



PROGRAMMES

FACULTY & RESEARCH

EVENTS

FOR ORGANISATIONS

Home > CEE > Open Enrollment Programmes > Certificate Programme In Business Analytics

"Certificate Programme in Business Analytics ranked No.1."

Retains top spot in the rankings by Analytics Vidhya, a knowledge portal for analytics.

[Know more >](#)

1 2 3 4

CBA PROGRAMME

INDUSTRY CONNECT

EXPLORING ANALYTICS

**Certificate
Programme in
Business
Analytics >**

PROGRAMME
OVERVIEW

ADMISSIONS

FACULTY

INDUSTRY
ENGAGEMENT

ALUMNI

CBA PROGRAMME



ANALYTICS IS THE use of modern data mining, pattern matching, data visualisation and predictive modelling tools to produce analyses and algorithms that help businesses make better decisions.

ChandraSekhar Katuri

According to a recent Wall Street Journal, companies, barraged with data from the Web and

The CBA is revamped as AMPBA from Class of 2020 (Winter)

[To access AMPBA website, click here.](#)

Quick Links

[CONTACT US](#)

Infographics

Certificate of Business Analytics



Data Visualization

Indian School of Business, Hyderabad Campus

June 12th, 2019

Sudhir Wadhwa,

Masters in Mathematics

Masters in Business Administration

sudhir.wadhwa@gmail.com

skwadhwa@scu.edu

Sudhir_Wadhwa@isb.edu

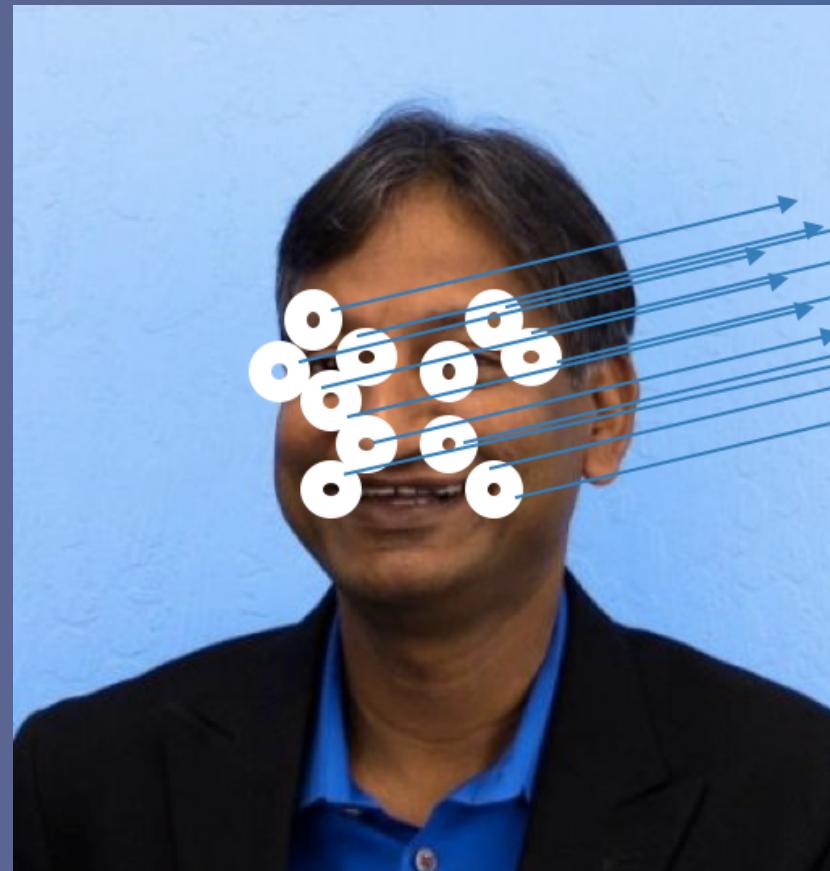
Teaching Assistants (TA)

Sadaf & Soumodip



Slides & Videos are not for Distribution. All slides are Intellectual property (IP) of DataTiles group & Tableau. Under Digital rights management (DRM) law, You are given access to these slides as Students of Indian School of Business. Please do not publish it on Social Media. Please do not Email / distribute.

Who is teaching this class?



CTO, Viridis
Learning

Machine learning advisor to
SanDisk, El Camino Hospital,
Xenous, Timeli, Reflektion,
and Kaidan group

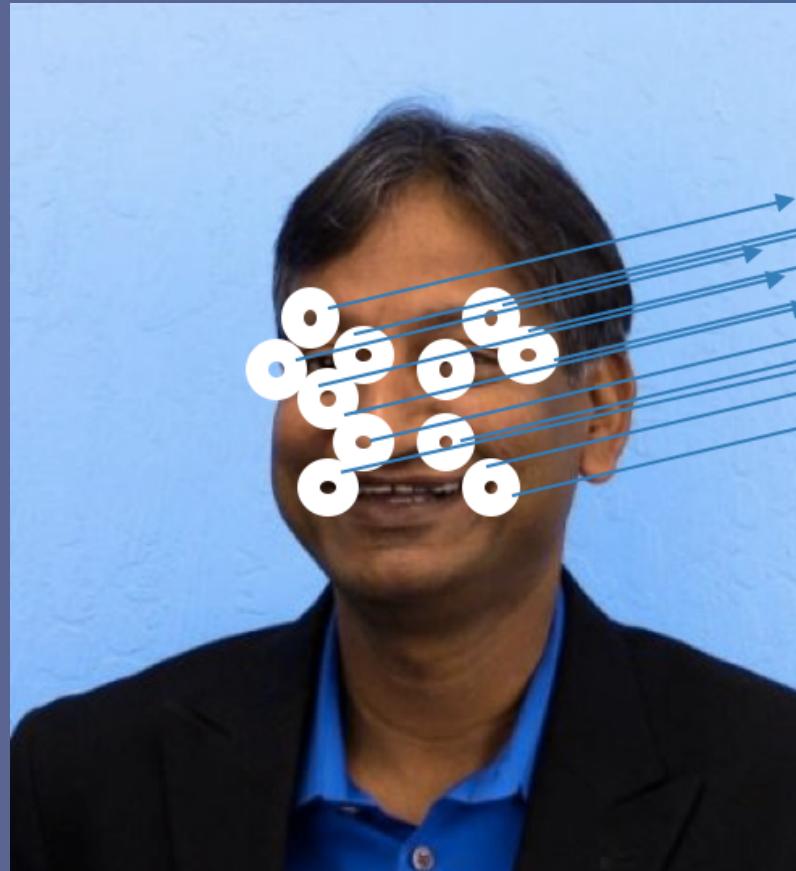
Countries lived and worked in:
USA, Germany, and Japan

Fastest way to get to me



WhatsApp : 408 398 0655

Gmail: sudhir.Wadhwa@gmail.com



D

Date	Session	Section B (AC3 LT)	Section A (AC7 LT)
12 th June	Session 1	7:45 AM to 9:45 AM	10:15 AM to 12:15 PM
13 th June	Session 2	7:45 AM to 9:45 AM	10:15 AM to 12:15 PM
14 th June	Session 3	7:45 AM to 9:45 AM	10:15 AM to 12:15 PM
15 th June	Session 4	7:45 AM to 9:45 AM	10:15 AM to 12:15 PM
16 th June	Session 5	7:45 AM to 9:45 AM	10:15 AM to 12:15 PM

ALL STUDENTS SUBMIT SURVEY



12th June

12th June

13th June

13th June

13th June

14th June

14th June

14th June

15th June

15th June

15th June

15th June

15th June

15th June

16th June

Introductions

Pipeline, Slides 10 to Slides 61

Slide 64 to Slide 80 – Showing end to end Salesforce example, BeautifulSoup

Start with Class 8 – Salesforce TWBX

Class3 STUDENTS INFORMATION SYSTEM

Class 1b – Top Charts – 182 to 211

Touch Coinbase – Class 2 BTC example - Class 4a - show 3 movies and crontab movie

Slides 130 to 202

Slides 289 to 309

Class 22 – Time Series,

Class11 R and Tableau, Slide 362

Class 9B – Geo Spatial Data – South African Bridges example

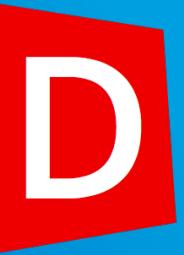
Class20 Supply Chain Hotel Example complete Product

Class 21 – Pygal , Flask Application

CLASS Presentation in front of judges

SUBMIT ALL YOUR PROJECTS FOR GRADING

ALL STUDENTS SUBMIT SURVEY



14th June

14th June

14th June

BTC Example complete

Class 9B – Geo Spatial Data – South African Bridges example

Working with PDF

15th June

Class 11 R and Tableau, Slide 362

15th June

Class 20 Supply Chain Hotel Example complete Product

15th June

Class 21 – Pygal , Flask Application

16th June

CLASS Presentation in front of judges

SUBMIT ALL YOUR PROJECTS FOR GRADING

ALL STUDENTS SUBMIT SURVEY



15th June

15th June

15th June

Working with PDF

Class 11 R and Tableau, Slide 362

Class 21 – Pygal , Flask Application

16th June

CLASS Presentation in front of judges

SUBMIT ALL YOUR PROJECTS FOR GRADING



- Welcome and Introduction of Sudhir
- Structure of the class
- Grading Policy
- Missing class Policy
- Logistics – Email address, My Office number
- Students will introduce at this point
- What do I expect in Students' Introduction (Show them the slide)
- Introduction to Data Science Pipeline and what is your role



Grades

- Group Project, 40 points
- Individual Project, 20 points
- Virtual Project , 10 points
- Team Presentation / Judges, 30 points



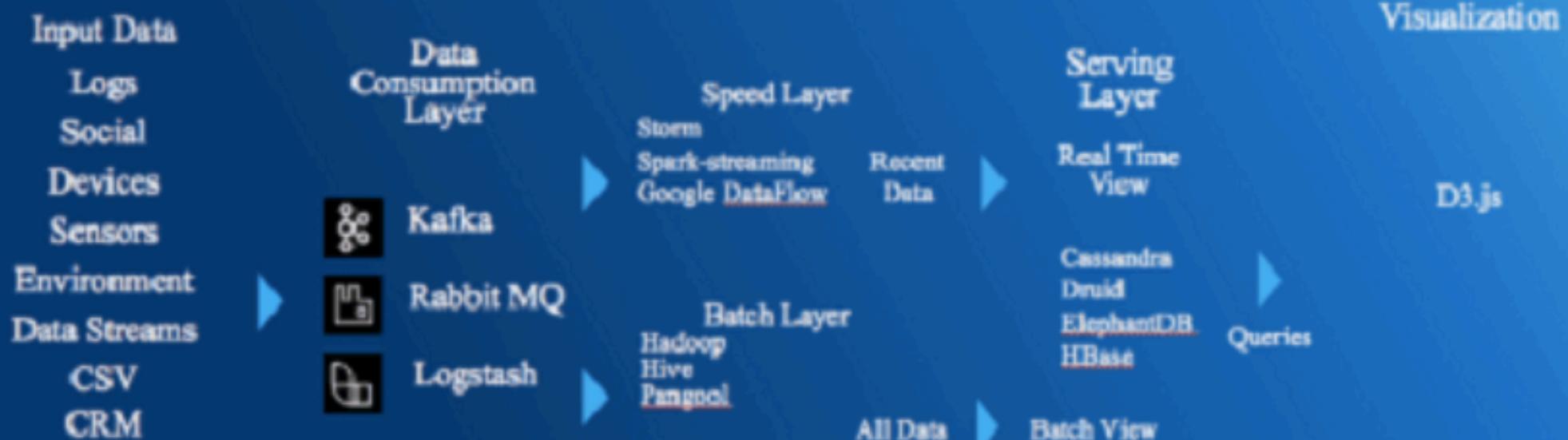
• Introductions

- Your Name
- Educational Experience
- Company Name (last company or current company)
- Current or last Project you completed
- Projects that interest you

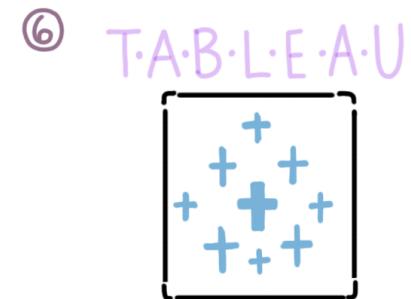
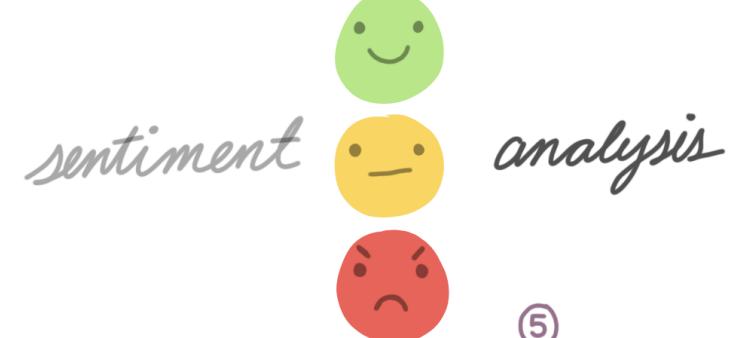
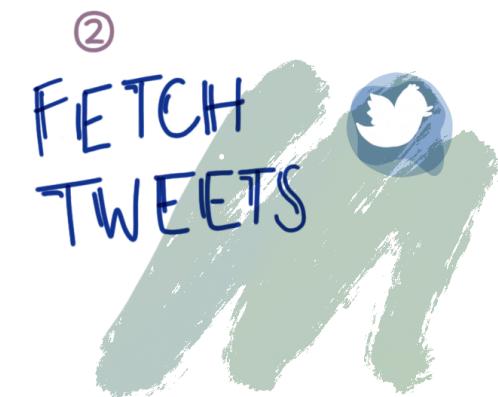
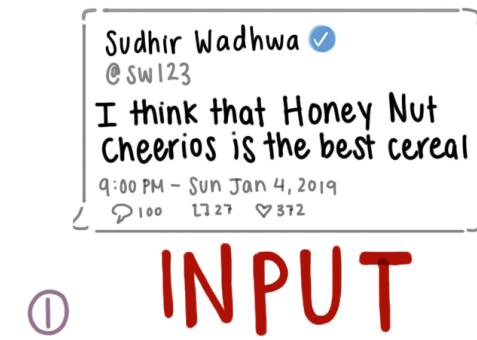
- Industry
- What kind of data is interesting from your domain (for example Employer Data in my case)
- How do you use the Data currently? Ingest?
- What are the Key Performance Indicators for your domain?

- Something unique about you (e.g. hobby)

Data Science Pipeline



© DataTiles.io, [DataTiles.ai](#)



int

Slides are not for Distribution





Power of Visualization



Let's look at some data

I		II		III		IV	
x	y	x	y	x	y	x	y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89



Let's look at some data

It turns out that these four data sets all have the same means, the same variances, the same x-y correlations, and even boil down to an identical linear regression.

In fact, it looks like these four data sets are practically identically based on this analysis.

I		II		III		IV	
x	y	x	y	x	y	x	y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89



Let's look at some data...with statistical analysis

I		II		III		IV	
x	y	x	y	x	y	x	y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89
9	--	9	--	9	--	9	--
11	--	11	--	11	--	11	--
--	7.5	--	7.5	--	7.5	--	7.5
--	4.122		4.122		4.122		4.122
0.816	0.816	0.816	0.816	0.816	0.816	0.816	0.816
$y = 3 + 0.5x$							

Mean of x
Variance of x
Mean of y
Variance of y
Correlation between x and y
Linear regression line

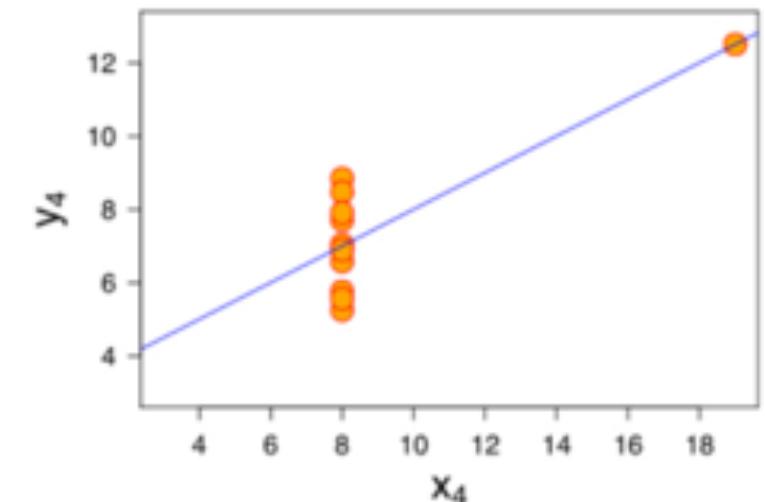
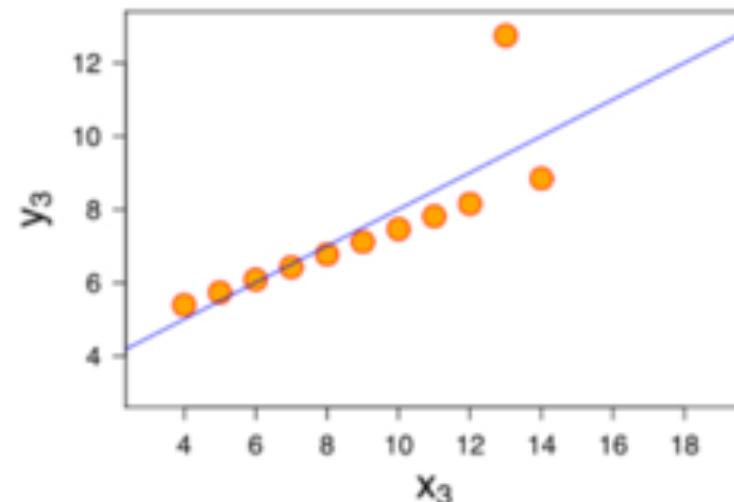
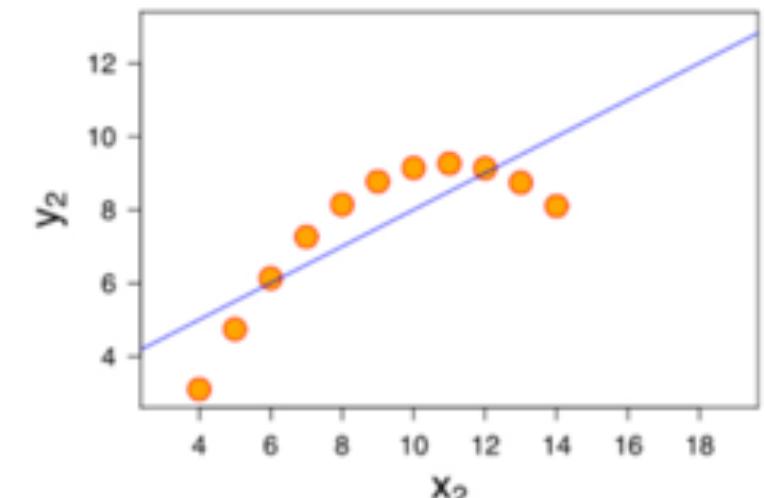
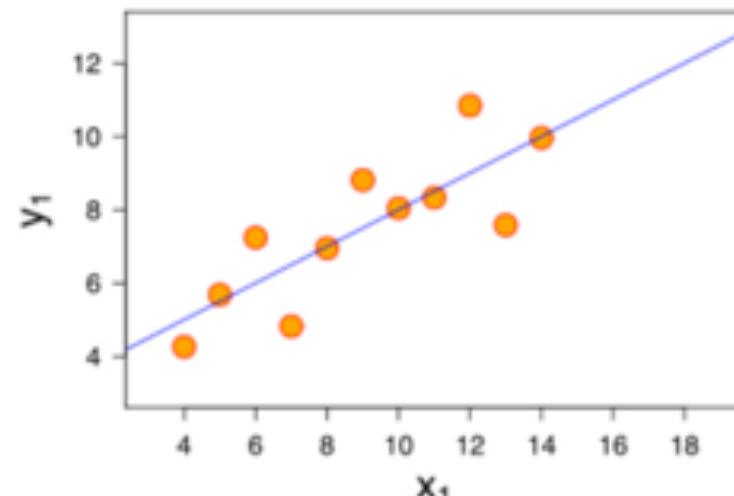
Let's look at some data...visually

Here are these same four data sets, plotted visually, with trend lines.

Anscombe's Quartet was created in 1973 by Francis Anscombe, a statistician. He developed this to show the importance of graphing data, as opposed to just applying statistical tests to analyze information.

This example is one of the basic reasons that visualization is so important – it helps us identify trends within our data that may be obfuscated by statistical tests, or would take us more time to realize.

"Anscombe's Quartet"



Source: Wikipedia



Why visualize?

The first step in understanding your data should always be to examine it visually. Visualization can play a critical role in helping you figure out what the interesting questions are.

Remember the picture superiority effect: pictures are retained at much higher rates than words.



Why visualize?

A good visualization reduces the time to insight. Here's a simple example:

*You receive a report of monthly revenues by product and region.
Which were your five highest-revenue product/regions last month?*

	Product A	Product B	Product C	Product D	Product E	Product F	Product G	Product H
Northeast	\$ 15,749.00	\$ 40,195.00	\$ 15,472.00	\$ 63,029.00	\$ 8,509.00	\$ 42,987.00	\$ 27,778.00	\$ 12,510.00
Southeast	\$ 48,044.00	\$ 20,741.00	\$ 40,643.00	\$ 15,687.00	\$ 12,342.00	\$ 23,297.00	\$ 10,401.00	\$ 10,522.00
Central	\$ 5,240.00	\$ 45,296.00	\$ 16,114.00	\$ 63,359.00	\$ 58,198.00	\$ 24,191.00	\$ 46,826.00	\$ 50,278.00
Northwest	\$ 12,860.00	\$ 11,548.00	\$ 43,134.00	\$ 19,331.00	\$ 60,563.00	\$ 51,475.00	\$ 28,954.00	\$ 14,405.00
Southwest	\$ 37,087.00	\$ 61,506.00	\$ 54,084.00	\$ 14,715.00	\$ 17,811.00	\$ 32,814.00	\$ 47,853.00	\$ 44,639.00



Why visualize?

Let's add color:

	Product A	Product B	Product C	Product D	Product E	Product F	Product G	Product H
Northeast	\$ 15,749.00	\$ 40,195.00	\$ 15,472.00	\$ 63,029.00	\$ 8,509.00	\$ 42,987.00	\$ 27,778.00	\$ 12,510.00
Southeast	\$ 48,044.00	\$ 20,741.00	\$ 40,643.00	\$ 15,687.00	\$ 12,342.00	\$ 23,297.00	\$ 10,401.00	\$ 10,522.00
Central	\$ 5,240.00	\$ 45,296.00	\$ 16,114.00	\$ 63,359.00	\$ 58,198.00	\$ 24,191.00	\$ 46,826.00	\$ 50,278.00
Northwest	\$ 12,860.00	\$ 11,548.00	\$ 43,134.00	\$ 19,331.00	\$ 60,563.00	\$ 51,475.00	\$ 28,954.00	\$ 14,405.00
Southwest	\$ 37,087.00	\$ 61,506.00	\$ 54,084.00	\$ 14,715.00	\$ 17,811.00	\$ 32,814.00	\$ 47,853.00	\$ 44,639.00

Color automatically focused your brain – no mental calculations/comparisons needed!

How many nines are there?



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

How many nines are there?



(There are 11 nines)

It's amazing how much easier it is to find and see the nines as soon as we use Color

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



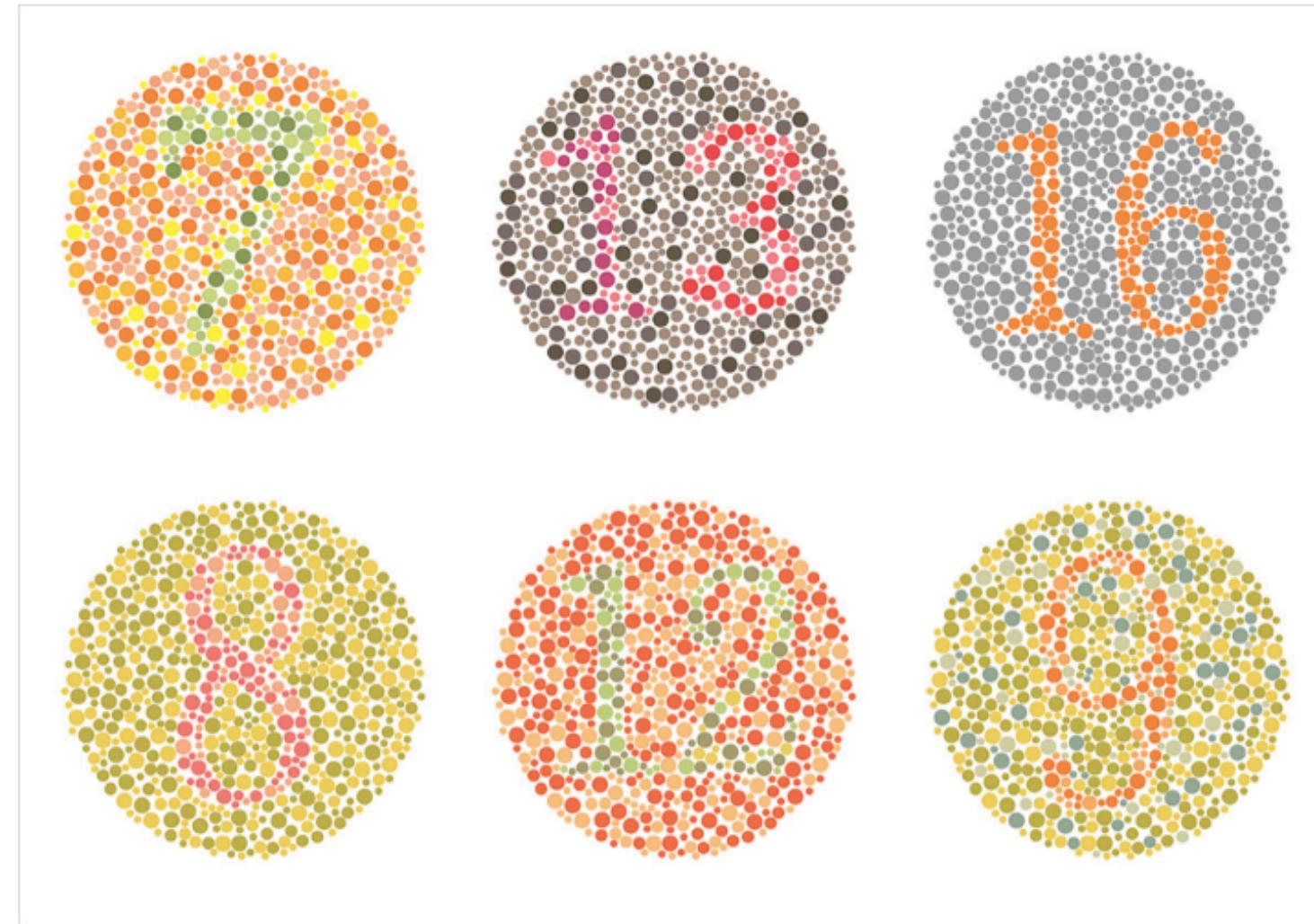
Using Color

Studies have shown that around 8% of men and 0.5% of women have color vision deficiency (CVD). This is more commonly referred to as colorblindness

- Avoid red/green palettes!
- Blue/orange is a good alternative.

Using Color

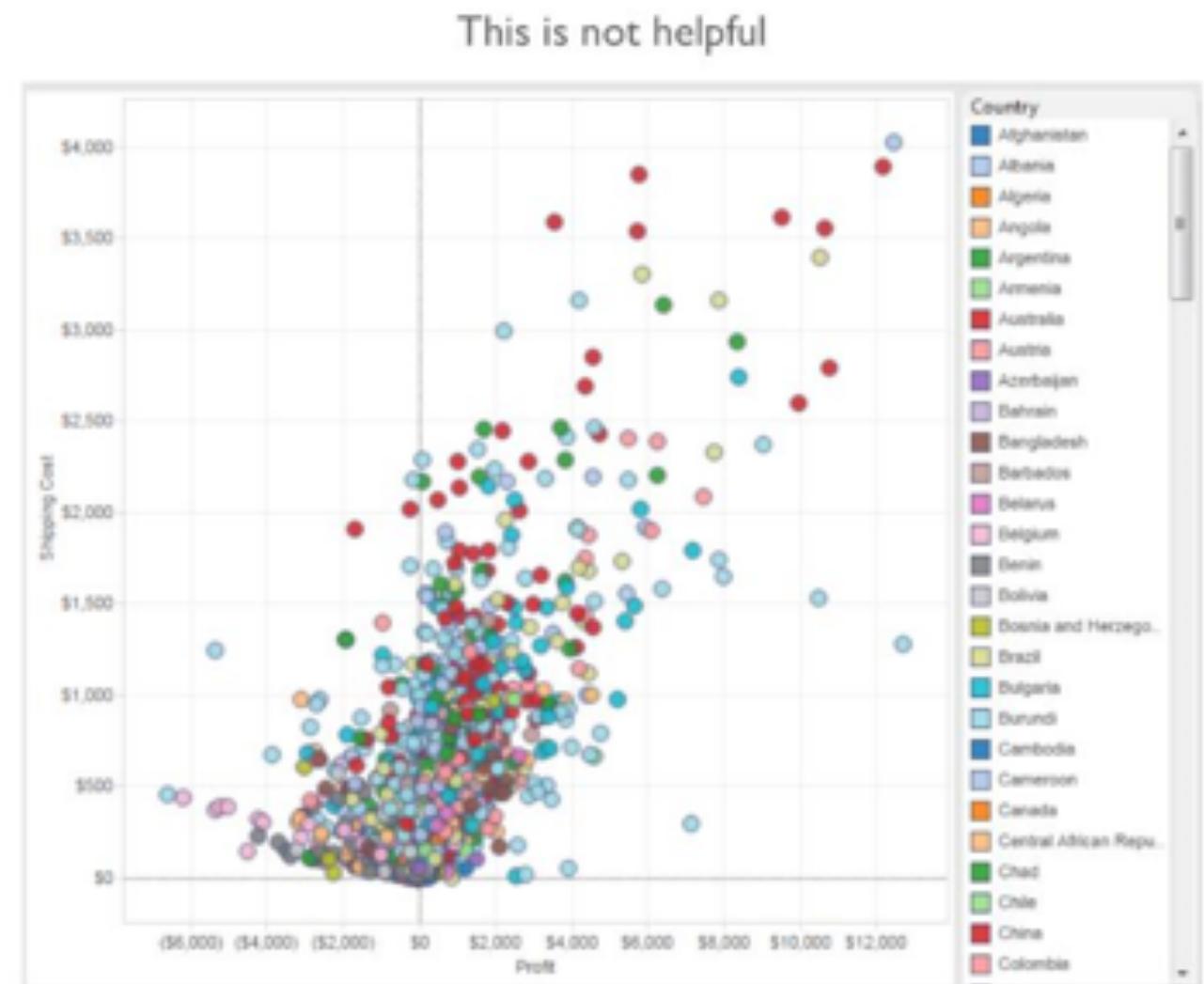
Avoid red/green palettes!



Using Color

Humans can only distinguish ~8 colors at a time

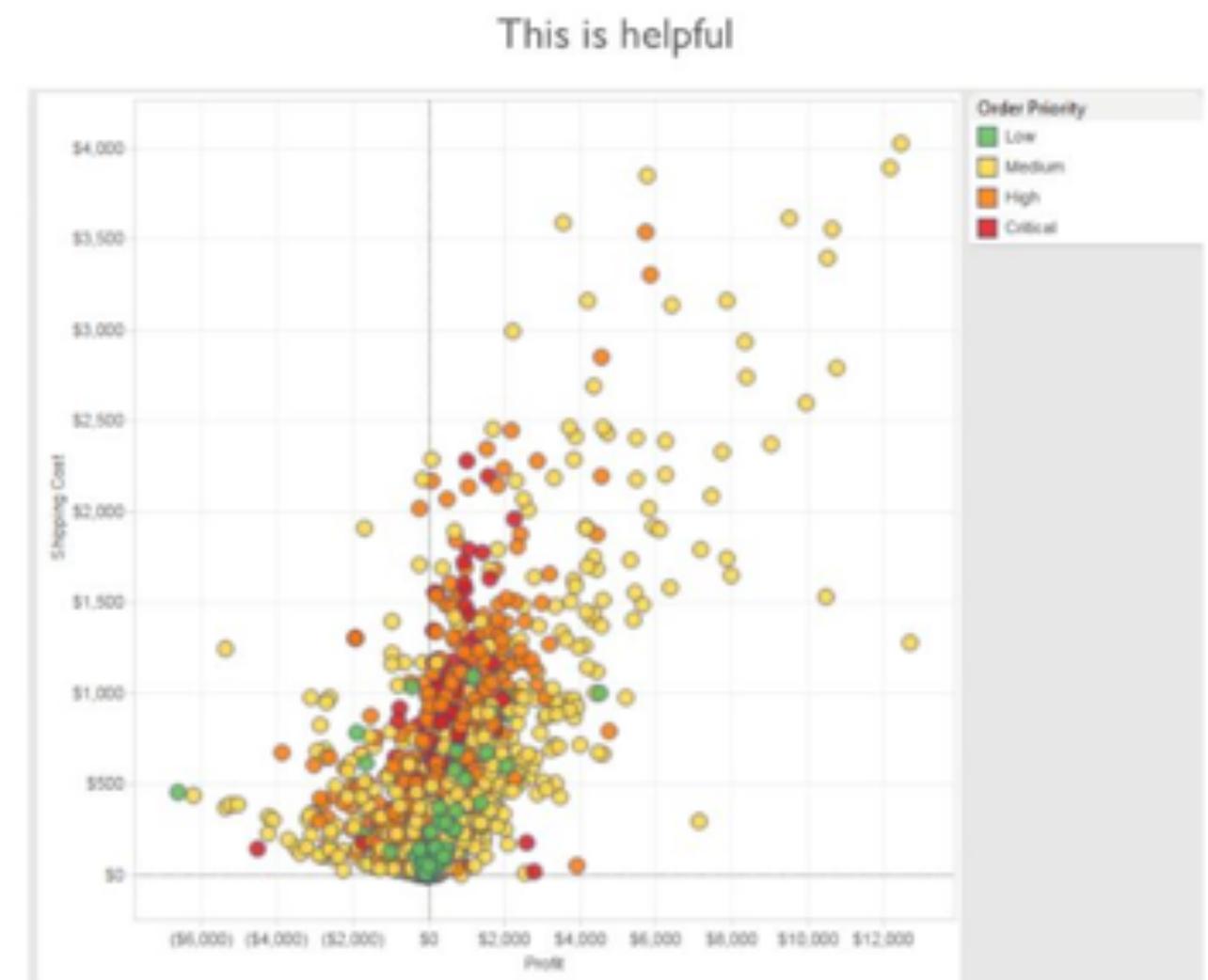
You have about 150 different data points and you're mapping them to color but this is bad because I can't tell if the blue is Albania or Cameroon



Using Color

Humans can only distinguish ~8 colors at a time

This is much better



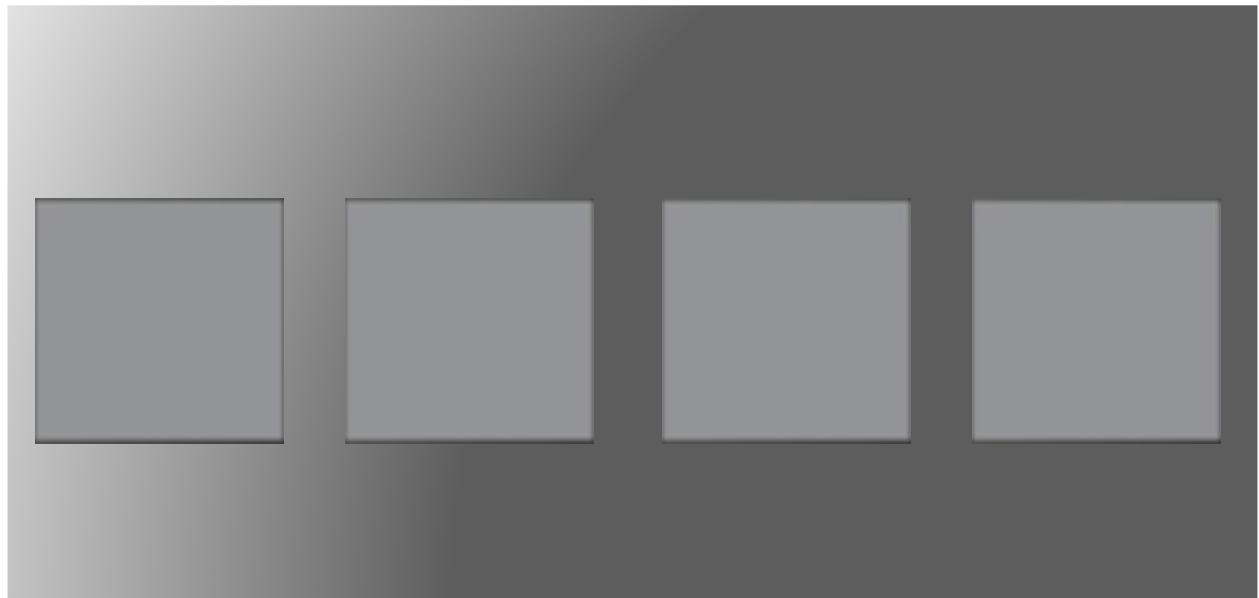


Color Me Impressed

Color perception is relative, not absolute

Color is relative, not absolute.

Which of these squares is lightest? If we're trying to encode meaning into color, we have to pay attention to how the viewer will perceive that color in context





Color Me Impressed

Color perception is relative, not absolute

They're all the same color but we didn't know that until we had a consistent background

When using color – provide your users with a consistent background so they can determine differences and non-differences. Knowing they'll sense a difference, help them to only perceive it if it's actually there.

It's not enough to just use color, be mindful of visual best practices to make the story accurate (that color in context)

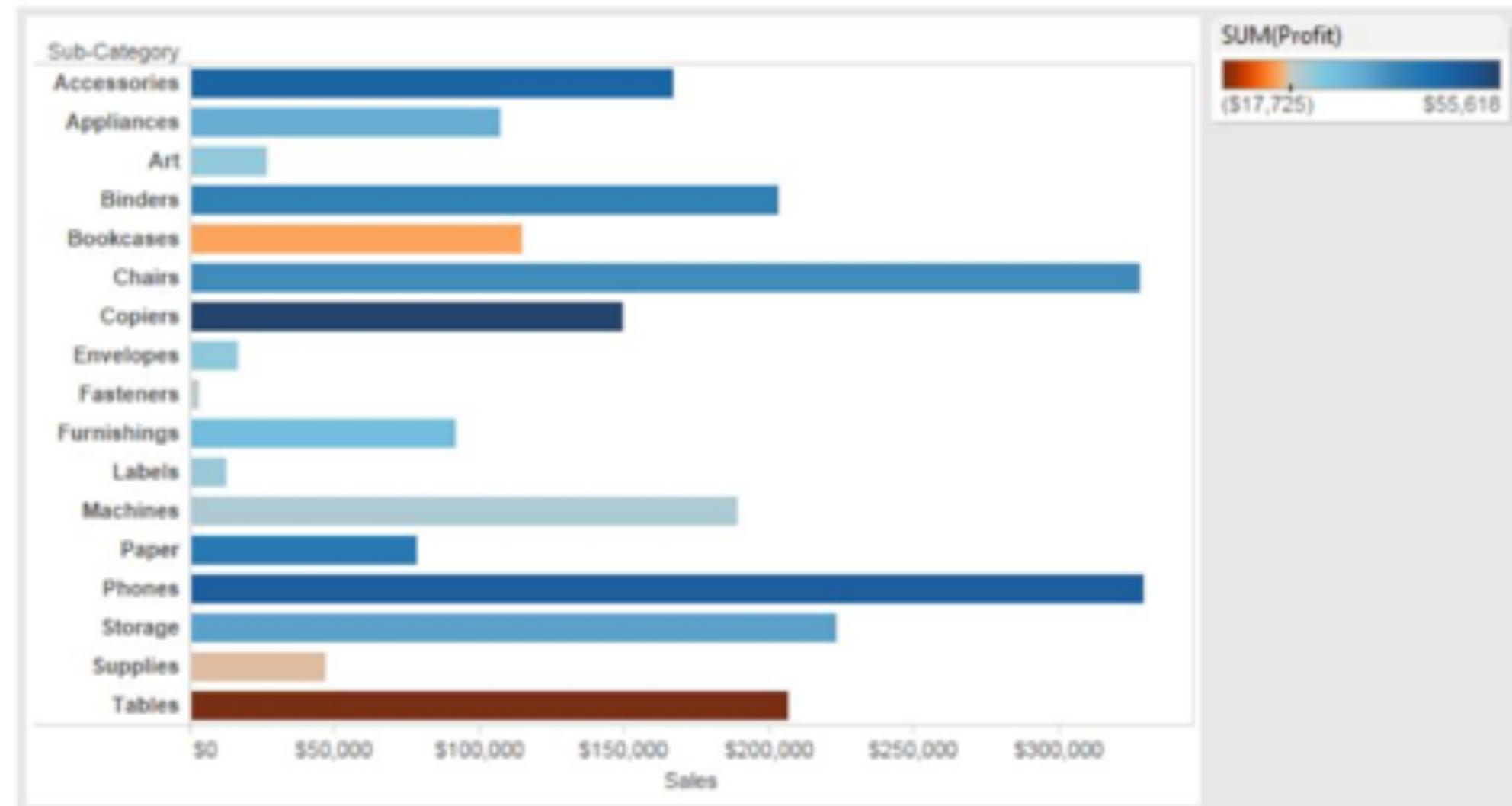


Color Me Impressed

For quantitative data, color intensity and diverging palettes work well

Using Orange and Blue
is color-blind friendly

Color helps user see
Bookcases are loosing
money for us





Visualization Best Practices

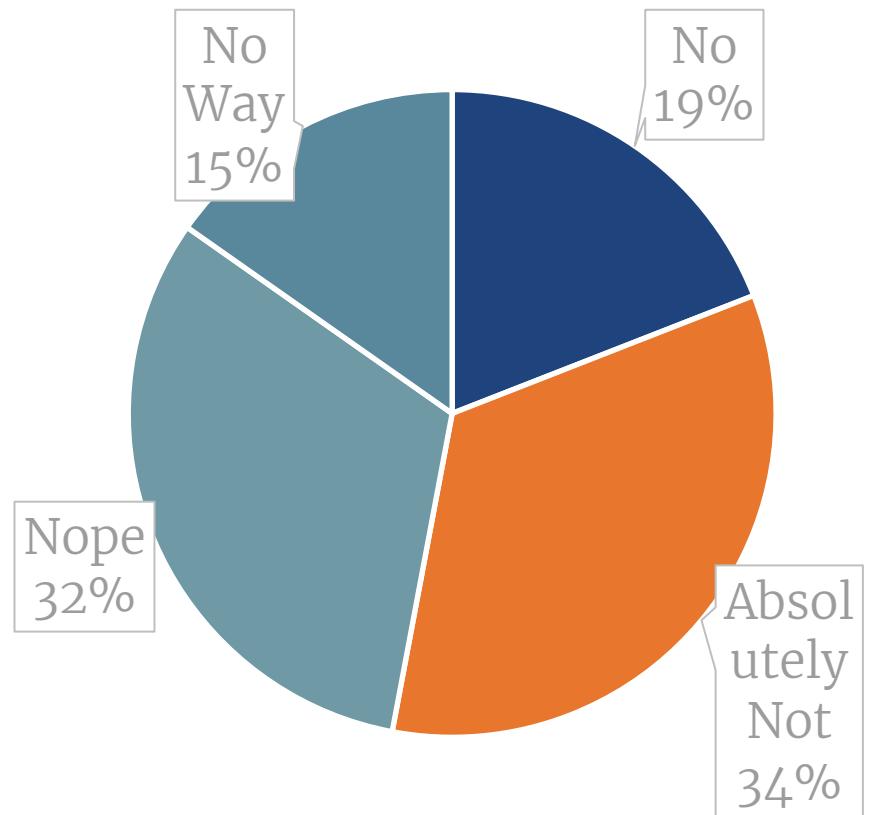


Your brain on autopilot (Kahneman's “System 1”)

Evryneoe can raed txet, eevn if the letetrs
are scrmeblad. The olny tnihg taht mtetras
is taht the fsrit and lsat leertts are in the
rgiht pcales. Tihs is one of the mnay enctixig
tgnihs you wlil laern in tihs cslas!

Pie Charts – NO!

Is a pie chart a good idea?





