DATA VISUALIZATION

FUNDAMENTALS

IS – Advanced Analytics
Data Visualization Fundamentals 4

IS – Advanced Analytics
Data Visualization Fundamentals 5

DATA VISUALIZATION PROJECT CHECKLIST	DEVELOPER	PUBLISHER
Research Data		
Define Fields		
Defines Filters		
Define Calculations		
Define KPI's		
6Ws.		
Connect to Data		
Clean and filter data before connecting		
Filter unnecessary data		
Rename Fields		
Assign correct data properties to fields		
Join new table		
repeat steps for each newly joined table		
Prepare Data Pane		
Creat Folders		
Create Hierarchies		
Create Groups		
Data Discovery(Create additional groups as needed)		
Crosstab Totals		
Crosstab Analysis		
Bar chart		
Side by Side Bar Chart		
Stacked Bar chart		
Pie chart		
Geo Map		
Line chart		
Scatter Plot		
Highlight Table		
Heat Map		
Tree Map		
Bubble Word Cloud		
Motion Chart		
Create Story		
Review 6 Ws		
Choose chart or charts that best answer 1 of the 6w's		
Add Filters and Interactivity		
Add Headline		
Organize Dashboard to visually be easy to understand		
Publish Dashboard to your site		

Data Visualization Fundamentals 8

TABLEAU PROJECT FLOW OUTLINE

1. RESEARCH
2. CONNECT
3. PREPARE DATA WINDOW
   a. MARKS CARD
4. DATA DISCOVERY
5. CREATE STORY/ANSWER QUESTIONS
6. PUBLISH

PROJECT FLOW OUTLINE IN DETAIL

RESEARCH DATA

- Data Source
- Understanding your data
- Relationships between fields

CONNECT TO DATA

- Connect to Data
- Filter
- Clean, Define, Rename
- Join
- Repeat
- Hide fields
- Live verses Extract

PREPARE DATA PANE

- Organize Data
    o Create Folders
    o Create Hierarchies
- Create Calculated Fields
- Create relationships with data
    o Create Groups

MARKS CARD

- Marks card
- Marks type
- Marks properties
- Using Color

DATA DISCOVERY

- Filters/Marks Card/Worksheet/Quick Filter/Menus
- Create a crosstab & crosstab with basic analysis
- Create (basic charts)
- Create a dashboard

CREATE A STORY/ANSWER QUESTIONS

- Create final charts
- Create final dashboard

PUBLISH

RESEARCH DATA

- Data Source
- Understanding your data
- Relationships between fields

DATA SOURCE

Data Visualizations add immense amounts of credibility. Decisions will be made off of your data. Getting to know your data source and the data you are using is where you should spend most of your time in a data visualization project.

With data coming from all directions you want to be sure to have the correct data source. Your accounting department may define the divisions of your company different from your buying department. How does that affect you? Well let's say you find a fantastic data source with sales information and begin to use it. Everything is going well until the new quarter rolls around. All of the sudden your numbers are off and everyone is upset. What happened? Your report was great just last quarter. Well, you may have used the wrong data source.

If you are looking at tables in a database and find one that appears to have all the data you want, you have to find out where the table is coming from for that table or what business area it is supporting. You can find this out by contacting your IS department and asking for documentation or sometimes it's in the name of the table.

The data source we will be using is comprised of; Store locations, Departments, Categories for Departments, Items, Sales in units and Sales dollars. A quick look at the dates will show this is for the first quarter of the fiscal year.

As you review a new project one of the first steps you will want to do is make a glossary for the, fields, filters and calculations.

Any field you see in a report or any calculation that is made you will want to define and log down as a reference. You really want to understand the report and data fluently. Most reports you find will be in some form of spreadsheet and have terms you may not know. Look at the example below.

US INCLUDING BD & RC (RTV CENTER)												
	TY	LY		TY	LY	TY	LY	TY	LY		TY	LY	TY
LOCATION	SALES$	SALES$	% CHG	WHSE D&D$	WHSE D&D$	RC D&D$	RC D&D$	TOTAL D&D$	TOTAL D&D$	% CHG	TOTAL D&D%	TOTAL D&D%	SPOILS$
NW 1 DISTRICT	257,256	247,894	3.78%	1,342	1,159	690	805	2,032	1,964	3.49%	0.79%	0.79%	(1,424)
NW 2 DISTRICT	256,986	237,418	8.24%	1,915	1,828	206	299	2,121	2,127	-0.26%	0.83%	0.90%	(1,556)
NW 3 DISTRICT	235,816	222,632	5.92%	1,580	1,287	540	581	2,120	1,869	13.48%	0.90%	0.84%	(1,315)
NW REGION	750,057	707,944	5.95%	4,837	4,274	1,437	1,685	6,274	5,959	5.28%	0.84%	0.84%	(4,294)

What the heck is RC D&D$ and all those % of CHG? What is going on and what value is this report offering? Where do you even begin?

GLOSSARY(FIELDS, FILTERS, CALCULATIONS)

FIELDS

Defining and understanding what fields mean is the best approach to understanding new report. Spell out any abbreviations and take as much time as you need to understand and define terms that related to how business is done.

<u>Fields</u>

TY = This year Sales

L Y = Last year Sales

RC = Return Center – Depots have a Return Center. The cost of returns are redistributed over all locations.

FILTER

Next look at any filters that are being used. This can show itself in many ways. The easiest to spot is filters that the user has created but what is not as obvious is the worksheets that are created in a workbook. A lot of times this is a form of filtering. You will want to make this part of your research.

Here is an example of tabs at the bottom of a worksheet and logging the filters.

<u>F</u>
<u>ilters</u>

Dollar amount (Net, Gross)

Regions (NW, BA, MW…)

CALCULATIONS

Now that you have a glossary of fields and filters, you need to find out what kind of math is being performed.

Sadly you may find that some of the math is incorrect or has a broken formula. You will want to work closely with your subject matter expert on this part of the project. Start by clicking into each cell and logging the formulas that you see.

Looking at something as simple as sales you will find that there is actually some rounding taking place and it's referencing another spreadsheet. What is on that other spreadsheet? It could be a data dump or it could be doing some math to come up with the sales number. You will want this information.

| C7 | | | fx | =SUMIF('Sales Query'!$A$1:$A$3000,$A7,'Sales Query'!$F$1:$F$3000)/1000 |

	B	C	D	E	F	G	H	I	J	
1									NET DAMAGED	
2										
3										
4	US INCLUDING BD & RC (RTV CENTER)									
5		TY	LY		TY	LY	TY	LY	TY	
					WHSE	WHSE			TOTAL	TO
6	LOCATION	SALES$	SALES$	% CHG	D&D$	D&D$	RC D&D$	RC D&D$	D&D$	D.
7	NW 1 DISTRICT	257,256	247,894	3.78%	1,342	1,159	690	805	2,032	1
8	NW 2 DISTRICT	256,986	237,418	8.24%	1,915	1,828	206	299	2,121	2
9	NW 3 DISTRICT	235,816	222,632	5.92%	1,580	1,287	540	581	2,120	1
10	NW REGION	750,057	707,944	5.95%	4,837	4,274	1,437	1,685	6,274	5

Look at the NW Region Row. That looks like a summarization.

| C10 | | | fx | =SUMIF('Sales Query'!$C$1:$C$3000,LEFT($B10,2),'Sales Query'!$F$1:$F$3000)/1000-SUMIF('Sales Query'!$E$1:$E$3000,1998,'Sales Query'!$F$1:$F$3000)/1000 |

	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1									NET DAMAGED AND DESTROYED - FY 2015 PER 11 PTD									
2											($ are in 000's)							
3																		
4	US INCLUDING BD & RC (RTV CENTER)																	
5		TY	LY		TY	LY	TY	LY	TY	LY		TY	LY	TY	LY		TY	LY
					WHSE	WHSE			TOTAL	TOTAL		TOTAL	TOTAL				SPOILS	SPO
6	LOCATION	SALES$	SALES$	% CHG	D&D$	D&D$	RC D&D$	RC D&D$	D&D$	D&D$	% CHG	D&D%	D&D%	SPOILS$	SPOILS$	% CHG	%	%
7	NW 1 DISTRICT	257,256	247,894	3.78%	1,342	1,159	690	805	2,032	1,964	3.49%	0.79%	0.79%	(1,424)	(1,383)	2.94%	0.55%	0.56
8	NW 2 DISTRICT	256,986	237,418	8.24%	1,915	1,828	206	299	2,121	2,127	-0.26%	0.83%	0.90%	(1,556)	(1,438)	8.21%	0.61%	0.61
9	NW 3 DISTRICT	235,816	222,632	5.92%	1,580	1,287	540	581	2,120	1,869	13.48%	0.90%	0.84%	(1,315)	(1,249)	5.25%	0.56%	0.56
10	NW REGION	750,057	707,944	5.95%	4,837	4,274	1,437	1,685	6,274	5,959	5.28%	0.84%	0.84%	(4,294)	(4,070)	5.51%	0.57%	0.57
11	RETURN CENTER	905	1,133	-20.13%	1,277	1,728	(1,437)	(1,685)	(160)	43	-471.71%	-17.67%	3.80%	0	0	0.00%	0.00%	0.00
12	NW INC RC	750,962	709,077	5.91%	6,114	6,002	0	0	6,114	6,002	1.86%	0.81%	0.85%	(4,294)	(4,070)	5.51%	0.57%	0.57

Sure enough it is! It looks like there is a lot of rounding going on with this row. It's important to note that Calculations can change as you go across or down a report.

It may seem overwhelming but having this Glossary of Fields, Filters and Calculations is the most important piece of any data analysis. Once you have completed this process meet with your business sponsor or subject matter expert to confirm that you understand the data you have been given.

Calculations

TY - Rounded to 1000

LY - Rounded to 1000

IS – Advanced Analytics
Data Visualization Fundamentals 14

% CHG - (TY minus LY) divided by LY

WHAT'S THE POINT?

Beyond the spreadsheet the important question you want to ask is:

"How does this report help you meet your goals?"

If you ask what is important they will say the entire report is. But if you ask them "How does this help you meet your goals?" You will find exactly what they are looking for. This will help you later when building your analysis and story.

Partial Example of a Glossary

Gross Margin Detail report or GMD report is used to make sure each location and department are within the correct range of GMD, not to high and not to low. The report is a description of what has happened and used to make managers aware of any problems or success they have.

Fields

     TY = This year Sales

     LY = Last year Sales

     RC = Return Center – Depots have a Return Center. The cost of returns are redistributed over all locations.

Filters

     Dollar amount (Net, Gross)

     Regions (NW, BA, MW…)

Calculations

     TY - Rounded to 1000

     LY - Rounded to 1000

     % CHG - (TY minus LY) divided by LY

Fields will relate in many ways.
- State and City relate together in a hierarchy.
- Sales dollars and Sales units relate in a correlation. The more units I sell, sales dollars increase.
- Fields can be used to join spreadsheets or tables together to share data between them.
- Fields can be used together in calculations to create new fields.

With data visualization, fields tell a story. Your job is to figure out what that story is and tell it.

No matter what the fields are, stories all have something in common. They answer a question and all questions come from the 6 W's

WHO, WHAT, WHEN, WHERE, WHY, HOW

After making your glossary look at each field and think of the 6 W's,

Then look at groups of fields and think of the 6 W's.

This is a great time to just write out every question that you can think of. As the old adage goes,

" There are no dumb questions."

Let the questions really fly here and you will find that a golden question will develop.

01 EXERCISE UNDERSTAND YOUR DATA

Find a basic report you or a coworker are currently working with. Build a glossary for it.

REVIEW OF GLOSSARY

Sales Report - Was unable to confirm with business sponsor what this report is used for. They requested to "see something cool".

Report

Department#	Long Dept Desc	Item#	DS.Item Description	DS.Catagory	DS.Catagory Description	Sales	Units Sold
13	FOOD	51	DM WHL PLD TOMATO #10	F	#10 TOMATO	0	0
13	FOOD	61	WHT HOMINY #10 TEASDALE	F	#10 VEGETABLES	314	90
13	FOOD	477	ARM&HAMR BAKING SODA 10#	D	BAKING COMPONENTS	0	0
13	FOOD	1069	DM GREEN BEANS #10	F	#10 VEGETABLES	3	1
13	FOOD	1071	DM CORN WHL KERNEL #10	F	#10 VEGETABLES	349	94
13	FOOD	1098	PUMPKIN BREAD MIX 4PK/64Z	D	BAKING MIXES	96,591	12,089
13	FOOD	1241	SESAME OIL 1 LITER	C	ALL OTHER OIL	12	1
13	FOOD	1399	STAR MARASCHINO CHERRIES	F	#10 FRUITS	0	0
13	FOOD	2251	PINEAPPLE CHUNK 10# DOLE	F	#10 FRUITS	492	93

Report Tabs

FOOD SUNDRIES / HARDLINE / SOFTLINE / ANCILLARY
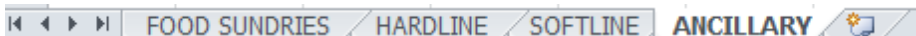
Fields

DS.Category = Category Code. Items are organized in a hierarchy of 3 category codes CAT1, CAT2, CAT3. CAT1 is the top of the hierarchy. Categories are grouped by departments. For this report DS.Category is equal to CAT1, the most commonly used Category Code.

Filters

Department Group Filter - Departments are grouped together and separated by worksheets. Groups are: FOOD SUNDRIES, HARDLINE, SOFTLINE, ANCILLARY

Calculations

There are no Calculations.

Questions

How many items are sold in each location?

What location is doing best or worse?

Where are people shopping?

What are the best categories in different regions or warehouses?

CONNECT TO DATA

- Connect to Data
- Filter
- Clean Data
- Join
- Repeat Filter, Clean, Join
- Live verses Extract

OVERVIEW – CONNECT TO DATA

CONNECTING TO .CSV
CONNECTING TO EXCEL
CONNECTING TO A SERVER

02 EXERCISE CONNECT TO DATA

1. Download the following 2 exercise files and save to your computer.

   Exercise file1

   Exercise file2 - Located P:\Corporate Trainers\MatthewMorris

   - Sales.txt
   - Dept and Cat descriptions.csv

2. Google may have converted the files to .xls or .xlsx. If it does, open each file and save as a .csv file.
3. Open Tableau. From the Main page, Under the Connect Menu, Select **Text file**.
4. Navigate to your exercise files and select **Sales.csv**.
5. Once connected the Data connection page will open.
6. Under files you should see Dept and Cat descriptions.csv & Sales.csv

IS – Advanc
Data Visual
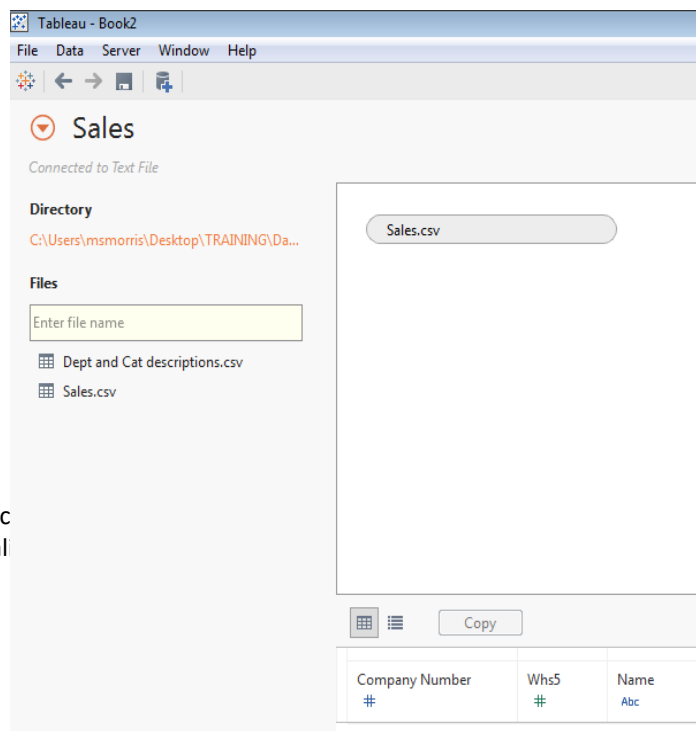
OVERVIEW FILTER DATA

Tableau is a data visualization tool that can perform all kinds of filtering. While it can perform these filters, it's best to have as much as possible filtered before using Tableau. You would not want to process a billion records in Tableau. You can, but the performance would be horrible and no one would use your data visualization except for you. It is important to keep your data streamlined for visualization rather than using the tool to clean data.

 Filter early and filter often.

Tableau gives you lots of options for filtering. The first is when connecting to your data.

