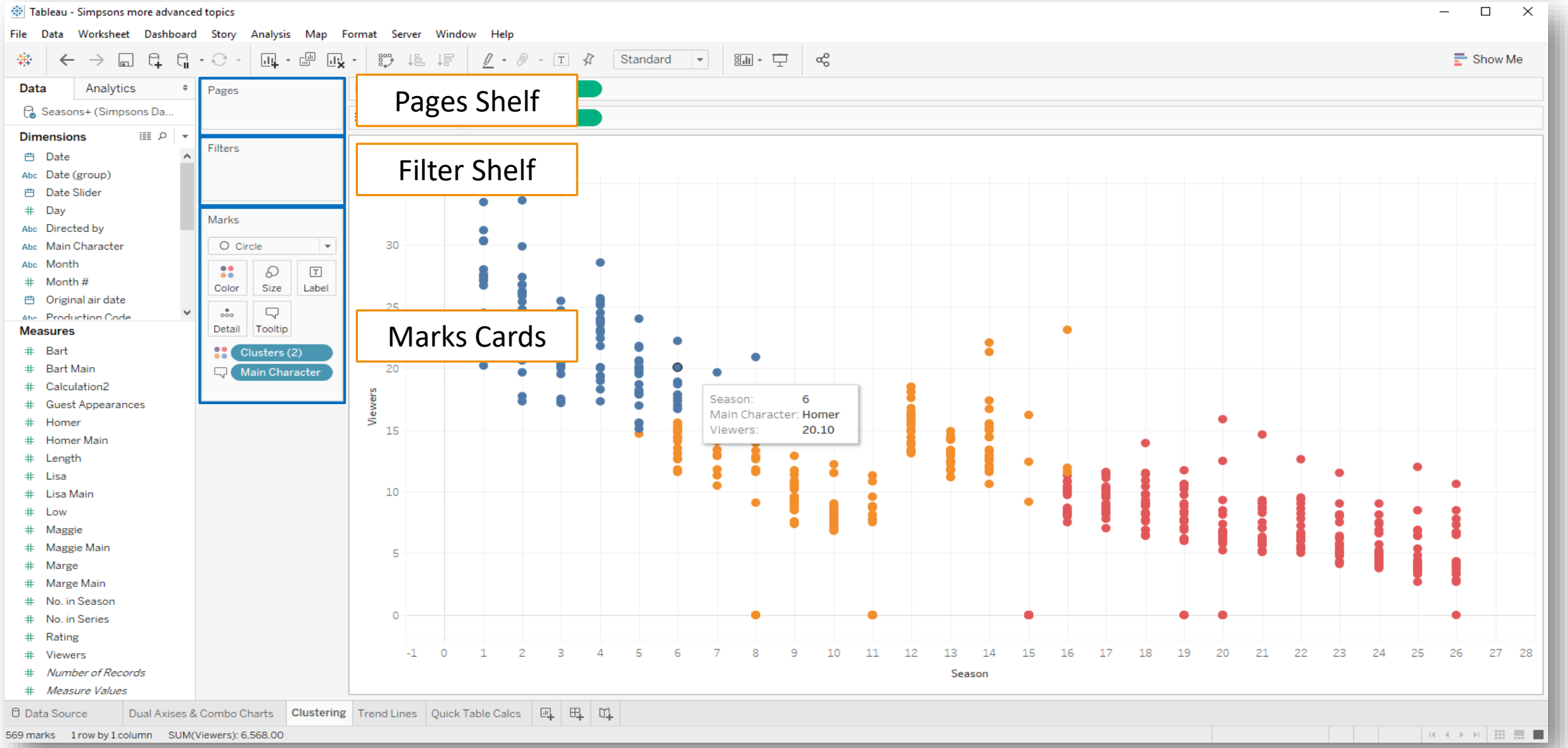
User Interface



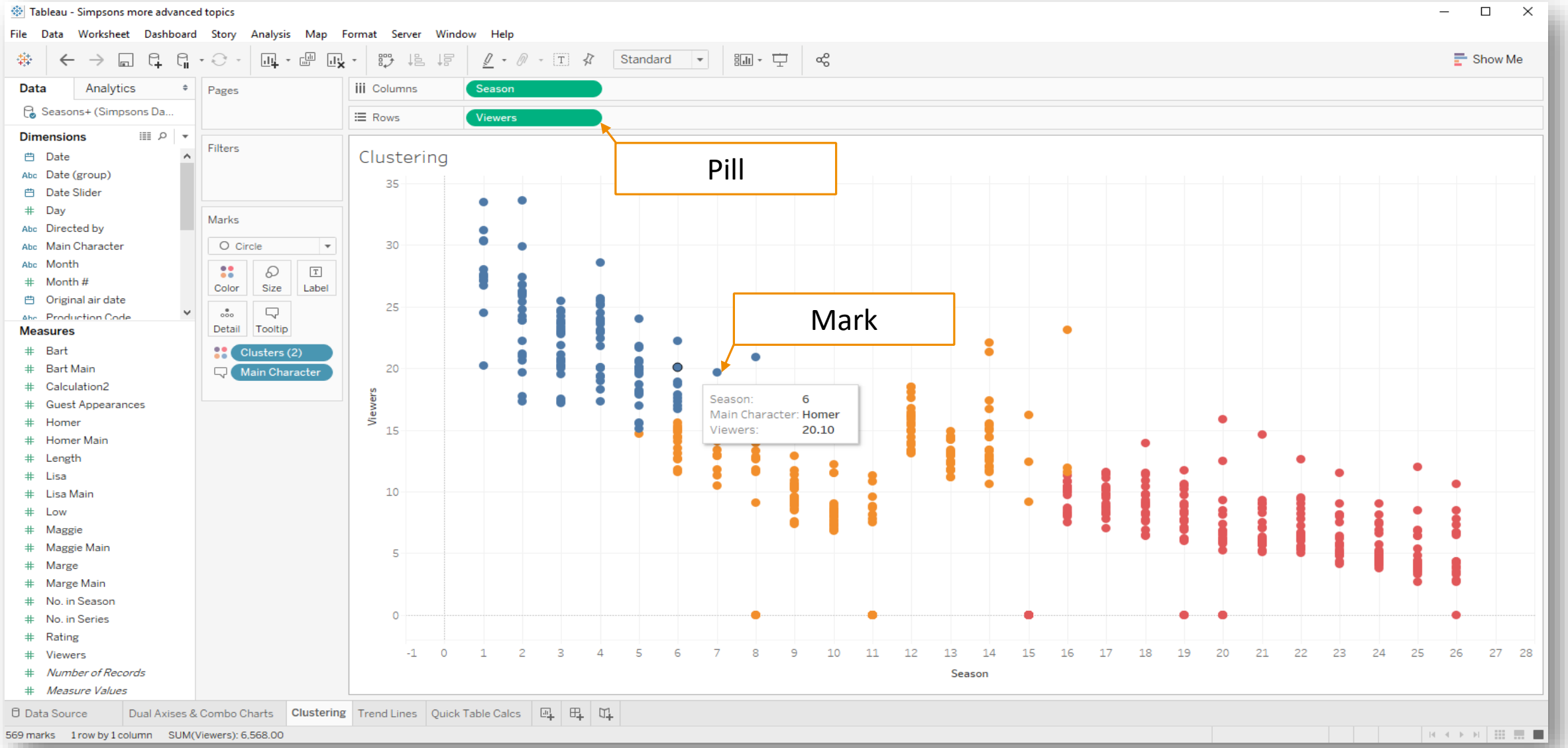
User Interface



User Interface



User Interface

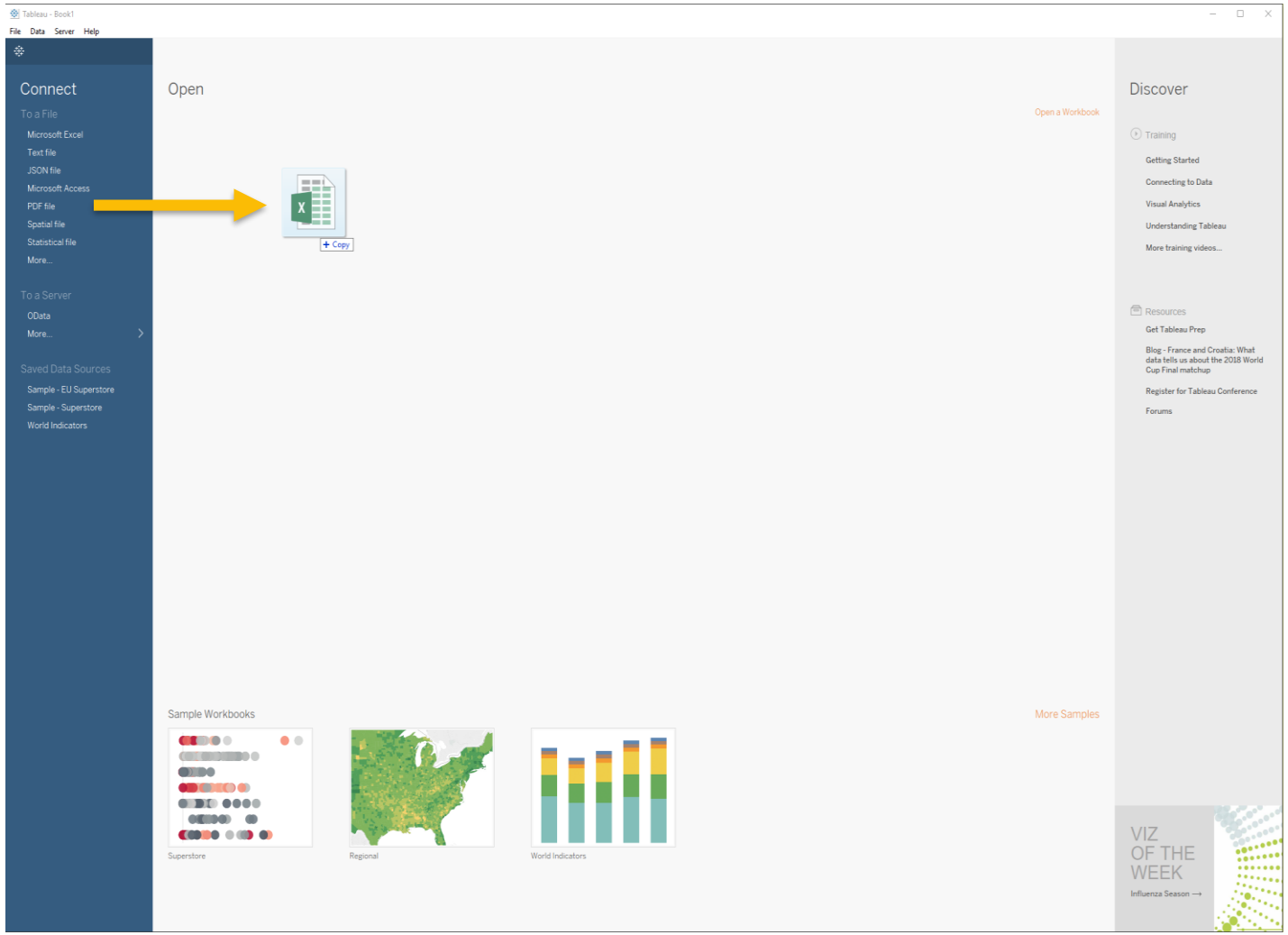


Getting started with Tableau

- Airbnb Dataset Demo

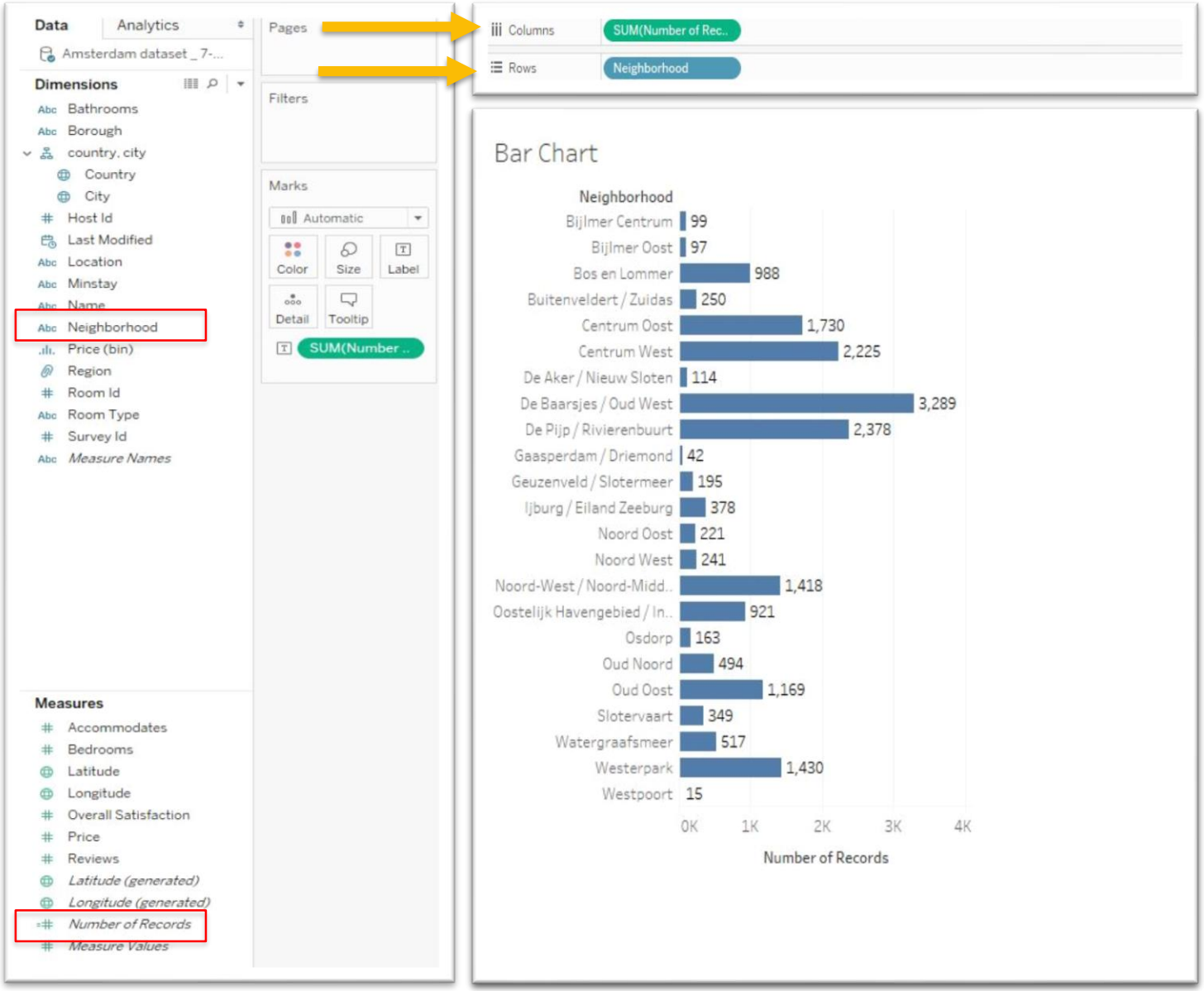
Airbnb: Listings per neighborhood

- First, connect to the Airbnb Excel file found in the Datasets folder by dragging and dropping it into Tableau



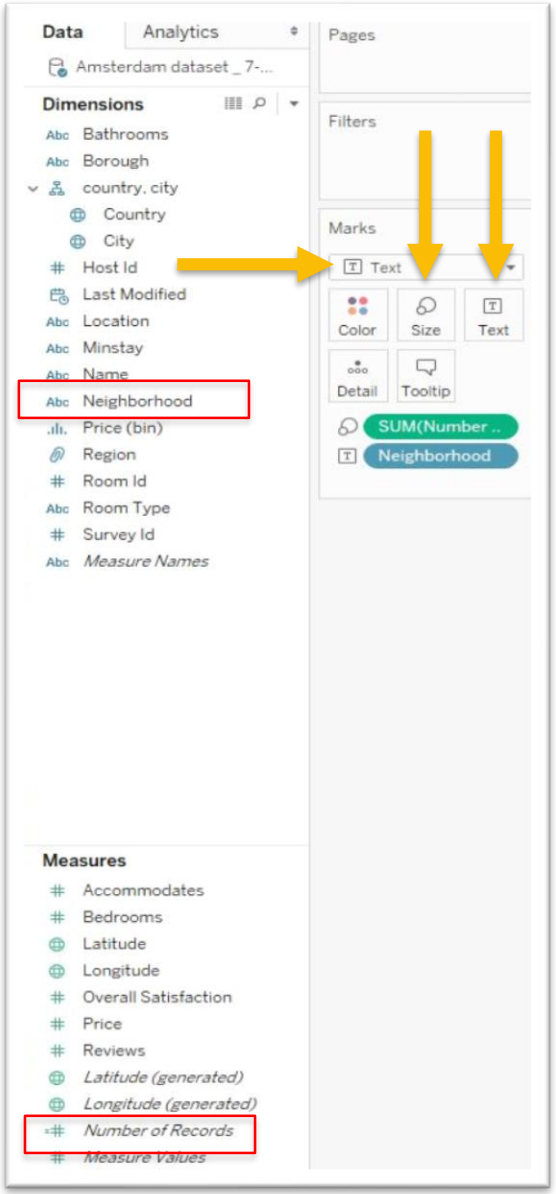
Airbnb: Listings per neighborhood

- First, connect to the Airbnb Excel file found in the Datasets folder by dragging and dropping it into Tableau
- Let's build a simple bar chart that shows the number of Airbnb listings per neighborhood in Amsterdam!
- Simply double click on **neighborhood** and **number of records**
- Tableau automatically **sums** the counts of Airbnb listings by **neighborhood** and will format the data as a **bar chart**



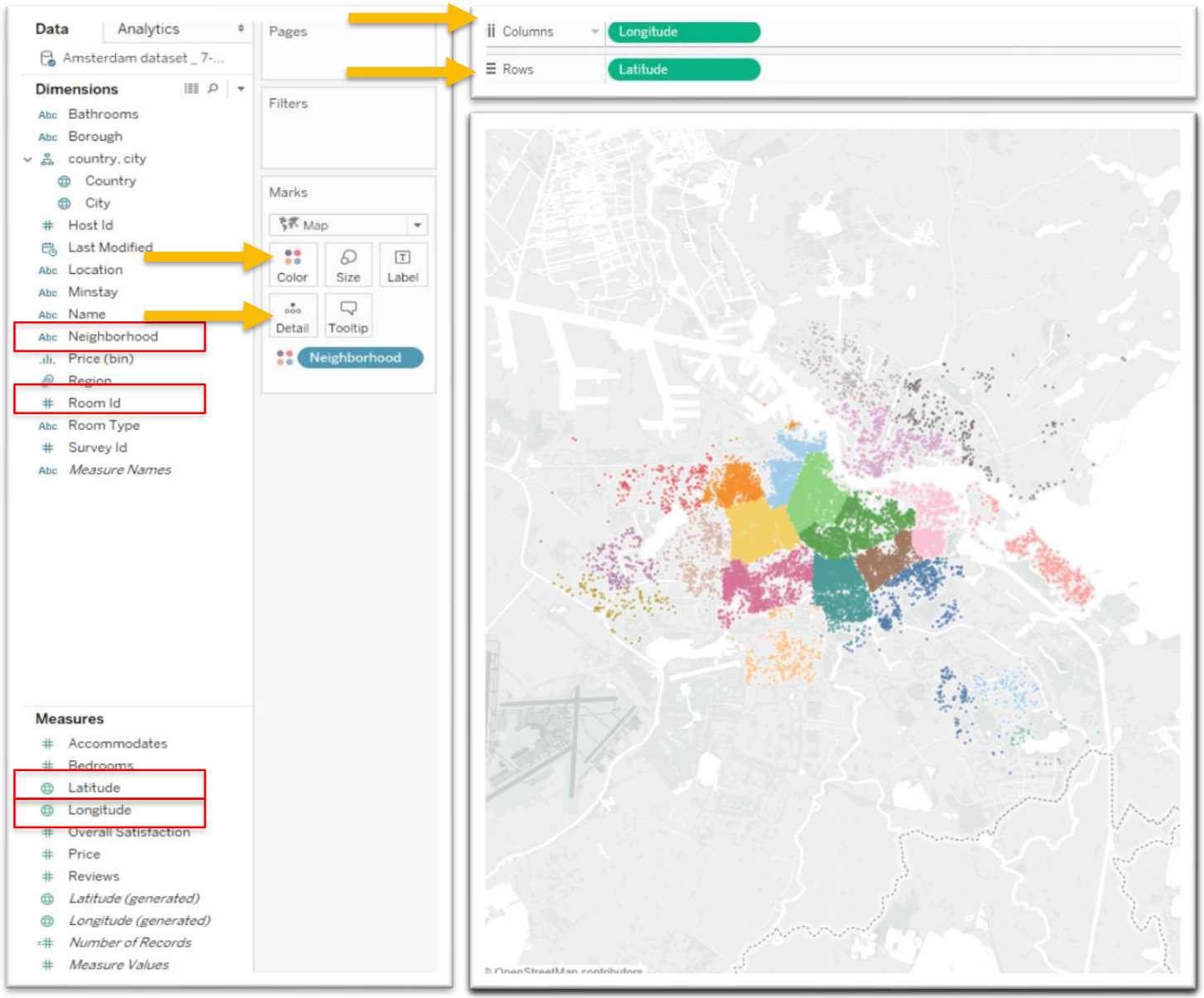
Airbnb: Listings per neighborhood

- Now let's represent the same information using a Word Cloud!
- In a new sheet, drag **neighborhood** onto **Text** in the **marks shelf**
- Next, drag **number of records** onto **Size** in the **marks shelf**
- By default, Tableau creates a **tree map**
- To change the **mark type**, simply select **text** in the dropdown menu on the **marks shelf**



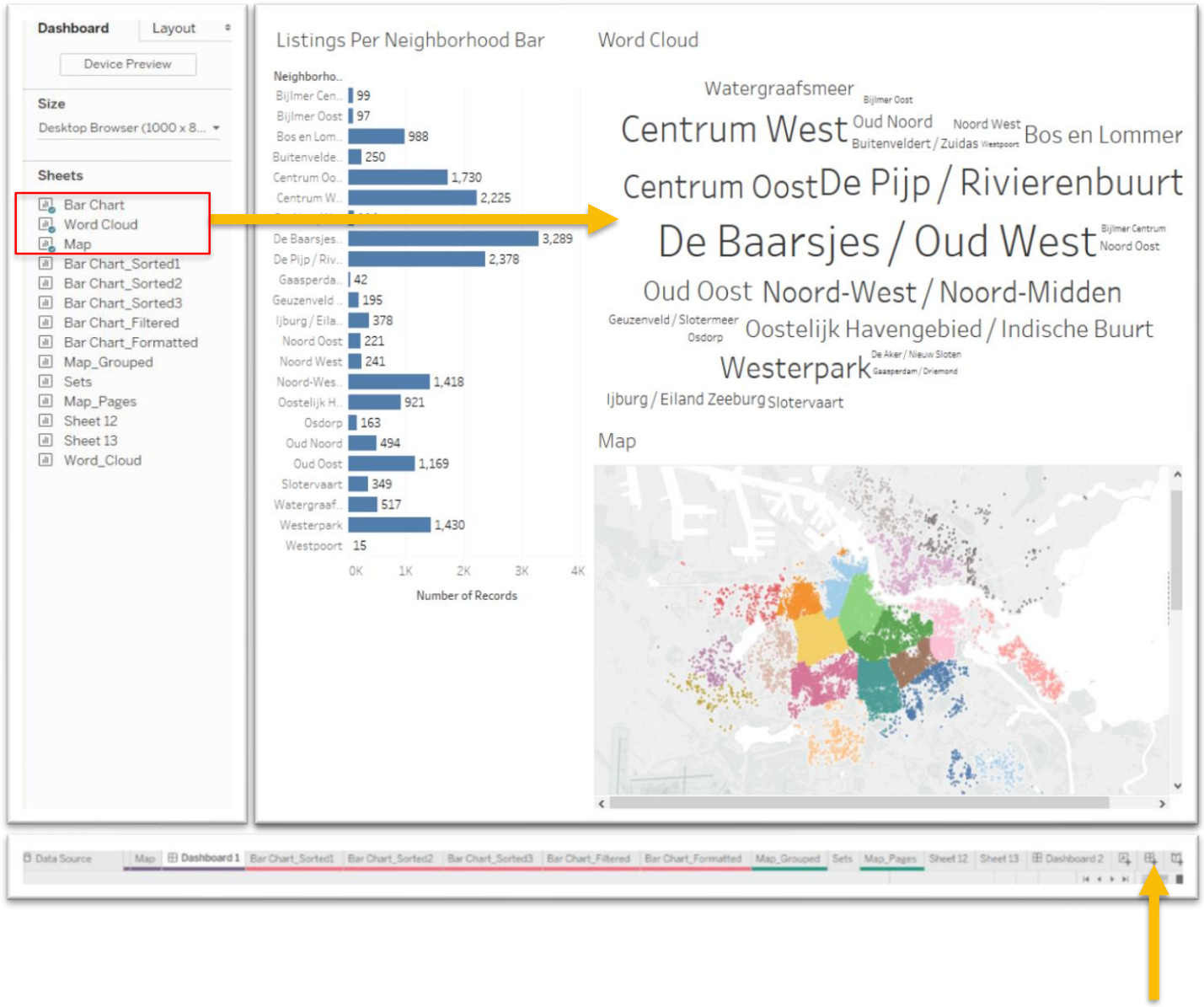
Airbnb: Listings per neighborhood

- A third way to show the data is geographically with a map
- In a new sheet, drag **longitude** onto **columns** and **latitude** onto **rows** and Tableau will automatically generate a map with a single point
- Drag **Room ID** onto **Detail** to display a point for each listing
- Drag **neighborhood** onto **Color** on the **marks shelf** to see the listings by neighborhood



Airbnb: Let's put it all together!

- Now we can create a **dashboard** to present our visualizations in a single view
- First create a new dashboard by right-clicking on the icon on the bottom of the user interface and select **new dashboard**
- Then simply drag the sheets into the view to place them in the dashboard

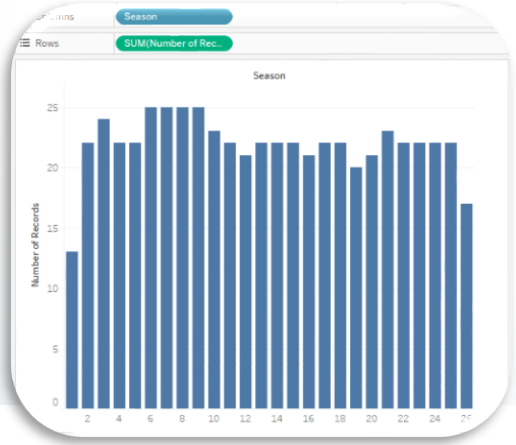


UNDERSTANDING DATA PARTS

What pill and where?

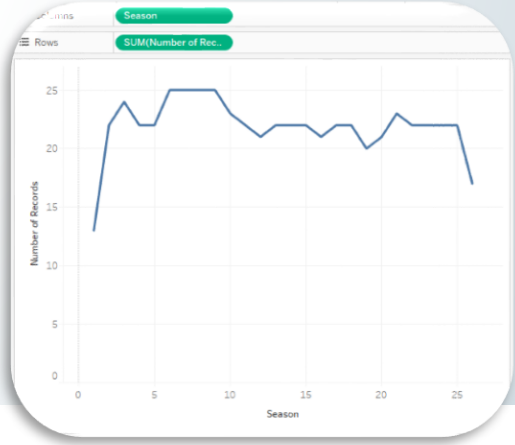
Discrete Variable

- Represented by **blue pills** ⊞ State
- Data is Discrete if the values are distinct and separate
- Usually categorical, but can be numeric or a date
- If numerical, can only be whole numbers
- Examples:
  - Neighborhood
  - Underweight/Normal /Obese classification
  - Counts



Continuous Variable

- Represented by **green pills** SUM(Profit)
- Data is Continuous if the values can be any value within a continuous range
- Can only be numeric or a date
- Examples:
  - Price
  - Weight



Dimension

- Tableau will automatically pick out fields it classifies as strings & Booleans and classify them as dimensions
- When you click and drag a field from the Dimensions area to Rows or Columns, Tableau creates column or row headers
- You can only turn a dimension into a continuous variable if it's a date
- Cannot be aggregated by a function such as sum() or avg()

Measure

- Tableau will automatically pick out fields it classifies as quantitative & numeric and classify them as measures
- When you drag a field from the Measures area to Rows or Columns, Tableau creates a continuous axis.
- Can be converted into discrete variables
- Can be aggregated or disaggregated

How To Distinguish Dimensions vs. Measures:

- Measures will have an **aggregation function** shown before the name when added to a shelf
  - SUM, MIN, MAX, … , AGG
- If there is no aggregation function before the variable name, it's a **dimension**

Exercise 1

Simpsons Part 1

- Exercise Objectives:
  - Understand some of the cues Tableau provides to help us understand our variables
  - Build intuition on different types of variables
  - Understand how Tableau uses aggregation on variables

Exercise Materials

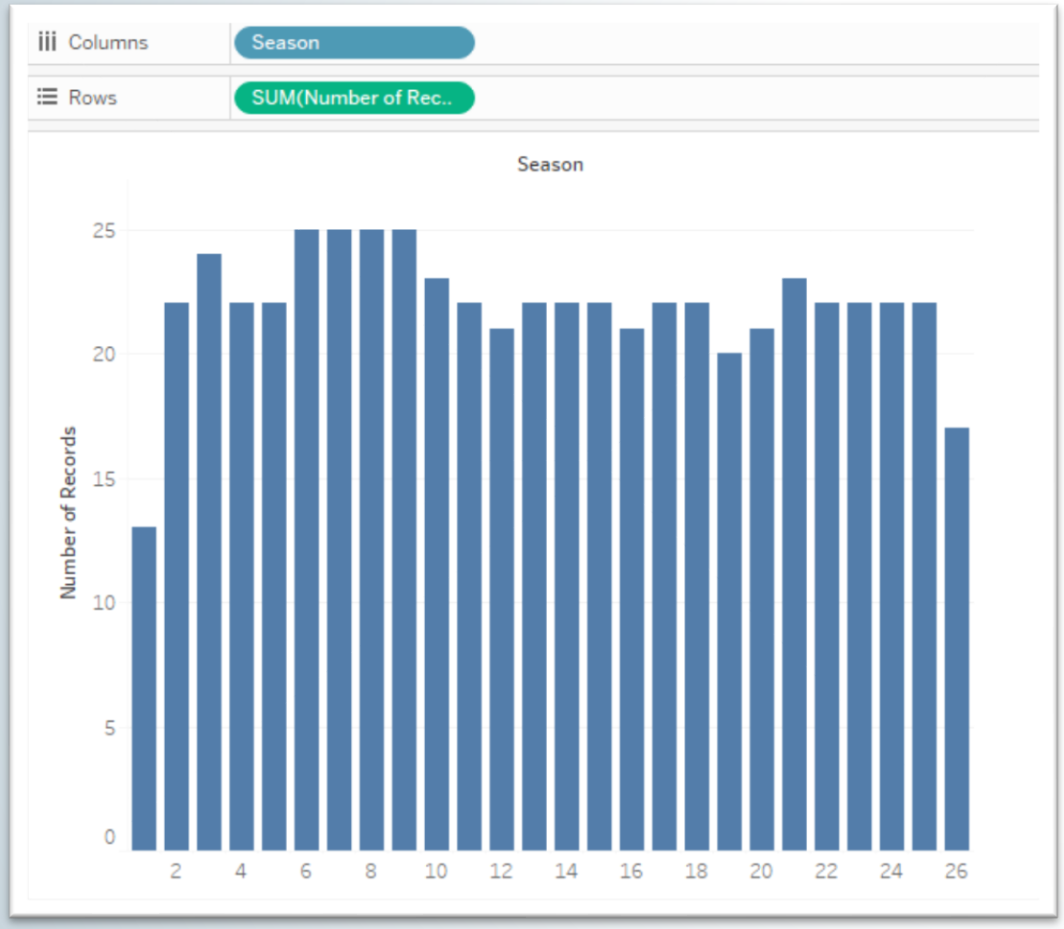
Exercise 1 Simpsons-Part 1-Starter.twbx

Exercise 1 Simpsons-Part 1-Solution.twbx

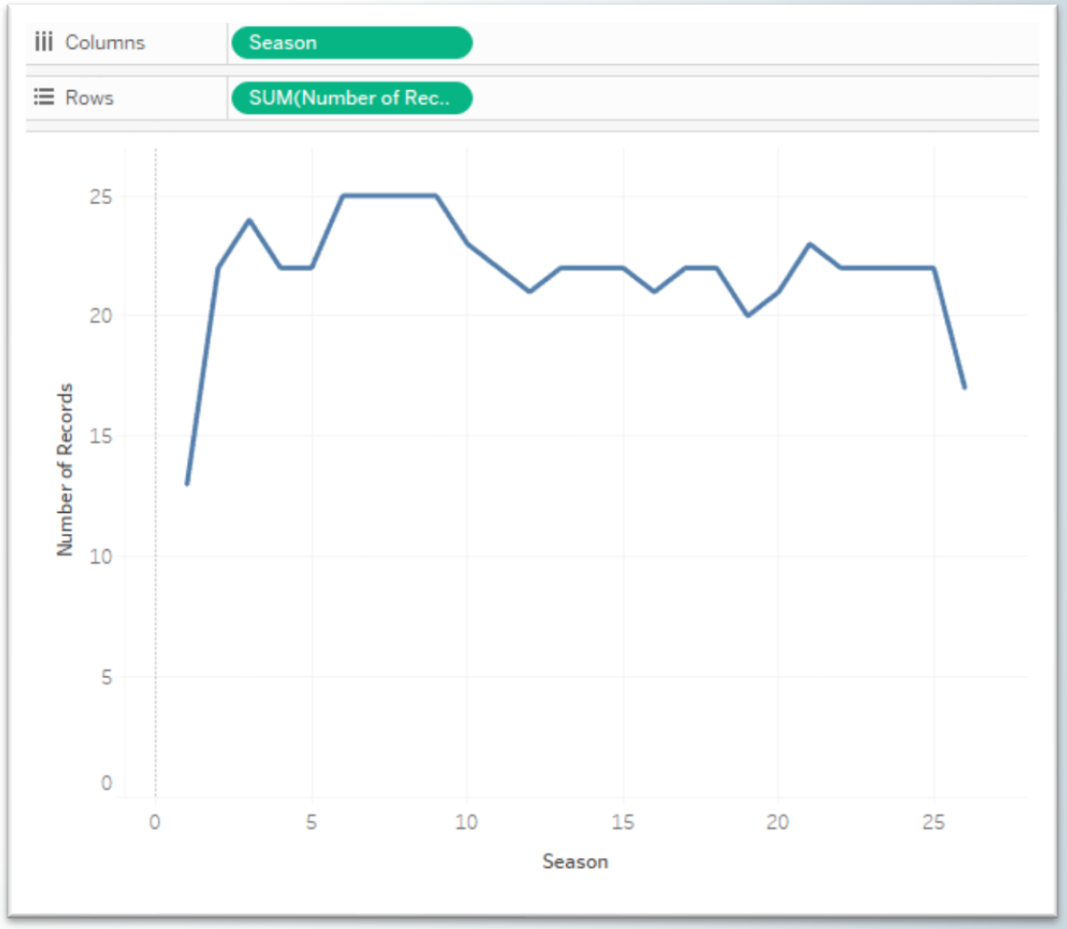
Check your Understanding

- Can a variable be either a dimension or a measure?
- Can a non-numeric variable be continuous?
- Can a numeric variable be discrete?
- What color does Tableau use to identify a Discrete variable?

