



QuantUniversity, LLC

[www.quantuniversity.com](http://www.quantuniversity.com)

QuantUniversity's  
**THE ANALYTICS CERTIFICATE  
PROGRAM**

REQUEST MORE DETAILS

# Visualization with Tableau

## Presented By:

Sri Krishnamurthy, CFA, CAP  
[www.QuantUniversity.com](http://www.QuantUniversity.com)  
[sri@quantuniversity.com](mailto:sri@quantuniversity.com)

Materials prepared for the  
Analytics Certificate program  
[www.analyticscertificate.com](http://www.analyticscertificate.com)

# Visualization with Tableau

## Introduction



# Visualization

- Visualization is considered the last step of knowledge discovery, in which representation techniques are used to present the minded knowledge to the user
- Discovered knowledge should be expressed in high-level languages, visual representations, or other expressive forms so that the knowledge can be easily understood and directly used by humans
- With the Power View tool, you should be able to gain insight by using charts and graphs to view your data in imaginative ways
- The purpose of data visualization is to portray data graphically to capture hidden trends and to show patterns clearly



## About Tableau software

- Tableau Software helps people see and understand data, and is also considered a very new and innovative approach to business intelligence
- Tableau helps anyone at any skill level to quickly analyze, visualize, and share information by just dragging and dropping the components of the data
- Spreadsheets, databases, and even big data sources are easily combined into visualization reports and dashboards by using Tableau
- Many individuals and businesses use Tableau to share data on their blogs and websites



# Visualization with Tableau

How to connect data:

- ✓ Connecting to datasets
- ✓ Data preparation
- ✓ Editing data connection and data sources
- ✓ Editing metadata and saving data sources



## Connecting to datasets

- Tableau can connect to several data sources
- How to pull up a file into a data connection window:
  - After loading and saving any dataset (e.g; Superstore) into your computer you can access your file by choosing the proper option from the Connect menu (Excel, Text File, Access, etc.)
  - Now Tableau brings us up to the data connection window
  - You can see different sheets and work with single or multiple tables
  - You may play with columns to rename, change the data type, split, etc.
  - Now you can move forward to the worksheet page by clicking on 'work sheets' at the bottom left side



# Connecting to datasets

Tableau Public - Book1

File Data Window Help

Orders (Superstore)

Connected to Excel!

Workbook Superstore.xls

Sheets Enter sheet name

Orders Returns Users

Orders

Data doesn't look right? Tableau Data Interpreter might be able to help. Turn on

Copy Show aliases Show hidden fields Rows 8,399

Row ID #	Order ID #	Order Date	Order Priority	Order Quantity #	Sales #	Discount #	Ship Mode	Profit #	Unit Price #
1	3	10/13/2010	Low	6	261.54	0.040000	Regular Air	-213.25	38.9
2	6	2/20/2012	Not Specified	2	6.93	0.010000	Regular Air	-4.64	2.0
3	32	7/15/2011	High	26	2,808.08	0.070000	Regular Air	1,054.82	107.5
4	32	7/15/2011	High	24	1,761.40	0.090000	Delivery Truck	-1,748.56	70.8
5	32	7/15/2011	High	23	160.23	0.040000	Regular Air	-85.13	7.9
6	32	7/15/2011	High	15	140.56	0.040000	Regular Air	-128.38	8.4
7	35	10/22/2011	Not Specified	30	288.56	0.030000	Regular Air	60.72	9.1

Data Source Sheet 1

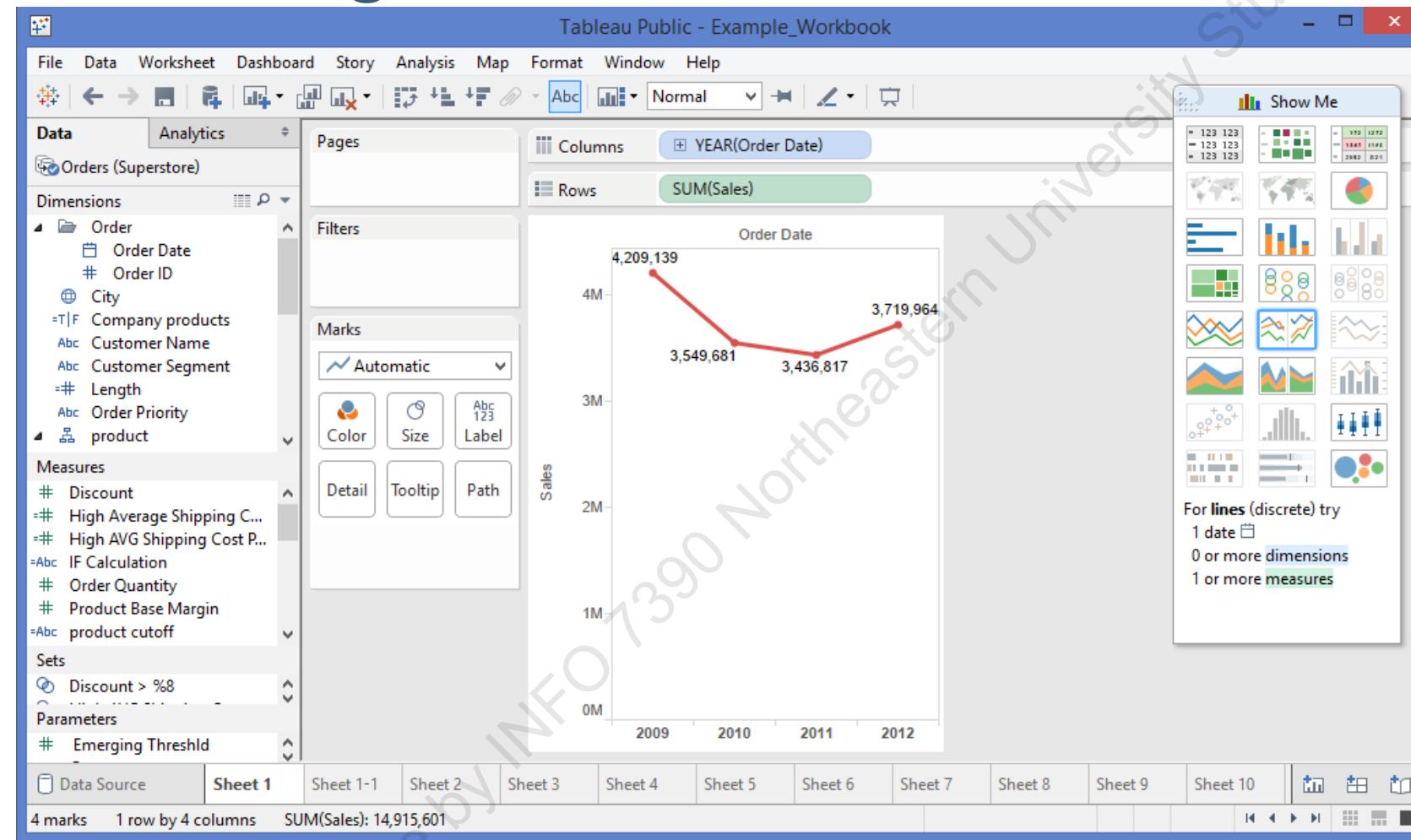


## Connecting to datasets

- Now you see your column names to the left and can start your analysis
- You can drop any rows and columns and play with features such as color, size, label, etc.
- As mentioned, you may connect different sheets in Tableau with different types of relationships such as inner or left. You can also change the join source. Right or full outer join may be available if the data source supports it
- If the dataset has lots of features you may search based on the name of the features by using the search option in worksheet page



# Connecting to datasets



See [Example.twbx—Sheet 1](#)



# Connecting to datasets

The screenshot shows a data connection window titled "Join". It displays four join types: Inner, Left, Right, and Full Outer. The "Left" join type is selected. Below the join type, there are two dropdown menus: "Data Source" and "Returns". The "Data Source" dropdown contains the following items:

- City
- Customer Name
- Customer Segment
- Discount
- Order Date
- Order ID
- Order Priority
- Order Quantity
- Product Base Margin
- Product Category
- Product Container
- Product Name
- Product Sub-Category
- Region
- Row ID
- Ship Date
- Ship Mode
- State
- Zip Code

The "Returns" dropdown contains the following items:

- # Order ID (Returns)
- Abc Status
- Abc Measure Names

A blue arrow points from the "Order ID" item in the "Data Source" dropdown to the "# Order ID (Returns)" item in the "Returns" dropdown.

Data connection window

The screenshot shows the Power BI desktop interface in the "Data" view. On the left, there is a navigation pane with the following sections:

- Data**: Shows "Orders (Superstore)" and "Superstore".
- Analytics**: Shows "City", "Customer Name", "Customer Segment", "Discount", "Number of Records", "Product Base Margin", "Product Category", "Product Container", "Product Name", "Product Sub-Category", "Shipping Cost", "Unit Price", and "Zip Code".
- Dimensions**: Shows "Orders\$" and "Returns\$".
- Measures**: Shows "Discount", "Order Quantity", "Product Base Margin", "Profit", "Sales", "Shipping Cost", and "Unit Price".

On the right, there is a search bar with the placeholder "Enter Text to Search" and a "q" icon. Below the search bar, the results show "City", "Customer Name", "Customer Segment", "Discount", "Number of Records", "Product Base Margin", "Product Category", "Product Container", "Product Name", "Product Sub-Category", "Shipping Cost", "Unit Price", and "Zip Code".

The screenshot shows the Power BI desktop interface in the "Analytics" view. On the left, there is a navigation pane with the following sections:

- Data**: Shows "Orders (Superstore)" and "Superstore".
- Analytics**: Shows "City", "Customer Name", "Customer Segment", "Discount", "Number of Records", "Product Base Margin", "Product Category", "Product Container", "Product Name", "Product Sub-Category", "Shipping Cost", "Unit Price", and "Zip Code".
- Dimensions**: Shows "Orders\$" and "Returns\$".
- Measures**: Shows "Discount", "Order Quantity", "Product Base Margin", "Profit", "Sales", "Shipping Cost", and "Unit Price".

On the right, there is a search bar with the placeholder "Enter Text to Search" and a "q" icon. Below the search bar, the results show "Ship Date", "Ship Mode", "State", "Zip Code", "Returns\$", "# Order ID (Returns)", "Abc Status", and "Abc Measure Names".

Search option



## Data Preparation

- Tableau prefers a tabular form of data, such as raw data which sleeps in databases for optimal analysis
- However there are several features in the Tableau desktop to reshape the text or excel file for better analysis. To do this:
  - The data interpreter can be applied by clicking on “turn on” button on data connector window
  - By clicking on ‘review results’ you will see the changes on your sheet
  - If you go to your specific sheet, you’ll see which fields have been used as header in red, and those considered as data in green
  - Also, you have the pivot option to do some aggregation



# Data Preparation

The screenshot shows the Tableau Data Interpreter interface for the 'Orders' data source. At the top, there's a header bar with the title 'Orders'. Below it, a message says 'Data Interpreter is on.' with buttons for 'Review results...' and 'Turn off'. A large blue arrow points from the 'Turn off' button towards the explanatory key on the right. The main area displays a table with columns: 'Field Name', 'Table', and 'Remote Field Name'. The table lists fields such as Row ID, Order ID, Order Date, Order Priority, Order Quantity, Sales, Discount, Ship Mode, and Profit. Most fields are mapped to the 'Orders' table, except for Sales, which is mapped to a separate 'Sales' table.

Field Name	Table	Remote Field Name
# Row ID	Orders	Row ID
# Order ID	Orders	Order ID
Order Date	Orders	Order Date
Order Priority	Orders	Order Priority
# Order Quantity	Orders	Order Quantity
# Sales	Orders	Sales
# Discount	Orders	Discount
Ship Mode	Orders	Ship Mode
# Profit	Orders	Profit

**Key for Understanding the Data Interpreter Results**

Use the key to understand how your data source has been interpreted.  
To view the results, click a worksheet tab.  
Note: Tableau never makes changes to your underlying data source.

**Key:**

- Data is interpreted as column headers (field names).
- Data is interpreted as values in your data source.
- Data is ignored and not included as part of your data source.
- Data has been excluded from your data source.

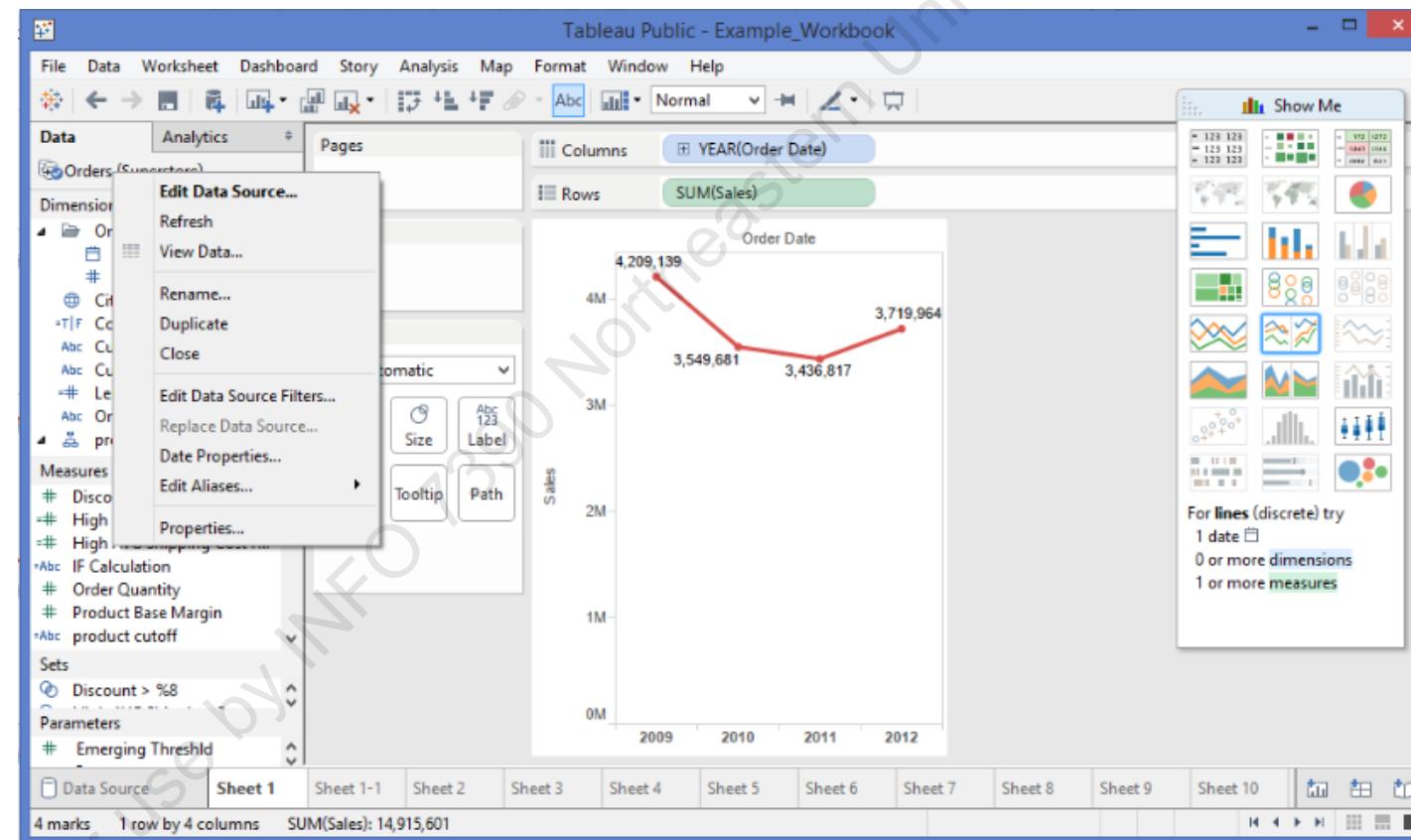
Note: To search for all excluded data, use CRTL +F on Windows or Command F on the Mac, and then type \*\*\*DATA REMOVED\*\*\*.