03 EXERCISE FILTER

1. In the upper right-hand corner, click on **Add…** link to Filters.
2. In the Edit Data Source Filters pop-out window, click on the **Add...** button.
3. Select the field you want filter by (in this exercise, select Dept), then click the **OK** button.
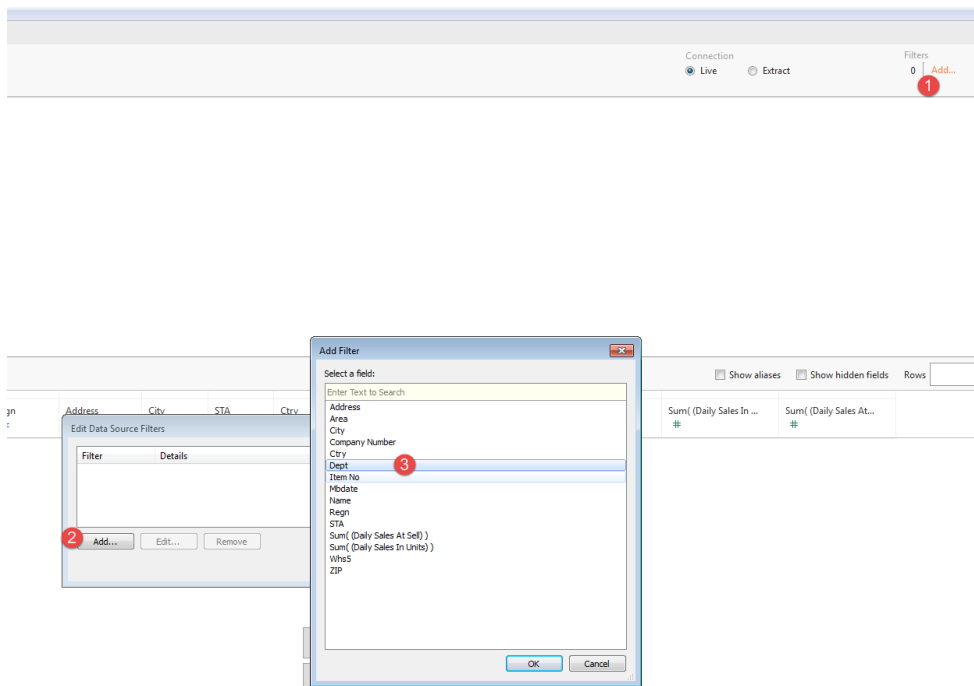
4. After selecting the field you want to filter by, the Filter edit window will display. Here you will have the option to choose; Range of values, At least (equal to or greater than), At most (Equal to or less than), or Special (working with nulls).
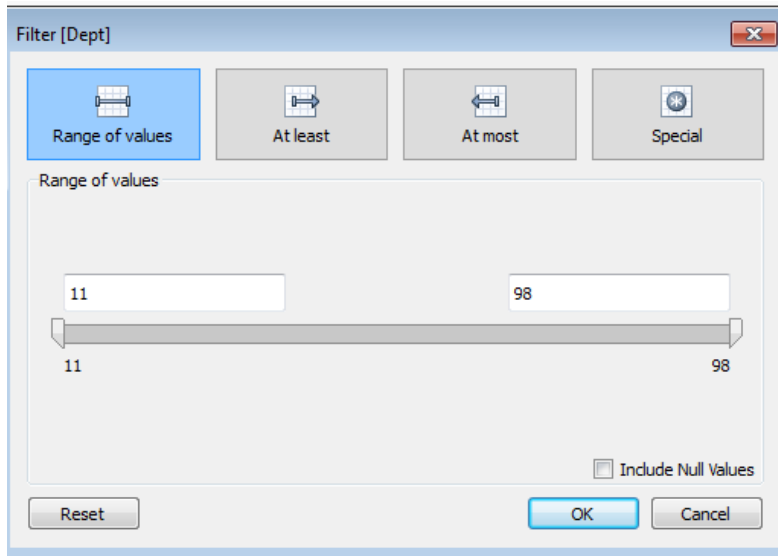
5. Because the data set is small enough we will skip the filter stage and move on to clean, define, rename, or remove fields. Click the **Cancel** button.

OVERVIEW – CLEAN DATA SET

Tableau is a data visualization tool. While it can perform other functions such as cleaning data, it is best to have this done before connecting Tableau to your data. However, for fine tuning and cleaning data, Tableau can help.

CLEANING DATA

Cleaning data is the act of removing, replacing, modifying data that is; incomplete, incorrect, inaccurate, or irrelevant.

Really it means to alter the data in any way that you want!

Remember it is best to do this with other extract tools before connecting tableau to your data. The cleaning you do here should be minimal. Tableau is a data visualization tool not a data cleansing tool.

04 EXERCISE CLEANING THE DATA SET

1. If the data isn't already visible, click on **Update Now** button.
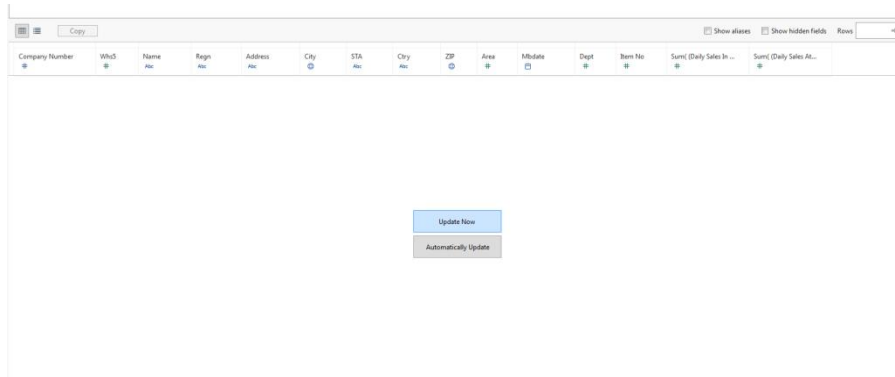
2. Do a visual scan of the data. How long did it take to load? If it took longer than five minutes you will need to put some filters on your data set.
3. Starting from left to right, look at each field.
4. Mouse over **Company Number** and click on the drop down arrow.
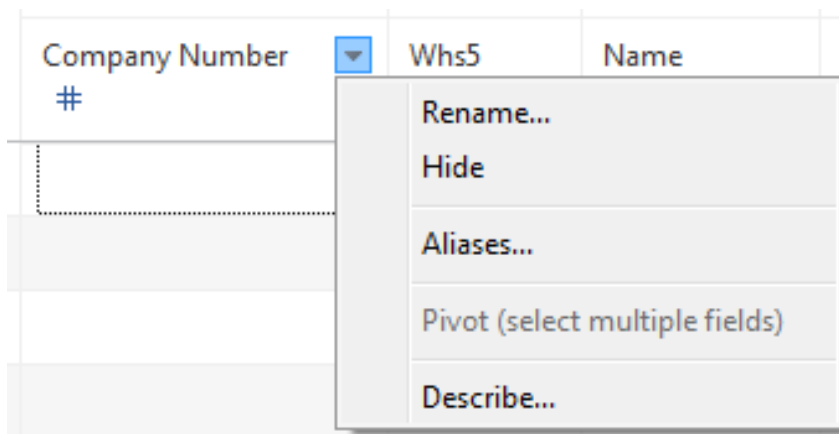
5. Click **Rename…,** then rename the field to **Company#.** Renaming fields can be some personal preference but keep in mind someone down the road will need to understand your field names. Using a field RF78345-QL4 is likely going to be difficult for someone else to know what you were thinking.
6. Look at the different symbols for each field. **Company number** has a # symbol to indicate it is a Number **Whs5** has a # to indicate it is a number. **Name** has Abc to indicate it is a string.
7. Click on the **Abc** symbol for **Regn**.
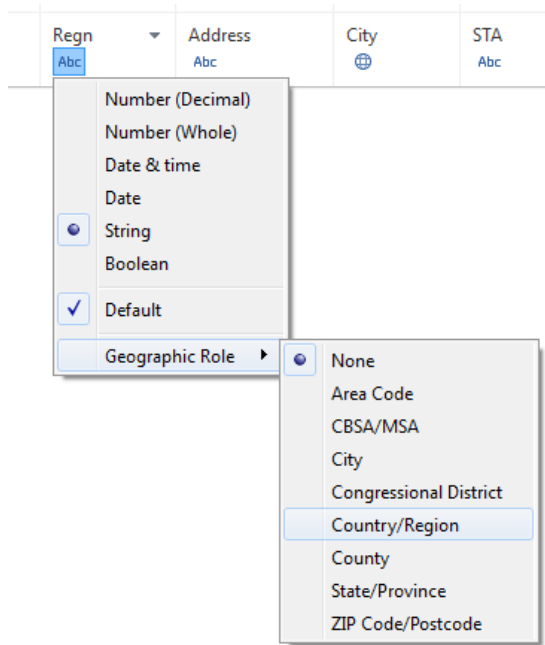
8. Click on **Geographic Role**.
9. Click on **Country/Region**. You have just altered the **Regn** field from string to a Geo coded field.
10. Using the table below rename old field names into new ones and match the data properties.

OLD FIELD NAME	NEW FIELD NAME	DATA PROPERTY
COMPANY NUMBER	Company#	Number Whole
WHS5	Warehouse#	Number Whole
NAME	Warehouse Name	String
REGN	Region	Geo
ADDRESS	Address	String
CITY	City	Geo
STA	State	Geo
CTRY	Country	Geo
ZIP	Zipcode	Geo
AREA	Area Code	Geo
MBDATE	Sales Date	Date
DEPT	Department#	String
ITEM NO	Item#	Number Whole
SUM([DAILY SALES IN UNITS])	Units Sold	Number Whole

SUM([DAILY SALES AT SELL])   Sales   Number Decimal

JOIN

Tableau is a data visualization tool that can perform joins. While it can perform joins its best to have your joins done before using Tableau.

When thinking of large sets of data and various tools, think of someone sculpting a marble block. At first it's a huge project and you need big hammers and chisels. As the sculpting project continues, you will switch to smaller hammers and chisels and finally towards the end, when you are doing the detail work to make your sculpture beautiful, you will use smaller tools. Tableau is a tool you use at the end of your data sculpting. You would not want to use this refined tool to chip away at a large marble block of data. It will take forever and there are much better tools for those kinds of jobs.

You have cleaned your data up but you need more data from another table. No problem you can link the tables together easily.

05 EXERCISE JOIN SALES TO DEPT AND CAT DESCRIPTION

1. Drag the **Dept and Cat descriptions.csv** into the white area next to **Sales.csv** this will create a default join.
2. Click on the two overlapping circles between the tables to review the join.
3. Under Data Source click in the cell that has the text "add new join clause" this will drop down a list of fields in the data source.
4. Choose **Department# (DEPT)** in both the data sources. Note that you will get an error message. Change the DEPT (Dept and Cat descriptions.csv) from Number(Whole) to String. Once this is done, the join will work.
5. Repeat this process to also join on **Item# (ITEM NO)**.
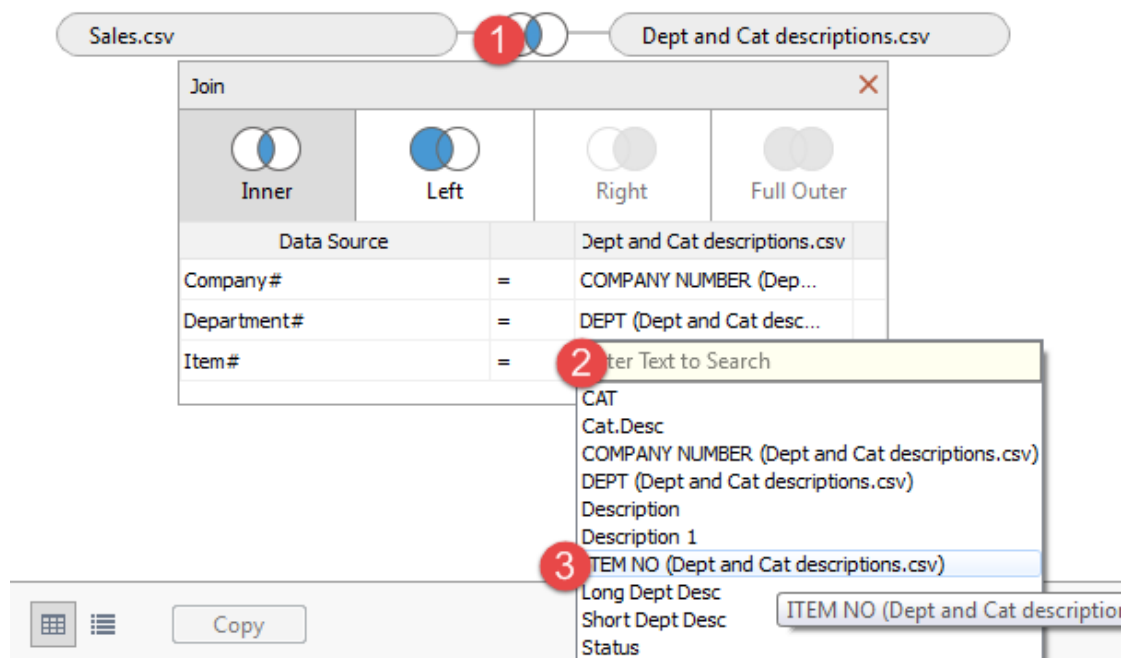
Repeat- FILTER, CLEAN, NEXT JOIN

The next steps after creating a join is to filter, then clean the data, and then move to the next join. Remember to filter early and to filter often. Do not create a lot of joins, then start your filters, and then do your cleaning. The order of operations is:

1. Add a table
2. Filter
3. Clean
4. Join
5. Repeat

Remember Tableau is a great data visualization tool, but you can do a lot of other functions with Tableau such as Cleaning and joining. It is best to prep your data as much as possible before opening Tableau.

06 EXERCISE FILTER, CLEAN, NEXT JOIN

Filter only active items

1. Under Filters click **Add….**
2. Click on the **Add** button.
3. Click **Status**, then Click OK.
4. Click into the box next to A.
5. Click OK and OK again to exit the filter window.

Clean Data

1. Rename the new fields and change the properties to match the table below:

OLD FIELD NAME	NEW FIELD NAME	DATA PROPERTY
COMPANY NUMBER	DS.Company#	Number Whole
DEPT	DS.Department#	Number Whole
LONG DEPT DESC	Long Dept Desc	String
SHORT DEPT DESC	Short Dept Desc	String
ITEM NO	DS.Item#	Number Whole
DESCRIPTION	Description	String
DESCRIPTION	Description 1	String
Status	Status	String

CAT	Category	String
CAT.DESC	Category Description	String

2. There are no more joins, you are ready to complete the last step. Removing unwanted fields. This makes it much easier to navigate your data when you are building your views.

   Remove fields that are irrelevant or redundant. Any fields that do not have multiple values can be removed and put in a header unless you plan on using them for some reference with another field. A good example of a field that can be hidden is **Company#** and **DS.Company#.** The data is filtered onto one company number rather than displaying this on rows of data. It can simply be placed in a header if needed.

   Hide fields: Company#, DS.Company#, DS.Department#, Short Department Description, DS.Item#, Description 1

LIVE VERSES EXTRACT

TIP

How can I change CYMMDD formats to something usable in Tableau?

You will need to prep this data outside of tableau either in your SQL(Query tool) or excel. Here is a quick run down on how to do this in Excel.

Convert Date - Some dates will need to be converted. CYMMDD dates will require some prep work either with excel or ETL before going into tableau. Here is how to change CYMMDD in Excel.

•Select the single-column range of "dates"
•Engage Text-to-columns
••••Check: Fixed width....click: Next
••••Set an insertion point between the 1st and 2nd digits....click: Next
••••Click the 1st column....Check: Skip this column
••••Click the 2nd column....Check: Date: YMD...Click: Finish

How many joins can I do?

Tableau is a data visualization tool if you find yourself stacking lots of joins together you will want to clean this data outside of Tableau first. The more joins you do the slower your visualizations will run.

What is Data Blending?

Data blending is a feature of Tableau that is seldom needed or used. For anyone that has used SQL this is like creating a subselect statement to another data source within your query. It's a way to join data into your data set without doing a join but more of a link

based on a field on an as needed basis. This will be covered in more detail in Intermediate and Advanced.

REVIEW

What are your thoughts on Tableau's ability to clean data?

What should you do if you see that you have created a 15 table join process in Tableau?

When connecting to your data if it takes five minutes to display what might be your next steps?

Place these in the correct order: ADD DATA SOURCE, JOIN, CLEAN, FILTER

Will you be able to make good decisions about joins if you have not completed the research phase of your project?

True or False, it is best to do all joins first then filter your data after?

PREPARE DATA WINDOW

Organize Data

Create Folders

Create Hierarchies

Create Groups

ORGANIZE DATA

Creating Folders

Setting up folders will help navigating and finding your data much easier. Folders are commonly named after a set of dimensions related together. Example of common folders are: LOCATION, PRODUCT, DATES, VENDORS, DEPARTMENTS

Creating Hierarchies

Hierarchies are a way to relate dependencies between data. it makes drilling through data much easier. An example of a common Hierarchy is: COUNTRY, REGION, STATE, CITY, ZIPCODE

Creating Groups

A group is a way to combine fields into a categories. For example this could be several products grouped as "Department".

In tableau you can create groups three different ways;

1. Data Window
2. View
3. Visual grouping

Groups sometimes are obvious and if they are, it is best to set them up while you are prepping the Data window. If you do not see any obvious groups, they may become obvious during the data discovery phase.

Here are three exercises to get you started with Folders, Hierarchies and Groups

07 EXERCISE CREATE FOLDERS

1. Click on the **Sheet1** tab.

IS – Advanced Analytics
Data Visualization Fundamentals 30

2. In the data pane Right click in the white space and choose **Group by folder**.
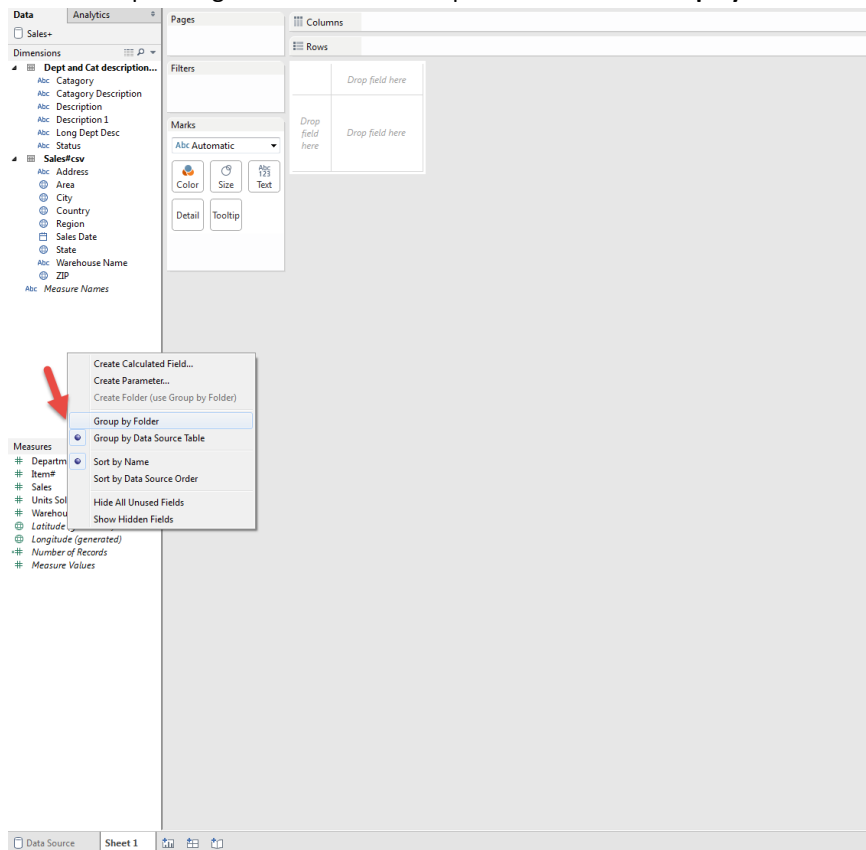
3. Dimensions will often be incorrectly placed in measures, drag and drop these measures anywhere in the dimensions part of the data pane.
    a. Department#
    b. Item#
    c. Warehouse#
4. Create Dimension folders by clicking in the whitespace of the dimensions pane and choosing create folder.
5. In the create folder prompt type **PRODUCT**.
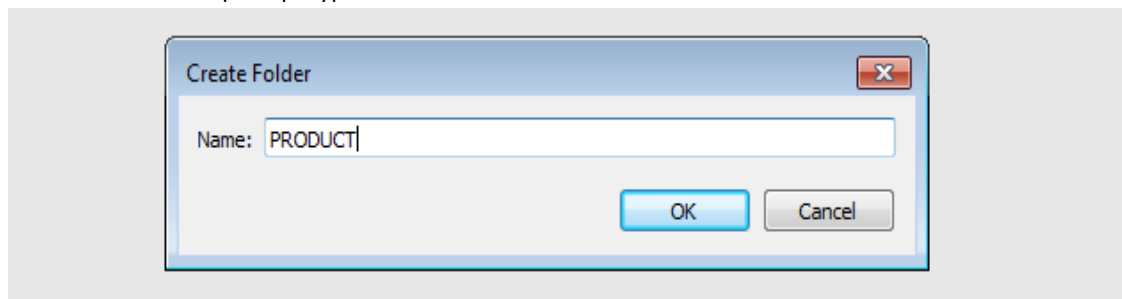
6. Create a DATE Folder and a LOCATION folder.