Discrete

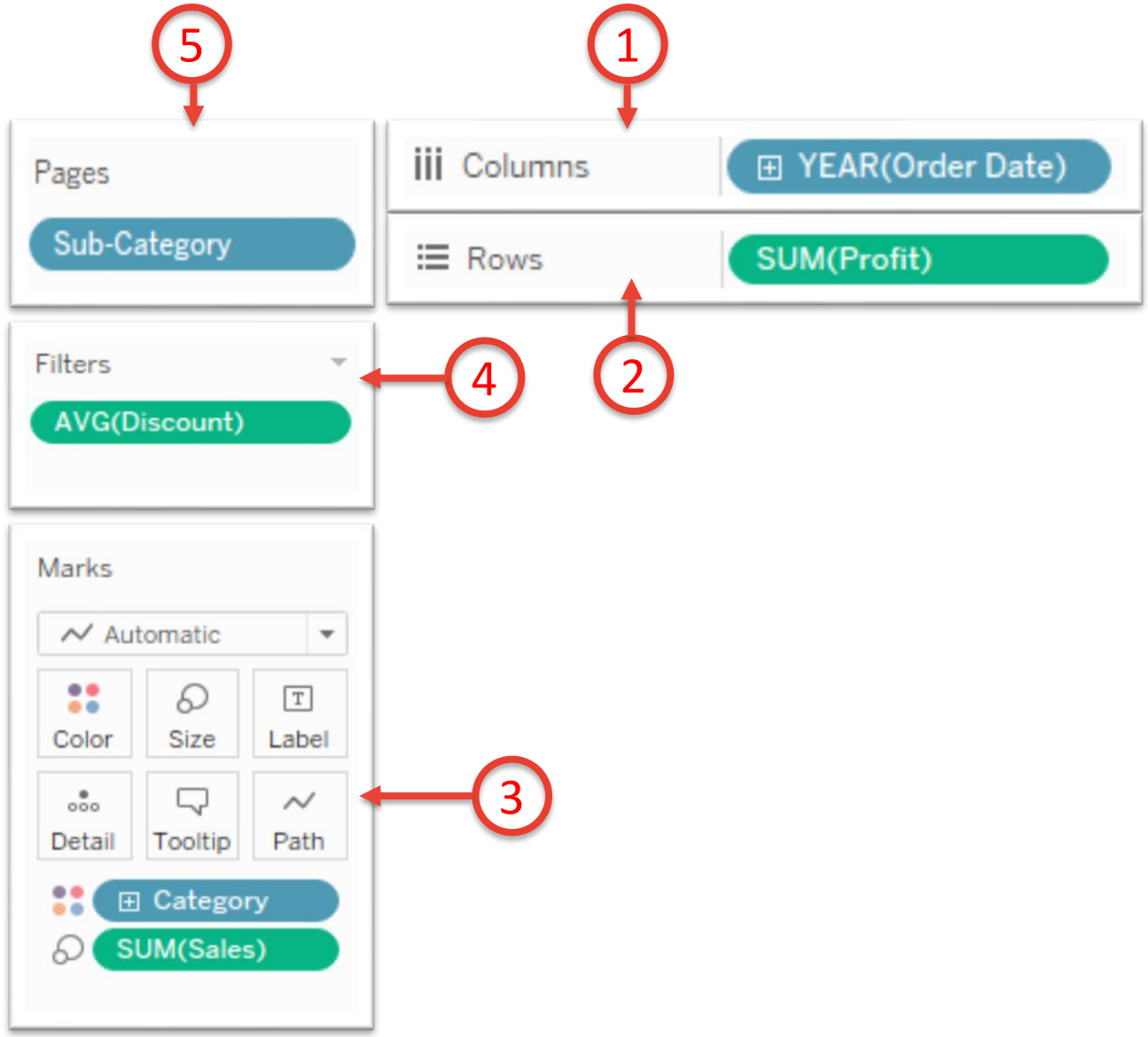


Continuous



Discrete vs. Continuous

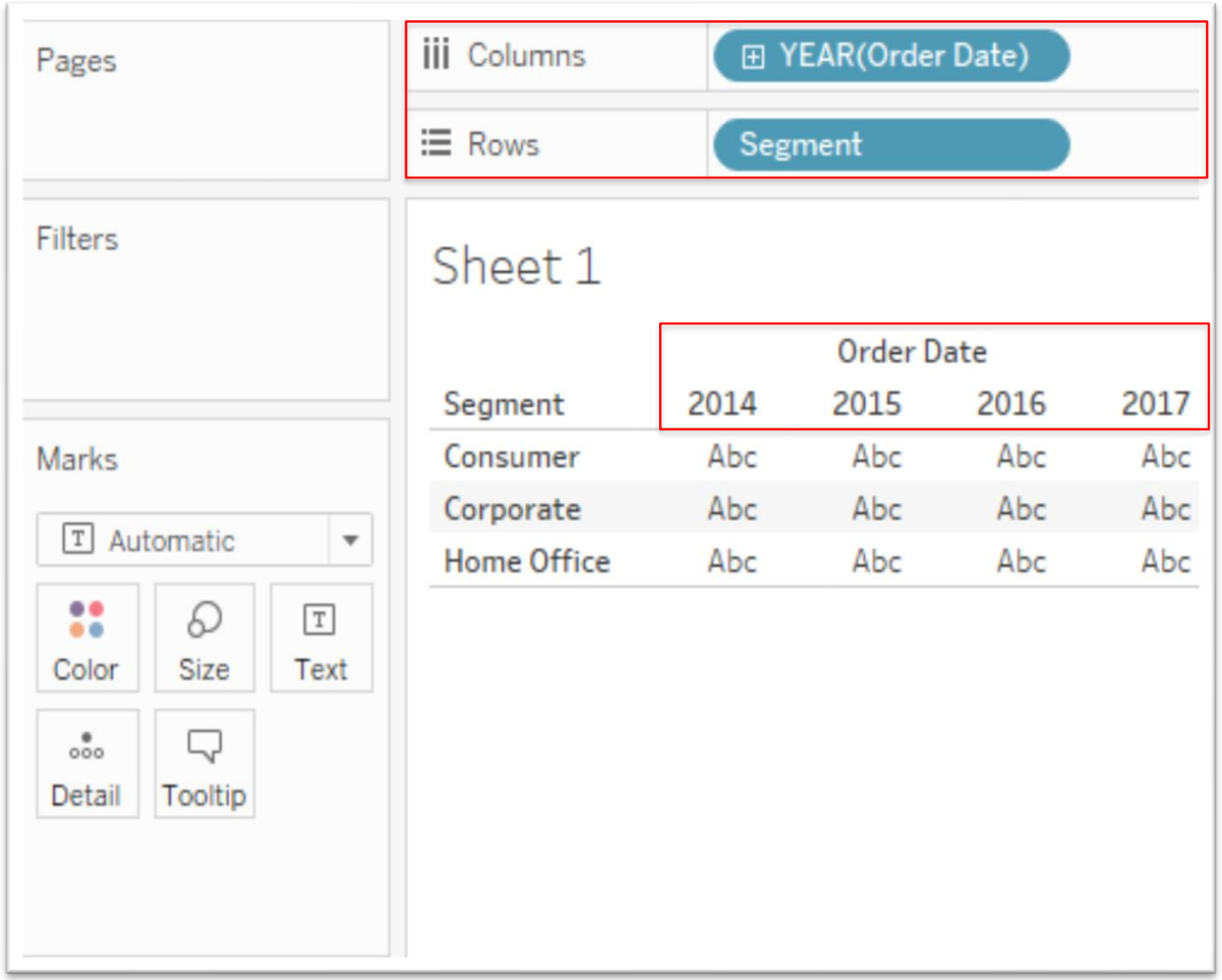
- Discrete variables are displayed as **blue pills**
- Continuous variables are displayed as **green pills**
- Where can we find them?
  1. Column Shelf
  2. Rows Shelf
  3. Marks Card
  4. Filters Card
  5. Pages Card



Discrete vs. Continuous

What happens when a Discrete Field is added to the columns or rows shelf?

Tableau creates a horizontal or vertical representation of the data displayed as a **Header**!



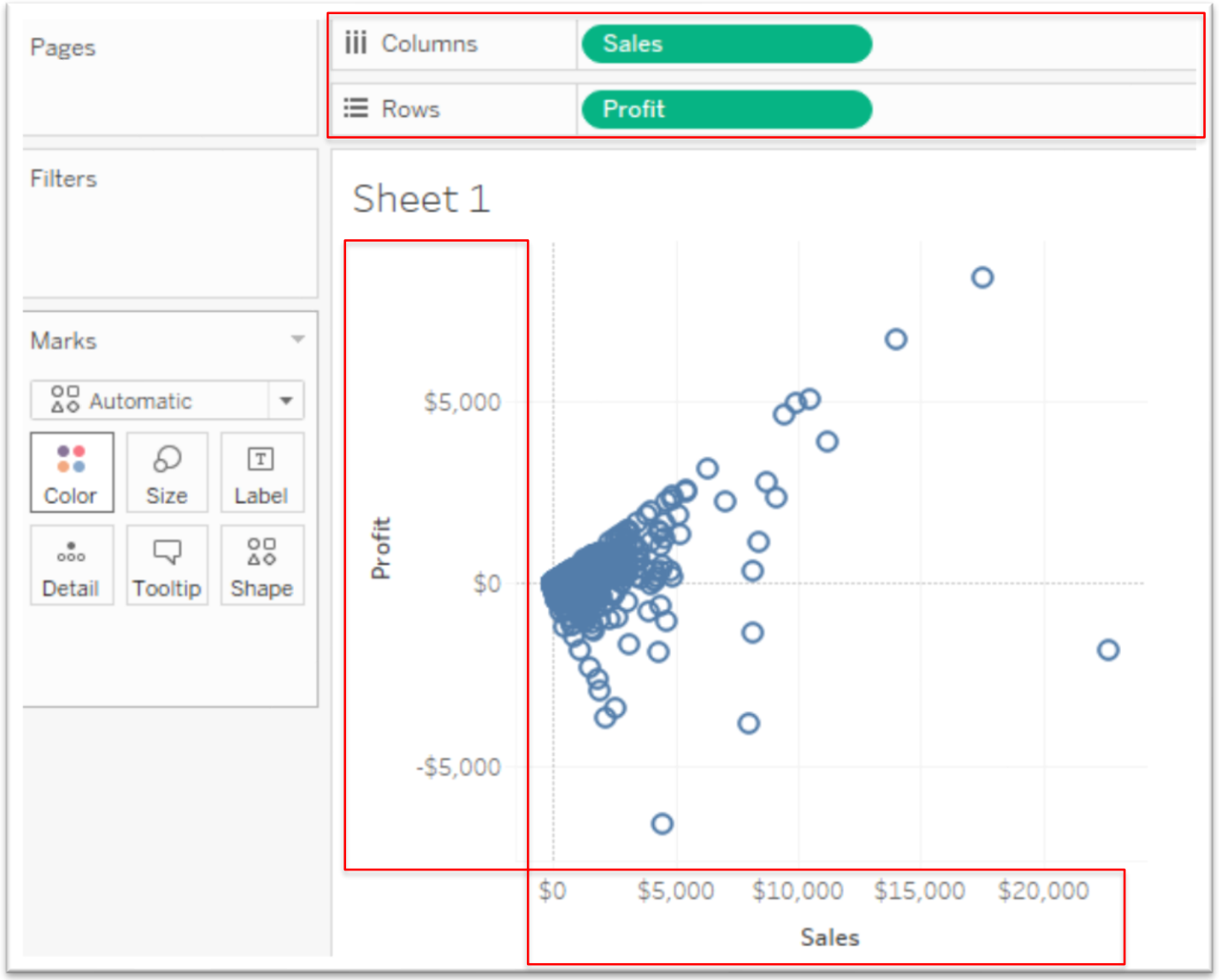
Discrete vs. Continuous

What happens when a Discrete Field is added to the columns or rows shelf?

Tableau creates a horizontal or vertical representation of the data displayed as a **Header**!

What happens when a Continuous Field is added to the columns or rows shelf?

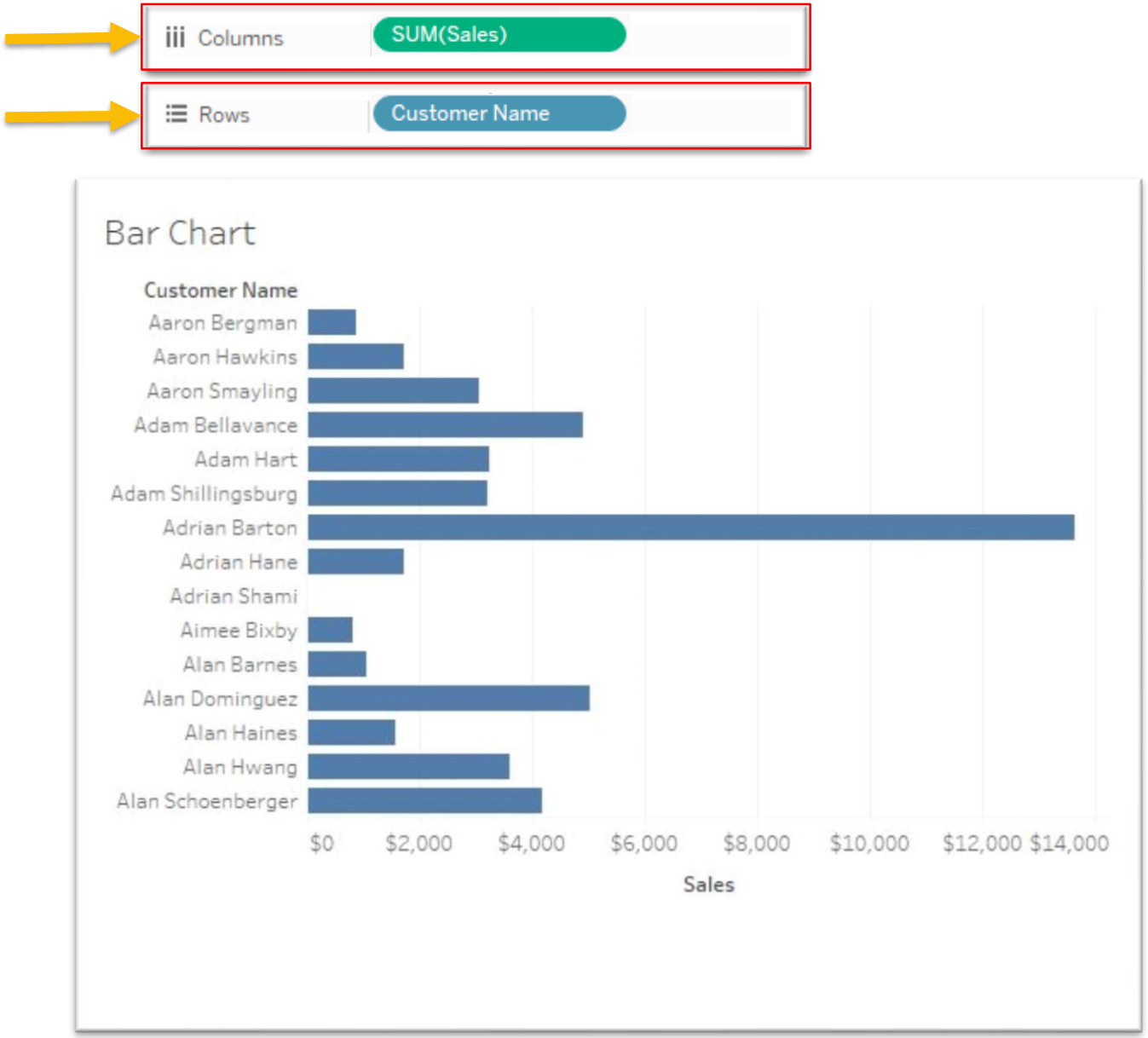
Tableau creates a horizontal or vertical representation of the data displayed as an **Axis**!



Check your Understanding

Q1: What Field Type is on the **Columns shelf**?

Q2: What Field Type is on the **Rows shelf**?



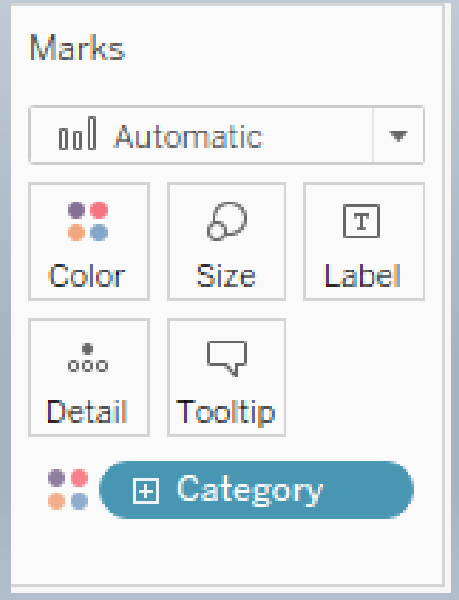
Discrete vs. Continuous

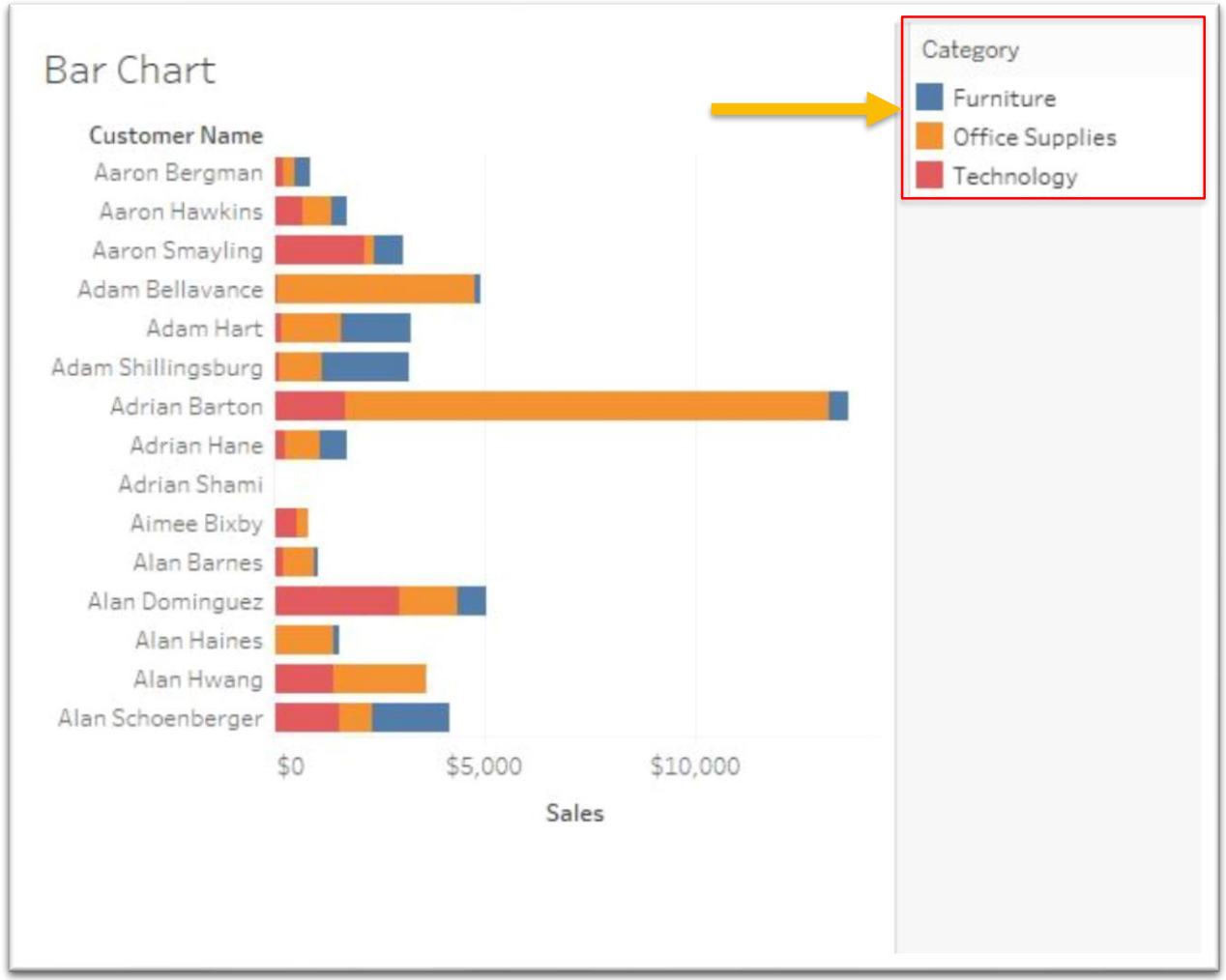
- …Throw some Color on It!

Discrete vs. Continuous

What happens when a **Discrete** Field is added to **Color** on the **Marks shelf**?

Tableau creates a **categorical** palette!

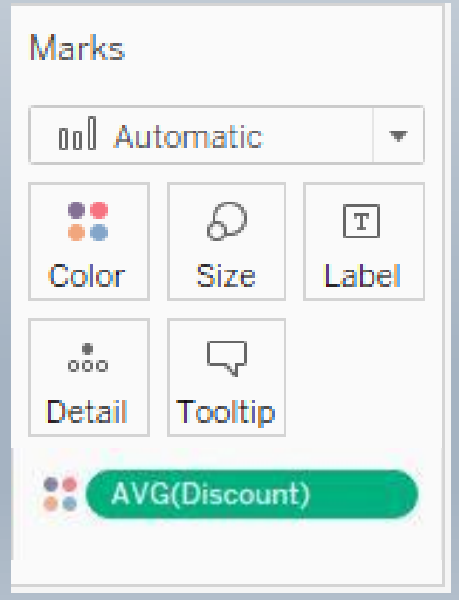


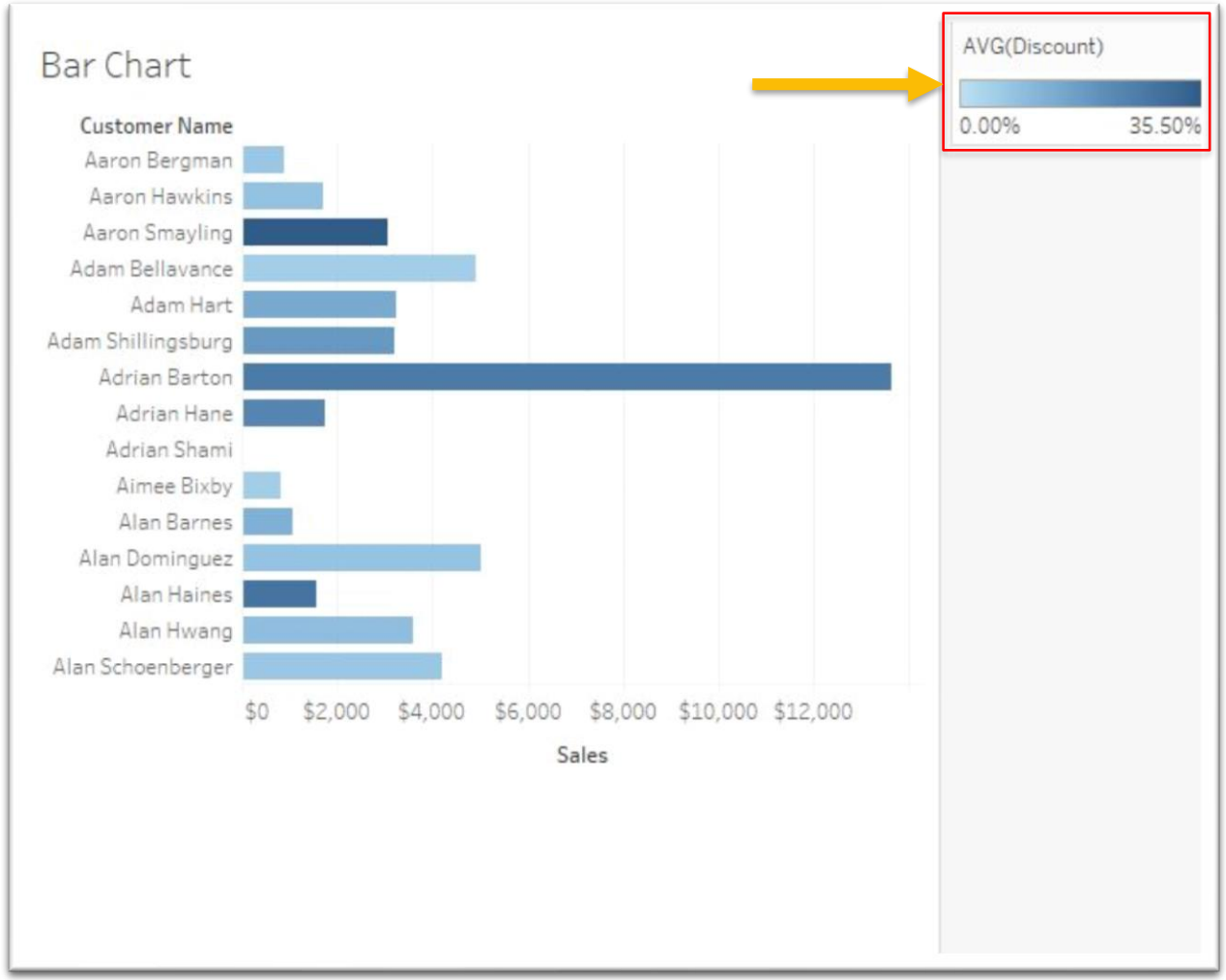


Discrete vs. Continuous

What happens when a **Continuous** Field is added to **Color** on the **Marks shelf**?

Tableau creates a **spectrum** palette!





What Pill and Where?

- …with a Heat Map

Check your Understanding

Q1: What kind of field is on the **Columns Shelf**?

Q2: What kind of field is on the **Rows Shelf**?

Q3: What kind of field is on the **Color** on the **Marks Shelf**?

Let's Review!

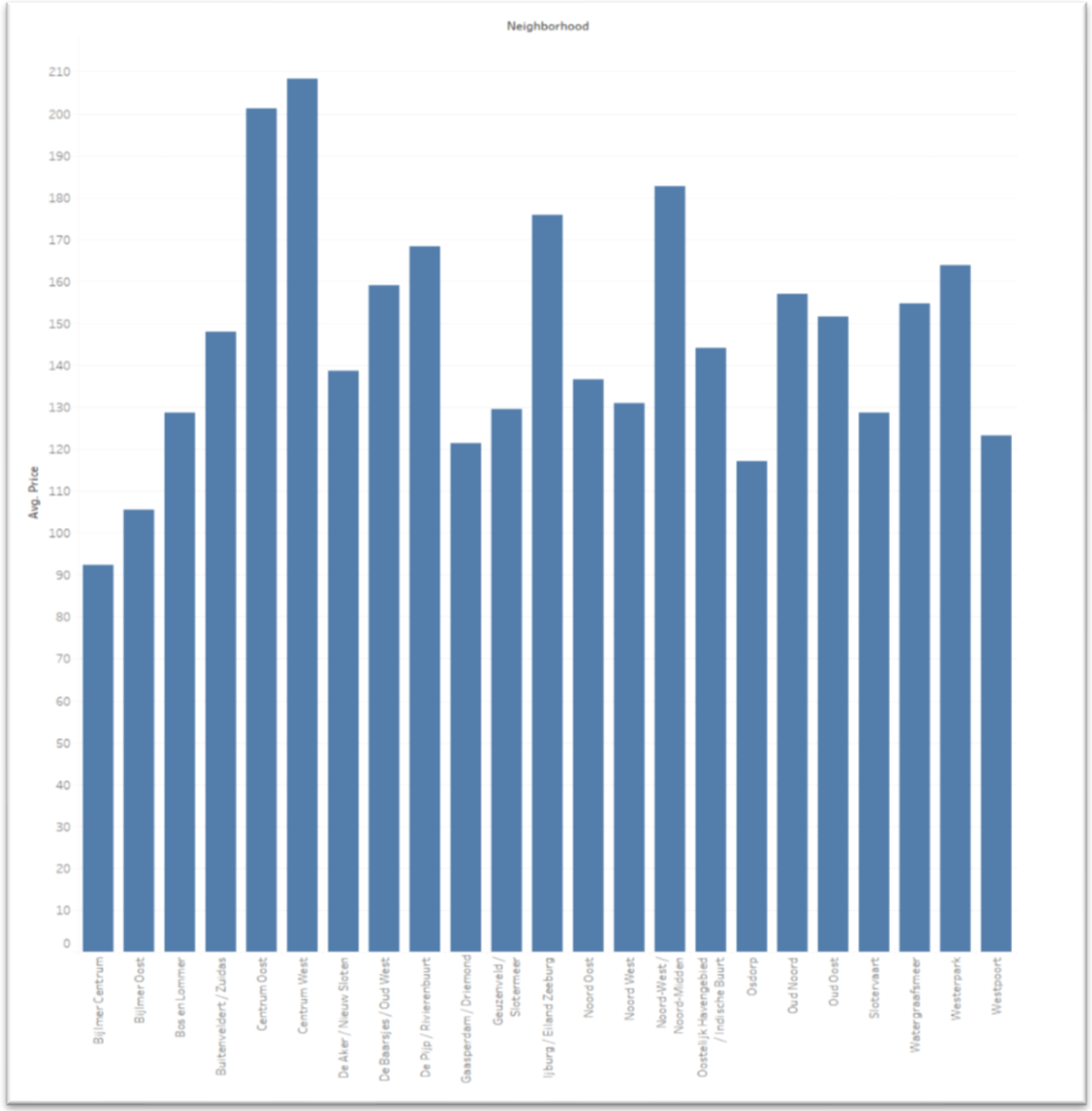
- Choosing dimension or measure affects how values are ***calculated***
- Choosing discrete or continuous affects how values are ***displayed***
  - Dimensions are *usually* discrete (blue)
  - Measures are *usually* continuous (green)
  - Fields can be treated differently in different visualizations

AGGREGATING YOUR DATA

Slicing and Dicing your Data

Aggregation

- **Aggregation** occurs when values of multiple rows are combined as the input to form a single value of more significant meaning
- **Marks** summarize the data by aggregating the data and summarizing the result
- Commonly used aggregate functions:
  - SUM
  - MIN/MAX
  - AVG
  - COUNT
  - AGG
- In this Airbnb example, Tableau is aggregating the Price by Neighborhood using AVG (average)



Let's Use Aggregation to Turn This Data Table…

Function	Subfunction	Year	Amount	GDP	CPI
National Defense	Defense-related activities	2013	16,034,000,000	16,202,700,000,000	2.3028
Energy	Emergency energy preparedness	2014	206,000,000	17,011,400,000,000	2.3392
Energy	Energy conservation	2014	1,987,000,000	17,011,400,000,000	2.3392
Energy	Energy information, policy, and regulation	2014	542,000,000	17,011,400,000,000	2.3392
Energy	Energy supply	2014	10,640,000,000	17,011,400,000,000	2.3392
Health	Consumer and occupational health and safety	2014	4,746,000,000	17,011,400,000,000	2.3392
Health	Health care services	2014	412,526,000,000	17,011,400,000,000	2.3392
Health	Health research and training	2014	33,523,000,000	17,011,400,000,000	2.3392

…Into This Summary Table!

Function	Year	SUM(Amount)	MAX(Amount)
Energy	2013	11,042,000,000	9,038,000,000
Health	2013	358,315,000,000	321,849,000,000
Energy	2014	13,375,000,000	10,640,000,000
Health	2014	450,795,000,000	412,526,000,000

How Does Tableau See Our Data?

- Tableau reads from tables of data
  - Each row represents one **record**
  - Each column represents one **field**

Function	Subfunction	Year	Amount	GDP	CPI
National Defense	Defense-related activities	2013	16,034,000,000	16,202,700,000,000	2.3028
Energy	Emergency energy preparedness	2014	206,000,000	17,011,400,000,000	2.3392
Energy	Energy conservation	2014	1,987,000,000	17,011,400,000,000	2.3392
Energy	Energy information, policy, and regulation	2014	542,000,000	17,011,400,000,000	2.3392
Energy	Energy supply	2014	10,640,000,000	17,011,400,000,000	2.3392
Health	Consumer and occupational health and safety	2014	4,746,000,000	17,011,400,000,000	2.3392
Health	Health care services	2014	412,526,000,000	17,011,400,000,000	2.3392
Health	Health research and training	2014	33,523,000,000	17,011,400,000,000	2.3392

Begin with the Federal Spending dataset

Function	Subfunction	Year	Amount	GDP	CPI
National Defense	Defense-related activities	2013	16,034,000,000	16,202,700,000,000	2.3028
Energy	Emergency energy preparedness	2014	206,000,000	17,011,400,000,000	2.3392
Energy	Energy conservation	2014	1,987,000,000	17,011,400,000,000	2.3392
Energy	Energy information, policy, and regulation	2014	542,000,000	17,011,400,000,000	2.3392
Energy	Energy supply	2014	10,640,000,000	17,011,400,000,000	2.3392
Health	Consumer and occupational health and safety	2014	4,746,000,000	17,011,400,000,000	2.3392
Health	Health care services	2014	412,526,000,000	17,011,400,000,000	2.3392
Health	Health research and training	2014	33,523,000,000	17,011,400,000,000	2.3392

Filtering Data Rows

Function	Subfunction	Year	Amount	GDP	CPI
~~National Defense~~	~~Defense-related activities~~	~~2013~~	~~16,034,000,000~~	~~16,202,700,000,000~~	~~2.3028~~
Energy	Emergency energy preparedness	2014	206,000,000	17,011,400,000,000	2.3392
Energy	Energy conservation	2014	1,987,000,000	17,011,400,000,000	2.3392
Energy	Energy information, policy, and regulation	2014	542,000,000	17,011,400,000,000	2.3392
Energy	Energy supply	2014	10,640,000,000	17,011,400,000,000	2.3392
Health	Consumer and occupational health and safety	2014	4,746,000,000	17,011,400,000,000	2.3392
Health	Health care services	2014	412,526,000,000	17,011,400,000,000	2.3392
Health	Health research and training	2014	33,523,000,000	17,011,400,000,000	2.3392

Database only retains data rows that match the criteria on the filter shelf

Selecting Columns (Fields)

Function	~~Subfunction~~	Year	Amount	~~GDP~~	~~CPI~~
Energy	~~Emergency energy preparedness~~	2014	206,000,000	~~17,011,400,000,000~~	~~2.3392~~
Energy	~~Energy conservation~~	2014	1,987,000,000	~~17,011,400,000,000~~	~~2.3392~~
Energy	~~Energy information, policy, and regulation~~	2014	542,000,000	~~17,011,400,000,000~~	~~2.3392~~
Energy	~~Energy supply~~	2014	10,640,000,000	~~17,011,400,000,000~~	~~2.3392~~
Health	~~Consumer and occupational health and safety~~	2014	4,746,000,000	~~17,011,400,000,000~~	~~2.3392~~
Health	~~Health care services~~	2014	412,526,000,000	~~17,011,400,000,000~~	~~2.3392~~
Health	~~Health research and training~~	2014	33,523,000,000	~~17,011,400,000,000~~	~~2.3392~~

Database only retains columns that are used on the visualization
(i.e. fields that appear on a shelf other than the filter shelf)

The Resulting Data

Function	Year	Amount
Energy	2013	9,038,000,000
Energy	2013	1,240,000,000
Energy	2013	217,000,000
Energy	2013	547,000,000
Health	2013	321,849,000,000
Health	2013	32,881,000,000
Health	2013	3,585,000,000
Energy	2014	206,000,000
Energy	2014	1,987,000,000
Energy	2014	542,000,000
Energy	2014	10,640,000,000
Health	2014	4,746,000,000
Health	2014	412,526,000,000
Health	2014	33,523,000,000

Partitioning Function by Year

Energy 2013

Function	Year	Amount
Energy	2013	9,038,000,000
Energy	2013	1,240,000,000
Energy	2013	217,000,000
Energy	2013	547,000,000

Energy 2014

Function	Year	Amount
Energy	2014	206,000,000
Energy	2014	1,987,000,000
Energy	2014	542,000,000
Energy	2014	10,640,000,000

Health 2013

Function	Year	Amount
Health	2013	321,849,000,000
Health	2013	32,881,000,000
Health	2013	3,585,000,000

Health 2014

Function	Year	Amount
Health	2014	4,746,000,000
Health	2014	412,526,000,000
Health	2014	33,523,000,000

Tableau partitions data rows by dimensions so each partition contains only rows with the same values (**members**) for each dimension

Aggregating the Energy Spending

Function	Year	Amount
Energy	2013	9,038,000,000
Energy	2013	1,240,000,000
Energy	2013	217,000,000
Energy	2013	547,000,000

Function	Year	SUM(Amount)	MAX(Amount)
Energy	2013	11,042,000,000	9,038,000,000

Single summary row representing all data rows in the Energy 2013 partition

Tableau aggregates the measures within each partition to yield a *single row* summarizing each partition, using an aggregation function e.g., SUM, AVG, MIN, MAX, COUNT, COUNTD …

Aggregating the Results

Function	Year	SUM(Amount)	MAX(Amount)
Energy	2013	11,042,000,000	9,038,000,000
Health	2013	358,315,000,000	321,849,000,000
Energy	2014	13,375,000,000	10,640,000,000
Health	2014	450,795,000,000	412,526,000,000

Summary Results Sent to Tableau
One row per partition

Aggregate Functions and Data Flow

Tableau

Construct (optimized) query

Aliases, Sorting, Table Calculations, Rendering

Query

Select Function, Year,
SUM(Amount), Max(Amount)
from …. Group By Function, Year

Query Results
(Summary Data)

Joins, Unions, Filtering, Selection, Partition, Aggregation, most calculations

Function	Year	SUM(Amount)	MAX(Amount)
Energy	2013	11,042,000,000	9,038,000,000
Health	2013	358,315,000,000	321,849,000,000
Energy	2014	13,375,000,000	10,640,000,000
Health	2014	450,795,000,000	412,526,000,000

DATA MANIPULATION

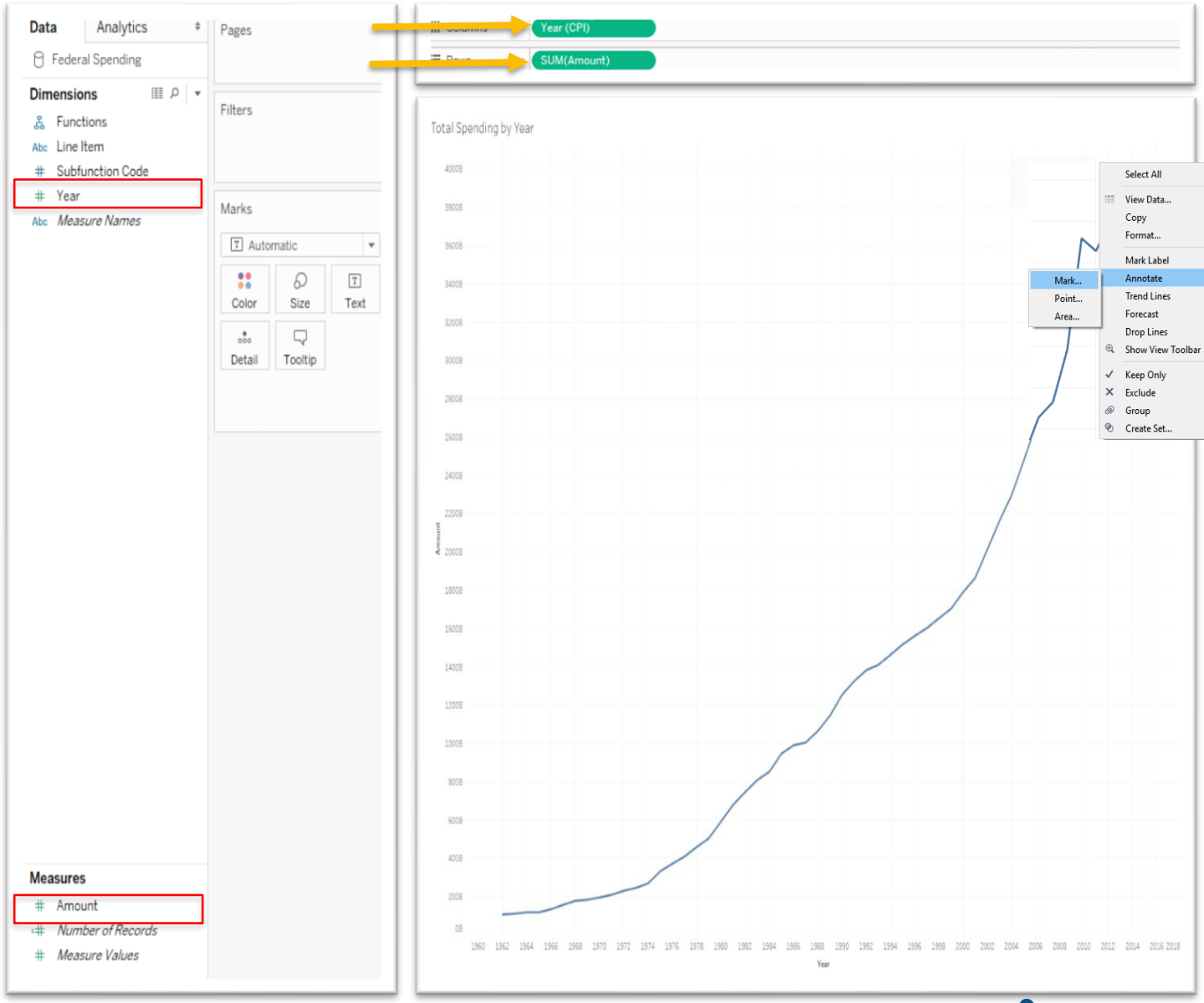
Federal Spending Dataset

Concepts

- **Join:** Joining is a method for combining related data from multiple **tables** based on a common **field**.
- **Calculation:** A new **field** (or **column**) in your data source, the values of which are determined by a **calculation** that you control. The new field is saved to your data source.
- **Quick Sort:** Tableau can quickly arrange dimensions in your visualization in ascending or descending order by hovering over the axis and clicking the sort icon.
- **Hierarchy:** Hierarchies allow us to organize and drill down into our data. Tableau will automatically create hierarchies. You can drag-and-drop to create custom hierarchies.
- **Group:** Use **groups** to combine **dimension** members into higher-level categories. **Groups** are marked with a **paper-clip** icon in the **Data pane**. **Groups**, however, are not dynamic.
- **Union:** Combine two or more tables by appending rows from one table to another. To union your data in Tableau, the tables must come from the same connection.