Row ID	Order ID	Order Date	Order Pric	Order Qua	Sales	Discount	Ship Mode	Profit
1	3	##### Low	6	261.54	0.04	Regular Ai	-213.25	
2	6	##### Not Speci	2	6.93	0.01	Regular Ai	-4.64	
3	32	##### High	26	2808.08	0.07	Regular Ai	1054.82	
4	32	##### High	24	1761.4	0.09	Delivery T	-1748.56	
5	32	##### High	23	160.2335	0.04	Regular Ai	-85.13	
6	32	##### High	15	140.56	0.04	Regular Ai	-128.38	
7	35	##### Not Speci	30	288.56	0.03	Regular Ai	60.72	
8	35	##### Not Speci	14	1892.848	0.01	Regular Ai	48.99	
9	36	##### Critical	46	2484.746	0.1	Regular Ai	657.48	
10	65	##### Critical	32	3812.73	0.02	Regular Ai	1470.30	
11	66	##### Low	41	108.15	0.09	Regular Ai	7.57	
12	69	6/3/2009 Not Speci	42	1186.06	0.09	Regular Ai	511.69	
13	69	6/3/2009 Not Speci	28	51.53	0.03	Express Ai	0.35	
14	70	##### Low	48	90.05	0.03	Regular Ai	-107.00	
15	70	##### Low	46	7804.53	0.05	Regular Ai	2057.17	
16	96	##### High	37	4158.124	0.01	Regular Ai	1228.89	
17	97	##### Medium	26	75.57	0.03	Regular Ai	28.24	
18	129	##### Low	4	32.72	0.09	Regular Ai	-22.59	
19	130	5/7/2012 High	3	461.89	0.05	Express Ai	-309.82	
20	130	5/7/2012 High	29	575.11	0.02	Regular Ai	71.75	
21	130	5/7/2012 High	23	236.46	0.05	Regular Ai	-134.31	
22	132	##### Medium	27	192.814	0.03	Regular Ai	-86.20	



# Editing data connection and data sources

- By right clicking on ‘data’ you can edit the data connection
- An important feature is the ability to replace the data source



## Editing metadata and saving data sources

- Tableau divides the columns into dimensions and measures, as descriptive and quantitative
- Tableau can't write changes to data sources. However, we can modify metadata for use in Tableau
- You can rename fields, create hierarchy, create folders, and also change the data type of any column
- Also, you can change the data type on the data source tab
- You can save all these changes to data sources as .tds. Note that it is just about the information of data connection



# Editing metadata and saving data sources (Hierarchies)

The screenshot shows the Tableau Public interface with the title "Tableau Public - Example\_Workbook". The Data pane on the left lists dimensions like Company products, Customer Name, and Product categories, along with measures such as SUM(Sales) and SUM(Profit). The Marks card indicates an automatic mark type with color, size, and label options. The main workspace displays a horizontal bar chart with three categories: Furniture, Office Supplies, and Technology. The Furniture category has a total sales value of 5,178,591 and a profit of 117,433. The Office Supplies category has a total sales value of 3,752,762 and a profit of 886,314. The Technology category has a total sales value of 5,984,248 and a profit of 117,433. The Show Me panel on the right suggests various chart types for scatter plots.

Product Category	SUM(Sales)	SUM(Profit)
Furniture	5,178,591	117,433
Office Supplies	3,752,762	886,314
Technology	5,984,248	117,433

See [Example.twbx—Sheet 1-1](#)



# Visualization with Tableau

## Visual Analytics:

- ✓ Working with marks and building views
- ✓ Drill down and hierarchies
- ✓ Sorting and grouping
- ✓ Groups and sets
- ✓ Creating sets
- ✓ Working with sets



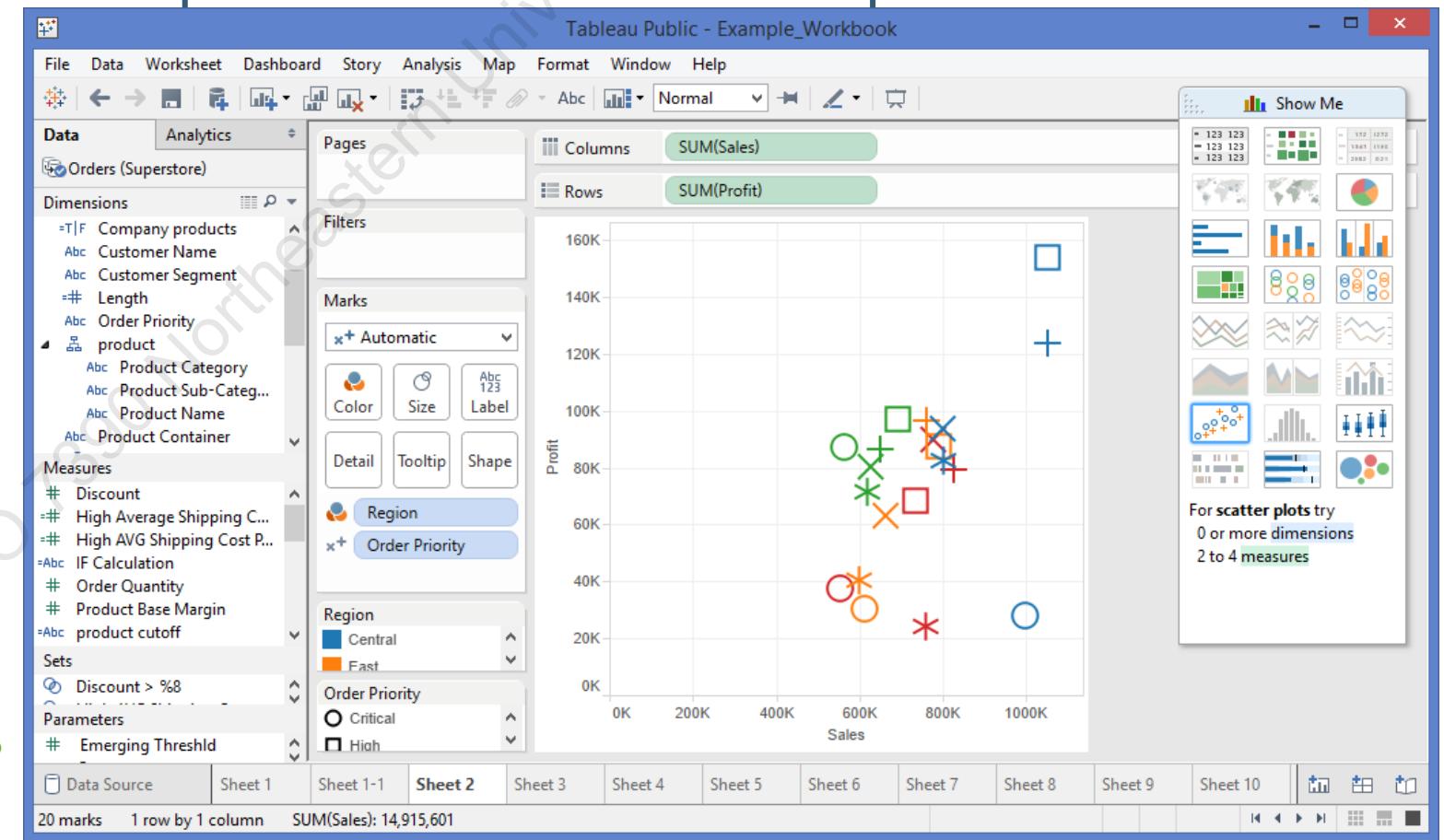
## Working with marks and building views

- There are two ways of analyzing data in Tableau:
  - Start exploring
  - Asking questions and try to answer them
- There are several mark types in the marks menu to analyze the data, Tableau automatically choose the best practice for primary analysis
- Marks allows you to build different views
- Dimensions define number of each mark
- Both dimensions and measures can be pulled up to the marks menu



# Working with marks and building views

- Scatter plot of sales vs. profit with different regions in different colors, and different order priorities in different shapes:



See [Example.twbx—Sheet 2](#)

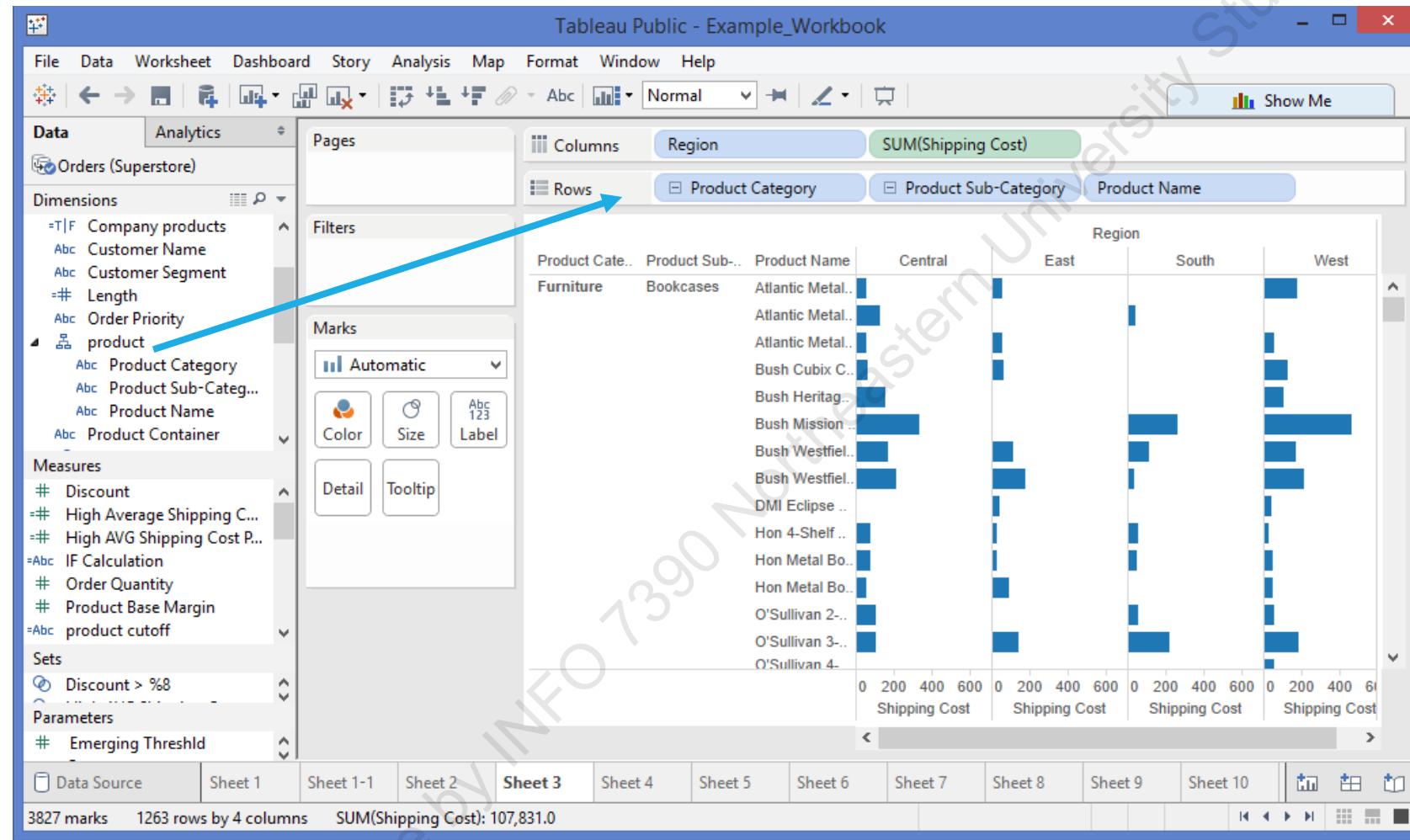


## Drill down and hierarchies

- By defining hierarchy and drilling down, we can go to the next step of details for analysis
- Date field has its hierarchy from year to quarter, month and week, by default
- Hierarchies can have multiple levels in Tableau
- We may have a separate field of hierarchy by easily copying and pasting the field outside of the hierarchy
- Dimensions define the number of each mark



# Drill down and hierarchies



See [Example.twbx—Sheet 3](#)



## Sorting & grouping

- There are several ways of sorting, the easiest way is to sort by access
- We may sort for a pill by using toolbar
- We can sort for discrete pill by right clicking on that and choose the ‘sort’ option
- By grouping we can combine different components of a field as a group
- In scatter plots we may choose different objects as a group
- We may use calculations and bins to create interactive and dynamic groups



# Sorting

Tableau Public - Example\_Workbook

The screenshot shows a Tableau Public dashboard titled "Example\_Workbook". The dashboard has a single sheet with a bar chart. The chart has "SUM(Sales)" in the columns shelf and "Product Category" in the rows shelf. The bars are color-coded by category. A tooltip for Furniture shows sales of 698,094. A tooltip for Office Supplies shows sales of 1,761,837. A tooltip for Technology shows sales of 1,889,314. A tooltip for Products shows sales of 2,168,697.

**Sort [Product Sub-Category]**

Sort order: Ascending

Sort by: Field (Sales)

Aggregation: Sum

Product Sub-Category	Sales
Furniture	698,094
Office Supplies	1,761,837
Technology	1,889,314
Products	2,168,697
Rubber Bands	1,889,314
Labels	2,168,697
Scissors, Rulers and Trimmers	
Pens & Art Supplies	
Envelopes	
Paper	
Appliances	
Binders and Binder Accessories	
Storage & Organization	
Computer Peripherals	
Copiers and Fax	
Telephones and Commu...	
Office Machines	

Sort product sub-category by sales

File Data Worksheet Dashboard Story Analysis Map Format Window Help

Data Analytics

Orders (Superstore)

Dimensions

Abc: Product Sub-Cat...  
Abc: Product Name  
Abc: Product Container  
Abc: Product Sub-Category ...  
Abc: Region  
# Row ID  
# Sales (bin)  
# Sales (bin) 2  
Ship Date  
Abc: Ship Mode  
State  
State (group)  
Zip Code  
Abc: Measure Names

Measures

#Abc: Profit Cutoff  
# Profit per Order ID  
# Rank of Profit  
# Running Sum (Sales)  
# Sales  
# sales - profit  
# Shipping Cost  
#Abc: Sign of profit  
# Time to ship  
# Unit Price

Sets

# Discount > %8  
# High AVG Shipping Cost

Parameters

# Emerging Threshld  
Abc: Company

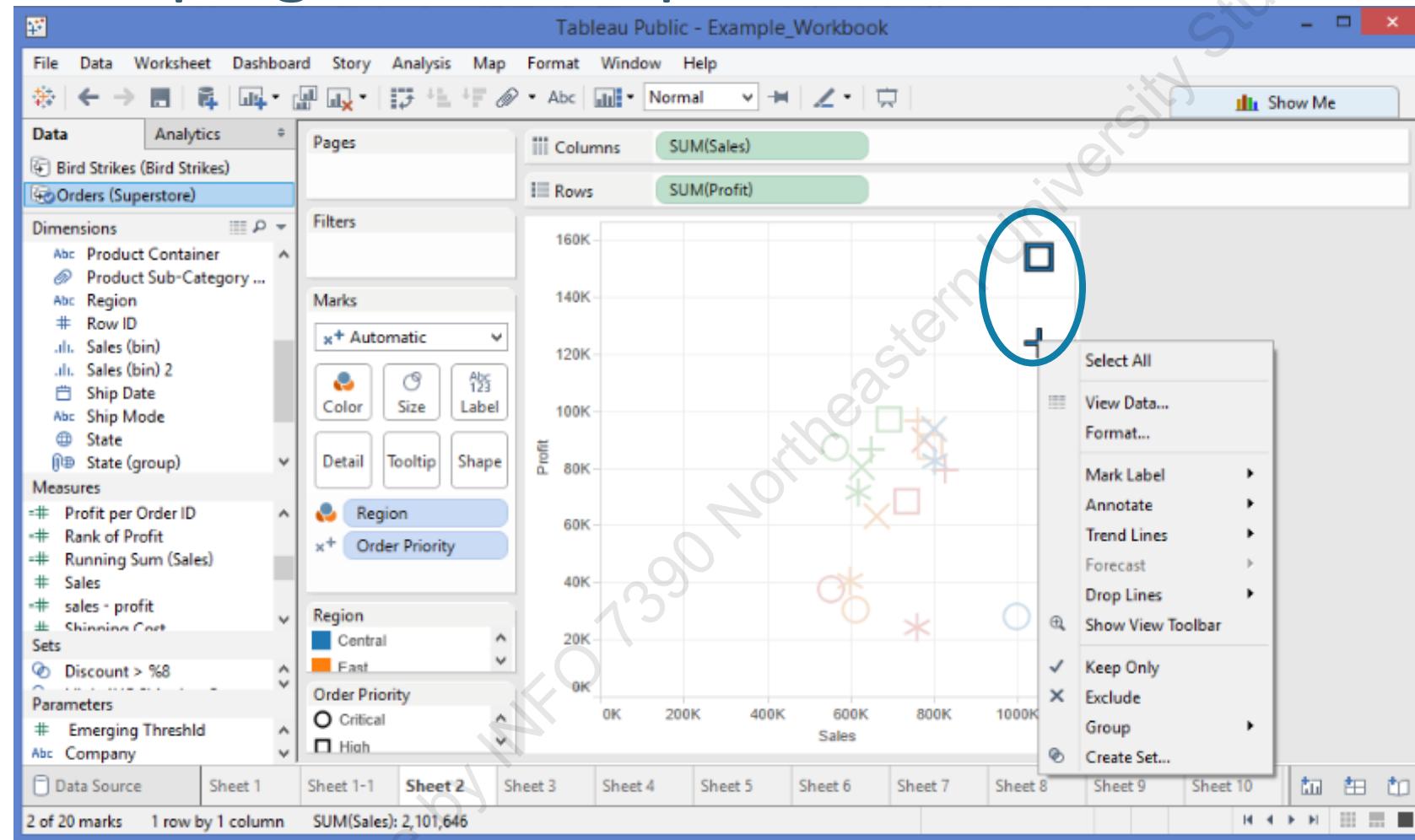
Data Source Sheet 1 Sheet 1-1 Sheet 2 Sheet 3 Sheet 4 Sheet 5 Sheet 13 Sheet 14 Sheet 15