Data	Analytics	
Sales+		

Dimensions

- DATE
- LOCATION
- PRODUCT
- Abc Address
- ⊕ Area
- Abc Catagory
- Abc Catagory Description
- ⊕ City
- ⊕ Country
- # Department#
- Abc Description
- Abc Description 1
- # Item#
- Abc Long Dept Desc
- ⊕ Region
- 📅 Sales Date
- ⊕ State
- Abc Status
- Abc Warehouse Name
- # Warehouse#
- ⊕ ZIP
- Abc *Measure Names*
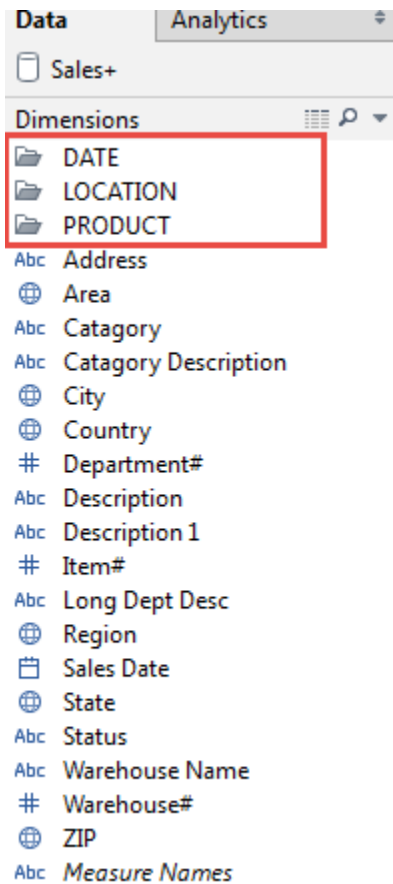
7. The following field need to be dragged and dropped into the DATE folder:
Sales Date

8. The following fields need to be dragged and dropped into the LOCATION folder:
Address, Area, City, Country, Region, State, Warehouse Name, Warehouse#, Zipcode

9. The following fields need to be dragged and dropped into the Product folder:
Category, Category Description, Department#, Description, Description1, item#, Department Description

Data   Analytics

Sales+

Dimensions

▲ 📂 DATE
      📅 Sales Date
▲ 📂 LOCATION
    Abc Address
    ⊕ Area
    ⊕ City
    ⊕ Country
    ⊕ Region
    ⊕ State
    Abc Warehouse Name
    # Warehouse#
    ⊕ Zipcode
▲ 📂 PRODUCT
    Abc Catagory
    Abc Catagory Description
    # Department#
    Abc Description
    Abc Description 1
    # Item#
    Abc Long Dept Desc
    Abc Status
  Abc *Measure Names*
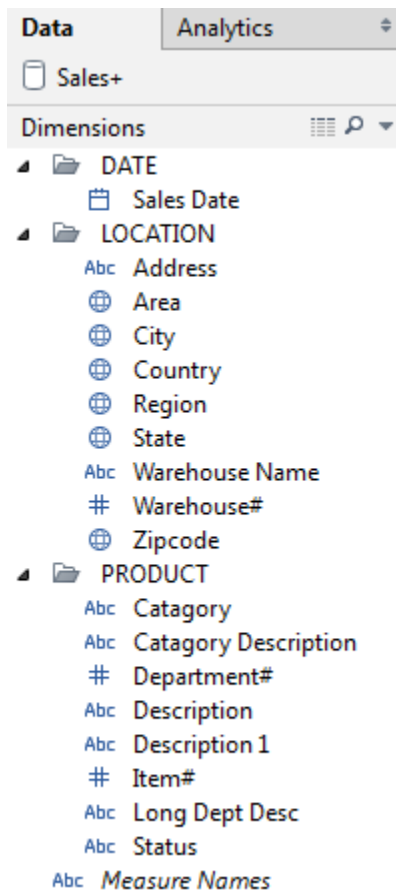
10. Using the arrows minimize each field to keep your data pane window easy to view and use.
   ▷ 📁 DATE
   ▷ 📁 LOCATION
   ▷ 📁 PRODUCT
     Abc *Measure Names*
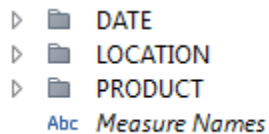
1. In the dimensions section of the data pane open the **LOCATION** folder.
2. Right click on **Country**.
3. Click on **Hierarchy**.
4. Click on **Create Hierarchy**.
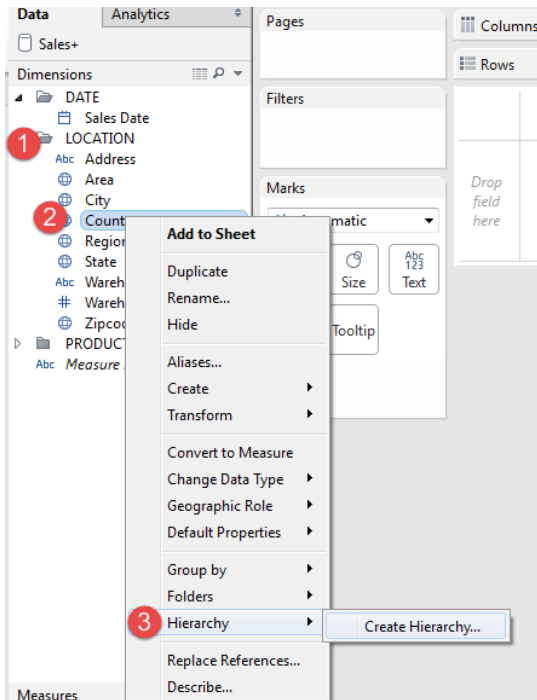
5. Name your hierarchy **Location.**
6. The following fields need to be dragged and dropped into the Location Hierarchy in this order: Country, Region, State, Area , City, ZipCode, Warehouse #, Warehouse Name, Address
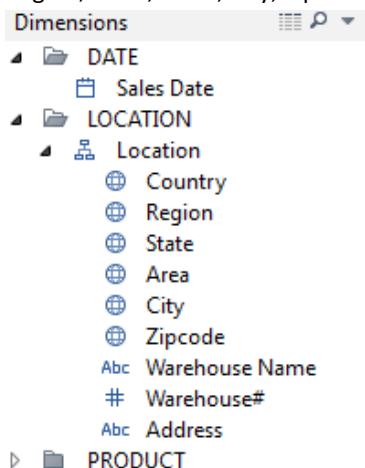
09 EXERCISE CREATE GROUPS

1. Right click the dimension **Department#**.
2. Click on **Create**.
3. Click on **Group**.
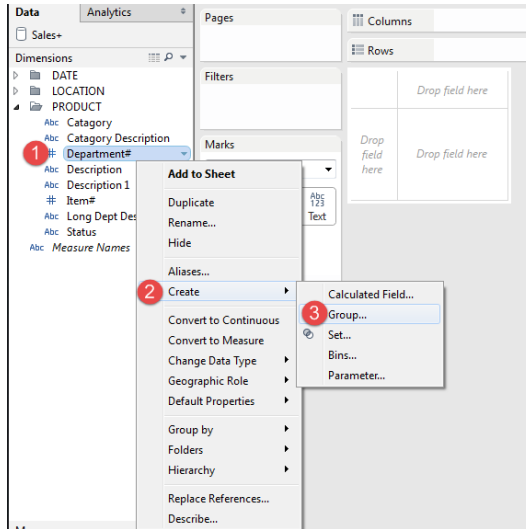
4. Holding the **Shift key** down select **11** and **19**.
5. Click on the **Group** button.
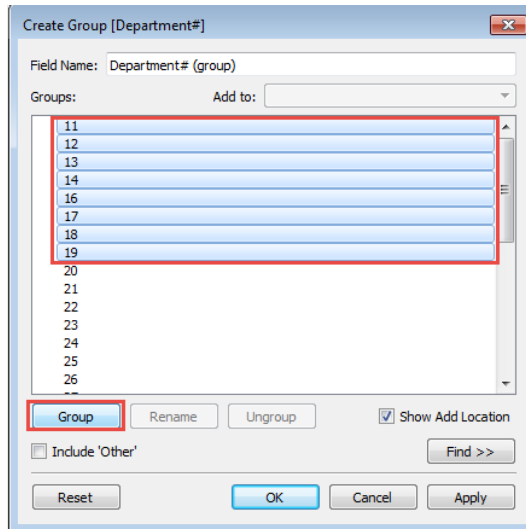
6. Rename Group **FOOD-SUNDRIES**.
7. Click **OK**.
8. Notice the New dimension Department#(group).

The rest of the departments will be grouped in the view later.

DATA DISCOVERY

The best way to discover your story, learn about your data and get insights is to view it. You can do this using: CROSSTAB TOTALS, CROSSTAB ANALYSIS, BAR CHART, STACKED BAR CHART, PIE CHART, GROUPS, LINE GRAPH, COMBINED AXIS, DUAL AXIS, SCATTER PLOT, SYMBOL MAP OR DOT MAP, FILLED MAP, HIGHLIGHT TABLE, HEAT MAP, TREE MAP, CONSTANT REFERENCE LINE WITH BAR CHARTS, BUBBLE WORD CLOUD and a MOTION CHART.

CHARTS

How do you know which chart to use when? The following definitions cover the charts you will be working on or modifying. The following charts and definitions come from Datavizcatalogue.com. The site has even more charts to explore and when to use them.

Bar Chart
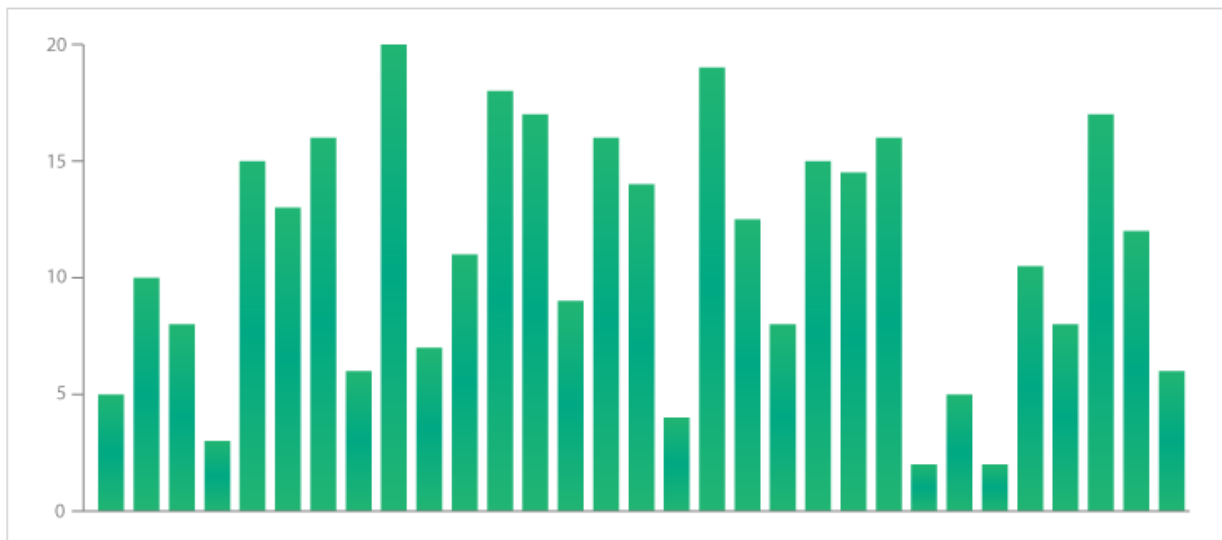
COMPARISONS, RELATIONSHIPS, PATTERNS

The classic bar chart uses either horizontal or vertical bars (AKA column chart) to show discrete, numerical comparisons among categories. One axis of the chart shows the specific categories being compared, and the other axis represents a discrete value.

Bars charts are distinguished from histograms, as they do not display continuous developments over an interval. Bar chart's discrete data is categorical data and answers the question of "how many?" in each category.

One major flaw with bar charts is labeling becomes problematic when there is a large amount of bars.

Stacked Bar Graph
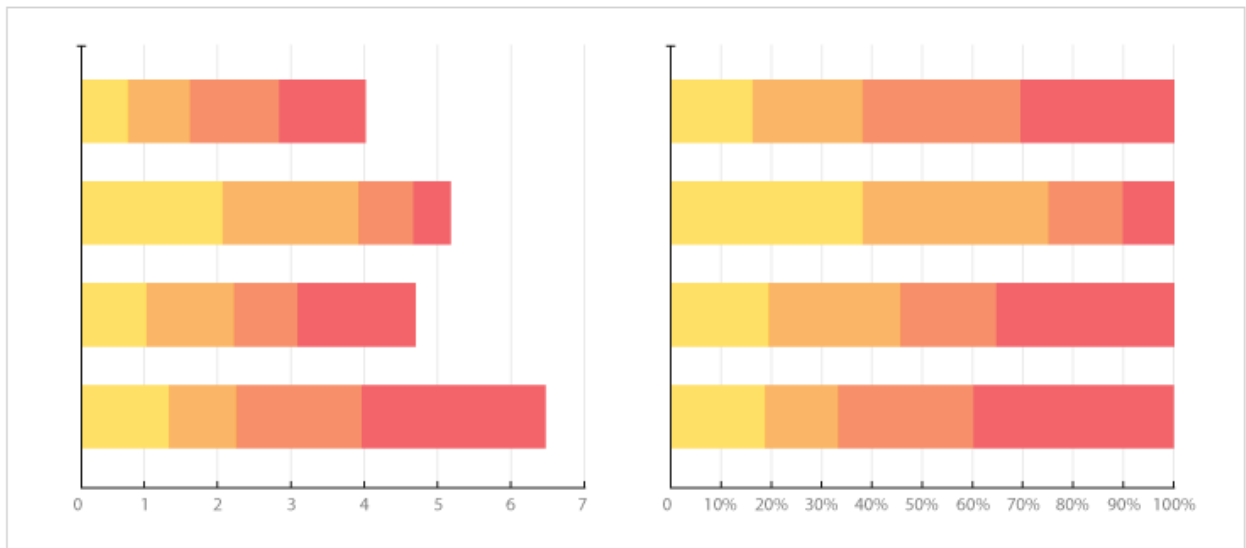
COMPARISONS, RELATIONSHIPS, PATTERNS, PROPORTIONS, PART-TO-A-WHOLE (When 100% Stacked Bar Graph)

Unlike a Multi-Set-Bar Graph which displays their bars side-by-side, Stacked Bar Graphs segment their bars of multiple datasets on top of each other. They are used to show how a larger category is divided into smaller categories and what the relationship of each part has on the total amount. There are two types of Stacked Bar Graphs:

Simple Stacked Bar Graphs place each value for the segment after the previous one. The total value of the bar is all the segment values added together. Ideal for comparing the total amounts across each group/segmented bar.

100% Stack Bar Graphs show the percentage-of-the-whole of each group and are plotted by the percentage of each value to the total amount in each group. This makes it easier to see the relative differences between quantities in each group.

One major flaw of Stacked Bar Graphs is that they become harder to read the more segments each bar has. Also comparing each segment to each other is difficult, as they're not aligned on a common baseline.

Pie Charts
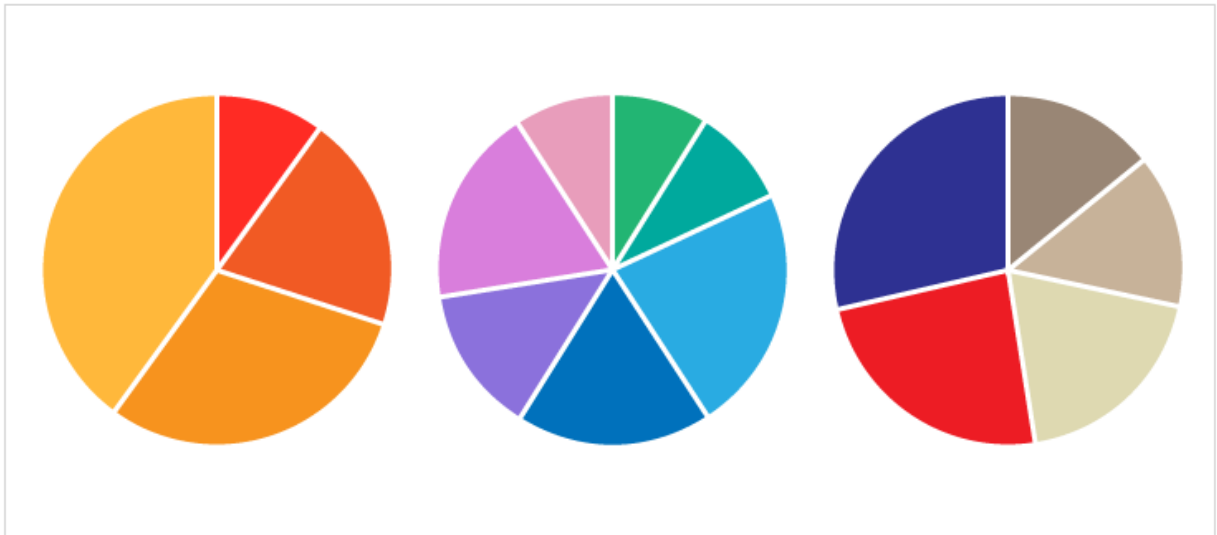
COMPARISONS, PART-TO-A-WHOLE, PROPORTIONS

Extensively used in presentations and offices, pie charts help show proportions and percentages between categories, by dividing a circle into proportional segments. Each arc length represents a proportion of each category, the full circle represents the total sum of all the data, equal to 100%.