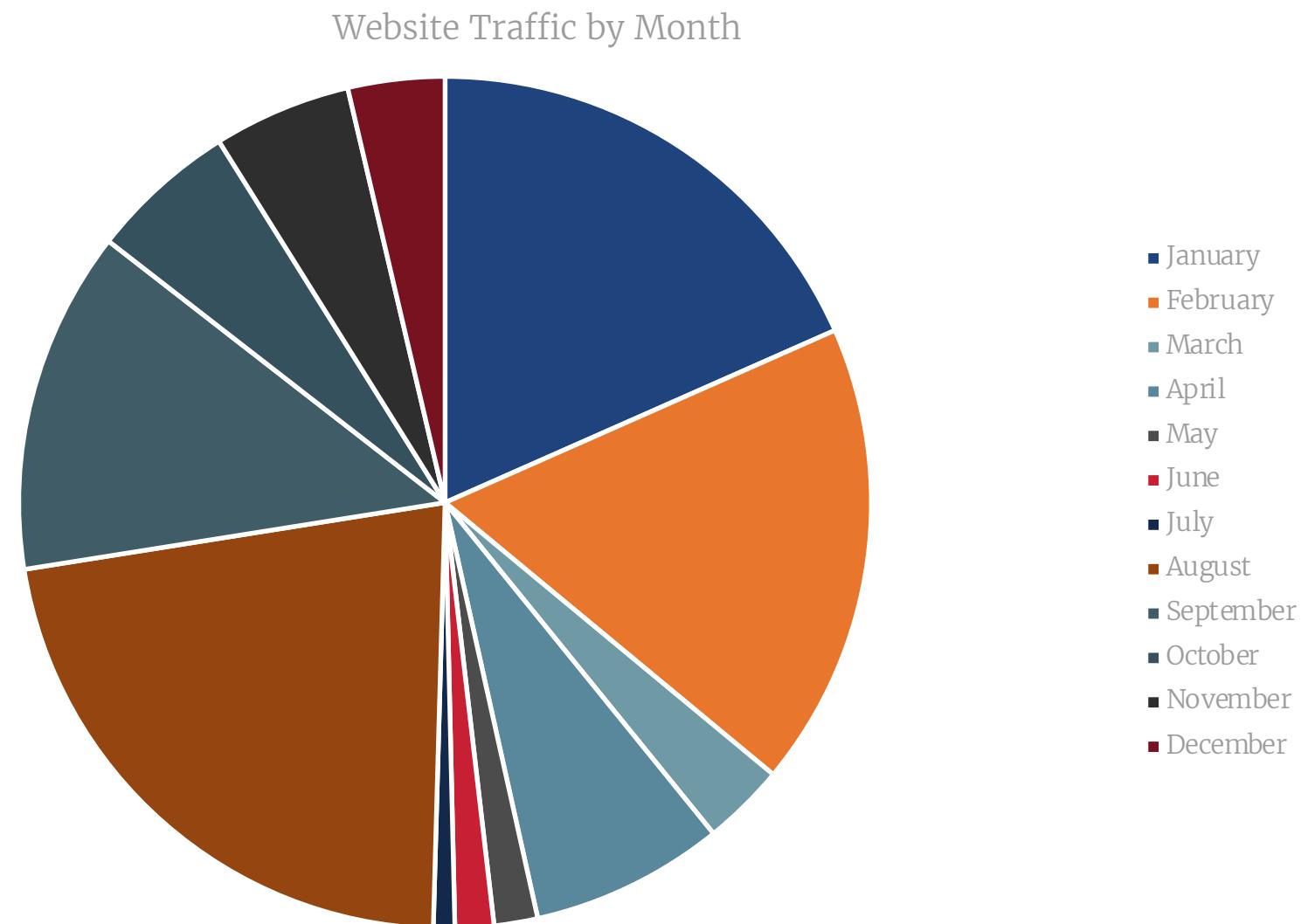
Pie Charts – NO!

From visualization guru Edward Tufte: “*the only worse design than a pie chart is several of them.*”

Pie charts are tempting because everyone understands what they are meant to convey (the various parts of a whole), **but it is much harder to compare slices of a pie** than it is to compare length or height (especially when slices become very thin):

There is extensive research showing that people are MUCH worse at interpreting the left-hand chart vs. the right-hand chart. This is a well-established best practice. It is shocking that you still see so many pie charts.

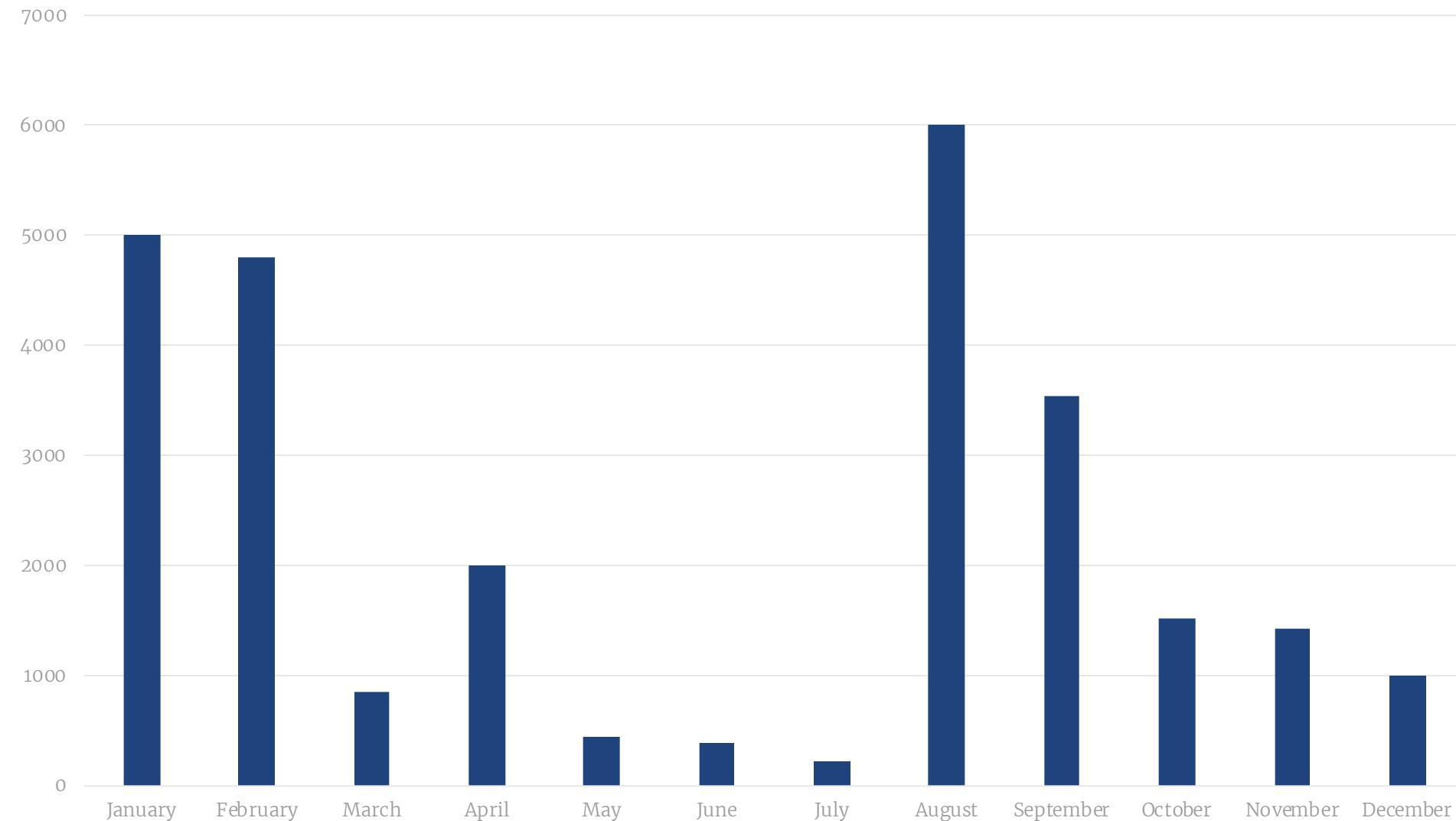
Pie Charts – NO!





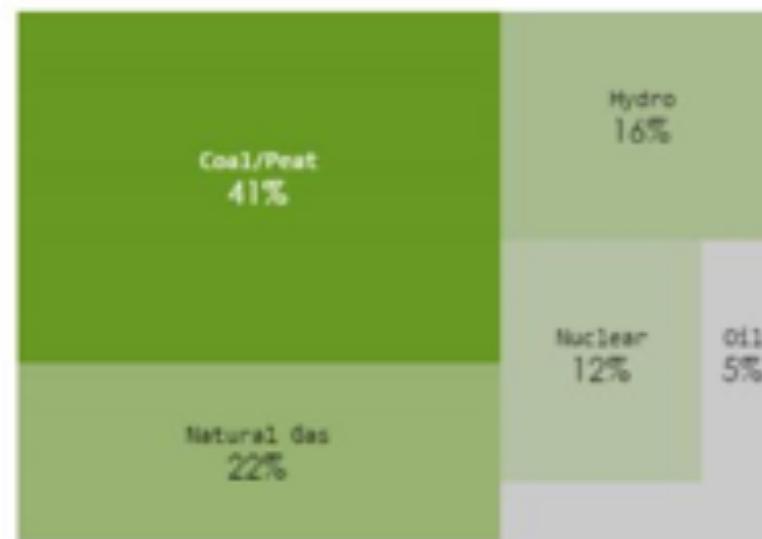
Pie Charts – NO!

Website Traffic by Month



Best Practices

Relative proportions:
treemap



Exploring relationships:
scatter plot





Type of Data

Qualitative Data (Categorical)

Arizona, Texas, Maine
Nancy, Kim, Jon, Jo
Starbucks, Peets

Qualitative Data (Ordinal)

Bronze, Silver, Gold
Excellent Health, Good Health, Poor Health
Love it, Like it, Hate it

Quantitative Data

Cost: \$100, \$500, \$0.50
Weight: 72 Kg, 45 Kg
Discount: 12%, 0.5%, 22%

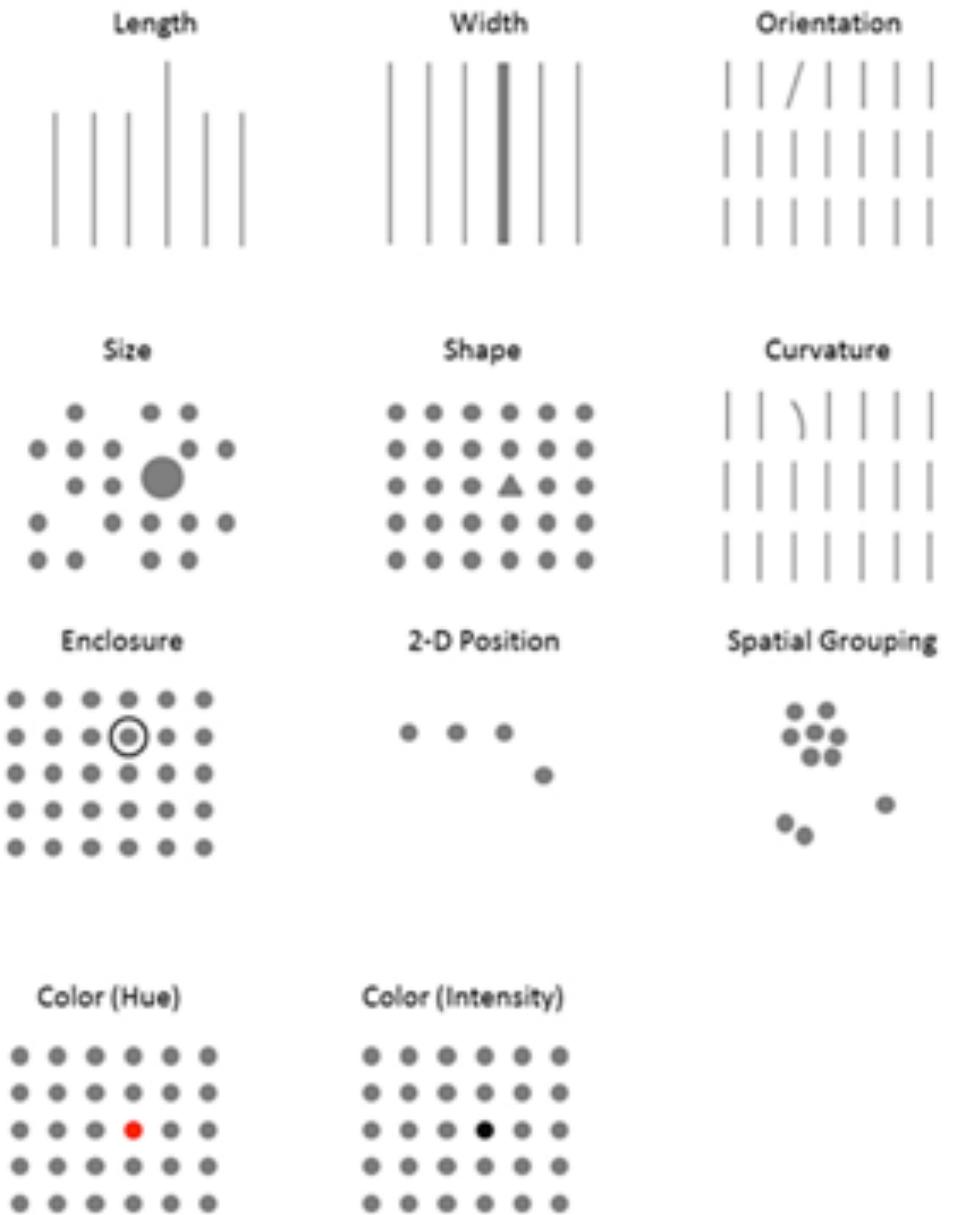
Qualitative, categorical – there is no clear order to the data and they don't have a clear relationship to each other, they don't contain numbers

Qualitative, ordinal – they have a relationship to each other but you can't measure the distance between them but you understand how they compare in relationship to one another

Quantitative – these data are numbers, you can do math with them, and you know exactly what the difference between them is (their relationship to one another – the difference between 5% and 10%). Quantitative data can be further broken down into **discrete** and **continuous** data. Discrete can be considered a “count” of something, continuous data is a measurement which can be broken down to lower levels of details.

Pre-attentive Visual Attributes

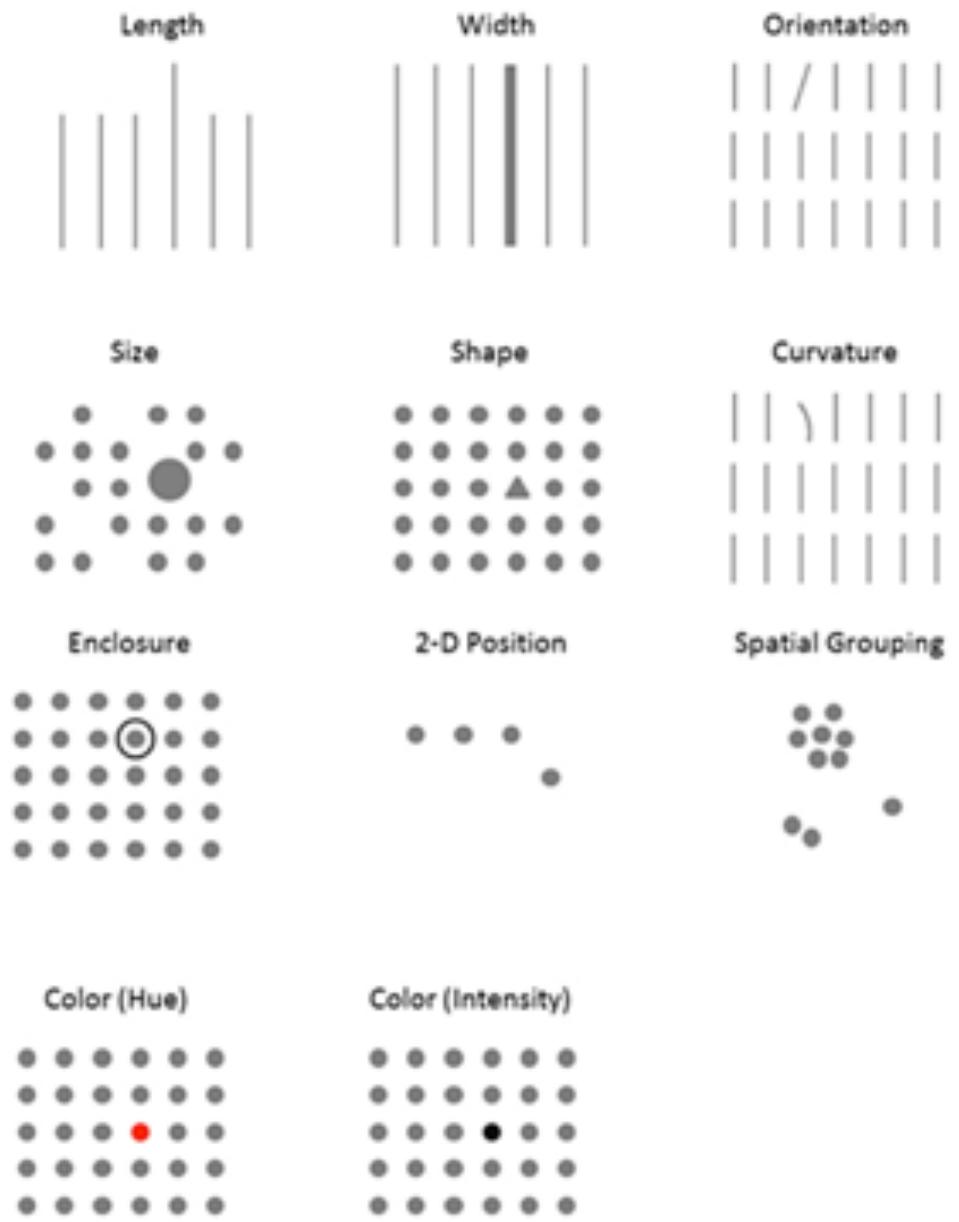
Length
Width
Orientation
Size
Shape
Curvature
Enclosure
2-D position
Spatial Grouping
Color(Hue)
Color(Intensity)



Pre-attentive Visual Attributes

By visualizing data, we are tapping into our ability to understand information pre-attentively.

Pre-attentive attributes are pieces of information we can process visually almost immediately, before sending the information to the attention processing parts of our brain. This is information we process and understand almost subconsciously. These are generally the best ways to present data, because we can see these patterns without thinking too hard.



Slides are not for Distribution



How do Humans Like Their Data?

Categorical

Position

Shape

Different Colors

Ordinal

Position

Size

Color Intensity

Shape

Quantitative

Position

Length

Size

Color Intensity



How do Humans Like Their Data?

Note that position is the best way across all 3 types of data. We can easily understand the relationship between two data points represented on an x and y axis.

Categorical

Position

Shape

Different Colors

Ordinal

Position

Size

Color Intensity

Different Colors

Shape

Quantitative

Position

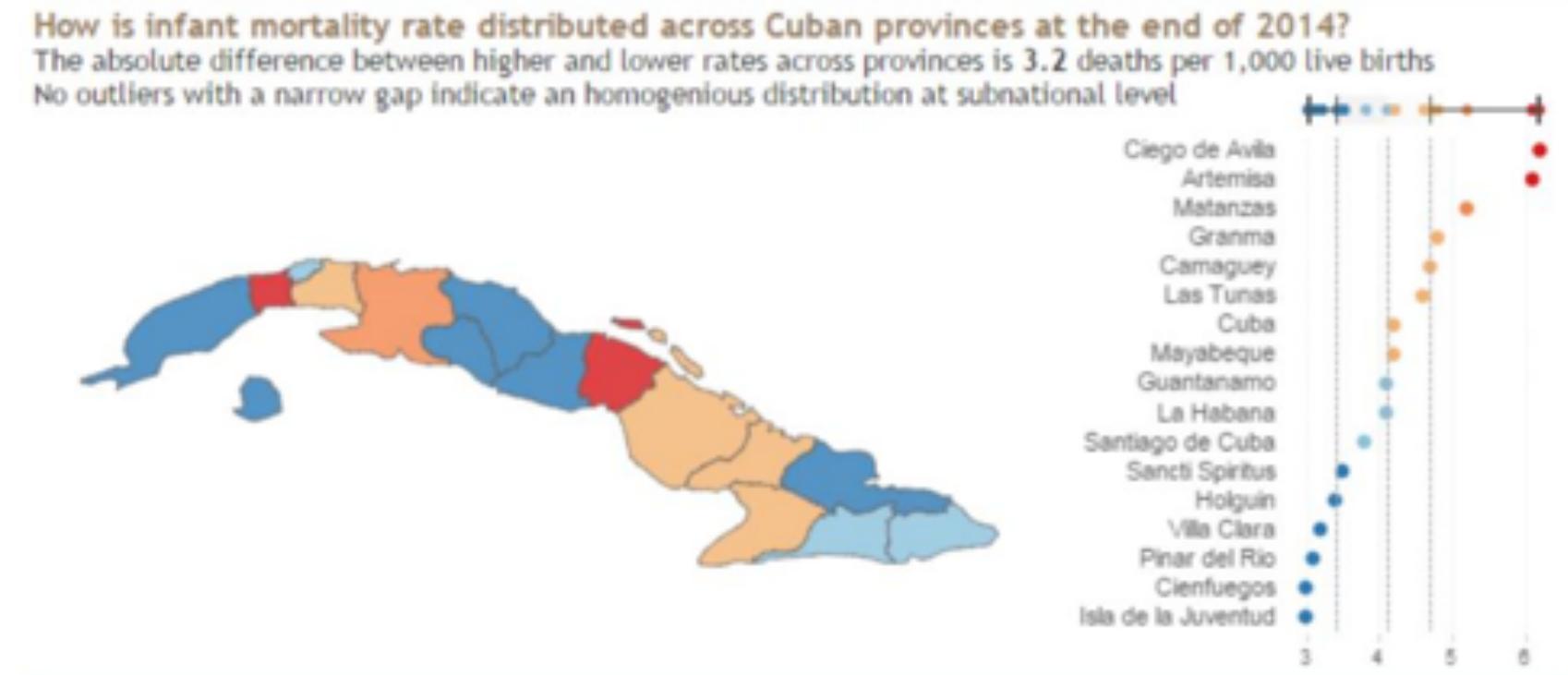
Length

Size

Color Intensity

Mapping to Insight

Use filled maps for defined areas and only ONE measure

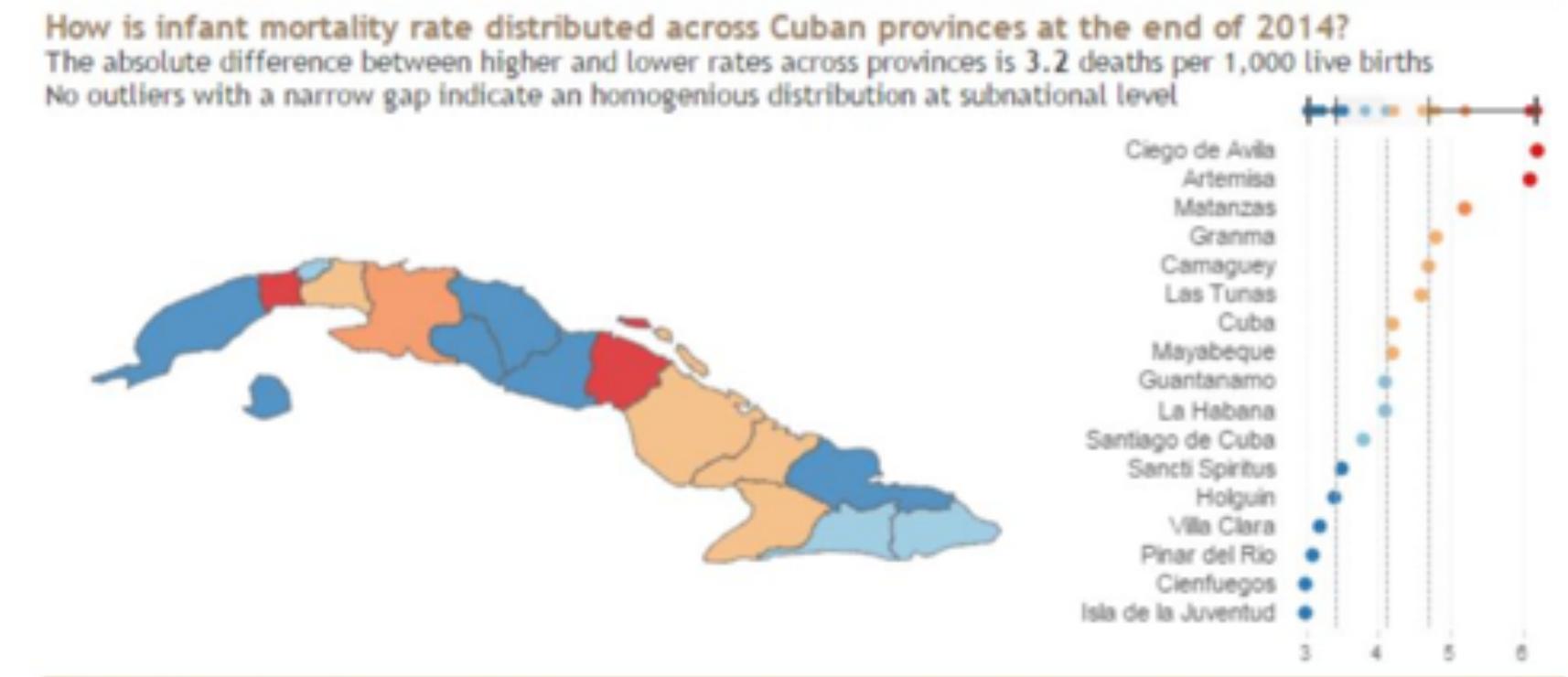


Mapping to Insight

Use filled maps for defined areas and only ONE measure

Filled maps work well when you have defined areas like states or cities and only one measure.

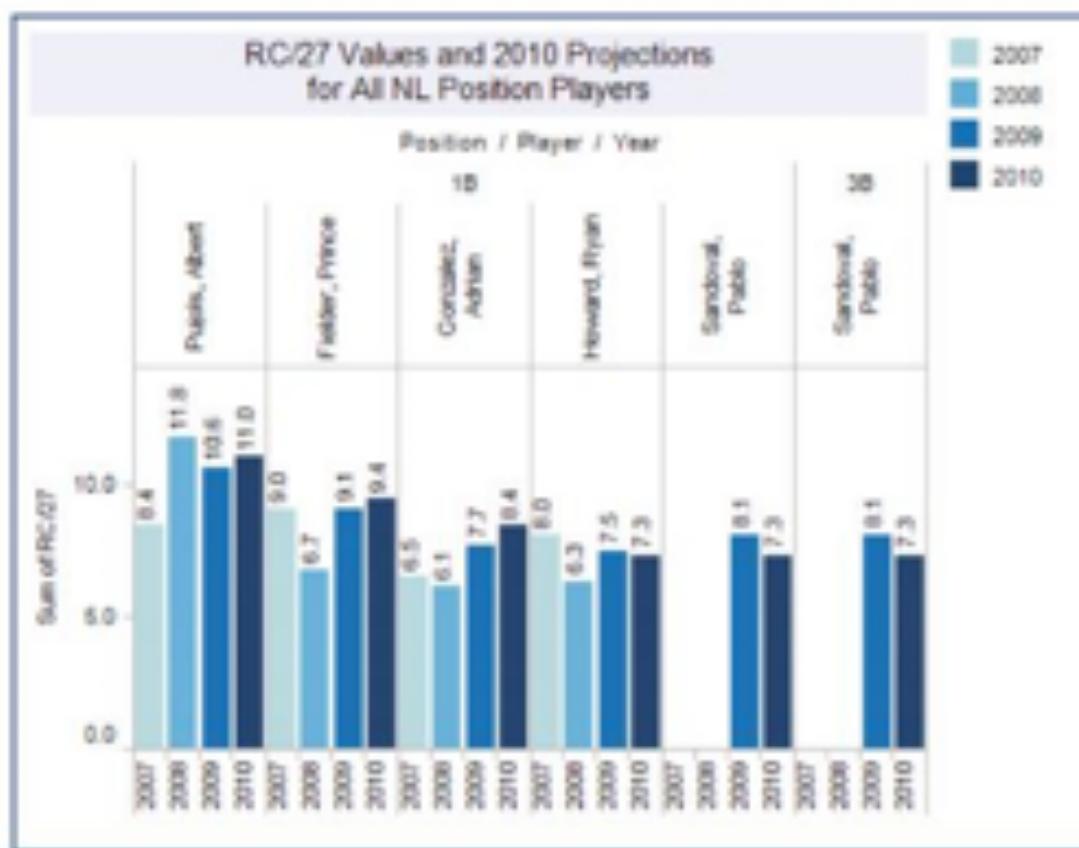
For this map, we are looking at infant mortality across Cuban provinces – see can see that there are two pockets of 4–6 deaths per 1,000 live births



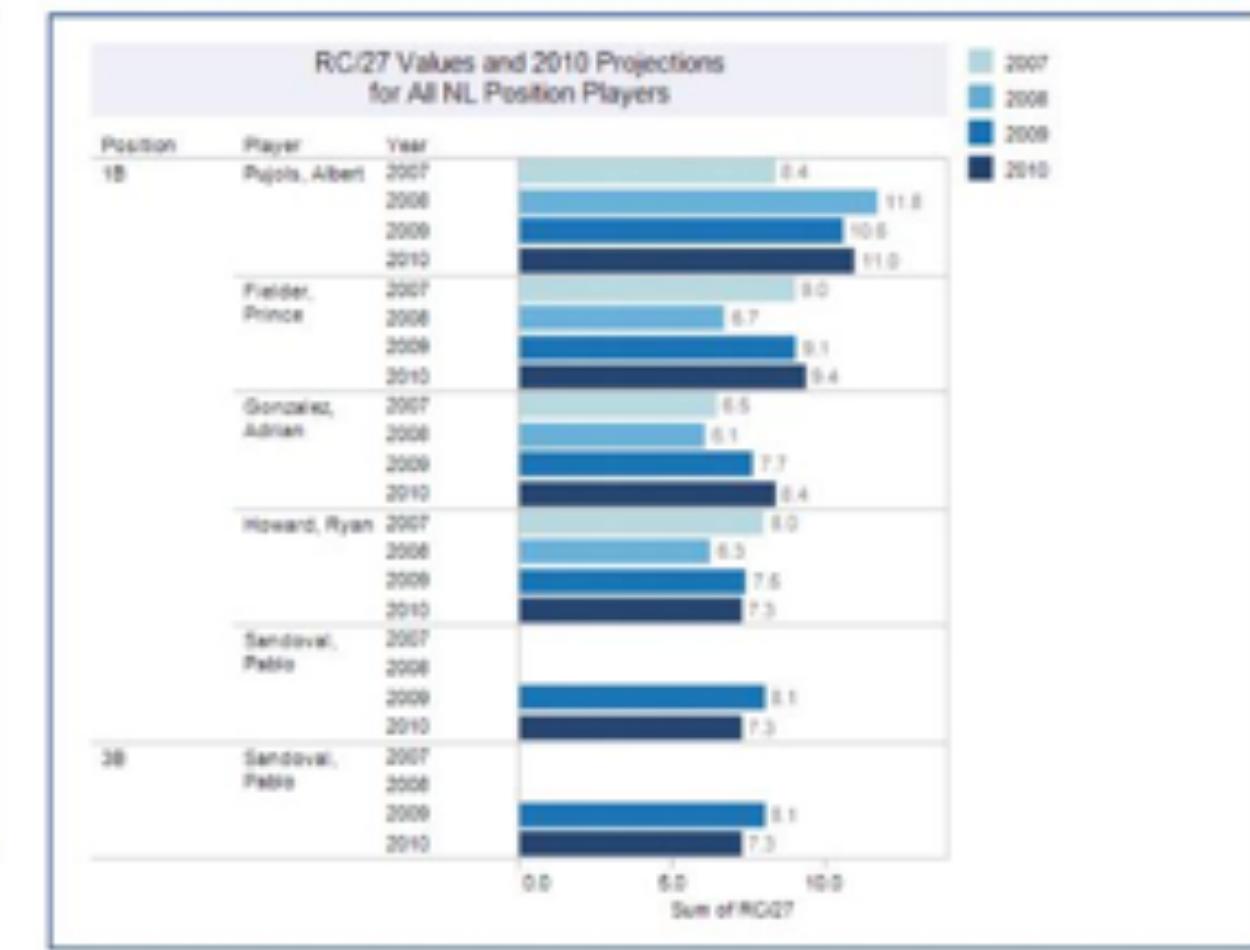
Best Practices

Orient data so people can read it

Good



Better

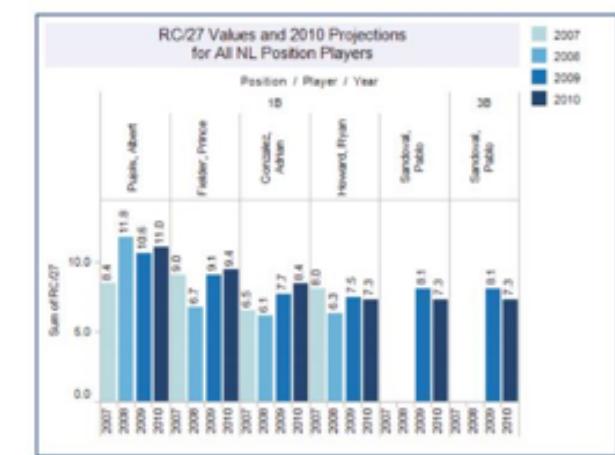


Best Practices

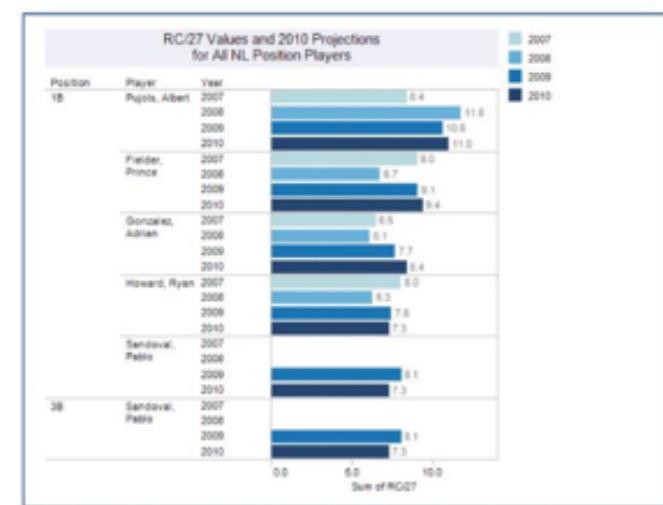
Orient data so people can read it

Orient data so people can read it easily. You read English from left to right so as much as possible, we want our data to go left to right.

Good

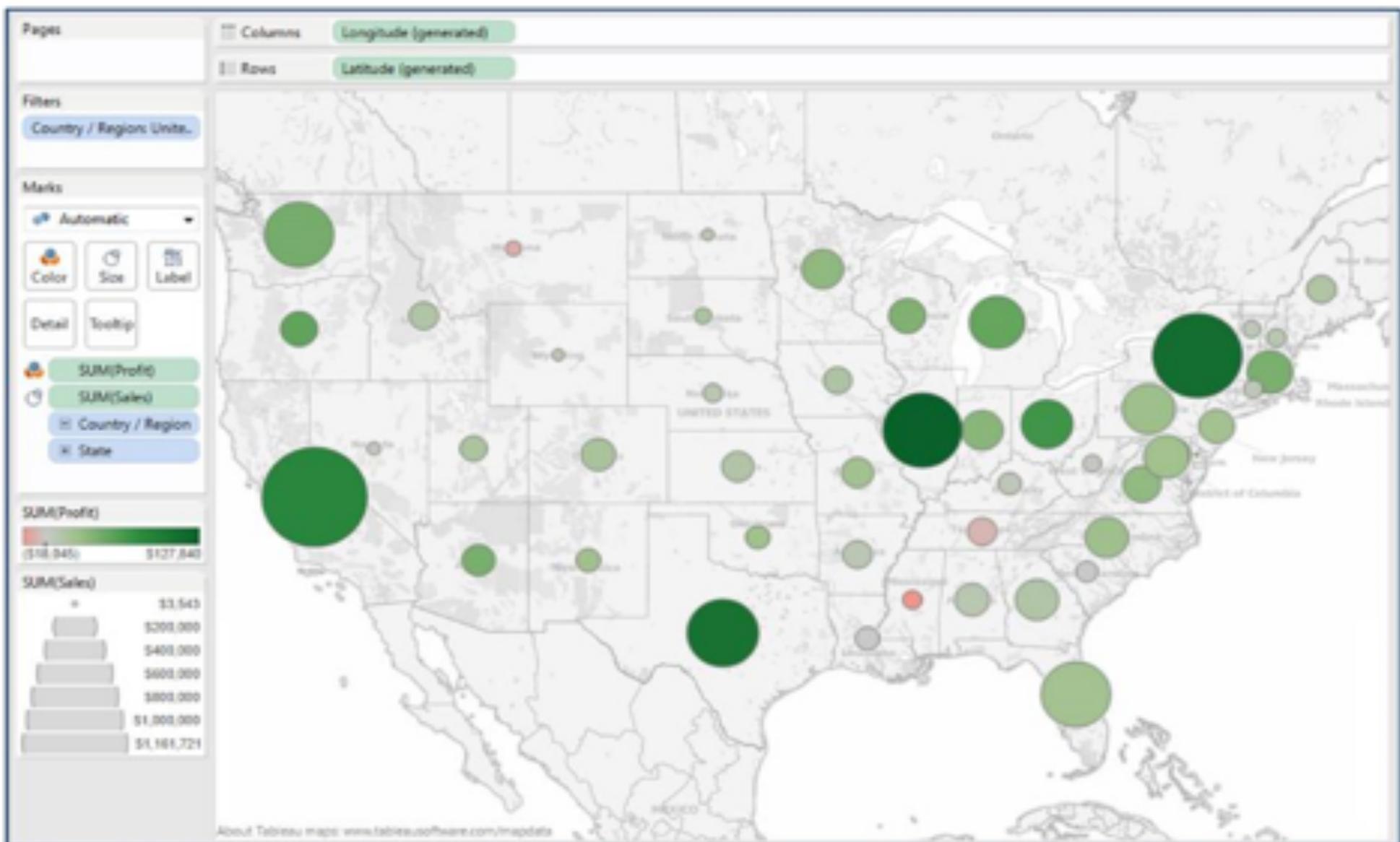


Better



Mapping to Insight

Filled maps won't work well for multiple measures



Dashboards

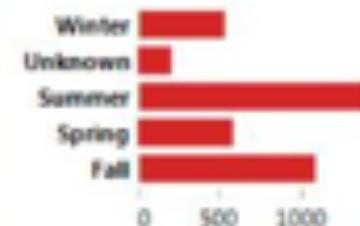
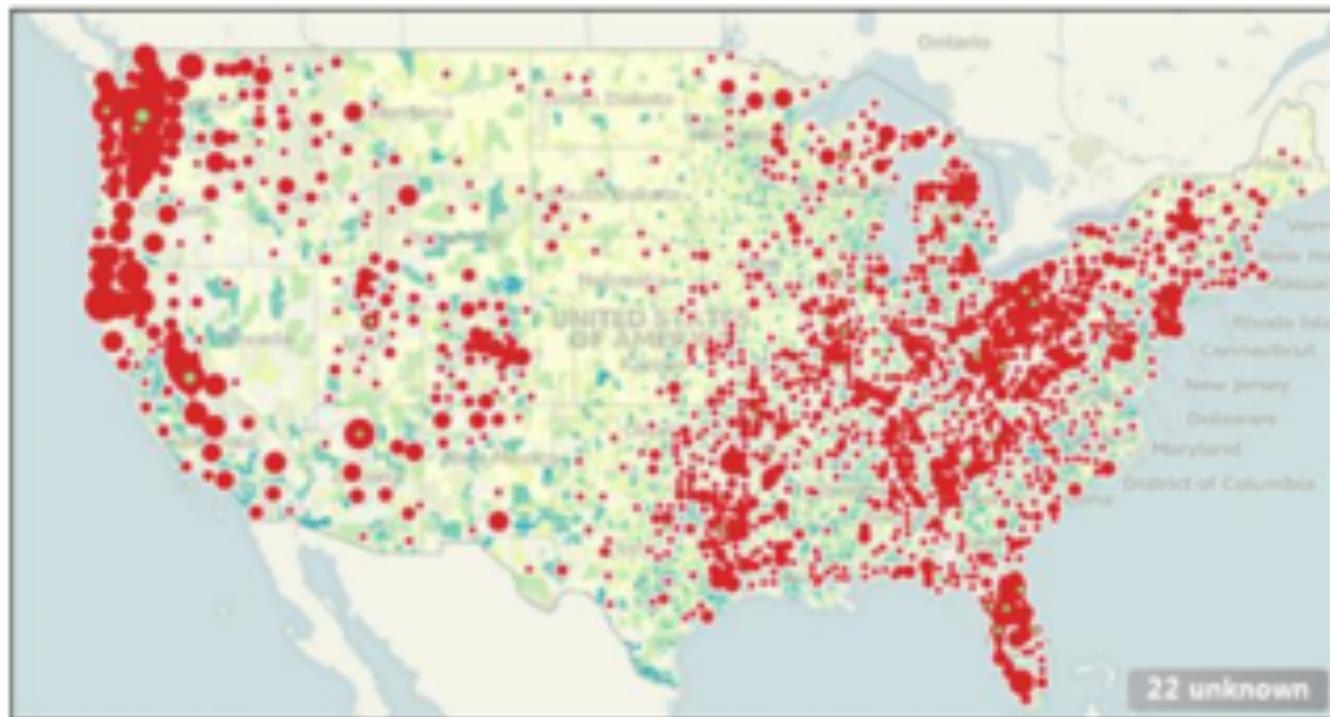
Dashboards should pass the five second test

What do you think the 5 second test is?

The goal is that you should know what is going on in the dashboard in 5 seconds.

This one is good b/c you know exactly what it is – good title, clear why map was used.

Finding Bigfoot



Data gathered from the official website of the "Bigfoot Field Researchers Organization" (BFRO).

The data was attempted to be scrubbed and cleaned to attain some type of normalcy, unfortunately the BFRO data submission process has no validation and fields are often used arbitrarily by submitters.

BFRO does the "Finding Bigfoot" Animal Planet TV show.

Click on ANY element of the visualization (location, season, year, detail field) in order to filter by that item.
Select the element AGAIN to go back to the full view.



The BFRO classifies sightings according to a system based on the sightings' "potential for misinterpretation".

Total Sightings	Class A	Class B	Class C	Unclassified
3,806	1,951	1,696	31	128

Dashboard : Five Second Test

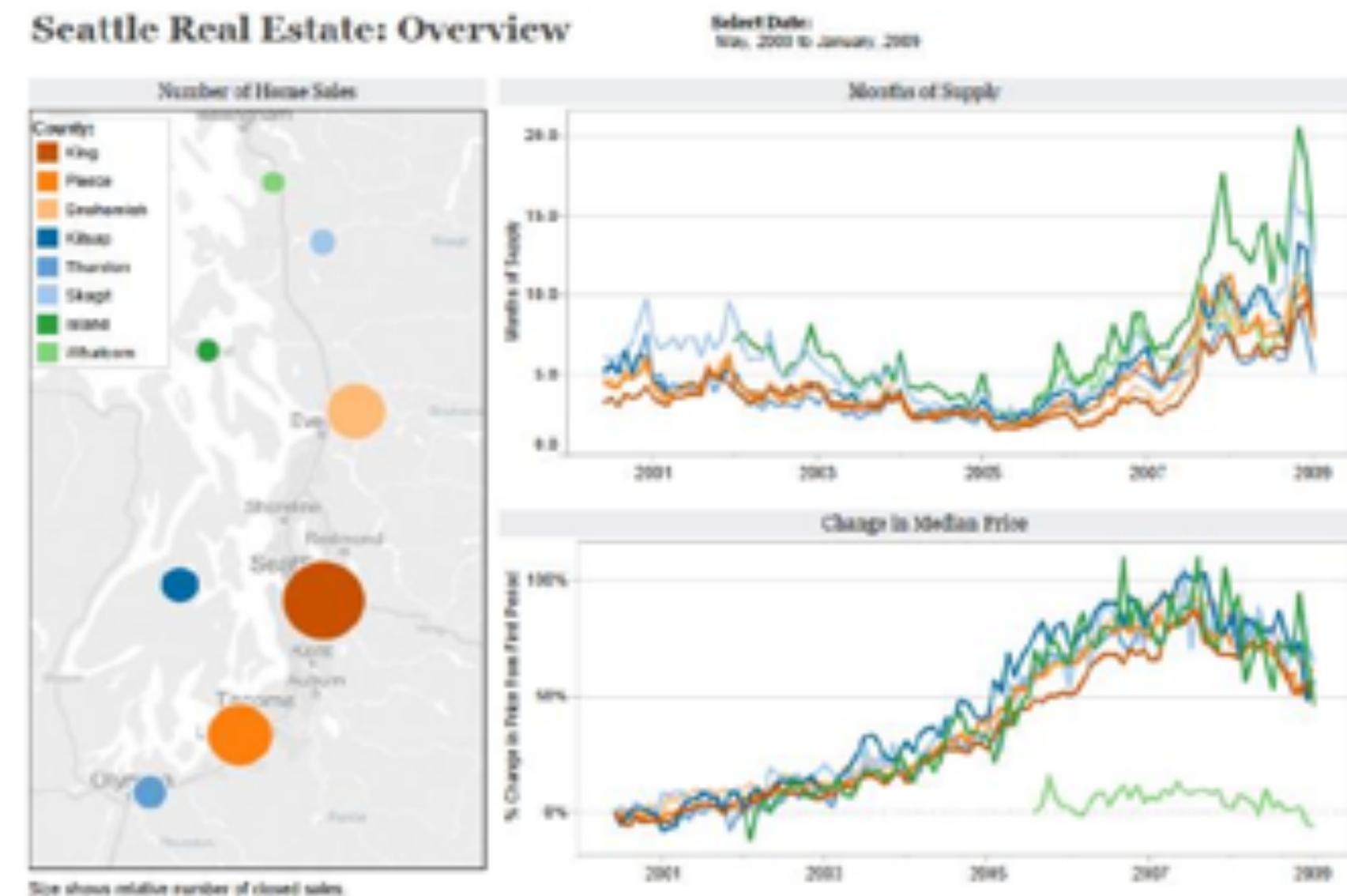
Most important view goes on top or top-left

Legends go near their views

Avoid using multiple color schemes on a single dashboard

Use 5 views or fewer in dashboards

Provide interactivity

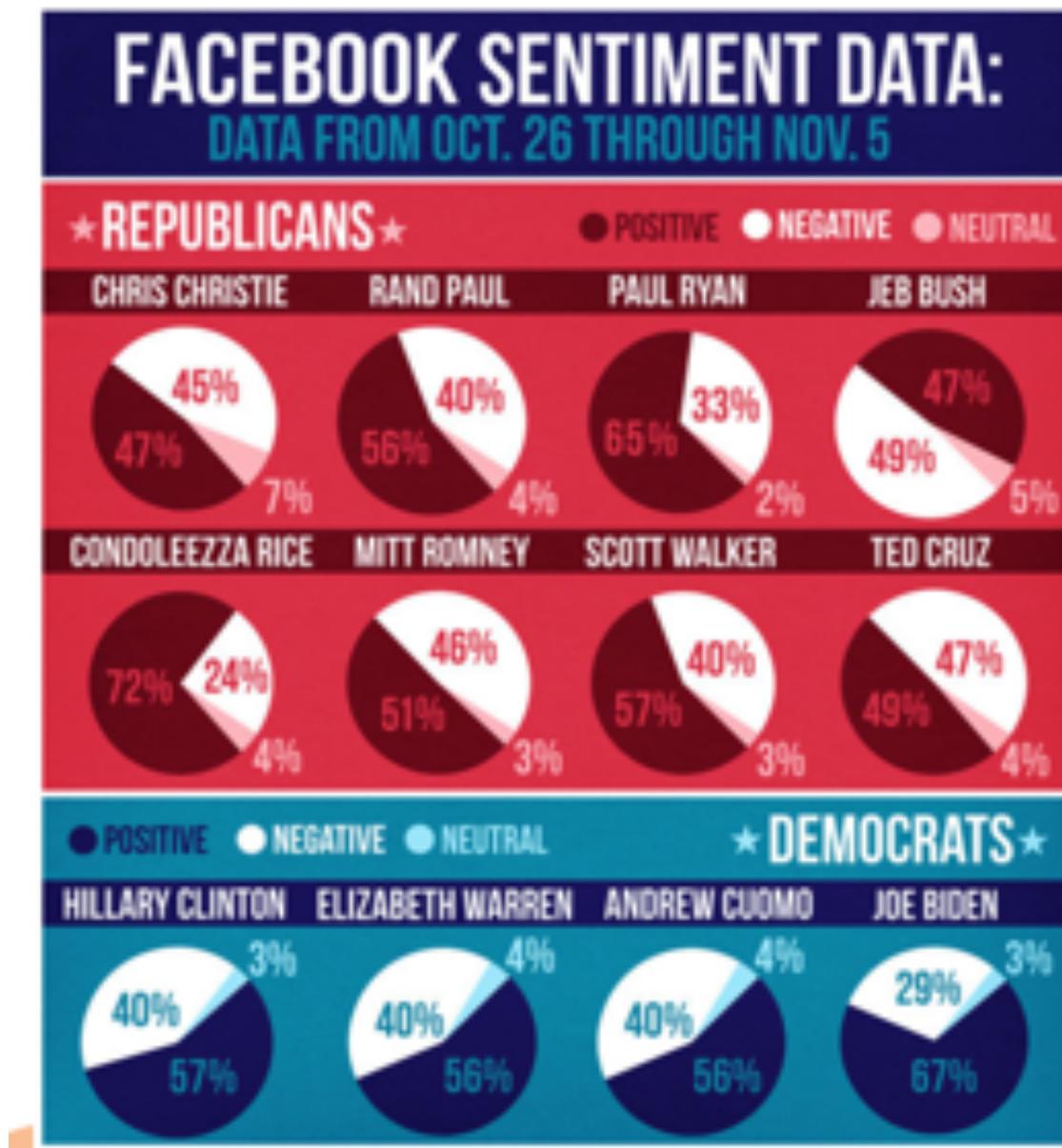




Five Questions to Consider in a Dashboard Review

1. What problem/question does this solve/answer?
2. Is this really the best way to display the information?
3. Does everything add value?
4. Is there functional interactivity?
5. Are there clear labels?