17 marks 17 rows by 1 column SUM(Sales): 14,915,601

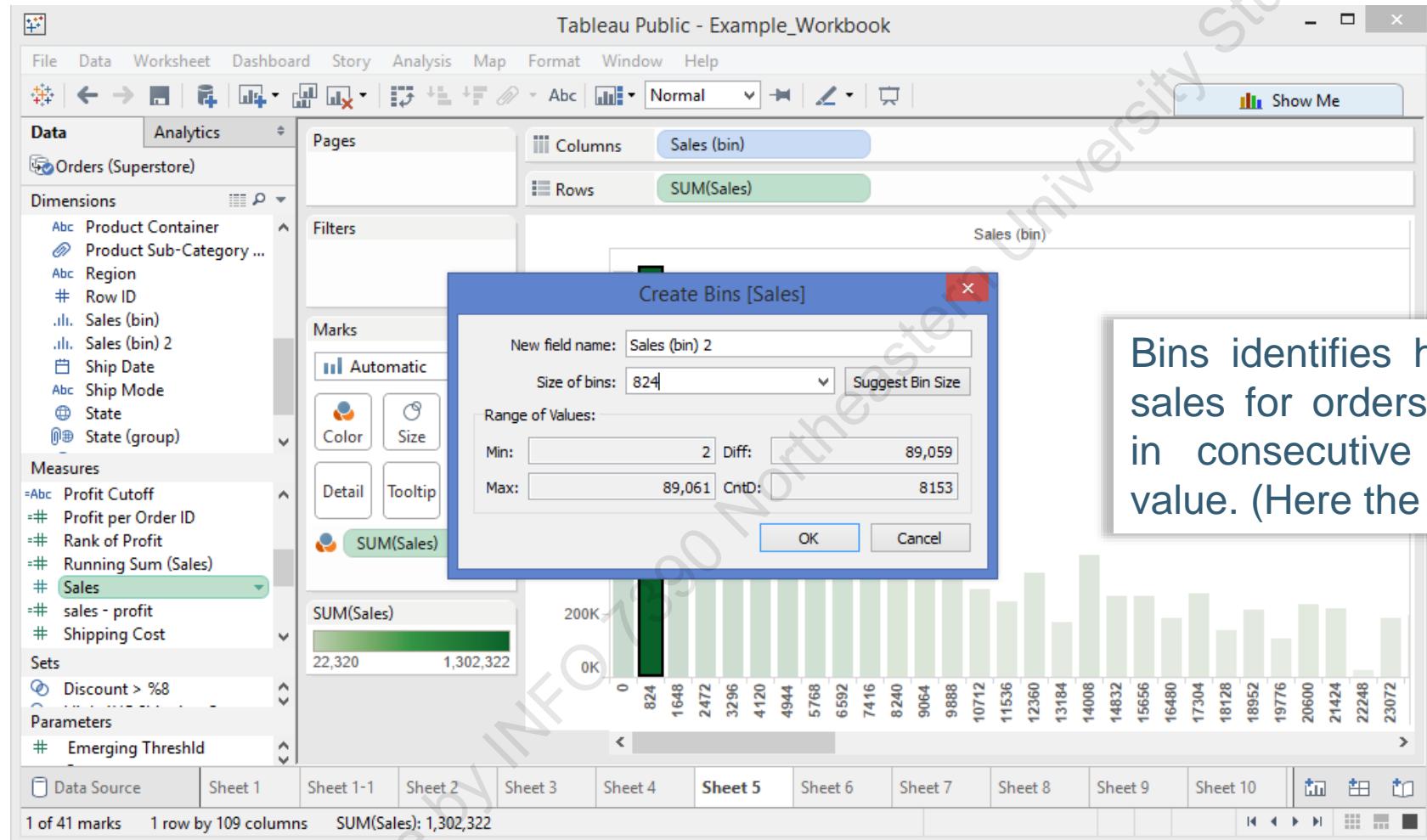
See [Example.twbx—Sheet 4](#)



# Grouping on scatterplot



# Grouping by bins



Bins identifies how much is the total sales for orders having sales amount in consecutive intervals of specific value. (Here the interval is 824)

See [Example.twbx—Sheet 5](#)



## Groups & sets

- Groups in Tableau can be part of Metadata definitions, while sets in tableau are a combination of data
- Groups can't be combined to create new groups, while sets can be combined to create new sets
- Groups can't be used in calculating fields, sets can be used in calculating fields
- Generally, sets are a more powerful type of groups

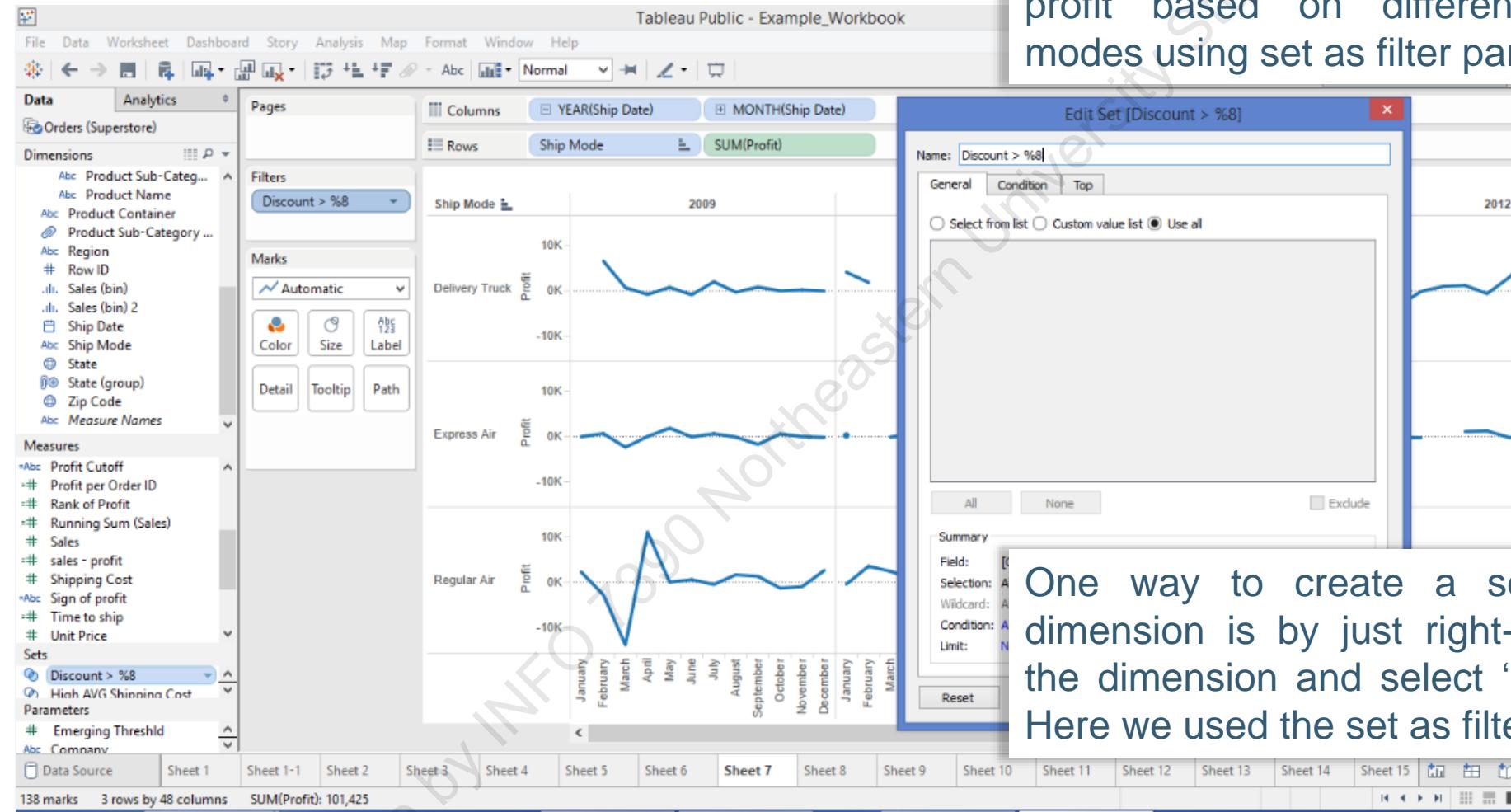


## Creating sets

- Sets are powerful analytical tools in visual analysis
- Sets in Tableau is a combination of different data, like filter results
- Groups can't be used in calculating fields, sets can be used in calculating fields
- Some questions should be answered by combining the sets
- Items in two combined set may be:
  - All members in both sets
  - Shared members in both sets
  - Members of first set not in second set
  - Members of second sets not in first set



# Sets



See [Example.twbx—Sheet 7](#)



## Working with sets

- Members can't be added to the existing set, but a second set can be created based on desirable members and these members can be added to a previous set
- Sets can be modified and saved by choosing 'edit set'
- Like fields, sets can be added to the marks
- Sets can be created from different filter results
- Like groups, sets can be added to hierarchies
- Unlike groups, sets can be also used in calculations



# Working with sets

Tableau Public - Example\_Workbook

The screenshot shows a Tableau dashboard with a line chart titled "Ship Date". The Y-axis is labeled "Profit" and ranges from -10K to 10K. The X-axis shows years from 2009 to 2012, with specific months visible for 2011 and 2012. The chart shows a general downward trend in profit over time, with a notable peak around January 2011. A tooltip for the calculation "IF [High AVG Shipping Cost] THEN [Profit]" is open, displaying the syntax and examples.

We can use a conditional set in IF statement.

Calculation1

```
IF [High AVG Shipping Cost] THEN [Profit] END
```

All

IF <expr> THEN <then> [ELSEIF <expr2> THEN <then2> ...] [ELSE <else>] END

Enter Text to Search

END

ENDSWITH

EXCLUDE

EXP

FIND

FINDNTH

Tests a series of expressions returning the <then> value for the first true <expr>.

Example: IF [Profit] > 0 THEN 'Profitable' ELSEIF [Profit] = 0 THEN 'Breakeven' ELSE 'Loss' END

Sheet 7

The calculation is valid.

Apply OK

138 marks 3 rows by 48 columns SUM(Profit): 101,425

See [Example.twbx—Sheet 7](#)



# Visualization with Tableau

## Visual Analytics:

- ✓ Ways to filter
- ✓ Parameters
- ✓ Formatting
- ✓ Trend lines
- ✓ Reference lines
- ✓ Forecasting

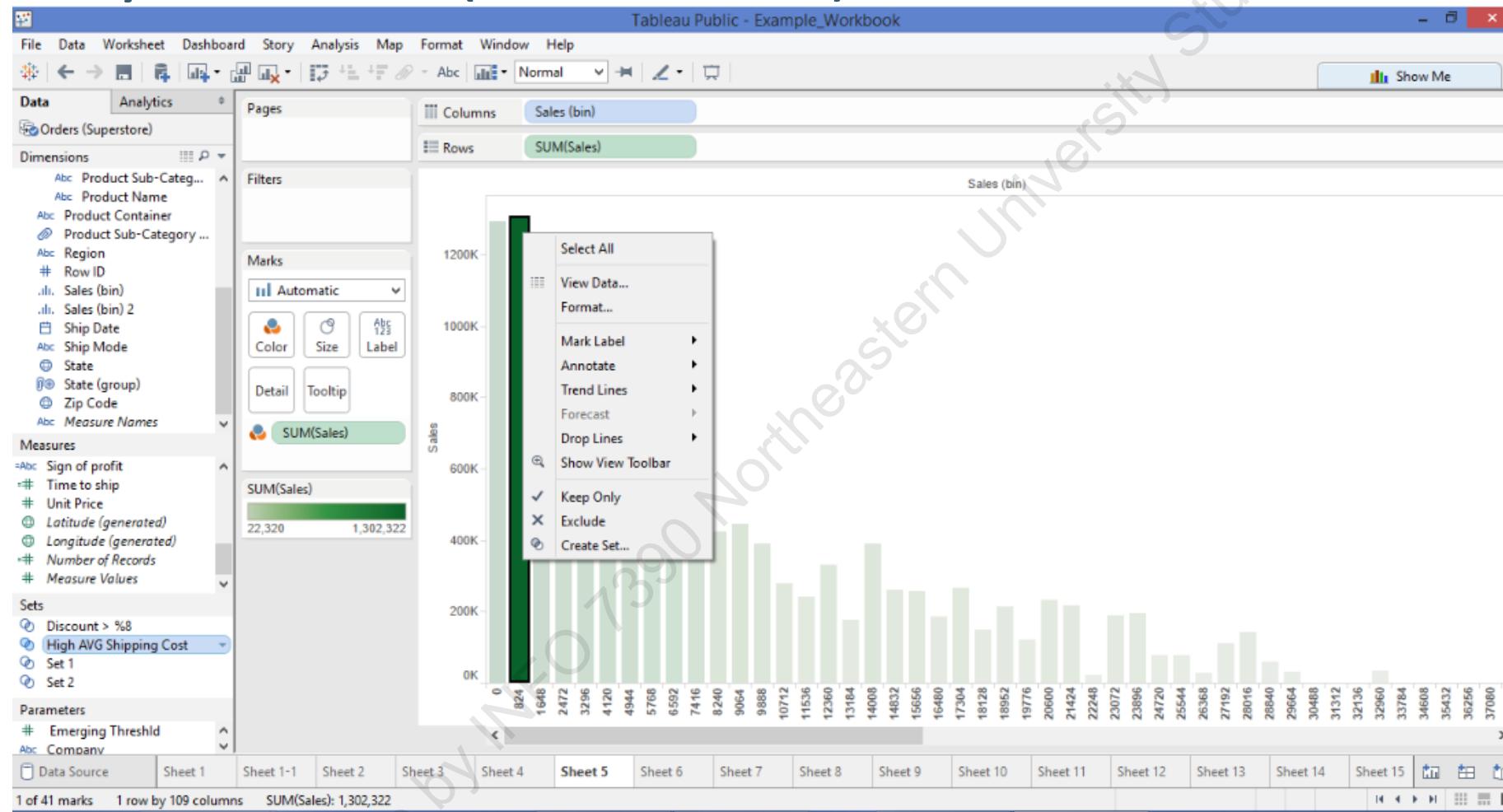


## Ways to filter

- Filter allows you to exclude members of a field based on specific characters
- Filters in Tableau are similar to the ‘where’ clause in SQL
- We can filter by :
  - Filter shelf
  - Quick filter
  - In the view
- To remove the filter, drag the pill off the shelf



# Ways to filter (in the view)



# Ways to filter (filter shelf)

Tableau Public - Example\_Workbook

File Data Worksheet Dashboard Story Analysis Map Format Window Help

Show Me

Data Analytics

Orders (Superstore)

Dimensions

- Customer Name
- Customer Segment
- Length
- Order Priority
- product
  - Product Category
  - Product Sub-Categ...
  - Product Name
  - Product Container
  - Product Sub-Category ...
- Region
- Row ID
- Sales (bin)
- Sales (bin) 2

Marks

- Automatic
- Color
- Size
- Label

Filters

Discount > %8

Pages

Columns

Rows

Ship Mode

Delivery Truck

Filter [Product Category]

General Wildcard Condition Top

None

By field:

Top 10 by Profit Sum

By formula:

Top 10 by Profit Sum

Reset OK Cancel Apply

Drag product category into filters

Ship Date

2011 2012

Express Air Regular Air

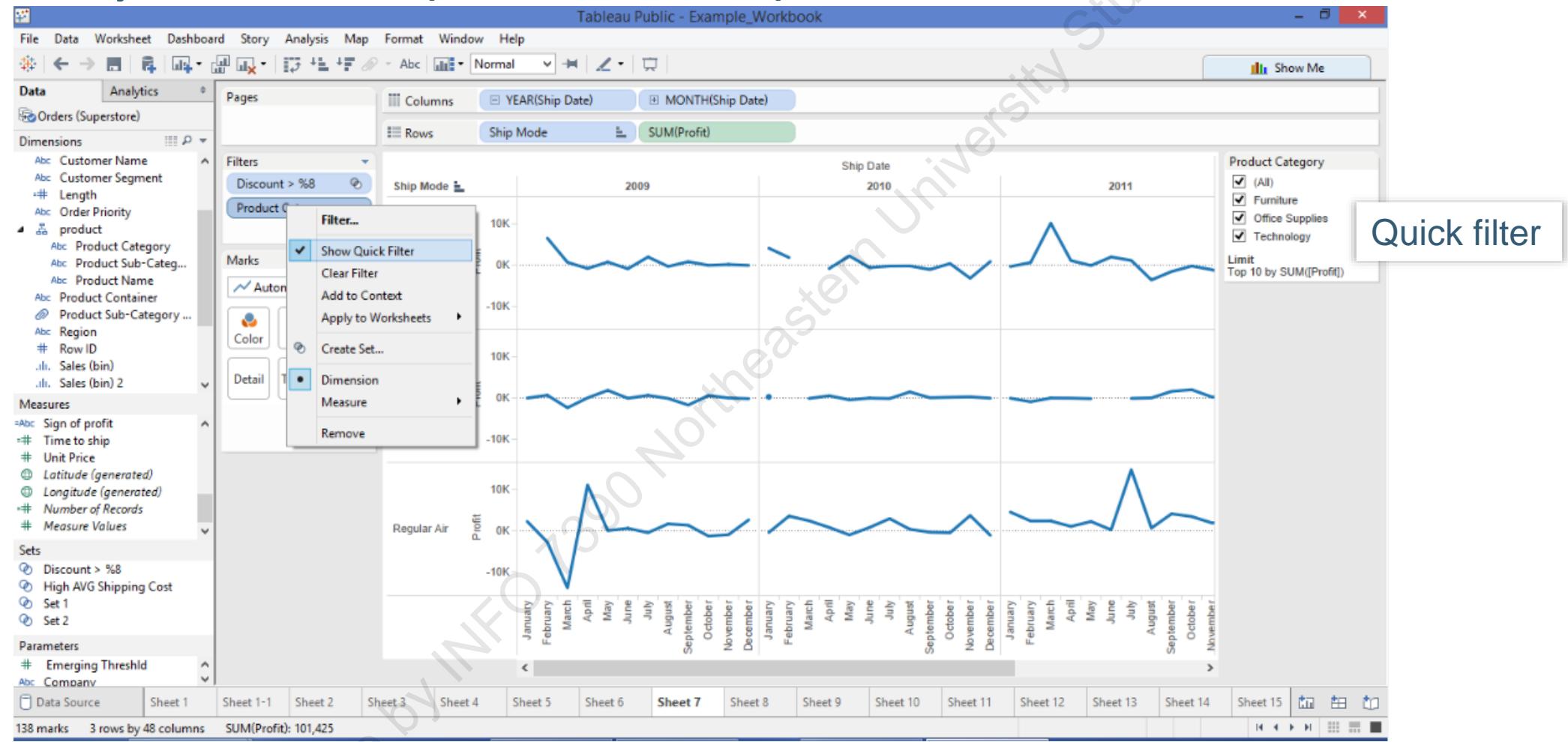
October November December January February March April May June July August September October November December January February March April May June July

138 marks 3 rows by 48 columns SUM(Profit): 101,425

For use by KMO7390 Northeastern University Students only



# Ways to filter (Quick Filter)



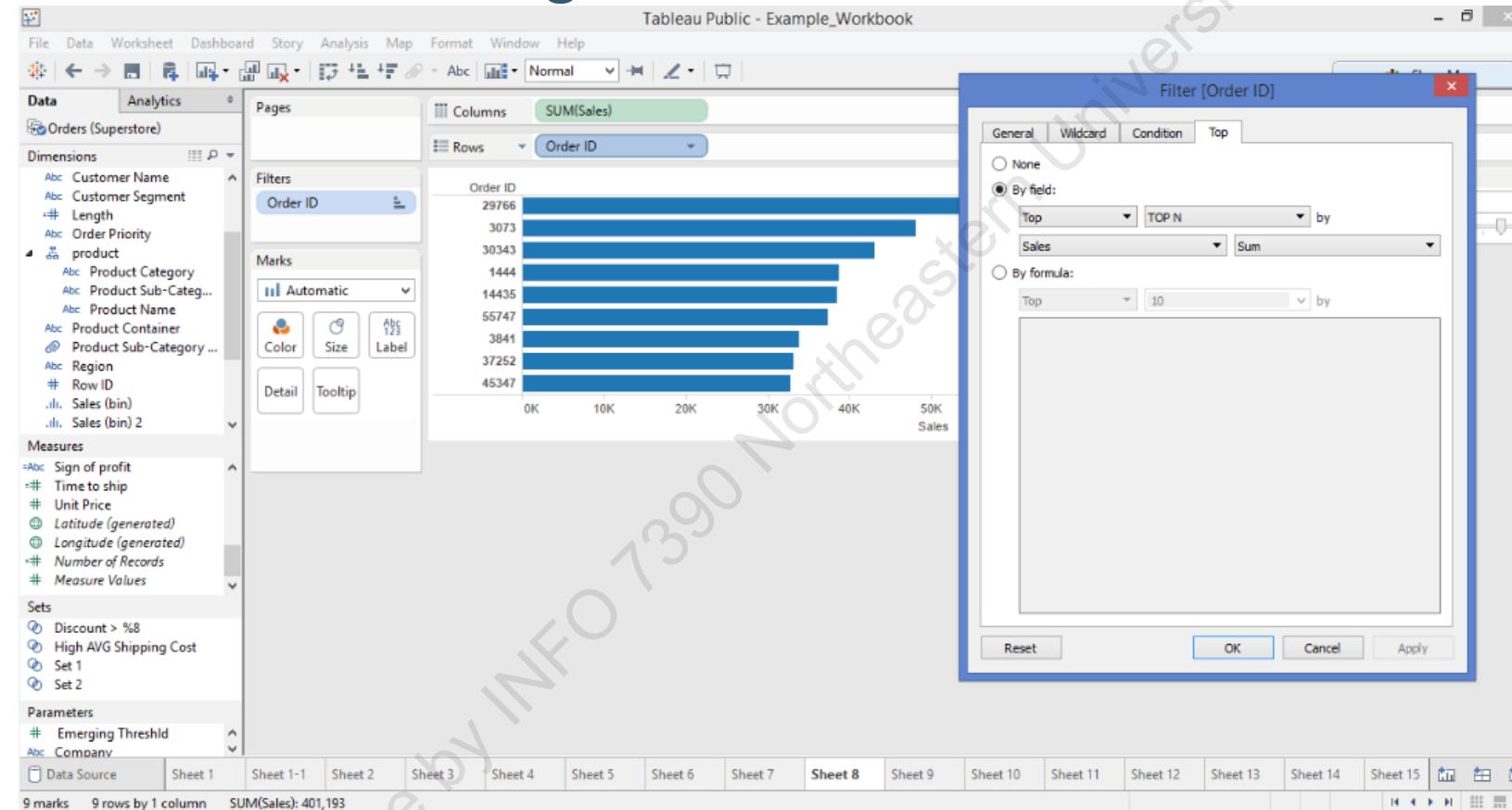
# Parameter

- We can think of parameters as variables in equations, in which their value can be controlled by the end user
- Parameters:
  - Add interactivity and flexibility to the workbook
  - Are not the same as filters, but can be used as filters
  - May also be used in definition of a set
  - Can be used as reference line
  - Supports variety of data types



# Parameter as filters

- The orders with largest sales:



See [Example.twbx—Sheet 7](#)



## Formatting

- Tableau provides different types of marks in mark shelf
- All properties of different marks can be adjusted
- In addition to provided marks, other features are available in Tableau, such as annotation
- You can modify the shape of the marks in Tableau such as color, text, format, and alignment
- Also, other components of view pane can be modified, such as filters, axis, mapping tools, etc.
- To do formatting you can also directly choose the “format” menu from the toolbar



# Formatting

Tableau Public - Example\_Workbook

File Data Worksheet Dashboard Story Analysis Map Format Window Help Show Me

Format Annotations

Text

Font: Arial, 8pt Alignment: Automatic

Box

Box: Four Sided Shading: None Border: Corners: Square

Line

Line: Line End: Arrow End Size:

Marks

Automatic

Color Size Label Detail Tooltip Path Region

Region

- Central
- East
- South
- West

Pages Columns MONTH(Order Date) Rows SUM(Sales)

Order Date

Month of Order Date: September  
Region: East  
Sales: 287,360

Sales

January February March April May June July August September October November December

48 marks 1 row by 12 columns SUM(Sales): 14,915,601

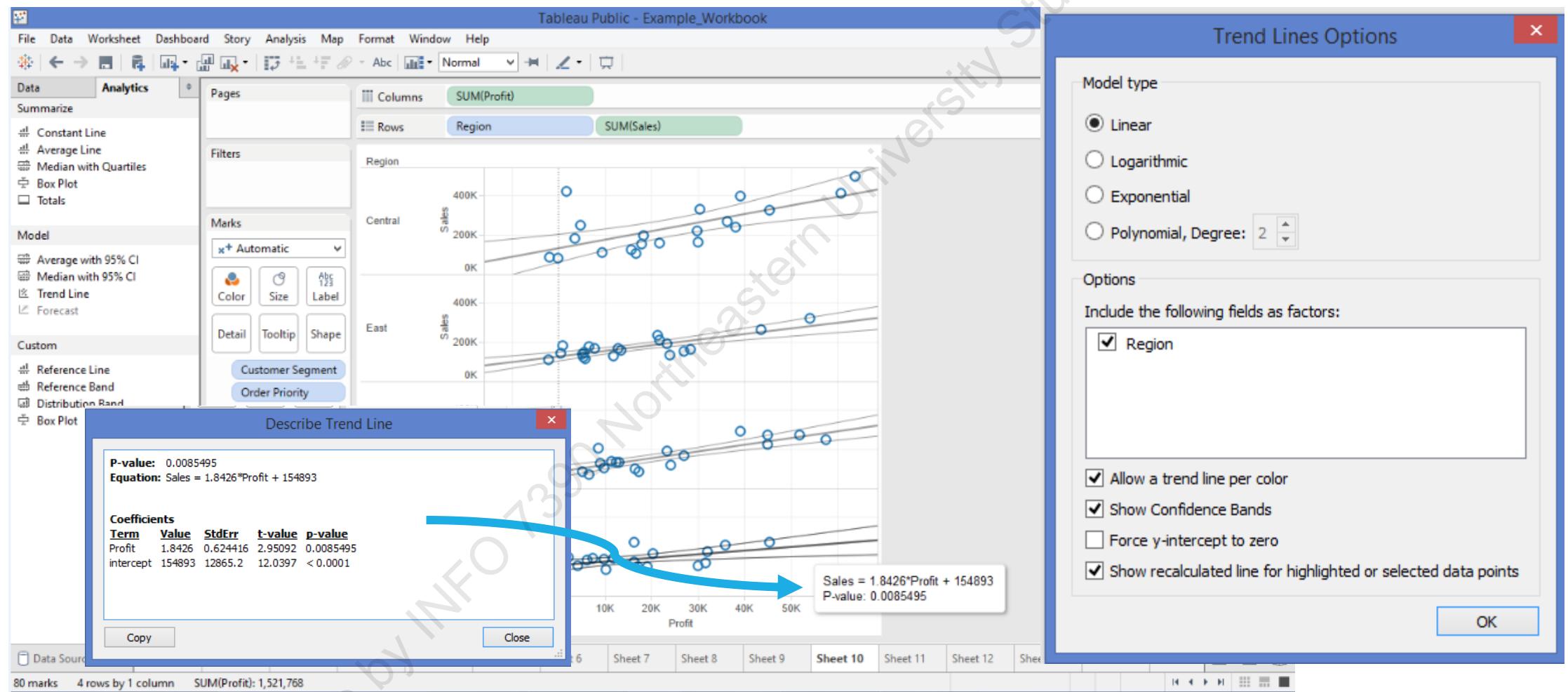
INFO 1300 Northeastern University

## Trend lines

- Trends can provide important insight to the analysis
- If you add more dimensions to view, Tableau allows you to have trend lines per pane or per scatterplot
- Go to “Trend Lines” under the “Analysis” menu to edit trend lines
- You can modify the model of trend line as linear, logarithmic, exponential, and polynomial
- You can also see the fitted equation of a trend line as well as model description in Tableau
- To do formatting you can also directly choose the “Format” menu from the toolbar



# Trend lines



See [Example.twbx—Sheet 10](#)



## Reference lines

- Reference lines do exactly what they seem to do. They add a point of reference to view
- Tableau allows you to modify all components of a reference line
- You can add a box plot and edit that for any desirable variable and edit the format of box, etc.
- Drop lines are also available in Tableau. They are very similar to the reference line, having reference for each axis
- Select “Reference Line” under the “Analysis” tab is an easy way to create reference lines



# Reference lines

Tableau Public - Example\_Workbook

File Data Worksheet Dashboard Story Analysis Map Format Window Help

Data Analytics

Summarize

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

Model

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

Custom

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

Pages

Columns Product Category Region

Rows SUM(Sales)

Marks

Automatic

Color Size Label

Detail Tooltip

Sales

1500K Average

1000K

500K

0K Central

Product Category

Region

West Central East South West

Technology

Edit Reference Line, Band, or Box

Line

Band

Distribution

Box Plot

Scope

Entire Table (radio button selected)

Per Pane

Per Cell

Line

Value: SUM(Sales) Average

Label: Computation

Line only 95

Formatting

Line:

Fill Above: None

Fill Below: None

Show recalculated line for highlighted or selected data points

OK

Data Source sheet 1 Sheet 1-1 Sheet 2 Sheet 3 Sheet 4 Sheet 5 Sheet 6 Sheet 7 Sheet 8 Sheet 9 Sheet 10 Sheet 11 Sheet 12 Sheet 13 Sheet 14 Sheet 15