Pie charts are ideal for giving the reader a quick idea of the proportional distribution of the data. However the major downsides to pie charts are:

  They cannot show more than a few values, because as the number of values shown increases, the size of each segment/slice becomes smaller. This makes them unsuitable for large amounts of data.

  They take up more space than their alternatives, like a 100% bar chart for example. Mainly due to their size and for the usual need for a legend.

  They are not great at making accurate comparisons between groups of pie charts. This being that it is harder to distinguish the size of items via area when it is for length.

In spite of that, comparing a given category (one slice) within the total of a single pie chart, then it can often be more effective.

Line Graph
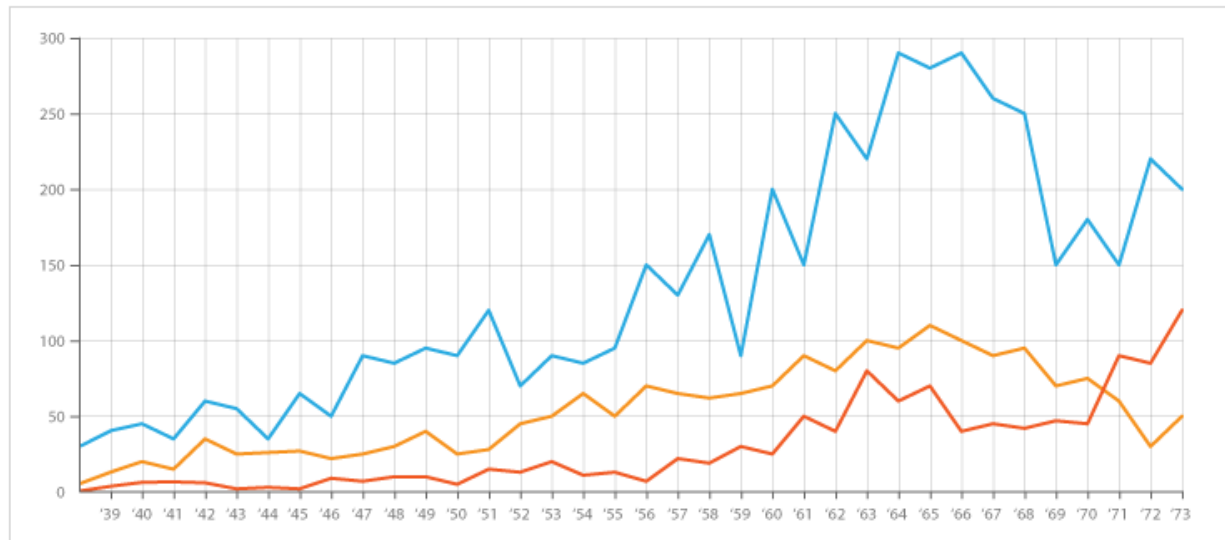
PATTERNS, DATA OVER TIME

Line graphs are used to display quantitative value over a continuous interval or time span. It is most frequently used to show trends and relationships (when grouped with other lines). Line graphs help to give a "big picture" over an interval, to see how it has developed over that period.

Line graphs are drawn by first plotting data points on a cartesian coordinate grid, then connecting a line between the points. Typically, the y-axis has a quantitative value, while the x-axis has either a category or sequenced scale. Negative values can be displayed below the x-axis.

Scatterplot

Positive Negative Null

Linear Exponential U-Shaped

Correlation Strength:

Strong Weak None
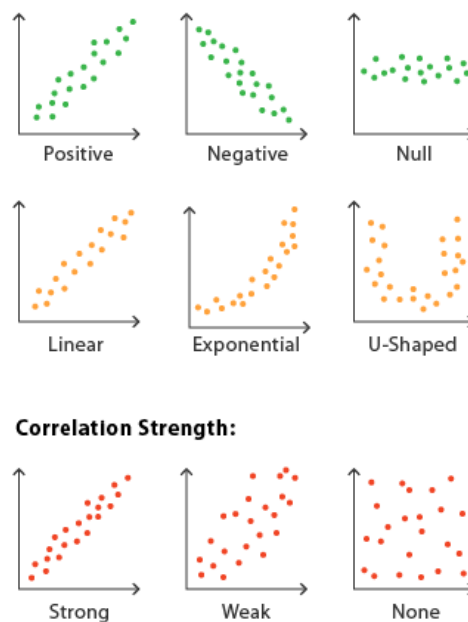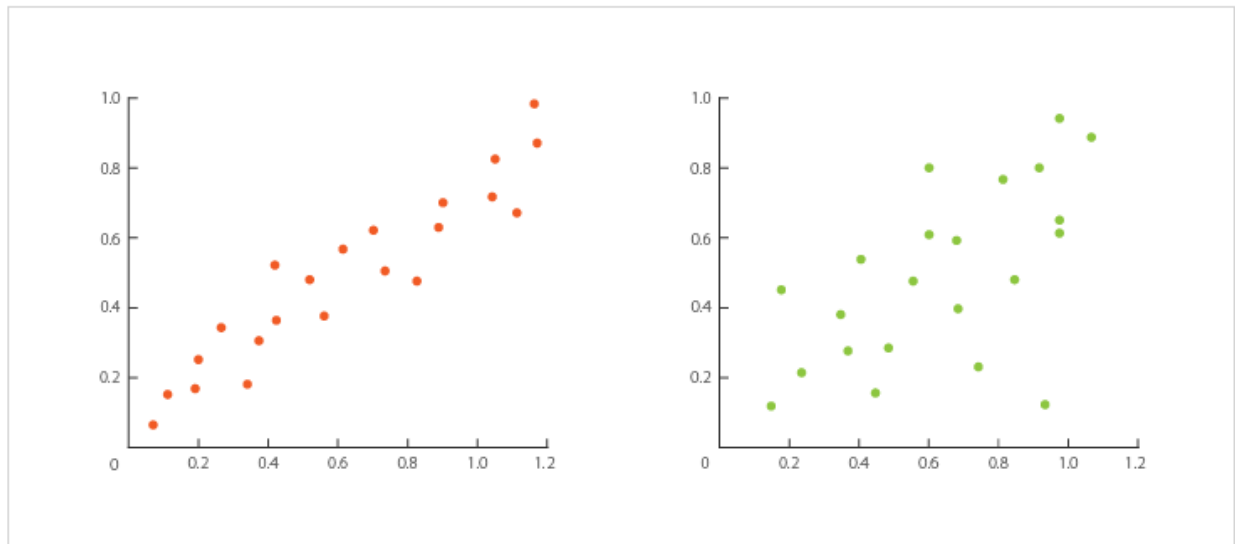
PATTERNS RELATIONSHIPS

Also known as: *scatter graph, point graph, X-Y plot, scatter chart or scattergram.*

Scatterplots use a collection of points placed using Cartesian Coordinates to display values from two variables. By displaying a variable in each axis, you can detect if a relationship or correlation between the two variables exists.

Various types of correlation can be interpreted through the patterns displayed on Scatterplots. These are: **positive**(values increase together), **negative** (one value decreases as the other increases), **null** (no correlation), **linear**, **exponential** and **U-shaped**. The strength of the correlation can be determined by

how closely packed the points are to each other on the graph. Points that end up far outside the general cluster of points are known as **outliers**.

Lines or curves are fitted within the graph to aid in analysis and are drawn as close to all the points as possible and to show how all the points were condensed into a single line would look. This is typically known as the **Line of Best Fit** or a **Trend Line** and can be used to make estimates via interpolation.

Scatterplots are ideal when you have paired numerical data and you want to see if one variable impacts the other. However, do remember that correlation is not causation and another unnoticed variable may be influencing results.

Dot Map
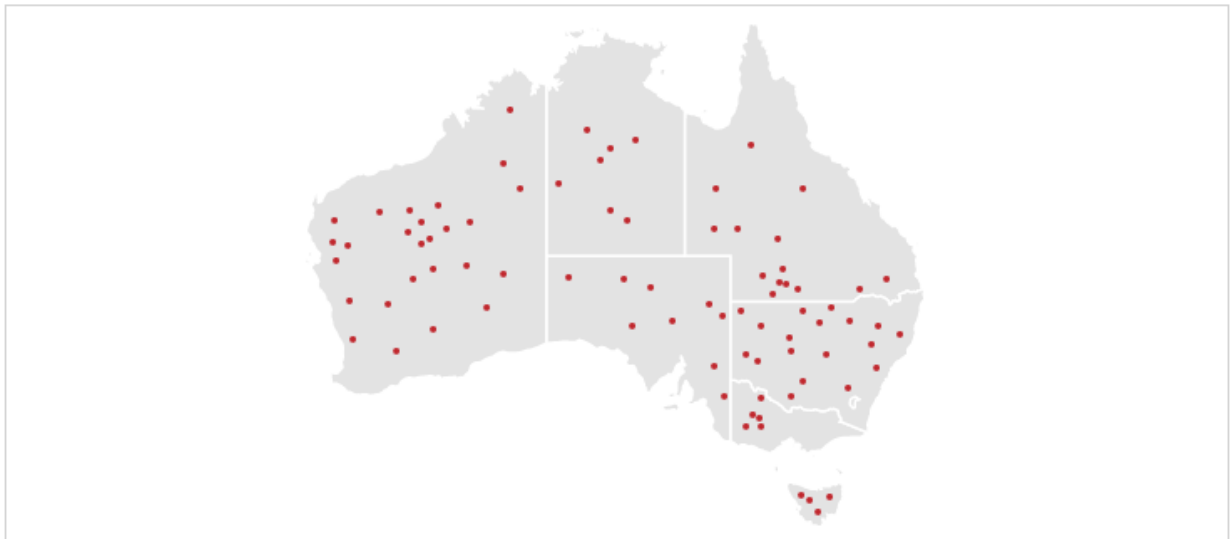
DISTRIBUTION, LOCATION, PATTERNS

Also known as: DOT DISTRIBUTION MAP, DOT DENSITY MAP.

Dot Maps are a way of detecting spatial patterns or the distribution of data over a geographical region, by placing equally sized points over a geographical region.

There are two types of Dot Map: **one-to-one** (one point represents a single count or object) and **one-to-many** (one point represents a particular unit, e.g. 1 point = 10 trees).

Dot Maps are ideal for seeing how things are distributed over a geographical region and can reveal patterns when the points cluster on the map. Dot Maps are easy to grasp and are better at giving an overview of the data, but are not great for retrieving exact values.

Treemap

COMPARISONS, HIERARCHY, PART-TO-A-WHOLE, PROPORTIONS

Treemaps are an alternative way of visualizing the hierarchical structure of a Tree Diagram while also displaying quantities for each category via area size. Each category is assigned a rectangle area with their subcategory rectangles nested inside of it. When a quantity is assigned to a category, its area size is displayed in proportion to that quantity and to the other quantities within the same parent category in a part-to-whole relationship. Also the area size of the parent category is the total of its subcategories. If no quantity is assigned to a subcategory, then it's area is divided equally amongst the other subcategories within its parent category.

The way rectangles are divided and ordered into sub-rectangles is dependent on the **tiling algorithm** used. Many tiling algorithms have been developed, but the **squarified algorithm** which keeps each rectangle as square as possible is the one commonly used.

Ben Shneiderman originally developed Treemaps as a way of visualizing a vast file directory on a computer, without taking up too much space on the screen. This makes Treemaps a more compact and space-efficient option for displaying hierarchies, that gives a quick overview of the structure. Treemaps are also great at comparing the proportions between categories via their size.

Word Cloud

Plato Republic

ANALYSING TEXT, DISTRIBUTION/FREQUENCY, PROPORTIONS

Also known as: *Tag Clouds*

A visualisation method that displays how frequently words appear in a given sample of text, by making the size of each word proportional to its frequency. All the words are then arranged in a cluster or cloud of words. Alternatively, the words can also be arranged in any format: horizontal lines, columns or within a shape etc.

Word Clouds can also be used to display words that have metadata assigned to them. For example, in a Word Cloud of all the World's countries, population could be assigned to each country's name to determine its size.

Colour used on Word Clouds is usually meaningless and is primarily aesthetic, but it can be used to categorise words or to display another data variable.

Typically, Word Clouds are used on websites or blogs to depict keyword or tag usage. Word Clouds can also be used to compare two different bodies of text together.

Although being simple and easy to understand, Word Clouds have some major flaws:

   Long words are emphasised over short words.

   Words whose letters contain many ascenders and descenders may receive more attention.

   They're not great for analytical accuracy, so used more for aesthetic reasons instead.

MARKS CARD

1. Marks card
2. Marks type
3. Marks properties
4. Using Color

MARKS CARD

The Marks card is where you can drag fields and alter how they are displayed in the view.

You can do this by changing the **MARKS TYPE** (Charts and Graphs) and by changing the **MARKS PROPERTIES** (color, size etc).
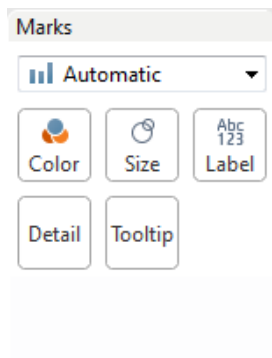
MARKS TYPE

Marks Type is yet another way to create a type of chart or graph. Mark types are available from the Marks card drop-down menu at the top of the Marks card.
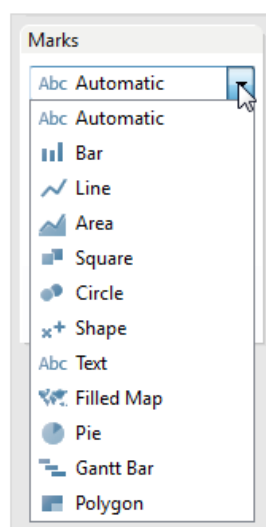
MARKS PROPERTIES
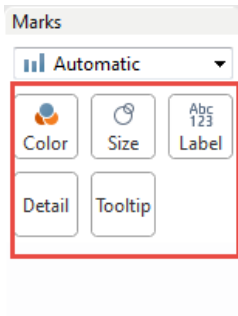
This is where you drag fields to control mark properties such as type, color, size, shape, and so on.

Depending on the MARKS TYPE you may see additional MARKS PROPERTIES. For example if you changed the MARKS TYPE to Pie you would see the new property of Angle.
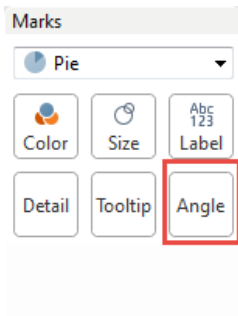
EDIT MARK PROPERTIES

By clicking on the various MARK PROPERTIES you can edit the settings.

Clicking on the Color Mark for example will give you the options to change; Color, Transparency and Border Effects.
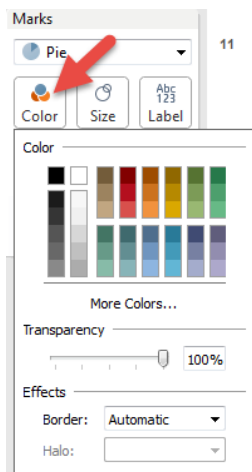
The fields on the Marks card are listed at the bottom of the card. Each field has an icon indicating the mark property it is setting. For example, the Marks card shown below has three fields: Segment is on Color, Region is on Shape, and Quantity is on Size.
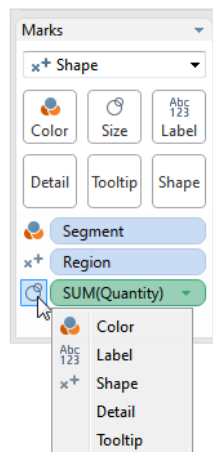
Notice the icons next to each field relate to each mark and can be changed by clicking on the icon and choosing another mark type. The Detail and Tooltip properties don't display with an icon. However, you can click the white space where the icon would display to access the menu.

If you want a field, like sales, to have multiple marks you can drag your sales field onto multiple marks.

Size and Shape can only have one field at a time.

Note: By default, dragging a new field to **Color** replaces the existing fields. To add a new field to color without replacing the existing field, hold the SHIFT key on your keyboard while dragging a new field to **Color** on the Marks card.

Without worrying about the data displaying familiarize yourself with the Marks card

1. Open a new Worksheet.
2. Drag Sales to Columns.
3. Drag Department#(group) to Rows.
4. On the Marks card change the Marks type to each of the different Marks type. Take note of any new Marks property tiles and what visuals you are seeing.
5. Set the Marks type to Bar.
6. Drag Units sold to the Color Marks property.
7. Edit the Colors Mark property by clicking on the Color tile.
8. Click on Edit colors.
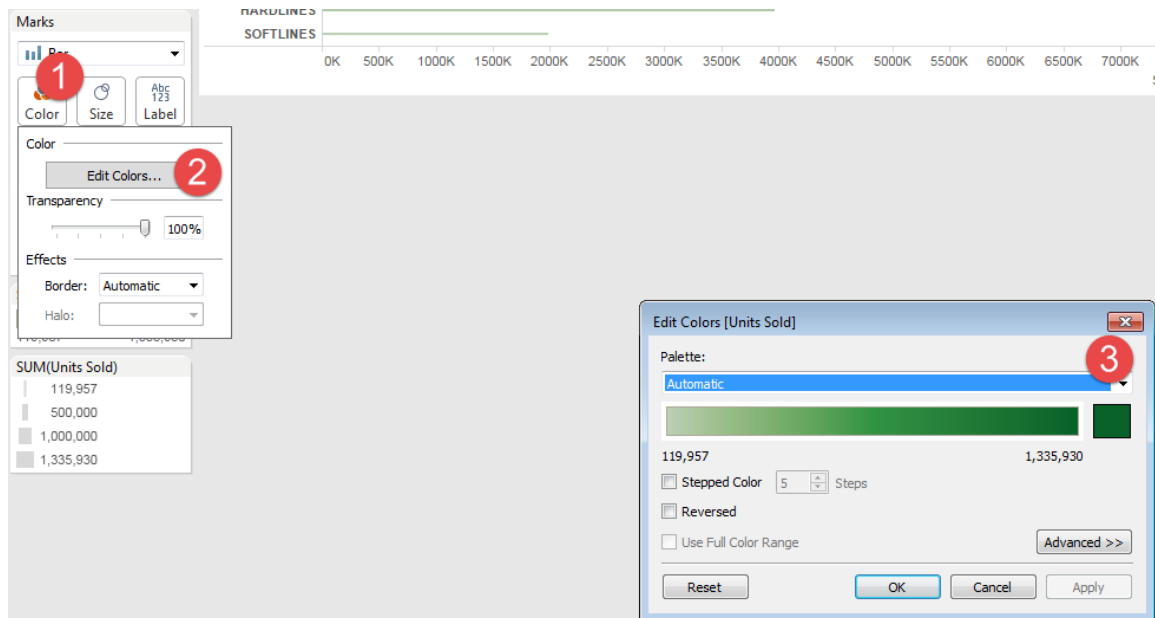
As you can see there are a lot of options for Editing Colors. Much of this is covered in great detail online, in forums and tableau training sites. For now it's fine to learn by doing,

9. Click drop down list and choose a color of your preference.
10. Click Apply and view your data.
11. Edit the color tile again and try various areas of the Edit Color window then view how it has changed your data.