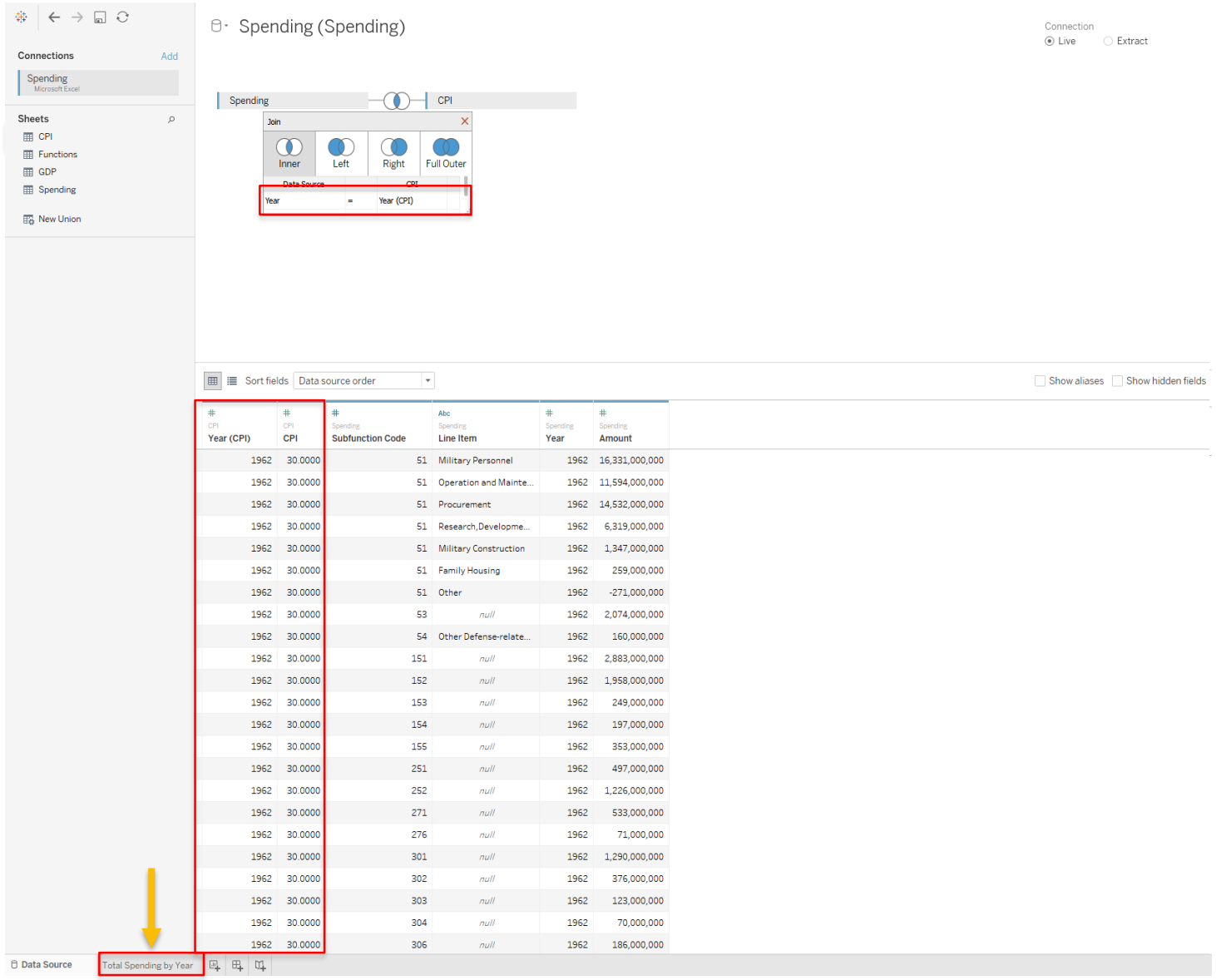
Federal Spending: Annotating your Viz

- Let's build and annotate this graph to emphasize total spending in the most recent year.
- First, connect to the Federal Spending Excel file found in the Datasets folder by dragging and dropping it into Tableau
- Navigate to a new sheet and start by dragging **Amount** onto **rows** and **Year** onto **columns**
- Display **Amount** for the latest year by right clicking the **2015** mark and create an **Annotation**



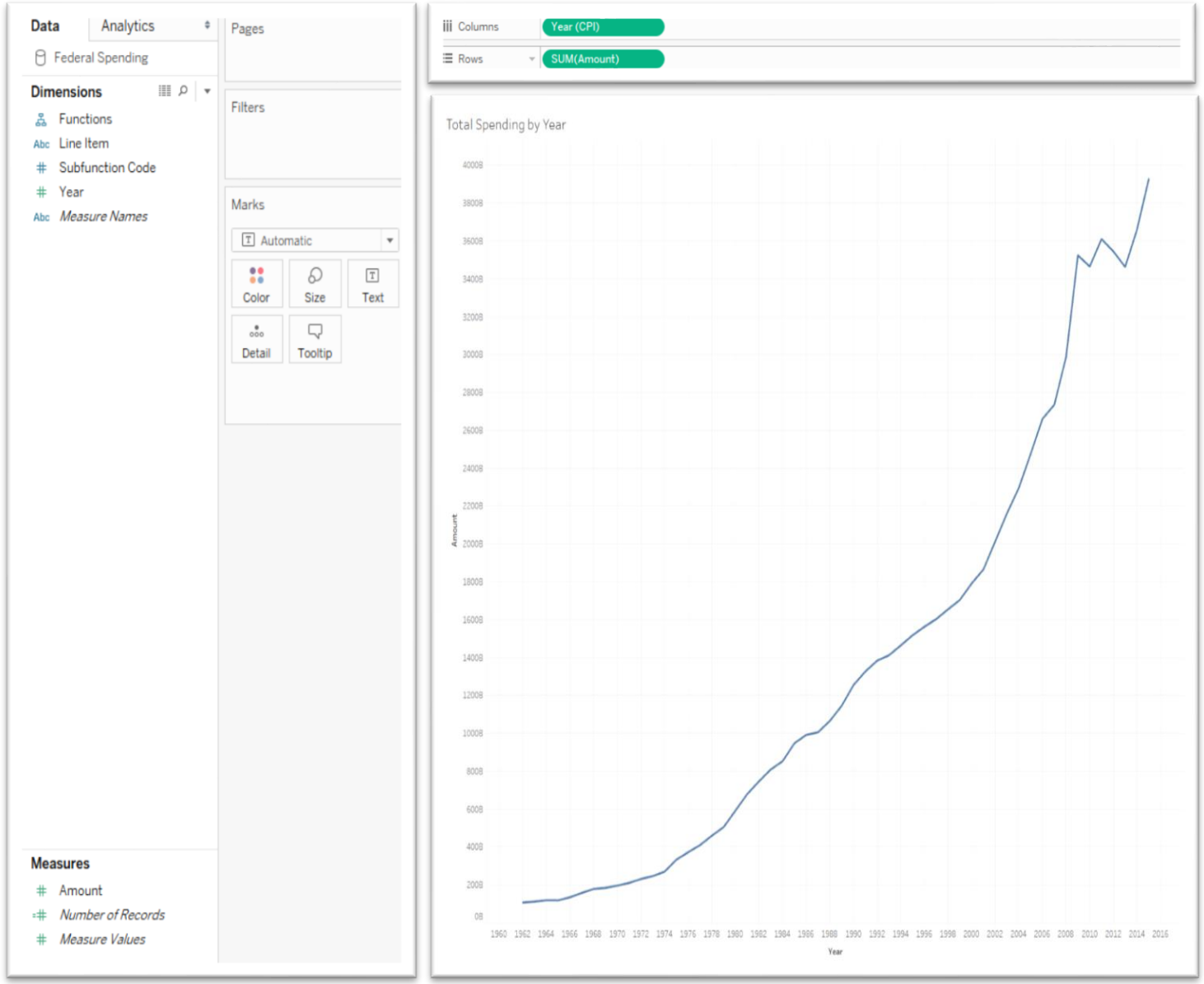
Federal Spending: Joining Data

- Now, let's account for inflation by navigating to the Data Source Tab and joining the **CPI** table
- Notice Tableau Automatically joined the data by **year**
- Navigate back to the Spending Sheet so we can now account for inflation!
- Where did the annotation go?
- Navigate back to the Data source tab and change the Join Type to **Left**



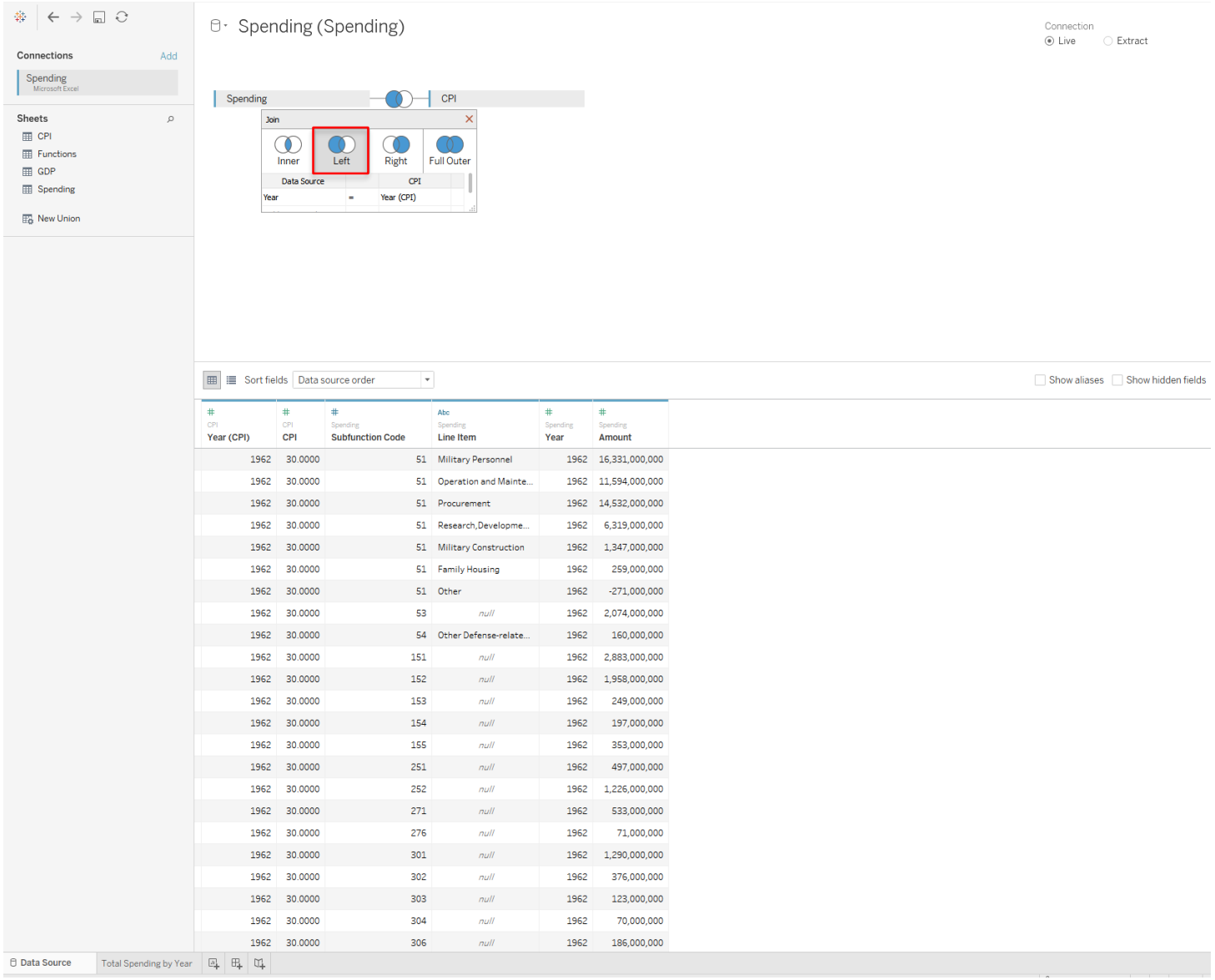
Federal Spending: Joining Data

- Now, let's account for inflation by navigating to the Data Source Tab and joining the **CPI** table
- Notice Tableau Automatically joined the data by **year**
- Navigate back to the Spending Sheet so we can now account for inflation!
- Where did the annotation go?
- Navigate back to the Data source tab and change the Join Type to **Left**



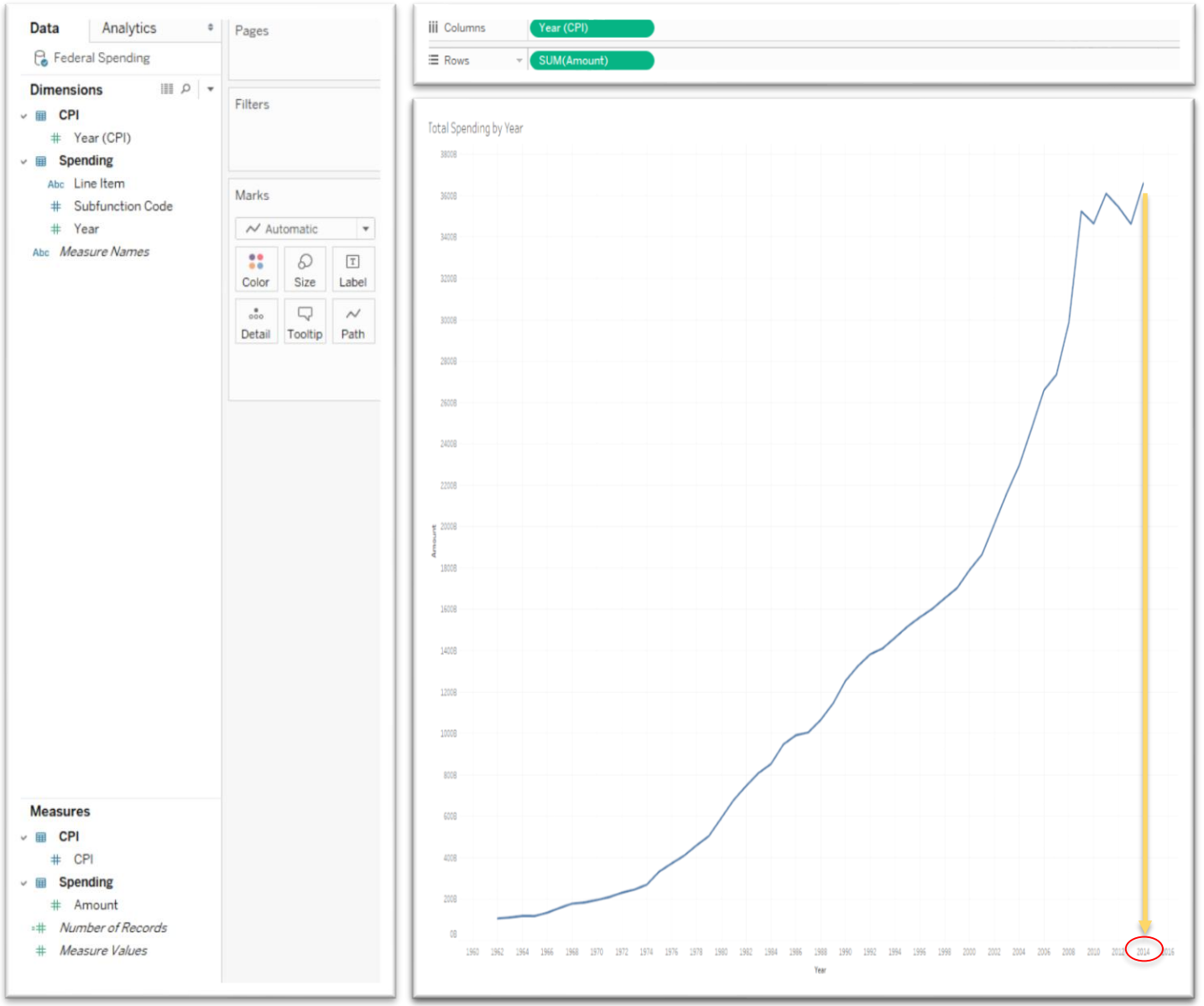
Federal Spending: Joining Data

- Now, let's account for inflation by navigating to the Data Source Tab and joining the **CPI** table
- Notice Tableau Automatically joined the data by **year**
- Navigate back to the Spending Sheet so we can now account for inflation!
- Where did the annotation go?
- Navigate back to the Data source tab and change the Join Type to **Left**



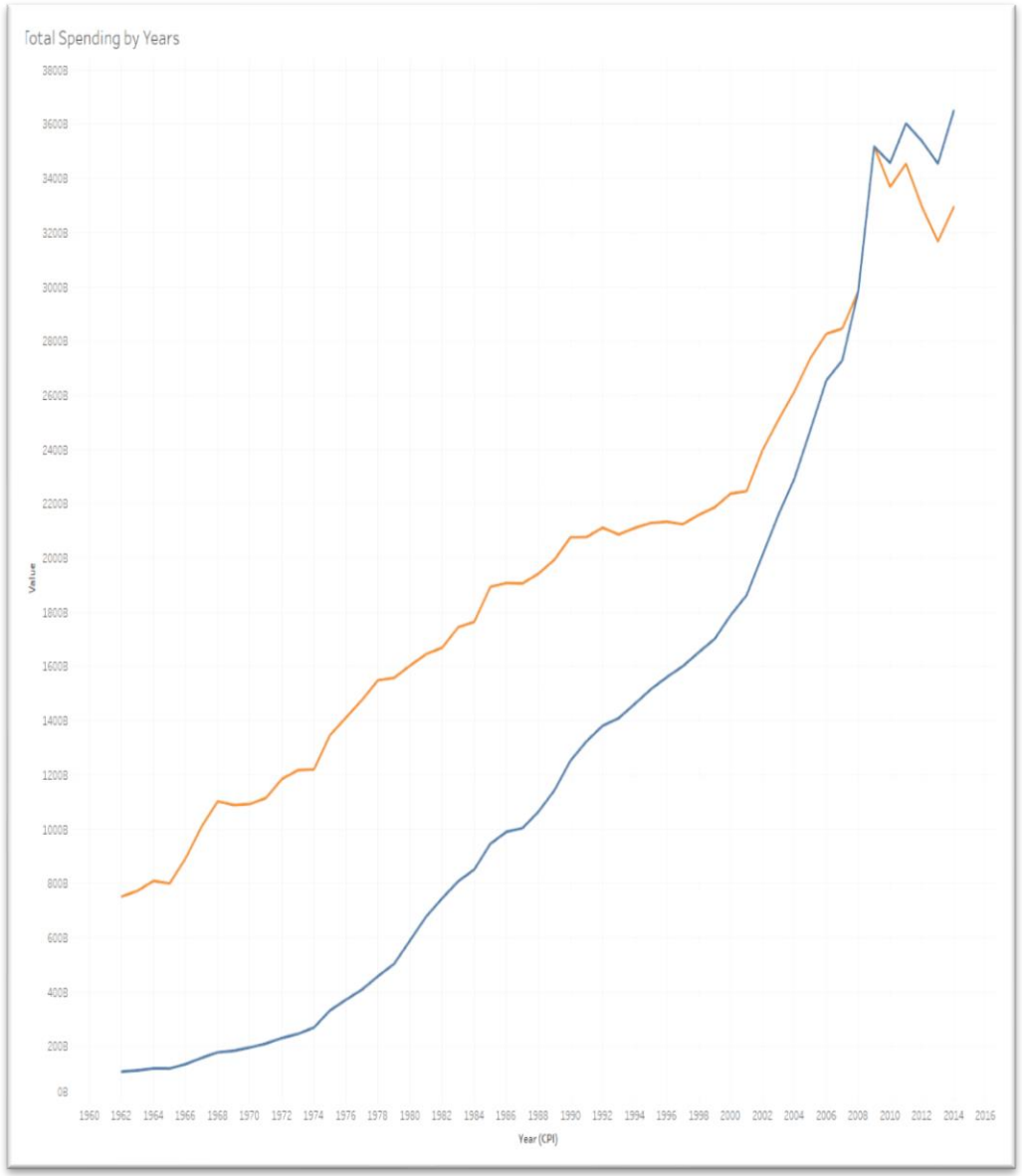
Federal Spending: Joining Data

- Now looking at our graph the Annotation is back! Why?
- Left Join : “Includes **all** values in **left** table and all matches from the right table. Members without matches will show up as nulls on the right.”
- Inner Join: “Includes **only** values with matches in **both** tables”
- It appears as though the CPI table did not include data for year **2015!**



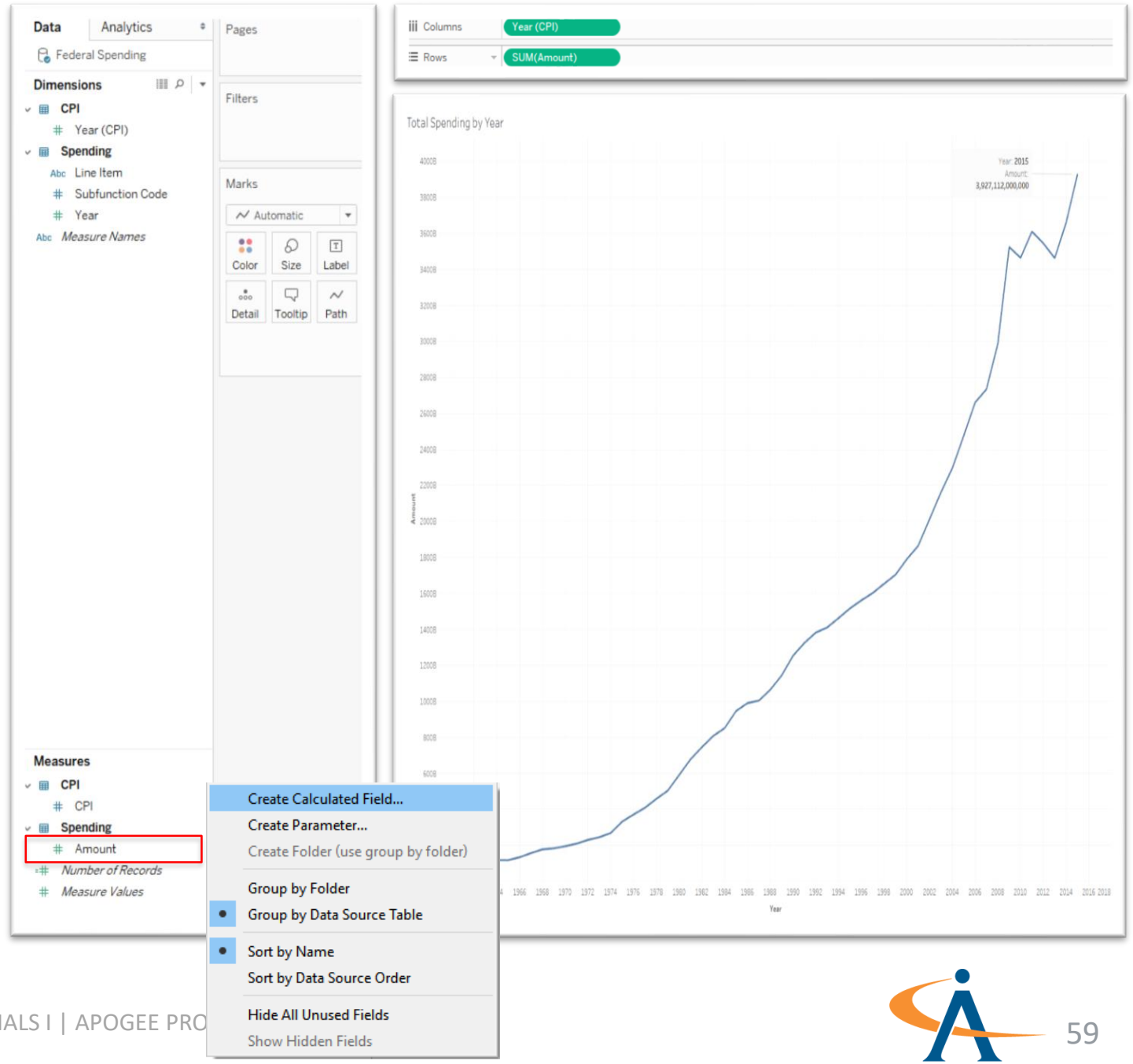
Federal Spending: Calculated Fields

- Create this viz to see how inflation affects Federal Spending
- First, right click the measures pane and create a **calculated field**
- Let's create the calculated field **"Amount In Constant Dollars"** to account for inflation
  - Don't worry too much about this formula, we're just becoming familiar with creating/applying calculations to a viz
- Drag **Amount in Constant Dollars** onto the **Amount** axis to visualize two Measure Values on a single graph



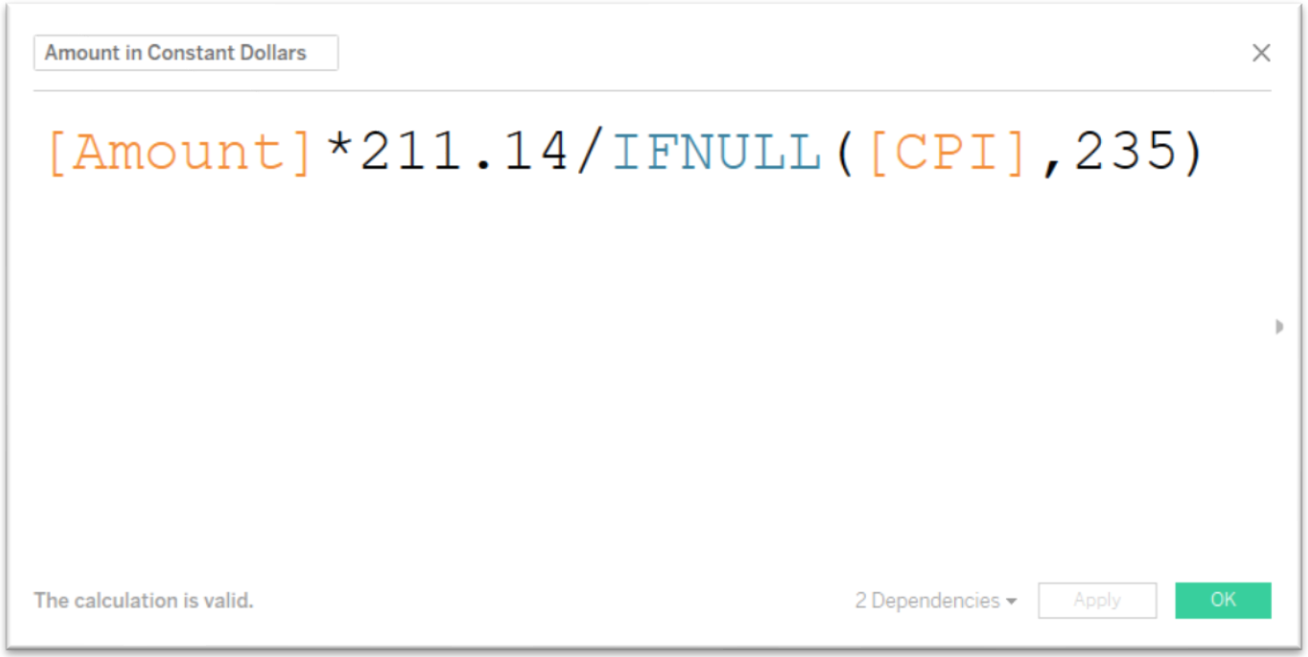
Federal Spending: Calculated Fields

- Create this viz to see how inflation affects Federal Spending
- First, right click the measures pane and create a **calculated field**
- Let's create the calculated field **"Amount In Constant Dollars"** to account for inflation
  - Don't worry too much about this formula, we're just becoming familiar with creating/applying calculations to a viz
- Drag **Amount in Constant Dollars** onto the **Amount** axis to visualize two Measure Values on a single graph



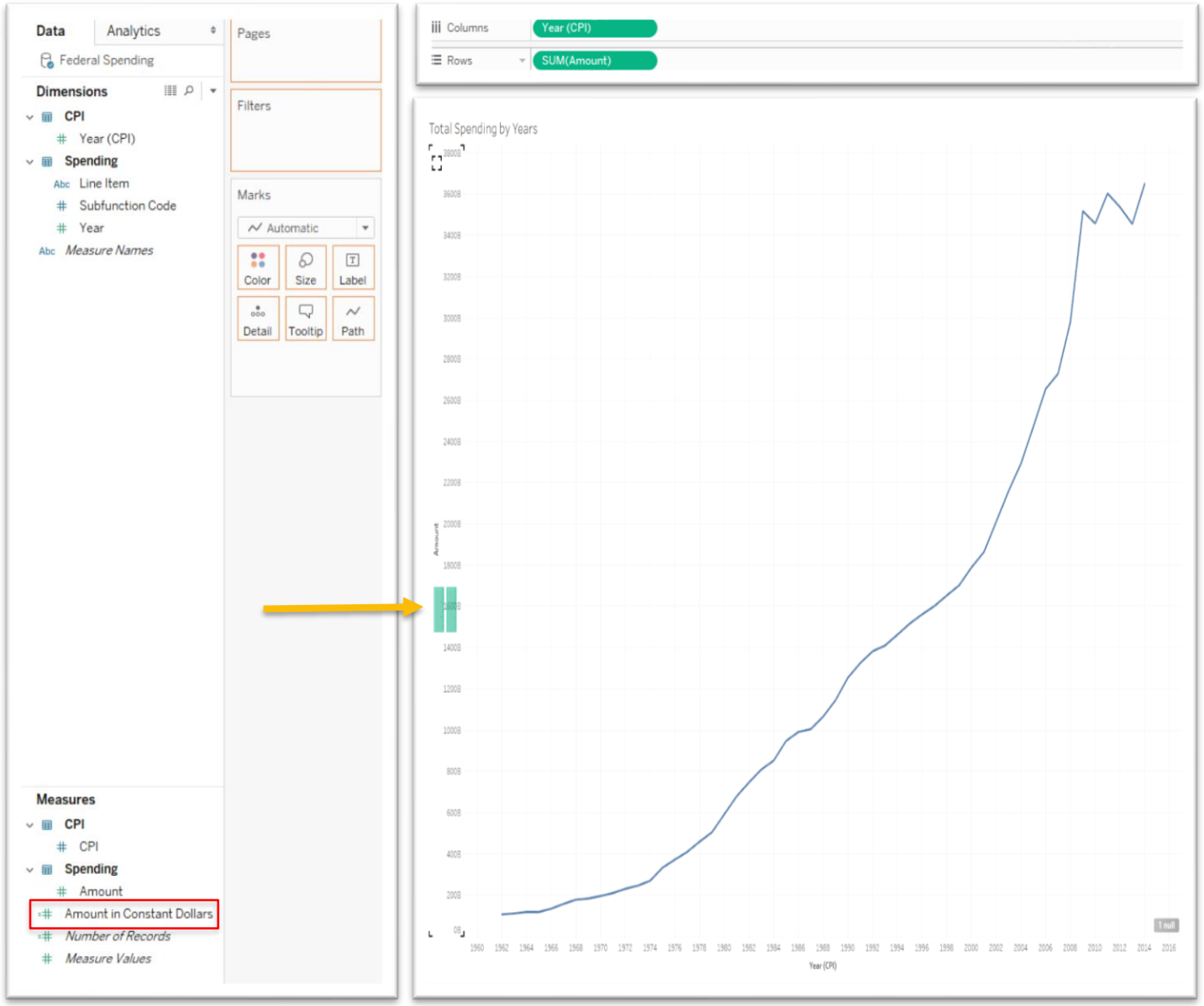
Federal Spending: Calculated Fields

- Create this viz to see how inflation affects Federal Spending
- First, right click the measures pane and create a **calculated field**
- Let's create the calculated field **"Amount In Constant Dollars"** to account for inflation
  - Don't worry too much about this formula, we're just becoming familiar with creating/applying calculations to a viz
- Drag **Amount in Constant Dollars** onto the **Amount** axis to visualize two Measure Values on a single graph



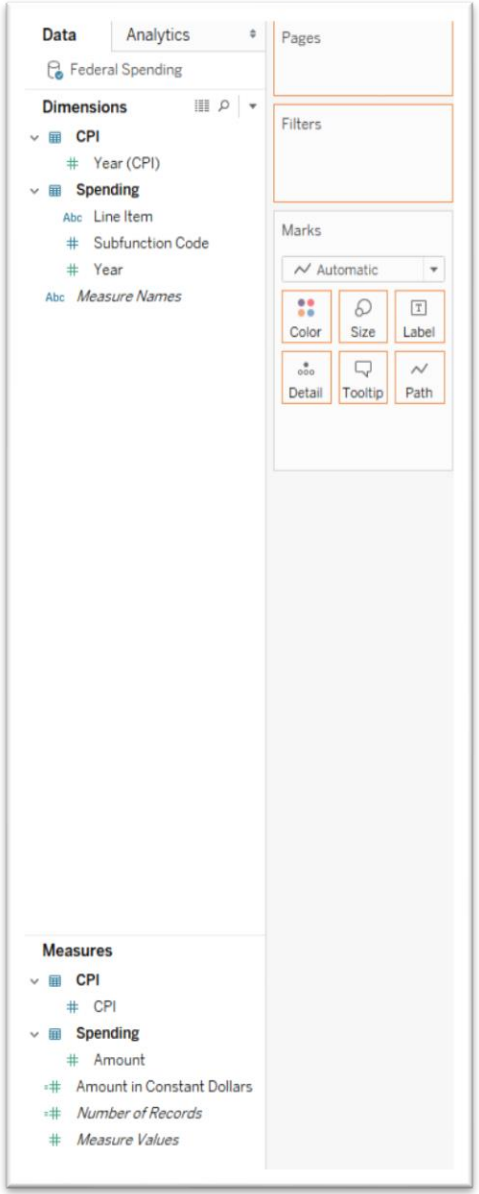
Federal Spending: Calculated Fields

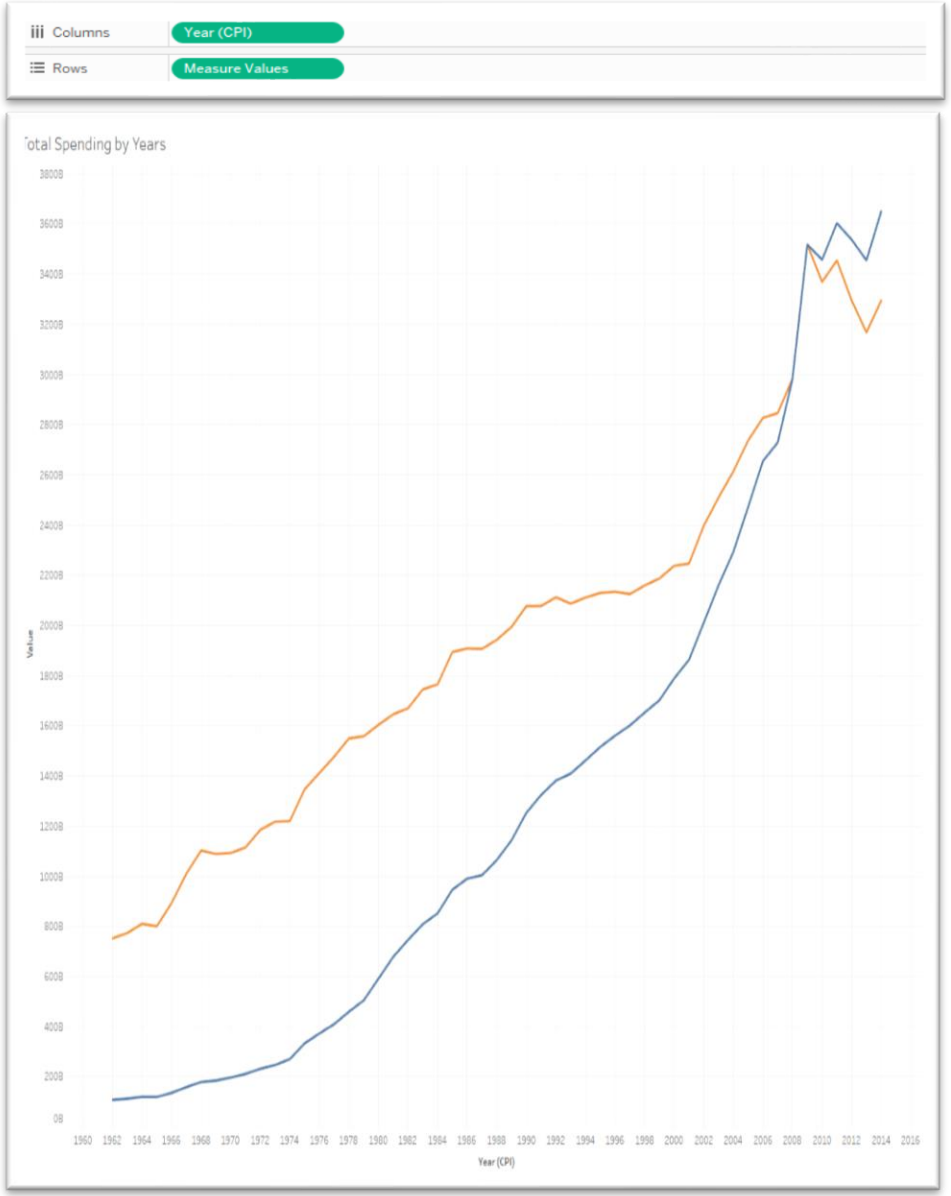
- Create this viz to see how inflation affects Federal Spending
- First, right click the measures pane and create a **calculated field**
- Let's create the calculated field **"Amount In Constant Dollars"** to account for inflation
  - Don't worry too much about this formula, we're just becoming familiar with creating/applying calculations to a viz
- Drag **Amount in Constant Dollars** onto the **Amount** axis to visualize two Measure Values on a single graph