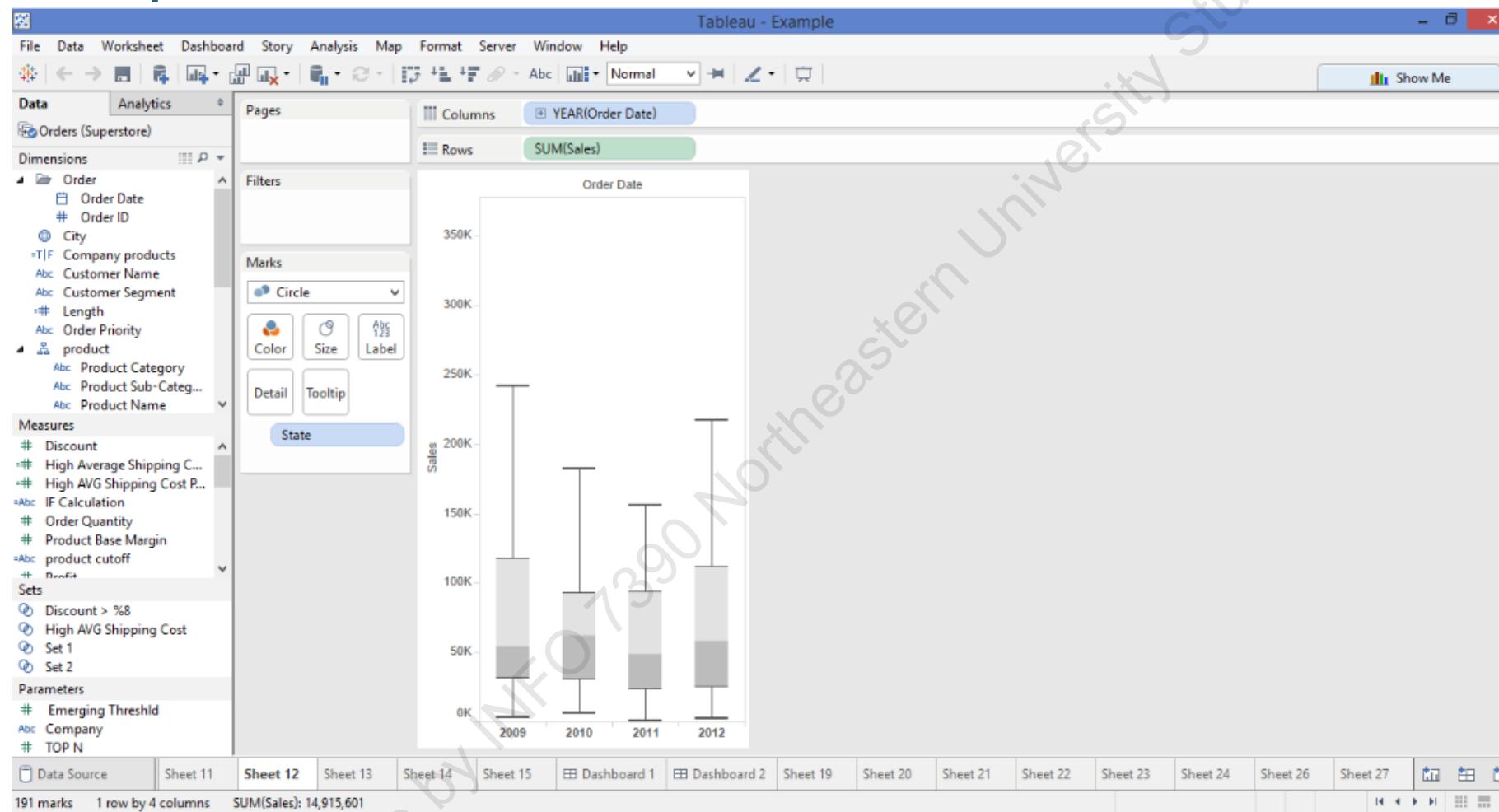
12 marks 1 row by 12 columns SUM(Sales): 14,915,601

For use by INFO310 students only

See [Example.twbx—Sheet 11](#)



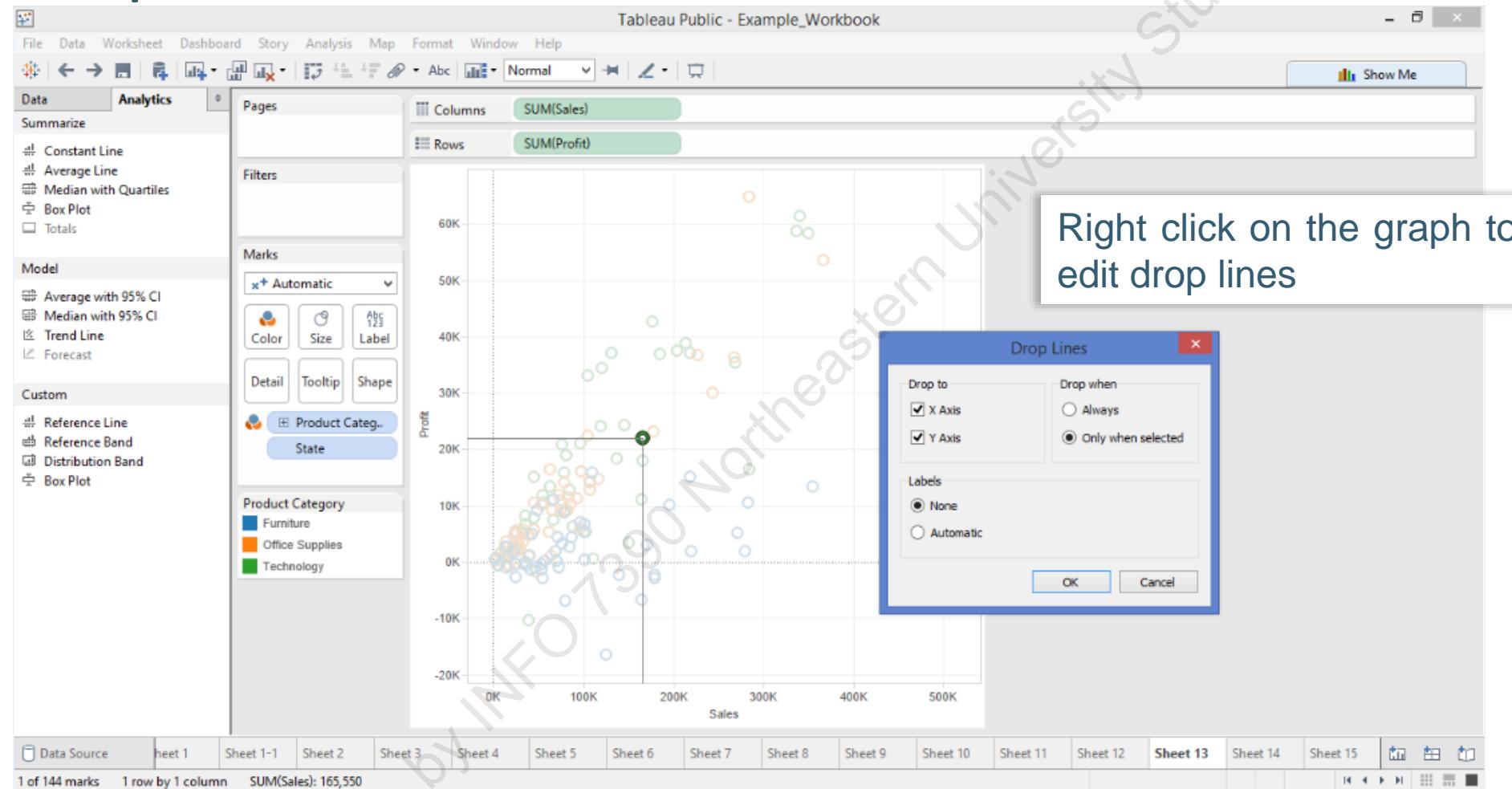
# Box plot



See [Example.twbx—Sheet 12](#)



# Drop lines



See [Example.twbx—Sheet 13](#)

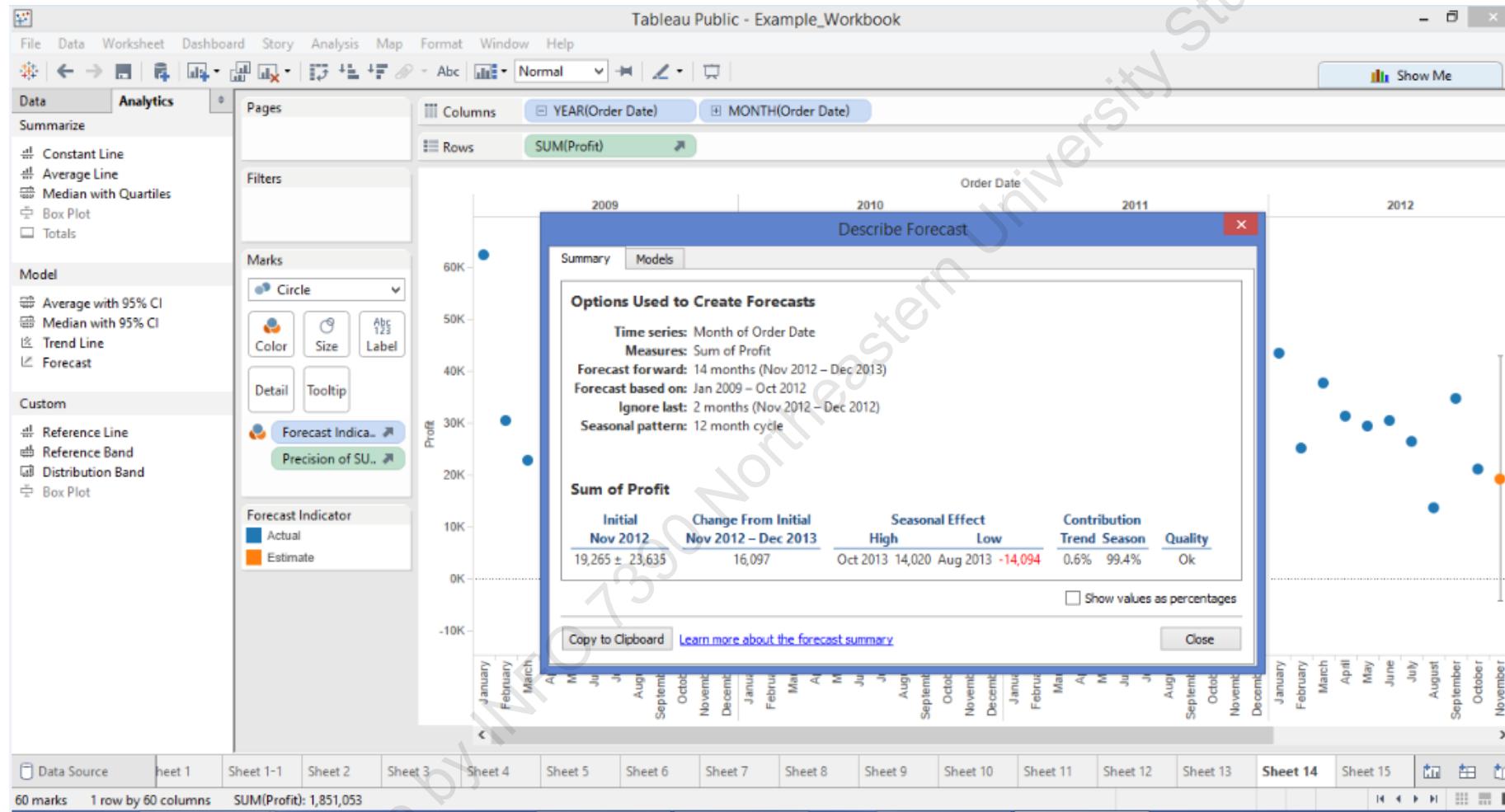


# Forecasting

- Forecasting means using statistical models for predicting future data based on historical data
- It is very easy to build a forecast in Tableau. All we need is a date field and the measure, and bring the forecast from the analytics tab
- Tableau uses exponential smoothing in its prediction
- Tableau allows you to tune some parameters of the model in the forecast option menu
- You can have the model description by choosing 'describe' from the forecast menu
- Forecasting in Tableau requires at least one date field, one measure, five data points, and two seasons of data if it is seasonal



# Forecasting



See [Example.twbx—Sheet 14](#)



# Visualization with Tableau

Dashboard and stories:

- ✓ Dashboard vs. stories
- ✓ Building a dashboard
- ✓ Dashboard interactivity with actions
- ✓ Story points



## Dashboard vs. stories

- Dashboards are great ways to present multiple views. They can include text, sheets, images, and webpages
- Stories in Tableau is a narrowed walk-through of one or more sheets of dashboards
- Dashboards tells you what, and stories tell you why
- Dashboards can update the sheet, while stories can't
- Any change in sheet or dashboard will be applied on story, while changes in story does not make change anywhere else
- Once you make a change on a specific field of story, other changes to this field on sheet and dashboard won't be applied to story

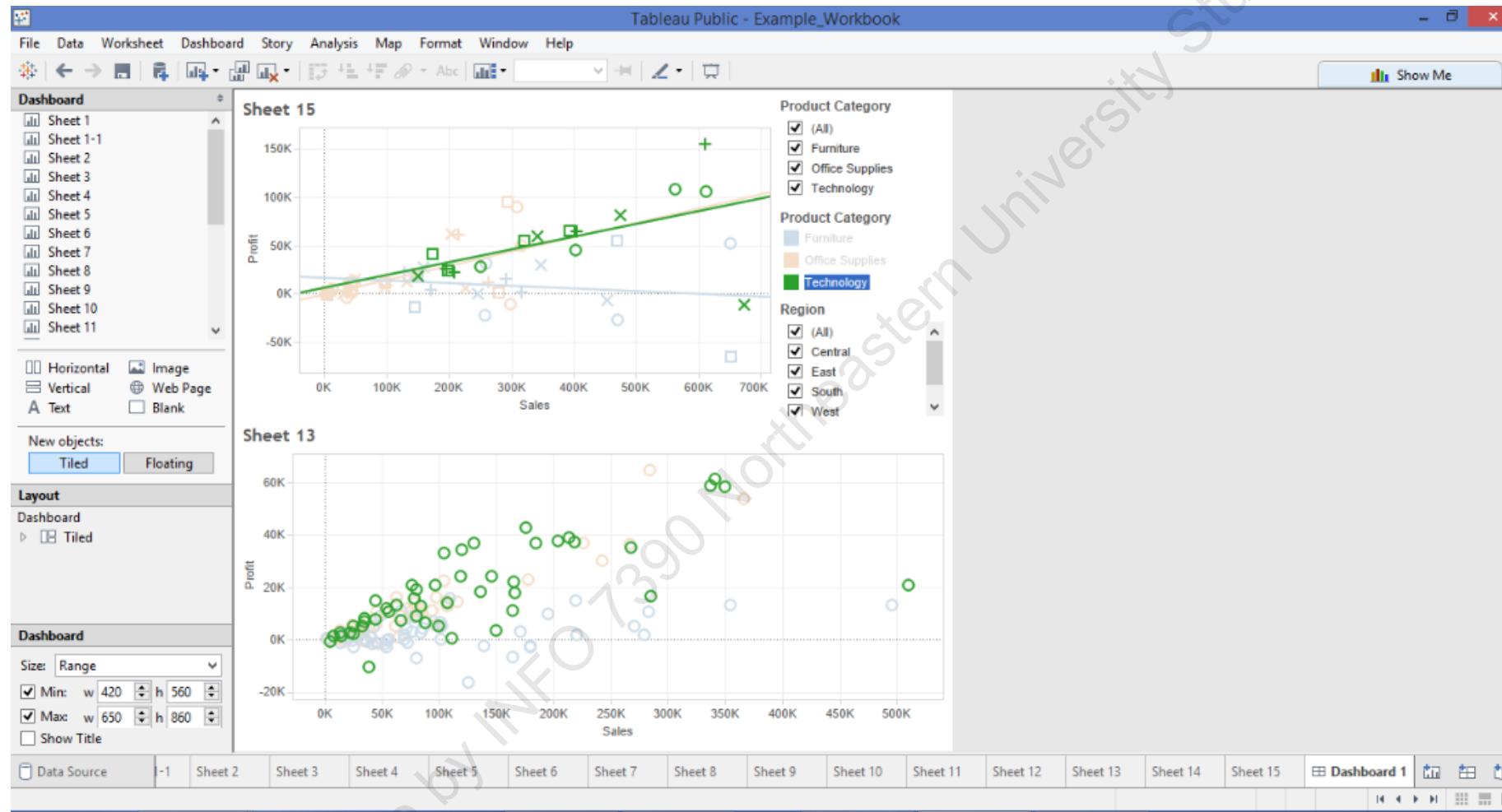


## Building a dashboard

- Dashboards can be created by clicking on the dashboard tab
- On dashboard, sheet data window is replaced by dashboard window
- You can modify layout and size of the dashboard by dashboard window
- When you add multiple views to a dashboard, the first view takes automatically the whole dashboard
- You can also apply filters on a specific view of a dashboard to all existing views of dashboards
- You may also change the layout of a dashboard by choosing ‘floating’ option rather than ‘tiled’