CHALLENGE Explore each of the Marks tiles by editing each one.

Using COLOR

How many nines can you find? How long did it take you to find them?

```
1 8 3 2 4 3 1 1 8 7 5 0 3
6 4 7 5 7 8 1 6 6 0 4 7 8
8 0 4 1 5 3 2 8 4 3 3 9 1
5 3 6 4 9 6 2 5 2 5 2 1 4
6 5 3 2 1 1 2 6 2 7 1 3 2
4 9 6 5 5 7 4 4 6 2 8 4 3
1 2 1 2 6 8 4 1 9 4 0 3 5
6 5 9 1 5 4 5 6 3 6 2 5 7
7 8 6 8 3 6 5 7 1 0 3 3 1
2 6 3 4 8 7 5 2 2 1 6 6 2
1 3 6 5 7 8 6 1 4 3 8 7 0
3 0 1 6 9 5 6 3 0 2 6 4 4
7 8 2 7 0 8 7 7 8 6 4 0 6
4 2 5 6 7 0 7 4 1 4 6 1 8
9 0 3 7 6 2 8 9 3 3 0 8 3
```

Using color is a great way to bring out the data we are looking for much quicker and more accurately. Asking ten people the above question you might get three or even ten different responses back on how many nines there are. Now multiply that by hundreds, thousands, and millions of records that people are processing either manually or through query tools. Not to mention the amount of labor to create all those reports, double checking and going back to fix numbers.

Color adds quick accuracy.

How many nines can you find? How long did it take you to find them?

While color is great, overuse can be a liability.

Here are some general guidelines when using color.

Use your companies' colors.

Use as few colors as possible or at the most between three and five colors unless using diverging colors or many hues of 1 color.

To avoid color blindness problems, here are a few handy rules and approaches to consider: Avoid the following color combinations, which are especially hard on color blind people: **Green** & **Red**; **Green** & **Brown**; Blue & **Purple**; **Green** & Blue; Light **Green** & Yellow; Blue & Grey; **Green** & Grey; **Green** & Black

BUILDING THE CHARTS IN TABLEAU

At this point you should feel familiar with:

1. Understanding data.
2. Various chart types and when to use them.
3. The marks card.

Data analysis is all about measuring something. What are we measuring? Dimensions! How are we measuring them? With Measures! The basic idea of what we are measuring is right in our data LOCATIONS, DEPARTMENTS, ITEMS, SALES, and DATES.

To keep it simple Dimensions are basically nouns (person, place or thing). Yes, it is more complicated than that, but at this level lets leave it there. We are measuring nouns, dimensions, events. Think of when you were a kid and you were growing. You probably measured your height. Height is a dimension and those tick marks are measurements.

To build a chart in tableau there are three areas you should be familiar with; Rows, Columns and the View.

1 Rows – Dimensions

2 Columns – Dimensions

3 View – Measures
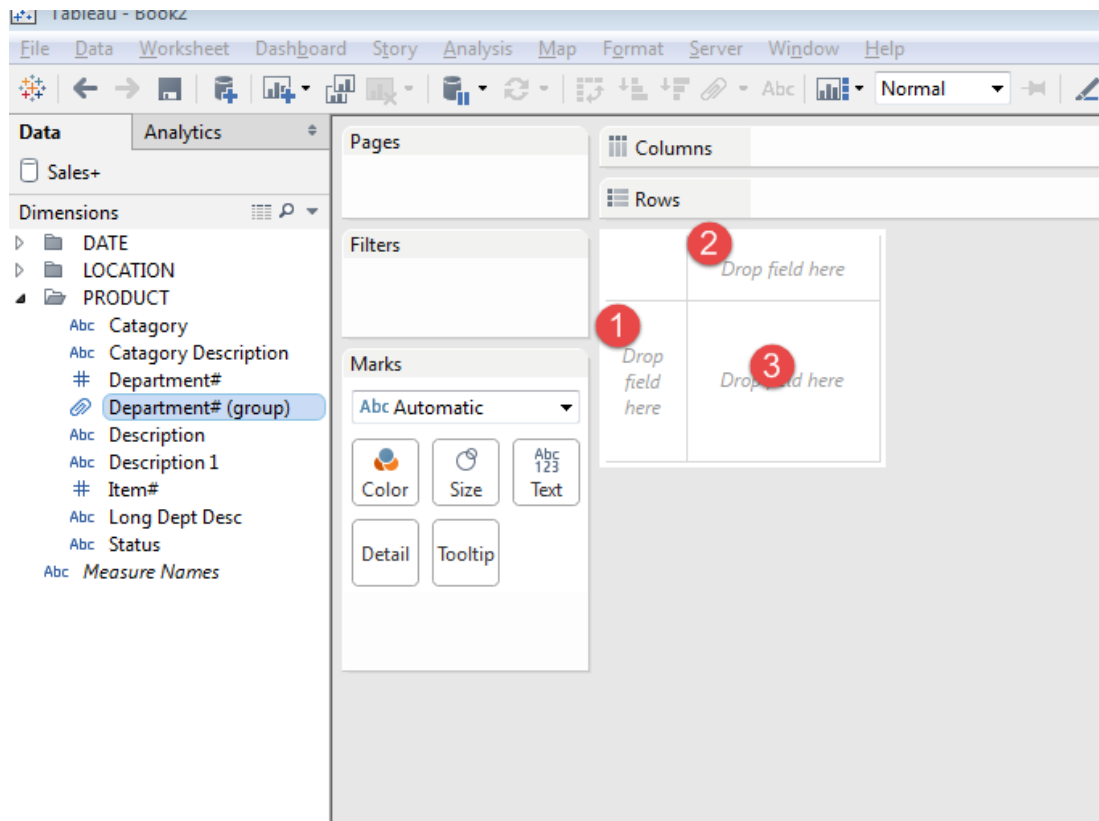
In our own small data set, the options are plentiful so where do we start? Start with the question the client has. All questions are a: Who, What, When, Why, How, Where scenarios. If your question is a where then start with location, if it is a When start with dates. The Type of question determines the Dimensions.

Remember those questions you developed when you were researching your data. That is also a good place to start. You are a data detective and you are following clues, some will lead to dead ends and some will lead to big answers, just keep investigating and asking questions.

11 EXERCISE – CROSSTAB TOTALS AND CREATING GROUPS IN THE VIEW

Let's start with a basic question, "Where are people shopping?" Where= location and shopping = sales.

1. From DIMENSIONS, drag REGION to COLUMNS.
2. From DIMENSIONS, drag DEPARTMENT to ROWS.
3. From DIMENSIONS, drag LONG DEPT DESCRIPTION to ROWS.

4. From MEASURES, drag SALES to TEXT on the Marks Card.
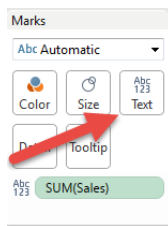
5. You should have something that looks like this:

Department#	Long Dept Desc	Region NW
11	TOBACCO	6,010
12	CANDY	1,864,958
13	FOOD	3,920,157
14	SUNDRIES	671,365
16	LIQUOR	269,254
17	COOLER	1,633,351
18	FROZEN FOOD	2,607,147
19	DELI	3,166,855
20	HEALTH & BEAUTY AIDS	60,513
21	OFFICE SUPPLIES	63,057
22	TIRES	184,462
23	HARDWARE	1,814,695
24	MAJOR APPLIANCES	1,140,456
25	AUTOMOTIVE	178,726
26	SPORTING GOODS	68,900

12 EXERCISE CREATE GROUPS IN THE VIEW
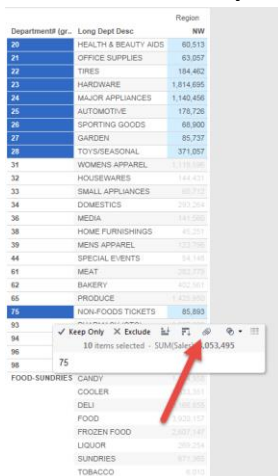
6. From Dimensions Drag and Drop Department#(group) on top of Department# in the rows. You will know the new pill will replace the old one when the old pills border highlights blue.
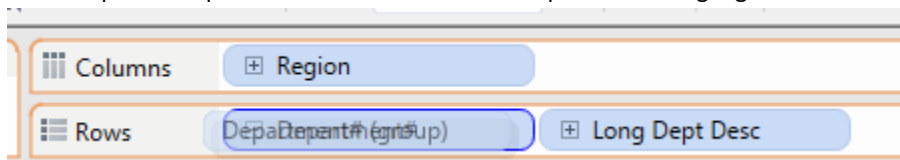
Your crosstab should now look like this.

Department# (gr..	Long Dept Desc	Region NW
20	HEALTH & BEAUTY AIDS	60,513
21	OFFICE SUPPLIES	63,057
22	TIRES	184,462
23	HARDWARE	1,814,695
24	MAJOR APPLIANCES	1,140,456
25	AUTOMOTIVE	178,726
26	SPORTING GOODS	68,900
27	GARDEN	85,737
28	TOYS/SEASONAL	371,057
31	WOMENS APPAREL	1,118,596
32	HOUSEWARES	144,431
33	SMALL APPLIANCES	60,712
34	DOMESTICS	293,264
36	MEDIA	141,560
38	HOME FURNISHINGS	45,251
39	MENS APPAREL	123,796
44	SPECIAL EVENTS	54,148
61	MEAT	282,779
62	BAKERY	402,561
65	PRODUCE	1,425,950
75	NON-FOODS TICKETS	85,893
93	PHARMACY (OTC)	3,227,002
94	OPTICAL	261,634
96	HEARING AIDS	54,799
98	PRINT SHOP	45,897
FOOD-SUNDRIES	CANDY	1,864,958
	COOLER	1,633,351
	DELI	3,166,855
	FOOD	3,920,157
	FROZEN FOOD	2,607,147
	LIQUOR	269,254
	SUNDRIES	671,365
	TOBACCO	6,010

7. Inside the crosstab hold down the shift key and click on **20** then click on **28** to select all departments between **20** and **29**.

8. Next Hold the **Ctrl key** down and click on **75**.

9. Options for the selected fields should pop up. Click on the **paperclip** group icon.

10. Your department should now be grouped together but the name is not useful. Right click on the new group and choose **Edit Alias…**

IS – Advanced Analytics
Data Visualization Fundamentals 52

11. Rename the alias to **HARDLINES**.
12. Repeat the above process for the next two groups:
    a. **SOFTLINES** (31-39,44, 45)
    b. **ANCILLARY**(remaining departments
13. When you are done it should look like the crosstab below.

Department# (gr..	Long Dept Desc	Region NW
ANCILLARY	BAKERY	402,561
	HEARING AIDS	54,799
	MEAT	282,779
	NON-FOODS TICKETS	85,893
	OPTICAL	261,634
	PHARMACY (OTC)	3,227,002
	PRINT SHOP	45,897
	PRODUCE	1,425,950
FOOD-SUNDRIES	CANDY	1,864,958
	COOLER	1,633,351
	DELI	3,166,855
	FOOD	3,920,157
	FROZEN FOOD	2,607,147
	LIQUOR	269,254
	SUNDRIES	671,365
	TOBACCO	6,010
HARDLINES	AUTOMOTIVE	178,726
	GARDEN	85,737
	HARDWARE	1,814,695
	HEALTH & BEAUTY AIDS	60,513
	MAJOR APPLIANCES	1,140,456
	OFFICE SUPPLIES	63,057
	SPORTING GOODS	68,900
	TIRES	184,462
	TOYS/SEASONAL	371,057
SOFTLINES	DOMESTICS	293,264
	HOME FURNISHINGS	45,251
	HOUSEWARES	144,431
	MEDIA	141,560
	MENS APPAREL	123,796
	SMALL APPLIANCES	60,712
	SPECIAL EVENTS	54,148
	WOMENS APPAREL	1,118,596

14. From the Analysis menu, click on **Totals**.
15. Click on **Show Row Grand Totals** and **Show Column Grand Totals**.
16. From the Analysis menu, click on **Totals**.
17. Click on **Add all Subtotals**.

IS – Advanced Analytics
Data Visualization Fundamentals 53
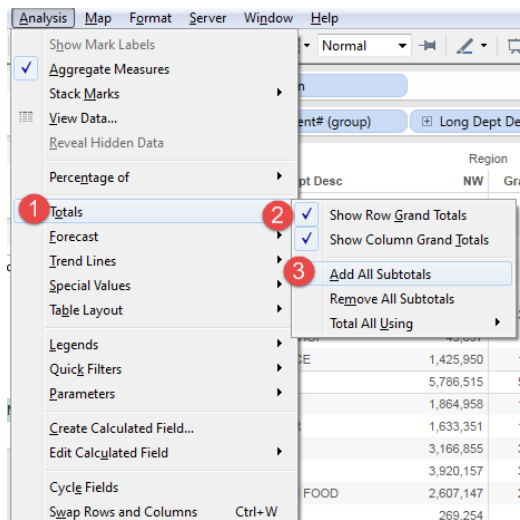
18. Next right click the worksheet tab.
19. **Click on Rename Sheet.**
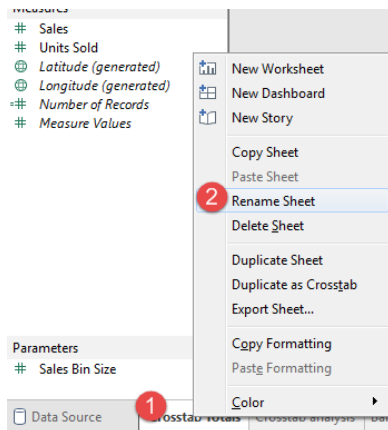
20. Rename the sheet Crosstab totals.

13 EXERCISE CROSSTAB ANALYSIS

Start a new sheet, build the below crosstab, and review the results:

1. From Dimensions drag Region to Columns.
2. From Dimensions drag Departement#(Group) to Rows.
3. From Dimensions drag Long Department Description to Rows.
4. From Measures drag Sales to Text on the Marks card.

Now we will create a Quick Table Calculation:

5. From Measures Drag and drop Sales into the Marks card under SUM(Sales).
6. Right click on the new **SUM(Sales)** in the Marks card.
7. Click on **Quick Table Calculation**.
8. Choose **Running Total.** The running total calculated field has been created (has the pyramid symbol).
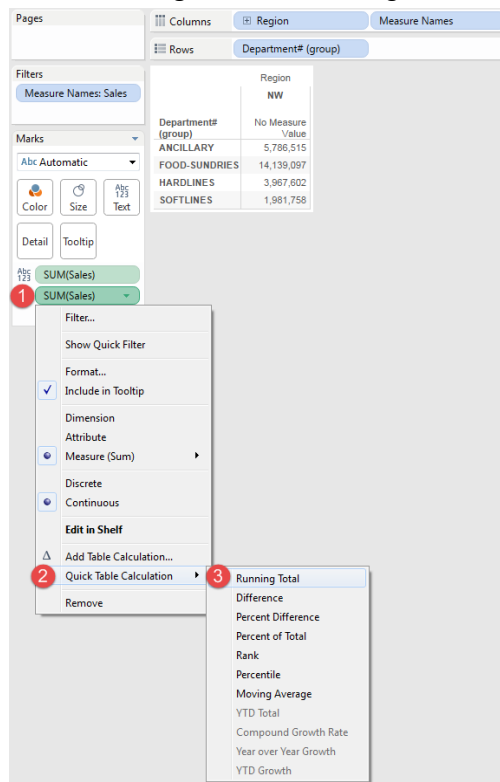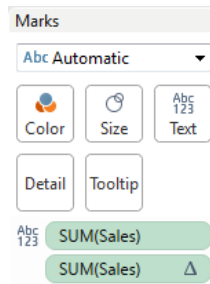
9. From the Marks card Drag & Drop the SUM(Sales) Calculated field into the worksheet next to the Sales numbers.
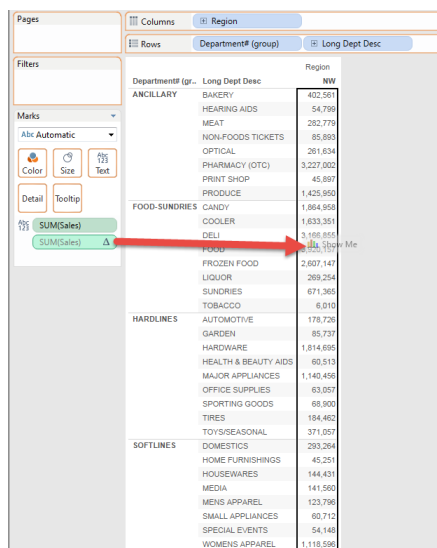
10. A new pane will appear called Measure Values. Change the order of the values by dragging Sum(sales) above the Sum(Sales) calculated field.
11. Make a copy of the calculation by holding the Ctrl key down while clicking on the new pill and drag and drop it into the Marks window.
12. Right click on the copied calculation, click Edit Table Calculation, change Calculation Type to Percent of Total.
13. Drag and drop the new pill into the view.
14. Repeat these steps 11-13 to build Quick Table Calculations for; Rank, Percentile, and Moving Average.

15. Review the numbers in your results.
16. By click on the headers you can drag and drop them to change the order. Organize the order to look like this.

17. Sort your data.
    a. Click on the header Percentile of Sales along Table(Down).
    b. Choose Sort Descending.

CHALLENGE – There are two fields in the view that can be excluded because the calculation does not make sense with the view.