Federal Spending: Calculated Fields

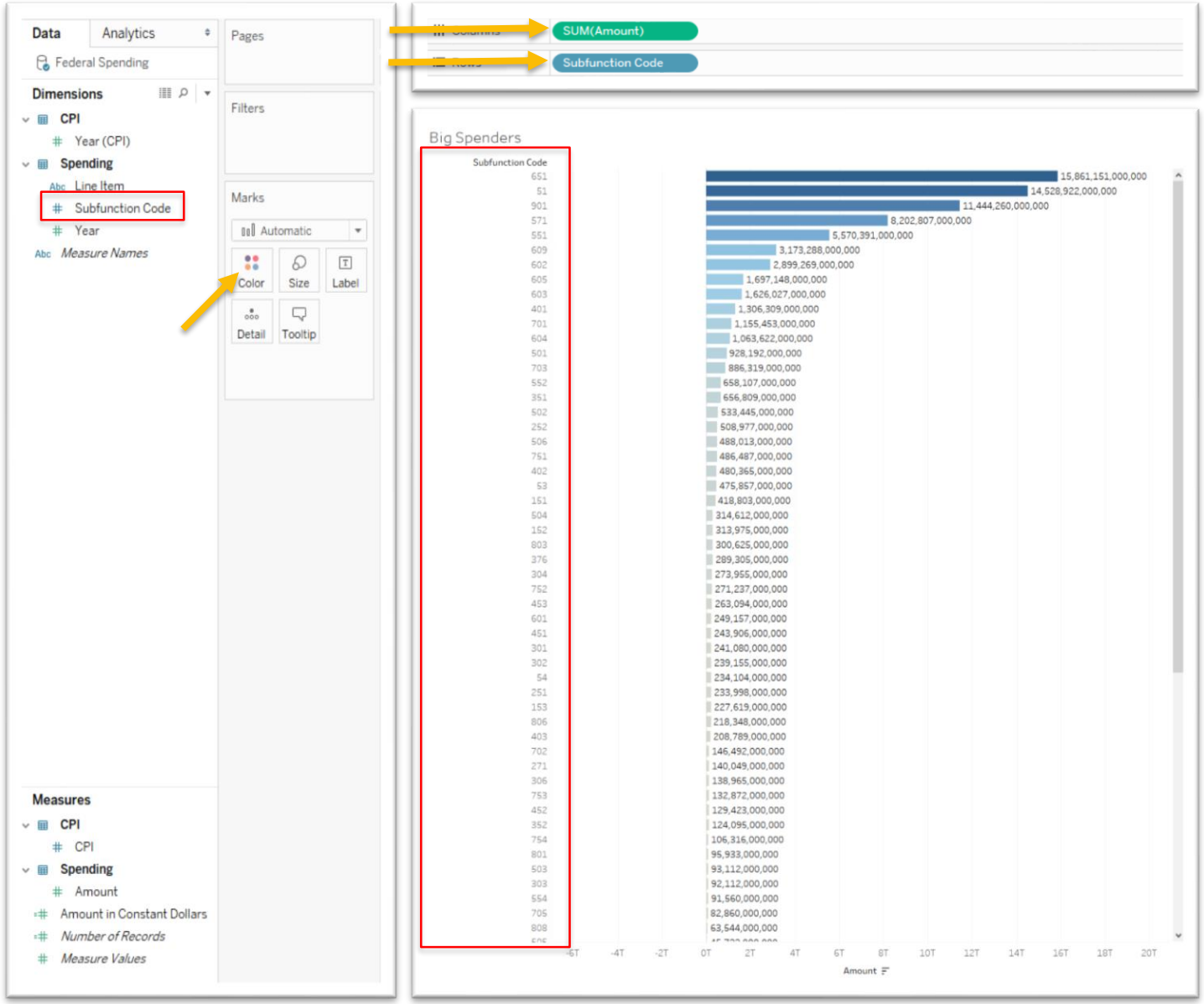
- Create this viz to see how inflation affects Federal Spending
- First, right click the measures pane and create a **calculated field**
- Let's create the calculated field **"Amount In Constant Dollars"** to account for inflation
  - Don't worry too much about this formula, we're just becoming familiar with creating/applying calculations to a viz
- Drag **Amount in Constant Dollars** onto the **Amount** axis to visualize two Measure Values on a single graph





Federal Spending:
Spending by Subfunction

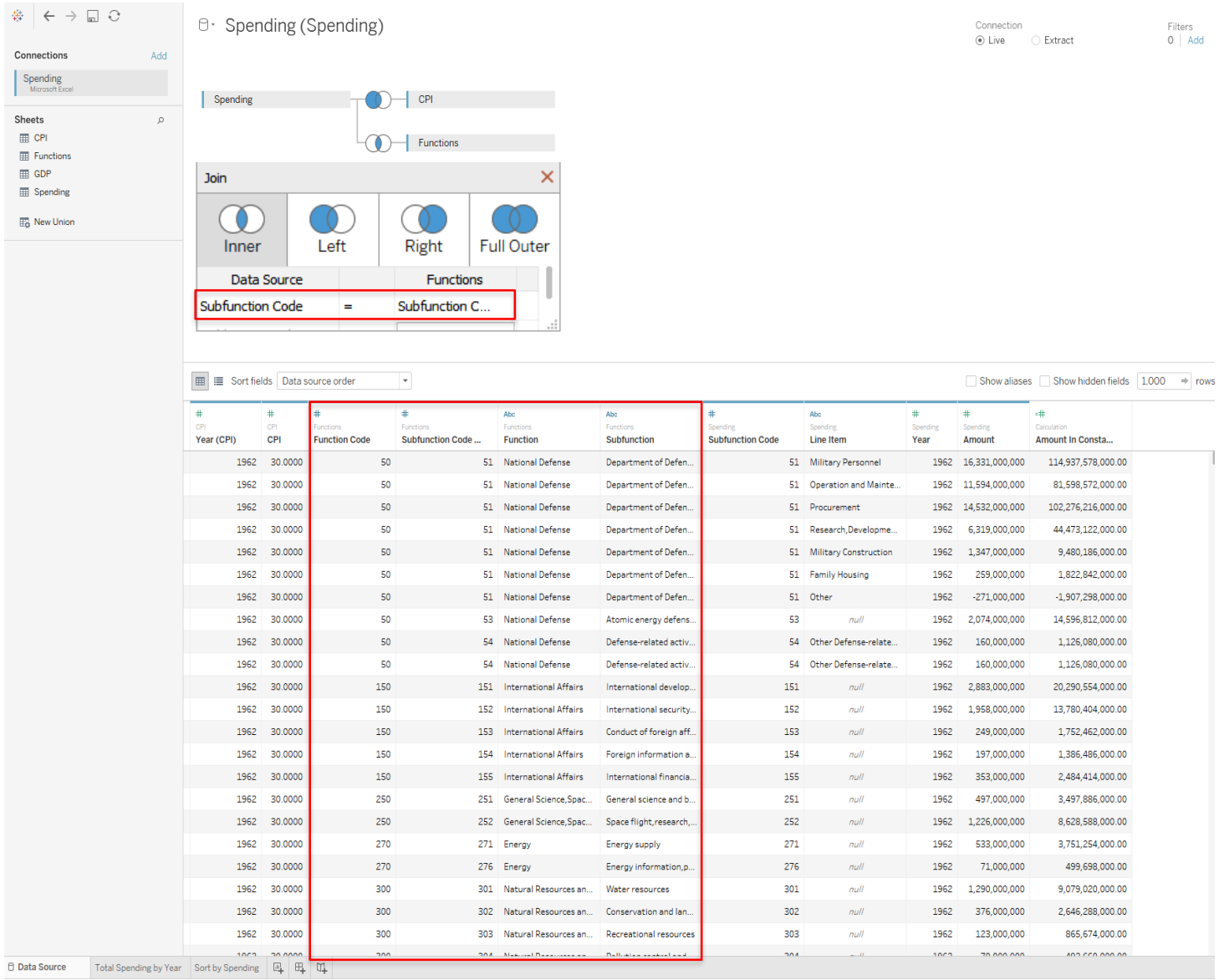
- Now, lets find out who the top spending contributors are
- Drag **Subfunction Code** to **rows** and **Amount** to **columns**
- Perform a quick sort and tableau automatically sorts **Amount** in descending order
- Drop Amount on **color** to generate a **diverging palette** centered around 0
- Looks like we are going to need more information on these spenders!



Federal Spending:

Spending by Subfunction

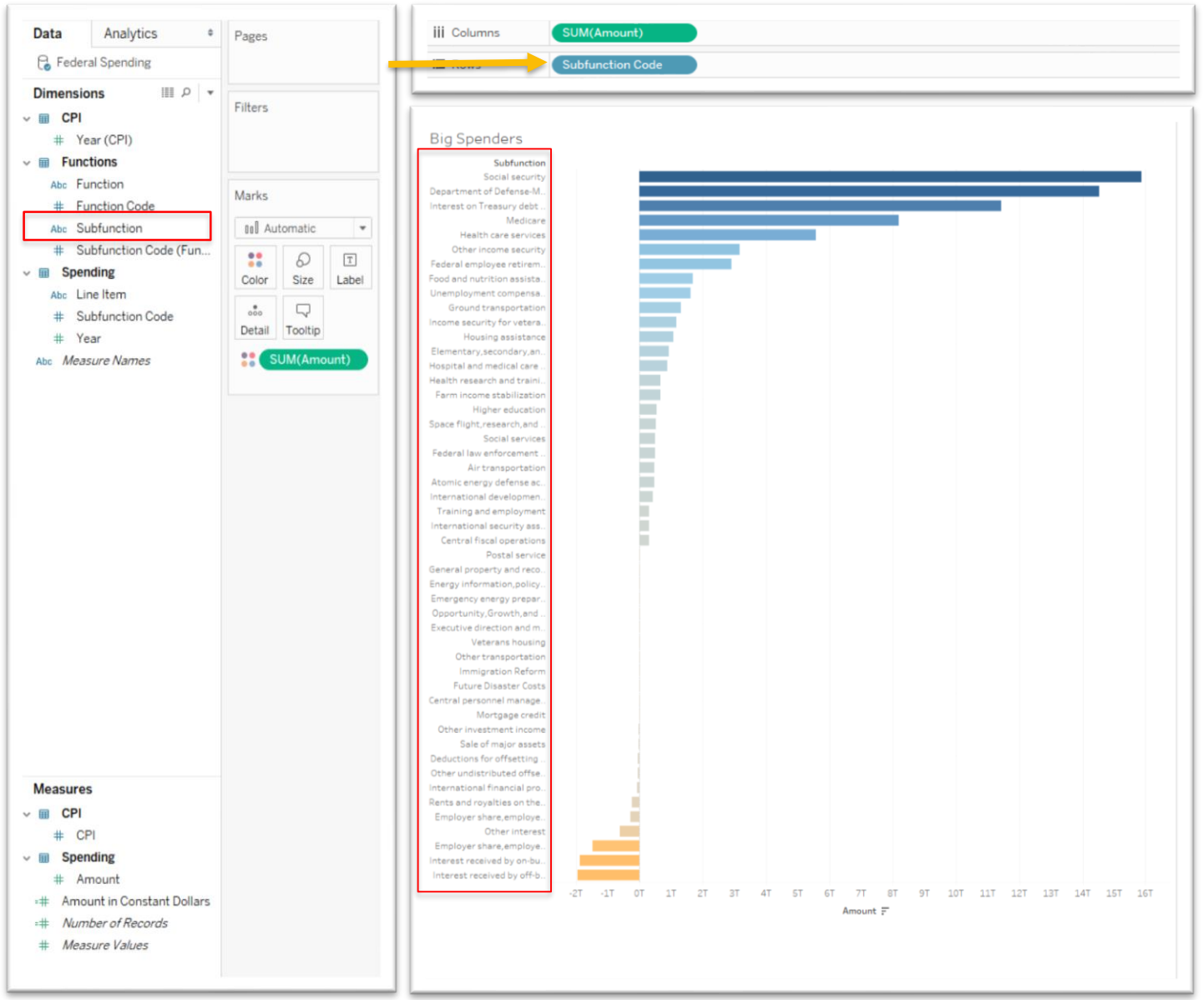
- Navigate back to the data source and join in the **Functions** sheet to identify Subfunction by name
- Tableau automatically creates a join using **Subfunction Code** as the **unique identifier**
- Navigate back to the Sort by Spending sheet and drag **Subfunction** to rows
- Spending can now be identified by subfunction



Federal Spending:

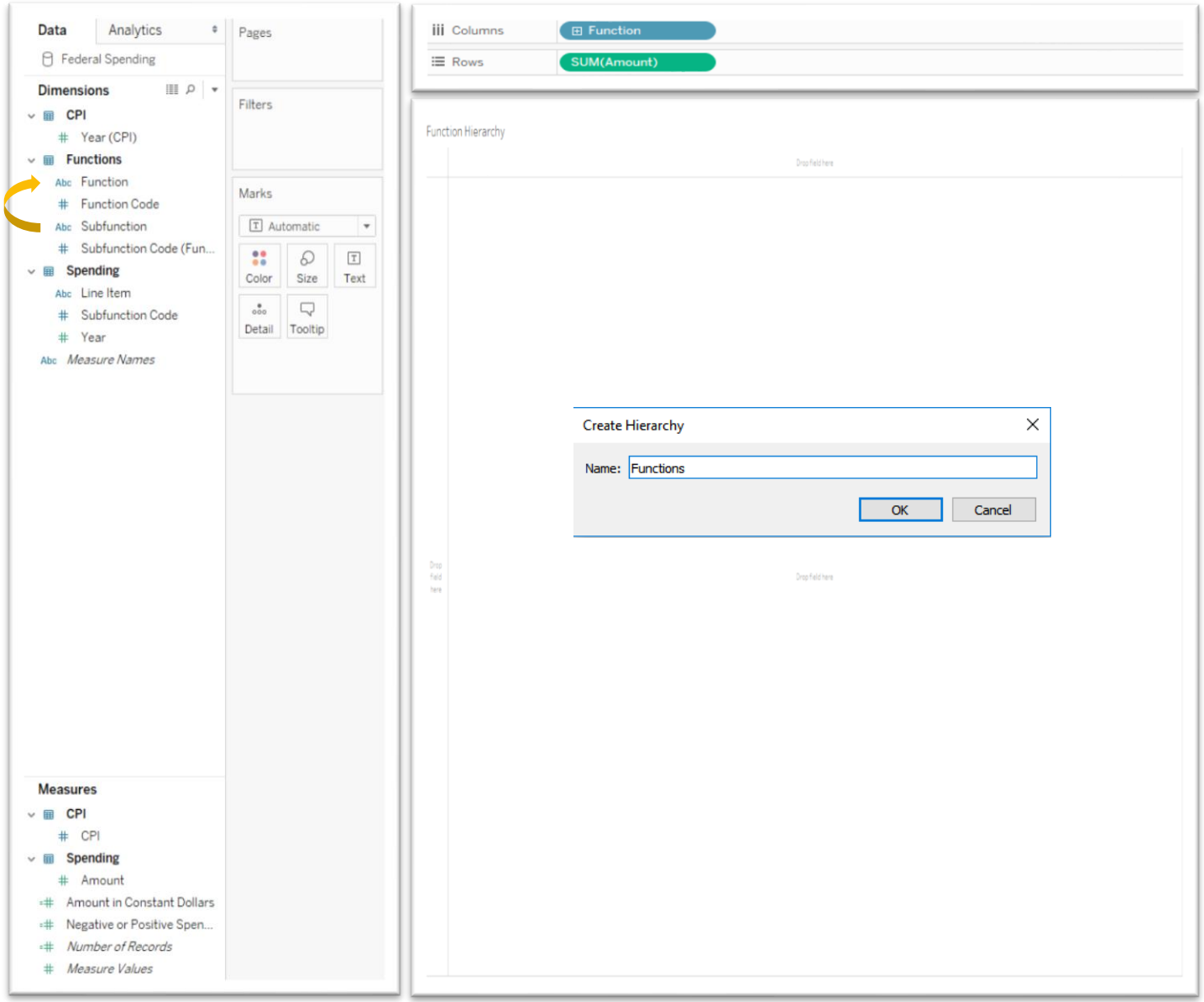
Spending by Subfunction

- Navigate back to the data source and join in the **Functions** sheet to identify Subfunction by name
- Tableau automatically creates a join using **Subfunction Code** as the **unique identifier**
- Navigate back to the Sort by Spending sheet and drag **Subfunction** to rows
- Spending can now be identified by subfunction



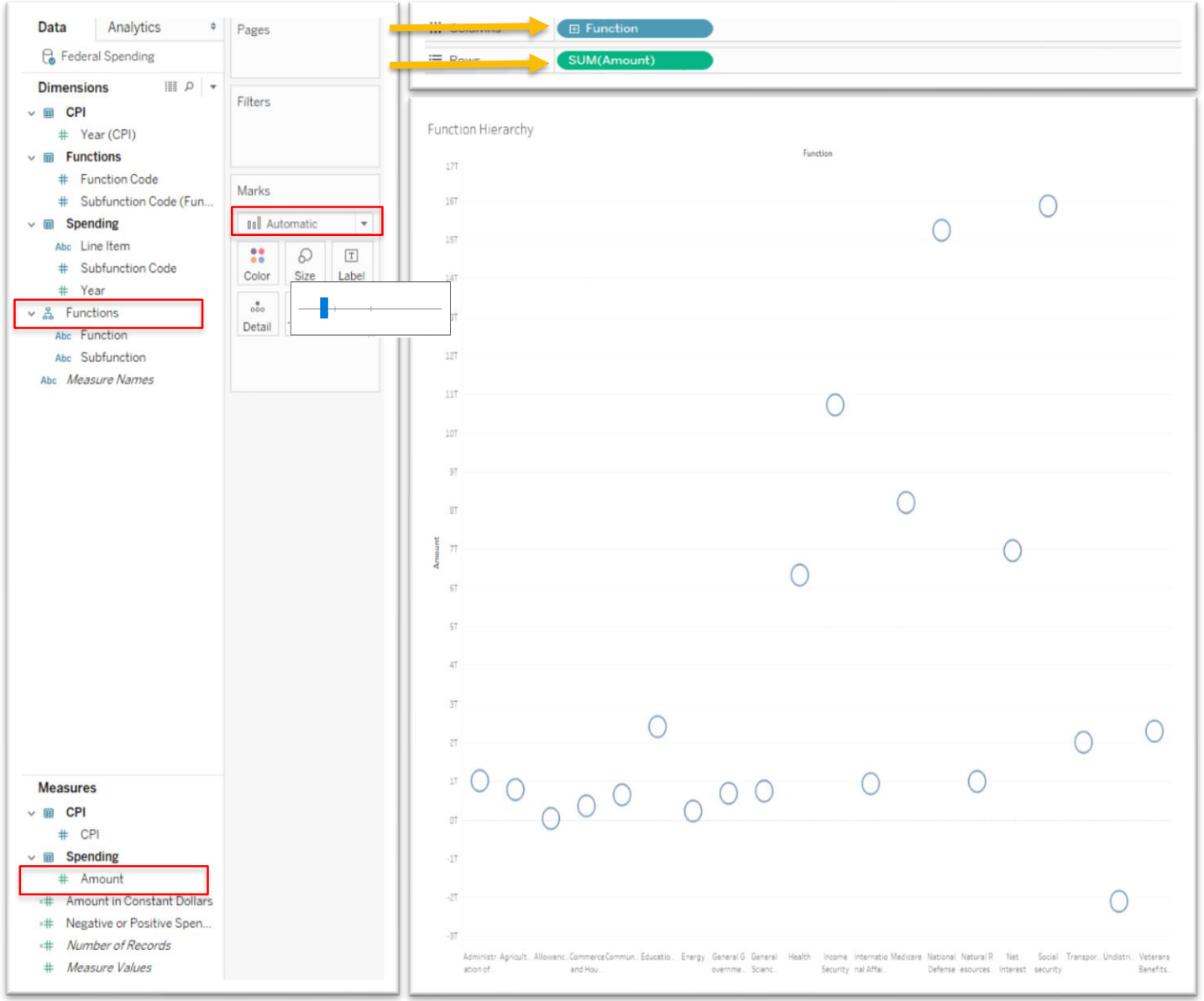
Federal Spending: Hierarchies

- Now lets create a hierarchy to see how “negative” spending affects our data.
- First, drag **Subfunction** onto **Function** in the dimensions shelf
- Rename this hierarchy to **“Functions”**
- Drag the **Functions** hierarchy to columns and **Amount** rows
- Create **hollow circles** by changing the mark type to **shape** and reducing the **size**



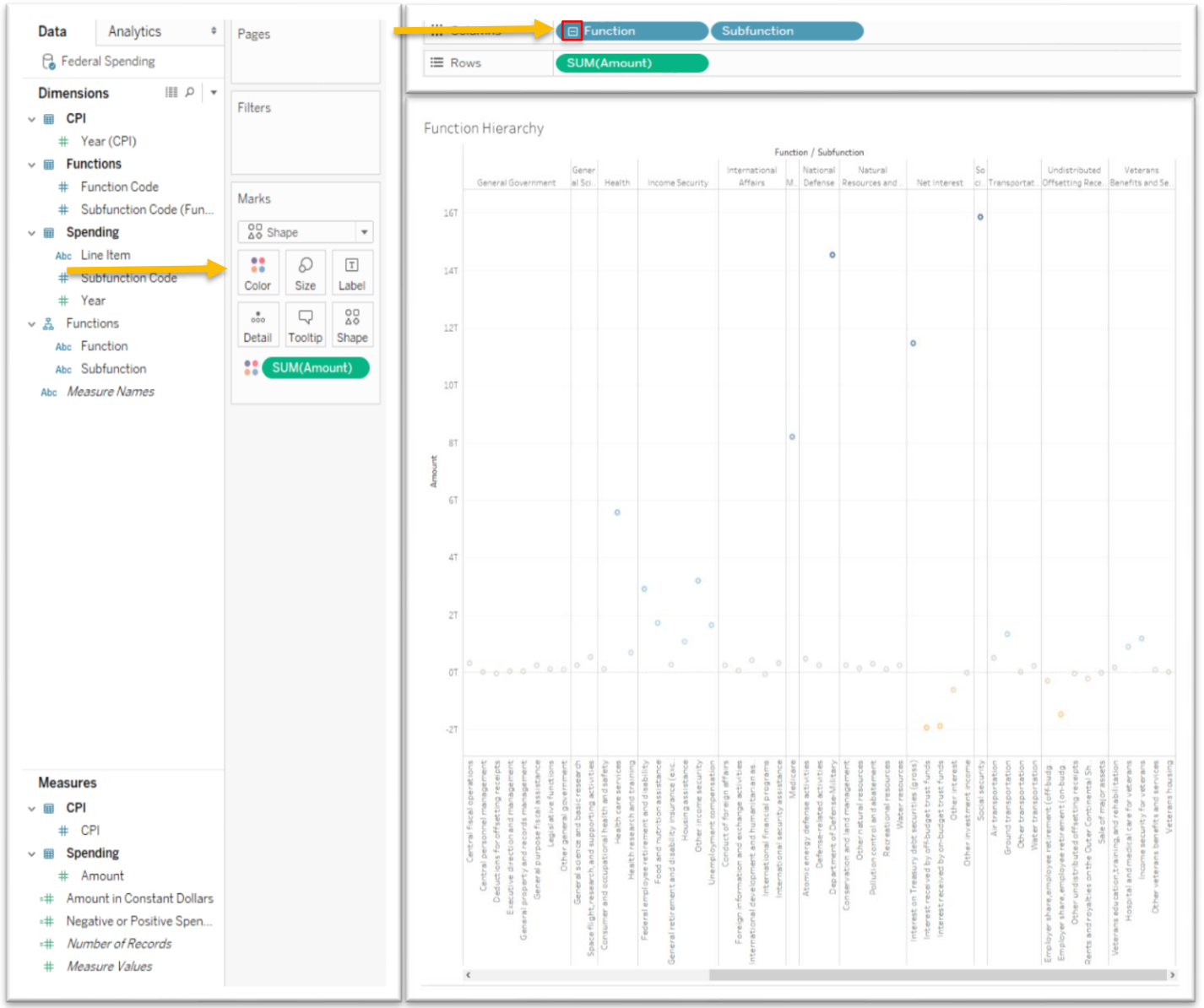
Federal Spending: Hierarchies

- Now lets create a hierarchy to see how “negative” spending affects our data.
- First, drag **Subfunction** onto **Function** in the dimensions shelf
- Rename this hierarchy to **“Functions”**
- Drag the **Functions** hierarchy to columns and **Amount** rows
- Create **hollow circles** by changing the mark type to **shape** and reducing the **size**



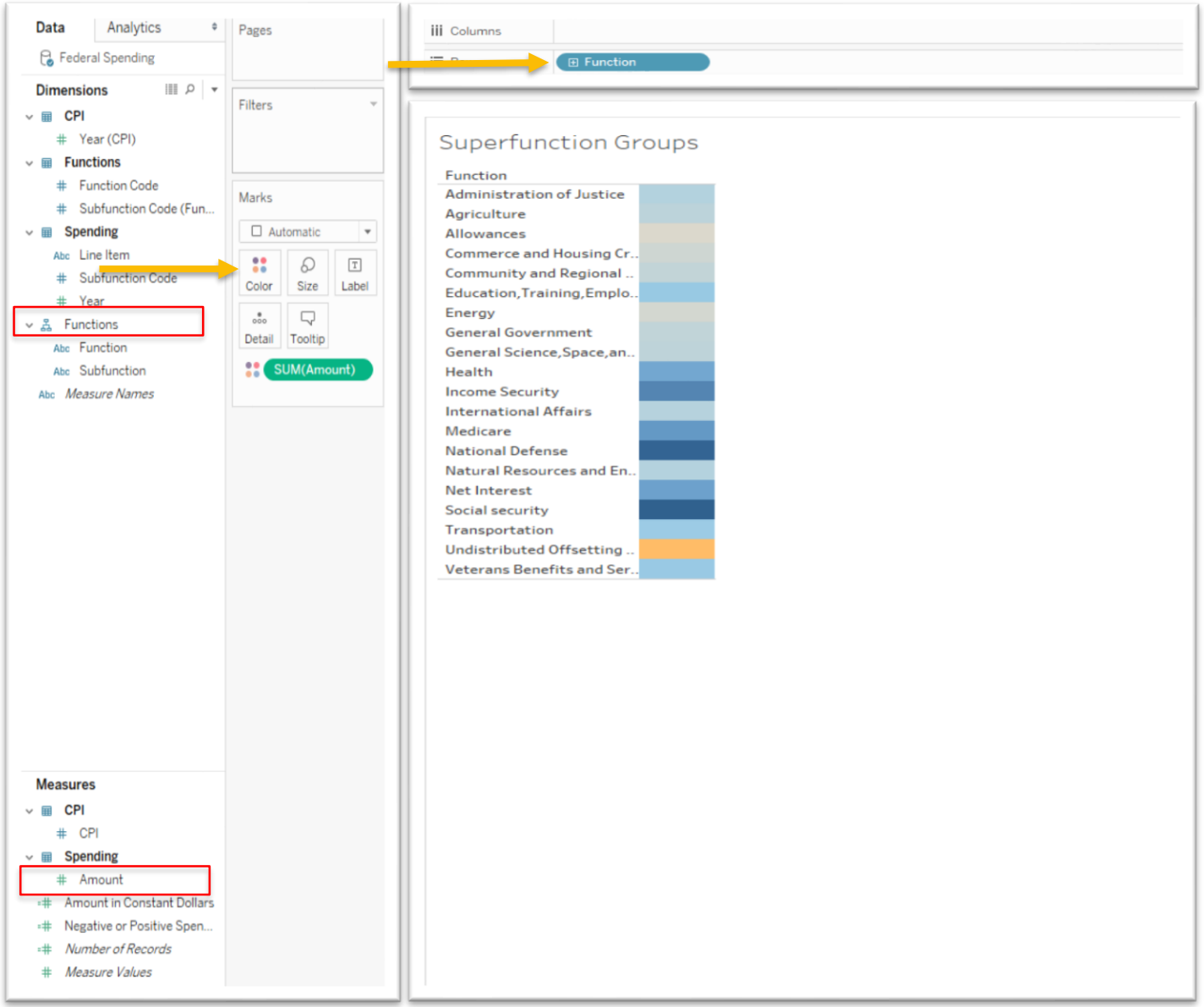
Federal Spending: Hierarchies

- Drag **Amount** to Color to create a diverging palate.
- Drill **Up** and **Down** the hierarchy to see different Levels of Spending



Federal Spending: Groups

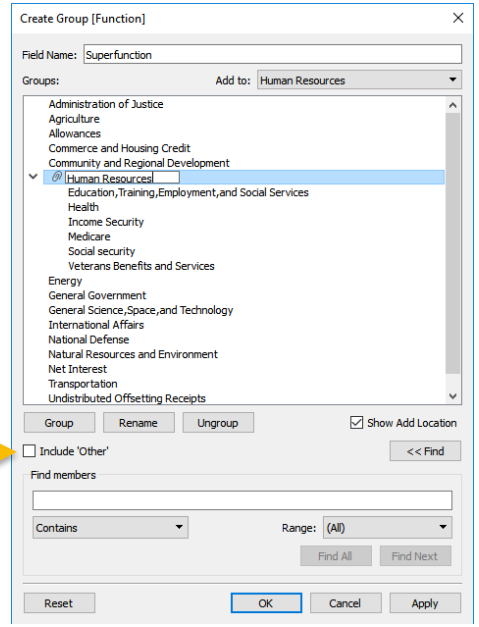
- Let's aggregate the data and group Functions into Superfunctions
- First, drag **Functions** to rows and **Amount** to the Colors Card



Federal Spending: Groups

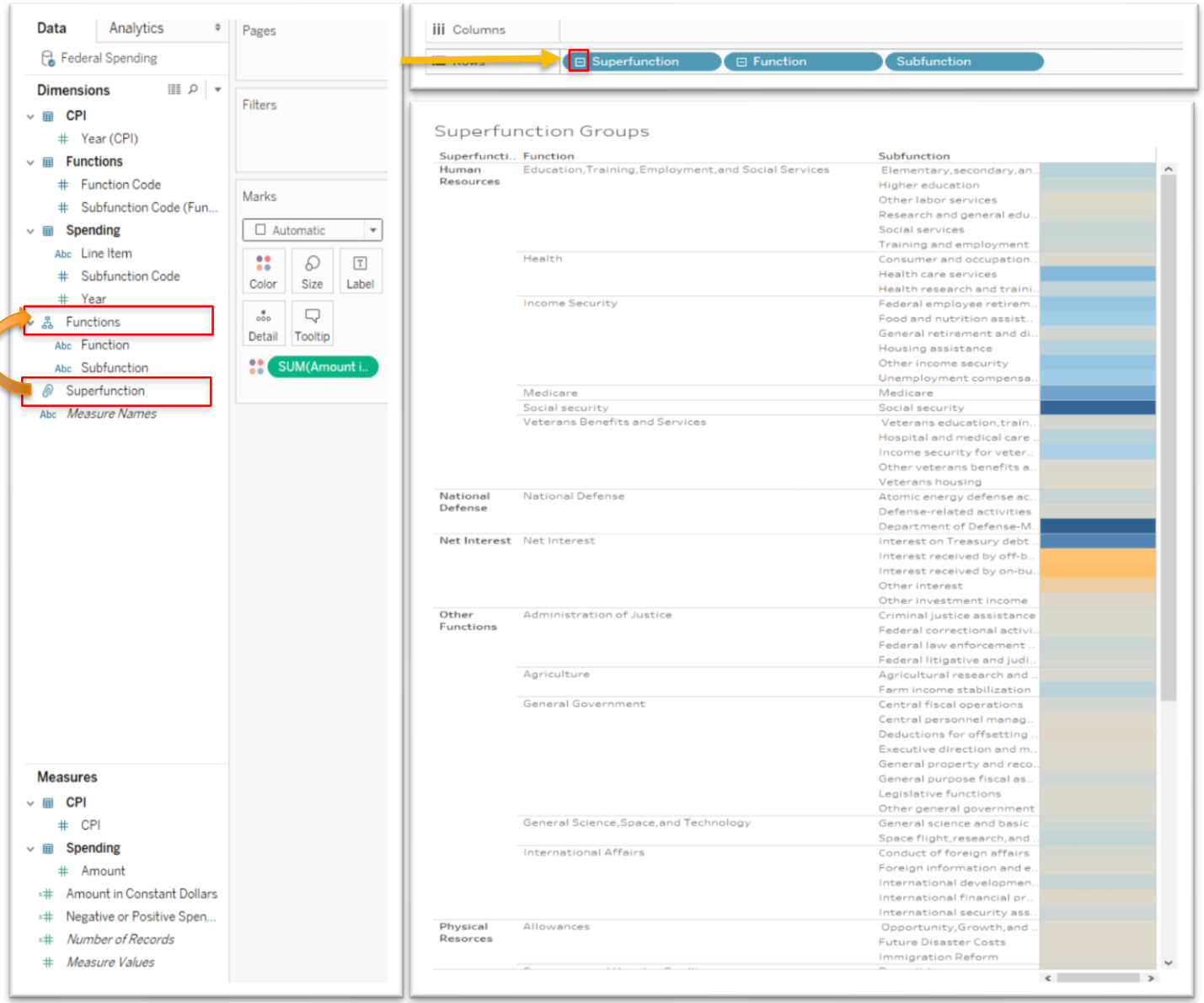
- Right click on **Function** in the dimensions pane and **Create a Group**
- Change the field name to "**Superfunction**"
- Now, group all **Functions** that belong to the **HR Superfunction** and rename it to **"Human Resources"**
- Create the Physical Resources **Superfunction** and then check the include other box to group the remaining functions

Superfunction	Function
Human Resources	• Education, Training, Employment, and Social Services • Health • Income Security • Medicare • Social Security • Veterans Benefits and Services
Physical Resources	• Allowances • Commerce and Housing Credit • Community and Regional Development • Energy • Natural Resources & Environment • Transportation



Federal Spending: Groups

- Locate **Superfunction** on the **Dimensions Pane** and drag it to the **Functions Hierarchy**
- Add the **Functions** hierarchy to the **rows** shelf. Then expand it to drill down into the data



Check your Understanding

- Does a left join exclude any data from the left table?
- Can you drag individual fields from a hierarchy to the view?

Let's Review!

- Joins allows deeper analysis when all of the data is not included in one table
- Calculations can be created to manipulate existing data
- Quick sorts allow for fast analysis
- Creating a hierarchy allows you to drill down data and find anomalies
- Grouping data allows for higher aggregation

Exercise 2

U.K. Car Accidents

- Exercise Objectives:
  - Practice sorting and grouping fields
  - Learn how to Union Data

Exercise Materials

Exercise 2 UK Accidents-Starter.twbx

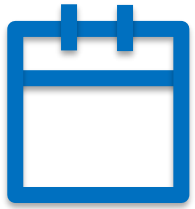
Exercise 2 UK Accidents-Solution.twbx

IT'S A DATE!

Date Formatting in Tableau

How does Tableau handle dates?