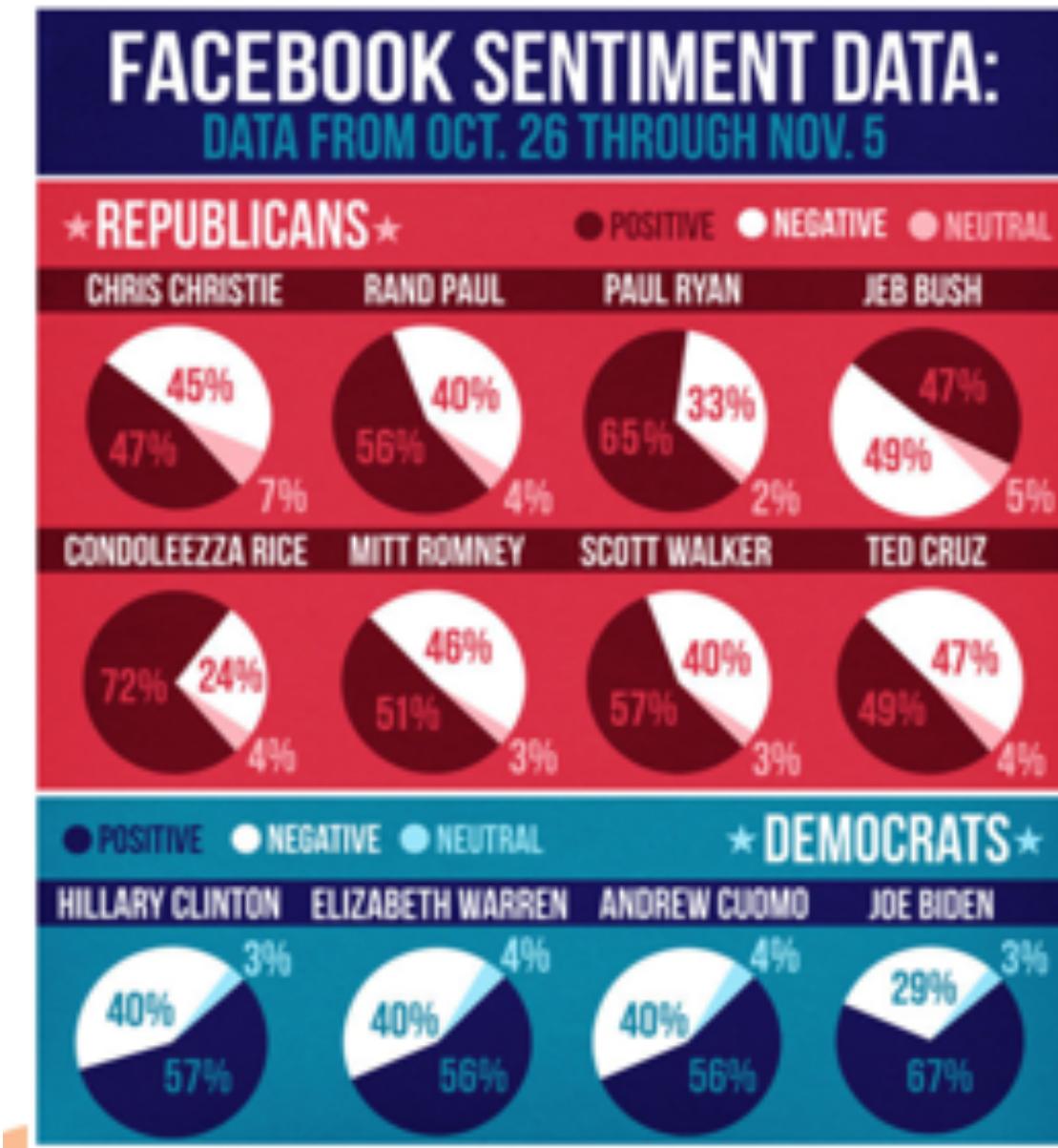
Dashboard Review Example: Before



(2014 poll data)
How can this improve?

1. What problem/question does this solve/answer?
2. Is this really the best way to display the information?
3. Does everything add value?
4. Is there functional interactivity?
5. Are there clear labels?

Dashboard Review Example: Before



Draw on the 5 questions from the viz review:

What question does this answer? In this case, the question is “how do Facebook users see 2016 political candidates”

Is the information displayed in the best way?

Pie charts are hard to compare, no sorting of the individual candidates

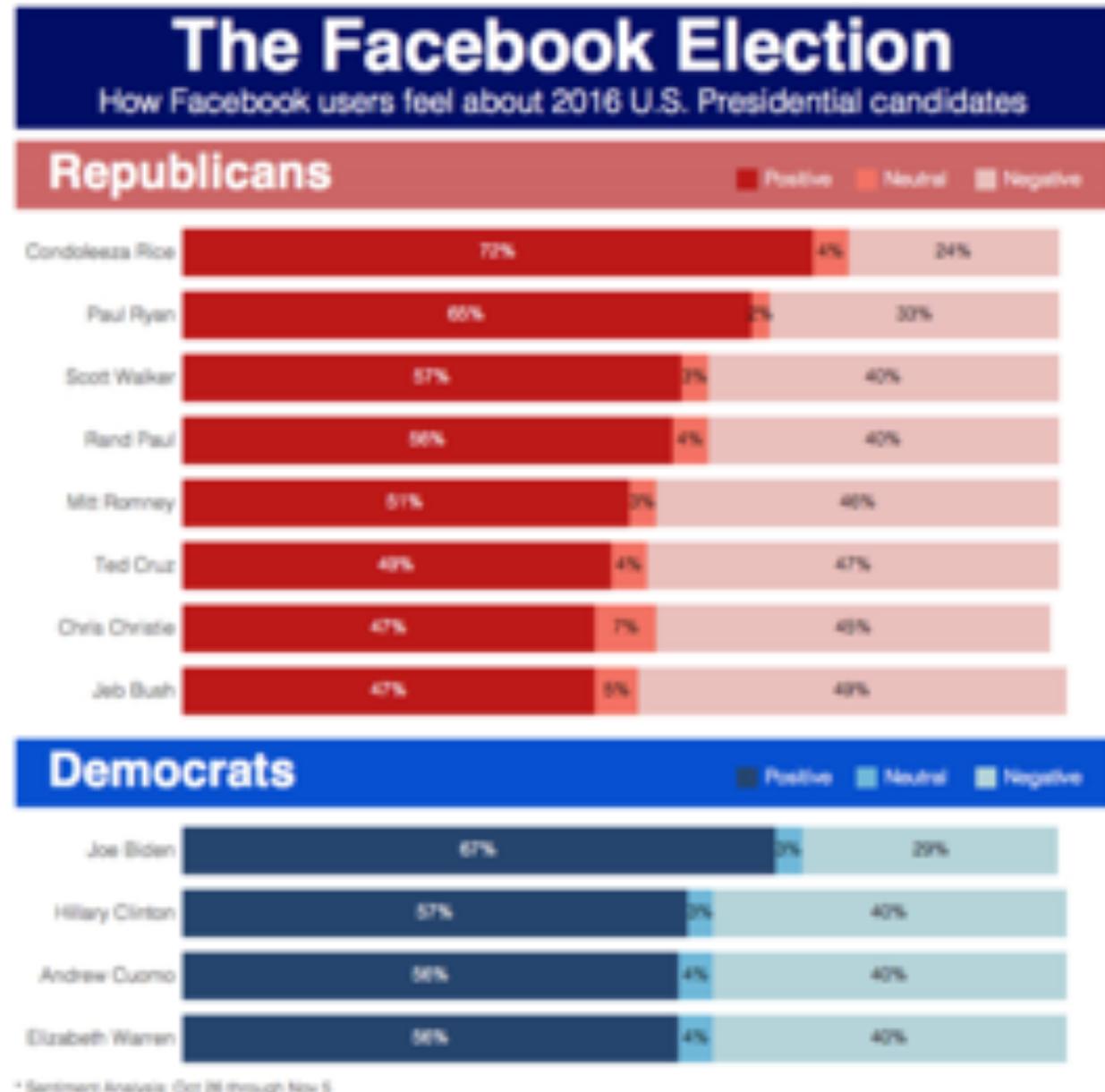
Second row of red labels could be ambiguous

Does everything add value? Alternating the republicans and democrats titles and color schemes doesn't help

Functional interactivity? no

Clear labels? Unclear title, All caps are hard to read

Dashboard Review Example: After



What's changed?

- Title clearly addresses question at hand
- Bar charts for cleaner comparison
- Labels – showing more with less
- Sorting of bars for stronger visualization of sentiment



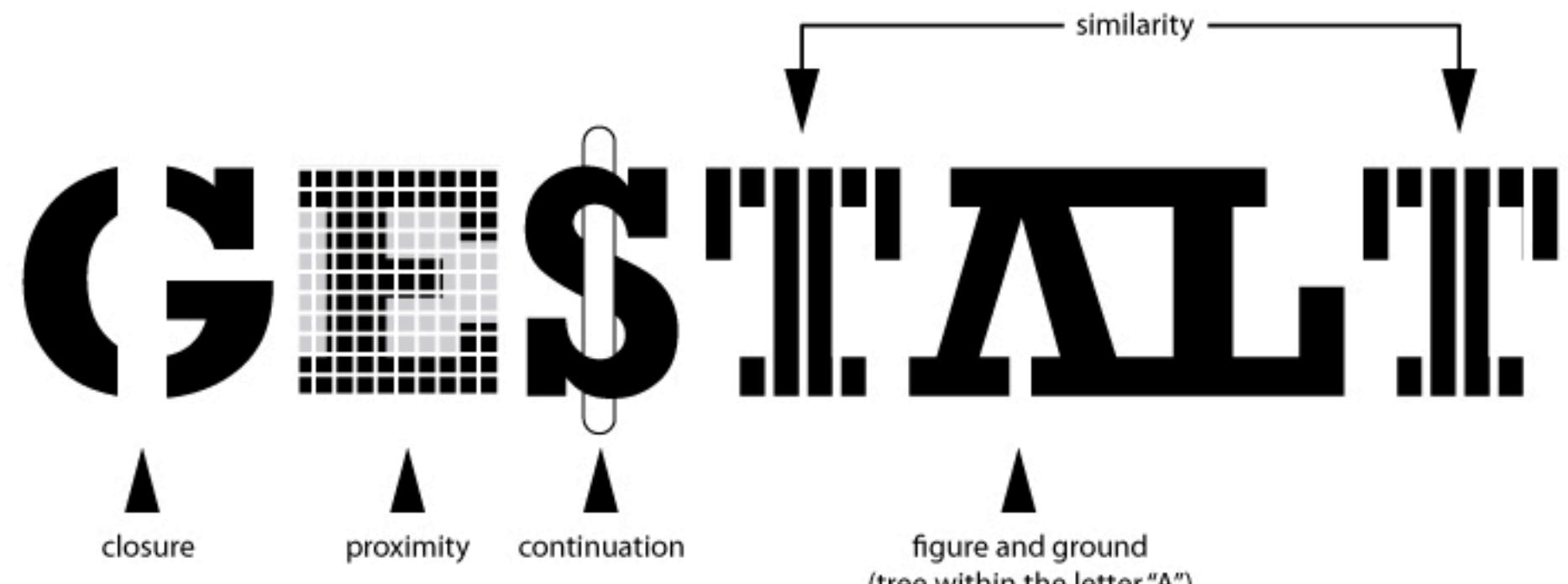
Visual Perception

What the Brain interprets what the eyes sees

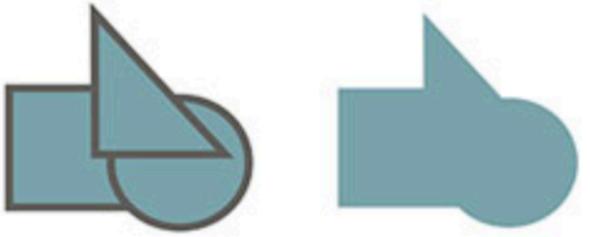
In visual communication, it is important to understand how a design is being perceived and interpreted.

Gestalt Principles

- Figure/Ground
- Proximity
- Similarity
- Parallelism
- Common fate
- Closure
- Continuity



"Gestalt" is a German word that pertains to a unified pattern, form, figure, or structure.



1

Simplicity

Combining simplicity with creativity can lead to stunning creations.

How to Master Simplicity:
Know how to balance simple shapes with visual stimulation. Give the eye a comfortable form that helps it interpret what it sees.

Source: <https://creativemarket.com/>

2



Figure-Ground

People can immediately identify which element is the figure, and which is the ground. Use these two related principles to make the most of the figure-ground effect:

- **Area** – The viewer's mind sees the smallest element as the figure and the larger one as the ground or background.
- **Convexity** – Convex elements are related to figures.

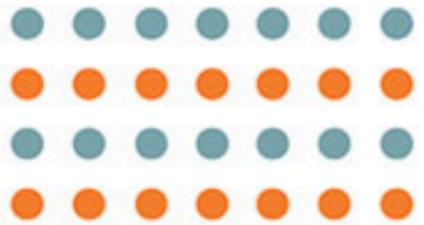
3



Proximity

Elements close to each other are perceived as part of the same group. Common Use Case: Kerning. Proper kerning helps readers snap up each word.

4



Similarity

Elements that look alike are perceived as part of a group.

The principle of similarity applies to:

- Color
- Size
- Orientation
- Shape
- Texture

5



Common Fate

Objects that seem to be moving in the same direction are often seen as a group.

Source: <https://creativemarket.com/>

6



Symmetry

The principle of symmetry applies to

- Mirrored shapes
- Balanced elements
- Parallel lines

7



Continuity

Objects that are plotted in a continuous pattern are grouped together by the mind. Smooth lines often make a unified figure.

Source: <https://creativemarket.com/>

8



Closure

The mind wants closure. A shape only needs to be implied for the mind to "fill in the gaps" and see what it wants to see

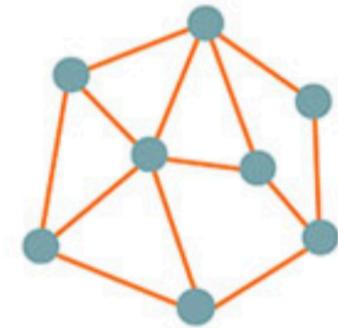
9

Common Region

Elements in one region are seen as one group. Badge designs are seen as unified, despite some containing text, banners, and other objects.

Source: <https://creativemarket.com/>

10



Connectedness

When there is a connection between elements, it is easy to see a unified figure.

Achieve connections with:

- Arrow
- Lines
- Illustrations (i.e. a rope)



Connect to Dataset

Tableau - Book1

Data Analytics Connect to Data

Dimensions

77

78

79

80

81

82

83

84

85 Measures

86

87

88

89

90

91

92

93

94

Sheet 1

Drop field here

Drop field here

Drop field here

Color Size Text

Detail Tooltip

Standard

Show Me

Columns

Rows

Pages

Data Source Sheet 1

Tableau interface showing the workspace for creating a visualization. The top navigation bar includes standard application icons and a 'Show Me' button. The left sidebar contains tabs for 'Data' (selected), 'Analytics', and 'Connect to Data', along with sections for 'Dimensions' (listing rows 77 through 94) and 'Measures'. The main workspace is titled 'Sheet 1' and features three blank drop zones labeled 'Drop field here'. A 'Marks' shelf on the left provides options for 'Automatic' marking and specific controls for 'Color', 'Size', 'Text', 'Detail', and 'Tooltip'. The bottom navigation bar includes icons for 'Data Source', 'Sheet 1', and other sheet management functions.

Craigslist

A Case Study

Inspect the page

Secure https://sfbay.craigslist.org/sfc/

Bookmarks P https://viridislearn... ML jo Cloud Reader Python Code for ML... U SFDC Tab DNLDs Product Viridis App

craigslist

post to classifieds
my account
search craigslist
search

event calendar

M	T	W	T	F	S	S
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30

help, faq, abuse, legal
avoid scams & fraud
personal safety tips
terms of use
privacy policy
system status

about craigslist
craigslist is hiring in sf
craigslist open source
craigslist blog
best-of-craigslist
craigslist TV
"craigslist joe"
craig connects
progressive directory
weather quake tide

city of san francisco

sfc sby eby pen nby scz

community

activities	local news
artists	lost+found
childcare	musicians
classes	pets
events	politics
general	rideshare
groups	volunteers

housing

apts / housing
housing swap
housing wanted
office / commercial
parking / storage
real estate for sale
rooms / shared
rooms wanted
sublets / temporary
vacation rentals

jobs

accounting+finance
admin / office
arch / engineering
art / media / design
biotech / science
business / mgmt
customer service
education
food / bev / hosp
general labor
government
human resources
internet engineers
legal / paralegal
manufacturing
marketing / pr / ad
medical / health
nonprofit sector
real estate
retail / wholesale
sales / biz dev
salon / spa / fitness
security
skilled trade / craft
software / qa / dba
systems / network
technical support
transport
tv / film / video
web / info design
writing / editing
[ETC]
[part-time]

nearby cl

bakersfield
chico
fresno
gold country
hanford
humboldt
inland empire
klamath falls
las vegas
los angeles
medford
mendocino co
merced
modesto
monterey
orange co
palm springs
redding
reno
roseburg
sacramento
san luis obispo
santa barbara
santa maria
siskiyou co
stockton
susansville
ventura
visalia-tulare
yuba-sutter

english

us cities

us states

canada

cl worldwide

SF bay area jobs "data engineer" × Craigslist-Soup/scrap.py at r SF bay area motorcycles/scoot × Sudhir

Secure https://sfbay.craigslist.org/search/mca

Bookmarks P https://viridislearn... ML jo Cloud Reader Python Code for ML... U SFDC Tab DNLDs Product Viridis App Other Bookmarks

CL SF bay area > all SF bay area > for sale > motorcycles

post account

motorcycles/scooters

search motorcycles/scooters

save search

all owner dealer

gallery

<< < prev 1 - 120 / 2500 next >

newest

- search titles only
- has image
- posted today
- bundle duplicates
- include nearby areas

MILES FROM ZIP
 miles from zip

PRICE
 min max

MAKE AND MODEL
 make / model

ENGINE DISPLACEMENT (CC)
 min max

MODEL YEAR
 min max

ODOMETER
 min max

- [▶ condition](#)
- [▶ fuel](#)
- [▶ paint color](#)
- [▶ title status](#)
- [▶ transmission](#)

[reset](#)

[update search](#)



★ Apr 9 2006 Custom Harley FXSTI \$8700
(healdsburg / windsor) 



★ Apr 9 Quad Artic 4x4 2004 \$3950 (fairfield / vacaville) 



★ Apr 9 Quad Artic 4x4 2004 \$3950 (fairfield / vacaville) 



SF bay area jobs "data engineer" Craigslist-Soup/scrap.py at SF bay area motorcycles/scoot Sudhir

Secure https://sfbay.craigslist.org/search/mca

Bookmarks P https://viridislearn... ML jo Cloud Reader Python Code for ML... U SFDC Tab DNLDs Product Viridis App Other Bookmarks

CL SF bay area > all SF bay area > for sale > motorcycles

post account

motorcycles/scooters

search motorcycles/scooters

all owner dealer

search titles only has image posted today bundle duplicates include nearby areas

MILES FROM ZIP

PRICE

MAKE AND MODEL

gallery

1 - 120 / 2500 << < prev next >

\$8700 

\$3950 

\$3950 

Elements Console Sources Network Timeline Profiles Application Security Audits Ember AdBlock

```

$8700
::after
</a>
►<p class="result-info">...</p>
```

Styles Computed Event Listeners DOM Breakpoints >

Filter :hov .cls +

element.style { }

@media only screen and (min-width: 480px) .search-legend.bottom { search.css?v=27e1d42...1 webkit-box-shadow: none; }

8700 1 of 4 ▲ ▼ Cancel

Console

top Preserve log

Slides are not for Distribution

SF bay area jobs "data engineer" Craigslist-Soup/scrap.py at SF bay area motorcycles/scoot Sudhir

Secure https://sfbay.craigslist.org/search/mca

Bookmarks P https://viridislearn... ML jo Cloud Reader Python Code for ML... U SFDC Tab DNLDs Product Viridis App Other Bookmarks

CL SF bay area > all SF bay area > for sale > motorcycles

post account

motorcycles/scooters search motorcycles/scooters save search

all owner dealer

search titles only has image posted today bundle duplicates include nearby areas

MILES FROM ZIP miles from zip

PRICE min max

MAKE AND MODEL make / model

gallery << < prev 1 - 120 / 2500 next > newest

\$8700  \$3950  \$3950 

Elements Console Sources Network Timeline Profiles Application Security Audits Ember AdBlock

```

$8700
::after
</a>
►<p class="result-info">...</p>
```

Styles Computed Event Listeners DOM Breakpoints >

Filter :hov .cls +

element.style { }

@media only screen and (min-width: 480px)
.search-legend.bottom { search.css?v=27e1d42...1
-webkit-box-shadow: none;

8700 1 of 4 Cancel

Console

top Preserve log

Slides are not for Distribution



```
1
2
3 # (C) DataTiles.io
4 # (C) DatatTiles.ai
5 # Sudhir Wadhwa , Jyoti Wadhwa , Feb , 2016
6 # For education / demo purpose only. Please do not use this code in production
7
8 import bs4 as bs
9 import csv
10 import requests
11
12
13 try:
14     # For Python 3.0 and later
15     from urllib.request import urlopen
16 except ImportError:
17     # Fall back to Python 2's urllib2
18     from urllib2 import urlopen
19
20 sourcehtml = urlopen("https://sfbay.craigslist.org/search/mca")
21 soup = bs.BeautifulSoup(sourcehtml,"lxml")
22
23
24 for cost in soup.findAll("span", { "class" : "result-price" }):
25     print(cost.text)
26
27
```

(base) sudhirwadhwa ~/Desktop \$ vi craigslist.py

```
import bs4 as bs
import csv
import requests

try:
    from urllib.request import urlopen
except ImportError:
    from urllib2 import urlopen

sourcehtml = urlopen("https://sfbay.craigslist.org/search/mca")
soup = bs.BeautifulSoup(sourcehtml,"lxml")

print (soup)

for cost in soup.findAll("span", {"class":"result-price" } ):
    print("====")
    print(cost.text)
```



LAB 4 : quotes : co-coding lab



You will need
scrapy for this lab

<https://scrapy.org/>



Quotes to Scrape

<http://quotes.toscra>

"The world as we have created it is a process of our thinking. It cannot be changed without changing our thinking."

by [Albert Einstein](#) (about)

Tags: [change](#) [deep-thoughts](#) [thinking](#) [world](#)

"It is our choices, Harry, that show what we truly are, far more than our abilities."

by [J.K. Rowling](#) (about)

Tags: [abilities](#) [choices](#)

"There are only two ways to live your life. One is as though nothing is a miracle. The other is as though everything is a miracle."

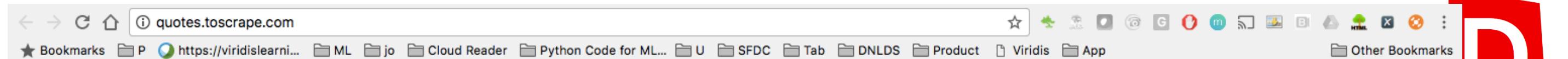
by [Albert Einstein](#) (about)

Tags: [inspirational](#) [life](#) [live](#) [miracle](#) [miracles](#)

"The person, be it gentleman or lady, who has not pleasure in a good novel, must be intolerably stupid."

by [Jane Austen](#) (about)

Tags: [aliteracy](#) [books](#) [classic](#) [humor](#)



Quotes to Scrape

[Login](#)

"The world as we have created it is a process of our thinking. It cannot be changed without changing our thinking."

by [Albert Einstein](#) ([about](#))

Tags: [change](#) [deep-thoughts](#) [thinking](#) [world](#)

"It is our choices, Harry, that show what we truly are, far more than our abilities."

Top Ten tags

love

inspirational

life

humor

The screenshot shows the Chrome DevTools interface. The Elements tab is active, displaying the DOM structure of a quote card. The quote itself is in a element, and the author's name is in a element. Below the author's name is a link to their profile page, represented by an [element. The right side of the DevTools shows the Styles tab, which contains the following CSS rules for 'a' elements:](#)

```
element.style {  
}  
a {  
    color: #4582ec;  
    text-decoration: none;  
}  
a {  
    background-color: transparent;  
}
```

The 'bootstrap.min.css:11' reference indicates these styles are from the Bootstrap library.

Quotes to Scrape

Login



"The world as we have created it is a process of our thinking. It cannot be changed without changing our thinking."

by [Albert Einstein](#) ([about](#))

Tags: [change](#) [deep-thoughts](#) [thinking](#) [world](#)

"It is our choices, Harry, that show what we truly are, far more than our abilities."

by [J.K. Rowling](#) ([about](#))

Tags: [abilities](#) [choices](#)

"There are only two ways to live your life. One is as though nothing is a miracle. The other is as though everything is a miracle."

by [Albert Einstein](#) ([about](#))

Tags: [inspirational](#) [life](#) [live](#) [miracle](#) [miracles](#)

"The person, be it gentleman or lady, who has not pleasure in a good novel, must be

Top Ten tags

love
inspirational
life
humor
books
reading
friendship
friends
truth
smile

The screenshot shows a browser developer tools window open over the Quotes to Scrape website. The main content area displays four quote cards. The bottom quote card's text is selected. The developer tools interface includes:

- Elements** tab: Shows the DOM structure of the selected quote text.
- Console** tab: Displays the selected text: `""The world as we have created it is a process of our thinking. It cannot be changed without changing our thinking."" == $0`.
- Sources**, **Network**, **Timeline**, **Profiles**, **Application**, **Security**, **Audits**, **Ember**, **AdBlock** tabs: Standard browser developer tools tabs.
- Styles**, **Computed**, **Event Listeners**, **DOM Breakpoints** buttons: CSS and DOM analysis tools.
- Filter**: A search bar for filtering results.
- Console** tab: Shows the command `$(text)` and the result `small`.
- Preserve log** checkbox: Toggled on.

```
import scrapy

class QuotesSpider(scrapy.Spider):
    name = "quotes"
    start_urls = [
        'http://quotes.toscrape.com/tag/humor/',
    ]
    1

    def parse(self, response):
        for quote in response.css('div.quote'):
            yield {
                'author': quote.xpath('span/small/text()').extract_first(),
                'text': quote.css('span.text::text').extract_first(),
            }
            2
            3
    4     next_page = response.css('li.next a::attr("href")').extract_first()
        if next_page is not None:
            next_page = response.urljoin(next_page)
            yield scrapy.Request(next_page, callback=self.parse)
```



scrapy runspider quotes_spider.py -o quotes.json

```
[  
  {  
    "text": "\u201cThe person, be it gentleman or lady, who has not pleasure in a good novel, must be intolerably stupid.\u201d",  
    "author": "Jane Austen"  
  },  
  
  {  
    "text": "\u201cThe reason I talk to myself is because I\u2019m the only one whose answers I accept.\u201d",  
    "author": "George Carlin"  
  },  
  {  
    "text": "\u201cI am free of all prejudice. I hate everyone equally. \u201d",  
    "author": "W.C. Fields"  
  },  
  {  
    "text": "\u201cA lady's imagination is very rapid; it jumps from admiration to love, from love to matrimony in a moment.\u201d",  
    "author": "Jane Austen"  
  }  
]
```



Dice.com

Inspect the page

Goal : Get Title and Location of Salesforce Developers in CA

Dice Tech Careers Career Center Insights Post Jobs Talent Solutions Login/Register

job title, skills, keywords or company name zip code, city or state Find Tech Jobs Advanced Search

Full-Time Part-Time Contracts Third Party

Recent Searches: Tableau Jobs In CA, Salesforce Developer Jobs In CA, Data Visualization Jobs In CA, Tableau Jobs In Palo Alto, CA

Salesforce Developer jobs in CA 1 - 30 of 5,013 positions

Sort by: relevance | date

MAP OUT YOUR NEXT OPPORTUNITY A new way to find jobs on the Dice Careers App. Discover

Sr Salesforce developer **** No Phone calls Please **** No Third Party Agencies Please**** Senior Sales... ZenFair Solutions Inc San Francisco, CA 5 days ago Easy Apply

Salesforce Developer Salesforce Developer Metrolink \$65,209.00 - \$101,896.00 Annually The SCRR... Metrolink Los Angeles, CA 2 days ago

Senior Salesforce Developer ICONIQ Capital is a global multi-family office/merchant bank, representing a grou...

?

**Output should be
csv file with 2
columns**

Title, Location

```
[sudhirwadhwa ~/Desktop/tbd/SCU $ more TableauJobsLocations.csv
Title,Location
Sr Salesforce developer,"San Francisco, CA"
Salesforce Developer,"Los Angeles, CA"
Senior Salesforce Developer,"San Francisco, CA"
Salesforce Developer,"Los Angeles, CA"
Senior Salesforce Developer - Specialist Senior,"Los Angeles, CA"
Senior Salesforce Developer - Specialist Senior,"Los Angeles, CA"
Sr. Salesforce Developer,"San Bruno, CA"
Salesforce Developer,"San Francisco, CA"
Salesforce Developer,"San Ramon, CA"
Salesforce Developer,"San Francisco, CA"
Senior Salesforce Developer,"Torrance, CA"
Senior Salesforce Developer - (building backend systems),"Costa Mesa, CA"
Sr. Salesforce Developer,"Santa Clara, CA"
Sr. Salesforce developer,"San Jose, CA"
Salesforce Developer,"Vista, CA"
SalesForce Developer,"San Francisco, CA"
Salesforce Developer,"San Rafael, CA"
Salesforce Developer,"San Francisco, CA"
Salesforce Developer,"San Diego, CA"
Senior Salesforce Developer - Wealth,"San Francisco, CA"
Salesforce Developer (Apex),"San Mateo, CA"
Senior Salesforce Developer,"Berkeley, CA"
Senior Salesforce Developer (API integration),"San Mateo, CA"
Salesforce Developer,"Mountain View, CA"
Sr. Salesforce Developer,"San Mateo, CA"
Salesforce Developers,"San Francisco, CA"
SalesForce Developer,"Santa Clara, CA"
Sr. Salesforce Developer,"San Francisco, CA"
Sr.Salesforce Developer (Having a LinkedIn is key),"San Francisco, CA"
Senior Salesforce Developer - Senior Consultant,"Los Angeles, CA"
Sr Salesforce developer,"San Francisco, CA"
Salesforce Developer,"Los Angeles, CA"
```



Dice® Tech Careers Career Center Insights Post Jobs Talent Solutions Login/Register

job title, skills, keywords or company name zip code, city or state Find Tech Jobs

Full-Time Part-Time Contracts Third Party Advanced Search

Recent Searches

- Tableau Jobs In CA
- Salesforce Developer Jobs In CA
- Data Visualization Jobs In CA
- Tableau Jobs In Palo Alto, CA

Salesforce Developer jobs in CA
1 - 30 of 5,013 positions

Sort by: relevance | date

Create Job Alert

Filter Results Reset

Company Segment + Sr Salesforce developer

MAP OUT YOUR NEXT OPPORTUNITY A new way to find jobs on the Dice Careers App. Discover

Elements Console Sources Network Timeline Profiles Application Security Audits Ember AdBlock

Styles Computed Event Listeners DOM Breakpoints

Filter :hover .cls

element.style { font-size: 18px; }

.serp-result-content job-search.min.css:1

.loggedInVisited:visited, .serp-result-content.loggedInVisited:visited .job-status-indicator.visited { color: #735aa5; }

.dice-btn-link { job-search.min.css:1 }

Console

Share your observation

Approach

Start coding

Slides are not for Distribution

```
1 ## (C) DataTiles.ai
2 ## (C) DataTiles.io
3 ## This is Proof of concept script, please do not use in production
4 ## Sudhir Wadhwa, Jyoti Wadhwa, January 2016
5
6 import bs4 as bs
7 import csv
8 import requests
9 holder = dict()
10
11 try:
12     # For Python 3.0 and later
13     from urllib.request import urlopen
14 except ImportError:
15     # Fall back to Python 2's urllib2
16     from urllib2 import urlopen
17
18 sourcehtml = urlopen("https://www.dice.com/jobs?q=Salesforce+Developer&l=CA")
19 soup = bs.BeautifulSoup(sourcehtml,"lxml")
20
21
22 with open('TableauJobsLocations.csv', 'w') as csvfile:
23     fieldnames = ['Title','Location']
24     jobwriter = csv.DictWriter(csvfile, fieldnames=fieldnames,dialect="excel",lineterminator='\n')
25     jobwriter.writeheader()
26     counter = 0
27     for a in soup.find_all('a', {"class": "dice-btn-link"}, href=True):
28         url = a['href']
29         if url.find('jobs/detail') > 0:
30             response=requests.get(url)
31             soup=bs.BeautifulSoup(response.text)
32             jobDesc = soup.find("div", { "id" : "jobdescSec" }).get_text().encode('ascii','ignore').upper()
33             holder['Title'] = soup.find("h1", { "class" : "jobTitle" }).get_text().encode('ascii','ignore').strip()
34             holder['Location'] = soup.find("li", { "class" : "location" }).get_text().encode('ascii','ignore').strip()
35             jobwriter.writerow(holder)
36             print ("....")
37             holder.clear
```

You have
Title and Location

Data is available.
You can use
Tableau or D3
now

```
[sudhirwadhwa ~/Desktop/tbd/SCU $ more TableauJobsLocations.csv
Title,Location
Sr Salesforce developer,"San Francisco, CA"
Salesforce Developer,"Los Angeles, CA"
Senior Salesforce Developer,"San Francisco, CA"
Salesforce Developer,"Los Angeles, CA"
Senior Salesforce Developer - Specialist Senior,"Los Angeles, CA"
Senior Salesforce Developer - Specialist Senior,"Los Angeles, CA"
Sr. Salesforce Developer,"San Bruno, CA"
Salesforce Developer,"San Francisco, CA"
Salesforce Developer,"San Ramon, CA"
Salesforce Developer,"San Francisco, CA"
Senior Salesforce Developer,"Torrance, CA"
Senior Salesforce Developer - (building backend systems),"Costa Mesa, CA"
Sr. Salesforce Developer,"Santa Clara, CA"
Sr. Salesforce developer,"San Jose, CA"
Salesforce Developer,"Vista, CA"
SalesForce Developer,"San Francisco, CA"
Salesforce Developer,"San Rafael, CA"
Salesforce Developer,"San Francisco, CA"
Salesforce Developer,"San Diego, CA"
Senior Salesforce Developer - Wealth,"San Francisco, CA"
Salesforce Developer (Apex),"San Mateo, CA"
Senior Salesforce Developer,"Berkeley, CA"
Senior Salesforce Developer (API integration),"San Mateo, CA"
Salesforce Developer,"Mountain View, CA"
Sr. Salesforce Developer,"San Mateo, CA"
Salesforce Developers,"San Francisco, CA"
SalesForce Developer,"Santa Clara, CA"
Sr. Salesforce Developer,"San Francisco, CA"
Sr.Salesforce Developer (Having a LinkedIn is key),"San Francisco, CA"
Senior Salesforce Developer - Senior Consultant,"Los Angeles, CA"
Sr Salesforce developer,"San Francisco, CA"
Salesforce Developer,"Los Angeles, CA"
```



Salesforce case study

End to End Pipeline with Salesforce Jobs example

Article on How final project looks like

Refer to Article

<https://software.intel.com/en-us/articles/using-visualization-to-tell-a-compelling-data-story>



Create Pipeline ,
Write Python Script to Ingest data from dice.com
Open Tableau,
Menu walk through
Connect with SalesforceJobs.csv
Create HeatMaps, ShowMe



Connecting to Tableau



Connecting to Tableau

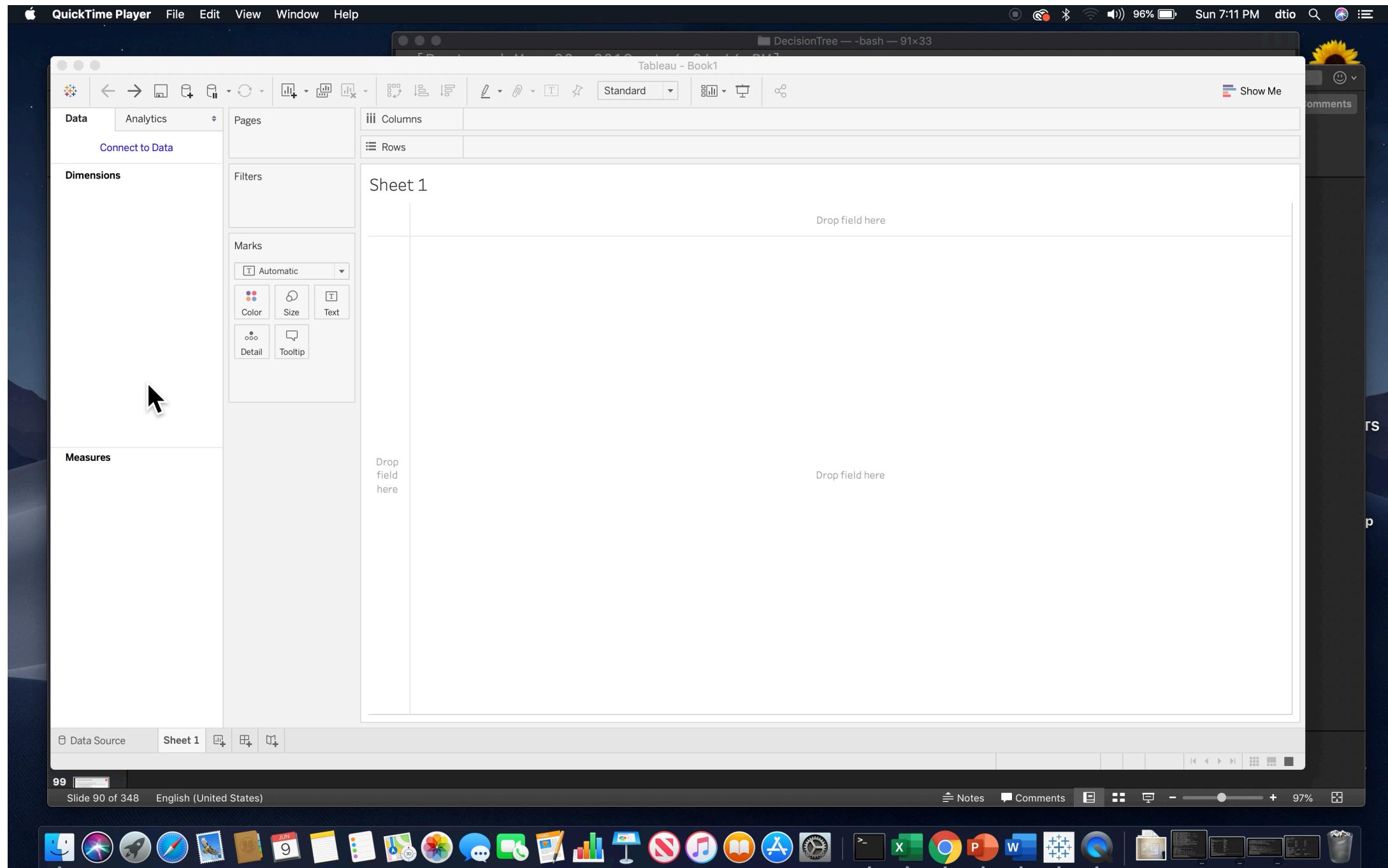
- Now Tableau brings us to the data connection window.
- Here we can see the name of the file – and here we can click to rename the connection if desired
- Down below, we can see all the sheets in the Excel file.
 - Sheets in Excel are treated the same as tables in databases, and we can choose to connect to a single table or join multiple tables.
- Simply drag a sheet into the data connection canvas.
 - Tables can be renamed simply by clicking on the name. The gear icon brings up options related to the data source.
- We can see the data down in the preview pane.



Connecting to Tableau

- There's a lot we can do on this screen (we will go over each of these in detail later in the class):
- If our column names aren't ideal, we can click on the drop down arrow to the right of the name and select rename.
- Clicking on the data type icon allows us to change the default data type for that column.
- If a field contains data that is concatenated, like Order ID and we'd like to split it, we can do that with the Split option – either custom or automatic.

Split





Live versus Extract

Live versus Extract



Something to consider before we begin analyzing our data is if we want to connect live or extract.

- Connecting live leaves the data in the database or source file.
- This is best when we want to leverage a high performance database's capabilities, or to get up-to-the-second changes in data visualized in Tableau.
- That being said, sometimes connecting live can result in a slow experience, depending on the database.



Live versus Extract

The other option is to extract the data into Tableau's high performance in-memory data engine.

- This can help when connecting to a slow database or to take query load off critical systems.
- We can also import only some of the data and bring in specific elements (to access those options, click Edit)
- For now, we'll connect live.

Some data sources may have limitations regarding connecting live or being extracted – for example, most OLAP data sources cannot be extracted, and some cloud-based data sources must be extracted.



Live versus Extract



Click on the first Sheet tab to open the canvas. If you ever want to go back to the data connection window, simply click on the Data Source tab.

Note that Dimensions are discrete and Measures are continuous.
We will talk more about this in class.



Quick Viz Example

Double-click Sales – this gives a single bar. Tableau has aggregated the data. What if you want to see more detail? If I want to see Sales at a more granular level, click **Analyze, uncheck Aggregate Measures** (If you want, you can reduce the size of the circles). Now we have more detail – this is an important feature! For example:

We now see how many marks there are

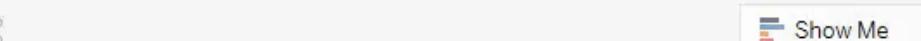
We can **right-click a mark** (or several marks), **View Data**, Underlying tab, and see the actual data (nice for investigating outliers, etc.)

Go back and **check Aggregate Measures**, and bring out **Category** (double-click). This gives **a total (SUM)**, which is probably what we want, but we can easily change it to see the average sales price in each category (change to **AVERAGE**). Sort the bar chart. Use Marks card to adjust color and size, add/remove labels, etc.

Finally, drag out Sales, and drag Number of Records to the shelf. It is automatically calculated for you.



File Data Worksheet Dashboard Story Analysis Map Format Server Window Help



Show Me

Data

Analytics

Returns (Global Superst...)