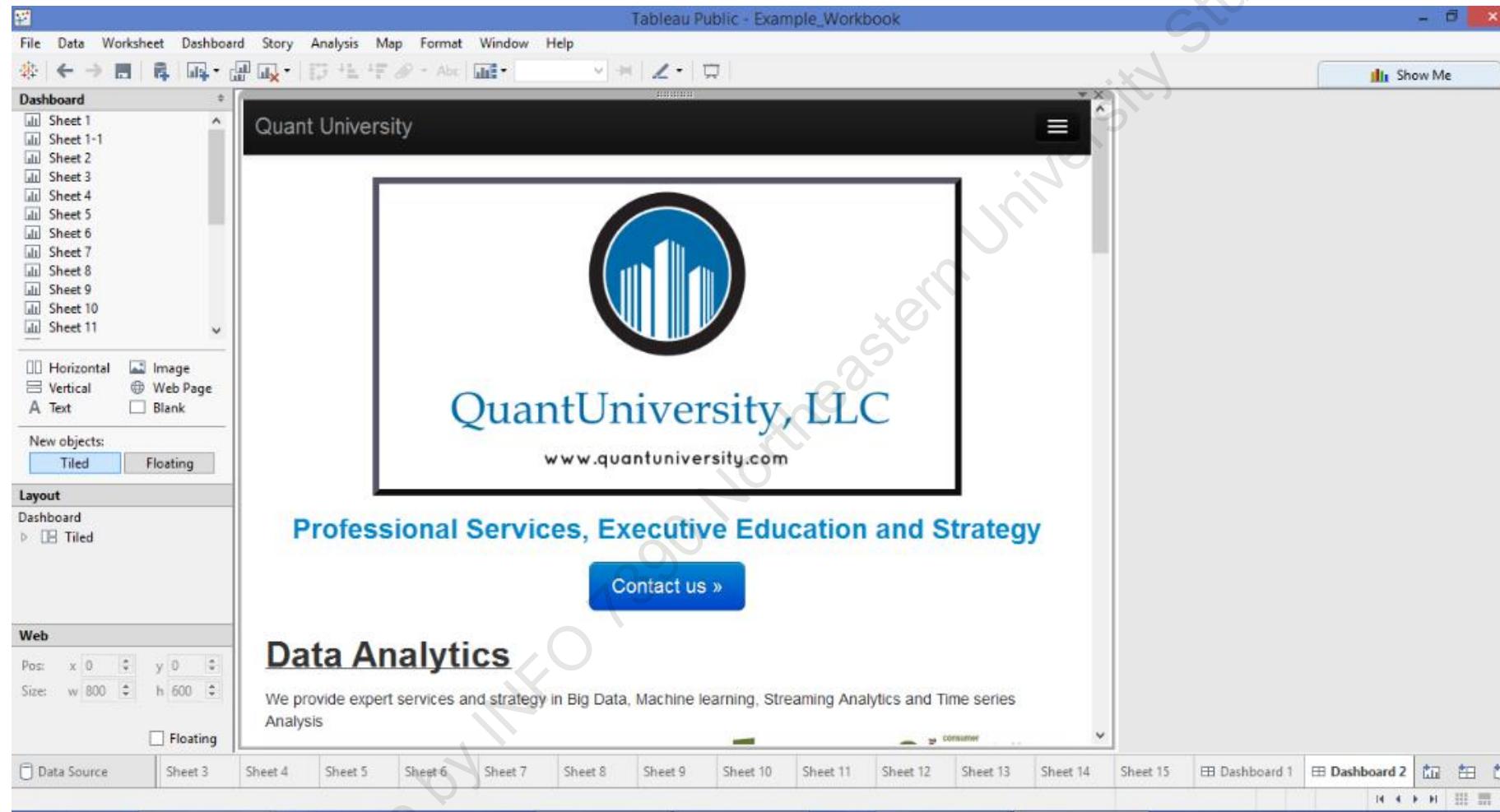
# Building a dashboard (from multiple sheets)



See [Example.twbx—Dashboard 1](#)



# Building a dashboard (from webpage)



See [Example.twbx—Dashboard 2](#)

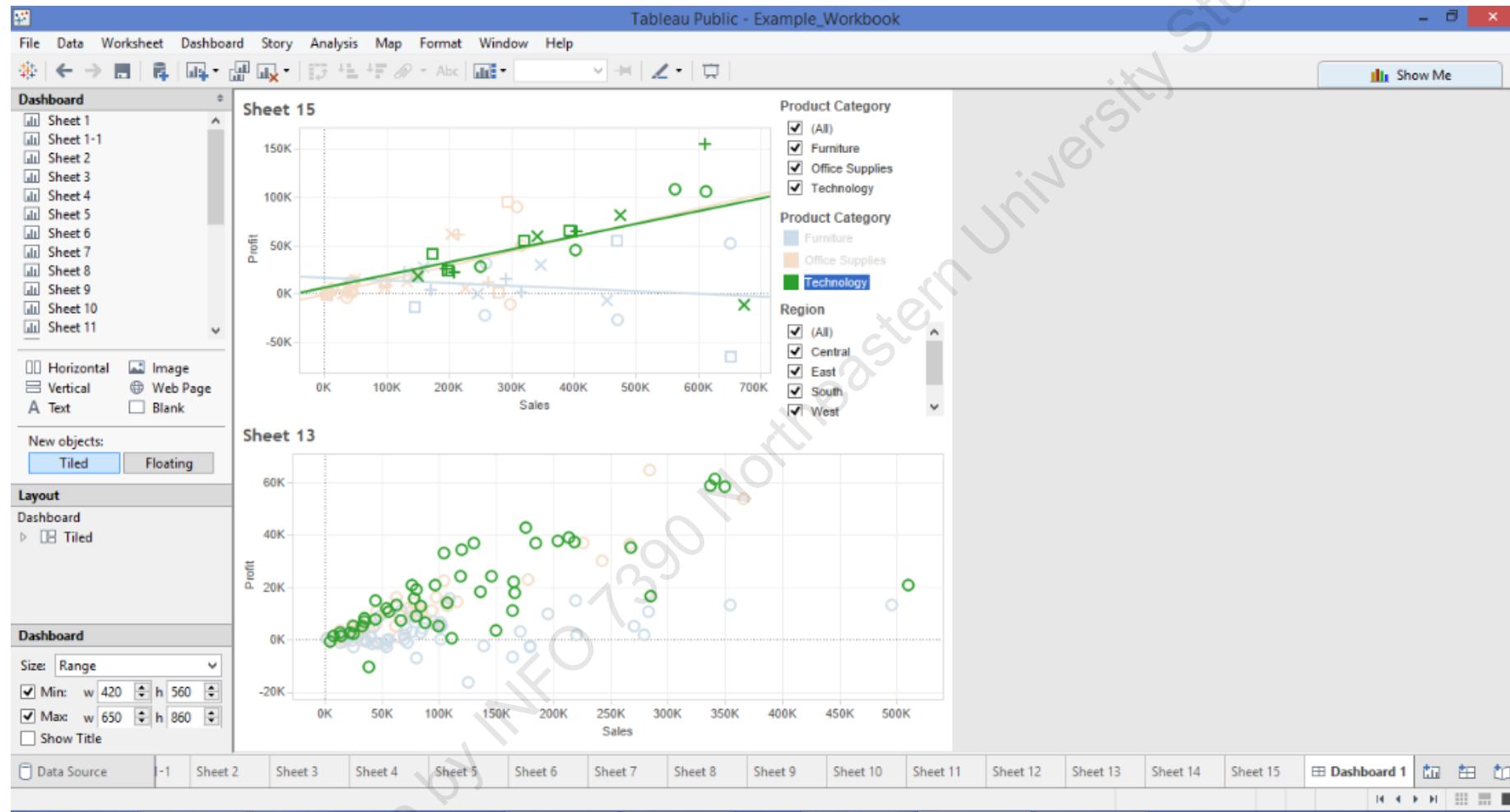


## Dashboard interactivity with actions

- Dashboard actions are powerful interactive elements within a dashboard derived from a worksheet
- They can be highlight action, filter action or URL action
- Actions on dashboard have several parts
- Highlight and filter actions have source and target
- Highlight and filter actions are used to control view of a dashboard
- URL actions can include field values as dynamic inputs
- Different actions can work together on a dashboard



# Dashboard interactivity with actions (highlight action)



## Story points

- Story points lets you create interactive data driven stories
- Stories consist of specific dashboards in sequential presentation
- Story points let you tell different scenarios to an audience while analyzing the data
- Again, dashboards can tell you what, and stories can tell you why
- The major benefit of story points is to preserve the specific state of a given visualization, like a filter or highlighting
- If you want to bring more than one visualization to a story point, you should bring it as a dashboard
- You can modify various elements of a story by going to 'story' menu and choosing 'format'



# Visualization with Tableau

## Calculations

- ✓ Calculation
- ✓ LOD (level of detail) expression
- ✓ Table calculation
- ✓ Aggregate calculation
- ✓ Date calculation

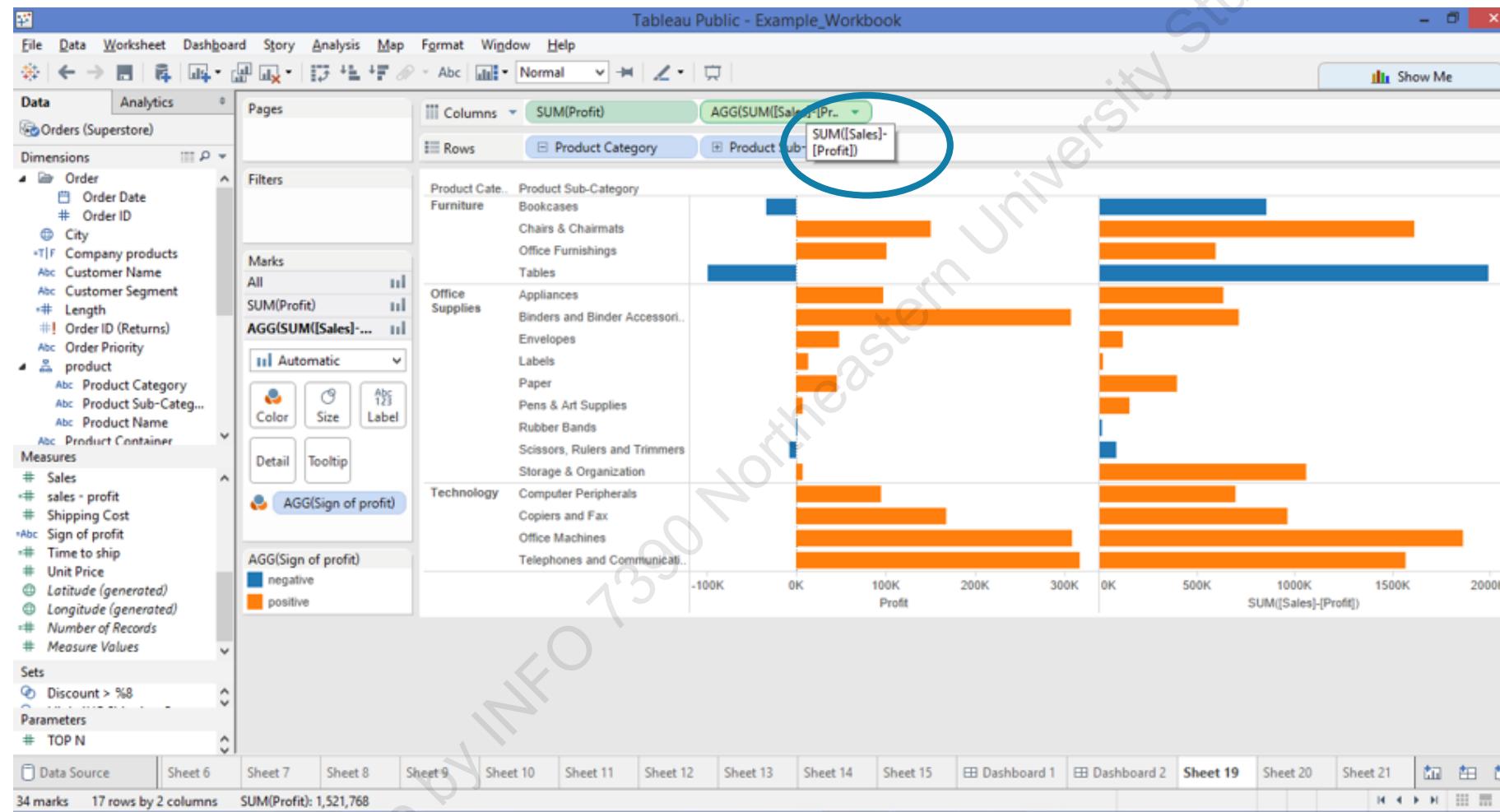


## Calculations

- Similarly to formulas in excel, calculations allow you to manipulate data in a number of ways
- Calculated field is created by defining a formula
- Calculation in Tableau can be applied to many different types of data, such as date, numbers, string, etc.
- Regular calculation is passed as part of a query that Tableau ask as a data source and computation is handled by a data source itself, with only the result set being returned to Tableau
- A table calculation is a secondary calculation that is performed on the top of the results set
- Tableau allows you to write syntax for different calculations



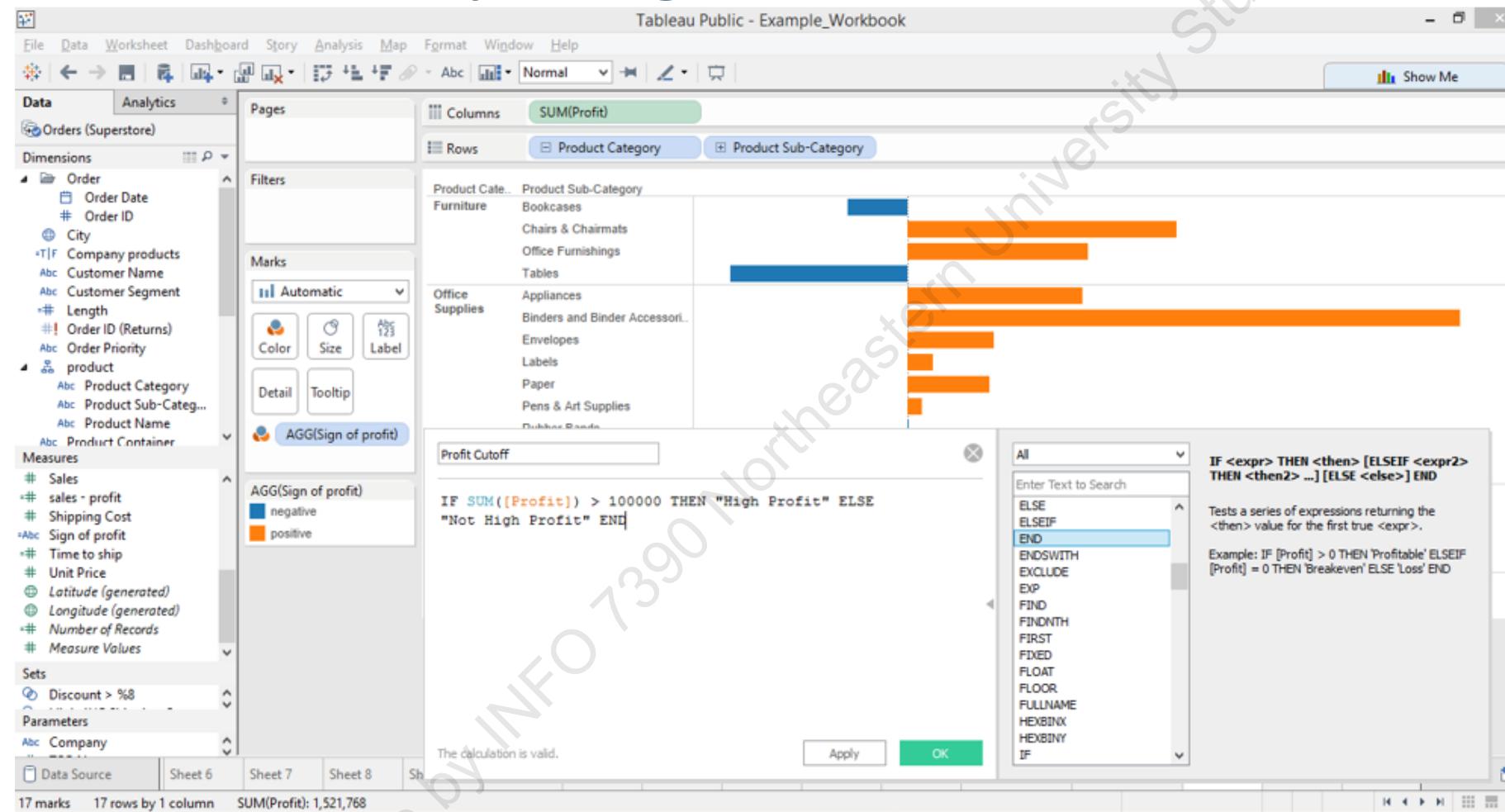
# Calculations



See [Example.twbx—Sheet 19](#)



# Calculations by using conditions



See [Example.twbx—Sheet 19](#)



## LOD (level of detail) expression

- Allows us to do some level of compute aggregation that is not at the level detail in view. Lets say we want to have the average of profit for all orders in a city. First we should sum at the level of order ID as below, then average

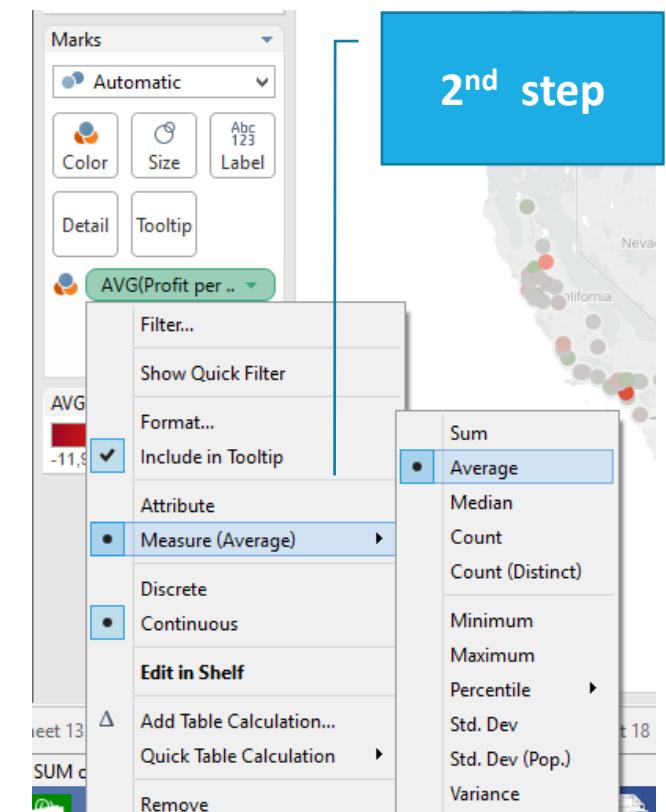
Profit per Order ID

```
{INCLUDE [Order ID] : SUM([Profit])}
```

The calculation is valid.