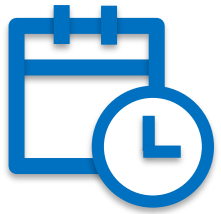
- Dates represent a point in time specific to the day level of detail.
- Every date contains three mandatory components: month, day, year
  - January 20, 2017 is a date
  - 1984 is **not** a date
  - December 1941 is **not** a date



- When connecting to Excel or text data sources, you may have to change the datatype from **String** to **Date** if Tableau couldn't be certain all the values are valid dates (or use the **DateParse()** or **MakeDate()** functions)

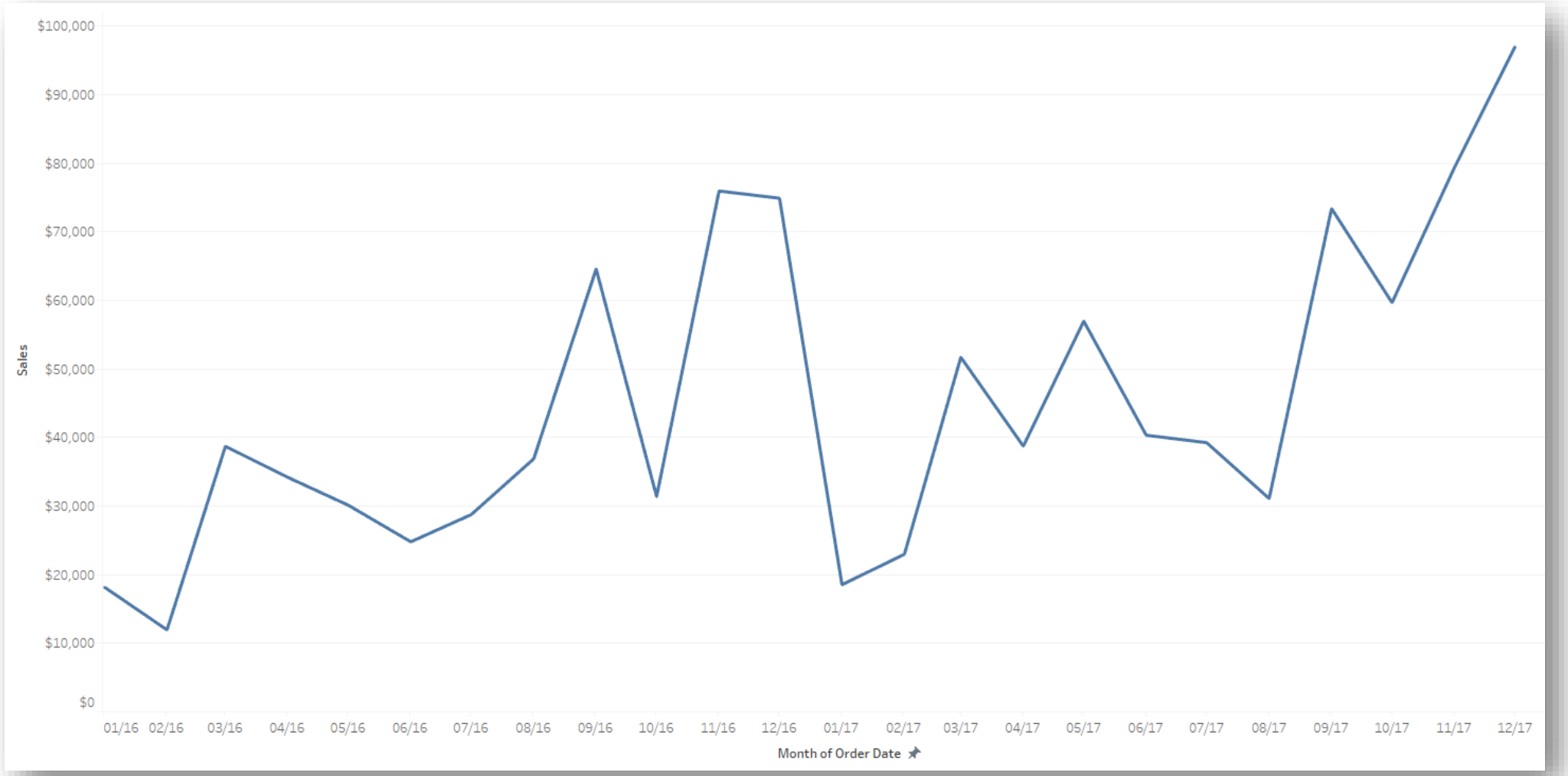
What is a DateTime?

- Date Times must have all the components of dates, but also include: **hour**, **minute**, **second**.
- Unspecified trailing components, such as seconds, are treated as zero.
  - January 20, 2017 10:30 **is** a **datetime**

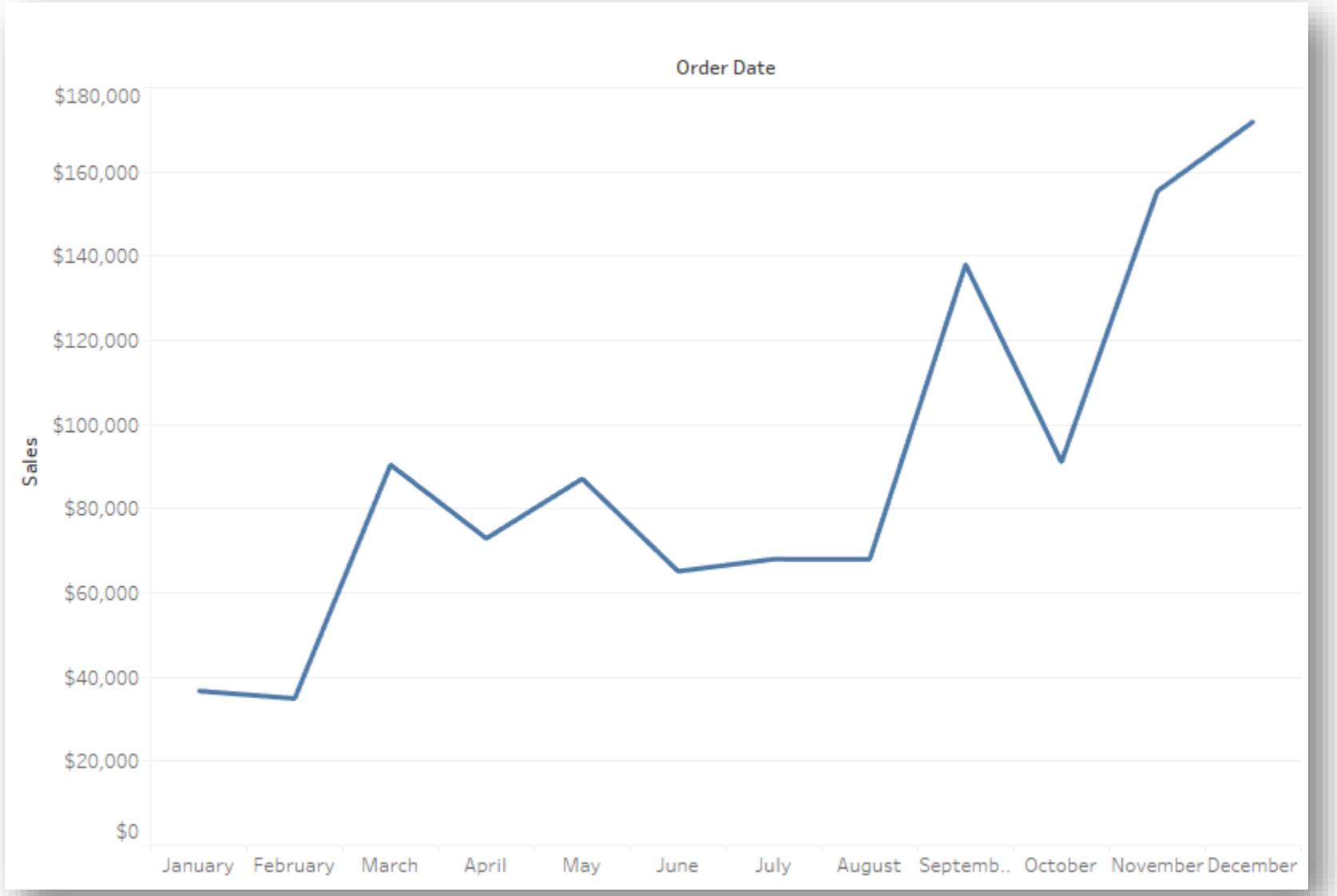


- If all the values lack time information (i.e. have a time of 00:00:00), then changing to a Date type is more efficient and avoids cluttering controls with time detail

Truncated Date

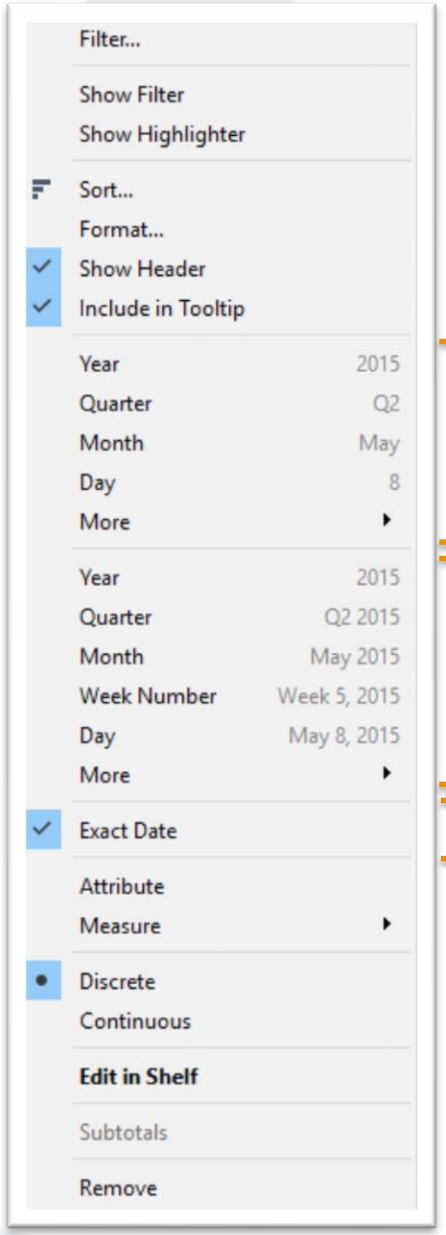


Date Part


Order Date
Sales
$180,000
$160,000
$140,000
$120,000
$100,000
$80,000
$60,000
$40,000
$20,000
$0
January
February
March
April
May
June
July
August
Septemb..
October
November
December

Date Menu

- When you place a date/datetime field on a shelf, you can specify to:
  - **extract** part of the date/datetime value
  - **truncate** the value to a particular level
  - or leave the value **unaltered** (exact date)
- **You also can treat the resulting value as discrete or continuous**
- Or you can treat the field as a measure using aggregation functions like min() and max()
- These combinations give you tremendous flexibility in how you handle dates

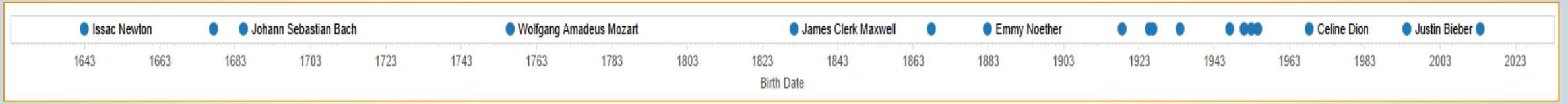


Date Parts
Act as Discrete Fields by default

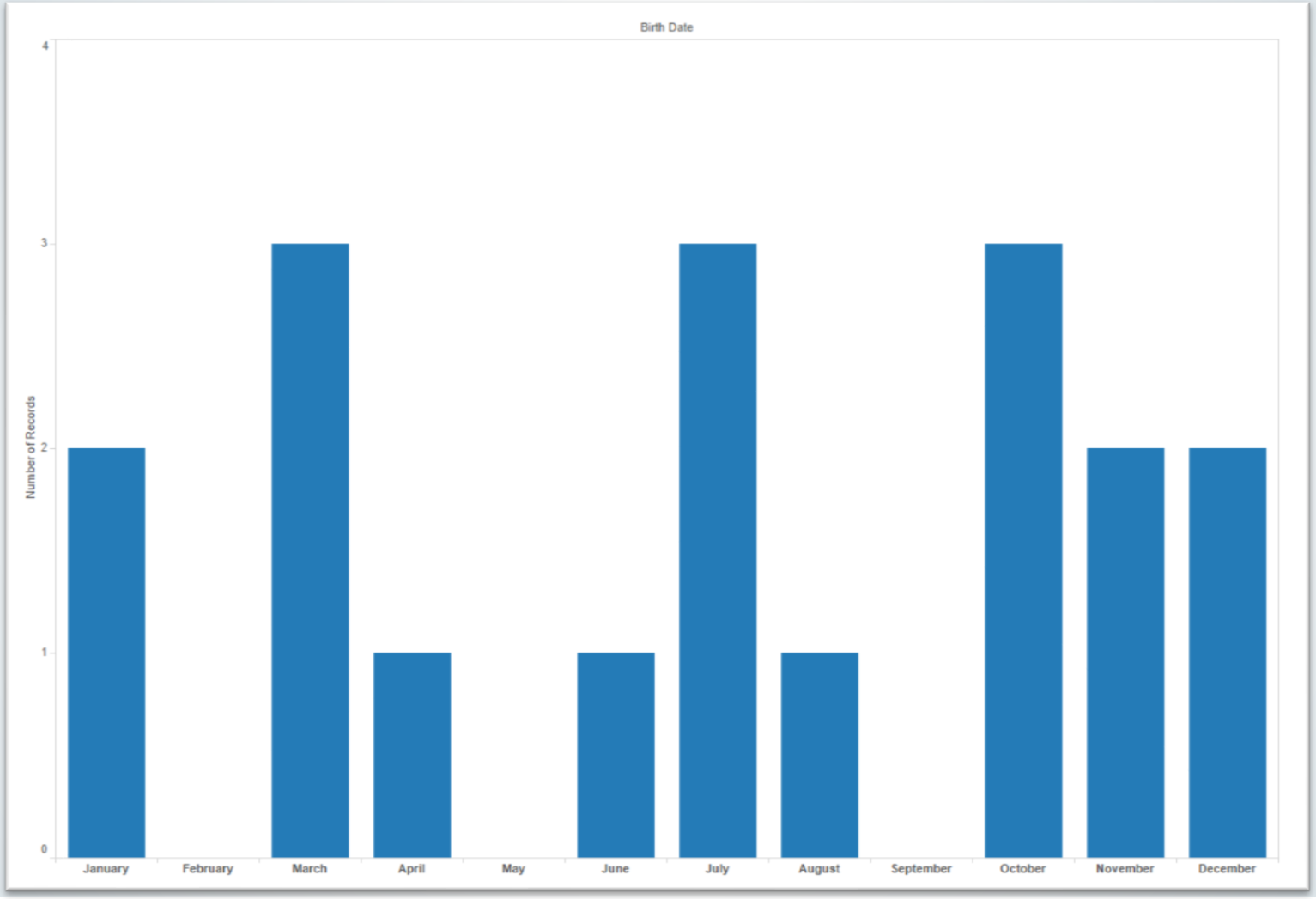
Truncated Dates
Act as Continuous Fields by default

Unaltered Date
Act as Continuous Fields by default

Quiz 1: Recreate this Viz!



Quiz 2: Recreate this Viz!



CREATING THE VIZ YOU WANT!

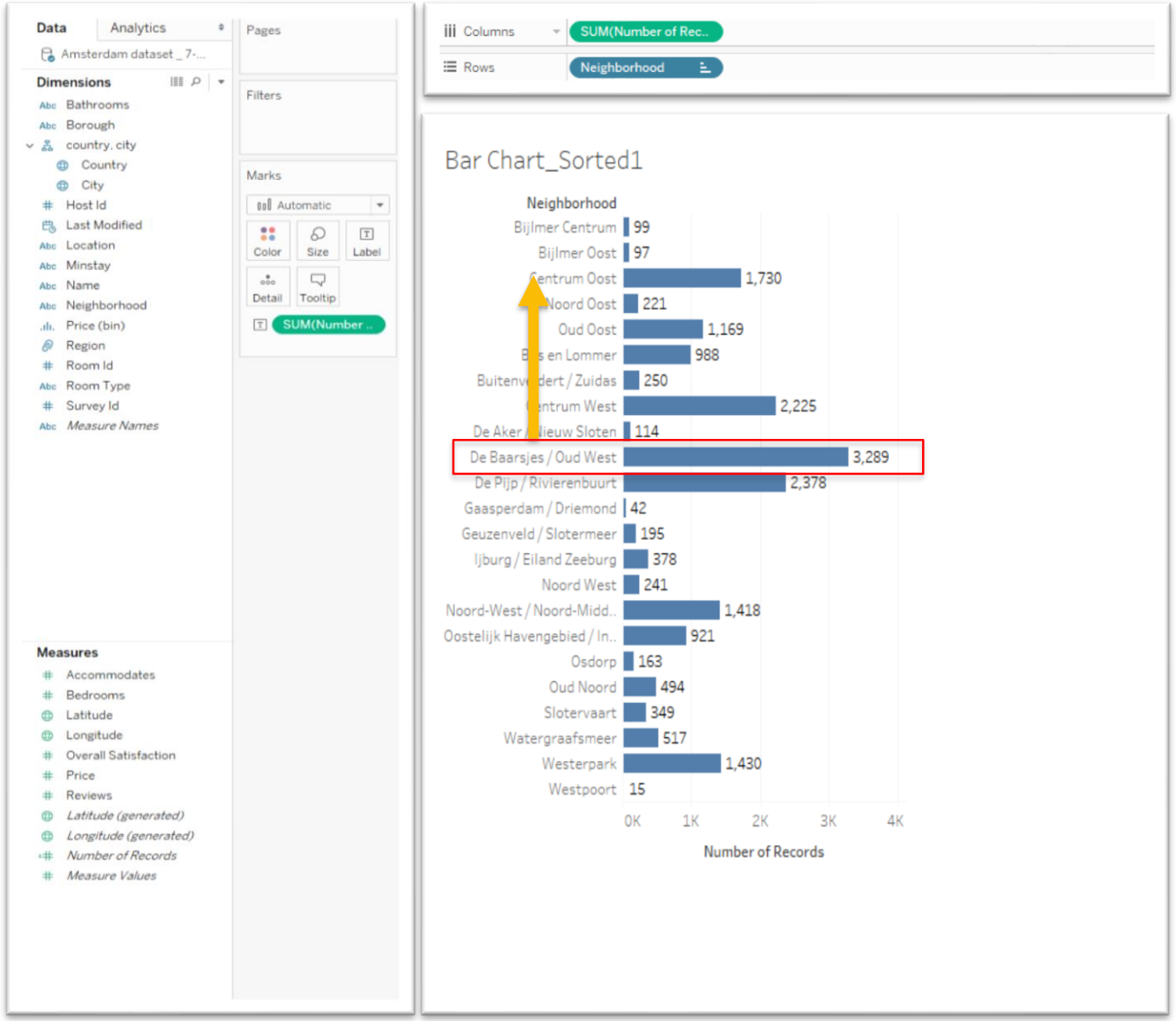
Formatting your Data

Concepts

- **Advanced Sort:** Tableau sorts data by arranging dimensions in a specified order. Each dimension that appears in the worksheet can be sorted independently of any other dimension. You can choose to **sort by** the data source order, alphabetic order, a field or manually from the sort dialog box.
- **Filters:** Tableau **filters** provide the user the ability to restrict the data that is available in individual views or even the entire data source by a **dimension**, **measure** or **set.**
- **Bins:** Tableau **bins** are useful in creating a **range** of continuous data, similar to 'buckets' that makes it possible to visualize a continuous field like a discrete field.
- **Sets:** **Sets** are custom fields based on dimensions that are used to subset your data. They can be created manually or computed. Either dimensions or measures can be used to determine what is included or excluded from a set using conditional logic
- **Pages:** Dropping a field on the **Pages Shelf** allows you to analyze your data based on individual values contained within a field. A common use of pages is to add a Date field to show how your visualization changes over time.

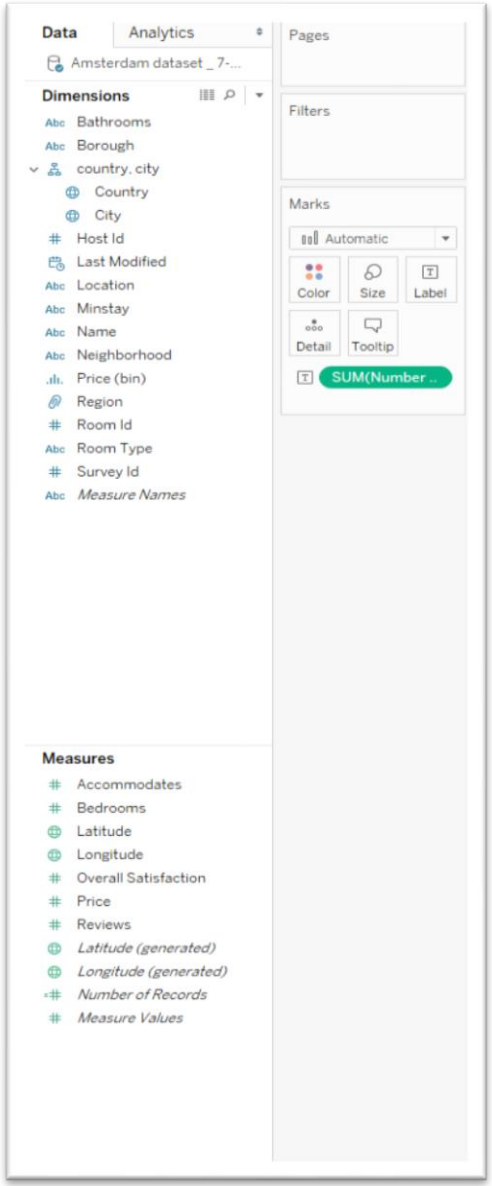
Airbnb: Sorting

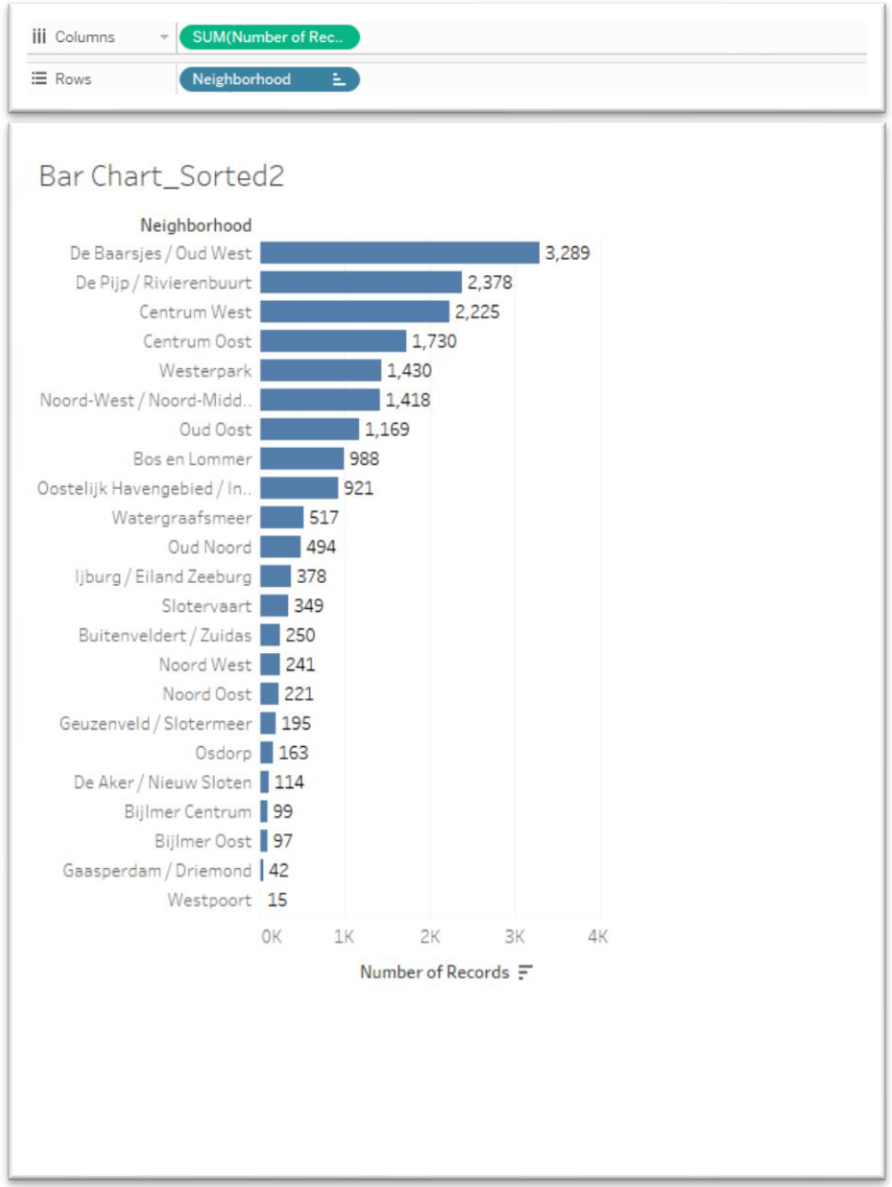
- There are several ways to sort dimensions in Tableau
- **Manually** sort by clicking on a neighborhood name and dragging it
- To create a **dynamic** sort, click on the **sort descending** icon on the task bar



Airbnb: Sorting

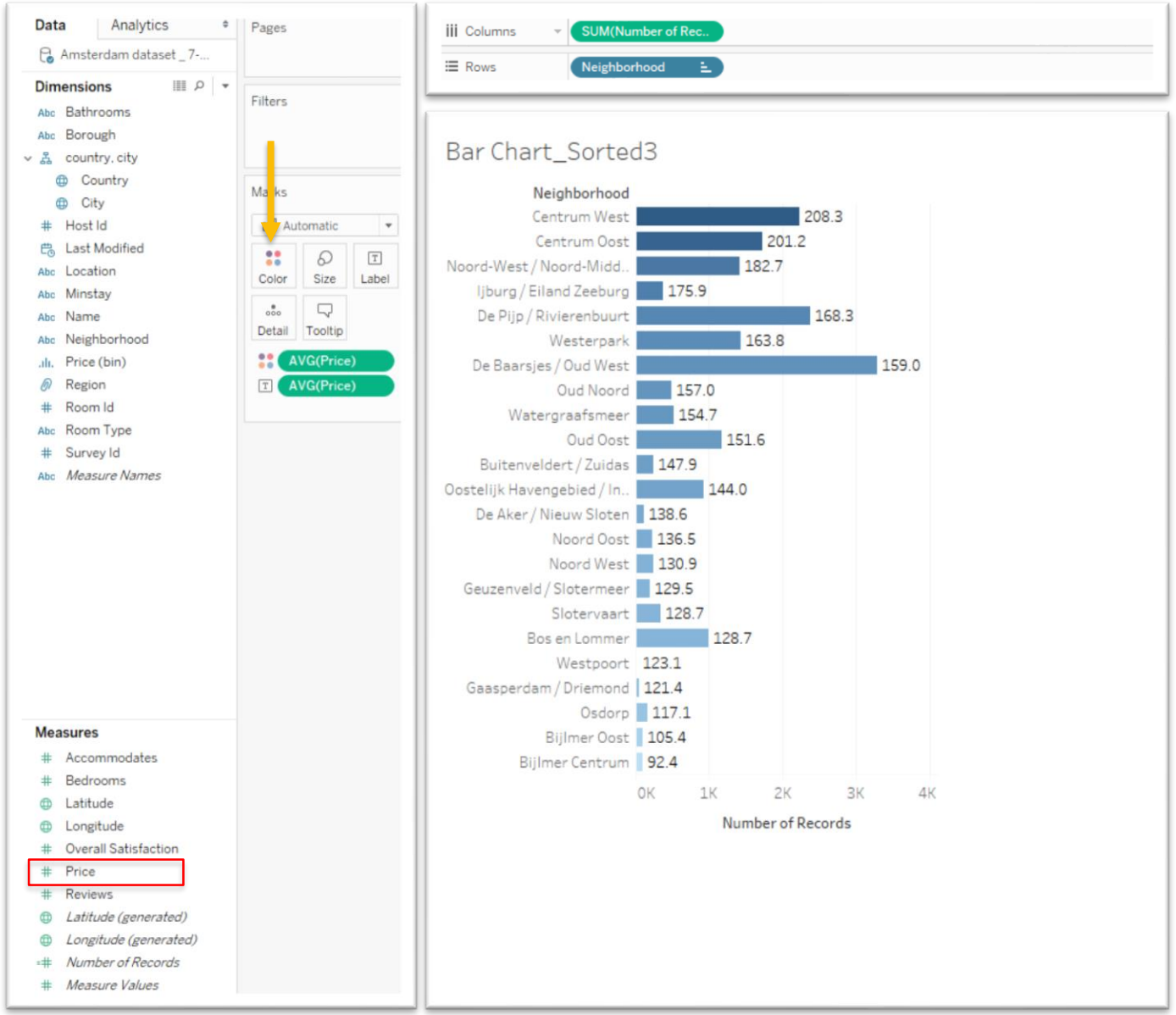
- There are several ways to sort dimensions in Tableau
- **Manually** sort by clicking on a neighborhood name and dragging it
- To create a **dynamic** sort, click on the **sort descending** icon on the task bar





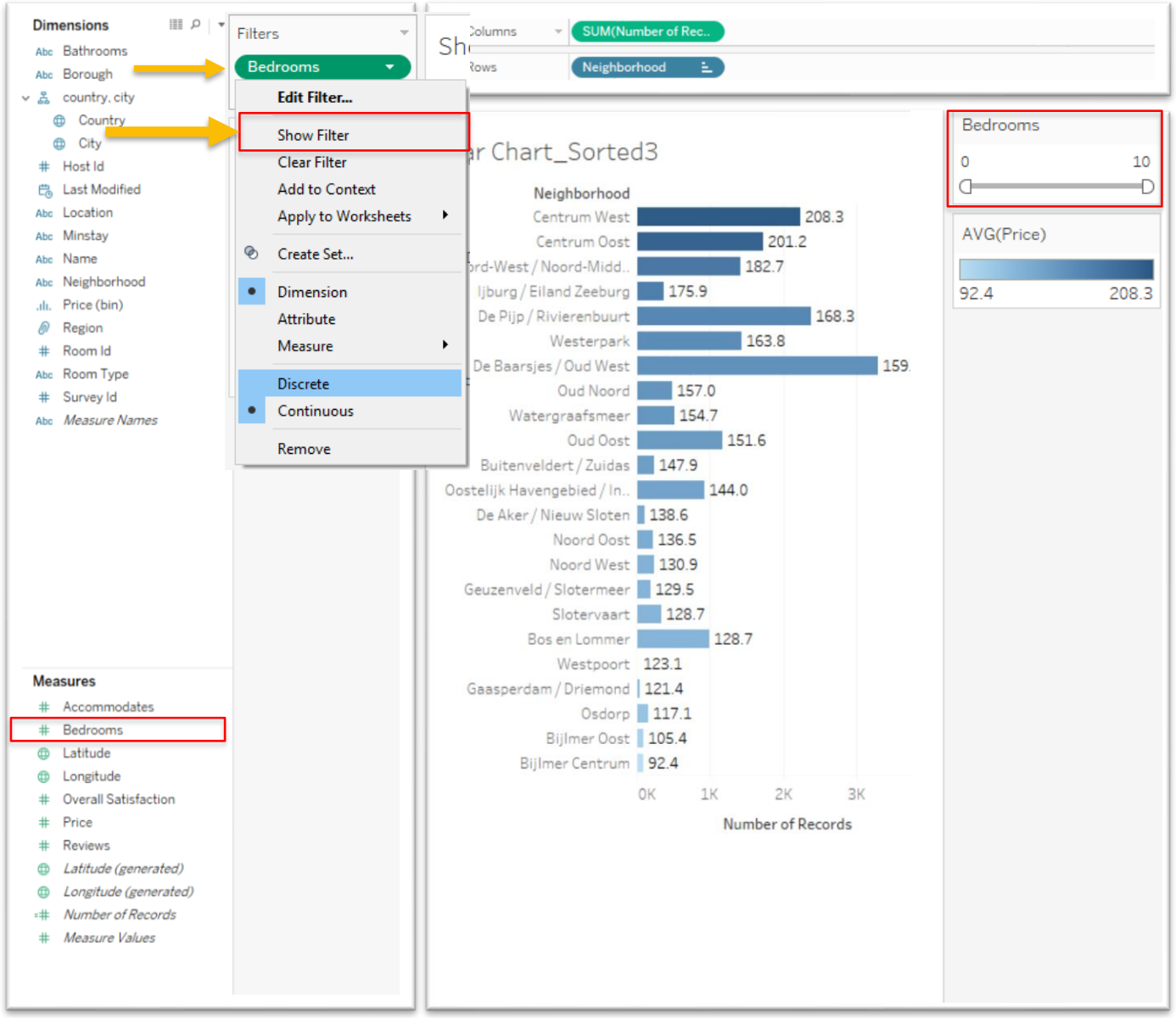
Airbnb: Sorting

- You can also use a **dynamic** sort to arrange the neighborhoods by **average price**
- First, drag **Price** onto the **Color shelf** and choose **Average** as the aggregation
- Next, right-click on the **Neighborhood pill** to open the **Sort** dialogue box and select **Price** as the **Field** and **Average** as the **Aggregation**



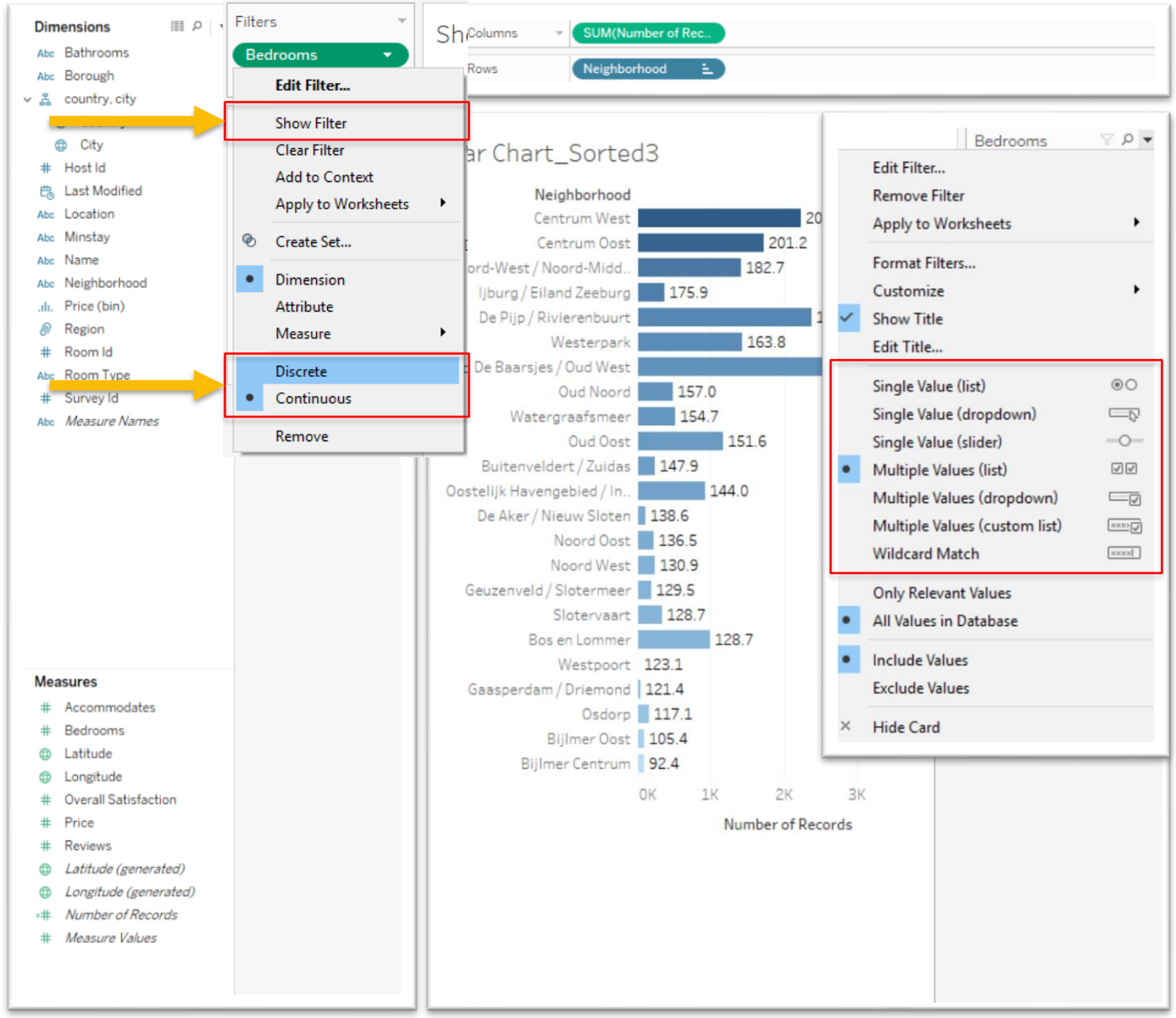
Airbnb: Filtering

- To filter **Neighborhoods** by # of **Bedrooms**, drag the **Bedrooms** field onto the **Filters shelf** and select “All Values”
- Right-click on the **Bedrooms** pill and select “Show Filter” from the drop-down menu
- A **slider** appears!