Search



Dimensions

- Abc Category
- ⊕ City
- ⊕ Country
- Abc Customer ID
- Abc Customer Name
- Abc Market
- ⊕ Order Date
- Abc Order ID
- Abc Order ID - Split 1
- # Order ID - Split 2
- # Order ID - Split 3
- Abc Order Priority
- ⊕ Postal Code
- Abc Product ID
- Abc Product Name
- Abc Region
- Abc Segment

Measures

- # Discount
- # Profit
- # Quantity
- # Row ID
- # Sales (highlighted)
- # Shipping Cost
- ⊕ Latitude (generated)
- ⊕ Longitude (generated)
- # Number of Records
- # Measure Values

Data Source

Sheet 1



Pages

Columns

Rows

Sheet 1

Drop field here

Drop field here

Drop field here



Select or drag data

Use the Shift or Ctrl key to select multiple fields



Slides are not for Distribution

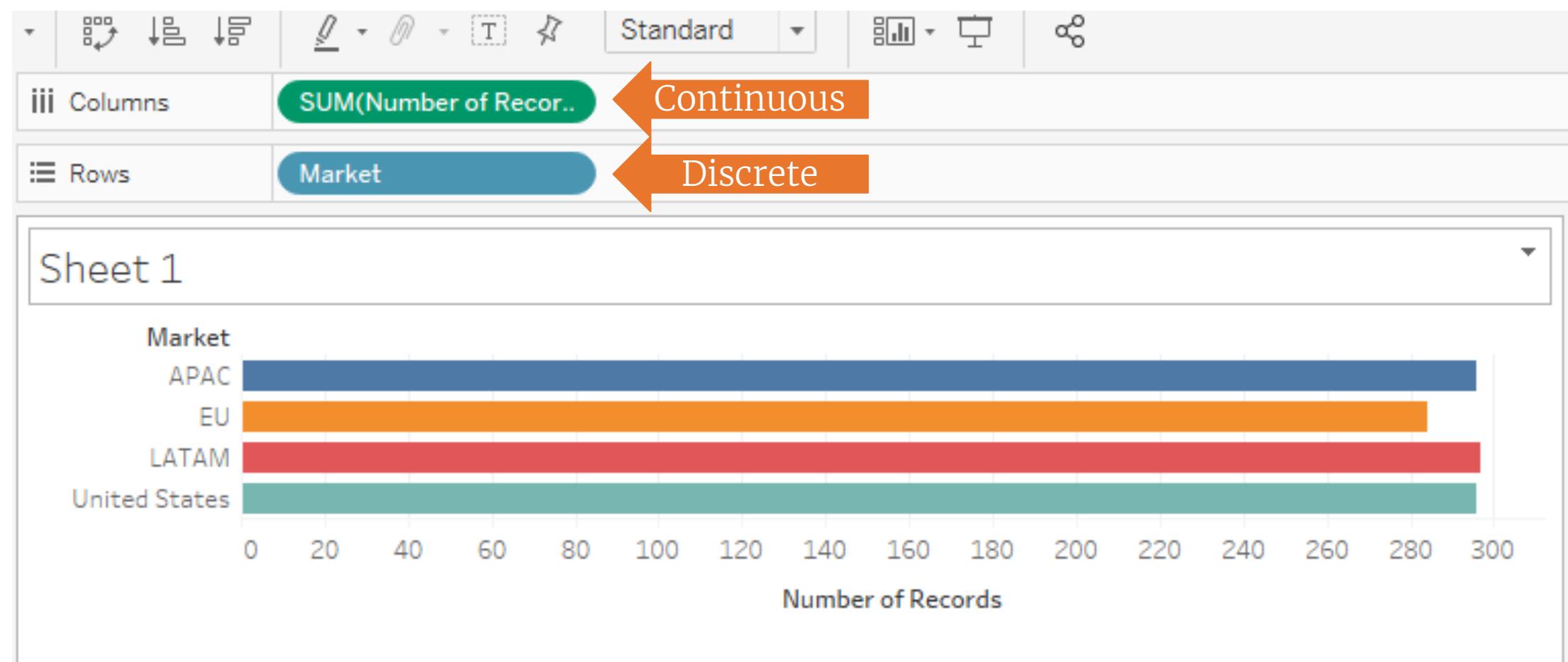


Connecting to Multiple Tables

Paying Attention to Pill Color

When we bring a field into the view from the Data Window, Tableau creates a pill.

The pill color indicates whether the pill is continuous or discrete. Discrete pills are **BLUE**, Continuous pills are **GREEN**.





Connecting to Multiple Tables

Next, what if we realize we need to bring in additional data?

- To add columns from *other tables* in the *same data source* (i.e. other worksheets in our Excel file), we need to edit our data connection.
- To do so, **click on the Data Source tab**.
- Let's join our *returns table* to the *orders table*.
 - Double click or drag out Returns.



Connecting to Multiple Tables

- Click on the join icon to show the details.
- The default join clause is shown here.
- Tableau has figured out that Market is a common field between these two tables.
- If desired, we could edit the join clause, or even create a new one
- For these data we could also select a left join (explain the difference).
- For now we'll change it to Left and close the dialog.

Again, we can view the data down here and verify the data we'll be connecting to.

It looks good, so we'll click back onto our sheet.

We now have columns from both Orders and Returns in our data window and are ready to start our analysis!

Navigation icons: Back, Forward, Home, Refresh, Undo, Redo, Standard View, Zoom In/Out, Share.

Toolbars: Data, Analytics, Pages, Columns, Rows, Show Me.

Search bar: Returns (Global Superst...).

Dimensions: Category, City, Country, Customer ID, Customer Name, Market, Order Date, Order ID, Order ID - Split 1, Order ID - Split 2, Order ID - Split 3, Order Priority, Postal Code, Product ID, Product Name, Region, Segment.

Measures: Discount, Profit, Quantity, Row ID, Sales, Shipping Cost, Latitude (generated), Longitude (generated), Number of Records, Measure Values.

Sheet 1: Category

Avg. Sales

Category	Avg. Sales
Technolo..	467.9
Furniture	416.2
Office Supplies	121.1

For horizontal bars try:

0 or more Dimensions

1 or more Measures

Data Source: Sheet 1

3 rows, 3 columns, SUM of AVG(Sales): 1,005.2





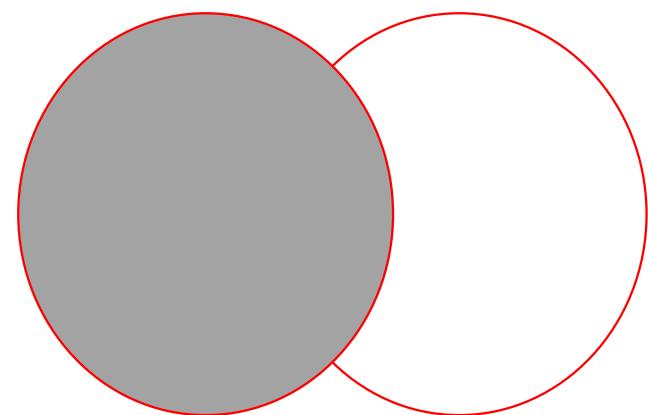
Left Join, Right join, Inner join and outer join

Student Name	class
Asha	MSIS 2626
Anna	MSIS 2603
Nancy	MSIS 2606
Adam	MSBA2639

Student Name	class
Asha	Business
Anna	Technology
Nancy	Finance
John	Finance
Sam	Business

Student Name	class
Asha	MSIS 2626
Anna	MSIS 2603
Nancy	MSIS 2606
Adam	MSBA 2639

Left Join



Student Name	Major_in
Asha	Business
Anna	Technology
Nancy	Finance
John	Finance
Sam	Business

Left Table Leads

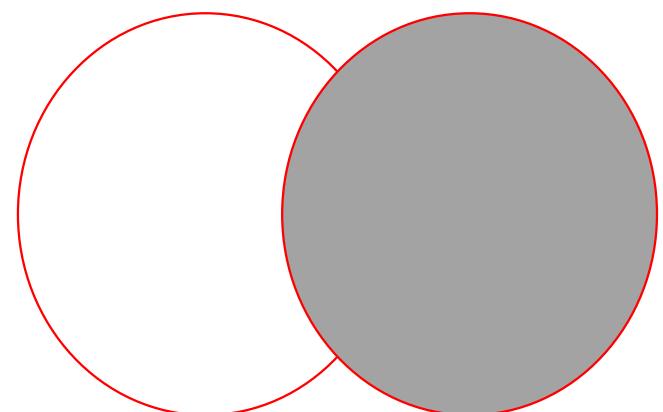
Student Name	class	Major_in
Asha	MSIS 2626	Business
Anna	MSIS 2603	Technology
Nancy	MSIS 2606	Finance
Adam	MSBA 2639	-

Student Name	class
Asha	MSIS 2626
Anna	MSIS 2603
Nancy	MSIS 2606
Adam	MSBA 2639

D

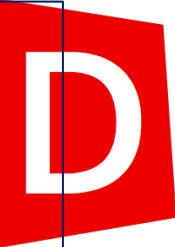
Student Name	Major_in
Asha	Business
Anna	Technology
Nancy	Finance
John	Finance
Sam	Business

Right Join



Right Table Leads

Student Name	class	Major_in
Asha	MSIS 2626	Business
Anna	MSIS 2603	Technology
Nancy	MSIS 2606	Finance
John	-	Finance
Sam	-	Business



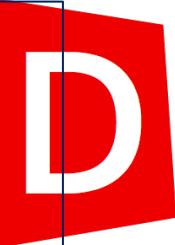
Student Name	class
Asha	MSIS 2626
Anna	MSIS 2603
Nancy	MSIS 2606
Adam	MSBA 2639

Student Name	Major_in
Asha	Business
Anna	Technology
Nancy	Finance
John	Finance
Sam	Business

Inner Join

INTERSECTION

Student Name	class	Major_in
Asha	MSIS 2626	Business
Anna	MSIS 2603	Technology
Nancy	MSIS 2606	Finance



Student Name	class
Asha	MSIS 2626
Anna	MSIS 2603
Nancy	MSIS 2606
Adam	MSBA 2639

Student Name	Major_in
Asha	Business
Anna	Technology
Nancy	Finance
John	Finance
Sam	Business

Outer Join

Everyone

Student Name	class	Major_in
Asha	MSIS 2626	Business
Anna	MSIS 2603	Technology
Nancy	MSIS 2606	Finance
Adam	MSBA 2639	-
John	-	Finance
Sam	-	Business



Poorly Formatted Data



Poorly Formatted Data

In our first example of connecting to data, we used an Excel file that was already formatted nicely – we just brought it straight into Tableau. In reality, data files are not always so well-formatted. Now we will practice using a file that needs some work before it is ready for analysis.

For this example we'll be using the “Data Prep – Flights” Excel file, which you can download from the course website. (Global Superstore, our main dataset, is too well-structured! This gives us a messier example to work with.)



Poorly Formatted Data

	A	B	C	D	E
1	Date	Employee	Resolved Incidents		
2	1/1/2014	B-002		4	
3	1/1/2014	E-055		1	
4	1/1/2014	E-075		14	
5	1/1/2014	B-066		4	
6	1/1/2014	C-025		17	
7	1/1/2014	E-030		2	
8	1/1/2014	C-001		14	
9	1/1/2014	E-038		4	
10	1/1/2014	C-054		2	
11	1/1/2014	A-081		3	
12	1/1/2014	B-031		14	
13	1/1/2014	D-019		2	
14	1/1/2014	E-096		2	
15	1/1/2014	D-026		0	
16	1/1/2014	F-022		3	

Open the file in Excel. Here we have a report in Excel, showing the number of resolved incidents per Employee per month. The “Ideal” tab shows how we wish the data would be formatted – like a database table.



Poorly Formatted Data

Flights Data Summary B1					
This report was generated on 1-1-15					
Employee	1/1/2014	2/1/2014	3/1/2014	4/1/2014	5/1/2014
B-002	4	1	5	2	3
E-055	1	2	1	3	4
E-075	14	17	16	15	18
B-066	4	4	5	2	5
C-025	17	13	17	18	17
E-030	2	2	1	1	0
C-001	14	14	14	14	13
E-038	4	1	0	4	0
C-054	2	5	4	4	2
A-081	3	2	4	5	2

However, sometimes we receive data that looks more like what we see in the “Resolved Incidents” tab. Luckily, there are several features in Tableau Desktop to help automatically reshape Text and Excel files to get them ready for analysis in Tableau.



Poorly Formatted Data

Let's connect to this Excel file and see if we can work with that poorly formatted sheet.

- In Tableau Desktop, click on Excel, navigate to where you saved the file and click open.
- Drag out the “Resolved Incidents” sheet.



Connect

To a File

Excel

Text file

Access

JSON file

PDF file

Spatial file

Statistical file

More...



To a Server

Tableau Server

MySQL

Oracle

Amazon Redshift

Microsoft SQL Server

More...

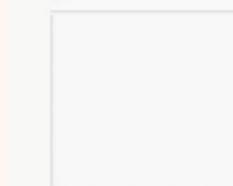


Saved Data Sources

Sample - Superstore

RECORDED WITH
SCREENCASTOMATIC

Open



Global Superstore



SUMMER TEST ...



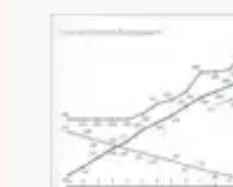
Enrollment Com...



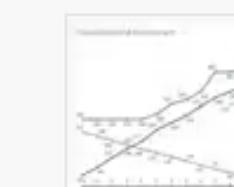
Academic Depar...



Academic Depar...



Course Assessm...



Course Assessm...



Analysis of Degr...



ECDB Sankey



12 Term Major C...



Program Array ...



Program Array ...



Enrollment Com...

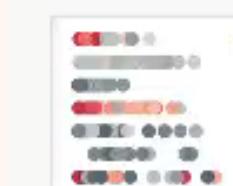


Enrollment Clas...



HG Grads - Post ...

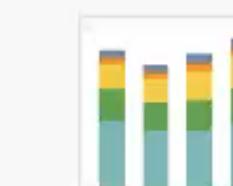
Sample Workbooks



Superstore



Regional



World Indicators

More Samples

Open a Workbook

Discover

Training

Getting Started

Connecting to Data

Visual Analytics

Understanding Tableau

More training videos...

Viz of the Week

Data Role T Models

Resources

Blog - Get Tableau 10.4 beta to better collaborate on trusted da

Tableau Conference 2017

Forums



Data Interpreter

Data Interpreter



Although Tableau can connect to this sheet, we can see here in the preview that there are some issues. There are no column names, the headers from Excel have a lot of nulls, and so on. Tableau has also recognized this and suggests the Data Interpreter (Tableau's built-in tool for preparing your data for analysis).

- Click turn on



Connections

Add

Copy of Data Prep - Flights
Excel

Sheets

 Use Data Interpreter

Data Interpreter might be able to
clean your Excel workbook.

Ideal

Irregular Delimiter

Resolved Incidents

Tiers

New Union

B - Resolved Incidents (Copy of Data Prep - Flights)

Connection

 Live Extract

Filters

0 | Add

Resolved Incidents

 Sort fields
 Data source order

 Show aliases
 Show hidden fields

Abc Resolved Incidents	Abc Resolved Inci...											
F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	
Flights Data Summary	null											
This report was gener...	null											
Employee	1/1/2014	2/1/2014	3/1/2014	4/1/2014	5/1/2014	6/1/2014	7/1/2014	8/1/2014	9/1/2014	10/1/2014	11/1/2014	
B-002	4	1	5	2	3	0	3	1	2	0	2	2
E-055	1	2	1	3	4	1	4	0	2	1	1	4
E-075	14	17	16	15	18	16	14	17	12	13	14	
B-066	4	4	5	2	5	0	0	2	0	1	0	0
C-025	17	13	17	18	17	17	12	15	17	17	14	
E-030	2	2	1	1	0	3	5	5	0	2	4	

RECORDED WITH

SCREENCASTOMATIC

Data Source

Go to Worksheet

Sheet 1



Data Interpreter



- Now we see that those headers and nulls have been stripped out, and our columns are properly identified!



Data Interpreter

- If we want more specifics on what the Data Interpreter did, we can click “Review Results” on the right. This will open an Excel file describing the changes.
- If we click to the tab we used, Resolved Incidents, we see which fields are being used as headers, in red, and which are considered data, in green



Resolved Incidents (Copy of Data Prep - Flights)

Connection
 Live
 Extract

Filters
 0 | Add

Connections Add

Copy of Data Prep - Flights
 Excel

Sheets o

Cleaned with Data Interpreter
[Review the results](#). (To undo changes, clear the check box.)

Ideal

Irregular Delimiter

Resolved Incidents

Tiers

New Union

Resolved Incidents

Sort fields Data source order ▼ Show aliases Show hidden fields 26 →

Abc	#	#	#	#	#	#	#	#	#	#
Resolved Incidents	Resolved In									
Employee	1/1/2014	2/1/2014	3/1/2014	4/1/2014	5/1/2014	6/1/2014	7/1/2014	8/1/2014	9/1/2014	Resolved In
B-002	4	1	5	2	3	0	3	1	4	1
E-055	1	2	1	3	4	1	4	0	4	0
E-075	14	17	16	15	18	16	14	17	17	17
B-066	4	4	5	2	5	0	0	0	0	2
C-025	17	13	17	18	17	17	12	12	15	15
E-030	2	2	1	1	0	3	5	5	5	5
C-001	14	14	14	14	13	18	17	17	14	14
E-038	4	1	0	4	0	2	5	5	0	0
C-054	2	5	4	4	2	3	0	0	5	5

Go to Worksheet x

RECORDED WITH

SCREENCAST

MATIC

Data Source

Sheet1





Data Interpreter

	A	B	C	D	E
1	Date	Employee	Resolved Incidents		
2	1/1/2014	B-002		4	
3	1/1/2014	E-055		1	
4	1/1/2014	E-075		14	
5	1/1/2014	B-066		4	
6	1/1/2014	C-025		17	
7	1/1/2014	E-030		2	
8	1/1/2014	C-001		14	
9	1/1/2014	E-038		4	
10	1/1/2014	C-054		2	
11	1/1/2014	A-081		3	
12	1/1/2014	B-031		14	
13	1/1/2014	D-019		2	
14	1/1/2014	E-096		2	
15	1/1/2014	D-026		0	
16	1/1/2014	F-022		3	

Before we go back to Tableau and our data connection, let's take one more look at that "Ideal" tab. Note that instead of having a column for each month with data underneath, in this format, there is a "Date" column and each row contains the number of resolved incidents for each unique combination of date and employee. This data is in the preferred format for analysis: taller, with more rows, rather than wider, with more columns. Let's see if we can do that in Tableau.

Slides are not for Distribution



Pivot

Pivot



Back in Tableau, we want to change the format from that column-per-month layout into a single date column and a single column for Resolved Incidents.

- To do this easily, we'll simply select all the date columns. Click on the first, scroll if necessary, then shift click on the last. We'll open the menu and select “Pivot”
- This pivot feature essentially merges the information from the original columns and rows into two new columns – Pivot field names, and Pivot field values.
- We can see that “Pivot field names” is actually our Date, so we can click to open the menu and select rename.
- Similarly, “Pivot field values” can be renamed “Resolved Incidents”



Connections Add

Copy of Data Prep - Flights
Excel

Sheets

 Cleaned with Data Interpreter
[Review the results](#). (To undo changes, clear the check box.)

Ideal

Irregular Delimiter

Resolved Incidents

Tiers

New Union

Resolved Incidents (Copy of Data Prep - Flights)

Connection
 Live
 ExtractFilters
0 | Add

Resolved Incidents

Sort fields Data source order ▾
 Show aliases Show hidden fields 26 ➔

Employee	1/1/2014	2/1/2014	3/1/2014	4/1/2014	5/1/2014	6/1/2014	7/1/2014	8/1/2014	9/1/2014
B-002	4	1	5	2	3	0	3	1	1
E-055	1	2	1	3	4	1	4	0	0
E-075	14	17	16	15	18	16	14	17	17
B-066	4	4	5	2	5	0	0	0	2
C-025	17	13	17	18	17	17	12	15	15
E-030	2	2	1	1	0	3	5	5	5
C-001	14	14	14	14	13	18	17	17	14
E-038	4	1	0	4	0	2	5	5	0
C-054	2	5	4	4	2	3	0	0	5

RECORDED WITH

SCREENCAST

MATIC

Data Source

Sheet 1



Split

Split



There's one more thing we can do to prepare this data. Note that the "Employee" field is actually two pieces of information – a location code, A, B, C, D, or E, followed by an Employee ID number. We can split the column based on the hyphen delimiter:

- Click to open the menu and select Split
- There are now two new fields – Employee – Split 1 and Split 2
- We'll use the Metadata Grid view (click the icon to the left of Sort Fields) to rename our split fields

Split



- Click on the name to edit in-line, Split 1 should be Location, and we'll hit tab, Split 2 should be Employee ID
- There's an Abc next to the Date field indicating this column is considered a String. We know it's actually a Date, though, so we can click on the Abc and select Date to update the data type.
- Now if we click on Sheet 1, we'll see nice tidy data ready for analysis!

Split

File Data Server Window Help

Connections Add

Copy of Data Prep - Flights Excel

Cleaned with Data Interpreter
[Review the results](#). (To undo changes, clear the check box.)

Ideal

Irregular Delimiter

Resolved Incidents

Tiers

New Union

Resolved Incidents (Copy of Data Prep - Flights)

Connection Live Extract

Filters 0 | Add

Resolved Incidents

Sort fields Data source order ▾

Show aliases Show hidden fields 312 →

Date	Resolved Incidents	Employee
1/1/2014	4	B-002
1/1/2014	1	E-055
1/1/2014	14	E-075
1/1/2014	4	B-066
1/1/2014	17	C-025
1/1/2014	2	E-030
1/1/2014	14	C-001
1/1/2014	4	E-038
1/1/2014	2	C-054

RECORDED WITH SCREENCASTOMATIC

Go to Worksheet ×

2017 Data Source Sheet 1

Slides are not for Distribution

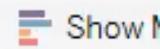


Custom Split

Custom Split



- Let's create a viz now: bring Employee ID to the view, Resolved Incidents to Columns, and sort it.
- It's clear from this view that there are really two groups of employees – some who resolve a much higher number of incidents than others. Looks like some employees are often able to get through more cases, and they have a Tier II designation.



Data Analytics

Resolved Incidents (Cop...)

Dimensions

Abc Date

Abc Employee

=# Employee ID

=Abc Location

Abc Measure Names

Measures

Resolved Incidents

-# Number of Records

Measure Values

Pages

Columns

Rows

Filters

Sheet 1

Drop field here

Drop field here

Drop field here

Marks

Automatic

Color

Size

Text

Detail

Tooltip



Custom Split

A	B	C	D	E	F	G	H	I	J	K	L	M	N
Employee	1/1/2014	2/1/2014	3/1/2014	4/1/2014	5/1/2014	6/1/2014	7/1/2014	8/1/2014	9/1/2014	#####	#####	#####	Header
B-002	4	1	5	2	3	0	3	1	2	0	2	5	Data
E-055	1	2	1	3	4	1	4	0	2	1	4	0	Data
E-075-II	14	17	16	15	18	16	14	17	12	13	14	12	Data
B-066	4	4	5	2	5	0	0	2	0	1	0	3	Data
C-025-II	17	13	17	18	17	17	12	15	17	17	14	15	Data
E-030	2	2	1	1	0	3	5	5	0	2	4	1	Data
C-001-II	14	14	14	14	13	18	17	14	13	18	15	14	Data
E-038	4	1	0	4	0	2	5	0	2	2	2	2	Data
C-054	2	5	4	4	2	3	0	5	5	5	3	5	Data
A-081	3	2	4	5	2	2	2	4	1	4	2	0	Data
B-031-II	14	14	14	14	15	13	15	14	12	16	12	18	Data
D-019	2	3	0	0	4	4	1	2	5	0	5	5	Data
E-096	2	0	4	4	5	3	3	0	5	4	2	0	Data
D-026	0	2	0	2	5	3	1	0	0	2	5	4	Data
E-022	3	3	4	3	4	2	0	3	2	3	3	1	Data
C-015	1	5	3	5	2	1	3	3	1	1	5	2	Data

If we look at our original data set in Excel, we see there's a tab called Tiers. This report adds a –II to the end of an employee ID if they're tier II. Because not all rows have this –II, a standard split won't work. Let's see if we can create a viz that incorporates this Tier designation.

[[side note: both Split and Custom Split require consistent delimiters. If our data has irregular delimiters, Tableau won't be able to split out the data using these options.]]

Custom Split



- Open a new Tableau file and recreate the viz, this time using the Tiers sheet from Excel.
- Remember to use the Data Interpreter, and Pivot the dates again
- Click on the Employee column to open the menu and select Custom Split
 - We can choose our delimiter, we'll use a hyphen
 - And now we can say we want to have 3 columns
 - This forces Tableau to break off that 3rd column with the tier II indicator
 - Now students should all try to finish the viz by themselves: rename all columns, and do your bar chart as above, and this time color the bars by Tier.

Custom Split

File Data Server Help

Connect

To a File

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

To a Server

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

Sample Workbooks

- Superstore
- Regional
- World Indicators

RECORDED WITH
SCREENCASTOMATIC

Open

- Step 1 Viz
- Global Superstore
- SUMMER TEST ...
- Enrollment Com...
- Academic Depar...
- Course Assessm...
- Course Assessm...
- Analysis of Degr...
- ECDB Sankey
- 12 Term Major C...
- Program Array ...
- Program Array ...
- Enrollment Com...
- Enrollment Clas...

Discover

Open a Workbook

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

Viz of the Week

Data Role T Models

Resources

Blog - Get Tableau 10.4 beta to better collaborate on trusted data

Tableau Conference 2017

Forums

© DataTiles Inc.



Complete Student Information System

Go to Folder > Class 3 STUDENTS INFORMATION SYSTEM

Each student will submit Dean Dashboard.



Top Charts

Complete Slides 182 to 211



Menus & Toolbars



Menus & Toolbar

- Let's use the Global Superstore Tableau workbook to explore various areas of the screen. You will have plenty of practice with this UI; this is just a quick orientation so you will know the basics and can start exploring for yourself.

New sheet tabs are found at the bottom. We can create sheets, dashboards, and stories with these tabs.

We can also do things like rename the sheets, drag to rearrange them, duplicate sheets, copy formatting, and many other things.

The screenshot shows the Tableau software interface. At the top, there's a menu bar with File, Data, Worksheet, Dashboard, Story, Analysis, Map, Format, Server, Window, and Help. Below the menu is a toolbar with various icons. On the left, there's a sidebar with sections for Data (Global Superstore), Dimensions (Market, Order ID, Returned, Measure Names), and Marks (Automatic, Color, Size, Text, Detail, Tooltip). The main workspace is titled "Sheet 1" and contains three empty columns labeled "Drop field here". At the bottom of the workspace, there are tabs for "Data Source" (circled in yellow) and "Sheet 1".

Sheet 1

Drop field here

Drop field here

Drop field here

iii Columns

iii Rows

Filters

Marks

Automatic

Color Size Text

Detail Tooltip

Measures

Number of Records

Measure Values

Data Analytics

Global Superstore

Dimensions

Market Order ID Returned Measure Names

Create a new sheet – this is a blank sheet. Sheets are where we build visualizations.

Data Source Sheet 1

RECORDED WITH

SCREENCASTOMATIC



Menus & Toolbar

- At the top, we have the **menus**. (The layout may look slightly different on a Mac.) The menus contain a lot of powerful controls – after today's class, every student should spend some time clicking through to see what options they contain.
- Below is the **toolbar**, with buttons like **undo** – there is no limit to how much you can undo, and this is a very important button that allows you to explore! (will be grayed out since we have not yet done anything)
- Here we also have **save** – there's no automatic save in Tableau, so make sure to save your work periodically.

Menus & Toolbar

Tableau - Book2

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Show Me

Data Analytics Global Superstore

Dimensions

- Market
- Order ID
- Returned
- Measure Names

Marks

- Automatic
- Color
- Size
- Text
- Detail
- Tooltip

Filters

Pages Columns Rows

Sheet 2

Drop field here

Drop field here

Drop field here

Measures

- Number of Records
- Measure Values

© DataTiles.io

RECORDED WITH SCREENCASTOMATIC

6:49 PM 7/24/2017

The screenshot shows the Tableau desktop application. The menu bar includes File, Data, Worksheet, Dashboard, Story, Analysis, Map, Format, Server, Window, and Help. The toolbar has various icons for navigation and data manipulation. The data pane on the left shows a connection to 'Global Superstore' with sections for Dimensions (Market, Order ID, Returned, Measure Names) and Measures (Number of Records, Measure Values). The visualization workspace is titled 'Sheet 2' and contains three blank drop zones labeled 'Drop field here'. A yellow circle highlights the cursor area in the bottom-left corner of the workspace. The bottom of the screen shows the Windows taskbar with icons for various applications like File Explorer, Edge, and Excel.



Data Window



Data Window

The screenshot shows the Data Window interface. At the top, there are tabs for 'Data' and 'Analytics'. The 'Data' tab is selected, indicated by a blue border. Below the tabs, a connection named 'Global Superstore' is listed. Under the 'Dimensions' section, there are four items: 'Market', 'Order ID', 'Returned', and 'Measure Names'. In the 'Measures' section, there are two items: 'Number of Records' and 'Measure Values'. The background of the window is light gray.

On the left of the screen is the **data window**. If we're on the data tab, the top lists all open data connections, and depending on which one is selected, the fields from that data source are listed below, broken out into dimensions and measures. The data window will also show any sets or parameters you may have.

If we open the map options or the format pane, these temporarily cover the data window. To get back to the data window, simply close the other pane.

The screenshot shows the Tableau interface with the 'Data' pane open on the left side. The 'Data' pane contains sections for 'Global Superstore', 'Dimensions' (Market, Order ID, Returned, Measure Names), and 'Measures' (Number of Records, Measure Values). The main workspace is titled 'Sheet 1' and has three empty drop zones labeled 'Drop field here'. A yellow circle highlights the cursor at the bottom center of the screen, which is positioned over the 'Data Source' tab in the bottom navigation bar. The bottom bar also includes icons for various applications like Excel, Word, and Powerpoint, along with system status indicators.

The data window can also be minimized like so, and then re-expanded.

This screenshot shows the Tableau interface with the 'Data' window minimized. The main workspace is labeled 'Sheet 1'. The 'Data' window is visible on the left side, containing sections for 'Dimensions', 'Filters', 'Marks', and 'Measures'. A yellow circle highlights the 'Data Source' tab at the bottom of the minimized window. The status bar at the bottom indicates 'RECORDED WITH SCREENCASTOMATIC'.

Tableau - Book2

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Data Analytics Global Superstore

Dimensions

- Market
- Order ID
- Returned
- Measure Names

Marks

- Automatic
- Color
- Size
- Text
- Detail
- Tooltip

Measures

- Number of Records
- Measure Values

Pages Columns Rows

Sheet 1

Drop field here

Drop field here

Drop field here

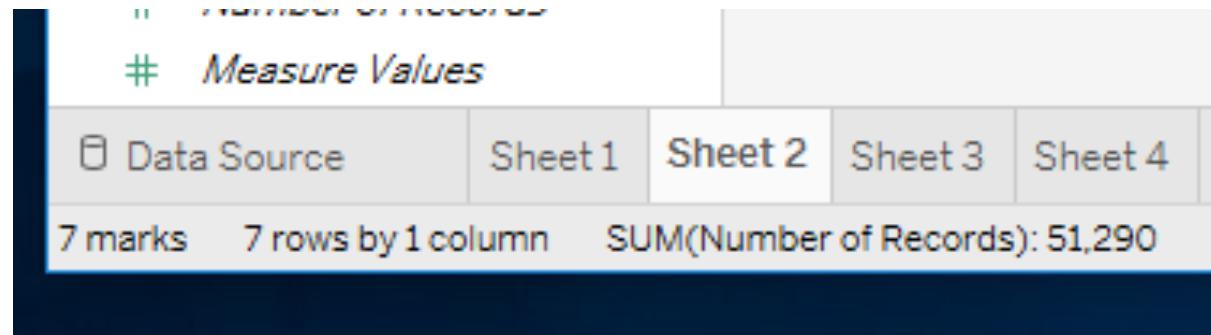
Data Source

Sheet 1 Sheet 2

RECORDED WITH SCREENCASTOMATIC

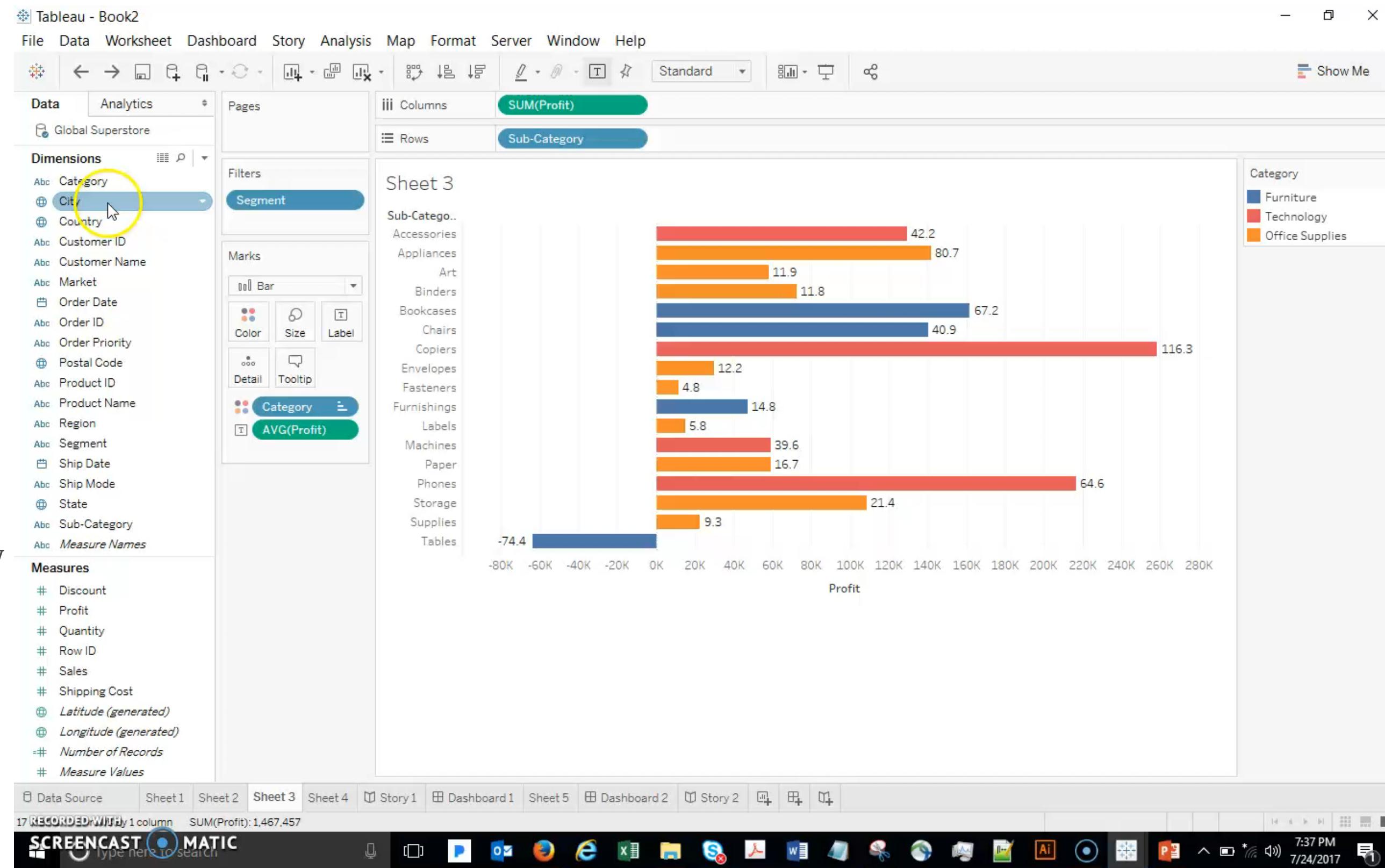
© DataTiles.io

Data Window



Down here at the bottom is the status bar. This shows the number of marks in the view as well as other summary information.

If we click to the Analytics tab, here, we're brought to an entirely new pane. Here we can bring out pieces of our analysis directly as drag and drop elements. We will get to these tools later in the course, especially things like trend lines and forecasts when we have time series data.





Shelves & Cards



Shelves & Cards

Finally, we have the shelves (or cards). A view can be built by dragging and dropping fields from the data window into the canvas directly, or onto the shelves.

- the Columns shelf and the Rows shelf (drag Profit to Columns and Sub-Category to Rows)
- the Pages shelf (topic for later)

Tableau - Book2

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Show Me

Data Analytics Global Superstore

Dimensions

- Abc Category
- Abc City
- Abc Country
- Abc Customer ID
- Abc Customer Name
- Abc Market
- Abc Order Date
- Abc Order ID
- Abc Order Priority
- Abc Postal Code
- Abc Product ID
- Abc Product Name
- Abc Region
- Abc Segment
- Abc Ship Date

Measures

- # Discount
- # Profit
- # Quantity
- # Row ID
- # Sales
- # Shipping Cost
- (#) Latitude (generated)
- (#) Longitude (generated)
- # Number of Records
- # Measure Values

Pages Columns Rows

Sheet 3

Drop field here

Marks

- Automatic
- Color
- Size
- Text
- Detail
- Tooltip

Drop field here

Drop field here

Data Source Sheet 1 Sheet 2 Sheet 3

Data Source

RECORDED WITH

Shelves & Cards



- Filter shelf – this is an important feature that you are likely to use a lot
 - Drag Segment to Filters as an example, and then **Show Filter** (formerly Show Quick Filter)
 - If you don't want the “All” option for some reason, just click the Quick Filter's drop-down menu and look under Customize.

Data Analytics

Global Superstore

Dimensions

- Abc Customer Name
- Abc Market
- Abc Order Date
- Abc Order ID
- Abc Order Priority
- Abc Postal Code
- Abc Product ID
- Abc Product Name
- Abc Region
- Abc Segment
- Abc Ship Date
- Abc Ship Mode
- Abc State
- Abc Sub-Category
- Abc Measure Names

Measures

- # Discount
- # Profit
- # Quantity
- # Row ID
- # Sales
- # Shipping Cost
- (#) Latitude (generated)
- (#) Longitude (generated)
- # Number of Records
- # Measure Values

Pages

Columns

SUM(Profit)

Rows

Sub-Category

Filters

Sheet 3

Sub-Catego..

Accessories

Appliances

Art

Binders

Bookcases

Chairs

Copiers

Envelopes

Fasteners

Furnishings

Labels

Machines

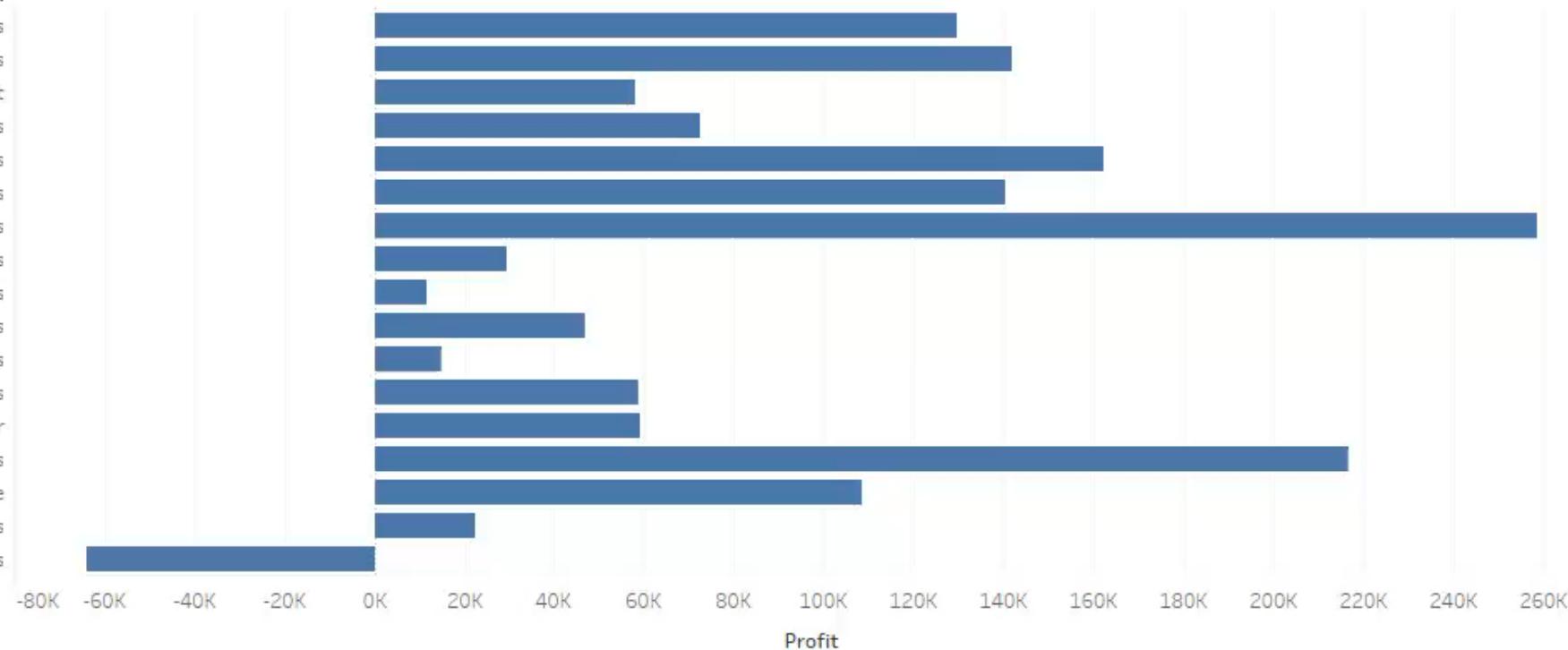
Paper

Phones

Storage

Supplies

Tables



Data Source

17 RECORDED WITH 1 column

SUM(Profit): 1,467,457

SCREENCAST MATIC

Type here to search





Shelves & Cards

- Marks Card
 - The Marks Card is made up of several other shelves, each of which can have fields placed on them and can be clicked on to edit their characteristics

Data Analytics

Global Superstore

Dimensions

- Abc Customer Name
- Abc Market
- Abc Order Date
- Abc Order ID
- Abc Order Priority
- Abc Postal Code
- Abc Product ID
- Abc Product Name
- Abc Region
- Abc Segment
- Abc Ship Date
- Abc Ship Mode
- Abc State
- Abc Sub-Category
- Abc Measure Names

Measures

- # Discount
- # Profit
- # Quantity
- # Row ID
- # Sales
- # Shipping Cost
- (# Latitude (generated))
- (# Longitude (generated))
- # Number of Records
- # Measure Values

Pages

SUM(Profit)

Rows Sub-Category

Filters

Segment

Sub-Catego..

Accessories

Appliances

Art

Binders

Bookcases

Chairs

Copiers

Envelopes

Fasteners

Furnishings

Labels

Machines

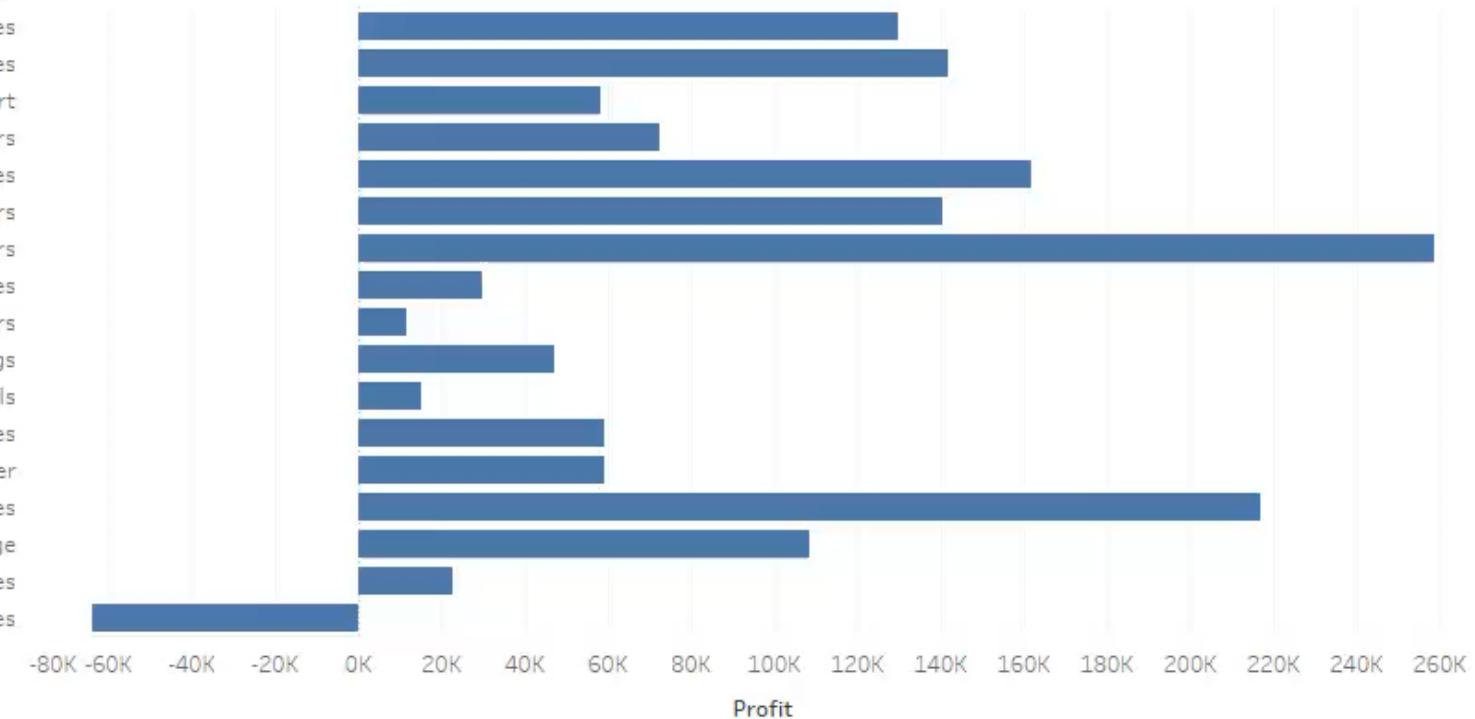
Paper

Phones

Storage

Supplies

Tables



Segment

(All)

Shelves & Cards

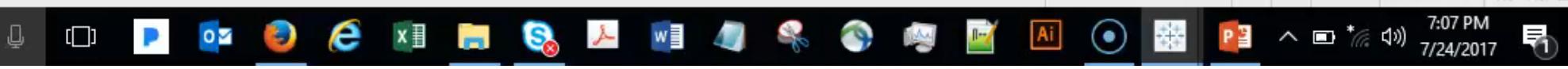
Data Source

17 RECORDED WITH 1 column

SUM(Profit): 1,467,457

SCREENCAST MATIC

Type here to search





Legends



Legends

Legends, such as for color, size, and shape, will automatically be created when a field is placed on the color, size or shape card. You can also change the order of fields in the viz by dragging them around in the legend.

Legends can be removed by clicking on the menu, and selecting hide card.

To bring a legend back, either right click in the space under the marks card, select Legend and choose which legend to bring back, or use the Analysis menu.



Data

Analytics

Global Superstore

Dimensions

- Abc Customer Name
- Abc Market
- Abc Order Date
- Abc Order ID
- Abc Order Priority
- Abc Postal Code
- Abc Product ID
- Abc Product Name
- Abc Region
- Abc Segment
- Abc Ship Date
- Abc Ship Mode
- Abc State
- Abc Sub-Category
- Abc Measure Names

Measures

- # Discount
- # Profit
- # Quantity
- # Row ID
- # Sales
- # Shipping Cost
- (#) Latitude (generated)
- (#) Longitude (generated)
- # Number of Records
- # Measure Values



Pages

Columns

SUM(Profit)

Rows

Sub-Category

Filters

Segment

Marks

Automatic

Color

Size

Label

Detail

Tooltip

Category

AVG(Profit)

Sheet 3

Sub-Catego..