18. Rename the sheet Crosstab Analysis and save your workbook.
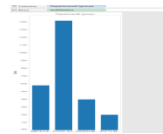
14 EXERCISE BAR CHART AND WORKING WITH THE MARKS CARD

1. Create a new sheet.
2. Drag Sales to columns.
3. Drag Department group to rows.
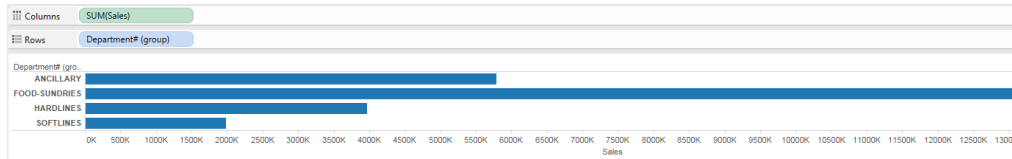
Switch the layout 2 ways!

1. Move Department#(group) to columns.
2. Move Sum(Sales) to rows.
3. Swap your columns automatically by holding down CTRL+W or simply click on the Swap symbol in the tool bar.
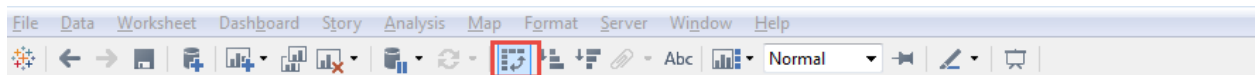

CHALLENGE -

Notice that you cannot drill into the department group. Add the Department group to your Department hierarchy so that it makes sense with your departments.

4. Once you have your Department Hierarchy set up with your Department#(Group) drill into show the department numbers.

IS – Advanced Analytics
Data Visualization Fundamentals 58

5. Review the data. While you can find the groups at the top of the worksheet the groups really don't pop out at you even with the grids.
6. Try the Marks card out to help distinguish the department groups quicker.
   a. Drag Department#(group) to Tooltip.
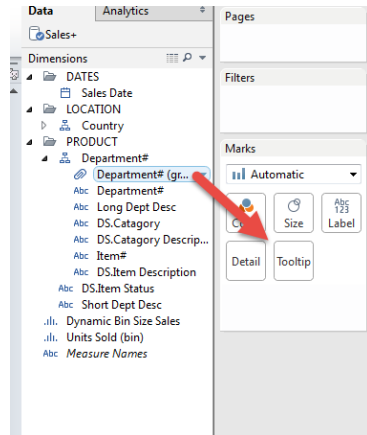
    b. This will produce an error, remove Department#(group) by dragging the pill into the blank space below the Marks card.
    c. Repeat these steps and try each Mark.
7. Rename sheet to Bar Chart and save the workbook.

15 EXERCISE CONSTANT REFERENCE LINE

1. Drag Department#(group) to Columns.
2. Drag Sales to Rows.
3. If you don't have a Bar chart change the Marks card drop down to the Bar chart Icon.
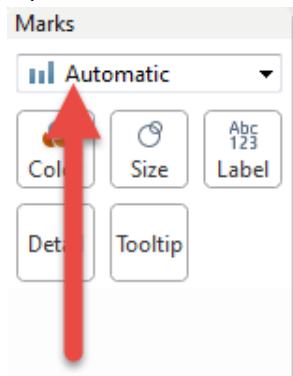
4. Right click the Sales Axis.
5. Click on Add Reference Line.
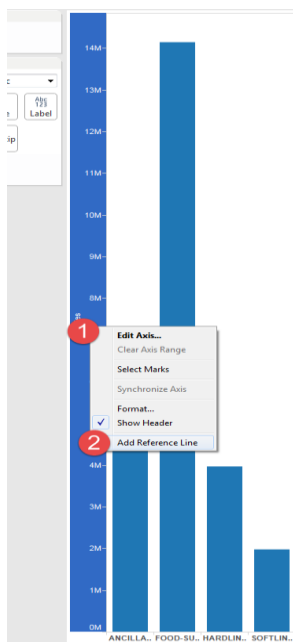
6. There is a lot in the Reference line Window for now just accept the defaults and Click OK. Adding a Reference line is as simple as right clicking into an axis and choosing Add reference line.

16 EXERCISE SIDE BY SIDE BAR CHART

1. Drag City to Columns Shelf.
2. Drag Measure names to Columns Shelf.
3. Drag Measure Values to Rows Shelf.
4. From the Measure Values card remove the Number of Records Pill.
   a. Click on number of Records pill.
   b. Drag and drop it into blank space.
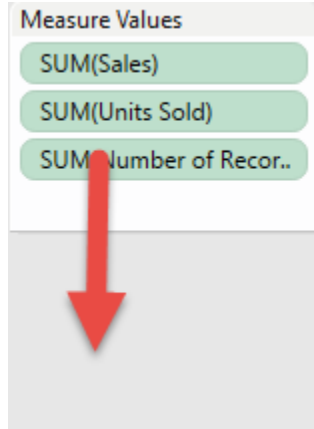
5. From the Data Window Drag and drop Measure Names to color on the marks card.
6. Rename sheet Side by Side Bar Chart and save workbook.

17 EXERCISE STACKED BAR CHART

1. Create a new sheet.
2. Drag Department#(Group) to columns.
3. Drag Sales to Rows.
4. Drag Units sold to opposite axis of Sales until dotted line appears then drop.
5. You should see a Stacked Bar chart, if not then use Marks card to select Bar chart.
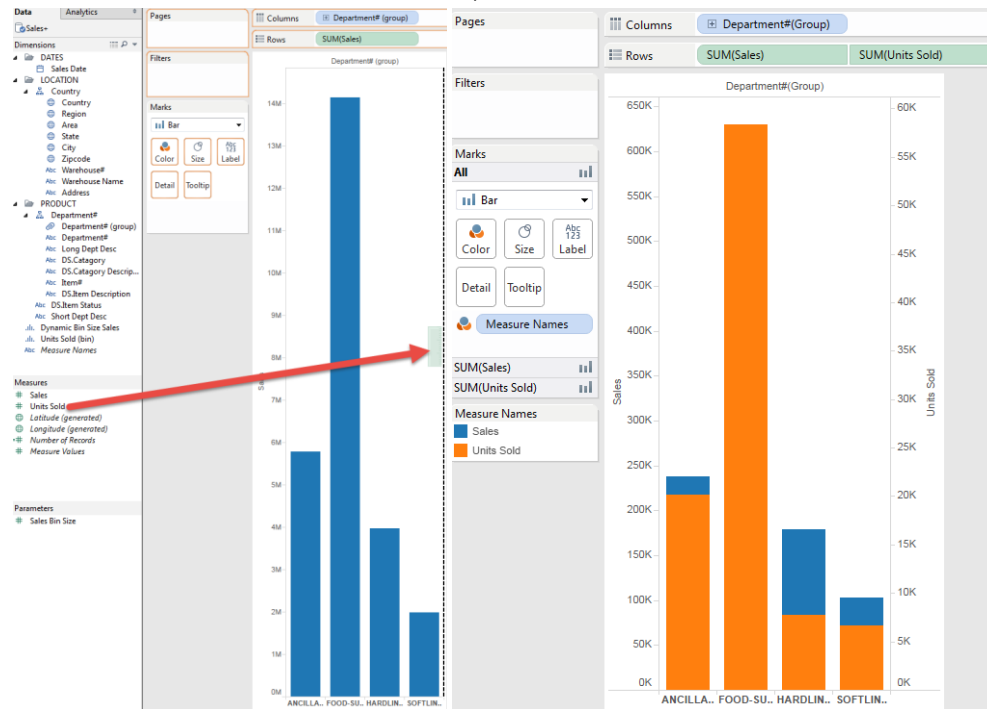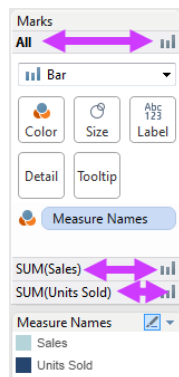
18 EXERCISE DUAL AXIS

You have already created a dual axis with the stacked bar chart.

1. Duplicate Stacked bar chart worksheet.
    a. Right click on **Stacked Bar chart** worksheet.
    b. Click on **Duplicate Sheet.**
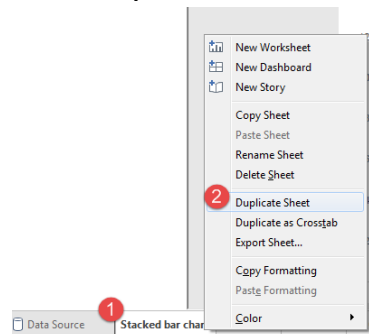
2. Review the Marks Card notice there are three Tabs : All, Sum(Sales), Sum(Units Sold).
   a. Take note of the symbol next to each of the tabs (it's the bar chart icon).

    b. When you have a dual axis, combined axis or any form of overlapping views or reports this is how you can navigate between them and edit each reports marks.
3. Click on Sum(Units Sold) tab in the marks card.

IS – Advanced Analytics
Data Visualization Fundamentals 64
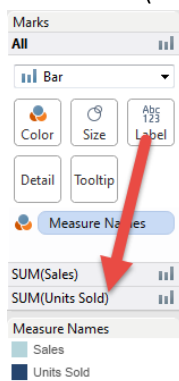
4. Use drop down arrow list box in the marks card to change the visual from a bar to a line graph for units sold.
5. If the data is not aggregated to the Department Group level, in the Columns Shelf, Click on the minus symbol next to Department#(group) to aggregate the data to the Department group level.
6. Rename sheet to Dual Axis and save workbook.
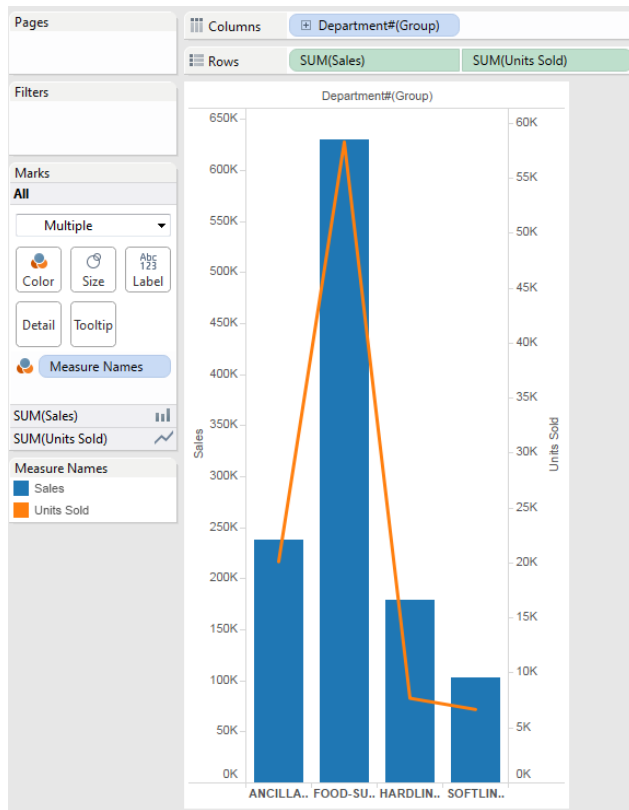
19 EXERCISE Pɪᴇ Cʜᴀʀᴛ & Wᴏʀᴋɪɴɢ ᴡɪᴛʜ ᴛᴏᴏʟᴛɪᴘs

1. Create a new sheet.
2. Drag Department#(group) to color on the Marks card.
3. From the Marks card drop down the list box and choose Pie.
4. Drag and Drop Sales onto Angle – Notice the Pie chart change.

ADD % of total sales to the tool tip

1. Drag Sales to the marks card under the current instance of Sales.
2. Right click on **Sales**.
3. Click on **Quick Table Calculation**.
4. Click on **Percent of Total**.
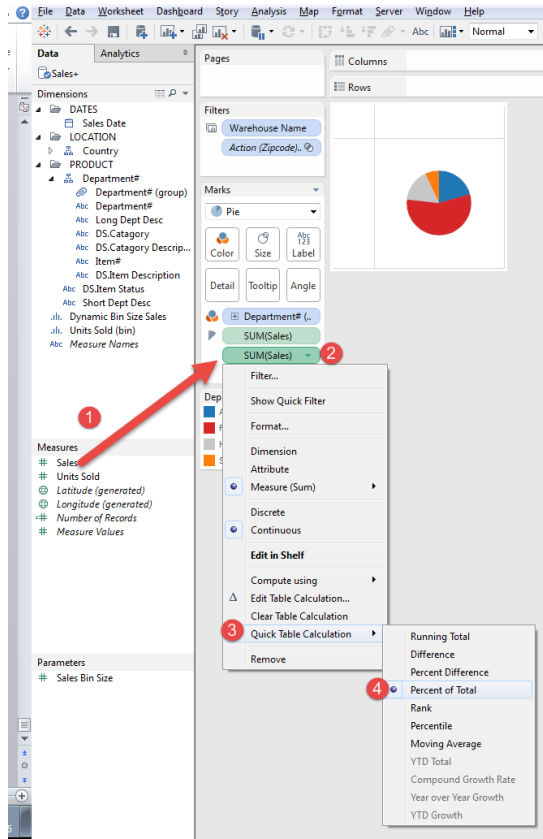
5. Hover your mouse over the Pie chart and take note of the tooltip that only displays the Department group. Wouldn't it be great if instead the tool tip looked something more like this?
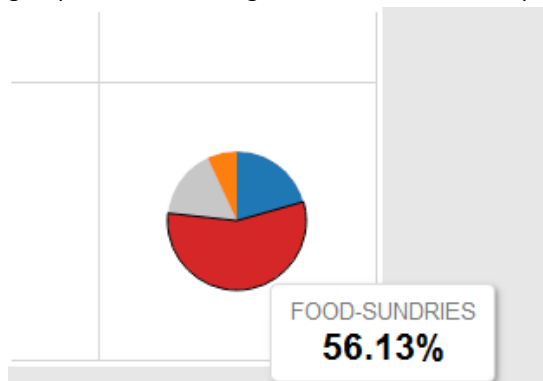

6. Click on the Tooltip Mark to Edit it.
7. Under <Department#(group)> add this script   <% of Total SUM(Sales)>.
    a. <Department#(group)> = will display the Department group name .
    b. <% of Total Sum(Sales)> = Will display the % of the total that the department takes up.
8. Adjust the font and color of these two fields then exit out of the tooltip editor.
9. Mouse over the Pie chart to view your changes.
10. Continue to Edit the tooltip until you are happy with how the information displays.
11. Rename sheet Pie Chart and save the workbook.

IS – Advanced Analytics
Data Visualization Fundamentals 66

1. Create a new sheet.
2. Drag Longitude to Columns.
3. Drag Latitude to Rows.
4. Drag Sales to Color in the Marks card.
5. Drag Units Sold to Size.
6. Adjust the size to the middle (click on Size and move the bar to the middle).
7. Drag State to Label, if you don't see anything, drag Region out of the Marks card.
8. Drill into State to the City level.
9. Rename sheet Geographic Mapping and save workbook.

21 EXERCISE FILLED MAP

1. Create a new sheet.
2. Drag Longitude to **Columns**.
3. Drag latitude to **Rows**.
4. Drag Units Sold to **Color**.
5. Drag Zipcode to the Marks Card under Units Sold, if you don't see anything, drag Region out of the Marks card.
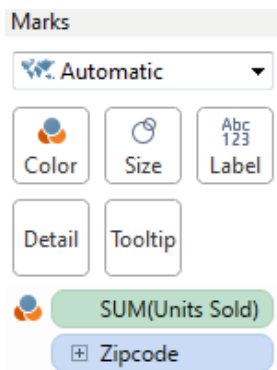
22 EXERCISE MAP WITH PIE CHARTS DUAL AXIS

1. Click the Zipcode field and then in the Show Me panel, select Filled Map.
2. Drag Sales to Color on the marks card.
3. Drag another instance of Latitude field to the rows shelf.
4. Right click the second Latitude field on the rows shelf and select Dual axis.
5. On the marks card, select the second Latitude field to bring it to the top.
6. On the drop down list of view types select Pie.
7. Drag Units to Size.
8. Drag Department to color.
9. Adjust Size to make pie charts viewable.
10. Rename sheet Map with Pie Charts Dual Axis and save workbook.

CHALLENGE from the view how would you create a group of only the warehouses in the Puget Sound Region?

## 23 EXERCISE LINE CHART AND CONTINUOUS OR DISCRETE DATES

First create a line graph to show sales overtime by Day

1. Create a new sheet.
2. Drag **Sales** to Rows.
3. Drag **Sales Date** to Columns.
4. Right click on **Sales Date** in columns.
5. Choose Day   May 8, 2015 (Continuous).
6. From the data window drag **Sales** to Color in the Marks card.
7. Rename sheet Line Chart and save workbook.
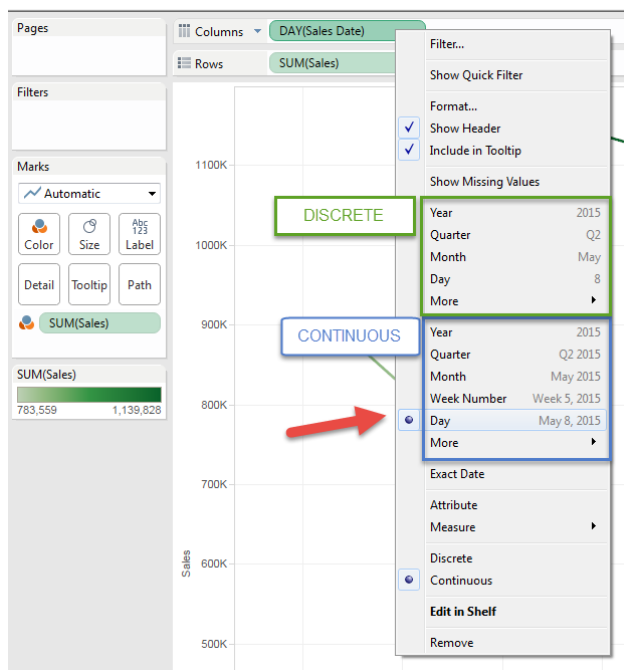
DISCRETE VS CONTINUOUS

DISCRETE DATES do not show linear time but rather as individual parts aggregated. For example if you were looking at three months worth of data and wanted to review it at a Daily level, Discrete would add up all the Mondays over the three months and show the total for all Mondays. Three months of data at a daily level would look like this.
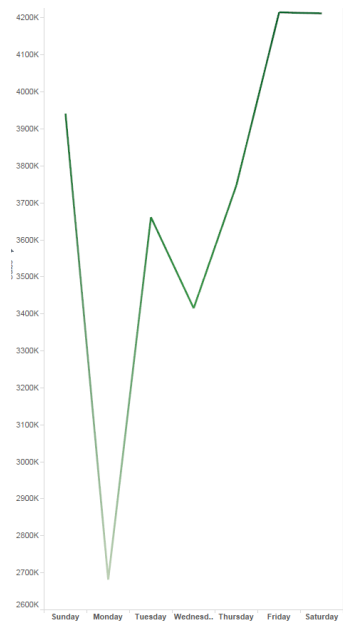
CONTINUOUS DATES are a linear progression of dates. If you wanted to see three months of data at the daily level you would not have aggregated data by day but rather each individual day in a linear progression. It would look like this
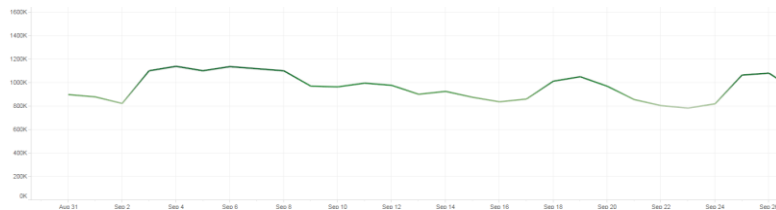
24 EXERCISE SCATTER PLOT

Scatter plots will show relationships between measurements.