Airbnb: Filtering

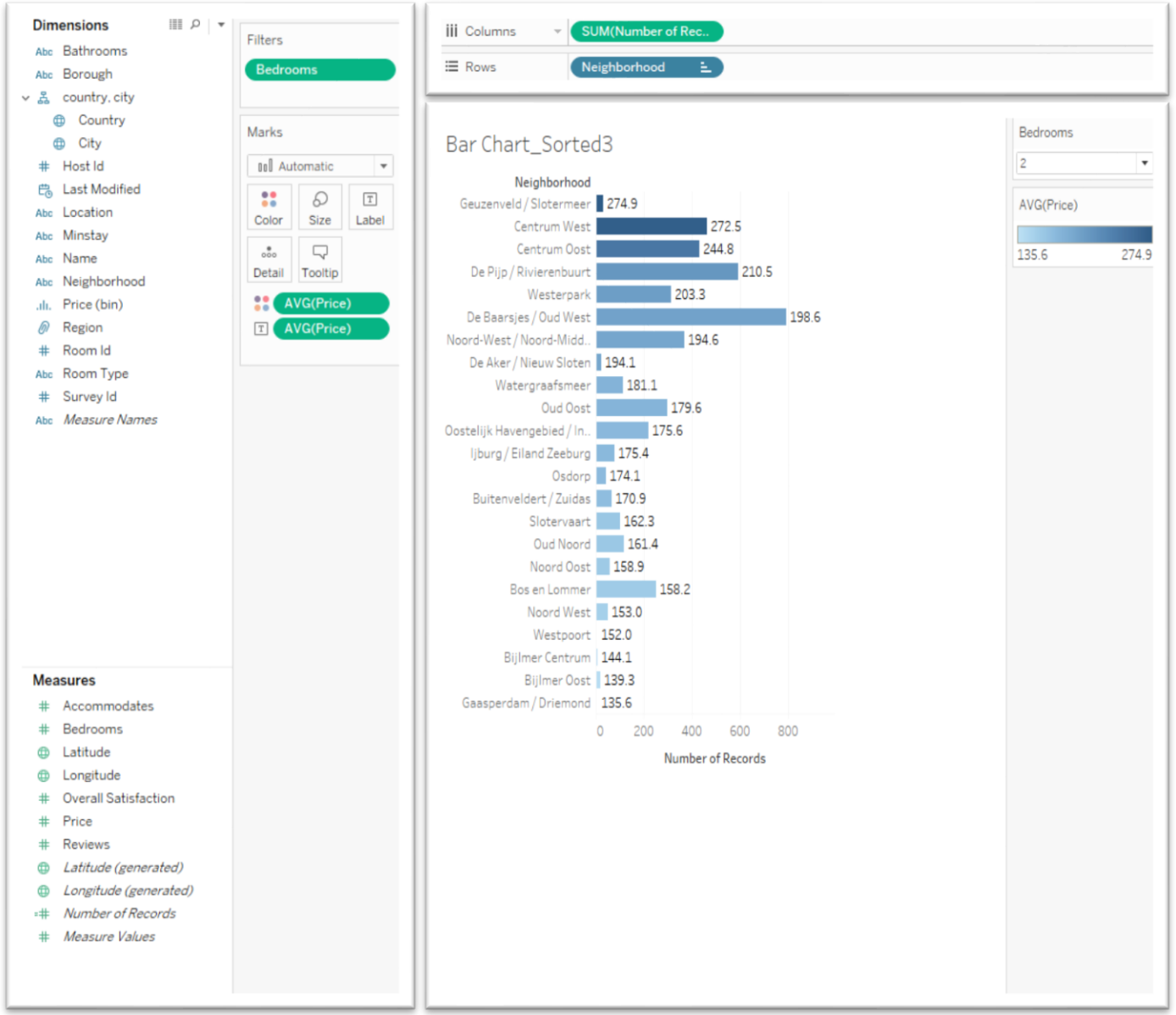
- To filter # of bedrooms one at a time, change **Bedrooms** to a discrete field in the drop down menu
- Select **Use All** in the dialogue box
- Again, select **Show Filter** in the drop-down menu
- To change the interactive filter menu, open the drop-down menu and select the desired filter menu



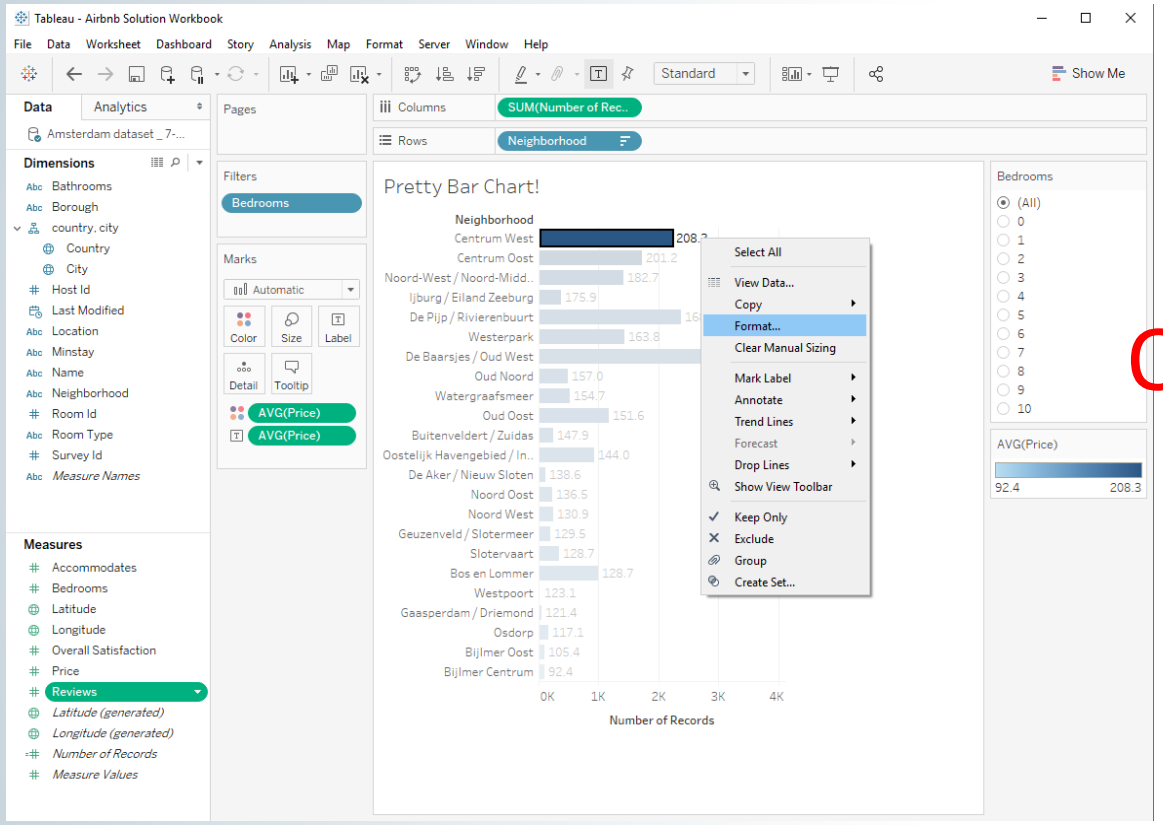
Airbnb: Filtering

- Which neighborhood has the most 2-bedroom listings?

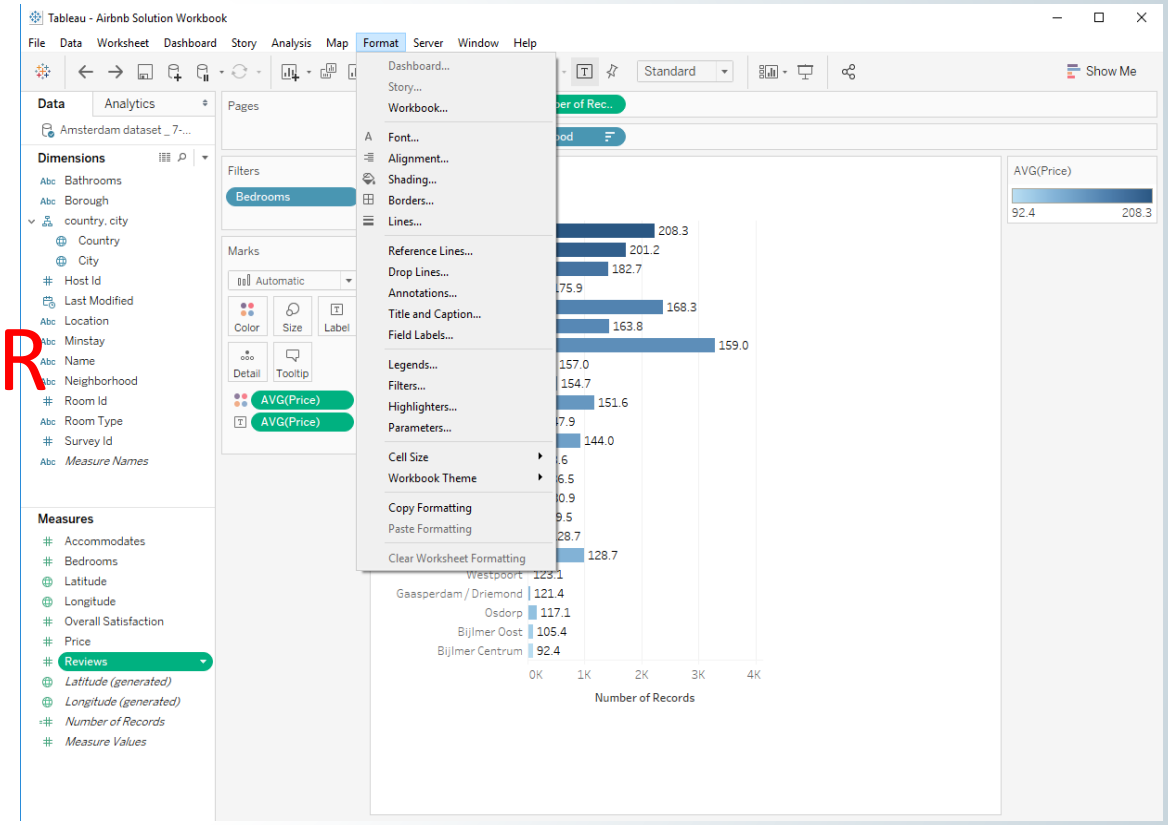
- In which neighborhood are 2-bedroom listings the most expensive on average?



Accessing the Format Bar

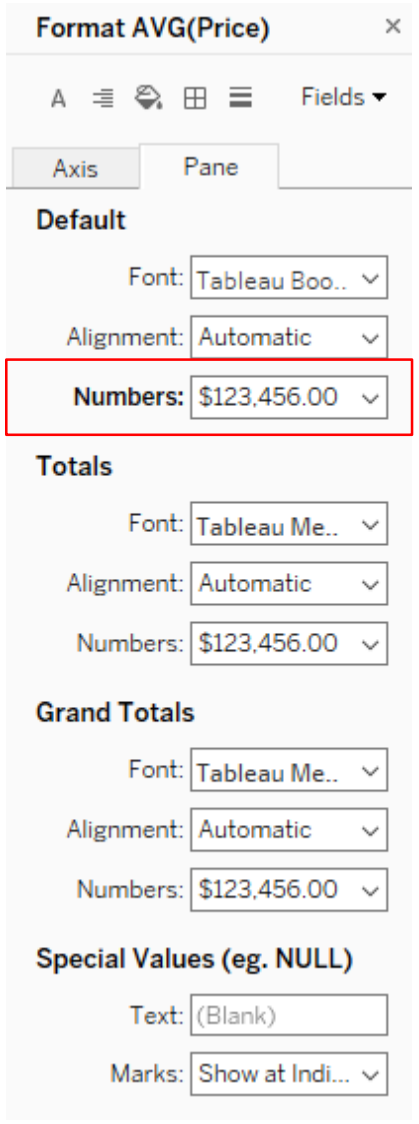


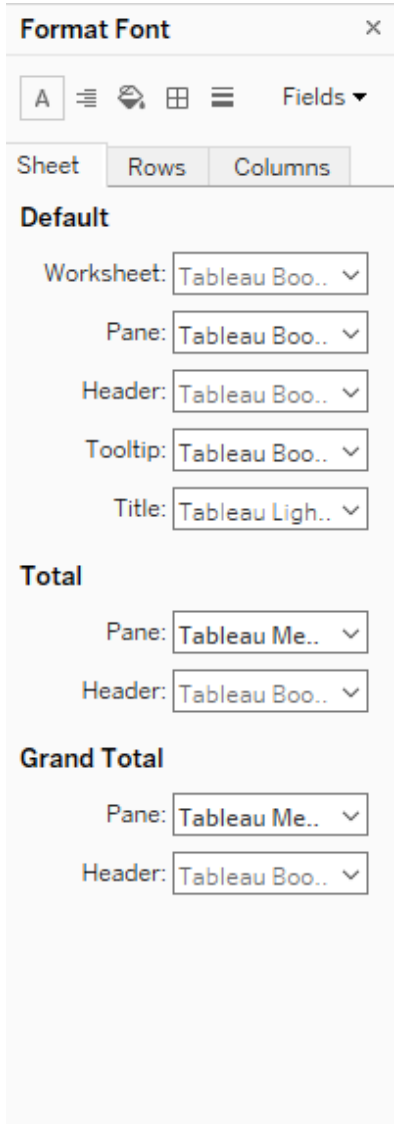
OR



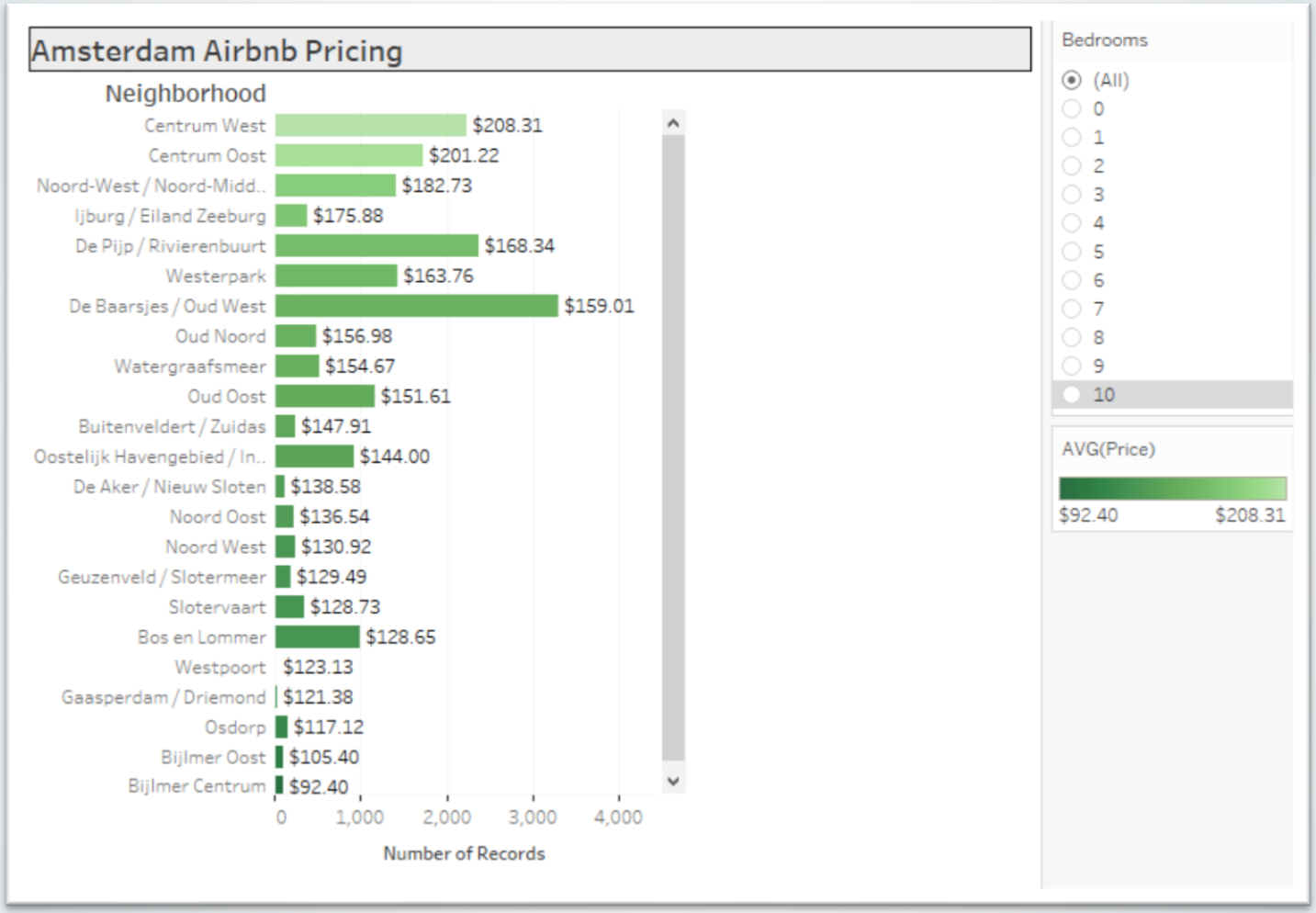
Format Bar

- What kind of format instruction we want to give
  - Font, Alignment, Shading, Borders, Lines, Fields [i.e. AVG(Price)]
- Tabs – scope of instruction
- Detailed Options





Recreate this Viz!



Exercise 3

Simpsons Part 2

- Exercise Objectives:
  - Apply & edit filters
  - Learn how to change & apply a color palette
  - Learn how to edit a title in a workbook, and how to insert variable names
  - Play around with the format pane!

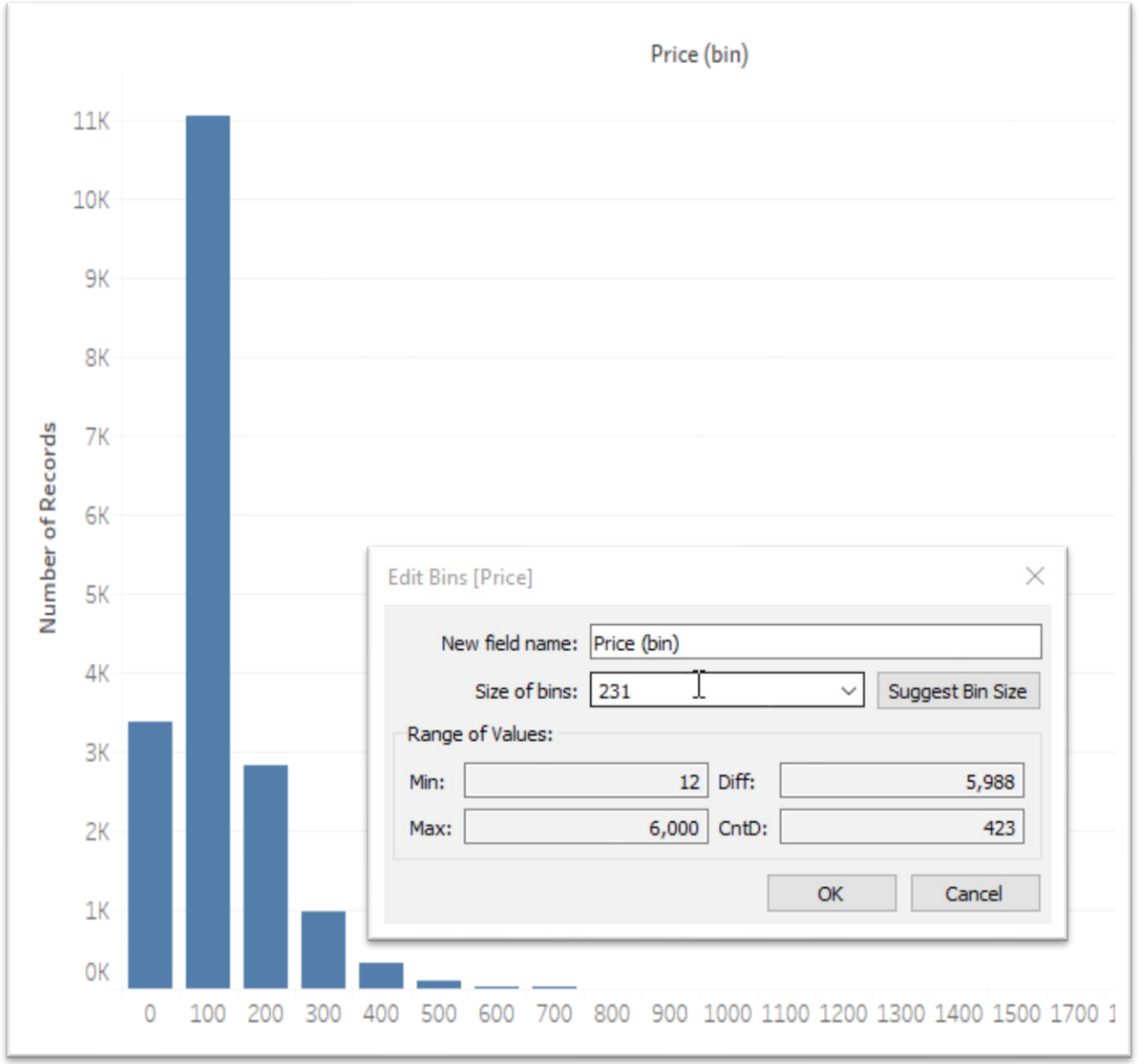
Exercise Materials

Exercise 3 Simpsons Starter 2.twbx

Exercise 3 Simpsons Solution 2.twbx

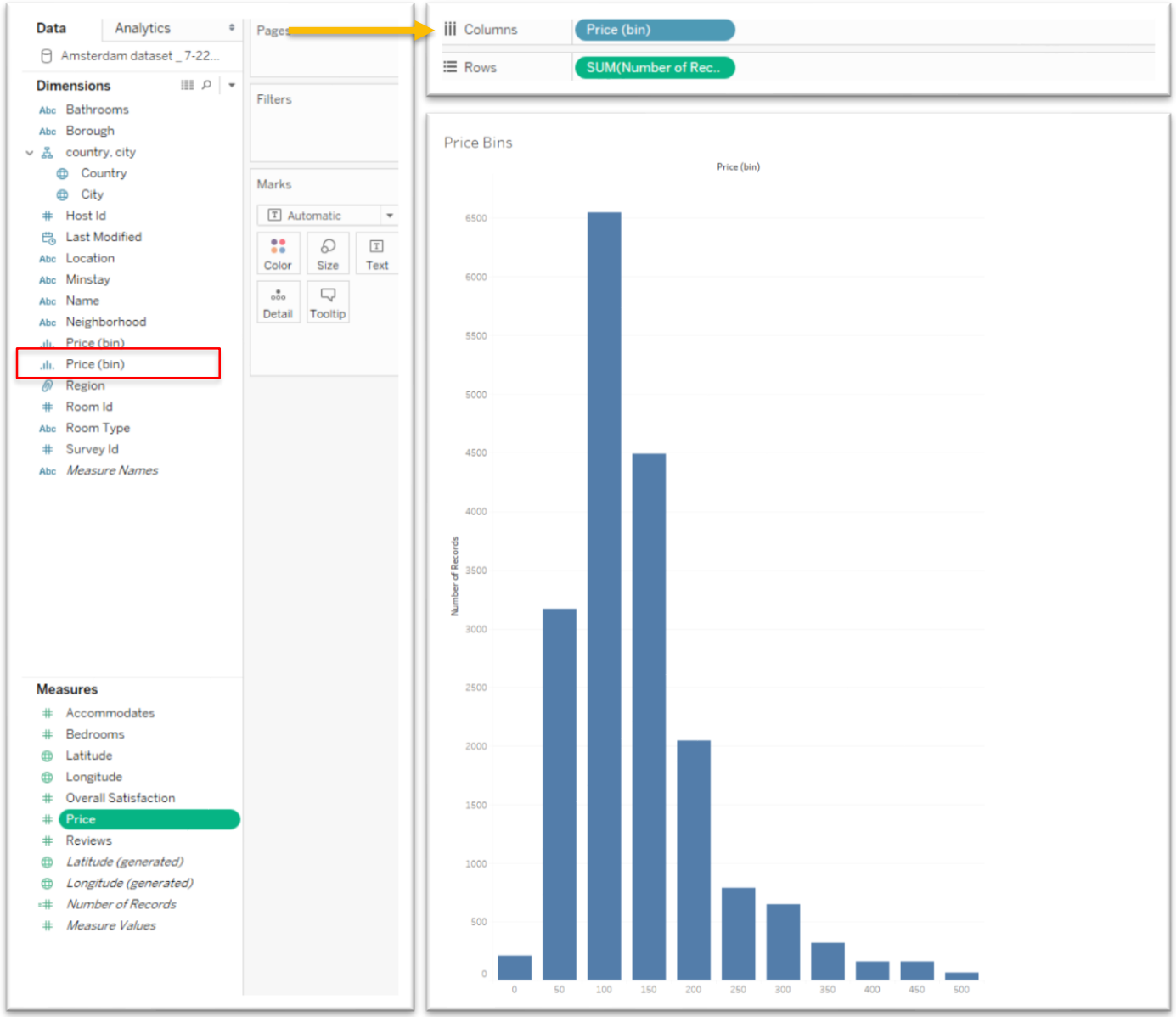
What are bins?

- A bin is a range of values that are grouped together
- A way to convert a continuous variable into a discrete variable
- Bins are often used in Histograms



Airbnb: Creating Bins

- Let's create **bins** for our **Price** measure
- Right-click on the field in the data pane and select **Create > Bins.**
- Edit the **Size of bins** to be 50 and click OK
- A new **dimension** is created
- Drag the new bin dimension to the **Columns shelf** and **Number of Records** on the **Rows shelf**



Exercise 4

Airbnb

- Exercise Objectives:
  - Create a **set** & understand how/when to use it
  - Explore **pages** and apply them to a viz to gain better insights in our data

Exercise Materials

Exercise 4 Airbnb Starter.twbx

Exercise 4 Airbnb Solution.twbx

ASKING YOUR DATA, "WHERE?"

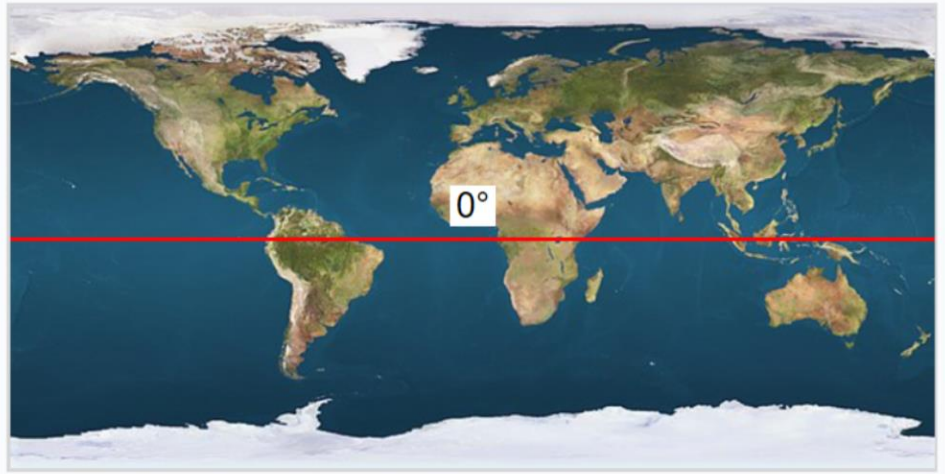
Mapping in Tableau

Why Use a Map?

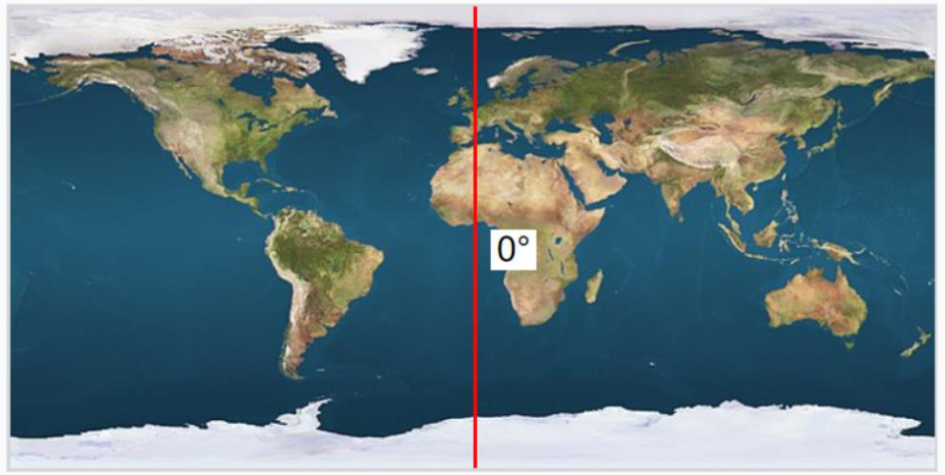
- Gives us the ability to answer the question “where?” when it comes to our data
- Provides geographical context to our data
- Gives new insight not immediately recognizable with a table, linear graph or pie chart
- Examples of geographic questions:
  - Where are teacher turnover rates the highest?
  - How many grocery stores are in each neighborhood?
  - What state has the most Big Foot sightings?

Latitude & Longitude

- Any point on a map can be represented with latitude and longitude points
- Latitude indicates distance north or south from the equator and Longitude indicates distance east or west from the Prime Meridian
- If your dataset includes latitude and longitude fields, Tableau can automatically plot them on a map
- If your dataset does not include latitude and longitude fields but has geographic place names such as city, country or state, Tableau can determine their coordinates with the generated latitude and longitude fields



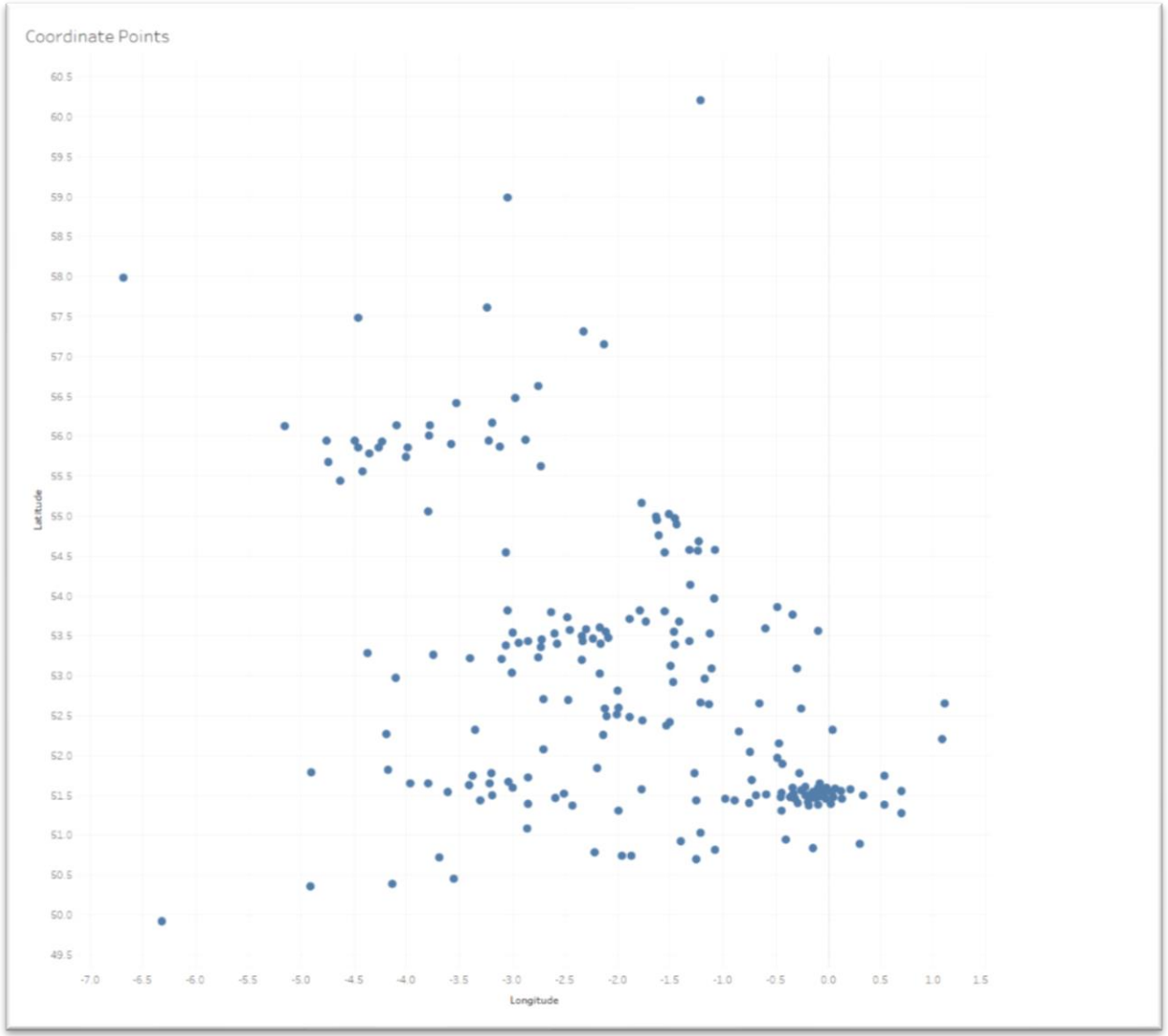
Equator



Prime Meridian

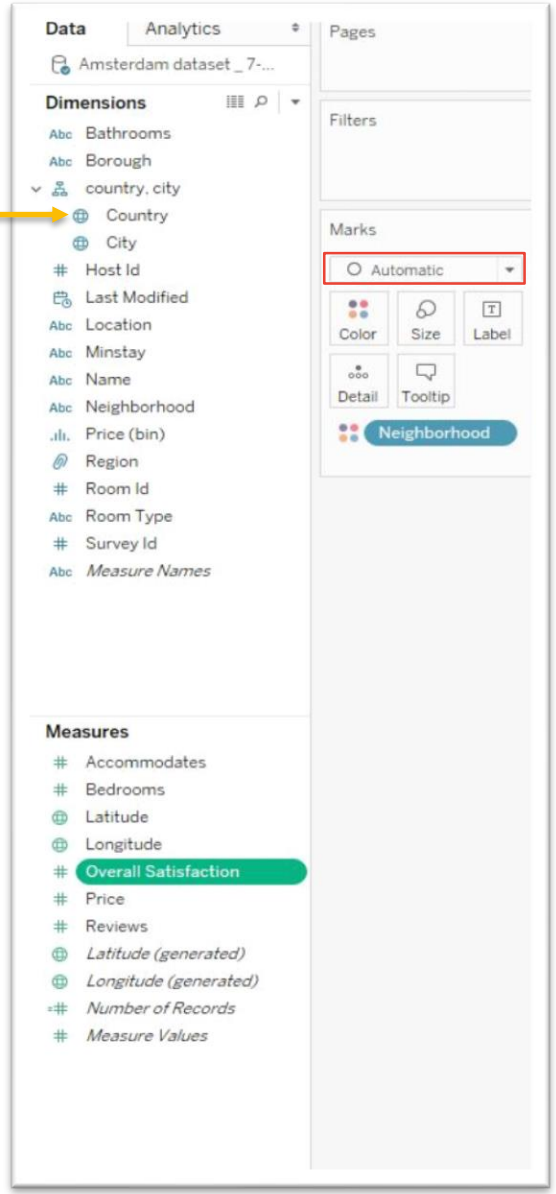
Coordinate Points

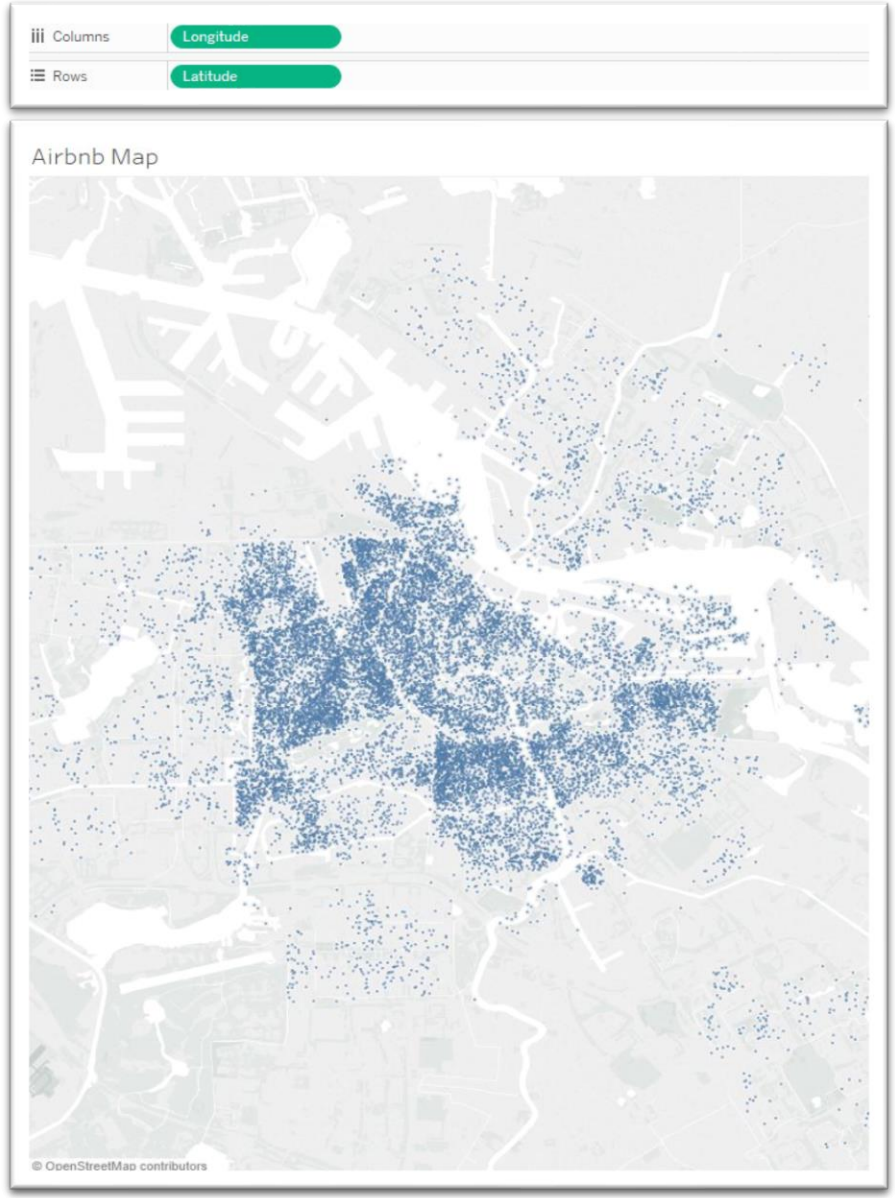
- Every point on the globe has unique latitude and longitude coordinates
- In Tableau, coordinates need to be in decimal format
- Tableau can automatically plot points using the latitude and longitude fields
- By default, Tableau will give an aggregate average of the latitude and longitude entries by whatever fields you have selected for your map
- If no other fields are selected, Tableau will generate a single coordinate point representing the average latitude and longitude