1<sup>st</sup> step

Apply      OK

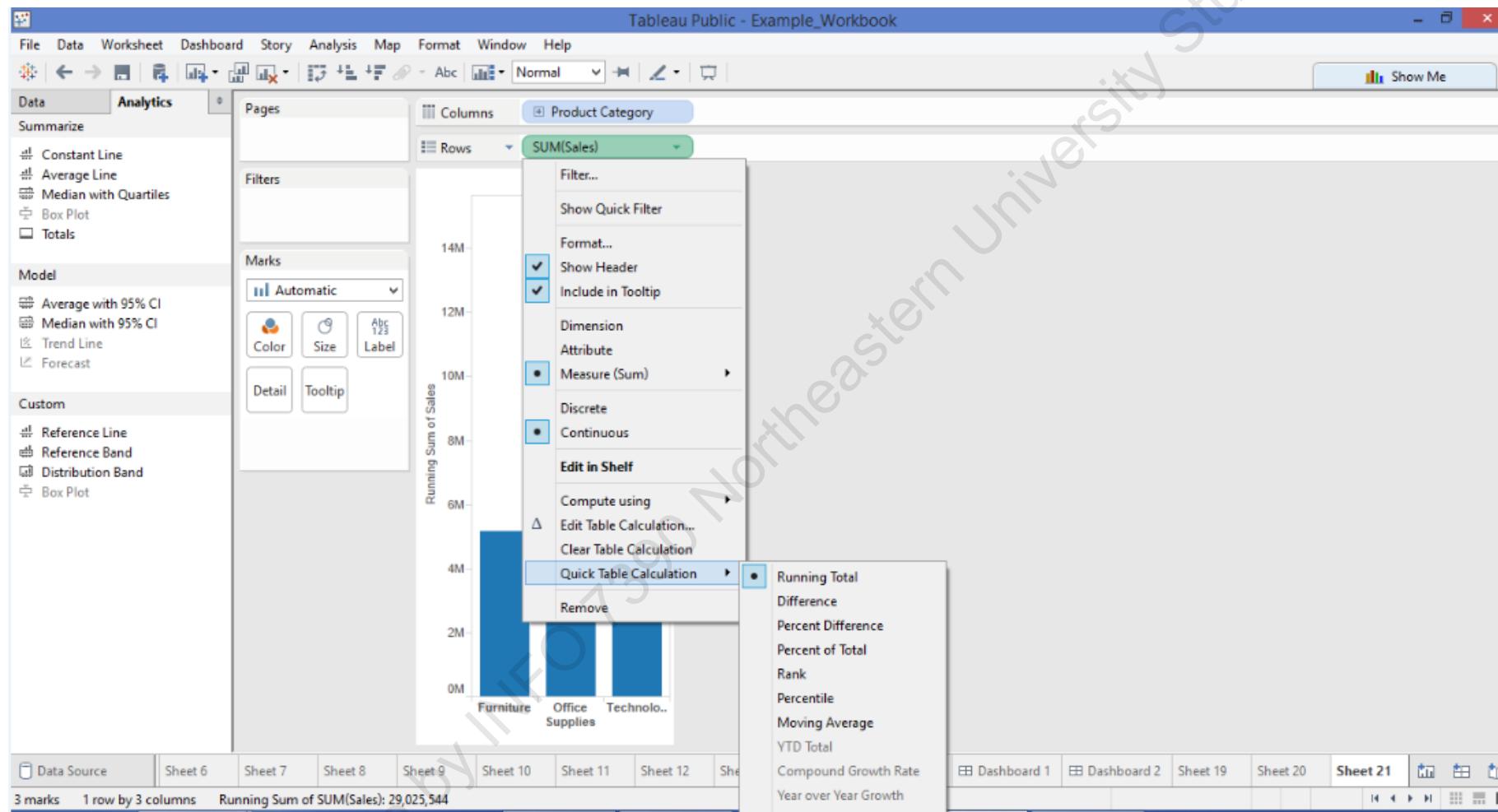


## Table calculation

- Table calculation allows us to ask questions, such as which product category has the most percentage of sales, or how much is the running sum of sales
- Non-table calculations are evaluated at the data source
- Table calculations are only performed on the data in the view
- Table calculations can be made by “Quick Table Calculation”
- In modifying a table, table specifies the scope and across specifies the direction
- Table calculation slices the data based on partitioning and addressing categories



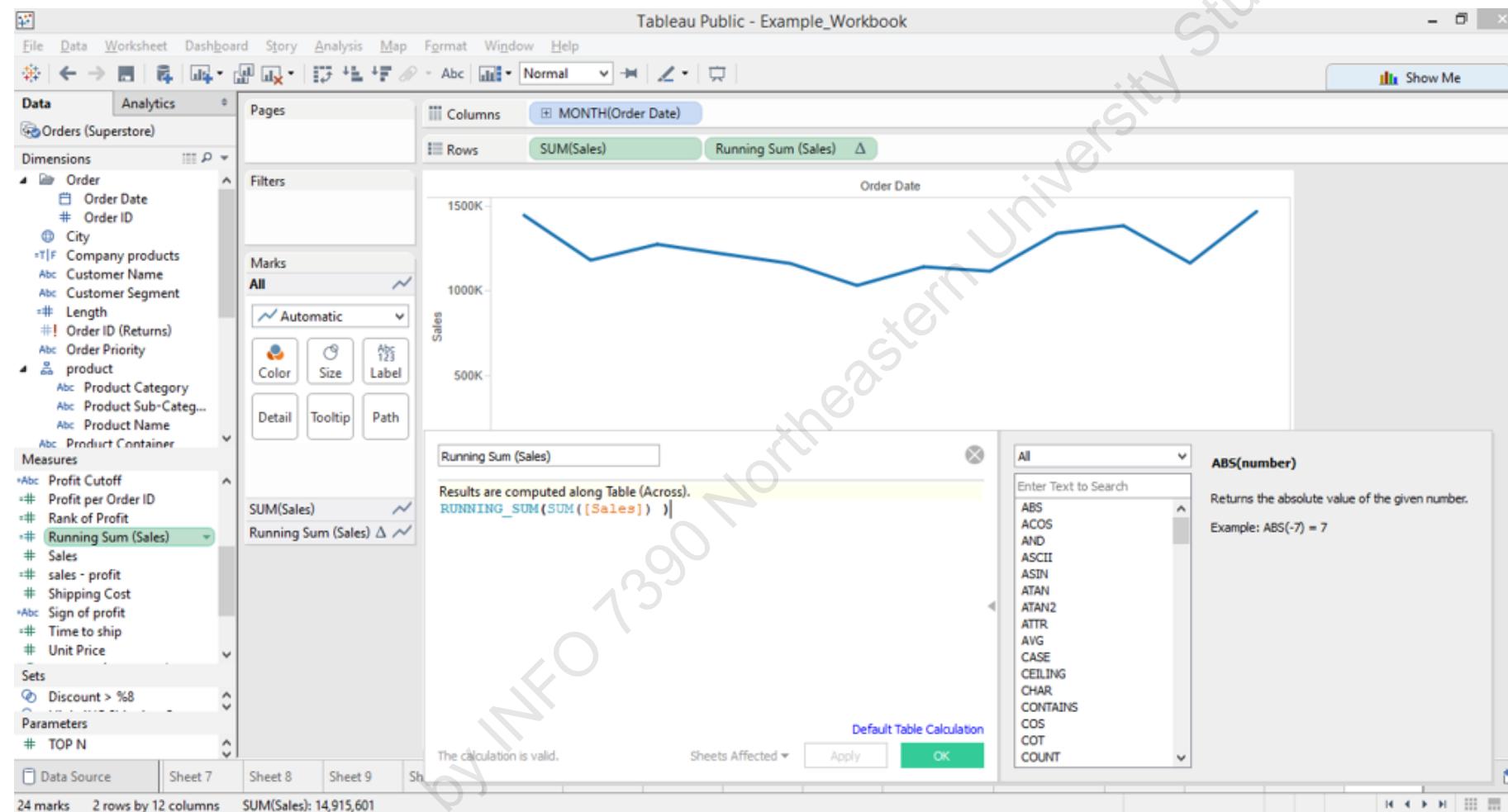
# Table calculation



See [Example.twbx—Sheet 21](#)



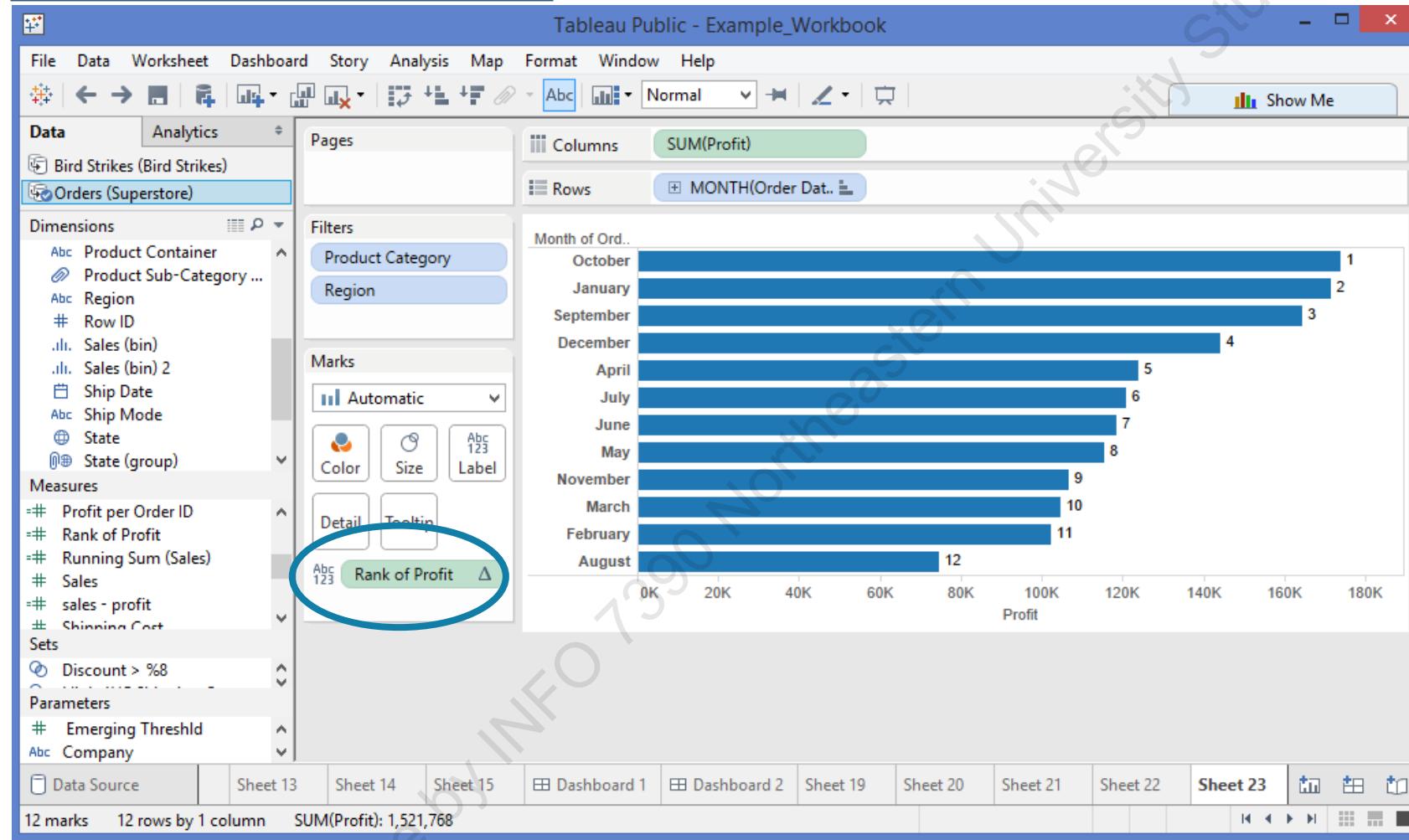
# Table calculation



See [Example.twbx—Sheet 22](#)



# Table calculation



See [Example.twbx—Sheet 23](#)



## Aggregate calculation

- AGG stands for aggregation and tells us aggregation is built into the calculation
- ATTR is another common reference seen in aggregate calculation
- If we are doing aggregation on different classes of a same dimension, we should use ATTR as attribute aggregation
- We may use aggregation at record level, such as average of sales at order level instead of product level



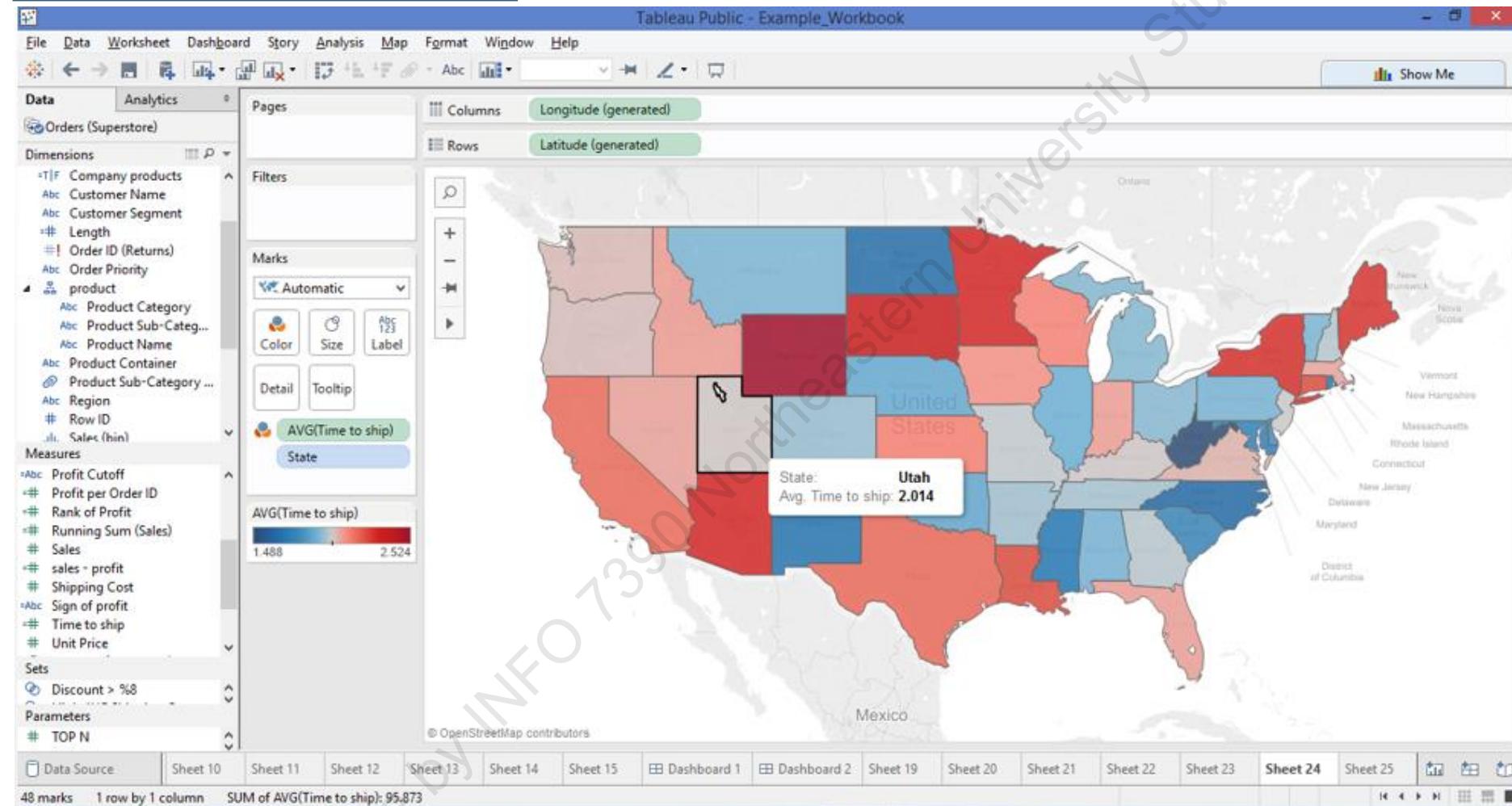
## Date calculation

- A common question could be something like, how much time did something take? Date calculation can easily answer that
- The function for doing this calculation is DATEDIF
- DATEDIF needs the unit of calculation like day, start date, and end date. Also MAKEDATE allows you to work with year, month and day