There is an obvious relationship between sales dollars and the number of units sold. Let's take a look at it.
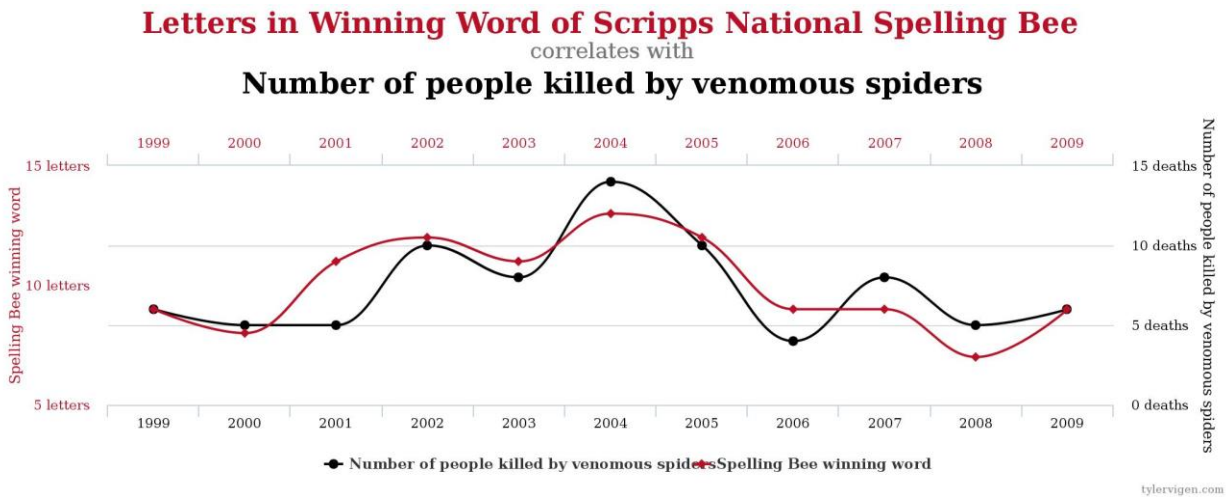
1. Create a new sheet.
2. Drag Units sold to the Columns shelf.
3. Drag Sales to the rows shelf. Every unit sold and every sales dollar has been aggregated to one number. In order to review individual units and how the sales correlate, this will need to be un-aggregated.
4. From the Analysis Menu Uncheck Aggregate measures.
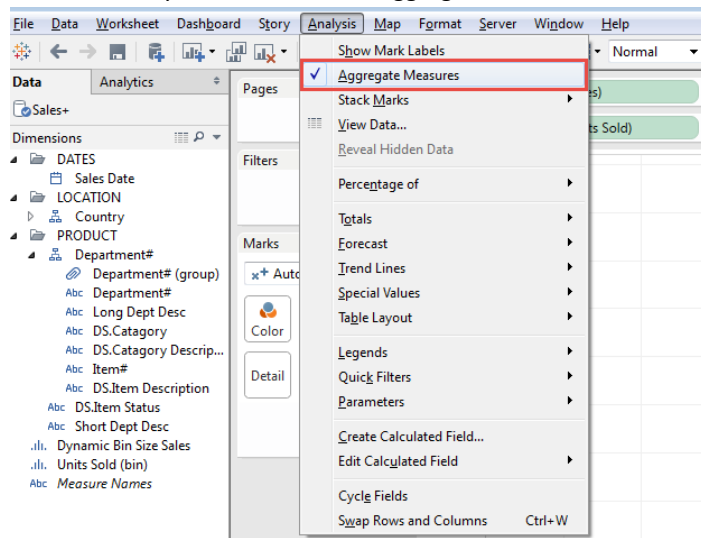
5. Mouse over each point to review the relationship between sales and units sold.
6. Rename the sheet Scatter Plot and save workbook.

Correlations are covered in more detail in the Intermediate and Advanced courses. It is important to know that not all correlations are causations. This means just because 2 things go up or down at the same time it does not mean one causes the other. A good example is venomous spider bites and spelling bees.

Letters in Winning Word of Scripps National Spelling Bee
correlates with
**Number of people killed by venomous spiders**

In this graph kids are either able to spell bigger words because of spider bites or the bigger the words that are spelled the angrier spiders get. Obviously neither of these statements are true. Yes , they both go up at the same time but one does not cause the other.

25 EXERCISE HIGHLIGHT TABLE – EDIT COLOR

1. Create a new sheet.
2. Drag Department#(Group) to Columns shelf.
3. Drag Warehouse Name to Rows shelf.
4. Drag Sales to the Colors Marks card. This should automatically build your Highlight table. If it does not, from the Marks card drop down the list box arrow and choose Square.
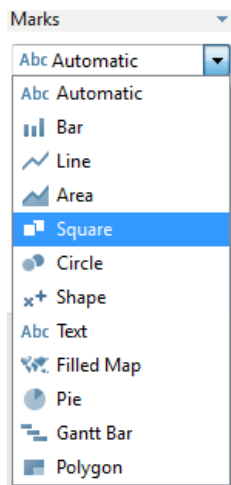
5. To format the color of the Highlight table, click on the sales legend listbox arrow and Choose **Edit Colors…**
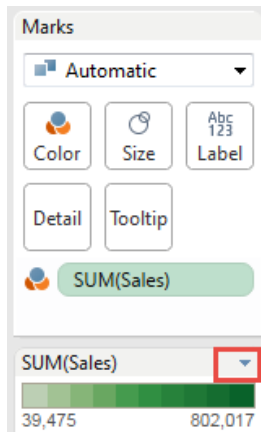
6. Click on the list box arrow.
7. Select Red-Blue Diverging.

8. Make sure **Use full color Range** is checked.
9. Click on **Apply**.
10. Click on **OK**.
11. Rename sheet Highlight Table and save workbook.
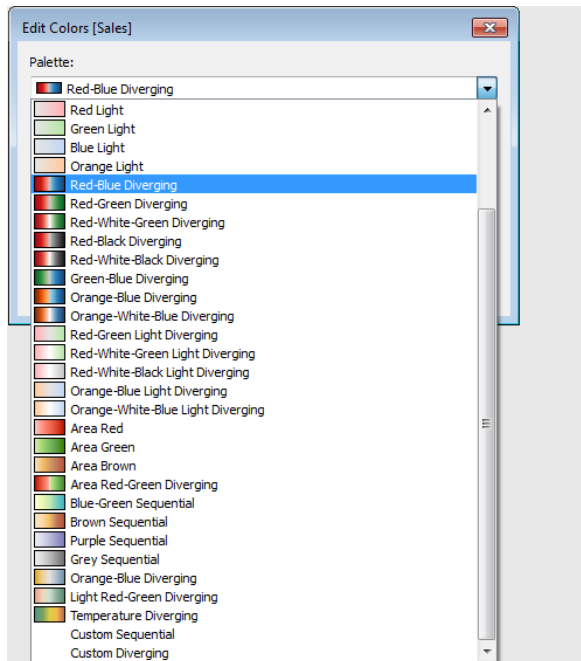
26 EXERCISE HEAT MAP – EDIT SIZE

1. Create a new sheet.
2. Drag **Department#(group)** to the Columns Shelf.
3. Drag Warehouse name to the Rows Shelf.
4. Drag Sales to Colors on the Marks card.
5. Edit Color to **Red Blue Diverging** using **Full Color Range**.
6. Drag **Units Sold** to Size on the Marks card.
7. Click on the Size Marks to display a slider.
8. Adjust the slider to change the size. The more the units sold the bigger the shape will appear in the worksheet.
9. Rename sheet Heat Map and save workbook.
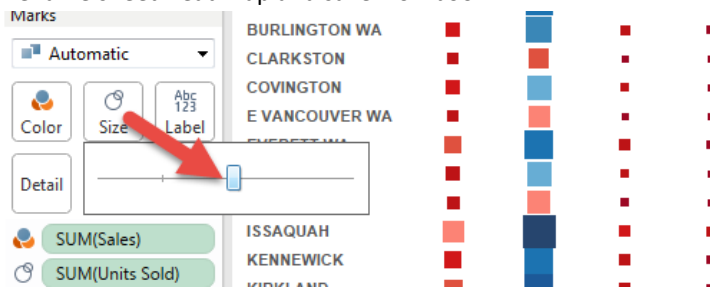
27 EXERCISE TREEMAP – SHOW ME

1. Create a new sheet.
2. Drag State to Color.
3. Drag City to Color.
4. Drag Sales to Size.
5. Drag City to Label.
6. Drag Warehouse# to Label.
7. Rename sheet Tree Map and save workbook.

Many charts can be created using **Show me** and then edited (different avenue to create a tree map).

1. Create a new sheet.
2. Click on the show me tab to drop down the list of charts.
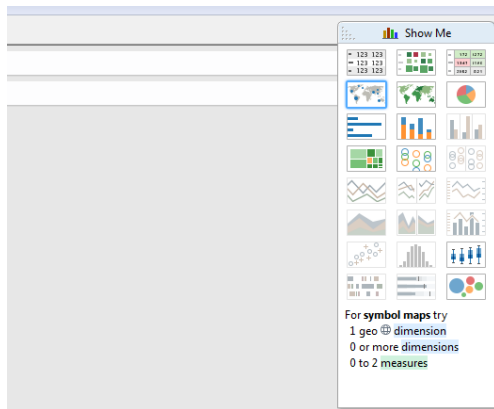
3. From the data window highlight State.
4. Hold down the Ctrl key and select Sales.
5. From the Show Me menu click on the TreeMap Icon.

1. Create a new sheet.
2. Drag Department#(Group) to Color. If Warning window appears, click Add all members.
3. Drag Units Sold to Size.
4. Drag Item description to Label.
5. Edit the tool tip.
    a. Remove everything except **<DS.Item Description> & <SUM(Units Sold)>.**
    b. Change the font size of <SUM(Units Sold)> to 8 and change its color to light gray.
    c. Type Units after the <SUM(Units Sold>tag.
    d. Center all text, then click OK.
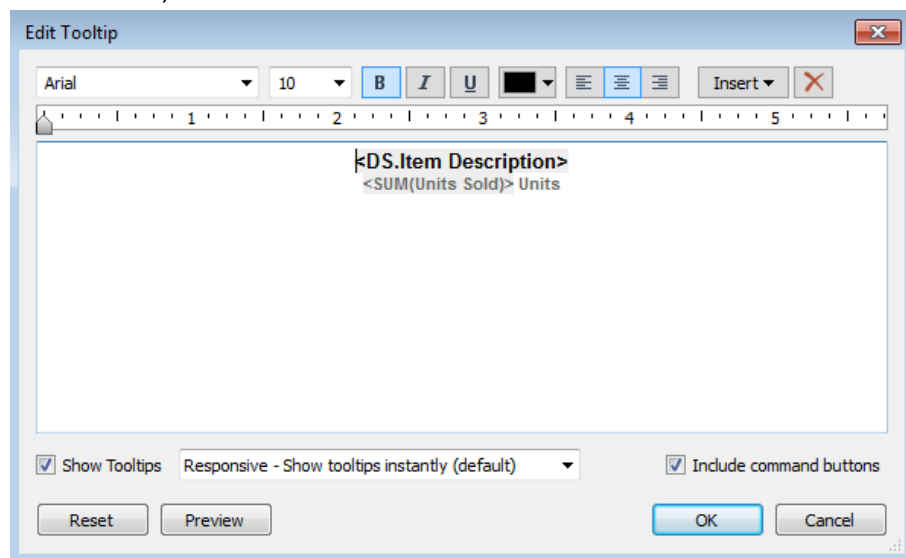
6. Go to top of Marks card, select drop down menu, select circles or text, to see the difference.
7. Mouse over the Word Bubbles to review data.
8. Rename sheet Word Bubbles and save workbook.

29 EXERCISE MOTION CHART

1. Create a new sheet.
2. Drag Sales Date to Columns and Choose Continuous Day.
3. Drag Sales to Rows.
4. Drag Department#(Group) to Color.
5. Drag Units Sold to Size.
6. Drag Sales Date to Pages.
7. Rename sheet Motion Chart and save workbook.

30 EXERCISE DASHBOARD 1

Click on the new dashboard icon.

1. Drag Crosstab Analysis onto dashboard.
2. Click on the white space and notice the border of the crosstab analysis.
3. Drag bar chart to the whitespace of the dashboard.
4. Drag another Chart into the dashboard.
5. Remove all charts you have added to the dashboard.
     a. Click on chart.
     b. Click on **X**.
6. Click on **Floating**
7. Drag and drop Crosstab Analysis into the Dashboard
8. Resize the Cross tab to take up over half of the dashboard.
9. Click the drop down arrow
10. Click **Fit**
11. Click **Entire View**
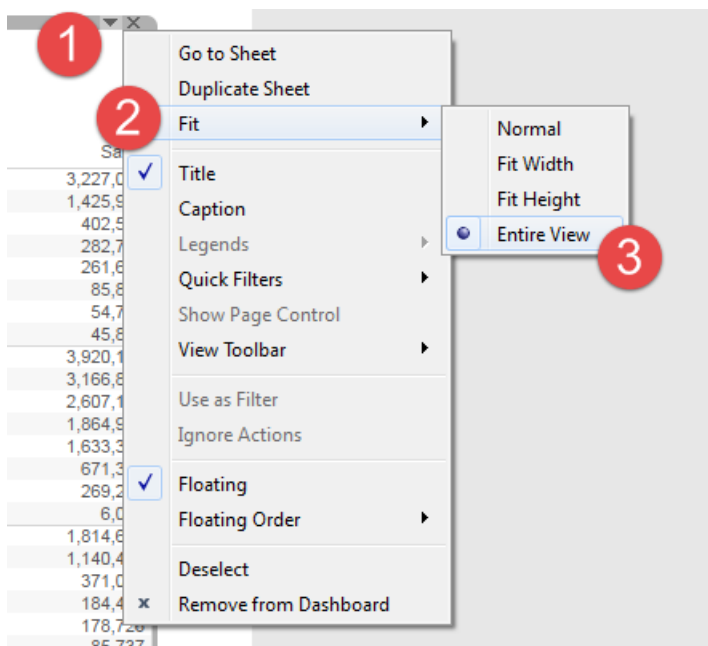
12. Drag Text Box to the Dashboard
13. Type "WHAT'S THE STORY?"
14. Click on Center icon to center the text
15. Drag text box to center of dashboard above the crosstab

IS – Advanced Analytics
Data Visualization Fundamentals 80

16. Save the workbook.

31 EXERCISE DASHBOARD 2

1. Create a new dashboard
2. Click on **Floating**
3. Drag Pie chart into the Dashboard workspace
4. Fit to Entire view
5. Uncheck title
6. On Legends click Drop down arrow and uncheck Show Title

7. Drag Motion chart into the worksheet
8. Fit to Entire view
9. Remove Motion chart legend
10. Filters and legends will Cascade
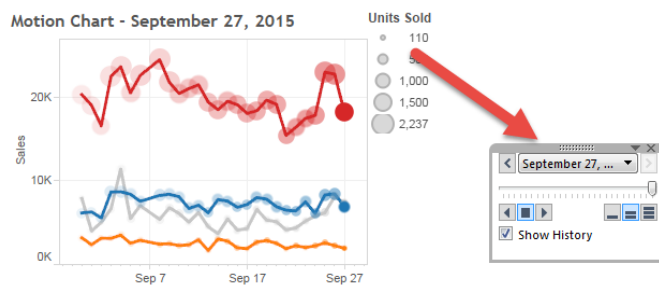
11. Separate the cascading legend and Pages filter
12. Remove the Legend for Units sold be clicking on it then clicking on the X
13. Fit to Entire view
14. Double click on Title Replace <Sheet Name> with Sales $
15. Drag Groups and barchart into the worksheet
16. Adjust the size and position of the charts to look something like the example below
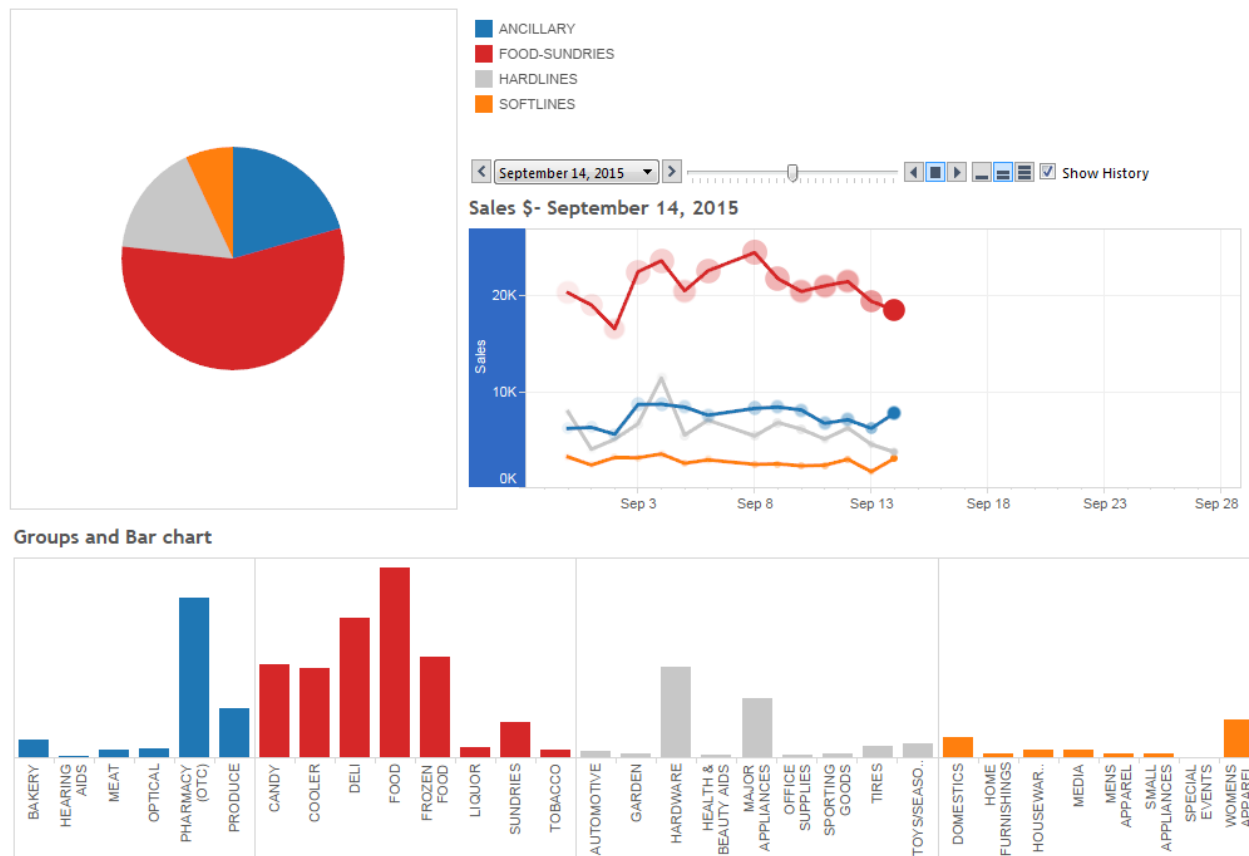
Looks great, but what are you looking at? What is the story? What decisions can be made from this?