Airbnb: Point Maps

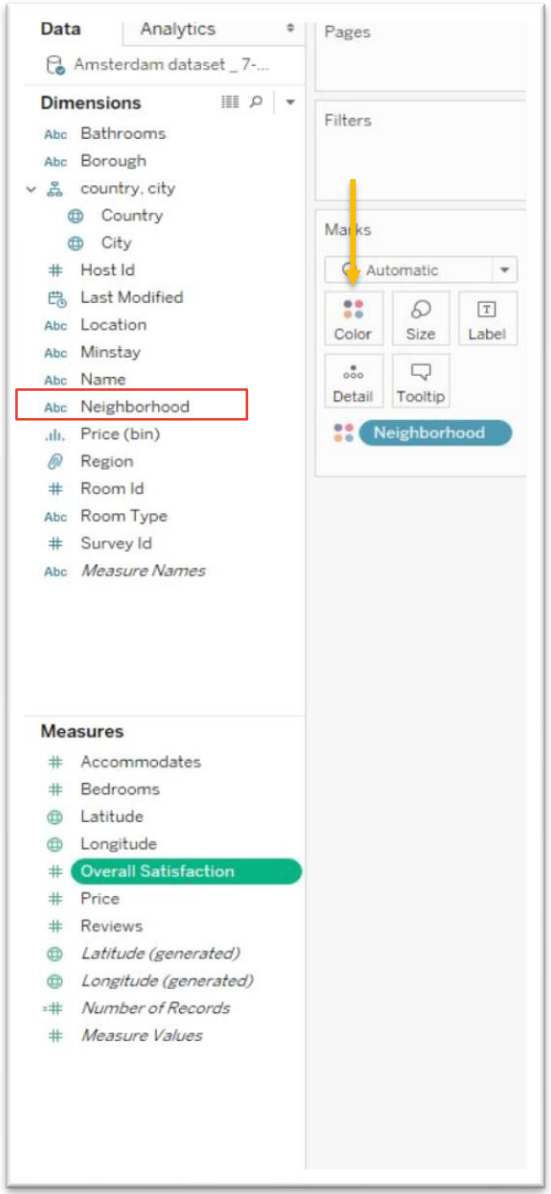
- Refer back to the Airbnb map we created in the beginning of the course
- Tableau automatically recognizes geographic field names like city, state and country
- A globe icon in the data pane represents a geographic field
- To plot a location double click or drag the field onto the sheet
- By default, Tableau will plot geographic locations with a symbol mark

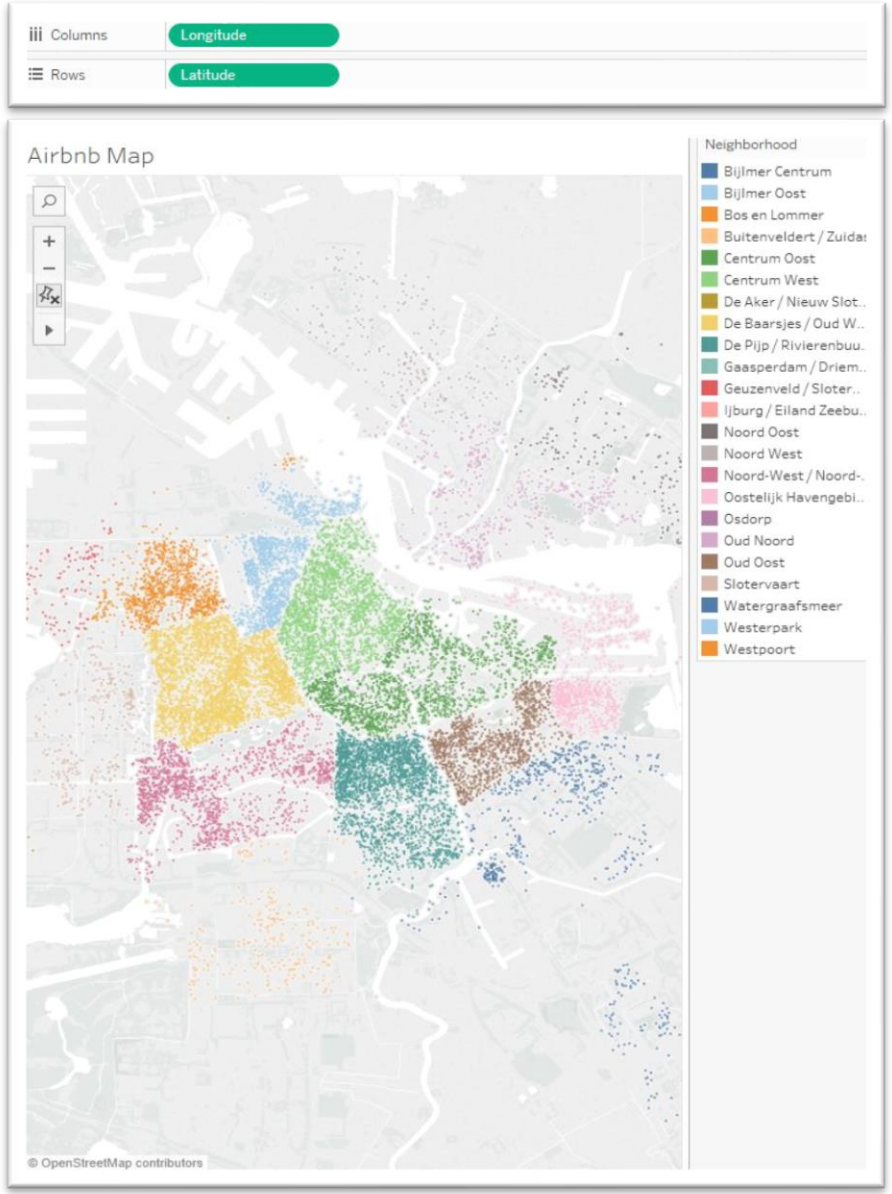




Airbnb: Point Maps

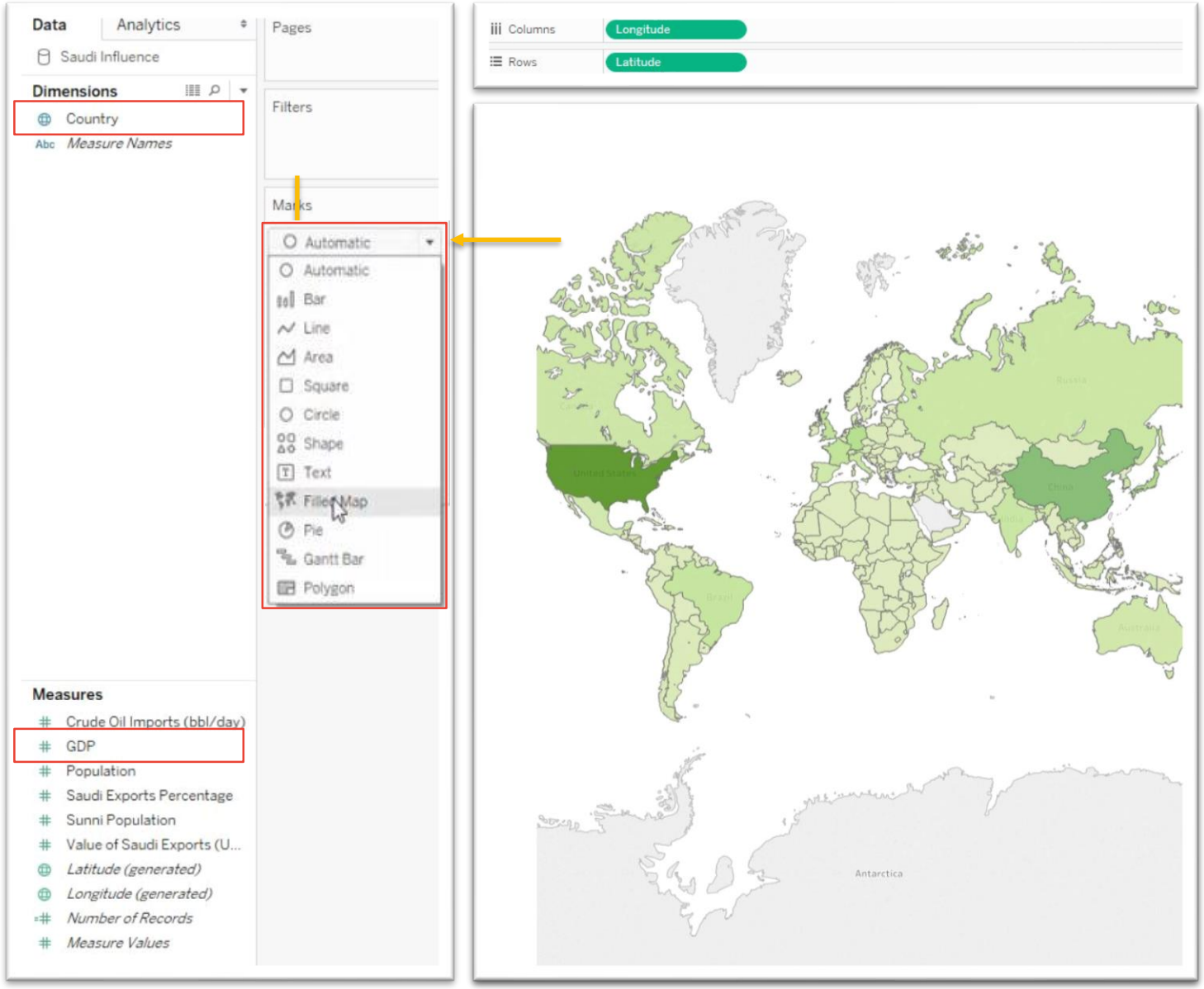
- Tableau automatically recognizes geographic field names like city, state and country
- A globe icon in the data pane represents a geographic field
- To plot a location double click or drag the field onto the sheet
- By default, Tableau will plot geographic locations with a symbol mark





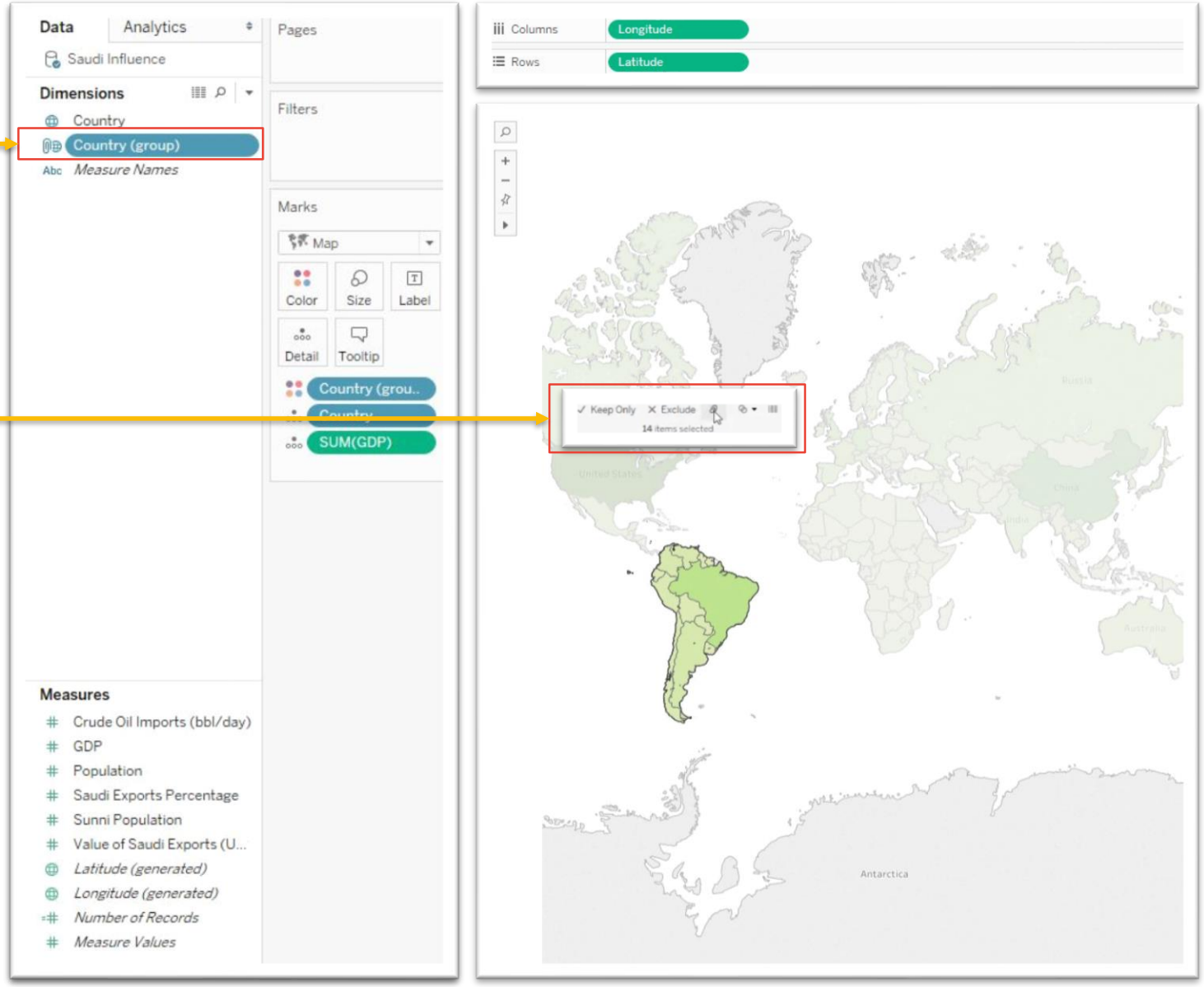
Saudi Influence: Filled Maps

- To begin, open the Tableau starter workbook **Indicators of Saudi Influence**
- Double-click on a geographic field such as city or country and Tableau creates a symbol map by default
- Drag a **measure** field onto **color** in the **marks shelf** and Tableau automatically converts to a filled map
- Drag a **dimension** onto **color** and Tableau reverts back to a symbol map
- To manually choose between a filled or symbol map, change the **mark type** in the drop-down menu



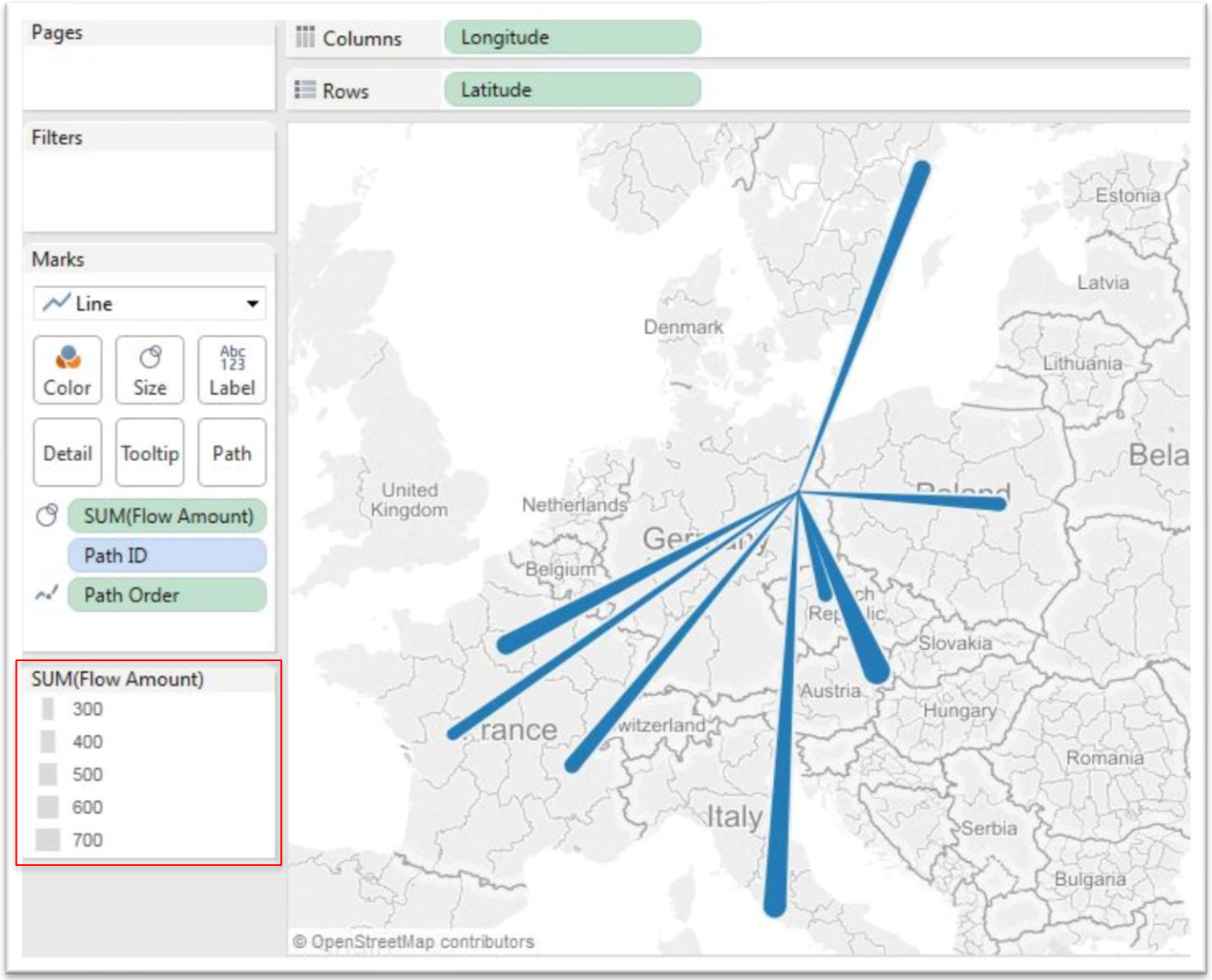
Saudi Influence: Filled Maps

- **Group** locations on a map to create custom territories (i.e. countries based continent)
- Select the locations (countries, states, etc.) that make up the territory and group them by clicking on the paper-clip
- Each group name will appear as a **dimension** and functions like any other group
- The original geographic field can be removed from the view to show only the custom territory



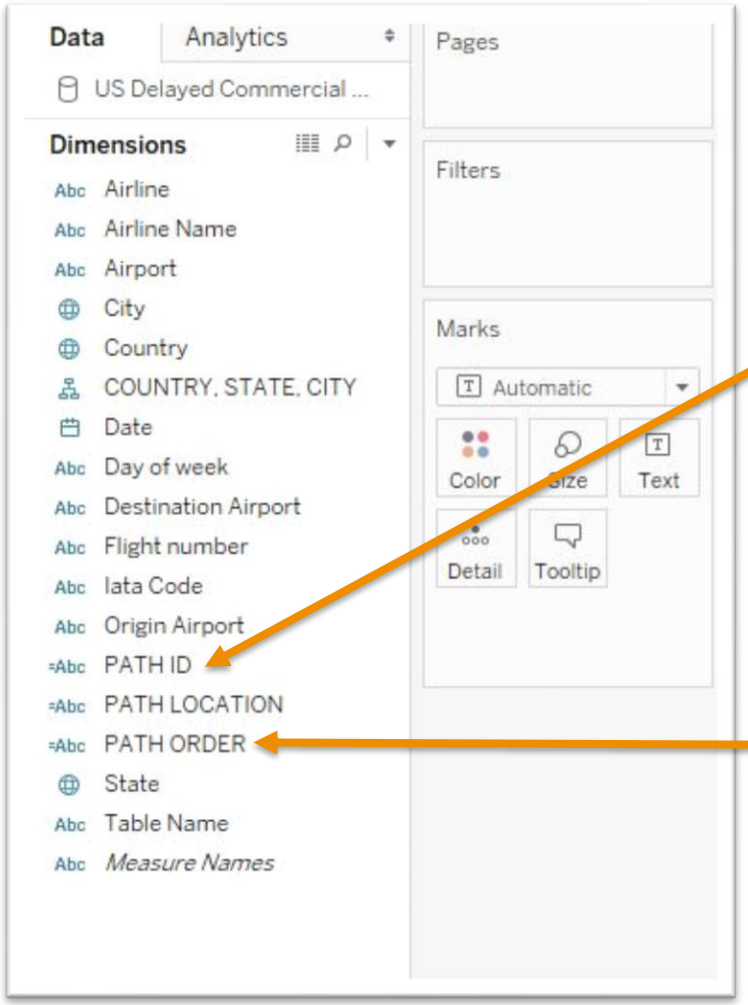
Path Maps

- Path maps geographically show the movement of information or objects from one location to another and their magnitudes
- The thickness of the **flow line** represents the magnitude, or amount, of movement
- **Flow lines** are drawn from one point of origin and branch out to an end point
- Arrows may be used to show direction or if the movement is ingoing or outgoing



Path Maps

- To create a **path** you will need the following fields:
  - **Path ID**
  - **Path order**
  - Location coordinates (latitude & longitude)
- **Path ID** and **Path Order** are calculated fields and typically do not exist in the original dataset

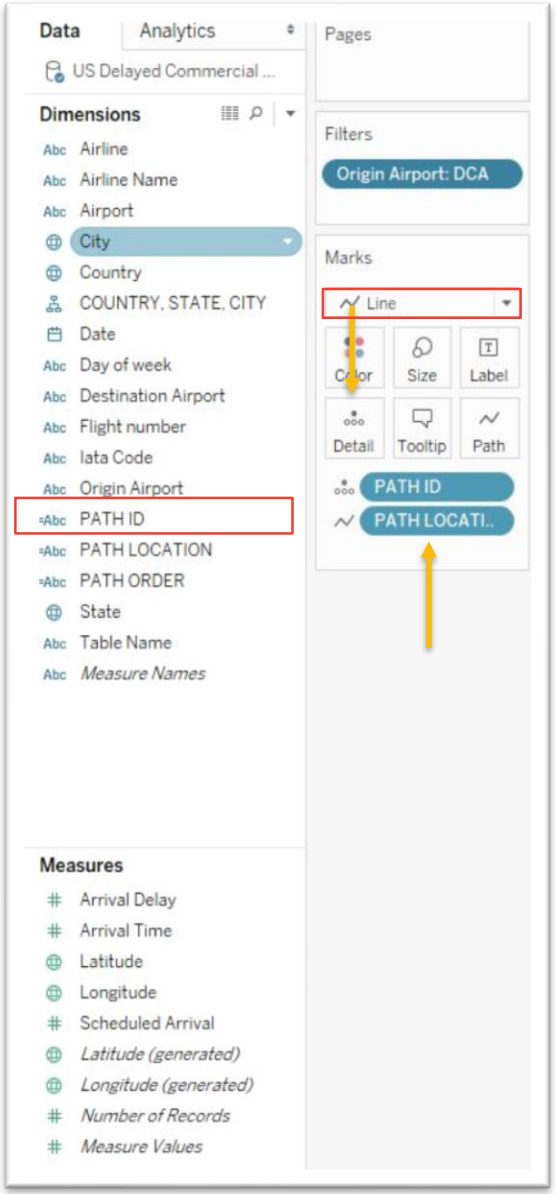


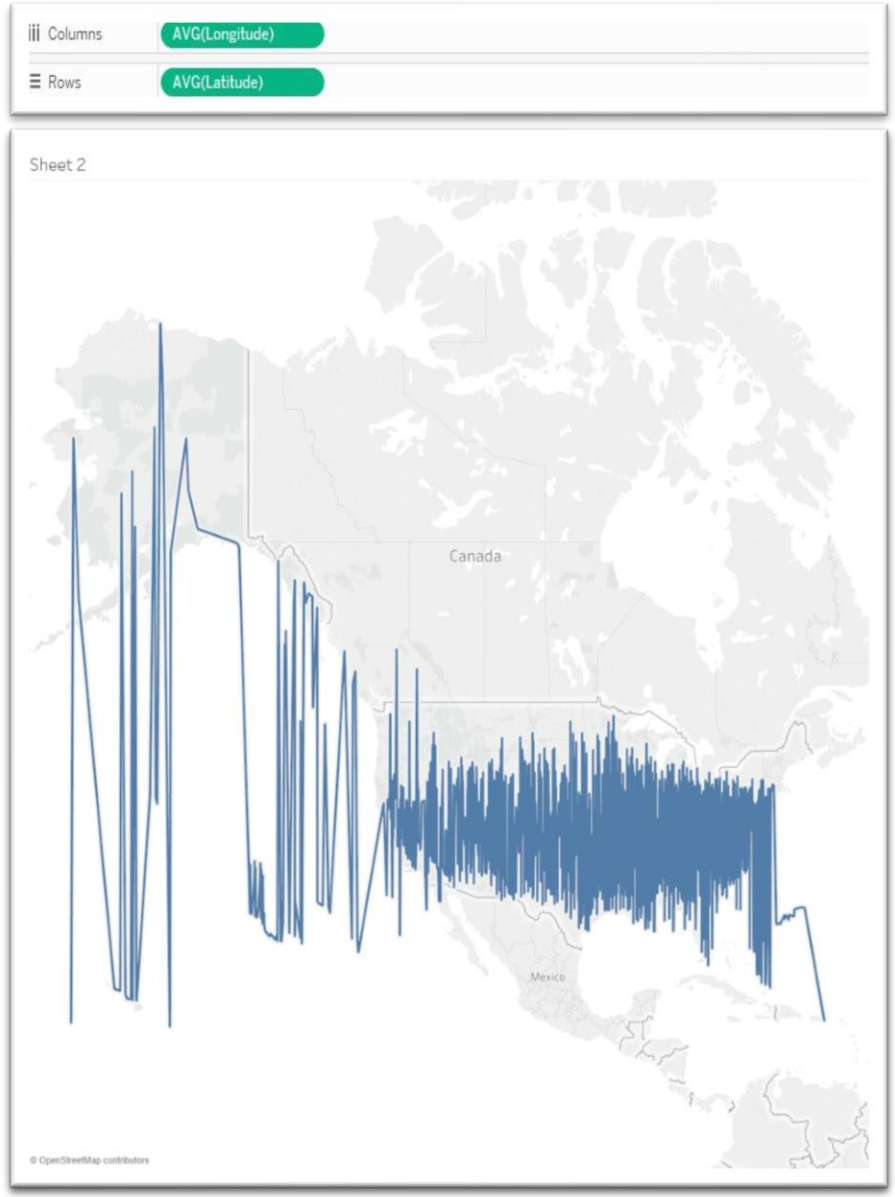
Path ID is a distinct value for each path in the data. This tells Tableau which locations group together along a path

Path Order is a field that assigns an order by which to connect the points along the **path** (for example, **date and time**)

Flight Delays: Path Maps

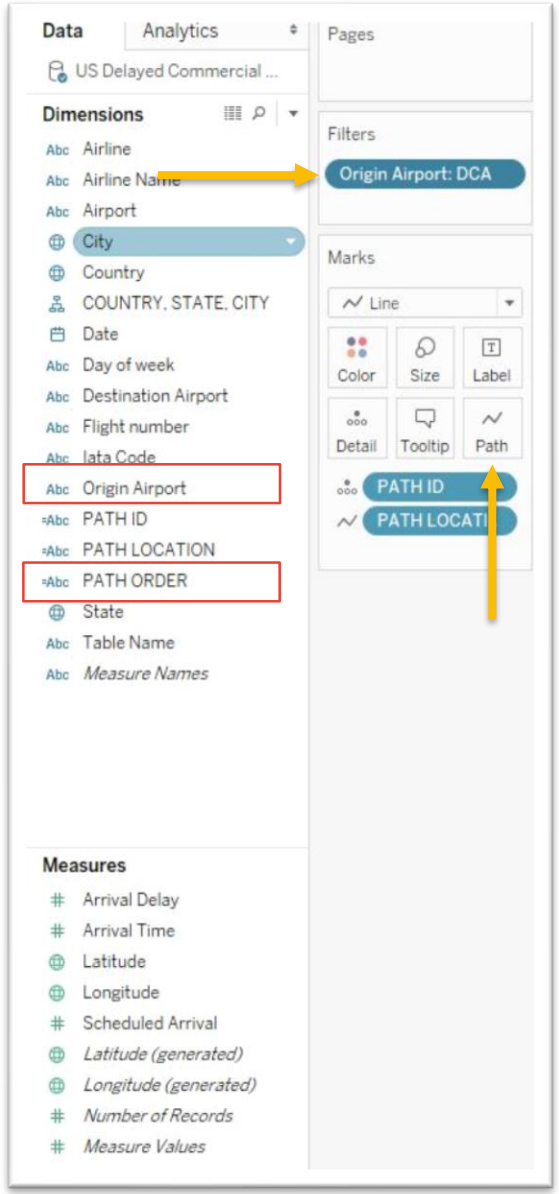
- To begin, open the Tableau starter workbook **Flight Delays** in the Starter Workbooks folder
- To create your path map visualization, start by double-clicking on your **latitude** and **longitude** fields to generate a map
- Next, drag your **Path ID** field onto **Detail** in your **Marks Shelf**
- By default, Tableau creates a point map. Select **Line** in the drop-down menu to create a path

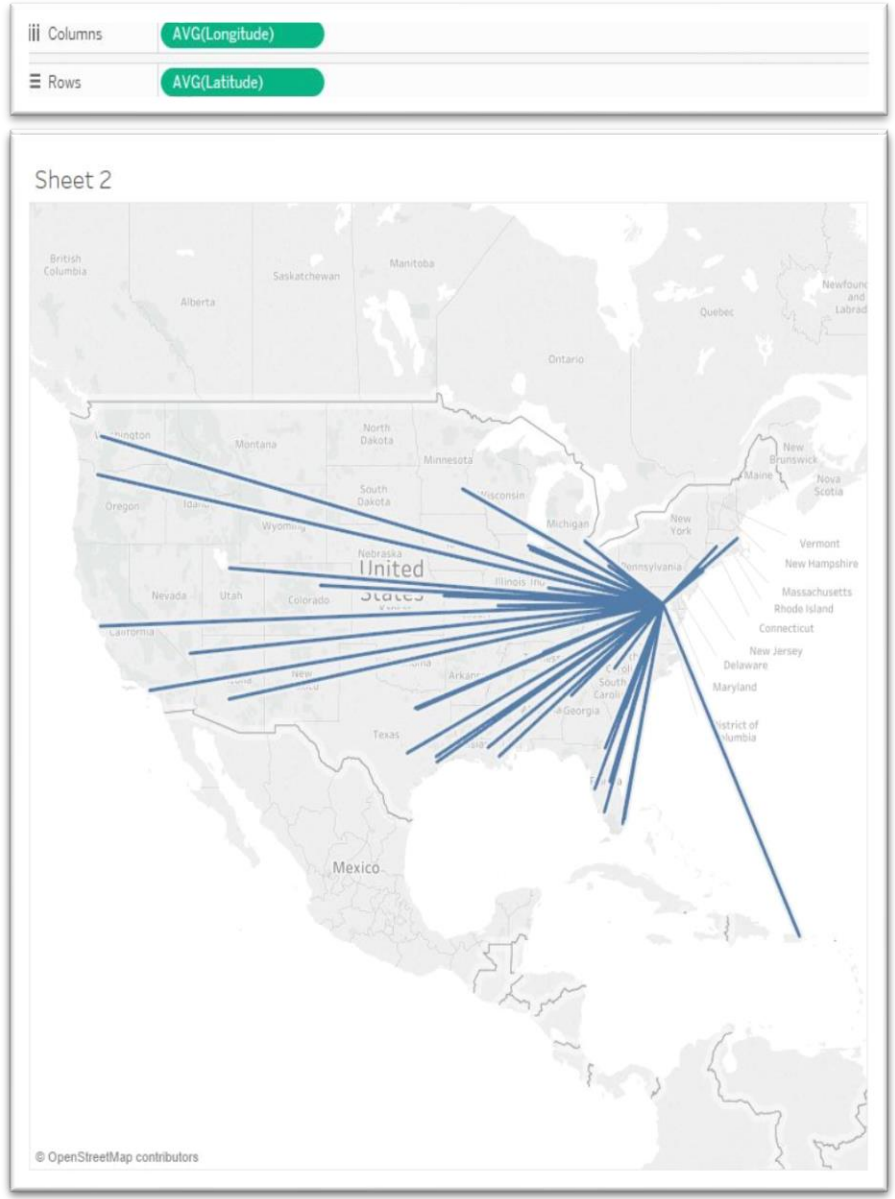




Flight Delays: Path Maps

- Finally, **drag** your **Path Order** field onto **Path** in your **Marks Shelf**
- You can filter fields such as **Origin** in this example to make your visualization more readable
- You can drag **measure** fields to **color** and **size** on the **Marks Shelf** to add more information to the visualization





Exercise 5

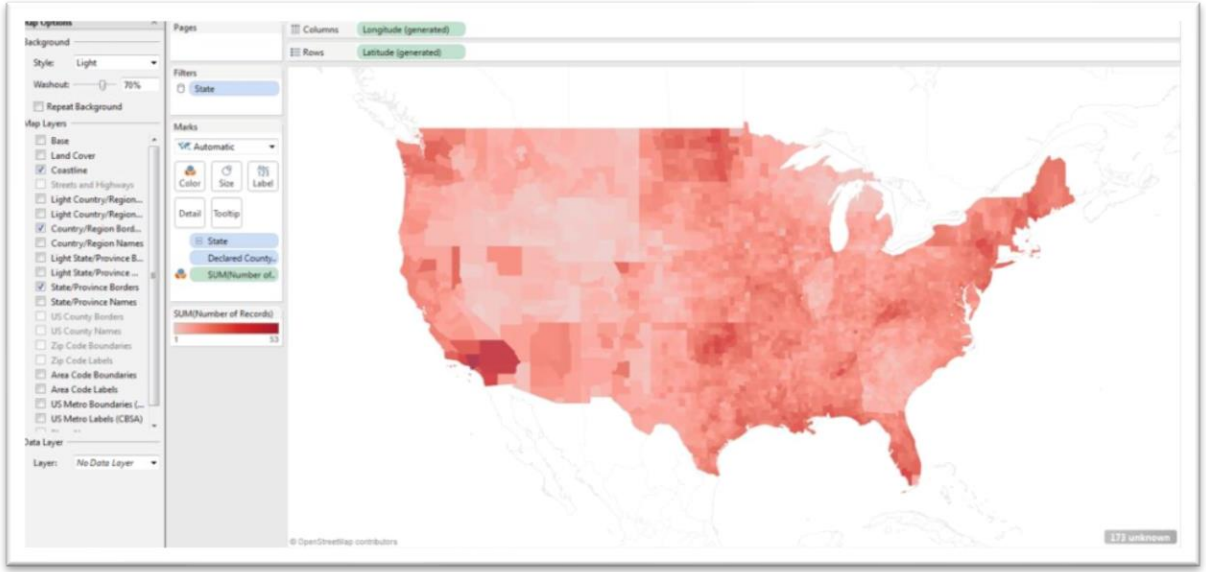
Path Maps

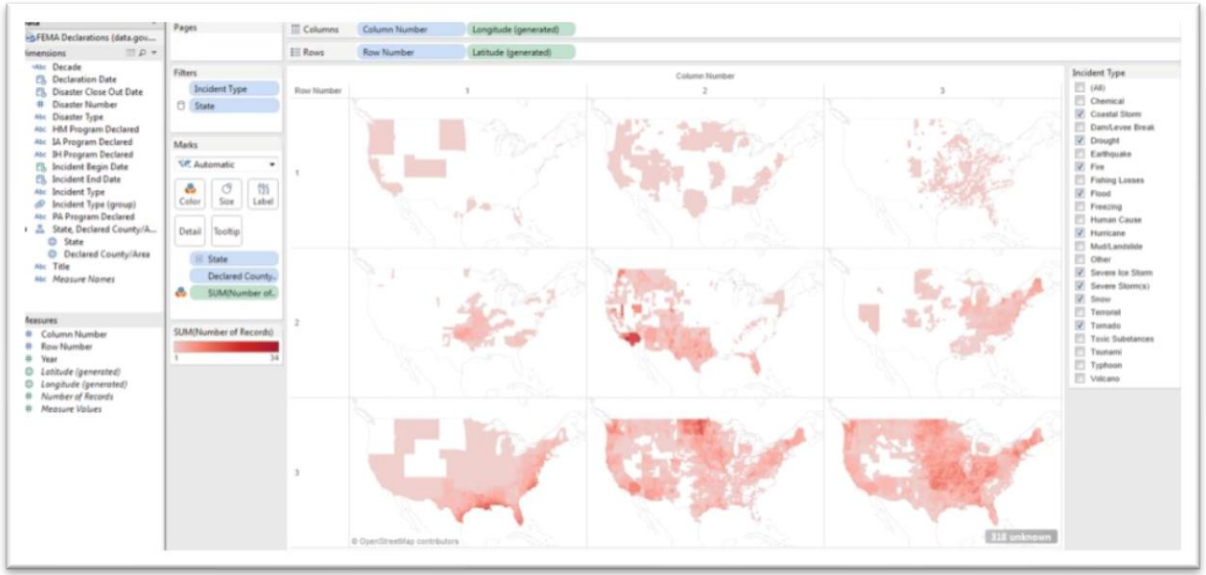
- Exercise Objectives:
  - Learn how to use map features to analyze track data
  - Learn how to leverage the Page shelf to animate data for presentation

Exercise Materials
Exercise 5 Belgian Trains Starter.twbx
Exercise 5 Belgian Trains Solution.twbx

Heat Maps

- When working with large datasets, points often become cluttered making the maps difficult to decipher
- Heat maps allow us to aggregate data points into surfaces representing different density ranges
- By rendering lots of point-level detail more intuitively as ‘hot-spots’, we can decipher broad geographic patterns more effectively