The screenshot shows a software interface for data analysis or reporting. On the left, a formula bar displays the text "Time to ship" above the formula `DATEDIFF('day', [Order Date], [Ship Date])`. Below the formula, a message says "The calculation is valid." At the bottom right of the formula bar are two buttons: "Apply" and "OK". On the right side of the interface is a help panel for the `DATEDIFF` function. The panel has a search bar at the top with the placeholder "Enter Text to Search". Below the search bar is a list of functions, with `DATEDIFF` highlighted in blue. To the right of the list is a detailed description of the `DATEDIFF` function, including its syntax: `DATEDIFF(date_part, start_date, end_date, [start_of_week])`. The description explains that it returns the difference between two dates where `start_date` is subtracted from `end_date`, expressed in units of `date_part`. If `start_of_week` is omitted, the week start day is determined by the start day configured for the data source. An example is provided: `DATEDIFF(month, #2004-07-15#, #2004-04-03#, 'sunday') = -3`.



# Date calculation



See [Example.twbx—Sheet 24](#)



# Visualization with Tableau

## Calculations

- ✓ Logic calculation
- ✓ Number calculation
- ✓ String calculation
- ✓ Type calculation
- ✓ How to integrate R and Tableau



## Logic calculation

- Logic calculation determines if a certain condition is TRUE or FALSE
- Also, logic calculation is known as Boolean logic
- A conditional filter is a filter that keep or exclude members based on the logical calculations
- IF and ELSE statements is the commonly used syntax for performing logic calculation



# Logic calculation

The screenshot shows a data visualization interface with two main components: a filter dialog on the left and a chart on the right.

**Filter [State] Dialog:**

- General Tab:** Selected.
- Condition Tab:** Selected.
- Top Tab:** Unselected.
- None:** Unselected.
- By field:** Selected.
  - City: Count
  - = 0
- Range of Values:**
  - Min: (empty)
  - Load button
  - Max: (empty)
- By formula:** Selected.  
Formula: `SUM([Sales]) > 50000 AND COUNTD([Order ID]) > 500`

**Chart View:**

- Columns:** CNTD(Order ID), SUM(Sales)
- Rows:** State
- Data:** California
- Summary:** Distinct count of Order ID: 543, Sales: 1,372,210

A large blue arrow points from the formula input field in the filter dialog to the "Distinct count of Order ID" value in the chart summary.



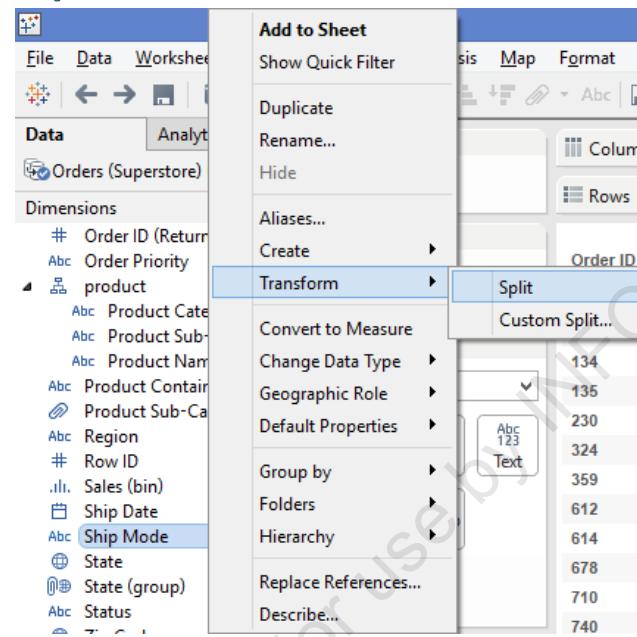
## Number calculation

- Sometimes there are some missing values, such as NULL, in many datasets
- To deal with these records, use the ZN function
- ZN returns zero when it faces NULL
- We can think of this function as a conditional statement;
  - Are there any values for the specific field at this record? If the answer is no, it assumes zero
- There are several other functions available that can be very useful in complex calculations from scratch

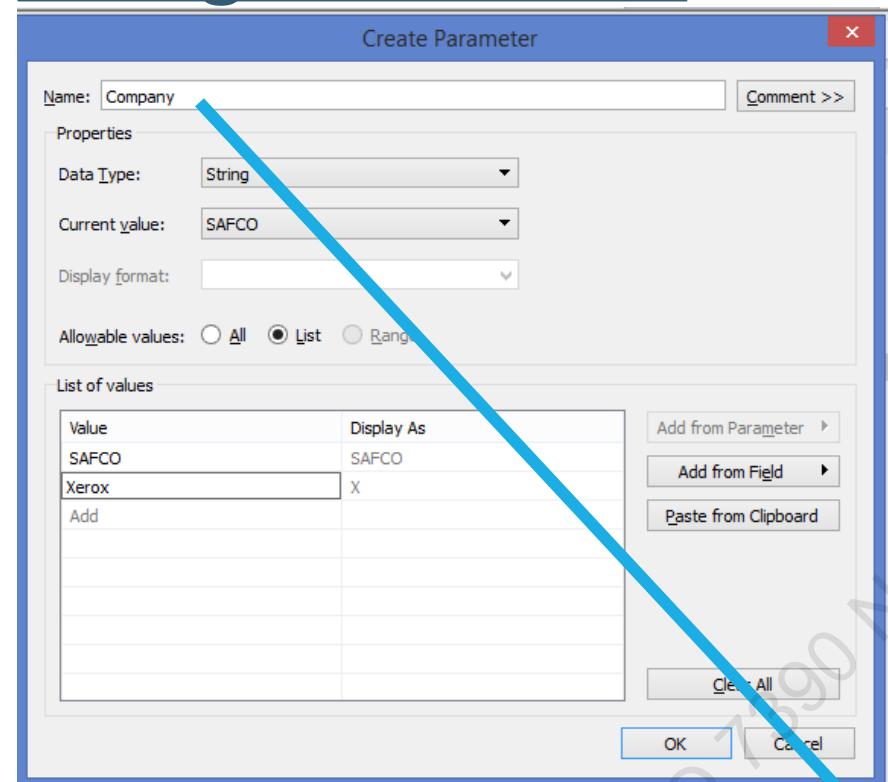


## String calculation

- String calculation allows you to work with string data types such as name in our calculation
- In some cases that consist of several parts, such as IDs and codes, the string function allows you to split them by right-clicking on the specific field choosing ‘transform’ and then ‘split’



# String calculation



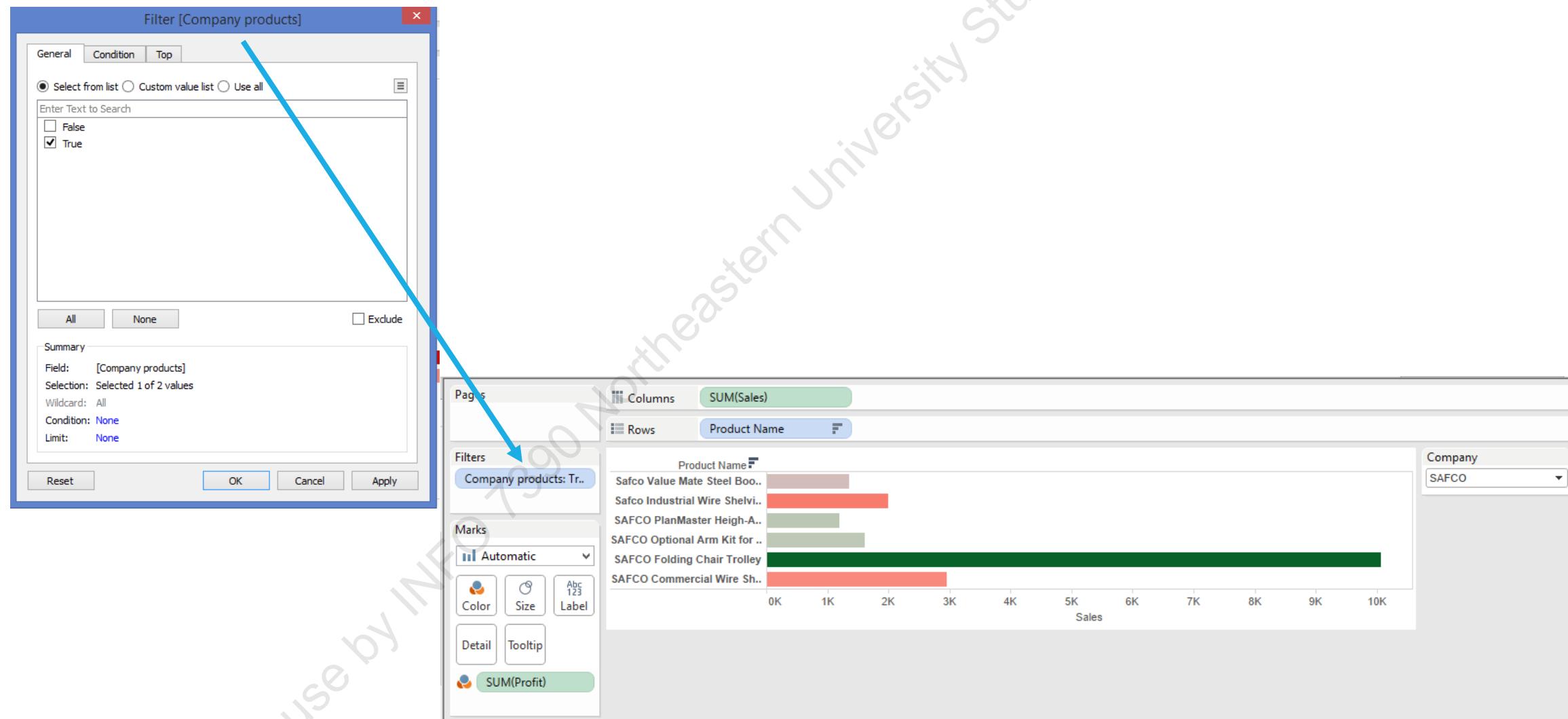
Company products

CONTAINS([Product Name], [Company])

The screenshot shows a search interface. A parameter named 'Company' is defined with a data type of 'String', current value 'SAFCO', and the string 'CONTAINS' is highlighted. The interface also displays 'Data type: String' and 'Current value: SAFCO'.

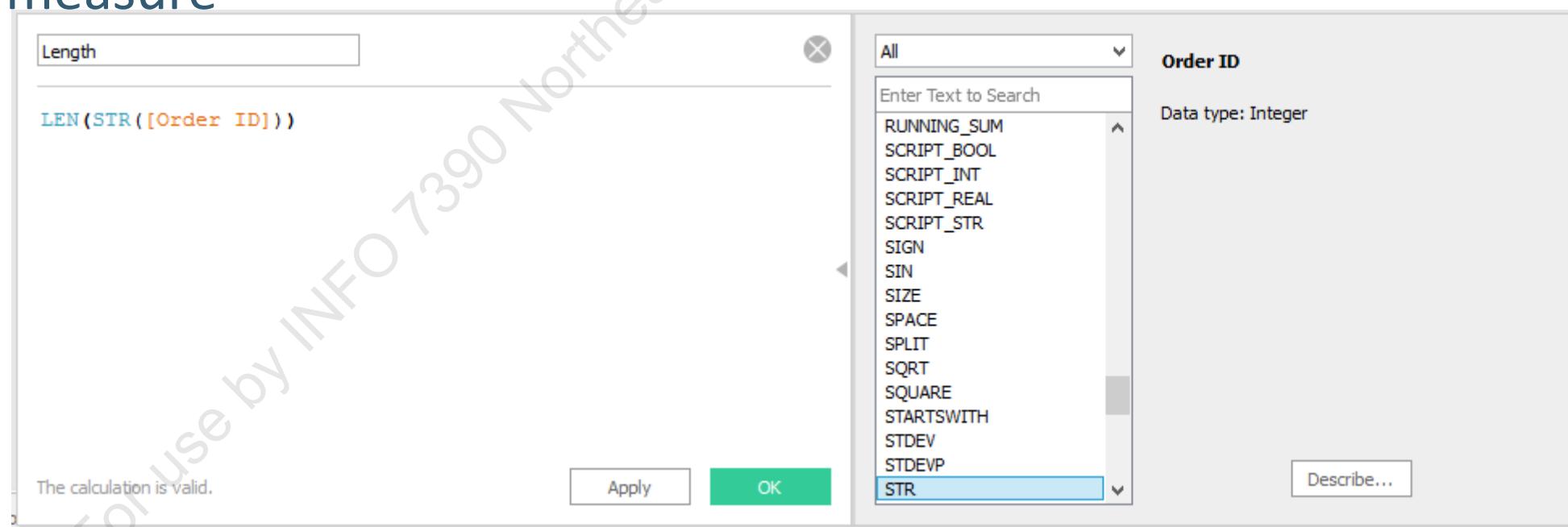


# String calculation



## Type calculation

- Sometimes you may need to converge the type of field
- The LEN function for example only works with string, so you may need to change the type of number to string. In such situations we should use LEN and move the calculated field to dimension rather than measure



## How to integrate R and Tableau

- Usually Tableau passes the data to R, where the computation takes place, then the results are returned to Tableau to get visualized like any other calculated field
- Using R with Tableau requires few steps:
  - First of all you should install R on your machine
  - Install the “Rserve” package , and call the library “Rserve” in R to make a connection between R and Tableau
  - The common workbook should be shared with R and Tableau, by choosing “Manage R connection” from setting of Tableau help
  - Any specific package that needs to be used in Tableau should be installed in R first



# Summary

We have covered	Feature engineering methods
Introduction	✓ What is visualization and Tableau software
How to connect data	✓ Connecting to data, data preparation, edit data connection/metadata
Visual analytics	✓ Working with marks and building views, hierarchies, sorting and grouping, groups and sets, how to create sets and work with them, filtering and ways to filter, parameters, formatting, fitting trend and reference lines and time series forecasting
Dashboards and stories	✓ Dashboard vs. stories, how to build a dashboard, how to use actions in dashboards and story points
Calculations	✓ How to create a calculation, level of detail in calculations, table calculation, aggregate calculation, different types of calculations such as date, logic, number, string, and type calculations and how to integrate R and Tableau



# Reference

- Documentations:
  - Jiawei Han, Micheline Kamber, Data Mining, Concepts and Techniques
  - S. Christian Albright, Wayne L. Winston, Business Analytics
- Videos:
  - Official Tableau Training & Tutorials



## Q&A



For use by INFO 7390 Northeastern University Students only





QuantUniversity, LLC

[www.quantuniversity.com](http://www.quantuniversity.com)

# Thank you!

## Contact

**Sri Krishnamurthy, CFA, CAP  
Founder and CEO  
QuantUniversity LLC.**

LinkedIn [srikrishnamurthy](#)

[www.QuantUniversity.com](http://www.QuantUniversity.com)