Accessories

Appliances

Art

Binders

Bookcases

Chairs

Copiers

Envelopes

Fasteners

Furnishings

Labels

Machines

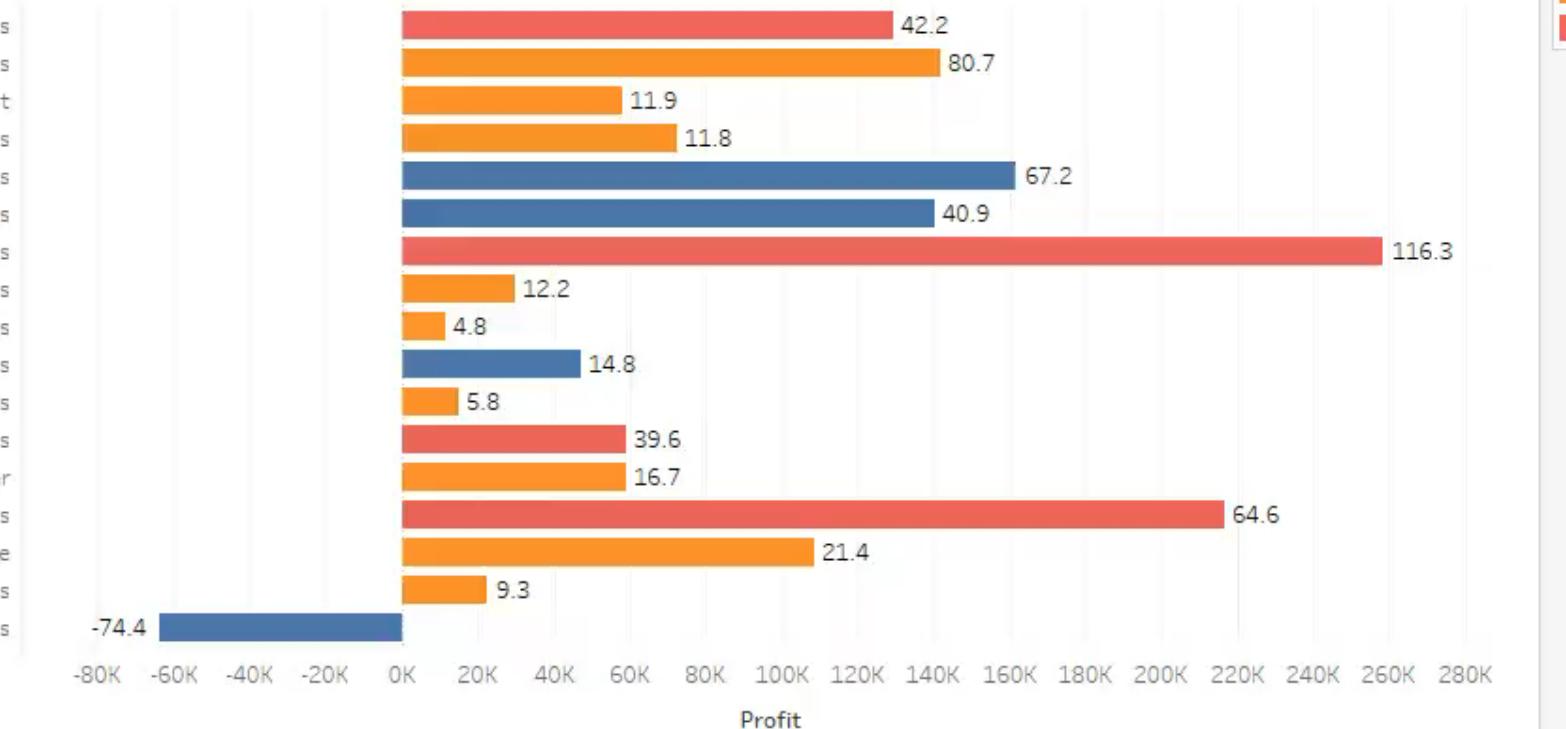
Paper

Phones

Storage

Supplies

Tables



Category

- Furniture
- Office Supplies
- Technology

Legends

Data Source

Sheet 1

Sheet 2

Sheet 3



17 RECORDED WITH 1 column SUM(Profit): 1,467,457



Add New Dashboards & Stories

Add New Dashboards & Stories



Note that we can also add a new dashboard or story instead of a simple sheet. These are ways to organize multiple sheets in a single view or sequence of views – we will cover them in detail later.



Data Analytics

Global Superstore

Dimensions

- Abc Customer Name
- Abc Market
- Abc Order Date
- Abc Order ID
- Abc Order Priority
- Abc Postal Code
- Abc Product ID
- Abc Product Name
- Abc Region
- Abc Segment
- Abc Ship Date
- Abc Ship Mode
- Abc State
- Abc Sub-Category
- Abc Measure Names

Measures

- # Discount
- # Profit
- # Quantity
- # Row ID
- # Sales
- # Shipping Cost
- (# Latitude (generated))
- (# Longitude (generated))
- =# Number of Records
- # Measure Values

Pages

Columns SUM(Profit)

Rows Sub-Category

Sheet 3

Sub-Catego..

Accessories

Appliances

Art

Binders

Bookcases

Chairs

Copiers

Envelopes

Fasteners

Furnishings

Labels

Machines

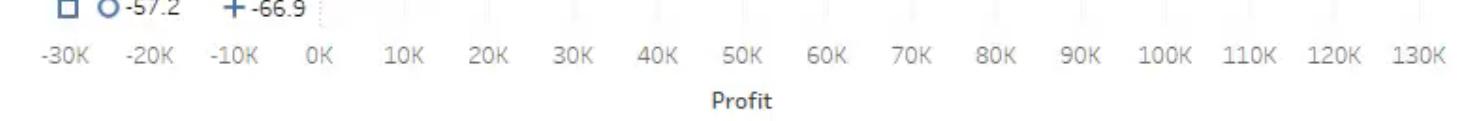
Paper

Phones

Storage

Supplies

Tables



Category

- Furniture
- Technology
- Office Supplies

Segment

- Corporate
- Consumer
- + Home Office

Add New Dashboards & Stories

Data Source

RECORDED WITH

1 column SUM(Profit): 1,467,457

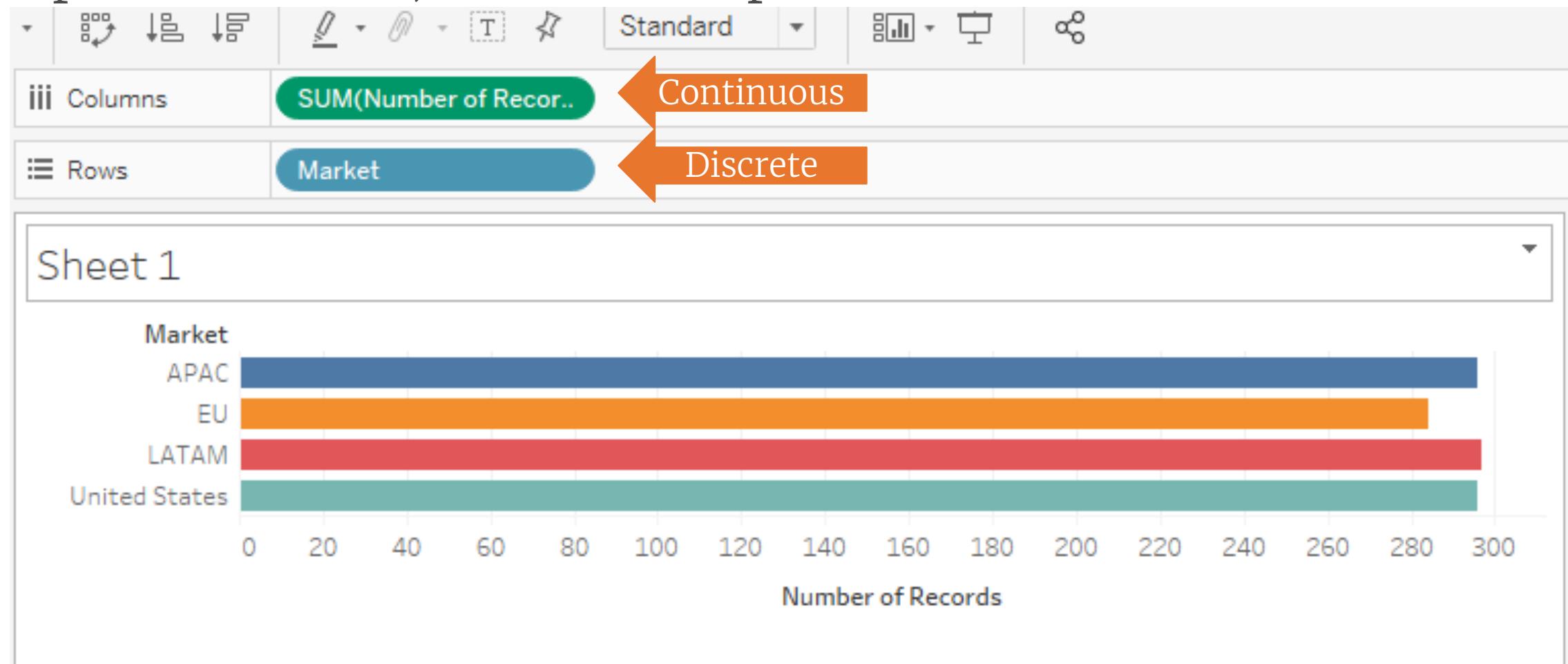




Paying Attention to Pill Color

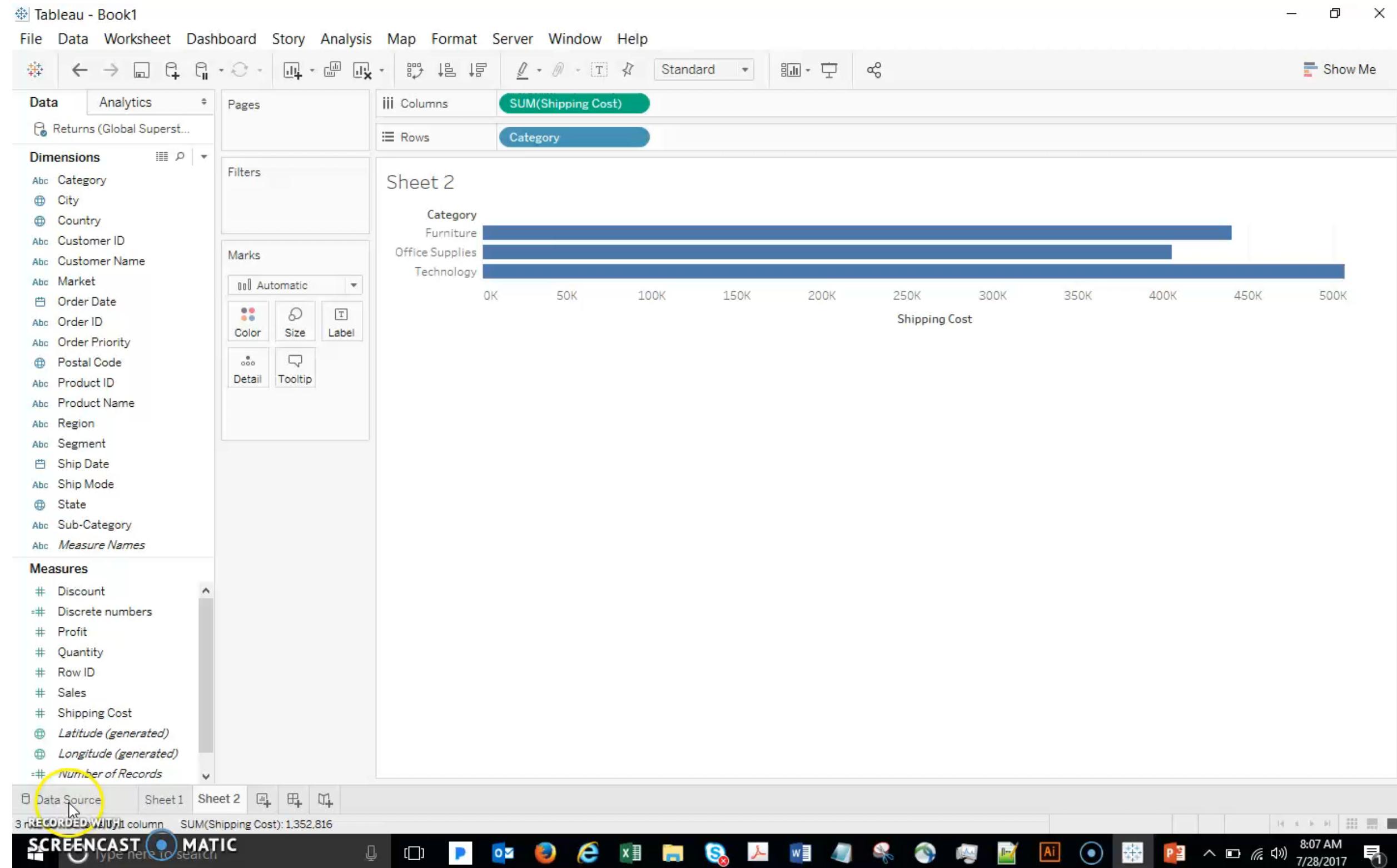
Paying Attention to Pill Color

When we bring a field into the view from the Data Window, Tableau creates a pill. The pill color indicates whether the pill is continuous or discrete. Discrete pills are BLUE, Continuous pills are GREEN.



Paying Attention to Pill Color

Text and categories are inherently discrete. Numbers can also be discrete if they can only take one of a limited set of distinct, separate values. On the other hand, numbers are continuous if they can take on any value in a range.





Dimensions VS Measures

Measures are the numeric values

1. Usually aggregated (sum, max, min, avg, SD)
2. Measure are meaningless without context. If I tell you 10 million, what does that mean

Dimensions provide the context – who, what, where and so on

1. Dimension provides slices
2. Hierarchies
3. Now I tell you – We sold 10 million \$ worth of laptops in east coast in Q1,2018 – this makes more sense



Dimensions & Measures

- In Tableau, dimensions set the granularity, or the level of detail in the view. Think of them as the things you group by or drill down by. Dimensions are usually (but not always) categorical fields such as Order Priority and City.
- It does not make sense to think about computing an average city or average priority (these are words) – it makes sense to slice or group the data using these fields. We typically want to view our data by some combination of these categories.
- What dimensions we use to build the view will determine how many marks we have – Order Priority has 4 categories, so it would give us 4 marks.



Dimensions & Measures

- Measures are usually numerical data like Shipping Cost. Inside of Tableau, measures are aggregations – they're aggregated up to the granularity set by the dimensions in the view. Think of them as the data elements that you want to perform calculations on.
- The value of a measure therefore depends on the context of the dimensions. For example, the result for the sum of Shipping Cost (double click it) is different if we have no dimensions in the view (just a single overall sum) versus when we add Order Priority (double click it) as a dimension – now we have a sum for each priority level.



Data Analytics

Returns (Global Superst...)

Dimensions

- Abc Category
- Abc City
- Abc Country
- Abc Customer ID
- Abc Customer Name
- Abc Market
- Abc Order Date
- Abc Order ID
- Abc Order Priority
- Abc Postal Code
- Abc Product ID
- Abc Product Name
- Abc Region
- Abc Segment
- Abc Ship Date
- Abc Ship Mode
- Abc State
- Abc Sub-Category
- Abc Measure Names

Measures

- # Discount
- =# Discrete numbers
- # Profit
- # Quantity
- # Row ID
- # Sales
- # Shipping Cost
- Abc Latitude (generated)
- Abc Longitude (generated)
- =# Number of Records

Data Source

RECORDED WITH

SCREENCASTOMATIC

Type here to search



Pages

Columns

Rows

Filters

Marks

Automatic

Color

Size

Text

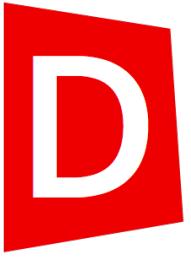
Detail

Tooltip

Drop field here

Dimensions & Measures

Dimensions & Measures



Remember:

- Dimensions come out onto the view as themselves
- Measures come out onto the view as aggregates



Axis vs. Label

Tableau - Book1

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Show Me

Data Analytics

Orders (Global Supersto...)

Dimensions

- Category
- City
- Country
- Customer ID
- Customer Name
- Market
- Order Date
- Order ID
- Order Priority
- Postal Code
- Product ID
- Product Name
- Region
- Row ID
- Segment
- Ship Date
- Ship Mode
- State
- Sub-Category
- Measure Names

Marks

- Automatic
- Color
- Size
- Text
- Detail
- Tooltip

Pages

Columns

Rows

Sheet 1

Drop field here

Drop field here

Drop field here

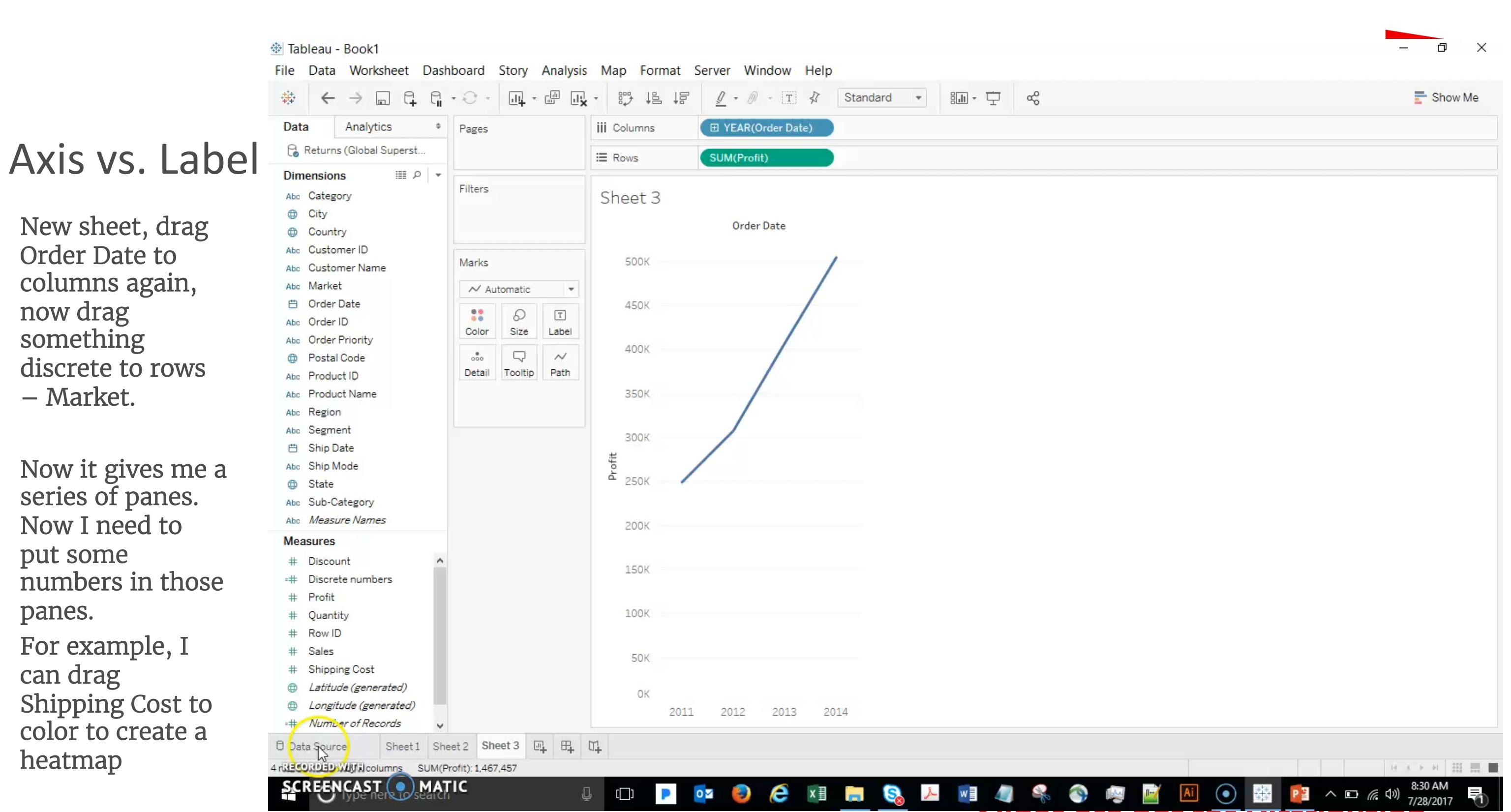
Measures

- Discount
- Profit
- Quantity
- Sales
- Shipping Cost
- Latitude (generated)
- Longitude (generated)
- Number of Records
- Measure Values

Data Source Sheet 1

RECORDED WITH SCREENCASTOMATIC

© DataTiles.io





Filtering



Filtering

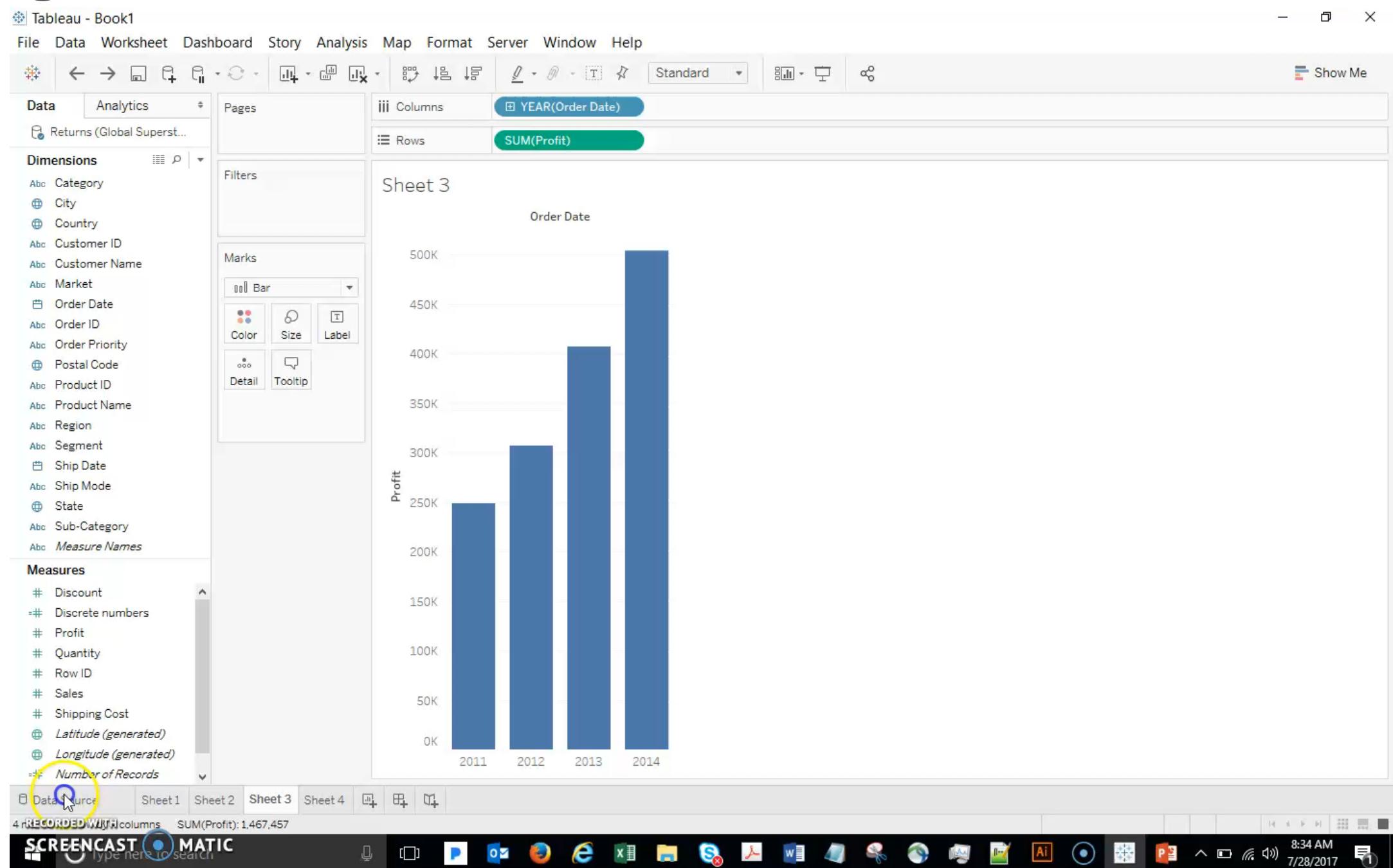
Go back to sheet with bar chart of profit by year. Quick preview here of filters, which we will be using a lot

Filtering on a discrete pill brings up options related to the specific list of values for that pill – drag **Market** to Filters shelf. Then Show Filter.

Filtering on a continuous measure pill asks first if we want to filter at the row-level or aggregate level, then brings up options for continuous ranges. Drag **Discount** to Filters shelf, select all values. Again, Show Filter.



Filtering



Slides are not for Distribution



Color & Maps

Color & Maps



- Let's also do a quick preview of a topic that we will spend more time on later – maps.
- A measure on color defaults to a filled map. A dimension on color defaults to a symbol map. If the pill is continuous Tableau will show a color gradient. If it is discrete Tableau will show a palette of distinct colors. This sounds more complex than it is – let's do a couple maps and see it in action.



Color & Maps

- New sheet, double click country to get a world map.
- Let's drag a numeric measure (Profit) to color – I get a filled map with gradient shading. Now Undo, and drag a discrete dimension (Region) to color – I get a map with a unique color for each region. We could add more data to this map, for example, by dragging a measure (Discount) to size. As I said, more on maps later, but this tells you how data type effects the appearance of your map.



Color & Maps

The screenshot shows the Tableau Data Catalog interface. On the left, the Dimensions pane lists various dimensions such as Category, City, Country, Customer ID, etc. The Measures pane lists measures like Discount, Profit, Quantity, Sales, etc. A yellow circle highlights the 'Data Source' button at the bottom of the navigation bar. The main area is titled 'Sheet 5' and contains three 'Drop field here' placeholder boxes.



Continuous vs. Discrete Dates



Continuous vs. Discrete Dates

With dates, it is very important to understand the difference between continuous and discrete, since dates can be brought into the view either way. If the date icon is blue in the data window, then the default when we bring out that pill will be discrete, otherwise it will be brought out as continuous

Useful trick: if we right click drag (option click drag on a Mac), we can select specifically what sort of date we want

Importantly, once a date is in the viz we can easily change its data type in the pill's menu – you need to remember that the options on top are discrete dates, the lower section are continuous dates.



Continuous vs. Discrete Dates

Now show it both ways:

New sheet. Start by double-clicking Profit, then bringing Order Date to columns. We can see by the blue pill that this is a discrete date. This means that each part of the date is treated like a category. So when you drill down (click the + to expand to years and quarters), you have a separate section of the viz for each year, rather than one continuous line.

Go to the Analytics tab and drag out a **trendline** – see that it does a separate line for each year. Discrete dates can be good because they allow you to treat each “date part” (year, month, etc.) as a separate variable. So, for example, if you want to filter sales by month (e.g. show me our sales in July), a discrete date works best.



Continuous vs. Discrete Dates

Now drag the Quarter pill out of the view so that you have only Year again, and remove the trendline. Right click and select the second year (discuss how the examples in the drop-down are slightly helpful...).

You can see how the line now becomes continuous, and when you click the + it simply adds resolution rather than adding new pills (since each part of the date is not considered a category anymore). Now drag a trendline out – single line for the entire (continuous) range.



Continuous vs. Discrete Dates

Tableau - Book1

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Data Analytics

Pages Columns Rows

Filters

Marks Automatic

Color Size Text

Detail Tooltip

Sheet 6

Drop field here

Drop field here

Drop field here

Dimensions

- Abc Category
- ⊕ City
- ⊕ Country
- Abc Customer ID
- Abc Customer Name
- Abc Market
- ⊕ Order Date
- Abc Order ID
- Abc Order Priority
- ⊕ Postal Code
- Abc Product ID
- Abc Product Name
- Abc Region
- Abc Segment
- ⊕ Ship Date
- Abc Ship Mode
- ⊕ State
- Abc Sub-Category
- Abc Measures

Measures

- # Discount
- # Discrete numbers
- # Profit
- # Quantity
- # Row ID
- # Sales
- # Shipping Cost
- ⊕ Latitude (generated)
- ⊕ Longitude (generated)
- # Number of Records
- + Measure Values

Data Source

RECORDED WITH SCREENCASTOMATIC

11:30 AM 7/28/2017



Bar Charts



Bar Charts

- We have already created some bar charts, but now let's dive more deeply into this very simple but very powerful chart type.
- Let's Analyze Sales (double click) by Category (double click) and drill down to Sub-Category (drag to Columns, left of Category).
- To make this easier to read let's rotate it (click Swap in the toolbar) and sort it (click Sort icon).
- This gives me a very nice look at Sales (in just a few seconds!). For example, Furniture (especially Chairs) seems to be selling well. But what about Profits?

Bar Charts

Tableau - Book1

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Show Me

Data Analytics

Pages Columns Rows

Sheet 8

Drop field here

Drop field here

Drop field here

Marks

Automatic

Color Size Text

Detail Tooltip

Dimensions

- Category
- City
- Country
- Customer ID
- Customer Name
- Market
- Order Date
- Order ID
- Order Priority
- Postal Code
- Product ID
- Product Name
- Region
- Segment
- Ship Date
- Ship Mode
- State
- Sub-Category

Measures

- Discount
- Discrete numbers
- Profit
- Quantity
- Row ID
- Sales
- Shipping Cost
- Latitude (generated)
- Longitude (generated)
- Number of Records
- Measure Values

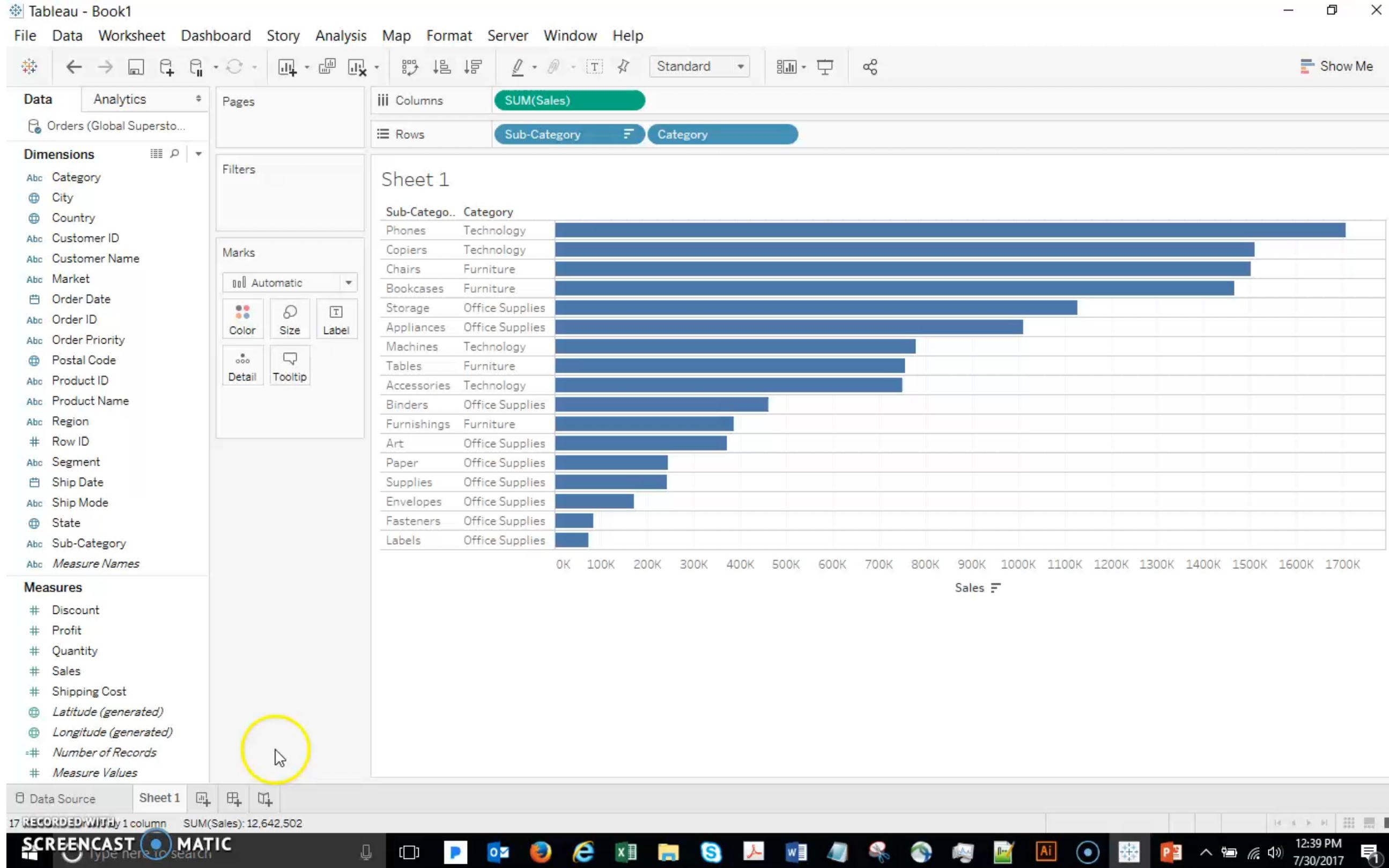
Data Source

RECORDED WITH SCREENCASTOMATIC

11:38 AM 7/28/2017

Bar Charts

Drag Profit out so you get the “show me” cursor. Drop it and Tableau will use it on the Color shelf. It’s clear that Tables are not profitable – and I’m selling lots of them! Not good.



Slides are not for Distribution

Bar Charts



- Reminder: the Show Me pane provides other possibilities for visualizing the data I have selected – we can click a couple to see other options, but then let's use the Back button to return to our bar chart.



Bar Charts

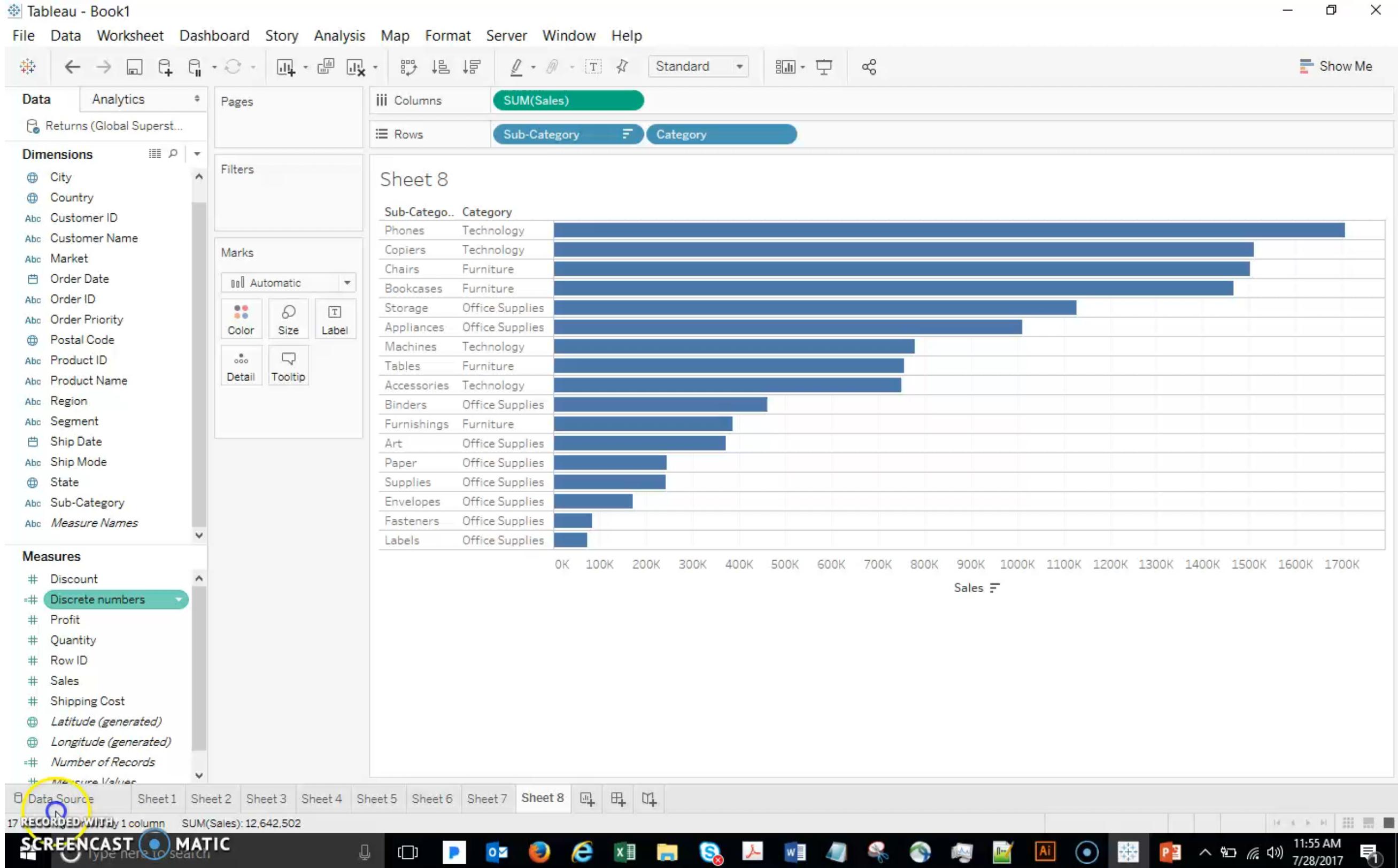
- Notice that Office Supplies has a number of smaller categories that are just cluttering up my analysis. This is a common issue in a dataset, and I don't want to remove these small categories because they may be significant in aggregate. Instead, let's select them all (hold down CTRL and select the words) and clicking the paper clip to Group them. Right-click and Edit Alias to rename the group.
- It's important to understand that this does not change the underlying data – the original sub-categories are still there. What this HAS done is created a new object (point it out in the dimensions pane).
- Now we see that there is a problem in Tables. Right-click Tables and select Keep Only.



Bar Charts

- Now that we're just looking at Tables, let's drill all the way down to the Product level. Drag out the **Product Name** to the viz and **Sort** descending.
- Finally let's look at the actual underlying data from our top selling product here. Click the first bar... **View Data**, **Full Data** tab. And now I have a nice list of transactions that I might want to investigate.

Bar Charts



Slides are not for Distribution



Scatterplots



Scatterplots

- Scatterplots are a very easy and powerful way to visualize the relationships between numeric variables (Measures). For example, I wonder if there is a relationship between **Discount** and **Profit** (seems like a reasonable possibility). Let's take a look by selecting those two fields and clicking **Scatterplot** on the **Show Me** menu. Go to **Analysis**, uncheck **Aggregate Measures**.
- Swap if needed to get **Discount** on the horizontal axis – looks like **Profit** tends to decrease as a larger **Discount** is given (confirms intuition).
- Remember that with Tableau, unlike many other tools, we don't need to stop here. It is very easy to keep adding data. For example, drag **Segment** to **Color**. This makes it easy to see what segment certain outliers belong to.



Scatterplots

- If we prefer to focus on just one segment (while leaving all the data displayed), we can click **Highlight Selected Values** in the Legend and select, for example, to focus our viz on the **Corporate** segment by highlighting it.
- Go back to viewing all the data. Go to the **Analytics pane** and drag out a **trend line** (choose **linear**). If you have had a statistics course, you probably realize that this is a regression line. We are not covering regression in this course, but for students who need a refresher, I have posted a document on the course website for you.

Scatterplots



This trend line does not look very impressive, but take a look at the y-axis scale. We need to investigate more before dismissing this as a negligible relationship. Hover over the trend line. The p-value is very small, so in fact the relationship between these variables is statistically significant. And we can see by the negative coefficient that the relationship is negative. However, the R-Squared value is somewhat low, so our simple trend line has not explained very much of the relationship. Drag the trendline out of the viz to remove it. [[Instructors might want to digress more into interpretations here if desired]]



Show Me

Data Analytics

Returns (Global Superst...)

Dimensions

- City
- Country
- Customer ID
- Customer Name
- Market
- Order Date
- Order ID
- Order Priority
- Postal Code
- Product ID
- Product Name
- Region
- Segment
- Ship Date
- Ship Mode
- State
- Sub-Category
- Sub-Category (group)

Measures

- # Discount
- =# Discrete numbers
- # Profit
- # Quantity
- # Row ID
- # Sales
- # Shipping Cost
- # Latitude (generated)
- # Longitude (generated)
- =# Number of Records
- + Measure Values

Pages

Columns

Rows

Sheet 9

Drop field here

Drop field here

Drop field here



Select or drag data

Use the Shift or Ctrl key to select multiple fields



Data Source

Sheet 1

Sheet 2

Sheet 3

Sheet 4

Sheet 5

Sheet 6

Sheet 7

Sheet 8

Sheet 9





Line Graphs



Line Graphs

- Like Scatterplots, line graphs are good for looking at the relationship between 2 variables (although a scatterplot is often better for this). Line graphs are often used for looking at how something changes over time. This is called time series analysis.
- Let's start by looking at sales over time. Double-click Sales, double-click Order Date. Tableau understands date fields and automatically breaks them down by Year, Quarter, etc.
- I actually need more detail than yearly sales results, so let's expand to Quarter and then Month. I don't really care about Quarter, so drag it out of the viz.
- Note that dual axis is very easy in Tableau – drag Profit to the right axis. Then Undo to get back to Sales only.

Tableau interface showing the Data pane and a blank sheet.

Data pane:

- Dimensions:** City, Country, Customer ID, Customer Name, Market, Order Date, Order ID, Order Priority, Postal Code, Product ID, Product Name, Region, Segment, Ship Date, Ship Mode, State, Sub-Category, Sub-Category (group).
- Measures:** Discount, Discrete numbers, Profit, Quantity, Row ID, Sales, Shipping Cost, Latitude (generated), Longitude (generated), Number of Records.

Sheet 10: A blank sheet with three drop zones labeled "Drop field here".

Bottom navigation:

- Data Source (highlighted with a yellow circle)
- Sheet 1, Sheet 2, Sheet 3, Sheet 4, Sheet 5, Sheet 6, Sheet 7, Sheet 8, Sheet 9, Sheet 10
- Standard
- Show Me



Line Graphs

- Remember that we can now drag things around and change the viz on the fly. We have discrete dates right now (we have already learned how, with discrete dates, each part of the date is treated separately). So for example we can drag **Year** to **Color** and see direct comparisons between the years. (Note that, if you liked the current format and just want each year to be a different color for clarity, you could CNTRL-drag **Year** to color).



Data Analytics

Returns (Global Superstore)

Dimensions

- City
- Country
- Customer ID
- Customer Name
- Market
- Order Date
- Order ID
- Order Priority
- Postal Code
- Product ID
- Product Name
- Region
- Segment
- Ship Date
- Ship Mode
- State
- Sub-Category
- Sub-Category (group)

Measure Names

- # Discount
- # Discrete numbers
- # Profit
- # Quantity
- # Row ID
- # Sales
- # Shipping Cost
- # Latitude (generated)
- # Longitude (generated)
- # Number of Records
- # Measure Values

Pages

Columns

YEAR(Order Date)

MONTH(Order Date)

Rows

SUM(Sales)

Sheet 10

Order Date

2011

2012

2013

2014

Sales

OK

200K

300K

400K

500K

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

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

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

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

Sheet 1 Sheet 2 Sheet 3 Sheet 4 Sheet 5 Sheet 6 Sheet 7 Sheet 8 Sheet 9 Sheet 10

Sheet 1 Sheet 2 Sheet 3 Sheet 4 Sheet 5 Sheet 6 Sheet 7 Sheet 8 Sheet 9 Sheet 10

Sheet 1 Sheet 2 Sheet 3 Sheet 4 Sheet 5 Sheet 6 Sheet 7 Sheet 8 Sheet 9 Sheet 10

Sheet 1 Sheet 2 Sheet 3 Sheet 4 Sheet 5 Sheet 6 Sheet 7 Sheet 8 Sheet 9 Sheet 10

48 RECORDED WITH 48 columns SUM(Sales): 12,642,502



Line Graphs

- Let's all get back to a separate line chart for each year. If I want to look at a common calculation for time series data, like year over year growth, it is very easy: Click on the green Sum(Sales) pill. Select Quick Table Calculation... Year over Year Growth. Now I can see that there were months in which sales went down YOY. Click the Back button.



Show Me

Data Analytics

Returns (Global Superst...)