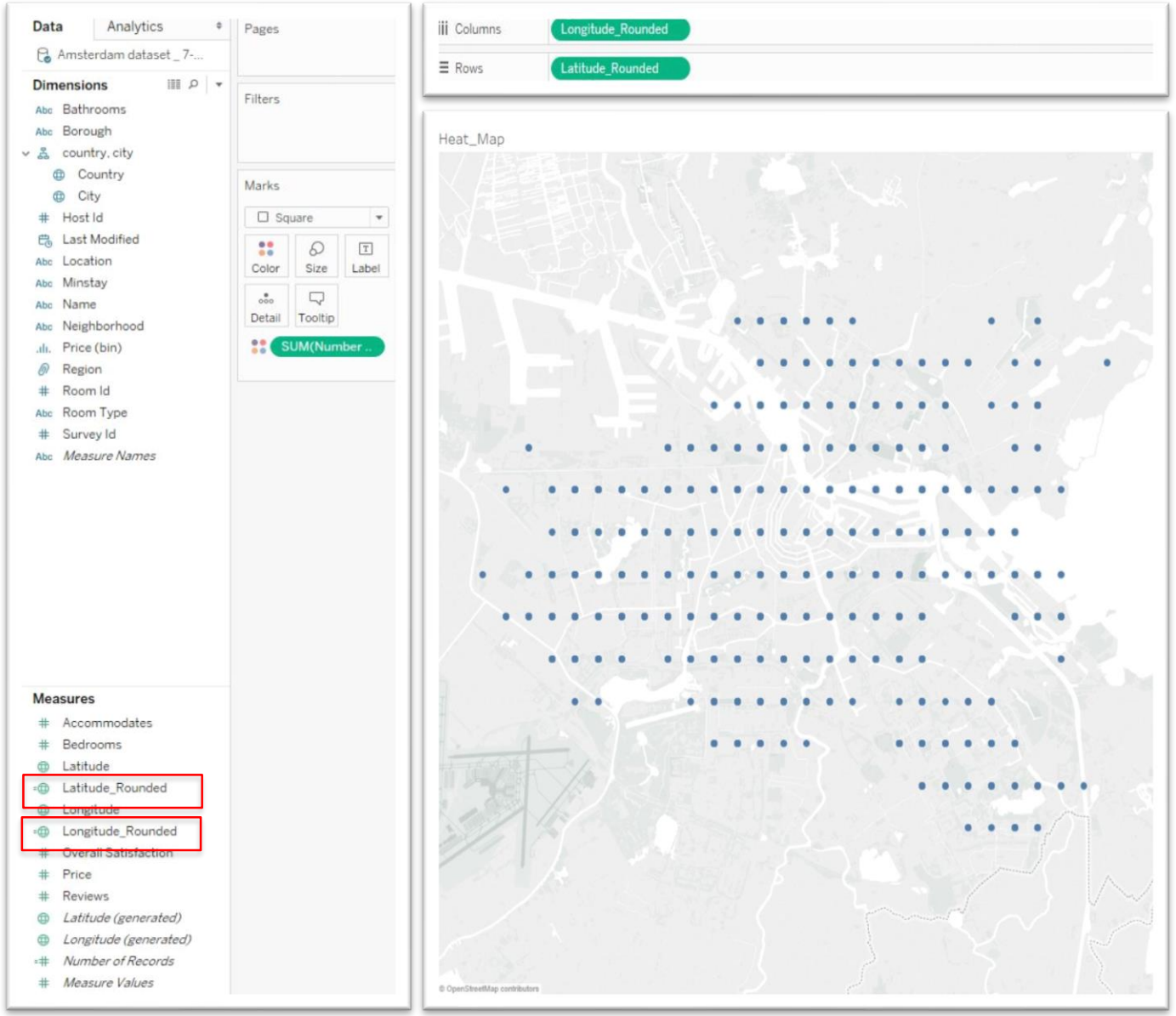
Images: dataremixed.com

Heat Maps

- First create a **calculated field** (one for each latitude and longitude field) to round the decimal values to the desired level of specificity

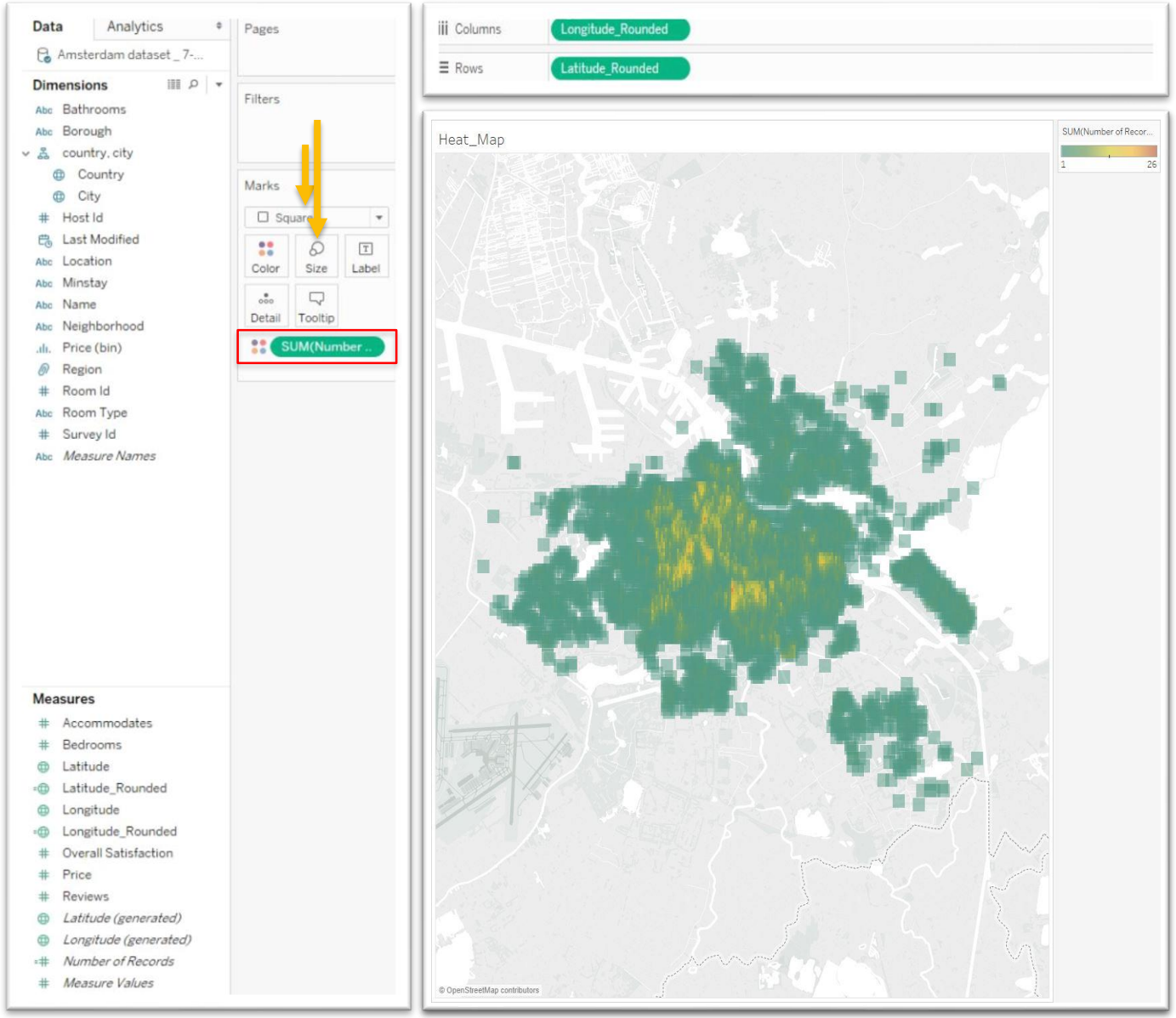
```
ROUND([Latitude],2)
```

- Change the **calculated fields** to **geographic** fields
- Add the calculated latitude and longitude fields to the sheet to create a grid of points and drag **Room ID** to detail



Heat Maps

- Change **Longitude_Rounded** and **Latitude_Rounded** to Dimensions
- Change the symbol to **square** on the **Marks Shelf** and increase the size so that the squares are touching
- Drag a **measure** field onto **color** and edit the colors and opacity as desired



Exercise 6

Heat Maps

- Exercise Objectives:
  - Learn how to create a geographic heat map
  - Apply zoom controls to analyze dense spatial data
  - Practice creating Calculated Fields

Exercise Materials

Exercise 6 Twitter HeatMap Starter.twbx

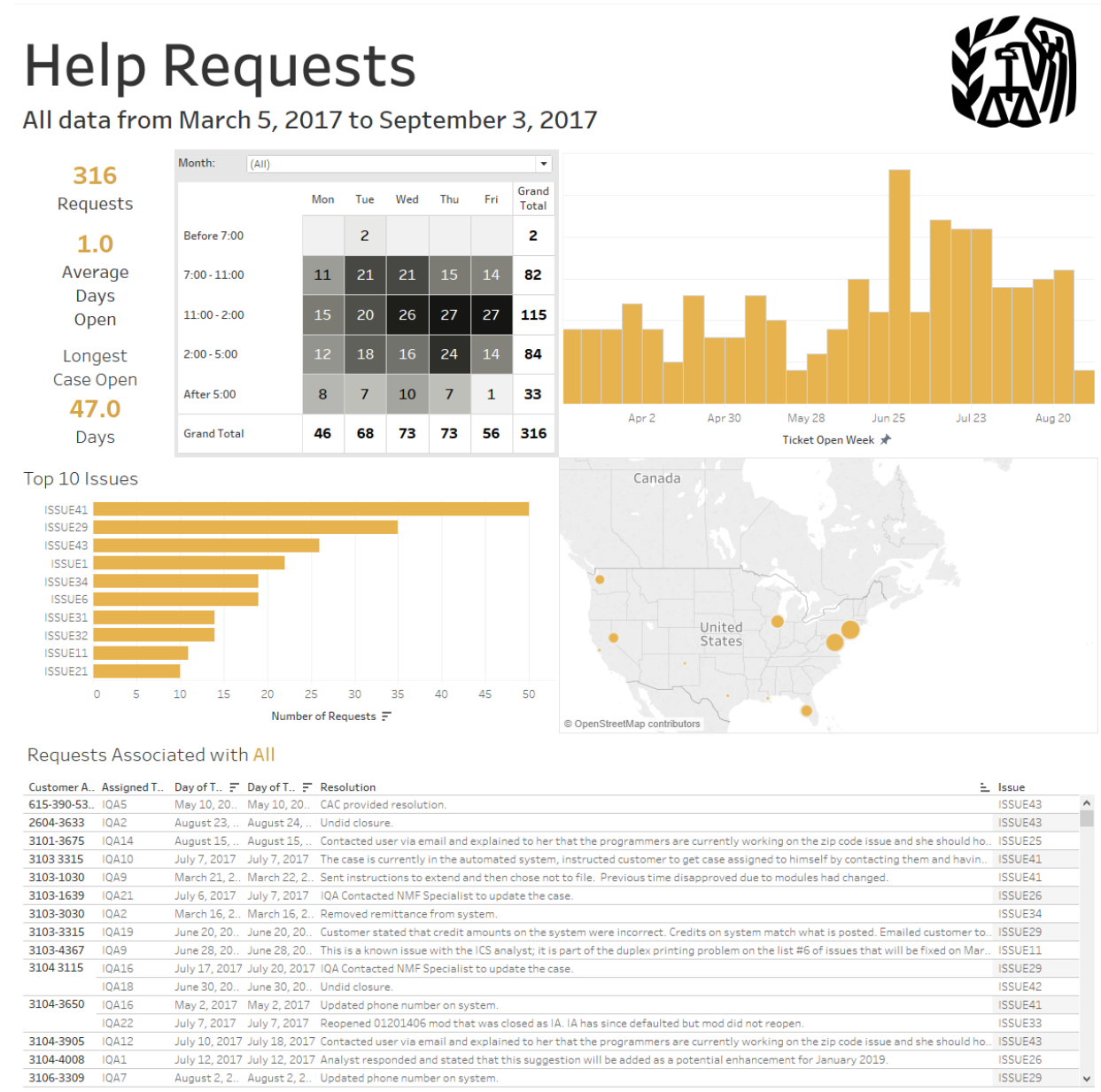
Exercise 6 Twitter HeatMap Solution.twbx

DASHBOARDS & STORIES

Creating a Narrative

What is a dashboard?

- A dashboard is a collection of several worksheets and supporting information shown in a single place so you can compare and monitor a variety of data simultaneously.
- When you create a dashboard, you can add views from any worksheet in the workbook.



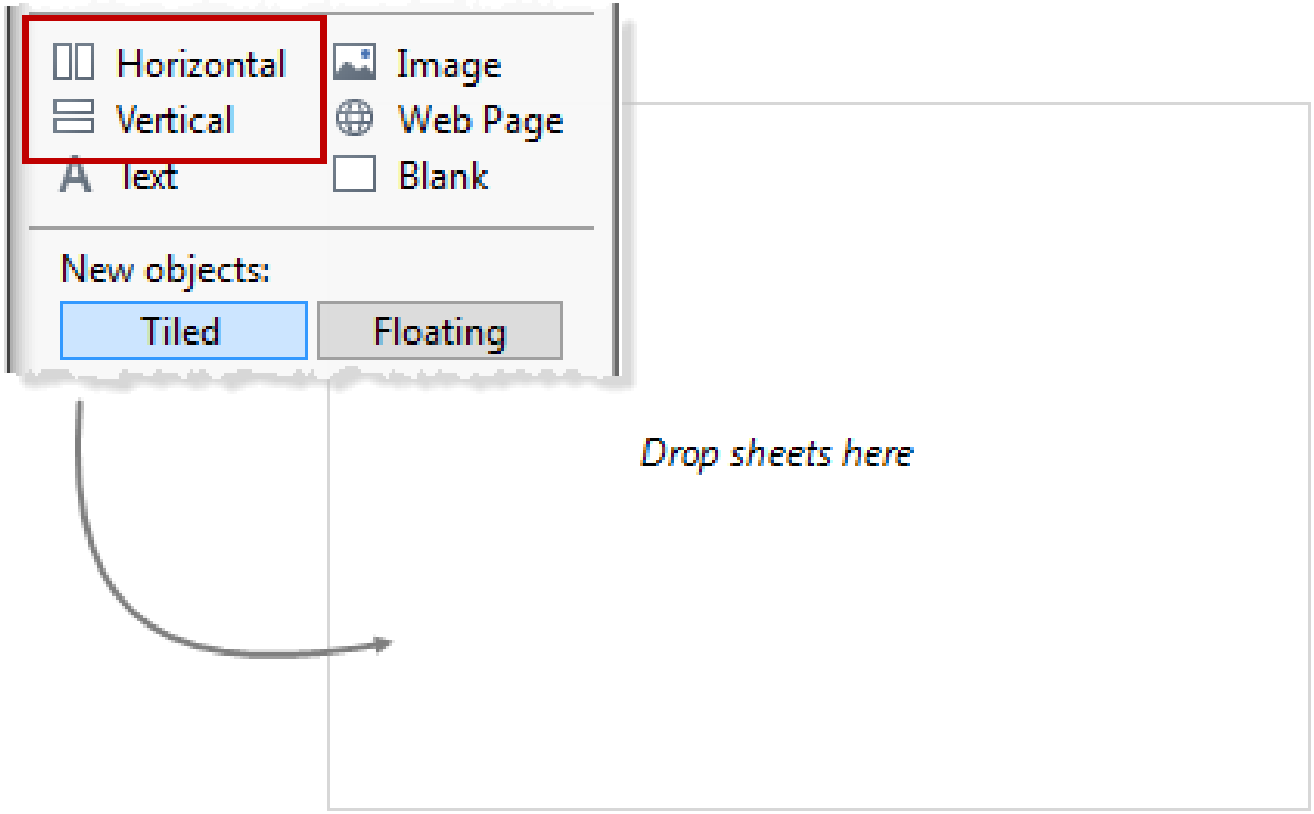
What is a story?

- In Tableau a story is a sequence of visualizations that work together to convey information.
- You can create stories to tell a data narrative, provide context, demonstrate how decisions relate to outcomes or to simply make a compelling case.

Image: Tableau Public

Layout Containers

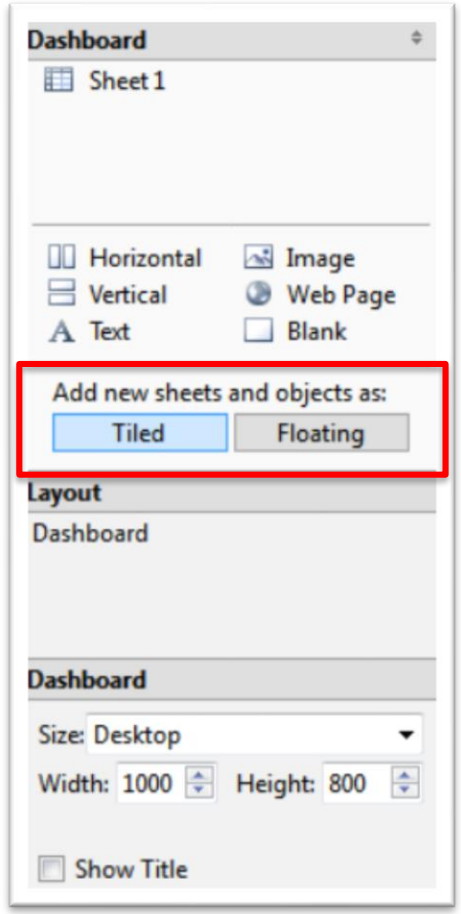
- A good way to organize your dashboard is with Layout containers!
- These containers create an area in the dashboard where objects automatically adjust their size and position based on the other objects in the container.



Layout Containers

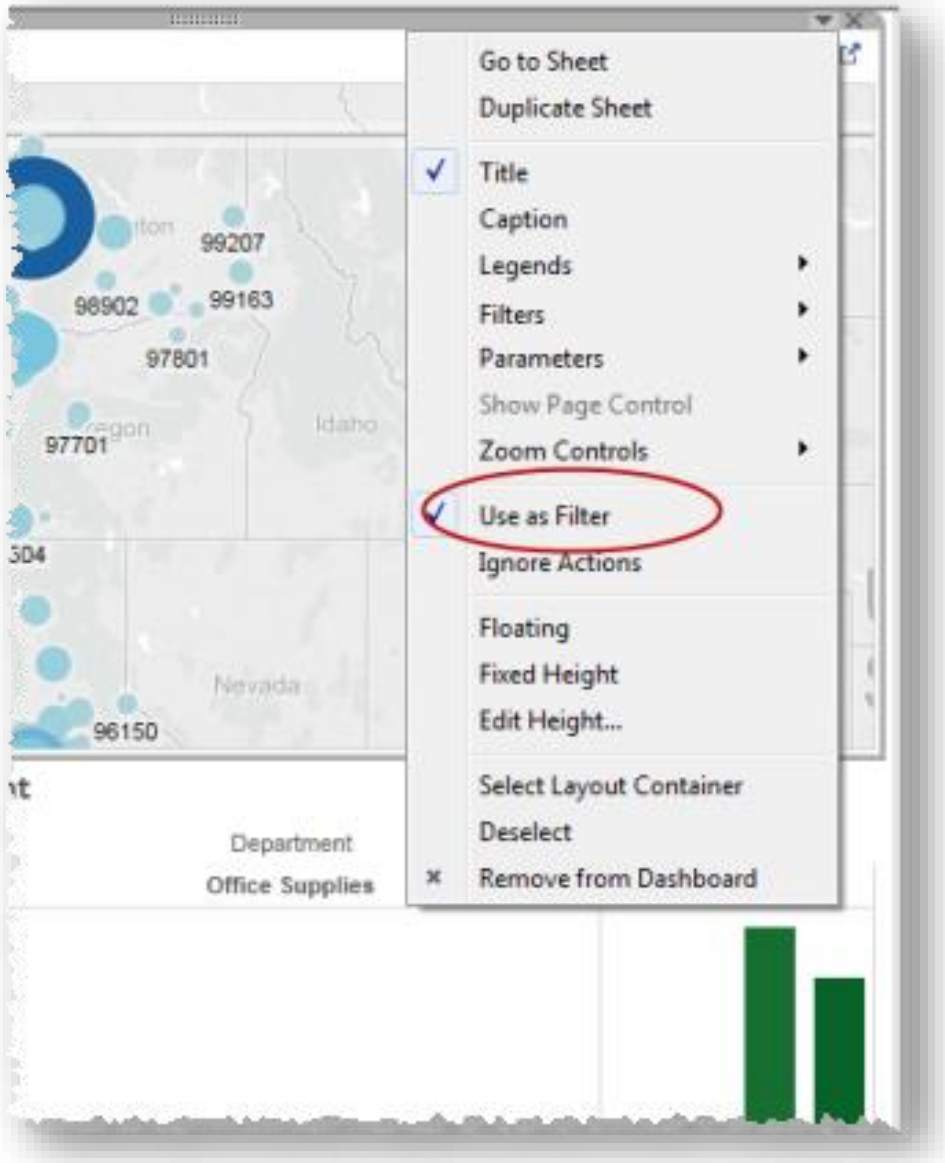
What if I simply want to drag and drop sheets onto the dashboard?

We can Switch to **Floating** Containers!



Filter Actions

- By default, filters only apply to the sheets they came from
- Let's change this!



Check your Understanding

What happens to the underlying sheets after a **Filter Action** has been applied to a view on a dashboard?

Dashboards…

- Dashboards allow you to present multiple sheets, text, images or web pages with some interactivity that ties all the views together
- When changes are made to a sheet on a dashboard, it changes the underlying sheet

Stories…

- Stories are narrated walkthroughs of a few sheets or dashboards that lead your audience through key insights from the underlying data
- When changes are made to a sheet on a story, it does not change the underlying sheet or dashboard

VISUALIZATION BEST PRACTICES

Above all else, show the data

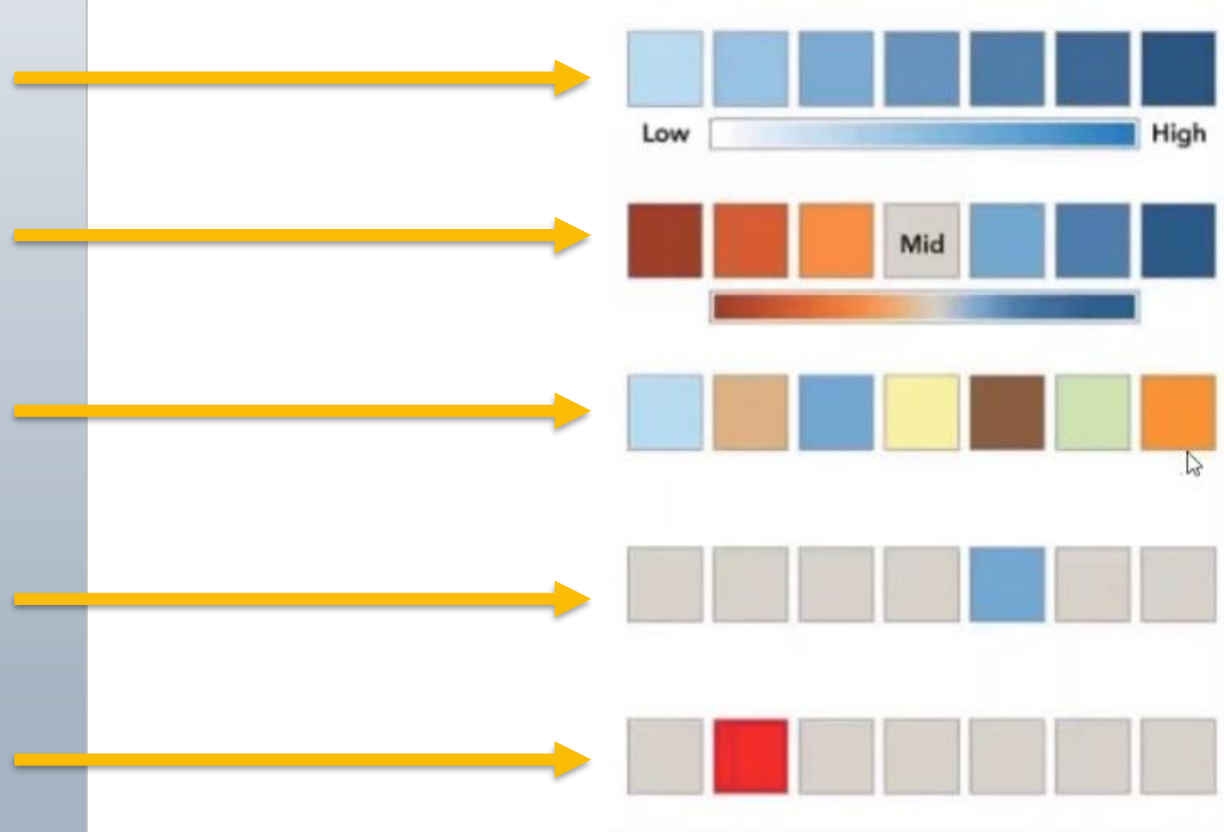
Concepts

- **Color:** Using color properly is an effective way to draw your viewers in and can help them better understand the underlying data
- **Preattentive Properties:** A preattentive visual property is one that is processed subconsciously by a viewer and can be harnessed to make it easier for a viewer to understand a visualization through design
- **Data-Ink Ratio:** The proportion of ink that is used to present actual data compared to the total amount of ink used in the entire display

Effective Use of Color

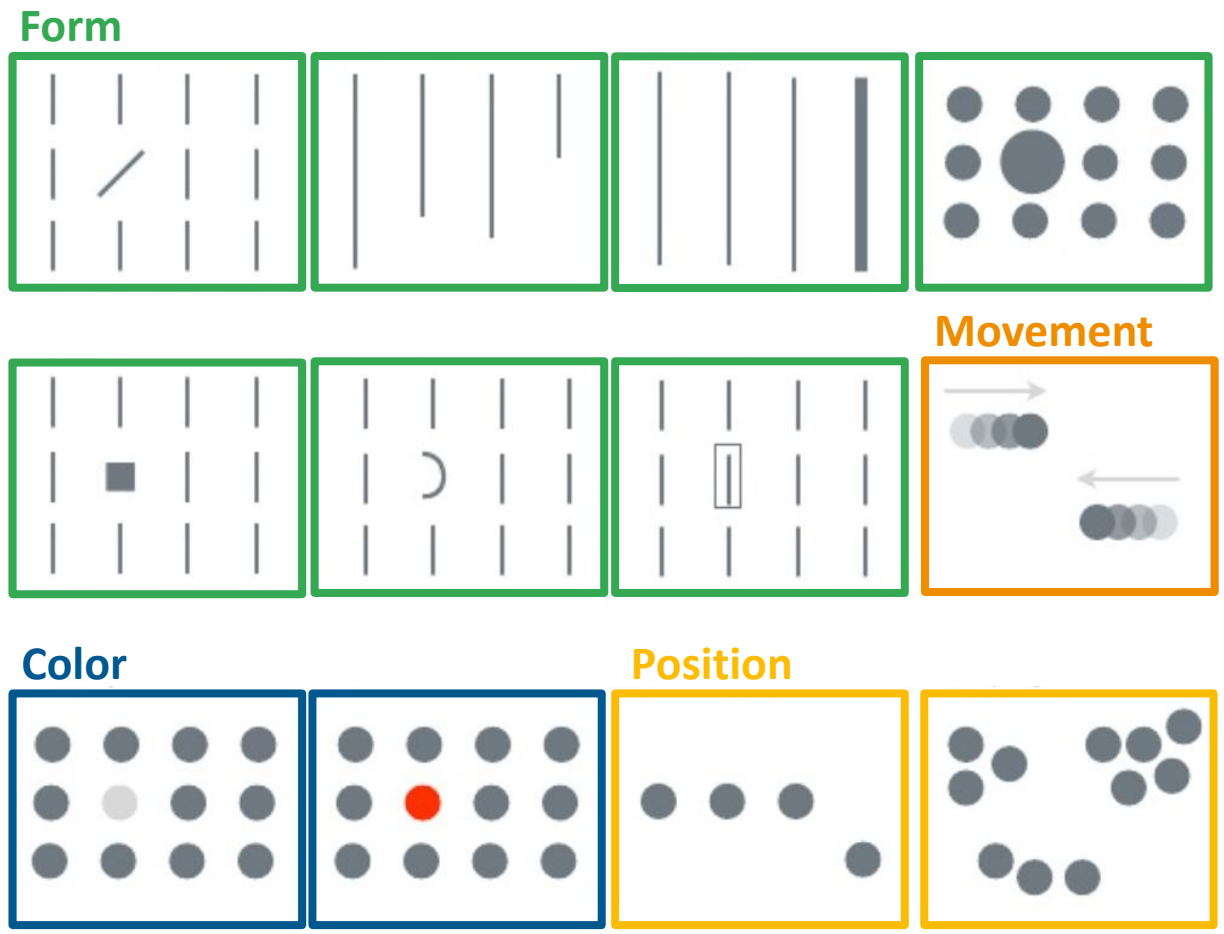
Effective Use of Color

- **Sequential**: Color is ordered from low to high
- **Diverging**: Two sequential colors with a neutral midpoint
- **Categorical**: Contrasting colors for individual comparison
- **Highlight**: Color used to highlight something
- **Alert**: Color used to get attention



Preattentive Properties

- A preattentive visual property is one which is processed in spatial memory without our conscious action
- There are four preattentive visual properties:
  - Color
  - Form
  - Movement
  - Spatial Positioning



Data-Ink Ratio

A large share of ink on a graphic should present data-information, the ink changing as the data change. Data-ink is the non-erasable core of a graphic, the non-redundant ink arranged in response to variation in the numbers represented.

Tufte, 1983

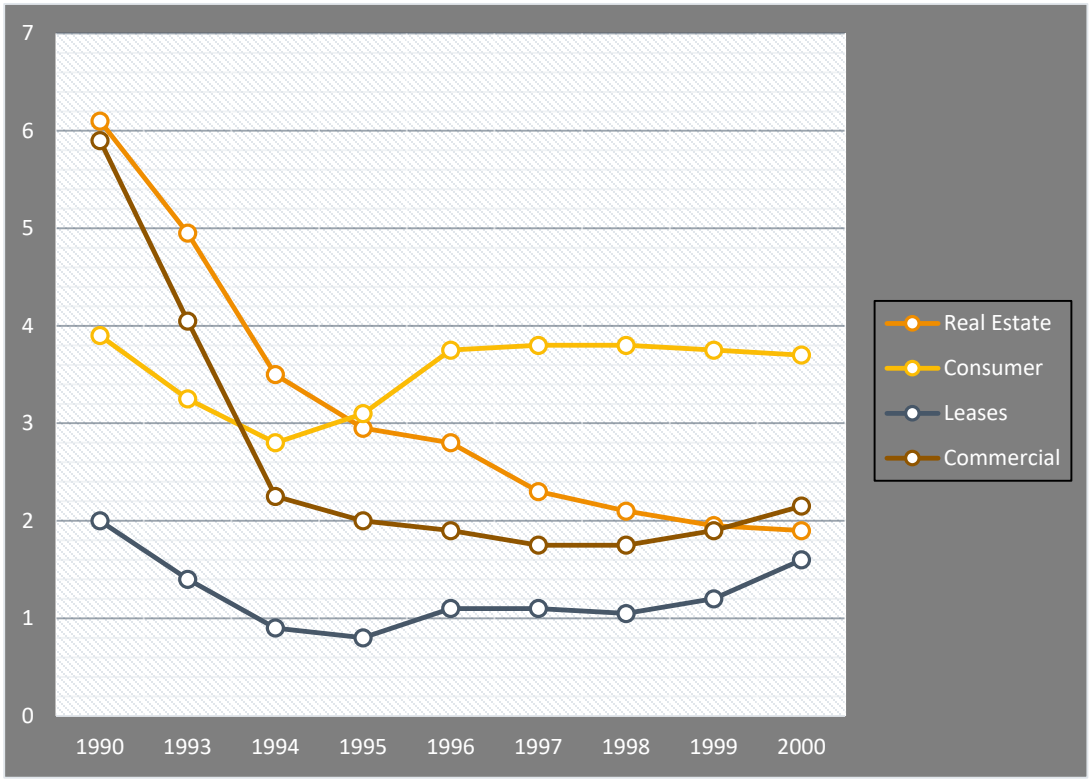
Data-ink ratio = data-ink / total ink (pixels)

- = proportion of ink (pixels) devoted to non-redundant display of information
- = 1.0 – proportion of a graphic that can be erased without loss of information

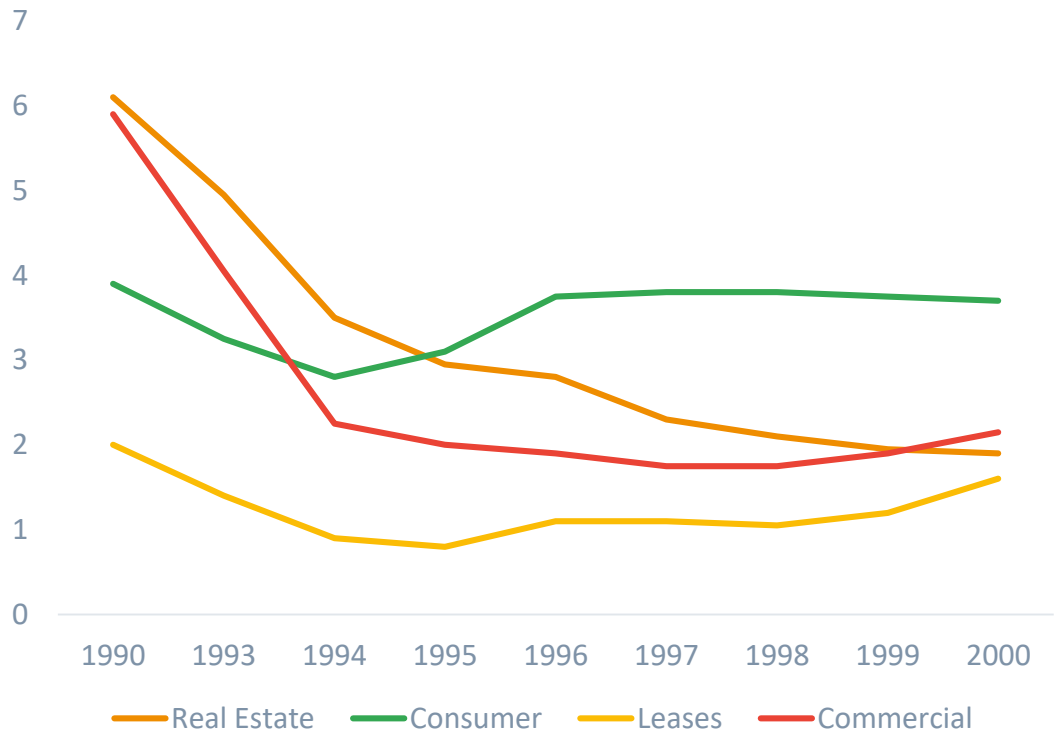
Maximize the data-ink ratio, within reason

- Erase non-data-ink, within reason
- Erase redundant data-ink, within reason

Low Data-Ink



High Data-Ink



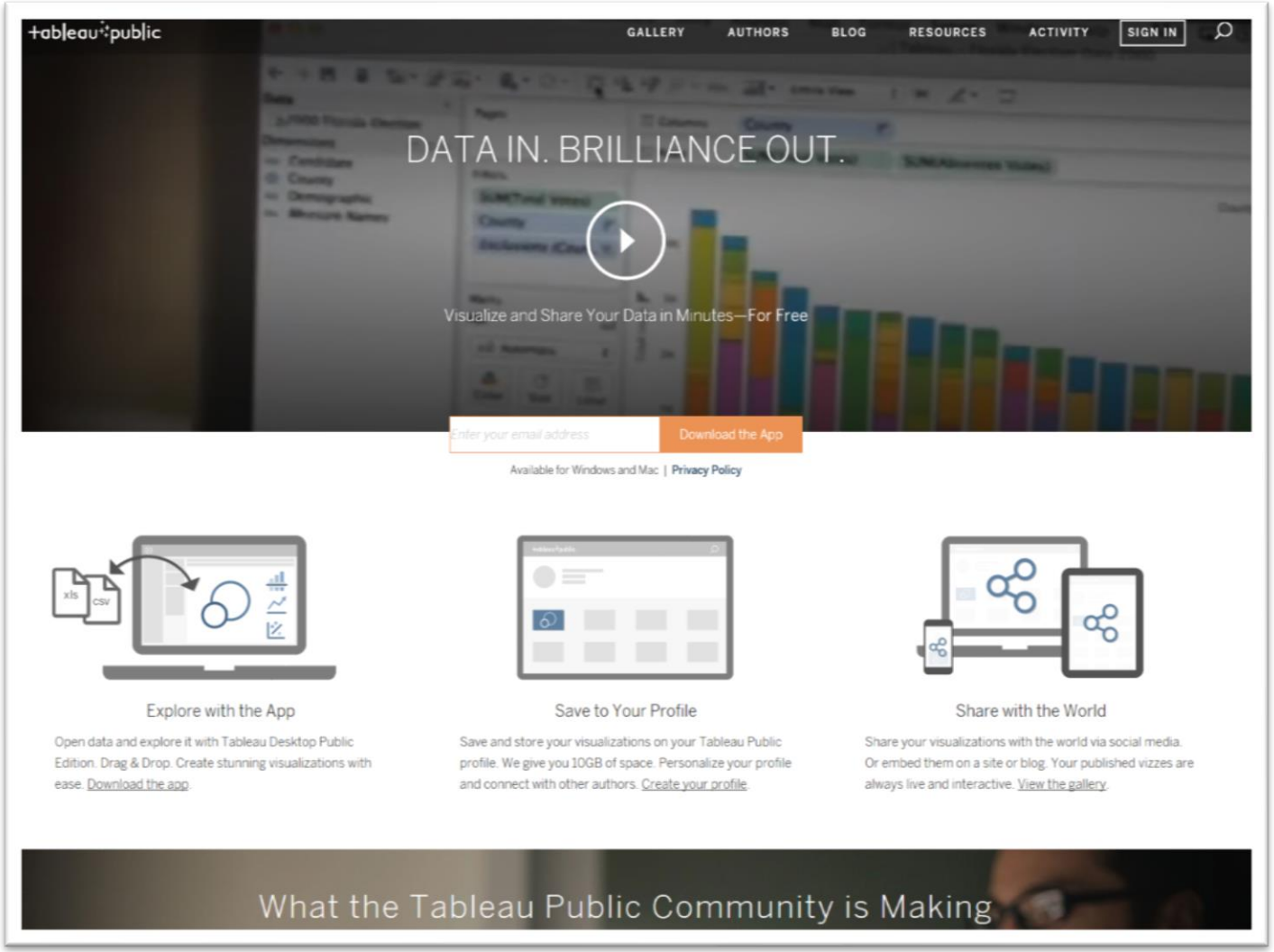
PARTICIPATE IN OUR COMMUNITY!

Tableau Public

Tableau Public

Click here to sign in or create a free account to access the Tableau community of professionals where you can publish work, browse others' work, ask questions and get inspired

Global Market Overview



What we've learned

- Connecting to your data
- Understanding data parts
- Creating basic calculations
- Joining data from multiple data sources
- Using the user interface to create impactful visualizations
- Develop dynamic dashboards & stories
- Leveraging mapping capabilities

Charts covered:

- Bar Chart
- Line Chart
- Word Cloud
- Dot Map
- Heat Map
- Timeline
- Path Maps
- Heat Maps

Continue Learning!

The following Apogee University Courses cover additional topics such as **field & chart types**, **calculations & LOD calculations**, **statistical modeling & forecasting**, **advanced mapping techniques**, **integrating R & Python**, and **MORE** to successfully prepare you to become a Tableau Desktop Qualified Associate.

- Tableau Essentials II
- Advanced Tableau I
- Advanced Tableau II

To learn more, visit our website.