Exploring the charts and graphs you made and just putting them in a dashboard will help you see relationships between your data and this is another step in discovery. Here you are a reporter or a detective and you are looking for a lead and then following it. The lead I see here is Food and Sundries is a big part of our business. Now I want to know. How big it is and where people are shopping.

32 EXERCISE DASHBOARD 3

After creating a few preliminary dashboards and discovery a story you want to explore, look through each of your worksheets to determine which worksheets would best help tell your story. The ones I chose I thought told the best story of where members are shopping in the warehouse. You may find different ones you like.

ADD THE HIGHLIGHT TABLE

1. Click through each of the worksheets you have created.
2. Start a new dashboard
3. Using floating Drag and drop the Highlight table into the view
4. Drag the highlight table to the lower right hand corner and resize to take up a larger portion of the view.
5. Remove Title
6. Drag legend to the top of the view

ADD THE CONSTANT REFERENCE LINE

1. Drag and drop the constant reference line Bar chart into the dashboard
2. Move the constant reference line to the upper left hand corner
3. Remove Title
4. Fit to entire View

ADD THE WORD BUBBLE CHART

1. Drag and drop the Word Bubble chart to the remaining white space.
2. Remove title
3. Fit to entire view
4. Click on the view drop down arrow
5. Click on Quick Filters (or Filters depending on the version)
6. Choose Department#(group)
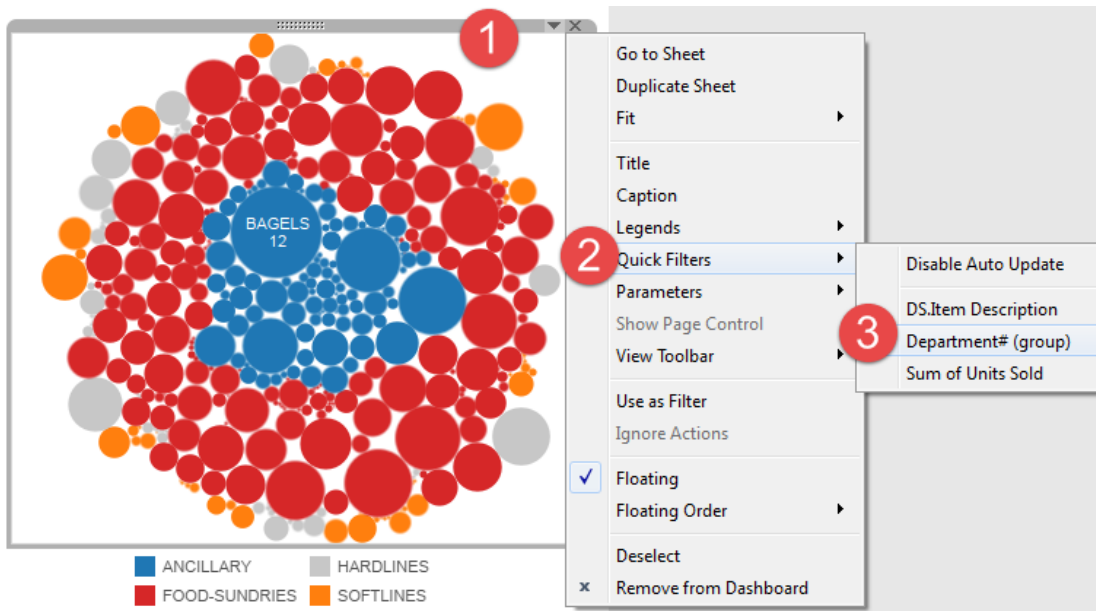
7. Adjust Filter size to display all of the Departments
8. Change the Department group filter to apply to all worksheets
   a. Click on list box arrow for filter
   b. Click on **Apply to Worksheets**
   c. Click on **All Using This Data Source**
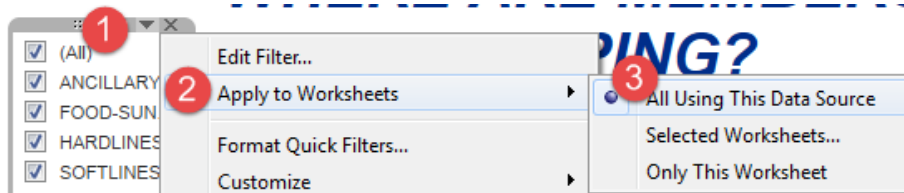
HEADERS AND LOGOS

1. Drag a text box into the view
2. Type "Where are Members Shopping?"
3. Change the font size to 26
4. Adjust the text box to fit the title on one line
5. Add Company Logo or image
   a. Any images you want to add to your Dashboard should be stored in C:\Documents\My Tableau Repository\Shapes\My Image
   b. With Floating checked, Drag the **Image** object from the **Dashboard** window to the Dashboard space
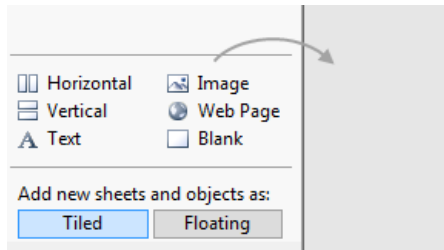
c. Navigate to the image you want to use, select it and click **Open**.

d. Resize the image and move to upper Left side of the screen
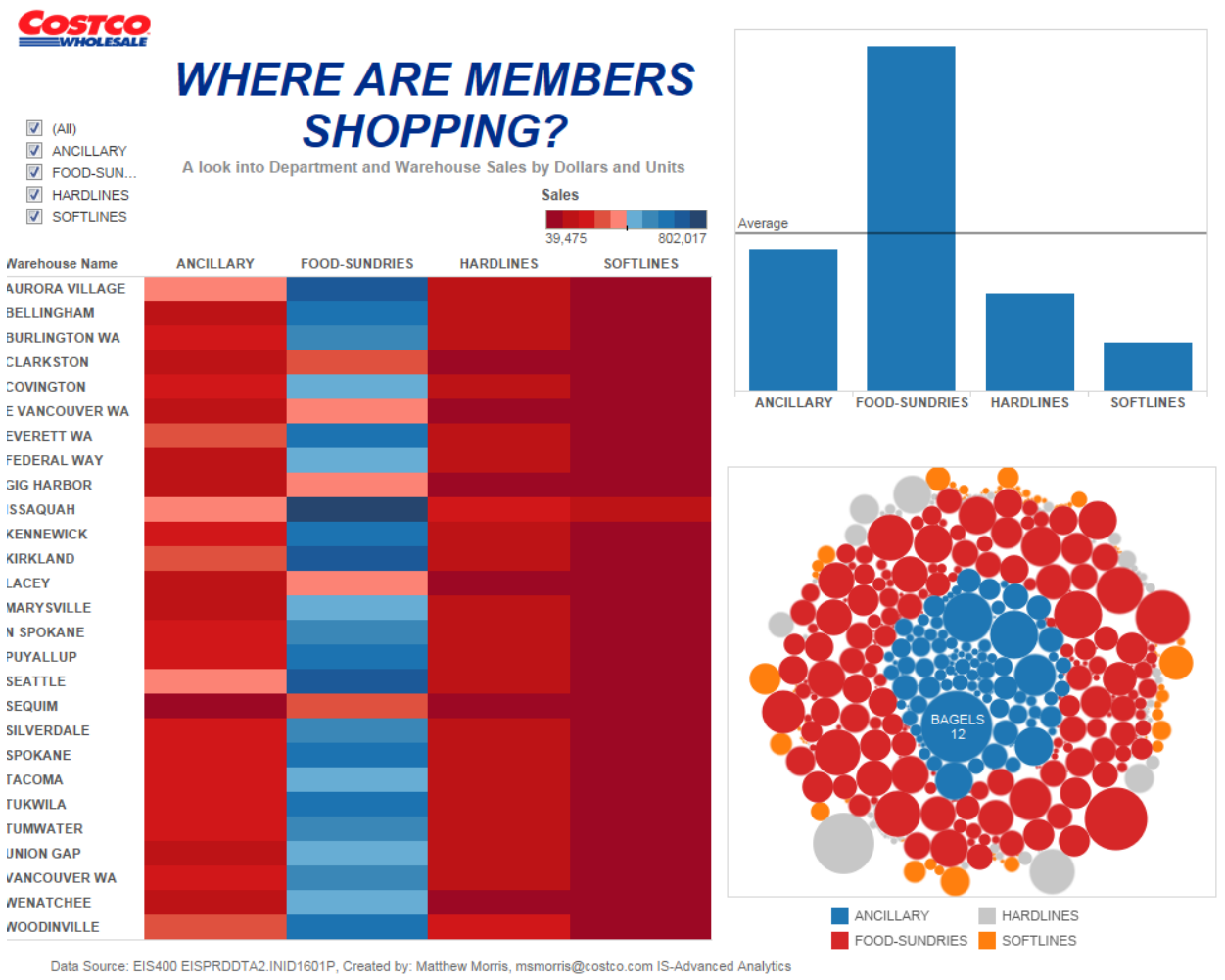
33 EXERCISE CREATE YOUR OWN DASHBOARD

Create your own dashboard using; Filled map, Pie chart, Motion chart and Grouped locations to show throughout the state how members distributed in departments.

CHALLENGE: As you create your dashboards, you may notice some confusion with colors. In some, Red indicates a category while in others, Red indicates low sales. Let's say for example you are limited on your color scheme because of company colors. How can you work with a color that has a double meaning (category & measurement)?

34 EXERCISE STORY

Stories are a way to show a linear narrative using charts and dashboards. The benefit of this is you can connect facts, provide context, show cause and effect or build a case.

These directions show you how to navigate creating a story but you will need to create your own.

In the example we have been working with the story that answers the question; "Where are members shopping?" Using Dashboards or worksheets you have created, develop captions that tell a story.

1. Click the New Story tab.

2. In the lower-left corner of the screen, choose a size for your story. Choose from one of the predefined sizes, or set a custom size:
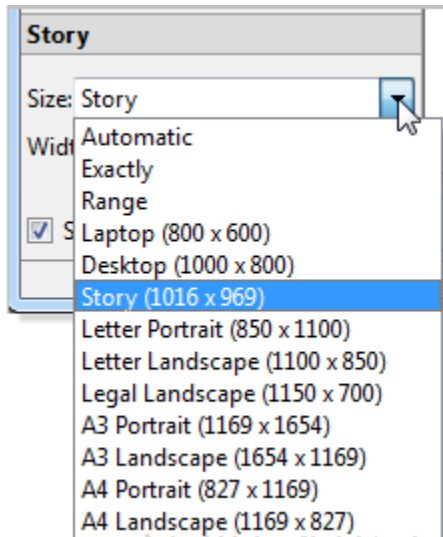
3. To add a title to your story, double-click **Story Title** to open the Edit Title dialog box. You can type your title in the dialog box, as well as choose a font, color, and alignment. Click **Apply** to view your changes.

4. Drag a sheet from the **Dashboards and Worksheets** area to the story, and drop it into the center of the view:

5. Click **Add a Caption** to summarize the story point.

Story Title

Add a caption

Once you have added 1 story point you will have the option to add another caption by choosing New Blank point or Duplicate. This will open a new Story pane.

New Blank Point

Duplicate
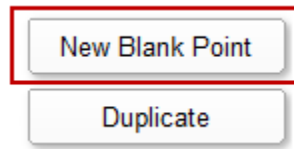
PUBLISH

This section is from Tableau.com

DECISIONS YOU MAKE DURING PUBLISHING

This section discusses choices you'll be given as you step through the publishing process. Take these into consideration before you initiate the process.

Sign in to Tableau Server or Tableau Online

To publish, you need to sign in to Tableau Server or Tableau Online. Make sure you know how your administrator set up authentication on your server or site. For more information, see Sign in to Tableau Server or Online.

Add your workbook or data source to a project

When you publish, you select a project in which you want to put your workbook or data source. Projects serve as containers for workbooks or data sources. Server or content administrators can use them to organize related content, as well as to set permissions on the workbooks and data sources within them. Before you start the publishing project, make sure you know which project you want to publish to.

Set permissions

As the owner of the published workbook, you can specify who can access it from the server and what they can do with it. Alternatively, in many environments, you can use the default permissions, and your administrator will manage content permissions on the server or site. For more information, see Setting Permissions.

Show or hide sheets when you publish

When you publish a workbook, you can specify which sheets to include. Showing or hiding sheets is useful when you want to publish a dashboard or story without publishing the worksheets that were used to create that dashboard or story.

Important: Hiding sheets is not a security measure. Anyone who has the **Download/Web Save As** permission can access the hidden sheets by opening the workbook on the server or downloading it from the server. Other editing permissions also can allow access to hidden sheets. For more information, see Control Access to Published Content in the Tableau Server Help.

Show sheets as tabs

When you publish a workbook that contains multiple sheets, you can specify how users navigate the published workbook.

- If you selected multiple sheets under **Sheets to Publish** and want to provide tab-based navigation across multiple sheets, select this check box.

- If you want to hide any sheets in the workbook, make sure the **Show Sheets as Tabs** check box is cleared.

  When you select this check box, all sheets are shown, even if you set sheet-level permissions.

Include external files when you publish

If your views contain any information that isn't available to the server, or to other users who access the workbook on the server, include the files that contain that information when you publish.

For example, you might use an Excel, CSV, or other data source local to your computer; image files; and so on.

If you are publishing to Tableau Online, and the workbook connects to a data source that Tableau Online cannot connect to directly, select the **Include External Files** check box.

For more information, see Keep Data Fresh in the Tableau Online Help.

Refreshing web data connector extracts

When you publish a workbook with a web data connector data source, you need to *import* the web data connector to the server before you can set up a refresh schedule. You can do this only on Tableau Server. For information, see Web Data Connectors in Tableau Server in the Tableau Server Help.

To use web data connector data sources on Tableau Online, you need to create a local extract and use it with the sync client to refresh the published data source. For information, see Schedule Refreshes Using the Sync Client in the Tableau Online Help.

PUBLISHING STEPS

1. Select **Server** > **Publish Workbook**.

   If you are not already signed in to a server, the **Tableau Server Sign In** dialog box appears.

2. Complete the sign-in process, as described earlier in Sign in to Tableau Server or Tableau Online.

   To sign in to Tableau Online, enter **https://online.tableau.com**.

3. In the **Publish Workbook to Tableau Server** dialog box, specify the following:

   o **Project**: Select the name of the project you want to publish into. The default project on Tableau Server is named Default.

   o **Name**: Provide a name for the workbook in the **Name** text box.

      Use the drop-down list to select an existing workbook to overwrite. To do this, you must have the appropriate permissions on the server.

   o **Description**: Provide a description that will help you and others know what analysis the workbook contains.

   o **Add Tags**: Enter keywords that describe the workbook. Tags help users find related workbooks when they browse the server.

      Separate tags using either a comma or a space. To add a tag that contains a space, put it in quotation marks (e.g., "Sales Quotes").

4. ou can specify permissions to allow or deny access to the workbook on the server.

   By default all users can interact with the workbook. As the publisher, you have all permissions.

5. Under **Sheets to Publish**, select the sheets you want to include in the published workbook.

6. Under **Options**, select the check boxes as appropriate:

   o **Show sheets as tabs** if you selected multiple sheets to publish and want to provide tab-based navigation.

   o **Show Selections** if you want selections you've made in the workbook to be published to the server.

   o **Include External Files** if your views contain any information from files that will not be available to other users who access the workbook on the server.

7. (Optional) Click **Authentication** (or **Scheduling & Authentication** to specify how users can access the data. For more information, see Embed Credentials to Streamline Signing In to a Published Workbook.

8. (Optional) If the workbook uses extract connections, you can add the workbook to a refresh schedule. For more information, see Refreshing Data Inside Workbooks.

9. Click **Publish**.

   Note: If you are publishing to a site that saves a revision history of workbooks, workbooks published with identical names will be saved as a revisions of the same workbook.

DATA VISUALIZATION PROJECT CHECKLIST	DEVELOPER	PUBLISHER
Research Data		
Define Fields		
Defines Filters		
Define Calculations		
Define KPI's		
6Ws.		
Connect to Data		
Clean and filter data before connecting		
Filter unnecessary data		
Rename Fields		
Assign correct data properties to fields		
Join new table		
repeat steps for each newly joined table		
Prepare Data Pane		
Creat Folders		
Create Hierarchies		
Create Groups		
Data Discovery(Create additional groups as needed)		
Crosstab Totals		
Crosstab Analysis		
Bar chart		
Side by Side Bar Chart		
Stacked Bar chart		
Pie chart		
Geo Map		
Line chart		
Scatter Plot		
Highlight Table		
Heat Map		
Tree Map		
Bubble Word Cloud		
Motion Chart		
Create Story		
Review 6 Ws		
Choose chart or charts that best answer 1 of the 6w's		
Add Filters and Interactivity		
Add Headline		
Organize Dashboard to visually be easy to understand		
Publish Dashboard to your site		