Dimensions

- City
- Country
- Customer ID
- Customer Name
- Market
- Order Date
- Order ID
- Order Priority
- Postal Code
- Product ID
- Product Name
- Region
- Segment
- Ship Date
- Ship Mode
- State
- Sub-Category
- Sub-Category (group)

Measures

- Discount
- Discrete numbers
- Profit
- Quantity
- Row ID
- Sales
- Shipping Cost
- Latitude (generated)
- Longitude (generated)
- Number of Records
- Measure Values

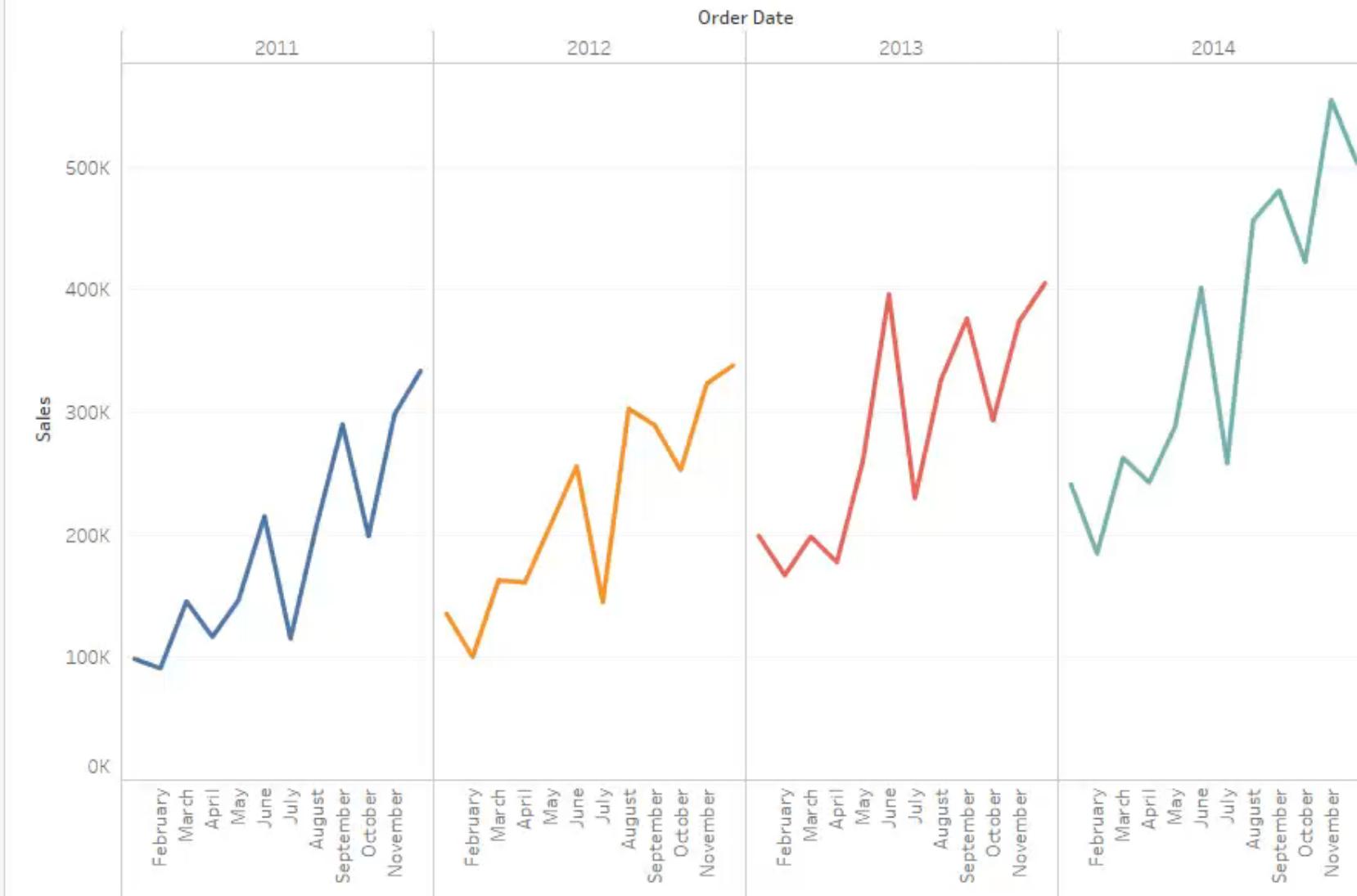
Data Source

48 RECORDED WITH 48 columns SUM(Sales): 12,642,502

Pages Columns YEAR(Order Date) MONTH(Order Date)

Rows SUM(Sales)

Sheet 10

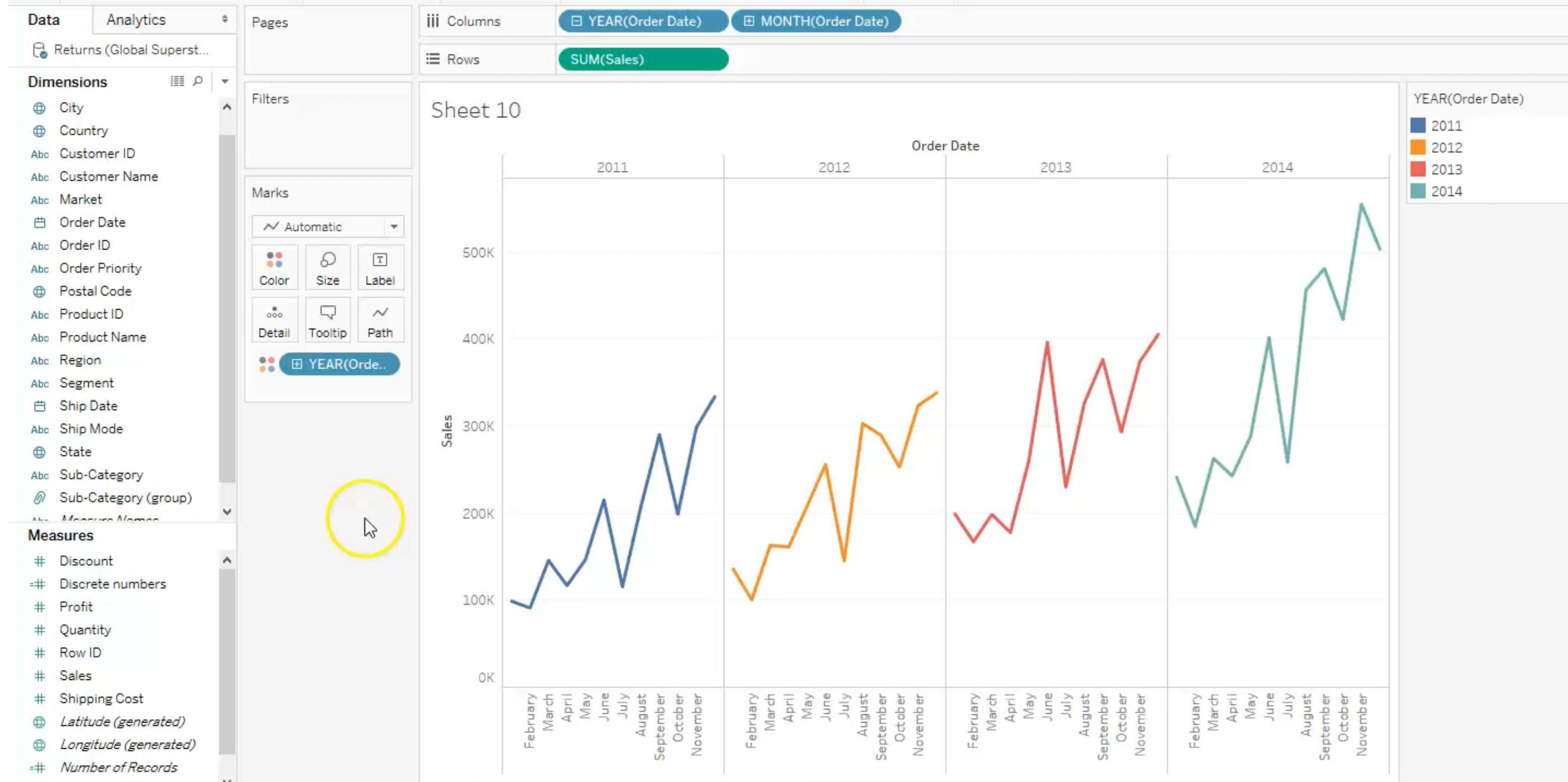


YEAR(Order Date)

- 2011
- 2012
- 2013
- 2014



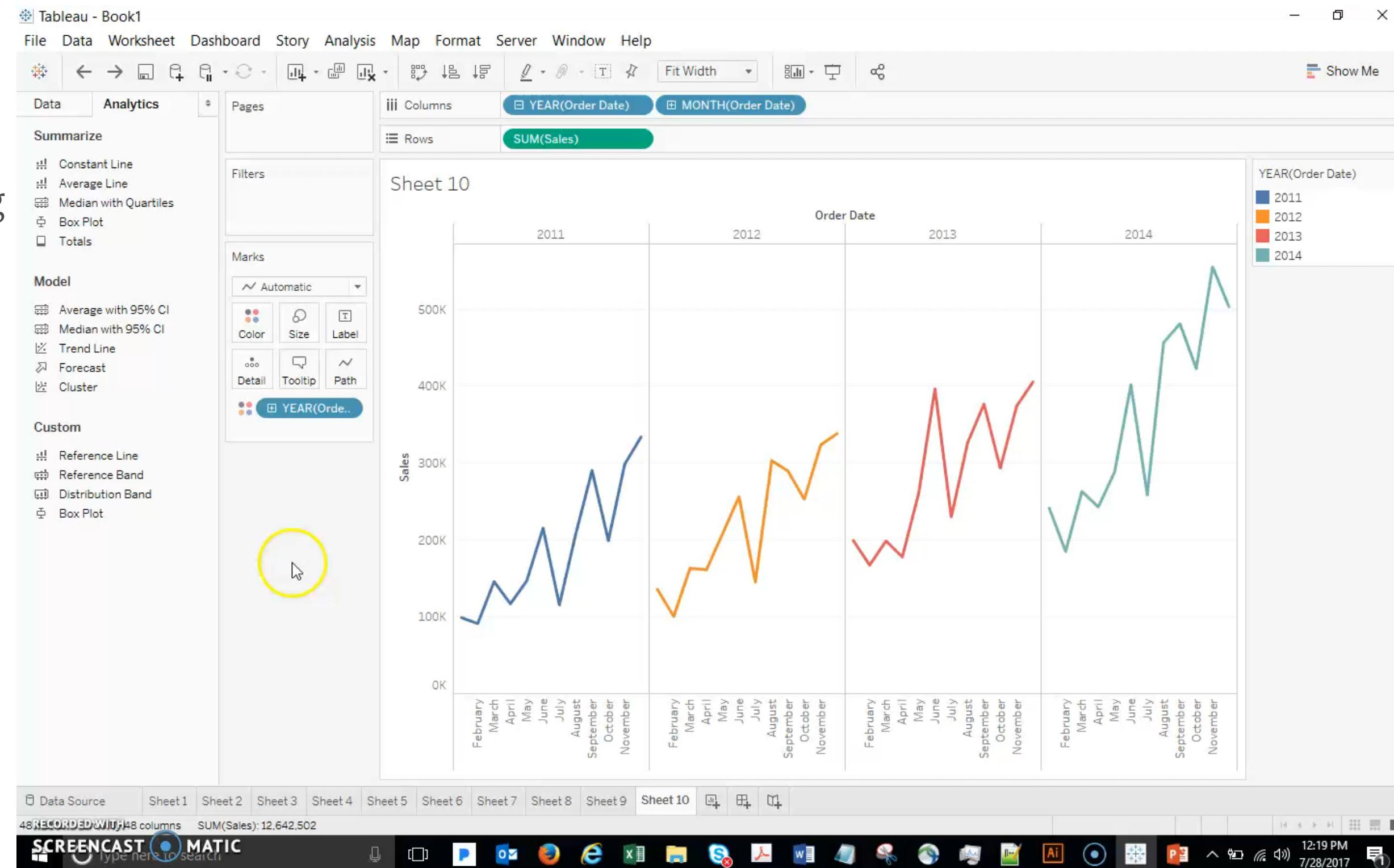
With
Tableau
it is also
easy to
add a
Forecast
. Go to
the
Analytics
pane
and
drag out
Forecast



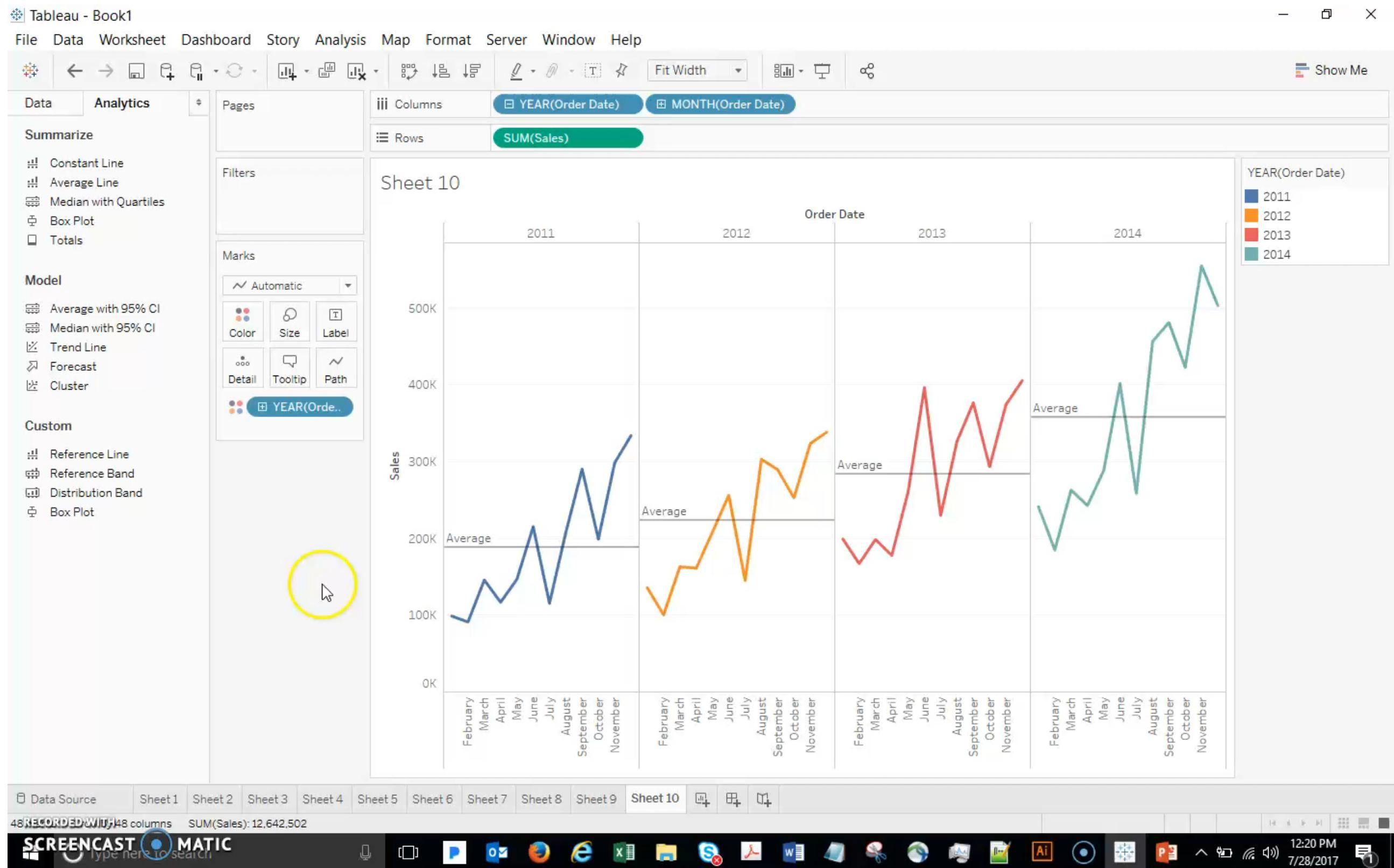
Click
the Back
button.

Another useful feature to add to a line chart is a reference line, for example: drag the Average Line onto the viz...drop it onto the Pane icon.

you can also drill down, and this average will dynamically adjust: select a few points in the line graph and note the new average line.



Although we are talking about line graphs, we want to briefly point out how easy it is to explore other options: click the drop-down in the Marks card and select Bar, then Circle, and then leave it as Area.



Slides are not for Distribution



Histograms



Histograms

- Double click Discount. The default is a bar chart, and we can add more data to break out discount by segment, etc. But what if we are just interested in learning more about the various discounts that we offer?
- Use Show Me to change to Histogram. Right-click the new Discount(bin) dimension, Edit, and change bins to 0.1. This is a very quick and easy way to get the view of the data that we want! Now we can see that most discounts are between 0 and 10%, and there are also a lot of discounts in the 10-20% range. We even have 300 instances in which the discount given was 70-80% (shown in the Tooltip). If we are suspicious about these, it just takes a couple clicks to get to the underlying data (right-click that bar and show the data).
- Remember that, since this is Tableau, we can add more data if we want – for example, drag Category to Color, then Undo this change.



Data Analytics

Returns (Global Superstore)

Dimensions

- City
- Country
- Customer ID
- Customer Name
- Market
- Order Date
- Order ID
- Order Priority
- Postal Code
- Product ID
- Product Name
- Region
- Segment
- Ship Date
- Ship Mode
- State
- Sub-Category
- Sub-Category (group)

Measures

- # Discount
- =# Discrete numbers
- # Profit
- # Quantity
- # Row ID
- # Sales
- # Shipping Cost
- Latitude (generated)
- Longitude (generated)
- =# Number of Records
- +# Measure Values

Data Source

RECORDED WITH

SCREENCASTOMATIC

Pages

Columns

Rows

Sheet 11

Drop field here

Drop field here

Drop field here

Select or drag data

Use the Shift or Ctrl key to select multiple fields





Heat Maps



Heat Maps

- Let's use heat maps to visualize the margin that we are making on Furniture items. Notice that margin is not in the data - let's do a quick preview of Calculated Fields to create the margin field (we will cover calculations in detail in a later topic):
- Analysis...Create Calculated Field. Name it Margin, formula is $\text{SUM}([\text{Profit}])/\text{SUM}([\text{Sales}])$
- Drag Category and Subcategory to Row
- Drag Order Date to Columns, and then change to Month (discrete)
- Drag Category to Filter shelf, and select only Furniture
- Finally, drag your new Margin field to the viz. We have a table of numbers. Not the most user-friendly...



Heat Maps

- Select **Heat Map** from Show Me, and then drag **AGG(Margin)** on the **Marks** card to **Color**. Change color to Orange-Blue Diverging. Drag edges if needed to make the viz larger. This is much better than the table of numbers! We immediately see some insights – e.g. maybe we are selling tables as a loss leader and making it up on chairs?

[File](#) [Data](#) [Worksheet](#) [Dashboard](#) [Story](#) [Analysis](#) [Map](#) [Format](#) [Server](#) [Window](#) [Help](#)

Show Me

[Data](#)[Analytics](#)[Returns \(Global Superst...\)](#)**Dimensions**[ABC Category](#)
[City](#)
[Country](#)
[Customer ID](#)
[Customer Name](#)
[Discount \(bin\)](#)
[Market](#)
[Order Date](#)
[Order ID](#)
[Order Priority](#)
[Postal Code](#)
[Product ID](#)
[Product Name](#)
[Region](#)
[Segment](#)
[Ship Date](#)
[Ship Mode](#)
[State](#)
[Sub Category](#)**Measures**[# Discount](#)
[# Discrete numbers](#)
[# Profit](#)
[# Quantity](#)
[# Row ID](#)
[# Sales](#)
[# Shipping Cost](#)
[@ Latitude \(generated\)](#)
[@ Longitude \(generated\)](#)
[# Number of Records](#)
[+ Measure Values](#)[Data Source](#)[Sheet 1](#)[Sheet 2](#)[Sheet 3](#)[Sheet 4](#)[Sheet 5](#)[Sheet 6](#)[Sheet 7](#)[Sheet 8](#)[Sheet 9](#)[Sheet 10](#)[Sheet 11](#)[Sheet 12](#)[Pages](#)[Columns](#)[Rows](#)[Filters](#)[Marks](#)[Automatic](#)[Color](#)[Size](#)[Text](#)[Detail](#)[Tooltip](#)[Drop field here](#)[Drop field here](#)[Drop field here](#)[RECORDED WITH](#)SCREENCASTOMATIC
Type here to search12:33 PM
7/28/2017



Creating a Basic Dashboard

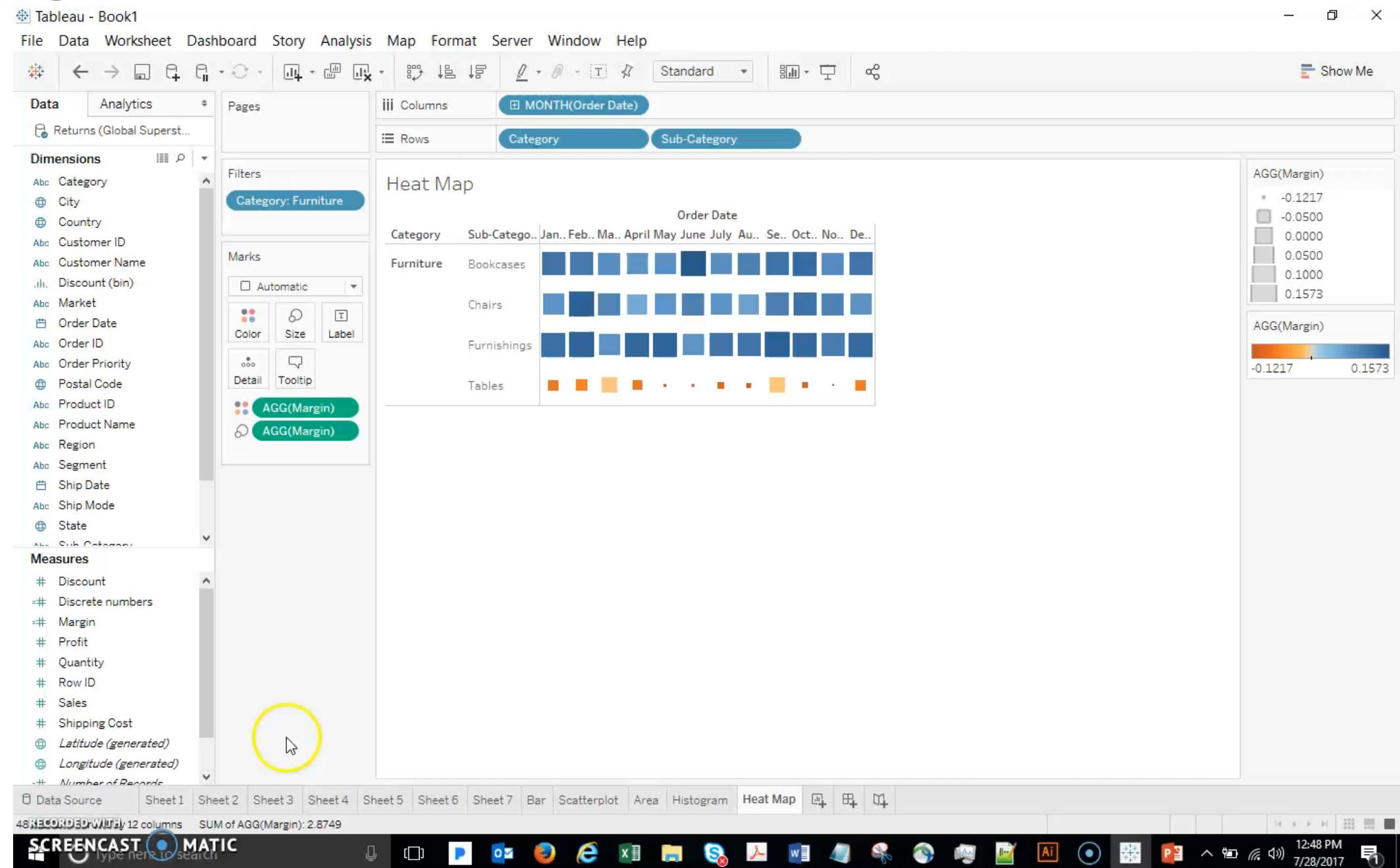
Creating a Basic Dashboard



Click the icon to create a new dashboard. Name it Initial Superstore Analysis. Go to the Dashboard drop-down and Show Title.

Notice these Dashboard options at the lower-left. You should always think about this before starting your Dashboard. We will leave it set to Desktop, but often Laptop is a good choice also. “Automatic” is OK sometimes, but not if you want to have floating legends (which we will show today), since these will move around as the screen resizes under the Automatic setting.

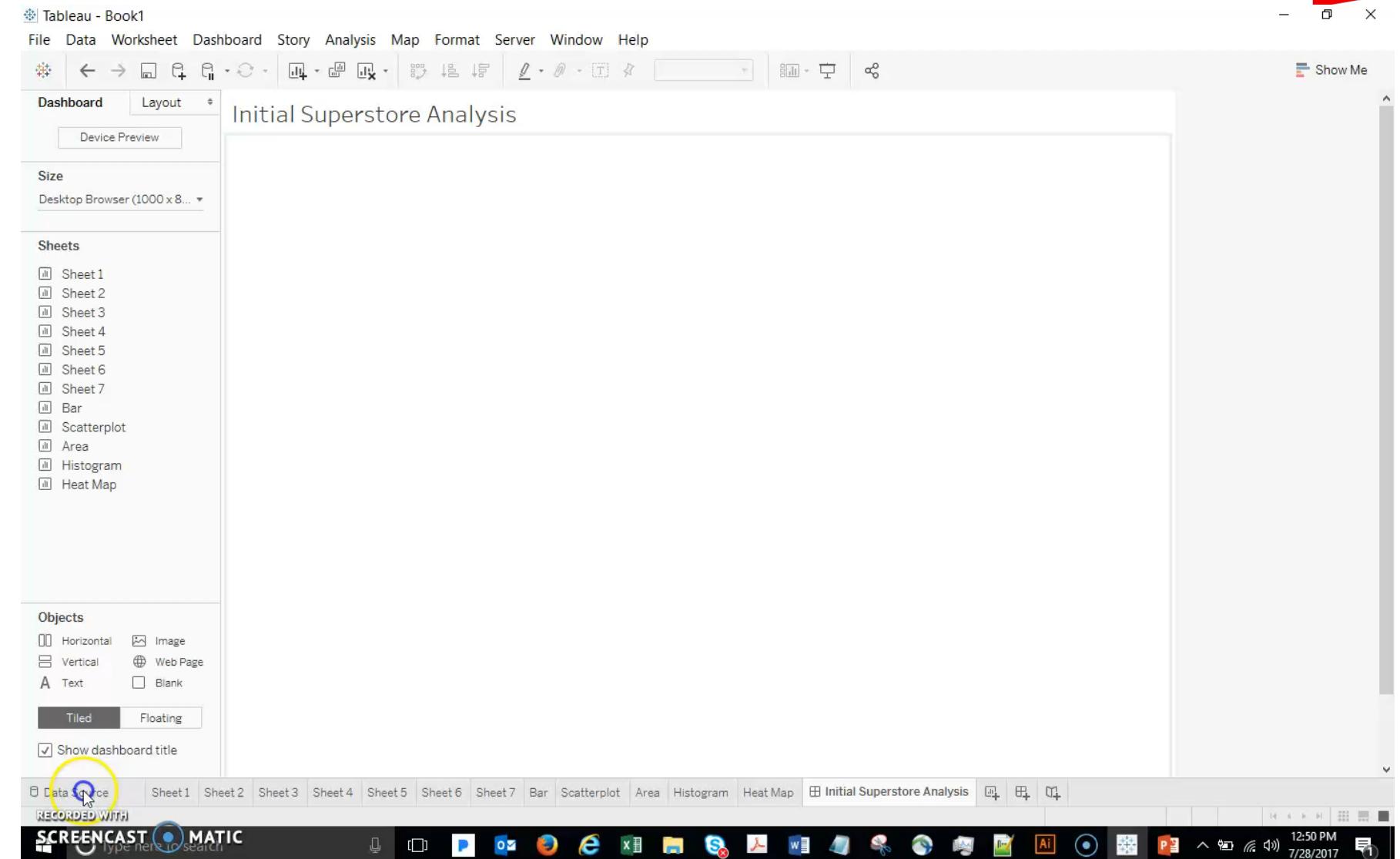
Creating a Basic Dashboard



Creating a Basic Dashboard

Before populating this new dashboard, let's add a Segment filter to the Scatterplot worksheet, and Show Filter. Now I am ready to populate my dashboard.

- Drag in the Bar, add the Scatterplot below it, and add the Histogram on the lower right.



Creating a Basic Dashboard



- Now let's explore some options to enhance this basic dashboard.
- **Float the Profit color legend over the Bar Chart** (Press shift+ drag to float it easily).
- **Float the Segment color legend over the Scatterplot.**
- Set the Segment filter to a single-value drop down.
- Let's have a closer look at that Segment filter. Notice that when I make changes to it, it only changes the worksheet that it was tied to. What if I want the entire dashboard to update based on my selection?



Creating a Basic Dashboard

Tableau - Book1

File Data Worksheet Dashboard Story Map Format Server Window Help

Dashboard Layout Device Preview

Size Desktop Browser (1000 x 800)

Sheets

- Sheet1
- Sheet2
- Sheet3
- Sheet4
- Sheet5
- Sheet6
- Sheet7
- Bar
- Scatterplot
- Area
- Histogram
- Heat Map

Objects

- Horizontal
- Vertical
- A Text
- Image
- Web Page
- Blank

Tiled Floating

Show dashboard title

RECORDED WITH SCREENCASTOMATIC

Initial Superstore Analysis

Bar

Sub-Catego.. Category Product Name

Tables	Furniture	Product Name	Sales
		Hon Conference Table, Adj.	14,500
		Barricks Conference Table..	13,000
		Bevis Conference Table, F..	13,000
		Chromcraft Computer Tab..	12,500
		Chromcraft Conference Ta..	12,000
		Bretford Rectangular Con..	11,500
		Lesro Conference Table, w..	11,000
		Bevis Computer Table, Ful..	10,500
		Bevis Wood Table, with B..	10,000
		Lesro Conference Table, R..	10,000
		Lesro Conference Table, A..	10,000
		Lesro Round Table, Recta..	10,000
		Round Conference Table ..	10,000

Profit

-3,650 3,009

Segment

- (All)
- Consumer
- Corporate
- Home Office

Segment

- Consumer
- Corporate
- Home Office

Scatterplot

Histogram

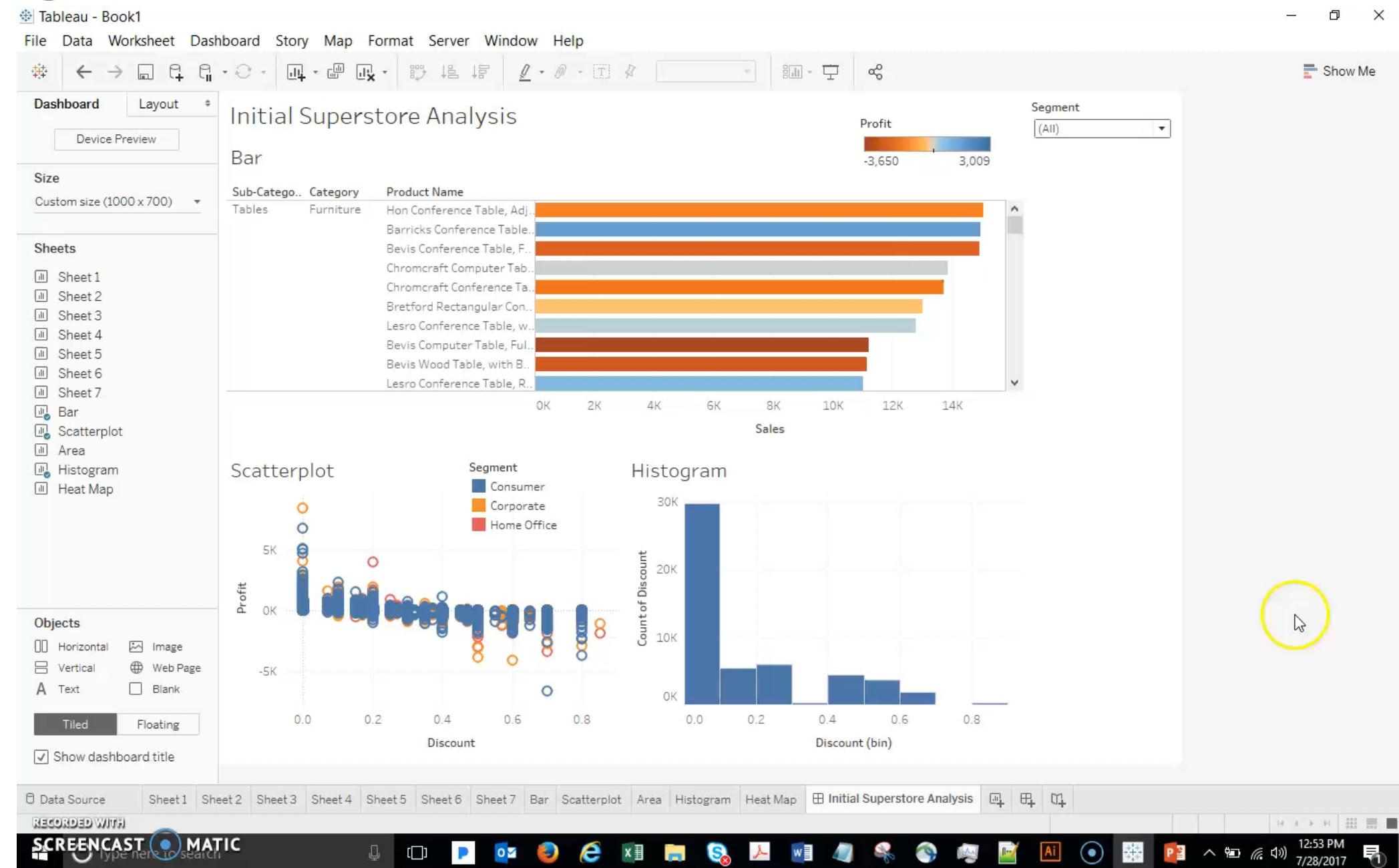
Count of Discount

Creating a Basic Dashboard



- Select the menu on the Segment filter, select Apply to Worksheets...All Using This Data Source.
- Make a couple of Segment selections to demonstrate how the change applies to the entire dashboard, finish with “All” categories selected.

Creating a Basic Dashboard



Slides are not for Distribution

Creating a Basic Dashboard

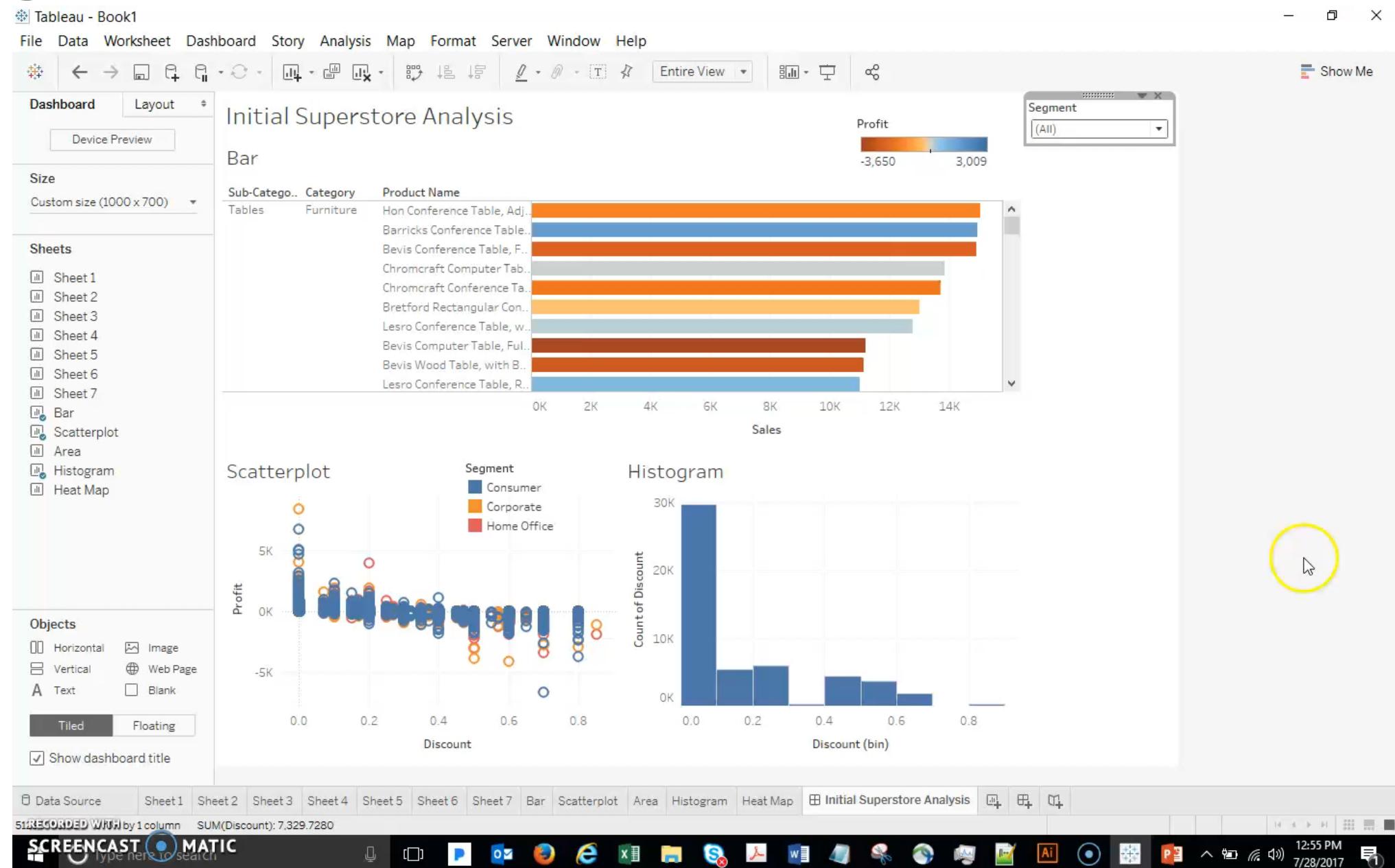


In addition to interacting with drop down filters like this, we can also use the **worksheets themselves as filters**. For example, I'd like to zero in on the different discount levels:

Select the ‘use as a filter’ icon on the Histogram (next to the drop-down arrow).

Click a few histogram bars and discuss

Creating a Basic Dashboard





Creating a Story

Creating a Story



Now click the **New Story** icon, and demonstrate how sheets or dashboards can be dragged in, captions written, etc. to develop a sequential walk through your analysis. Our current Tableau file might not be ideal for creating a Story, since we did a number of unrelated analyses, but this gives you an idea of what a Story can do – it is very powerful and simple!



Creating a Story

Tableau - Book1

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Dashboard Layout Device Preview

Size Custom size (1000 x 700)

Sheets

- Sheet 1
- Sheet 2
- Sheet 3
- Sheet 4
- Sheet 5
- Sheet 6
- Sheet 7
- Bar
- Scatterplot
- Area
- Histogram
- Heat Map

Objects

- Horizontal
- Vertical
- Text
- Image
- Web Page
- Blank

Tiled Floating

Show dashboard title

Initial Superstore Analysis

Segment (All)

Profit

-3,650 3,009

Bar

Sub-Catego.. Category Product Name

Tables	Furniture	Product Name
		Hon Conference Table, Adj.
		Barricks Conference Table..
		Bevis Conference Table, F..
		Chromcraft Computer Tab..
		Chromcraft Conference Ta..
		Bretford Rectangular Con..
		Lesro Conference Table, w..
		Bevis Computer Table, Ful..
		Bevis Wood Table, with B..
		Lesro Conference Table, R..

Sales

Scatterplot

Segment

- Consumer
- Corporate
- Home Office

Profit

Discount

Histogram

Count of Discount

Discount (bin)

RECORDED WITH SCREENCASTOMATIC

12:58 PM 7/28/2017



Labels



Labels

- Start by making a line chart – Sales by month (continuous) of Order Date, Market onto color
- Click the Label shelf to see the options:
- Checking this box will turn labels on in the view, the same as clicking the “T” label button in the toolbar.
- Drag Profit to the label shelf. This adds it to the label (turn it off with the toolbar). Click Label shelf, we can click on the ellipses here will bring up the full text editor to customize the label. (Only fields that have been added to the label shelf are available to use here.) Labels can be turned on for specific actions (such as highlighting) or for specific points (such as minimum/maximum), and other options here.

D

Labels

The screenshot shows the Tableau Data Catalog interface. On the left, the 'Measures' section is expanded, listing various metrics. The 'Sales' measure is highlighted with a yellow circle and has a cursor icon over it. The 'Sheet 8' workspace is visible on the right, showing two empty 'Drop field here' placeholder boxes.

Tableau - Global Superstore 10-3

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Data Analytics

Global Superstore

Dimensions

- Customers
 - Customer ID
 - Customer Name
 - Segment
- Orders
 - Order Date
 - Order ID
 - Order Priority
- Shipping
 - Ship Date
 - Ship Mode
- Location
 - Market
 - Region
 - Country
 - State
 - City
 - Postal Code
- Products
 - Category
 - Sub-Category
 - Product Name

Marks

- Automatic
- Color
- Size
- Text
- Detail
- Tooltip

Measures

- Discount
- Profit
- Quantity
- Sales
- Shipping Cost
- Latitude (generated)
- Longitude (generated)

Sheet 8

Drop field here

Drop field here

Drop field here

Data Source Sheet 8 Sheet 9

RECORDED WITH SCREENCASTOMATIC

1:13 PM 7/28/2017

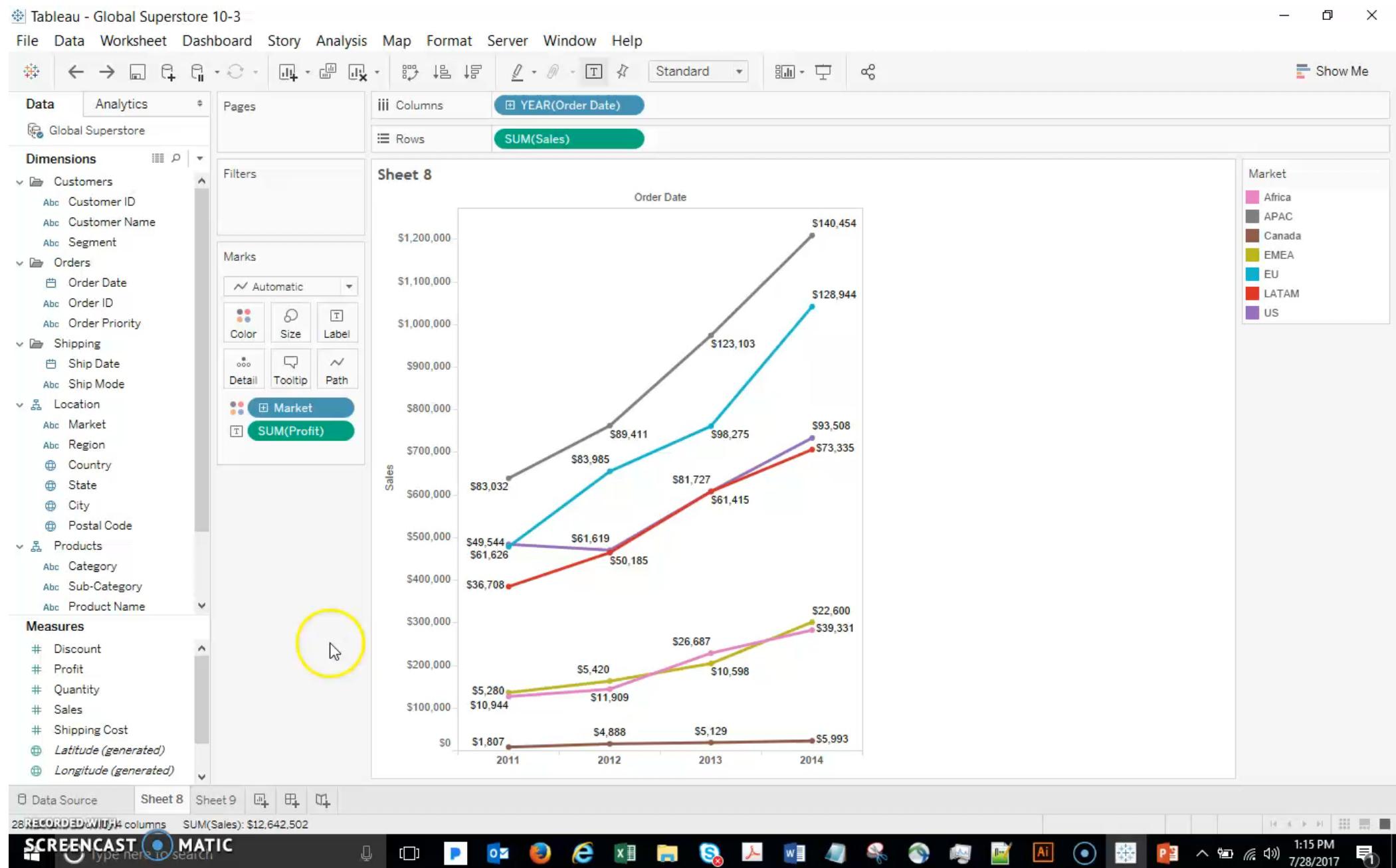


Labels

- If you want to always (or never) see a specific label, just **right click** on the point, go to **Mark Label**, and choose the desired behavior, such as Always Show.
- But you often don't have much space in the view for labels. Often, the best way to add additional information is through the use of annotations or Tooltips.

D

Labels





Annotations



Annotations

- Annotations are text boxes used to call out something specific in the view.
- A mark annotation stays with the mark itself, regardless of whatever changes are made on the view.
- A point annotation stays in the same relative position on the view, independently of the mark, as does an area annotation.
- Let's start with a Mark annotation – right-click a mark select **Annotate...Mark**
- These are not simple text boxes; like labels, annotations can have fields added to them as dynamic inputs.

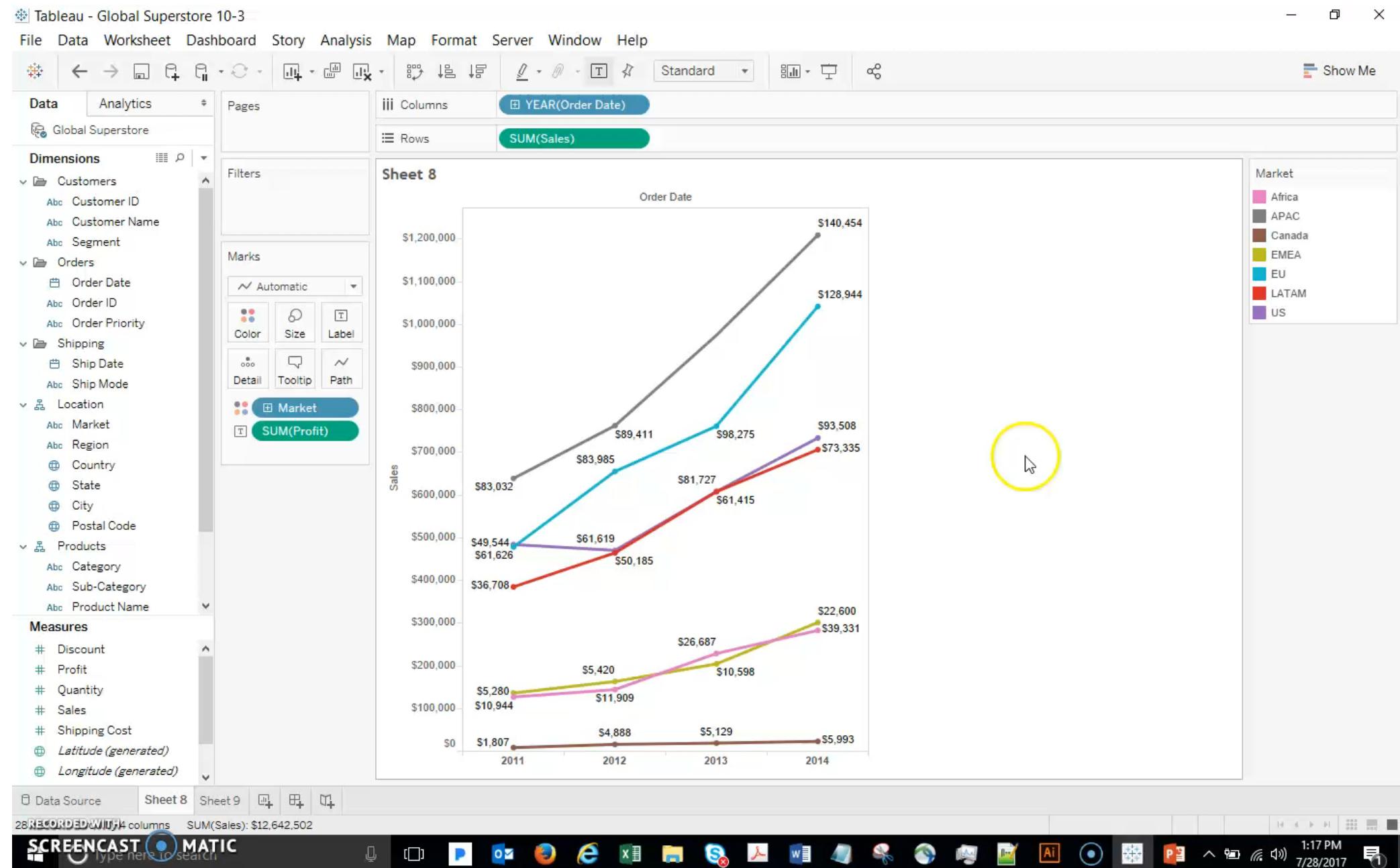


Annotations

- The appearance of annotations is easy to customize:
 - Single click to move or resize it.
 - Double click to format the text.
 - Right click and select **Format** to format the annotation box or line.
- A point annotation is similar, but it will stay in a particular place in the view rather than be tied to a mark, which can move as you interact with the view.



Annotations



Slides are not for Distribution



Tooltips



Tooltips

- To show Tooltips, let's create a new worksheet:
- Select Sales and Profit, choose Scatterplot from Show Me
- Order Priority to Color, then Number of Records and Market to Detail.
- Change the Mark to Circle, and increase the size.
- Hover – the standard Tooltip here has a lot of information. But with just a little work we can make it more user-friendly and engaging:
- Click Tooltip on the Marks card.



Tooltips

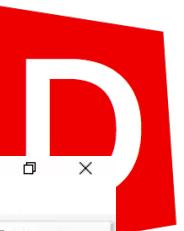
- Change it to read something like (be sure to change the size and color of some words):

<Market> Market's <Order Priority> priority orders
total <SUM(Sales)> in sales
with <SUM(Profit)> overall profit.
This circle represents <SUM(Number of Records)> records



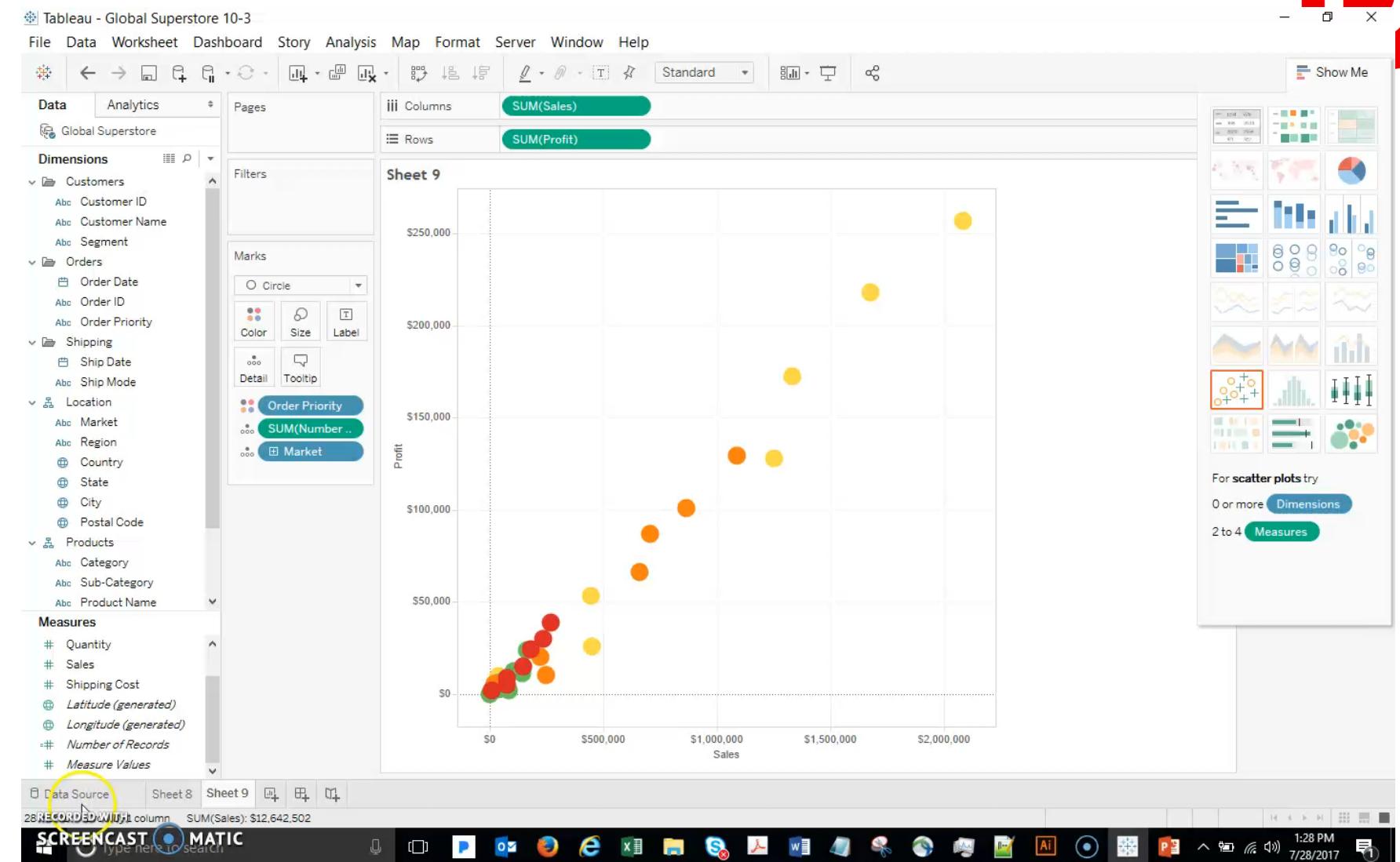
ToolTips

The screenshot shows the Tableau Data Editor interface. On the left, the Data pane displays various dimensions and measures from a 'Global Superstore' data source. The 'Dimensions' section includes Customers, Orders, Shipping, Location, and Products. The 'Measures' section includes Quantity, Sales, Shipping Cost, Latitude (generated), Longitude (generated), Number of Records, and Measure Values. In the center, the 'Sheet 10' workspace has two 'Drop field here' placeholder boxes. On the right, the Marks shelf is open, showing settings for 'Automatic' marks, with icons for Color, Size, Text, Detail, and Tooltip. A yellow circle highlights the 'Tooltip' icon. The bottom of the screen shows the Windows taskbar with various application icons.



Tooltips

- Remember that if you want to include a field in the tooltip that is not in your viz, just drag it to the Tooltip shelf
- You can click Insert within the Tooltip window to bring up the list of fields available to the tooltip



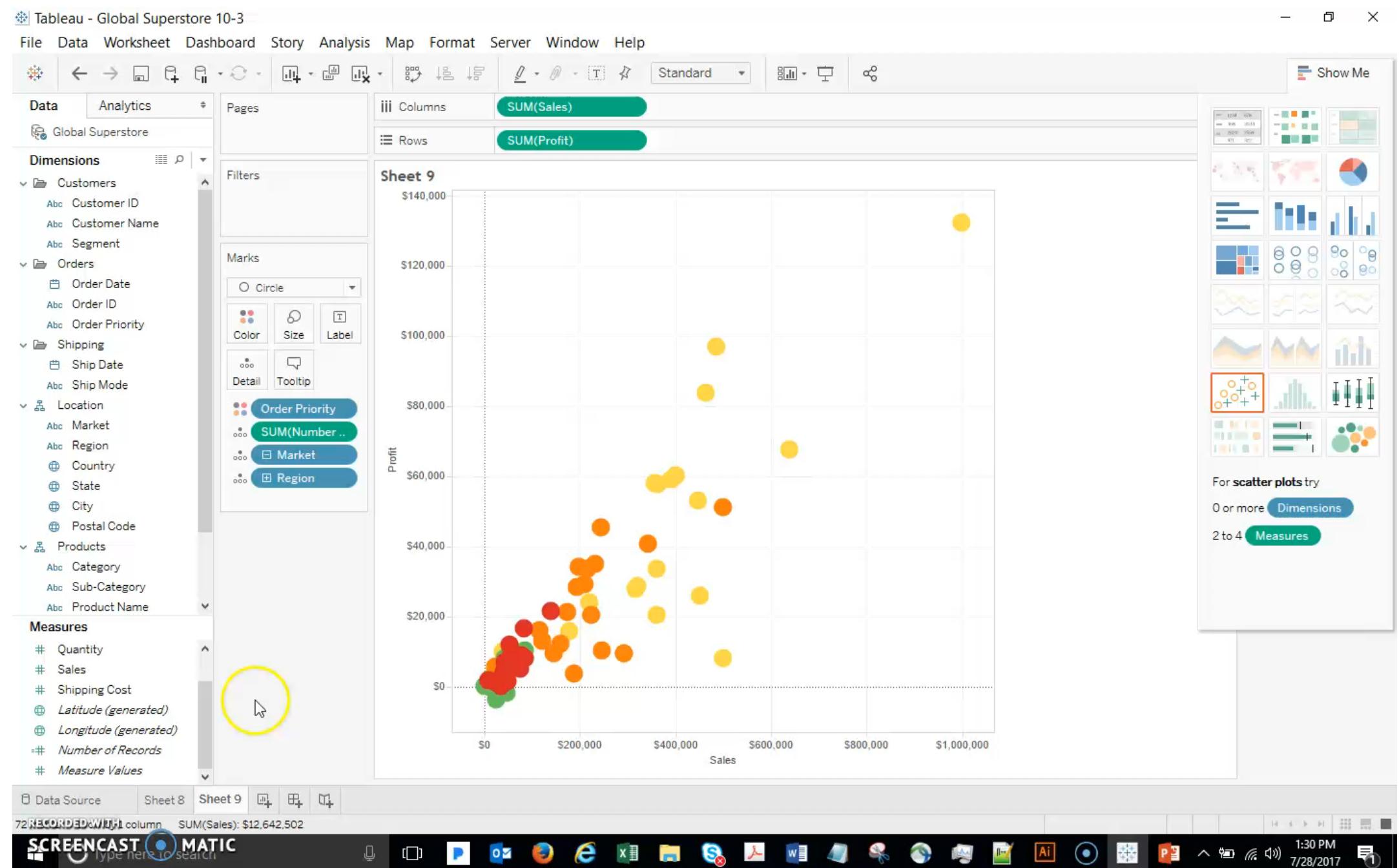


Tooltips

- In addition, if you prefer not to see the command buttons (keep only, group, view underlying data, etc), you can turn them off with a checkbox in the Tooltip window.
- With every viz you make, don't forget to pay attention to the Tooltips! They are often an important part of the user experience.



ToolTips





Formatting Specific Parts of the View

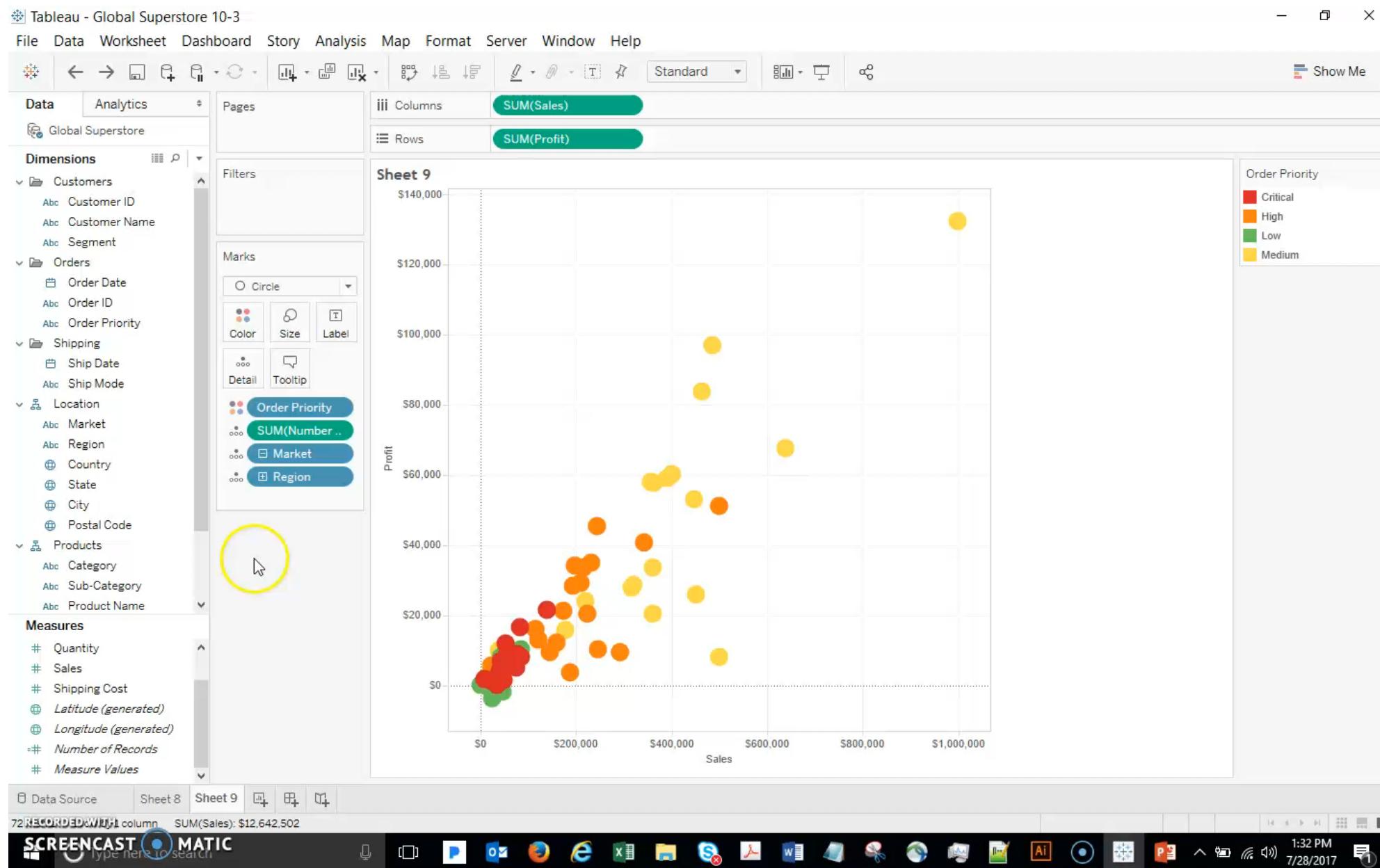
Formatting Specific Parts of the View



- Many parts of the view can be independently formatted. To see if a field, number, legend, or other part of the view can be formatted, right click or bring up the menu to see if the format option is available.
- For example, first create and show a Filter for Order Priority
- Let's format it. Click the drop down in the Filter and click Format Filters
 - The Format pane opens, offering us various options. Change the Font of the Body as an example. Just click the X to close the Format pane.



Formatting Specific Parts of the View





Editing & Formatting Axes

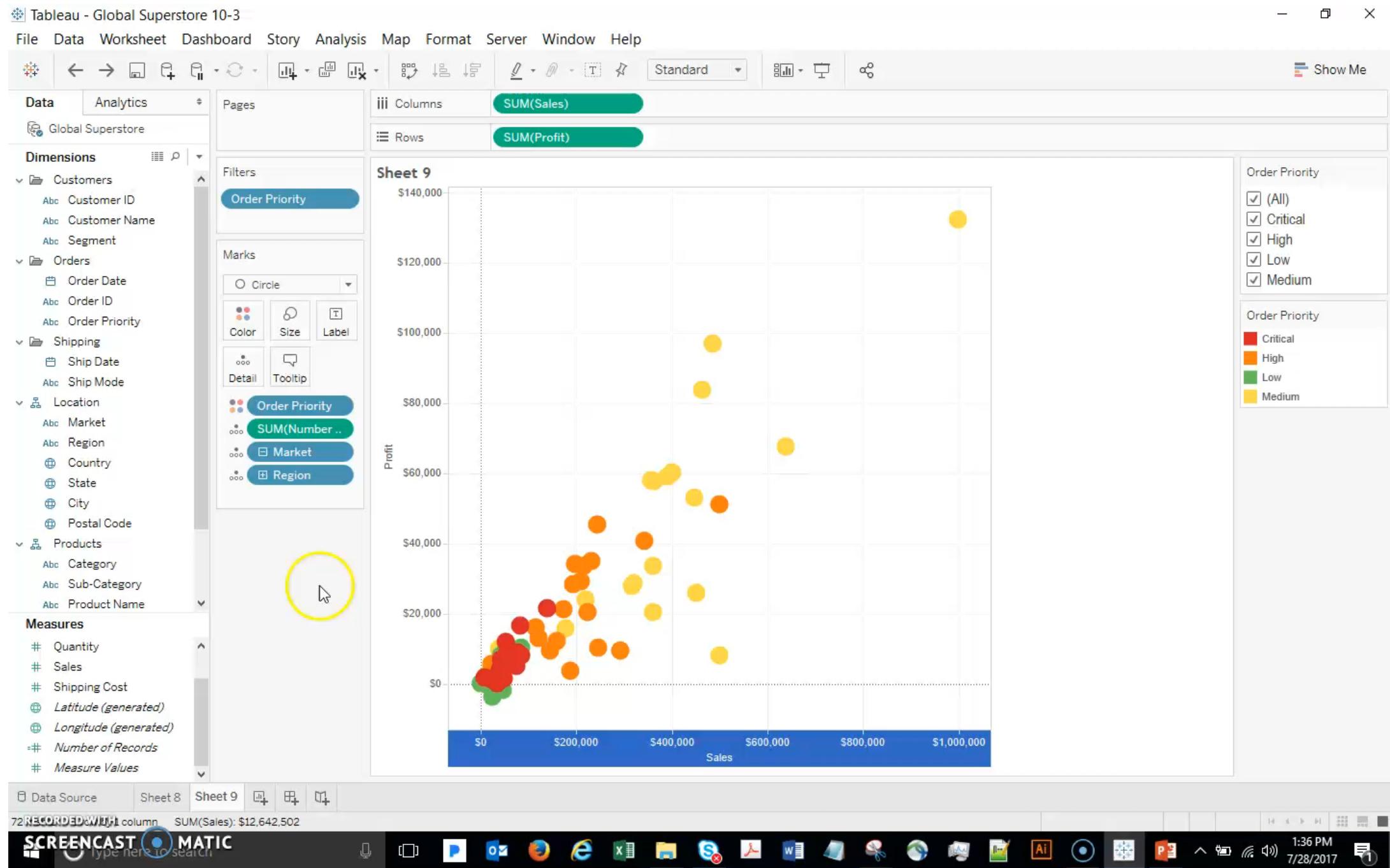


Editing & Formatting Axes

- To adjust anything having to do with the layout of the axis, right click and select **Edit Axis**.
- This dialog controls the range, tick marks, and the title.
- We can edit the range, to a **fixed range** that we prefer. You can change the tick marks, axis title, etc. also
- Additionally, this pin icon has appeared – this indicates the axis is no longer automatic.
 - Double click the pin and change it back to **Automatic axis**.
 - To hide an axis entirely, right click on the axis and uncheck **Show Header**.



Editing & Formatting Axes





Dynamic Titles



Dynamic Titles

The user-friendliness of a dashboard can often be improved by inserting a dynamic value into the view's title.

- Go to **Shipping Dashboard** and discuss is briefly
- For **Shipping Cost** bar chart, Right Click, Edit Title.
- If we click on Insert, we'll get a list of things that can be put in the title. This depends on what fields are in use in the view
- Here, we see that Ship Mode is an option.
- Let's rename this “Average Cost for” then insert Ship Mode and follow up with “Shipping”.
- Now this title dynamically changes as we change the Ship Mode filter.

Dynamic Titles



Tableau - dashboard_formatting_example 10dot3

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Dashboard Layout Device Preview

Size Automatic

Sheets

- Average Shipping C...
- Shipping Cost
- Running Total Shippi...

Objects

- Horizontal
- Vertical
- Text
- Tiled
- Floating
- Show dashboard title

Priority

- Critical
- High
- Medium
- Low

Ship Mode

- (All)

Market

- (All)

Running Total Shipping Costs

Same Day First Class Second Class Standard Class

Running Sum of Shipping Cost

Month of Ship Date

Shipping Cost

Ship Mode	Same Day	First Class	Second Class	Standard Class
Priority	\$62	\$60	\$56	\$30
High	\$40	\$39	\$32	\$29
Medium	\$28	\$29	\$22	\$16
Low	\$0	\$0	\$0	\$27

Avg. Shipping Cost

Average Shipping Cost by Country

© OpenStreetMap contributors

Data Source Average Shipping Cost by Count... Shipping Cost Running Total Shipping Costs Shipping Dashboard

RECORDED WITH SCREENCASTOMATIC

12 RECORDING WITH 1 column SUM of AVG(Shipping Cost): \$441 Highlighting on Order Priority

2:00 PM 7/28/2017



Format Menu



Format Menu

- Select **Format...Dashboard** from the top menus to open the format pane. We can customize a lot of things here.
- One thing to note is that if we **change the shading**, say to yellow, only dashboard items are shaded.
- Views have their own formatting, so if we want everything yellow we would need to edit each view individually.

D

Format Menu

Tableau - dashboard_formatting_example 10dot3

File Data Worksheet Dashboard Story Map Format Server Window Help

Dashboard Layout Device Preview

Size Automatic

Sheets Average Shipping C... Shipping Cost Running Total Shippi...

Objects Horizontal Image Vertical Web Page Text Blank Tiled Floating Show dashboard title

RECORDED WITH SCREENCASTOMATIC

Running Total Shipping Costs

Same Day First Class Second Class Standard Class

Average Cost for All Shipping

Same Day First Class Second Class Standard Class

Avg. Shipping Cost

Average Shipping Cost by Country

© OpenStreetMap contributors

Data Source Average Shipping Cost by Count... Shipping Cost Running Total Shipping Costs Shipping Dashboard

2:02 PM 7/28/2017



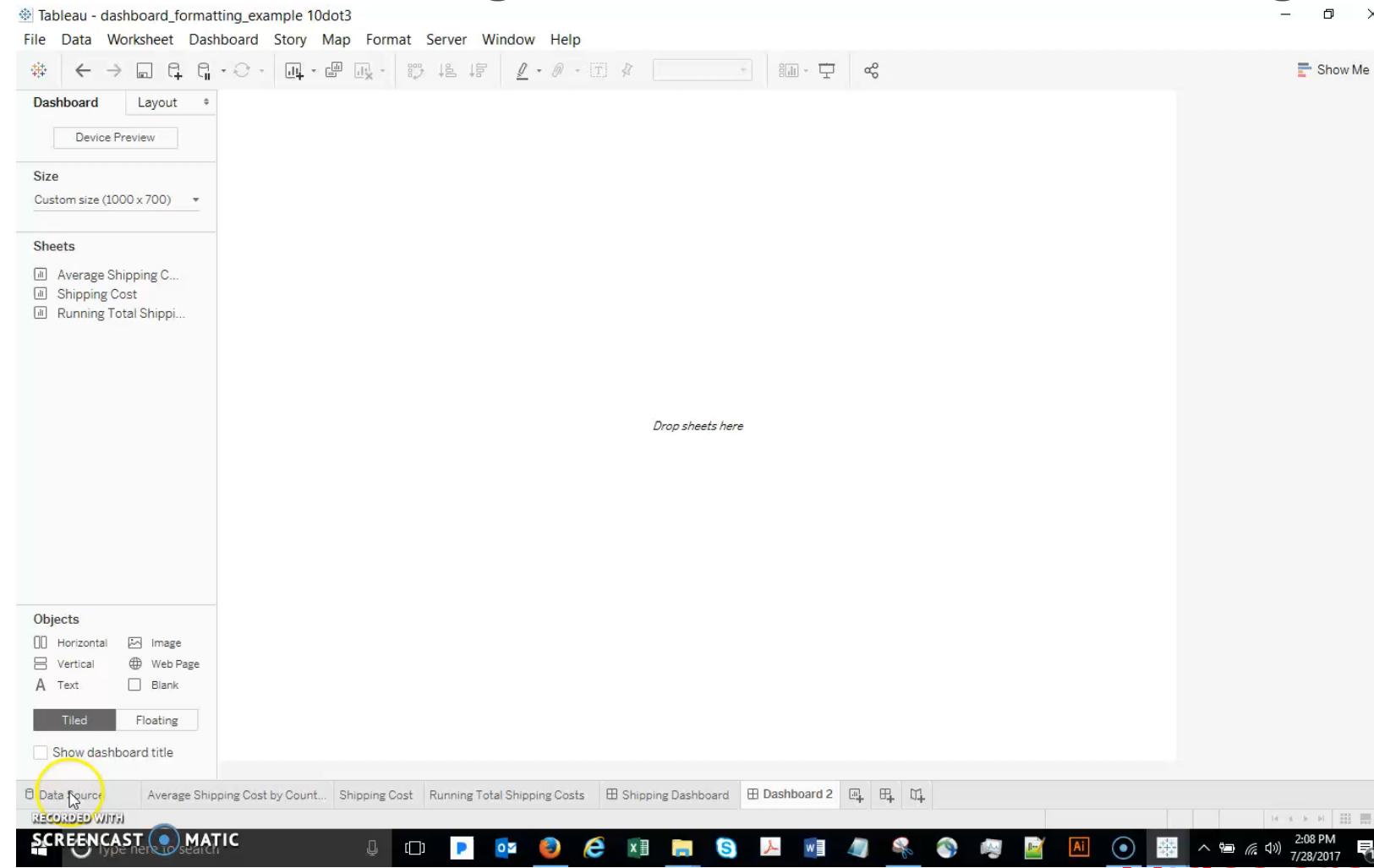
Images, Blank Objects, & Web Pages



Images, Blank Objects, & Web Pages

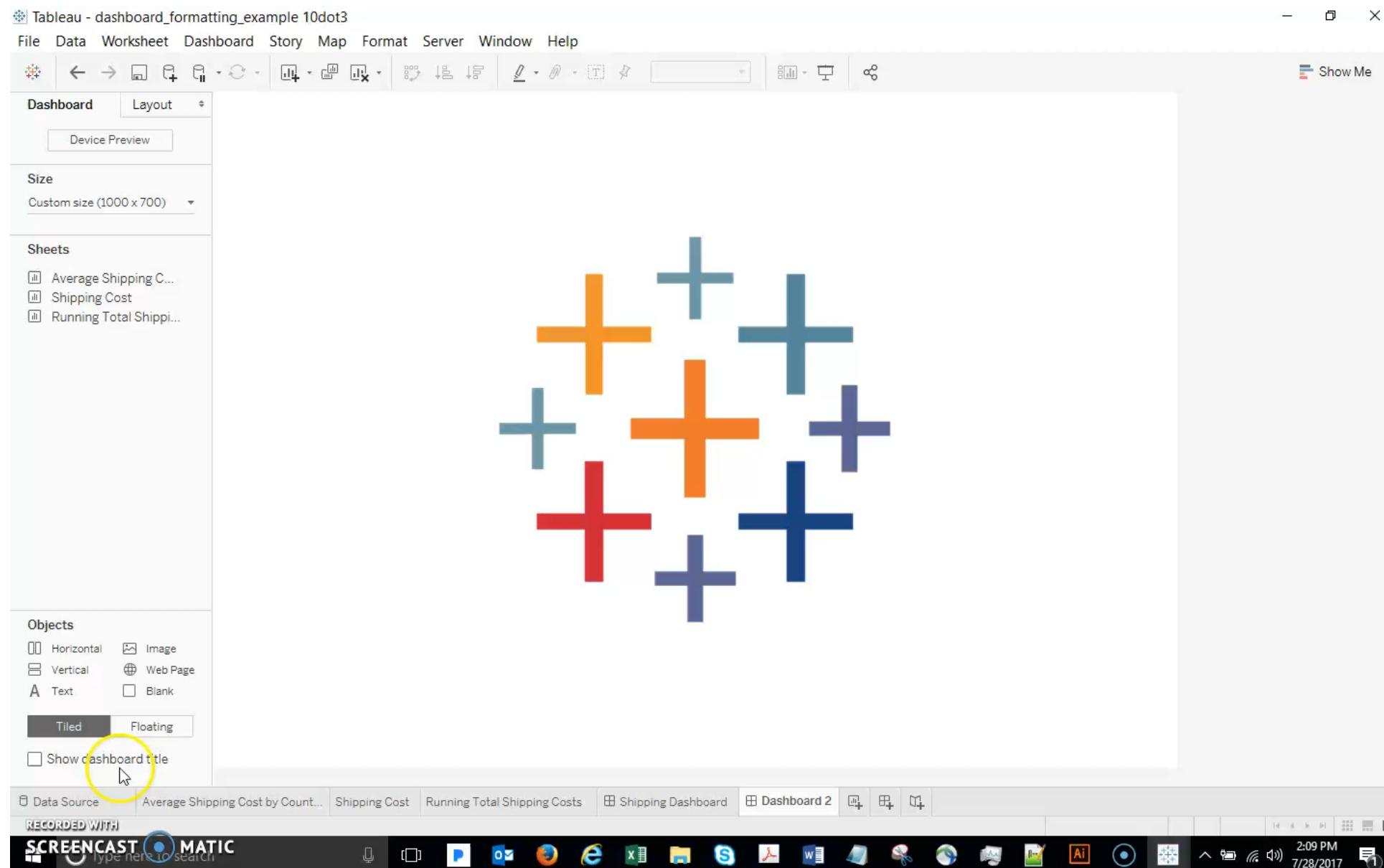
Create a new Dashboard.

- If you have an image (say a logo) that you want to add, you can drag out Image from the left and just navigate to where your image is stored.
 - You also have the option through the menu to make the image be a hyperlink to a URL



Images, Blank Objects, & Web Pages

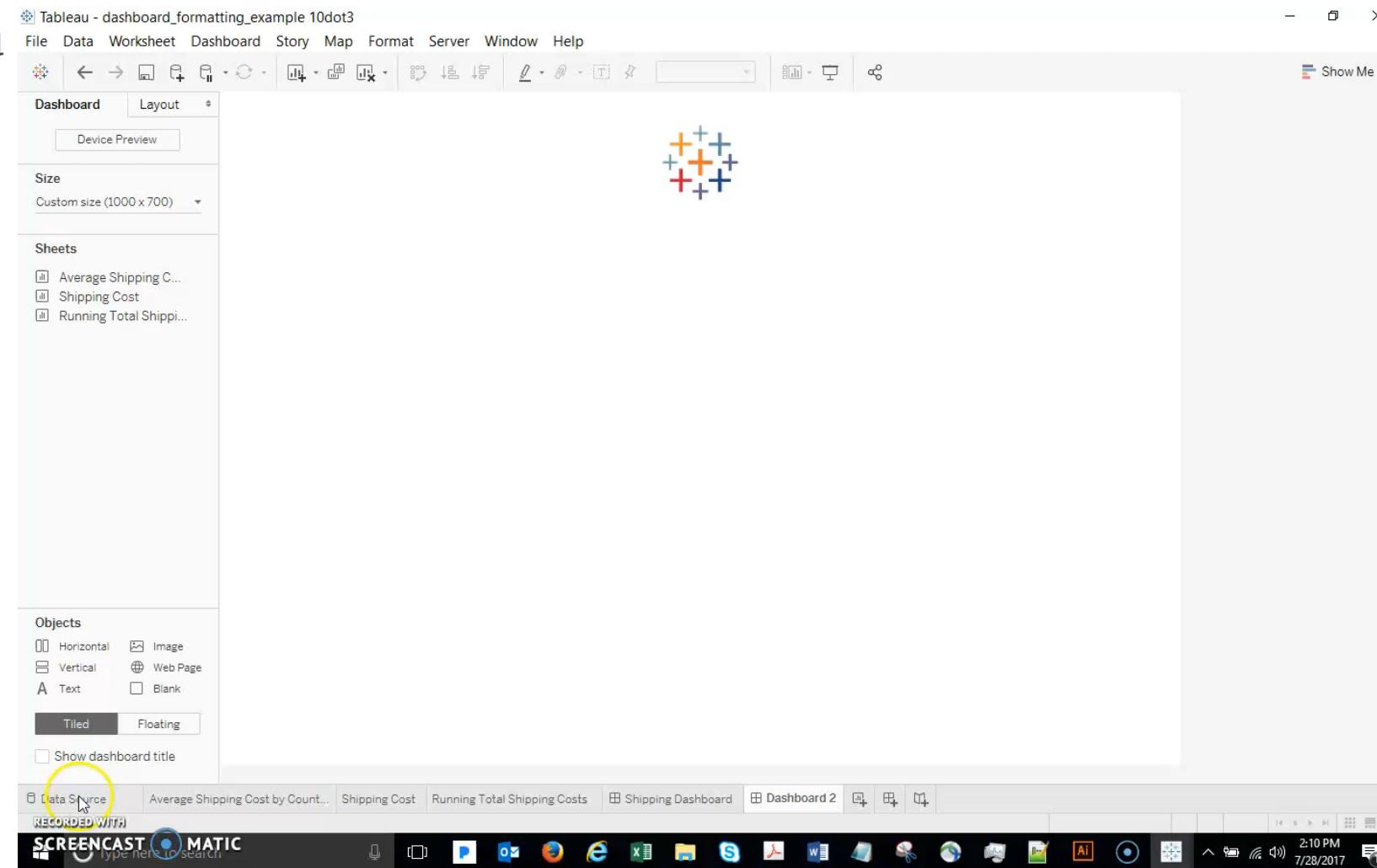
- Blank objects are useful if you want to add space to the viz, push an item to one side, etc.



Slides are not for Distribution

Images, Blank Objects, & Web Pages

- If we want to add a web page, we simply bring out the web page part
 - For the URL, I'll put google.com (or your school's URL)
 - This is treated just like any other part of the dashboard (drag out a worksheet, resize things)
 - And it's still a



Slides are not for Distribution



Simple Calculations in Tableau



Simple Calculations in Tableau

- Start with Global Superstore
- Let's create a simple profit viz – Profit to Columns, Market and Subcategory to Rows, Filter Category to just Furniture.



Simple Calculations in Tableau

The screenshot shows the Tableau Data Source view for a Global Superstore 10-3 dataset. The left sidebar lists dimensions and measures. A yellow circle highlights the cursor over the 'Sales' measure in the Measures section.

Dimensions:

- Customers:
 - Customer ID
 - Customer Name
 - Segment
- Orders:
 - Order Date
 - Order ID
 - Order Priority
- Shipping:
 - Ship Date
 - Ship Mode
- Location:
 - Market
 - Region
 - Country
 - State
 - City
 - Postal Code
- Products:
 - Category
 - Sub-Category
 - Product Name

Measures:

- # Discount
- # Profit
- # Quantity
- # Sales
- # Shipping Cost
- (Latitude (generated))
- (Longitude (generated))

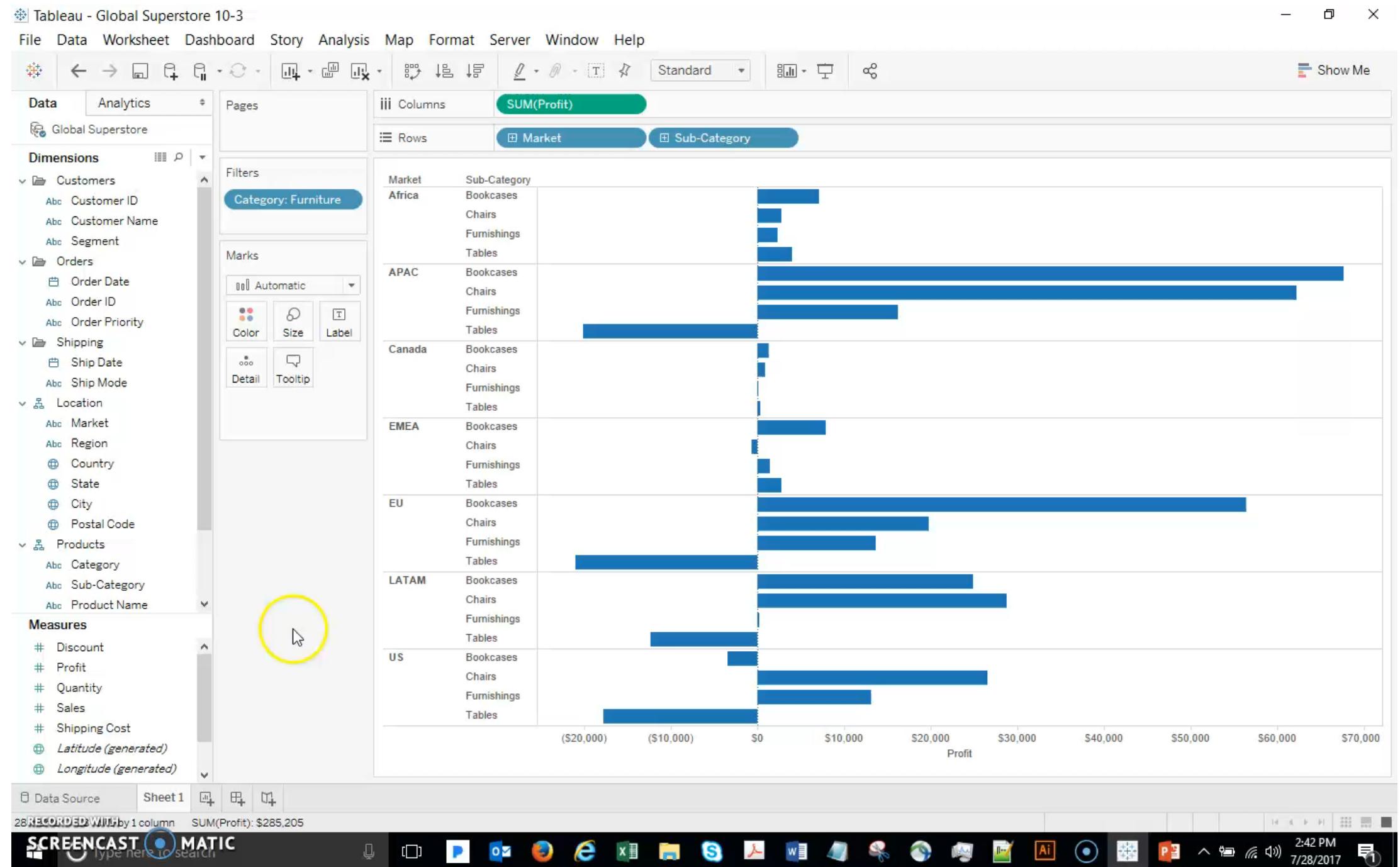


Simple Example 1:

- What if we want, for emphasis, to have one distinct color for positive and one for negative?
- Calculated fields are created by defining a formula:
- Drop-down Analysis...Create Calculated Field
- **IF SUM([Profit]) > 0 THEN "positive" ELSE "negative" END**
 - Explain the SUM here – remember that Measures need an aggregation level
 - Note that we are ignoring if Profits are actually zero for simplicity's sake
 - name it “sign of profit”, make a comment in your code using //
- Note the little equal sign in front of the field that indicates it's a calculated field, not from the data source
- Bring this calculated field to color



Simple Example 1:

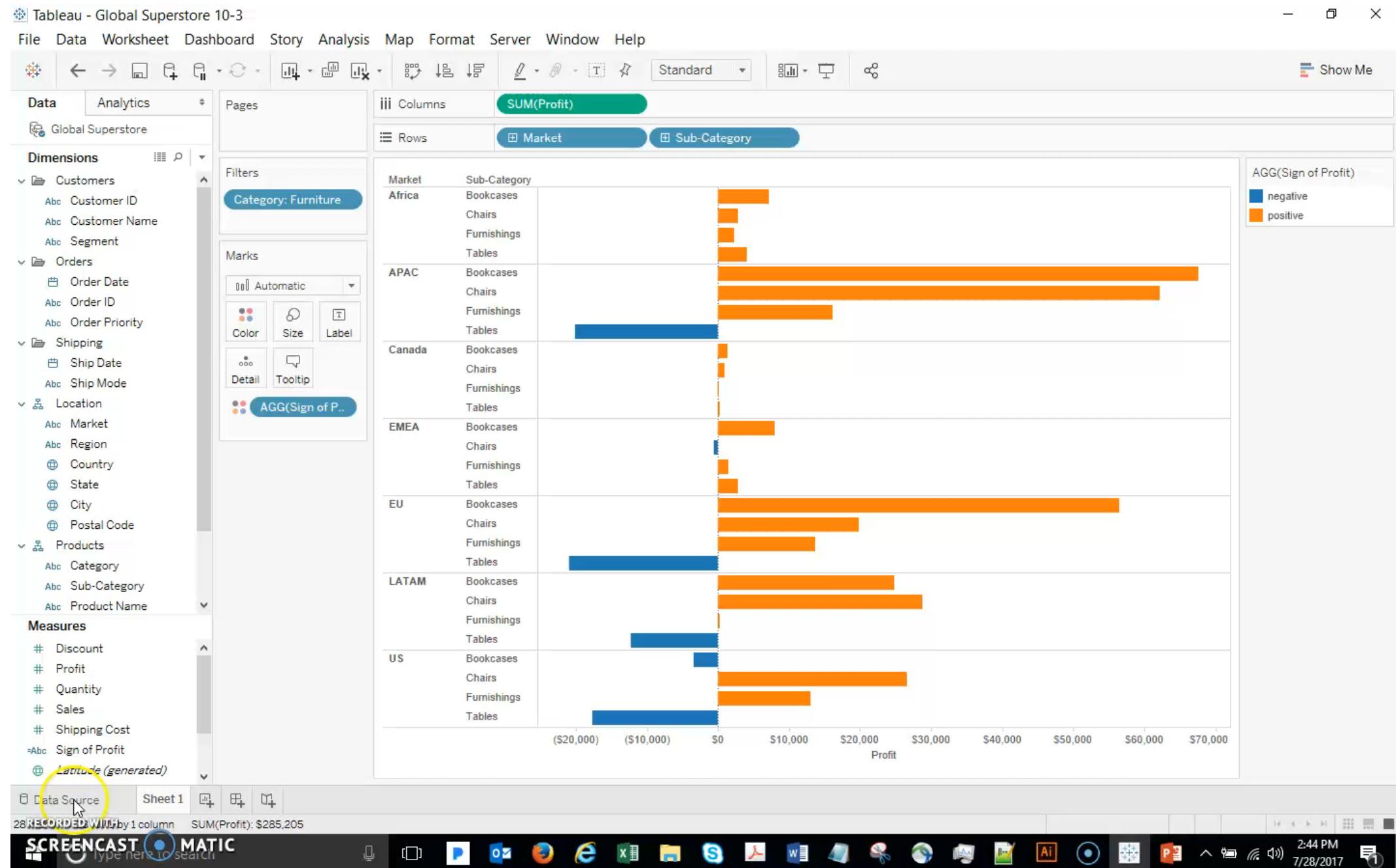




Simple Example 2:

- There might be a field of interest to me that is not in the data. For example, maybe I want to visualize Cost, defined as Sales - Profit.
- Add a Calculated field for that: $\text{SUM}([\text{Sales}]) - \text{SUM}([\text{Profit}])$. Name it “Cost”, drag Profit out of view and replace with Cost.

Simple Example 2:





Time between Dates



Time between Dates

- A common question could be something like “How much time did something take?” when you have a start and an end date in your data. A date calculation can handle that.
- Let’s look at the average time it takes for an order to ship
 - Right click in the data window and select Create Calculated Field
 - Name it “**Time to Ship**”
 - We’ll use a function called DATEDIFF to get the amount of time between Order Date and Shipping Date.

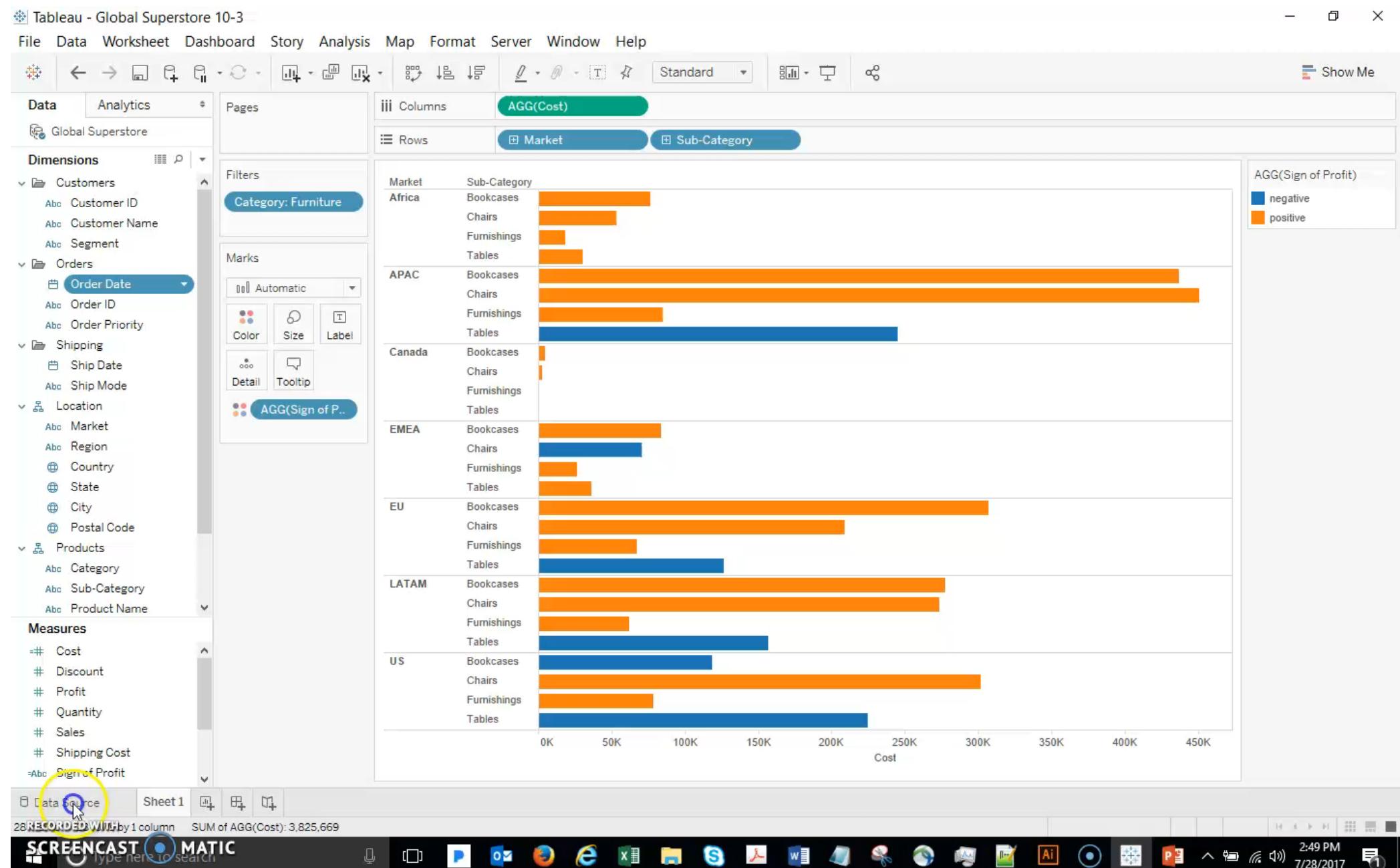


Time between Dates

- Search for DATEDIFF; the first argument needed is the date part. Type 'day' and a comma
- Next, drag out Order Date, another comma, and Ship Date
 - `DATEDIFF('day', [Order Date], [Ship Date])`
- Be sure to check that the calculation is valid, then click OK



Time between Dates



Slides are not for Distribution



Time between Dates

- Now we can color our countries by using Time to Ship
 - Select Country and Time to Ship, select Filled Map from Show Me
 - We'll change the aggregation to Average
 - You can edit colors if you prefer – maybe Grey Sequential for an easy interpretation of “darker is worse”



Time between Dates

Tableau - Global Superstore 10-3

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Show Me

Data Analytics

Global Superstore

Dimensions

- Customers
 - Customer ID
 - Customer Name
 - Segment
- Orders
 - Order Date
 - Order ID
 - Order Priority
- Shipping
 - Ship Date
 - Ship Mode
- Location
 - Market
 - Region
 - Country
 - State
 - City
 - Postal Code
- Products
 - Category
 - Sub-Category
 - Product Name

Marks

- Automatic
- Color
- Size
- Text
- Detail
- Tooltip

Pages

Columns

Rows

Sheet 2

Drop field here

Drop field here

Drop field here

DataSource Sheet 1 Sheet 2

RECORDED WITH SCREENCASTOMATIC

© DataTiles.io

2:51 PM 7/28/2017

Slides are not for Distribution



Introduction to Table Calculations

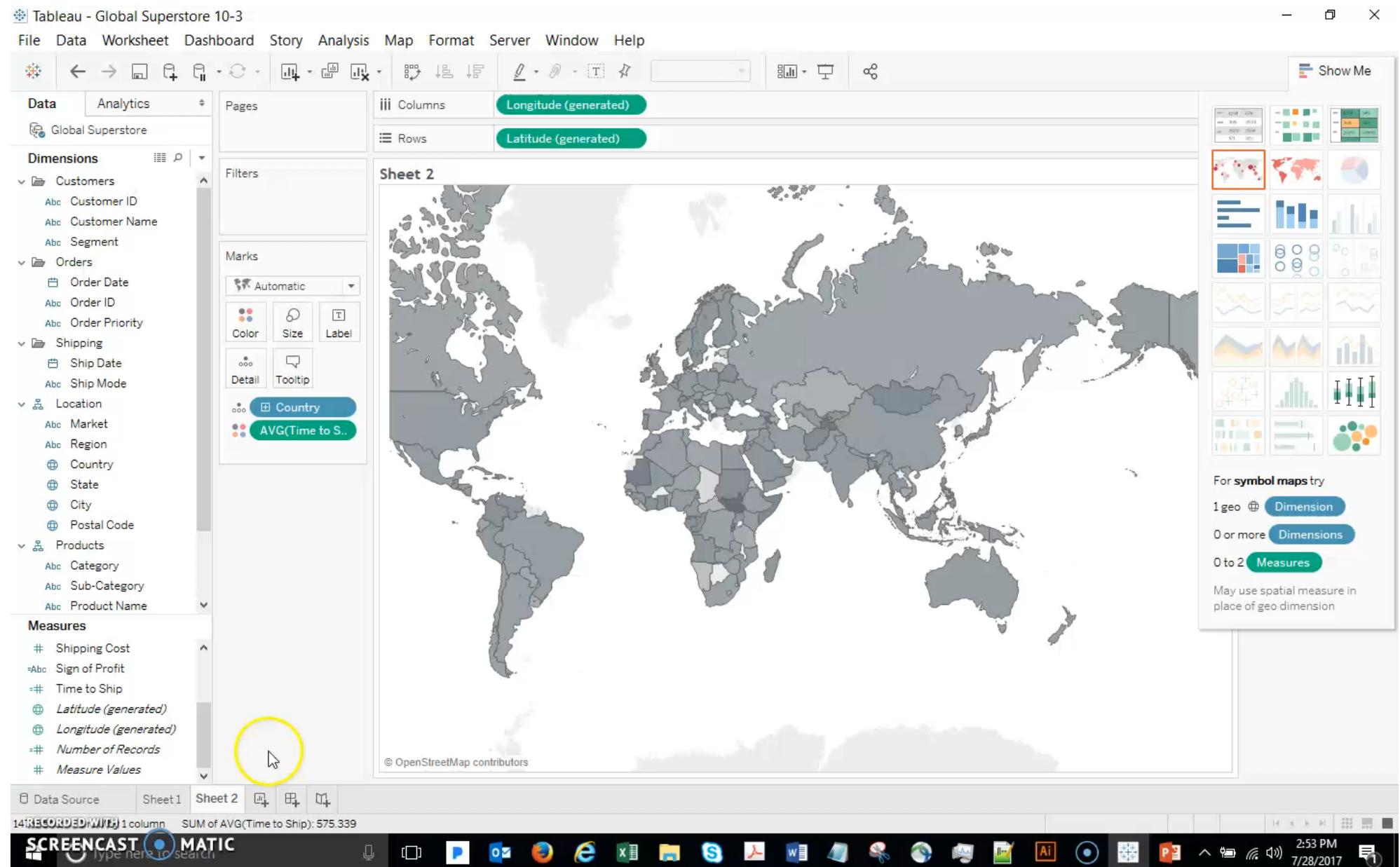


Introduction to Table Calculations

- A Table Calculation is a secondary calculation that is performed on top of a returned result set. Tableau includes a set of pre-defined, commonly used computations called Quick Table Calculations. These include options like Running Total, Percent of Total, and Year over Year Growth. Let's try one.
- New Sheet, Category to Columns, Sales to Rows. Order Date to Columns, Make it a continuous Date, Expand to Quarters. Sales to Rows again, drop-down and select the Running Total Quick Table Calculation. Now you have a set of graphs with quarterly sales as well as a running total.



Introduction to Table Calculations



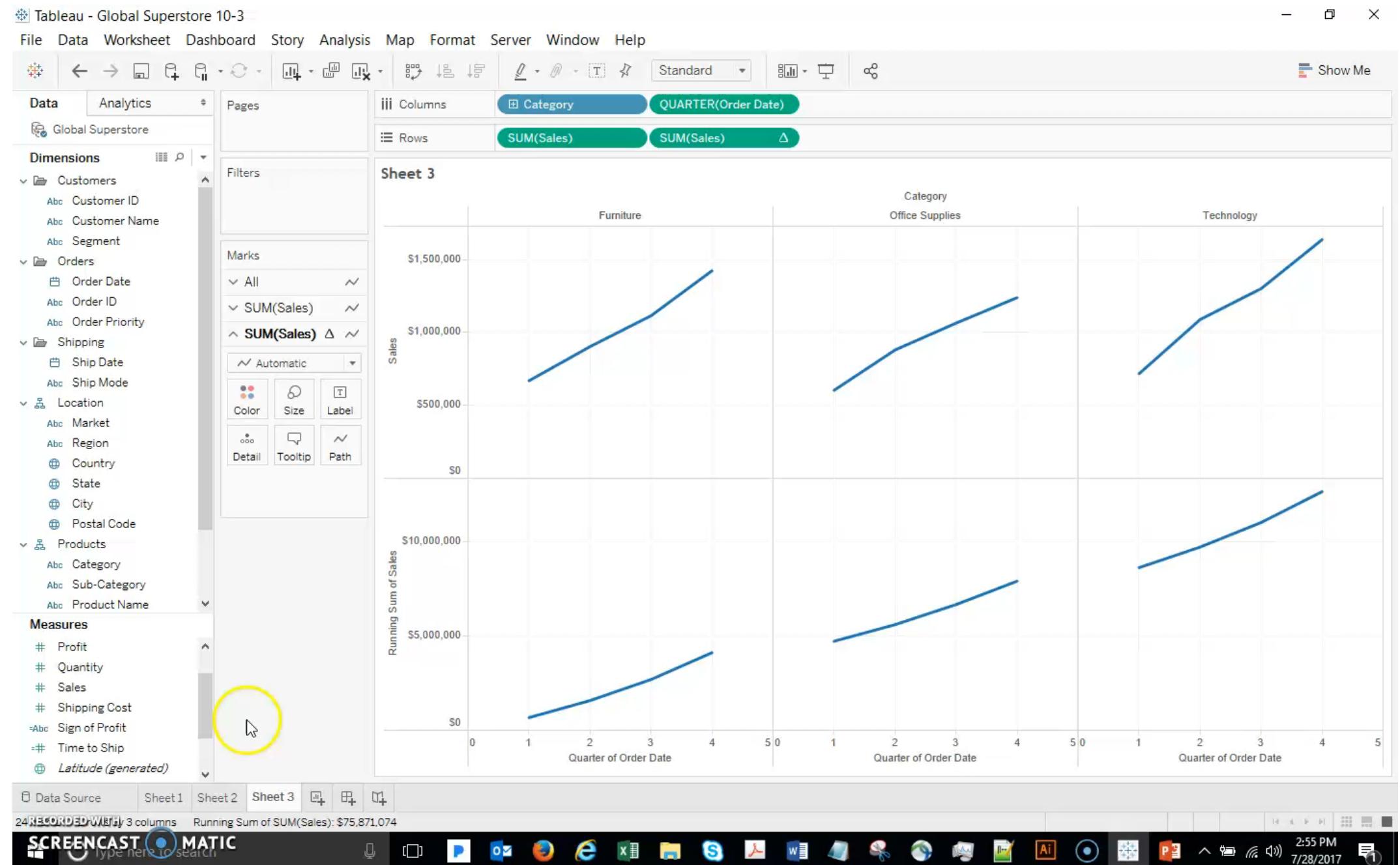
Introduction to Table Calculations



- If you want to the graphs on top of each other, you can drag your Table Calc to the right axis for a dual axis chart, then you could right click the axis and synchronize axis. But this is probably much less useful than separate charts in this case.

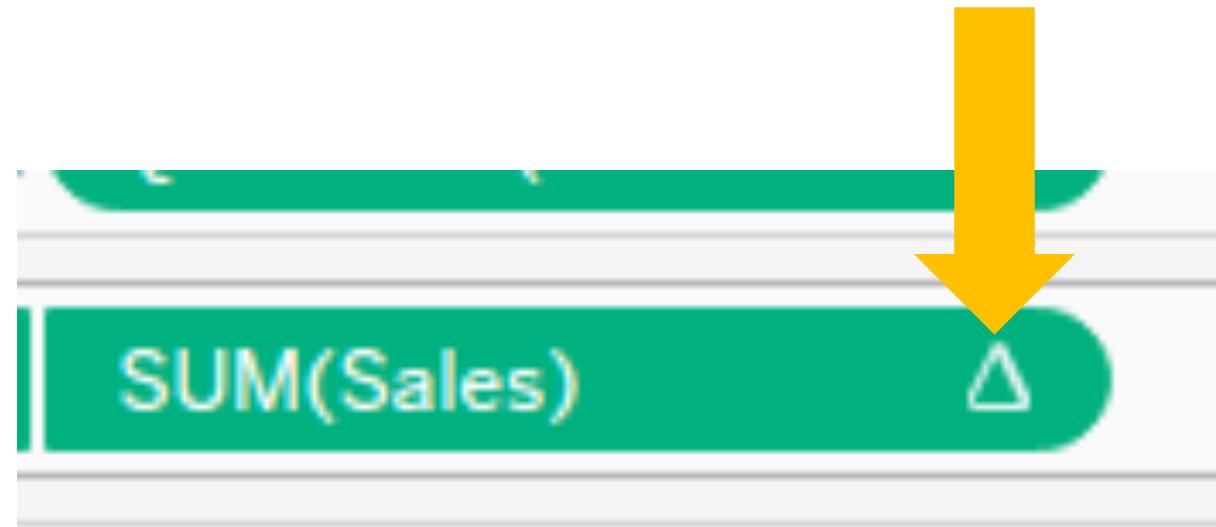


Introduction to Table Calculations



Introduction to Table Calculations

- You can tell that something is a table calculation by the delta symbol on the pill





Data Validation - Nulls to Zeroes

Data Validation - Nulls to Zeros



- A common issue with real-world datasets is the presence of null values, or blanks, in the data. This can make data analysis challenging and frustrating (as you may know, some statistical techniques will not work at all if there are nulls in the data). Luckily, with Calculations we can easily remove Nulls from our data:
- Now open **Sales_Quotas**. In this dataset, we have sales quotas for some, but not all of our countries.
- **Double-click Country** to get a dot everywhere we do business. I want to see how we are doing relative to quotas everywhere. But I have included an issue in the data that needs to be fixed: Right-click somewhere in the map and **view data**. There are several nulls for Quotas.



Data Validation - Nulls to Zeros

The screenshot shows the Tableau interface with the following details:

- Title Bar:** Tableau - Sales_Quotas10dot3
- Menu Bar:** File, Data, Worksheet, Dashboard, Story, Analysis, Map, Format, Server, Window, Help
- Toolbars:** Standard, Show Me
- Left Panel (Data Shelf):**
 - Dimensions:** Sales Quotas, Country
 - Measures:** Quotas, Sales, Latitude (generated), Longitude (generated), Number of Records, Measure Values
- Middle Panel (Marks Shelf):** Marks dropdown set to Automatic, with options for Color, Size, Text, Detail, and Tooltip.
- Right Panel (Canvas):** A single sheet titled "Corrected Quotas". The canvas has three drop zones:
 - Top: "Drop field here"
 - Left: "Drop field here"
 - Bottom: "Drop field here"
- Bottom Panel (Toolbar):** Data Source, Corrected Quotas, Difference from Quota, and a toolbar with various icons.
- Taskbar:** RECORDED WITH SCREENCASTOMATIC, a search bar, and system icons.
- System Status:** 2:59 PM, 7/28/2017

Data Validation - Nulls to Zeros



- Our main goal is to compare the sales to their quotas, but we cannot compare a number to a null. We need to convert our nulls to zeros so that we can do a comparison.
- right-click in the data window and select create calculated field.
- Name it “Corrected Quota”.
 - And type ZN(
 - The ZN function returns a zero wherever there is a null. We can think of this function as asking “Is there a quota for this country? If not, give it a zero.”

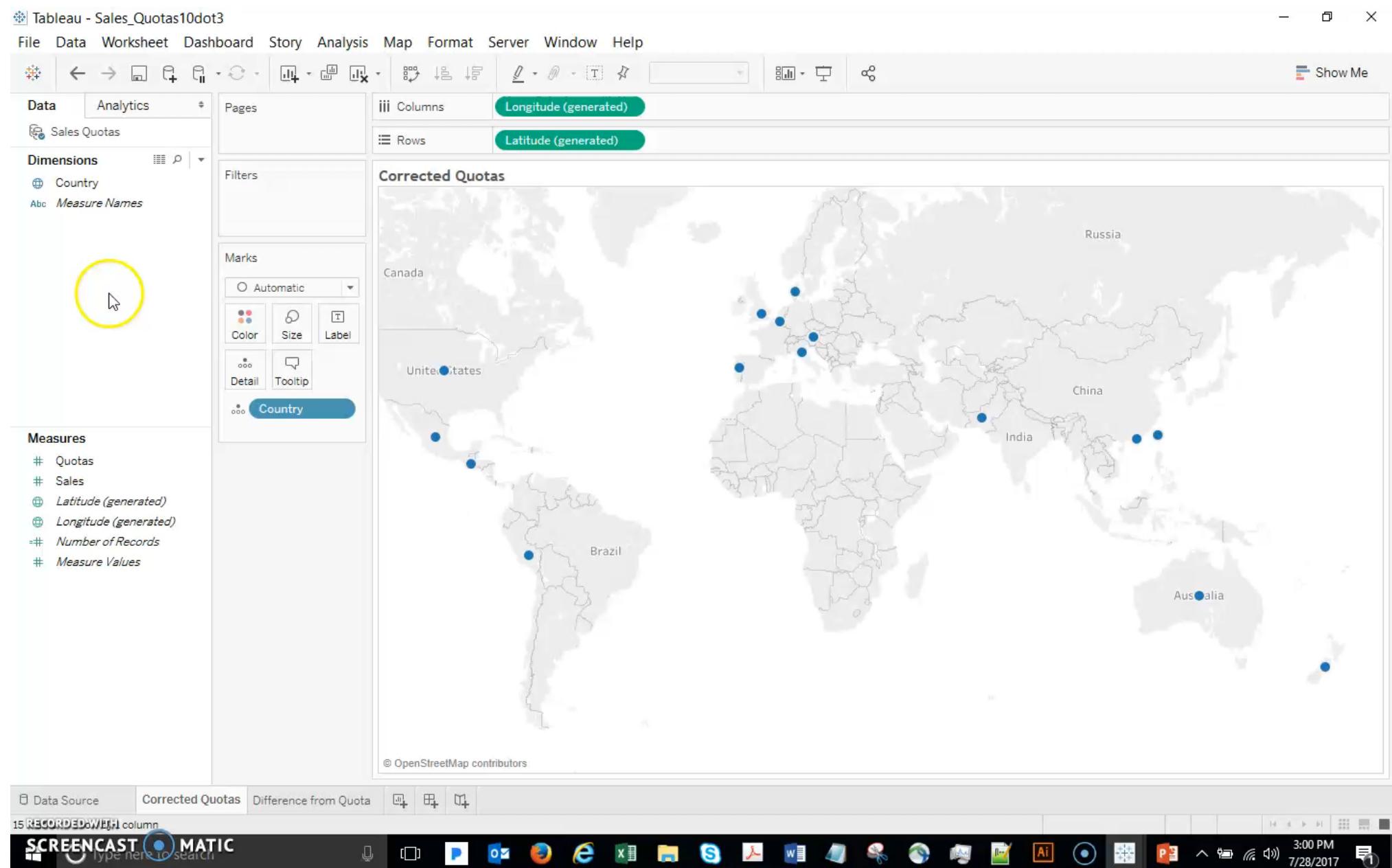
Data Validation - Nulls to Zeros



- Drag out Quota, and close the parentheses.
- Now if we replace Quota with our Corrected Quotas we see zeros in lieu of the blanks we had before.
- Click OK



Data Validation - Nulls to Zeros

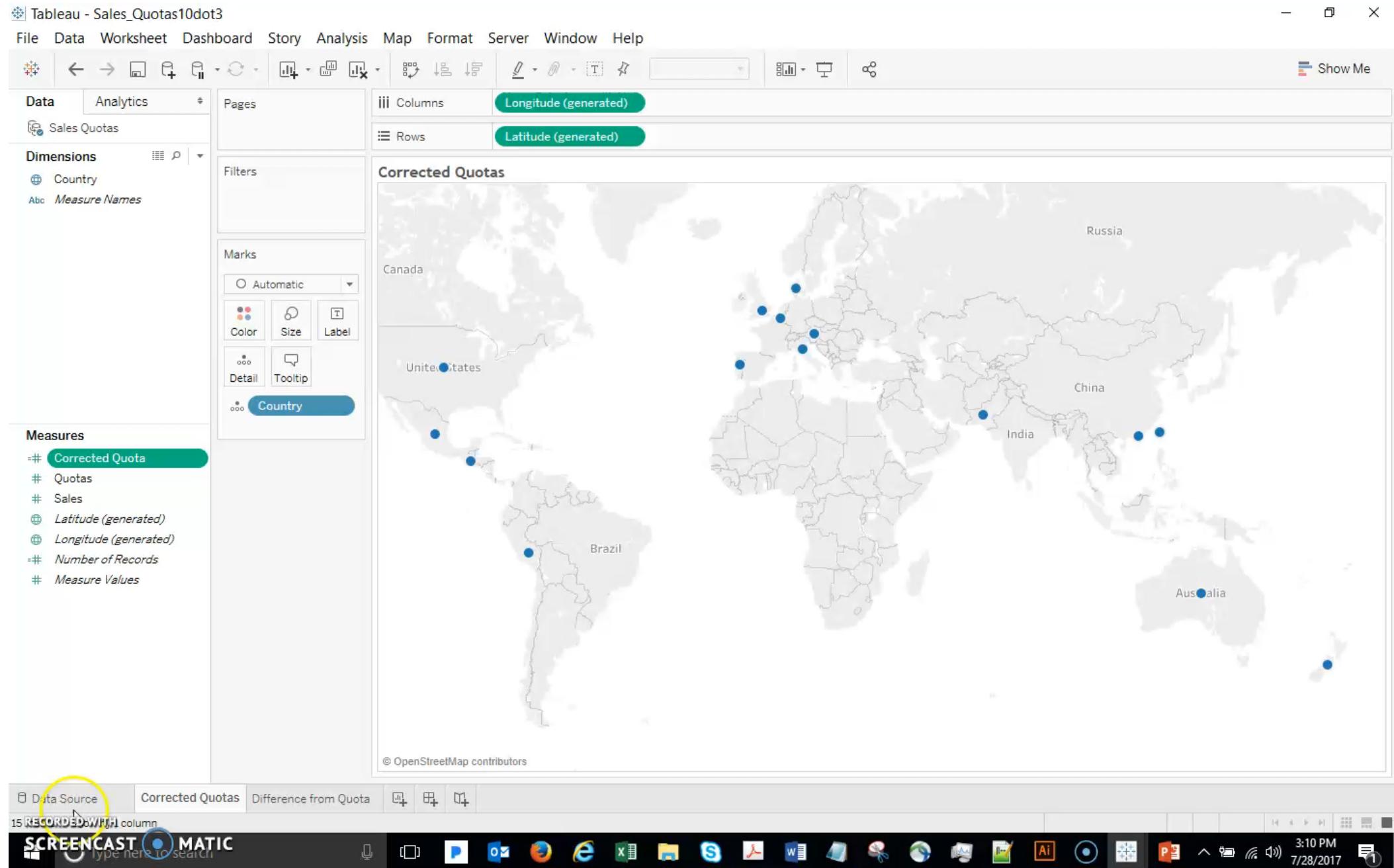


Data Validation - Nulls to Zeros



- Now we can visualize how different our sales are from our quotas. Again, we can right-click and create a Calculated Field
 - We'll call this our "QuotaCompare"
 - $\text{SUM}(\text{Sales}) - \text{SUM}(\text{Corrected Quota})$
- Double-click QuotaCompare to bring it into the view
- Taiwan, Portugal, and Austria look pretty good. To get more data into the viz, add Corrected Quota to the Tooltip. Why do these 3 countries look so strong?

Data Validation - Nulls to Zeros







Introduction to Maps



Working with Maps



Working with Maps

- Open a new worksheet. Double click State. This shows states across the entire globe. Let's focus on the US – Drag Country to Filter shelf, uncheck US, check Exclude, click OK.
- What this map tells me is that we've done business in these states. But, how much business? How much profit have we generated in each state?



Working with Maps

The screenshot shows the Tableau Data Source view for "Global Superstore 10-3". The interface includes a top navigation bar with File, Data, Worksheet, Dashboard, Story, Analysis, Map, Format, Server, Window, and Help. On the left, a sidebar displays the data hierarchy under "Data" and "Analytics". The "Dimensions" section lists Customers, Orders, Shipping, Location, and Products. The "Measures" section lists Discount, Profit, Quantity, Sales, Shipping Cost, and generated Latitude and Longitude fields. A yellow circle highlights the cursor over the "Measures" section. The main workspace is titled "Sheet 8" and contains two "Drop field here" placeholder areas.

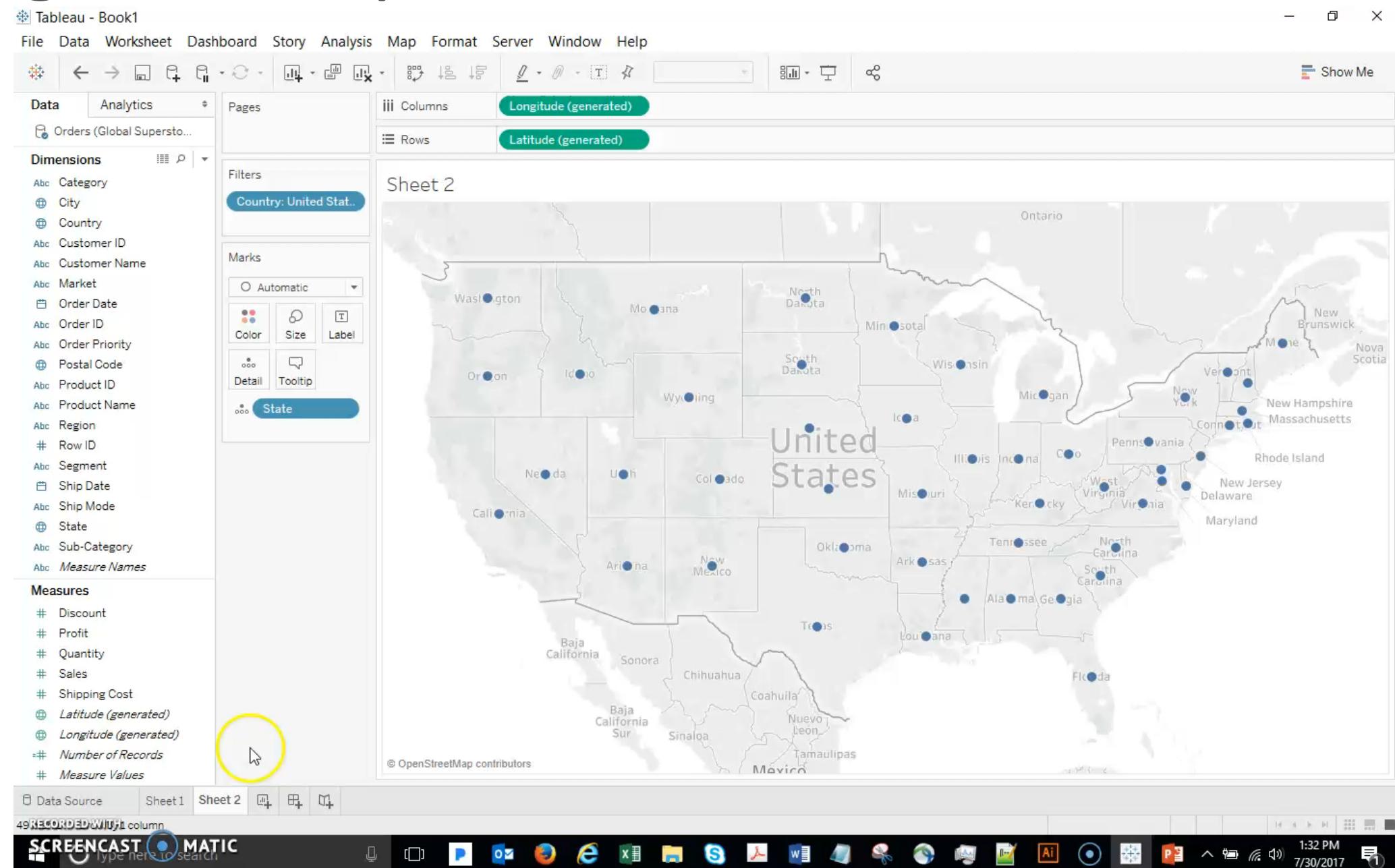


Working with Maps

- Drag Profit field to Color – remember to change from red/green to orange-blue!
- Overall, Texas looks like a challenging market for us. Now let's say we want to drill down and just look at the furniture category:
- Right-Click on Category....Show Filter. Select just the Furniture category.



Working with Maps





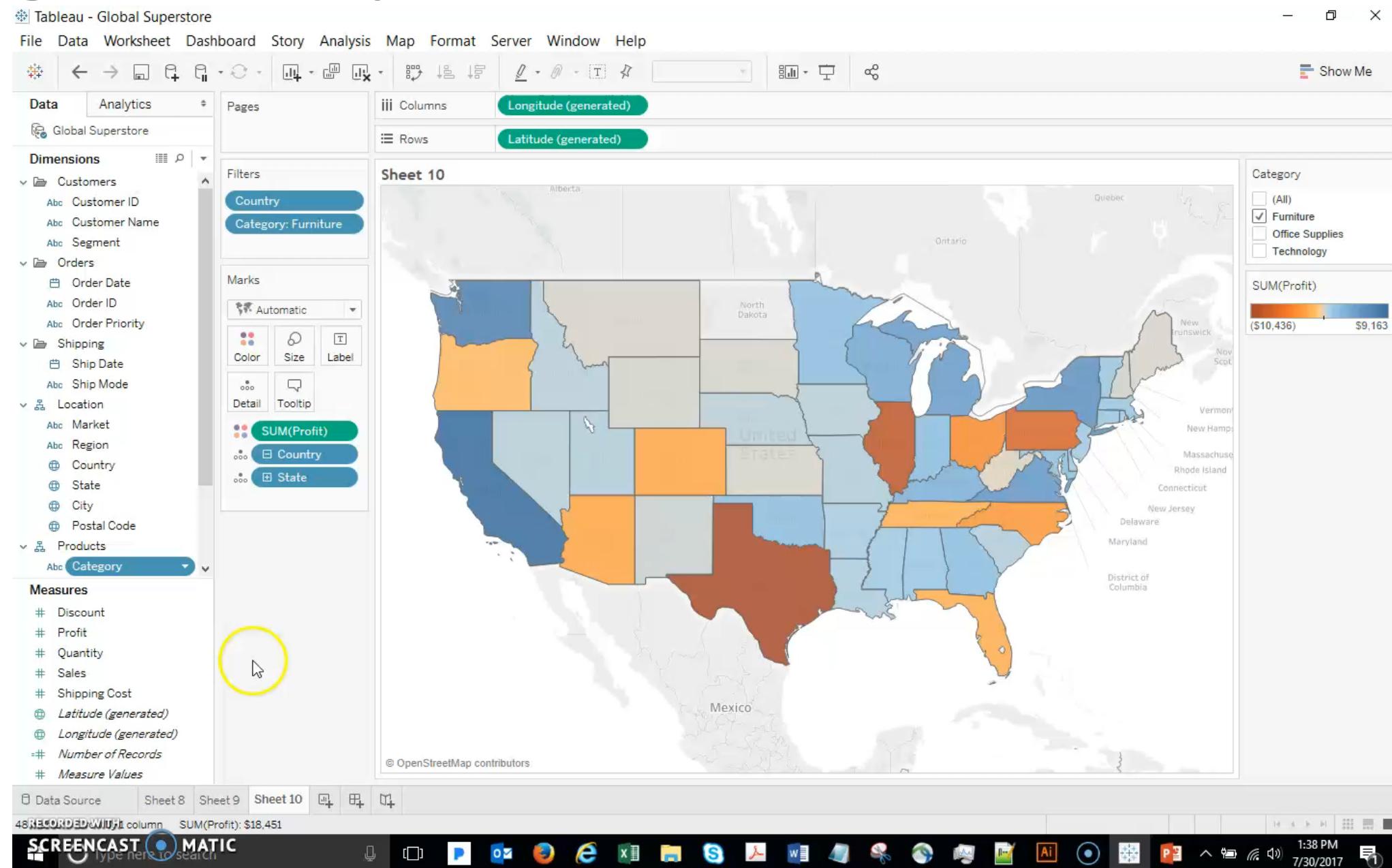
Working with Maps

Filled maps like this are great when you just want to see one measure on your map. But I would like to see both profits and sales. Only one attribute can be on color, so let's drag Sales to Size on the Marks card. Now we see them both together (click Size to make the bubbles larger).

Finally, it is easy to explore different levels of geography: click the plus by State on the Marks card to switch from State to City bubbles, then Undo.



Working with Maps





Map Options

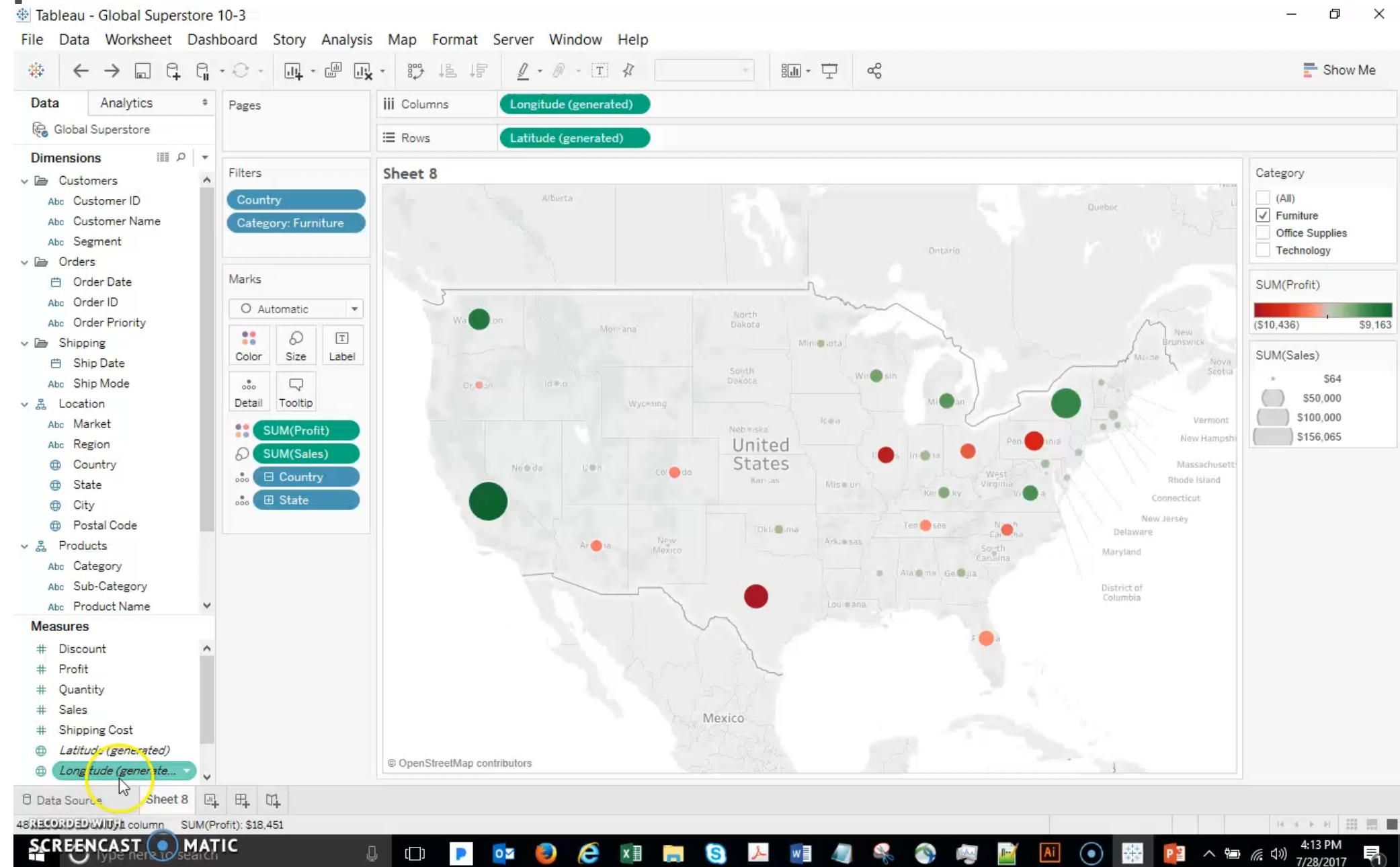


Map Options

- When working with maps, there are map-specific formatting options available.
- Map drop-down, then select **Map Layers**.
- This pane opens over the data window just like the formatting pane.
 - Here, we can change our map style – Dark, Light, or Normal
 - Control the washout
 - And control map layers, such as borders and names.
 - Some options, such as streets and highways, can be turned on only when the map is sufficiently zoomed in (currently greyed out)
 - Hover over the map and click the magnifying glass. Let's search for Houston, and now we can turn on Streets and Highways.



Map Options



Slides are not for Distribution



Dual Axis Maps

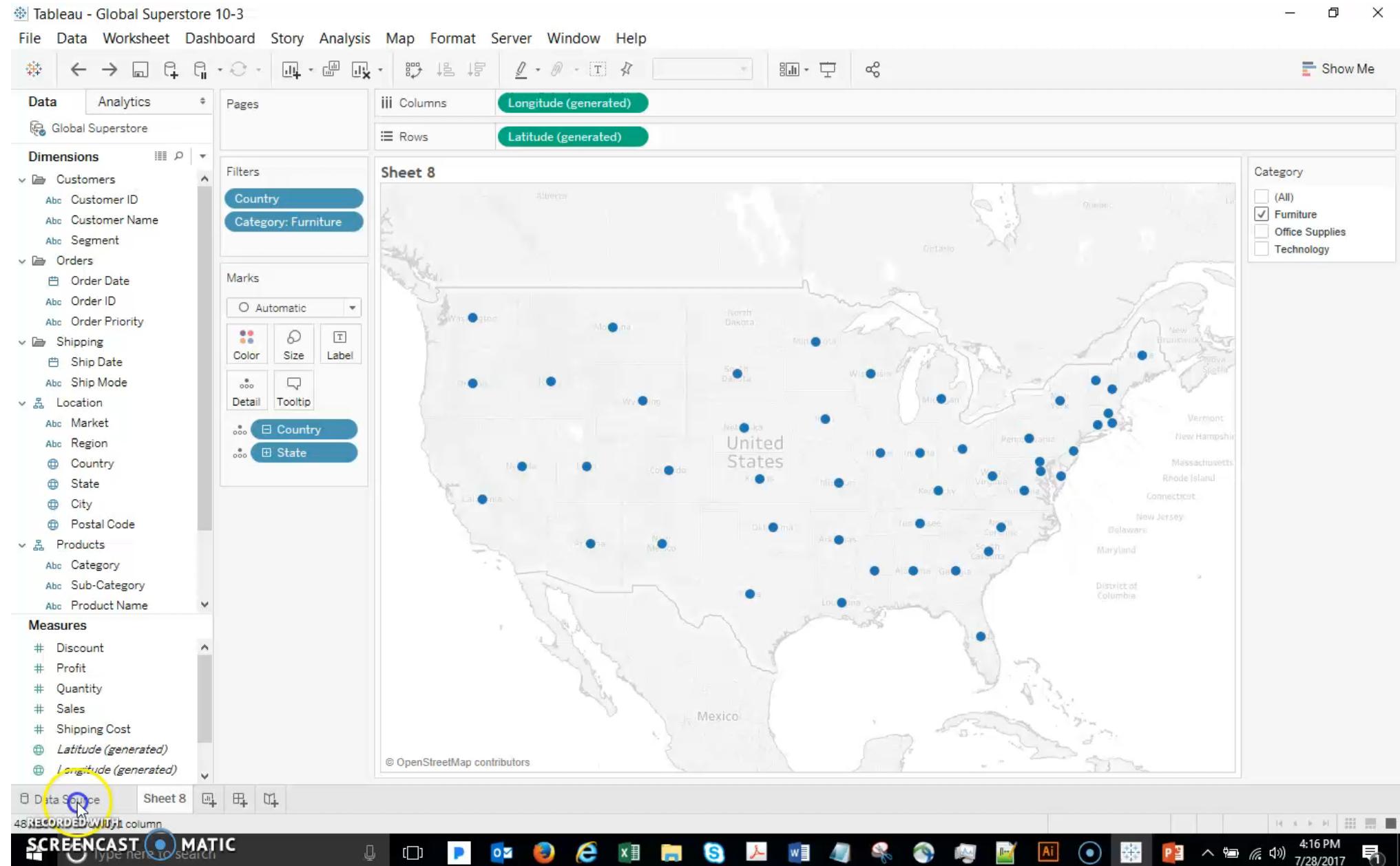


Dual Axis Maps

- If you want to put multiple data fields onto your map, a more flexible option can be to create a dual axis map. Open a **new sheet**. Go through the steps from the previous example to recreate your map with a simple dot for each state in the US.
- Drag **Longitude** to **Columns** again – now you have two identical maps.
- Select a point on the left map, then drag **Sales** to **Color** – now that map is shaded by sales
- Select a point on the right map, change the **Mark** to a **Square**. Make the color **Black**.
- Now click the drop-down on second **Longitude** pill in **Columns** and select **Dual Axis**.



Dual Axis Maps





Using Map Services



Using Map Services

- Mapbox enables a Tableau user to incorporate a very wide variety of background maps. This can be useful when you want landmarks, etc. to appear on your map, or when a specific map feature not included with Tableau (e.g. topography) is relevant to your analysis.
- Go to [Mapbox.com](https://mapbox.com) and sign up for a free account.
- From your Mapbox account, copy your account token (go to Account...API access tokens, and there is a place to copy it).
- From the Maps menu, choose Background Maps > Map Services.
- On the Map Services dialog, click Add > Mapbox Maps.



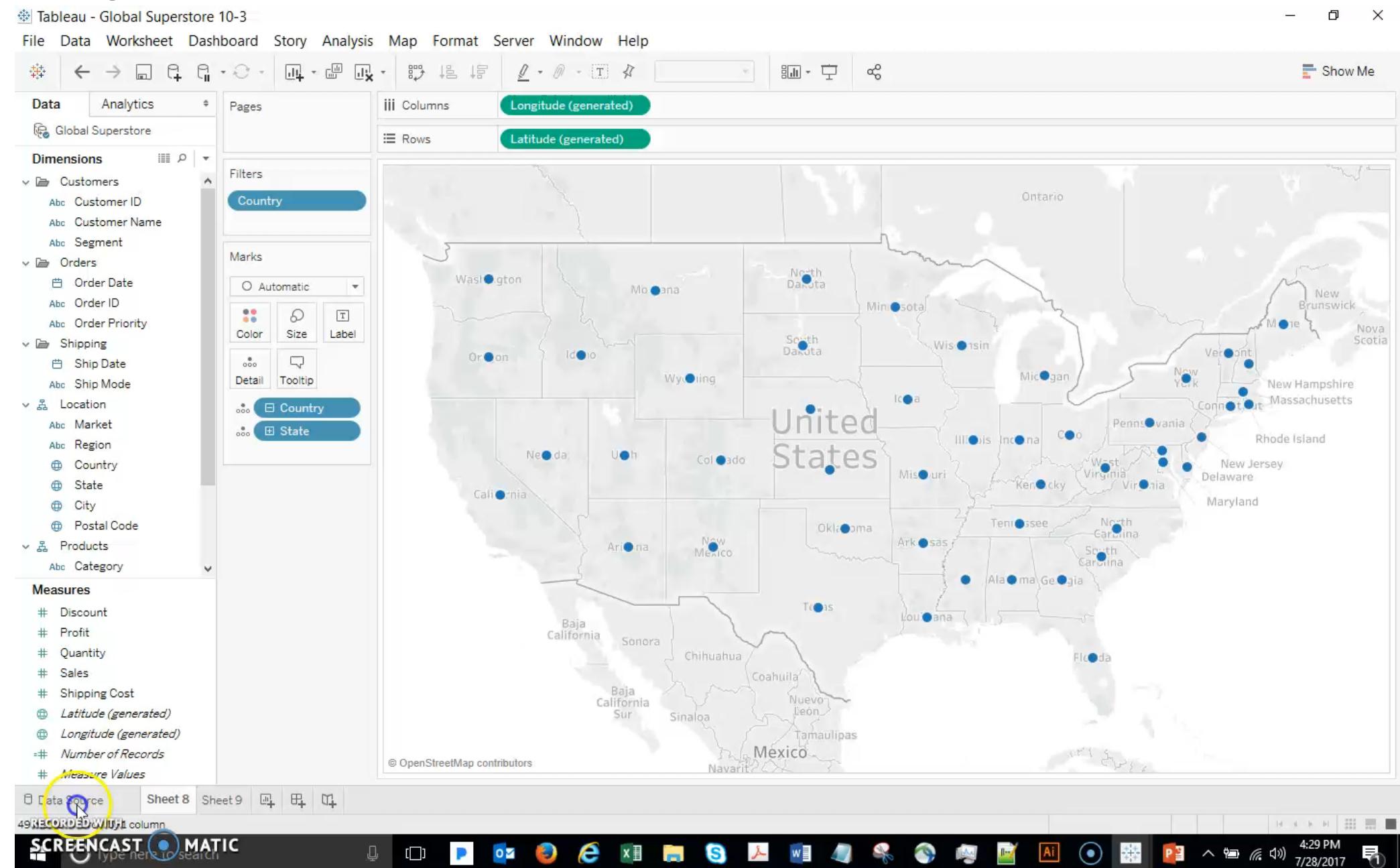
Using Map Services

- Choose a name for your style, and paste in your Mapbox account token.
- Pick a preset map – let's try Streets.
- Now **zoom** to Atlanta. Our background map now includes landmarks such as Universities, parks, etc. Students should experiment with other preset maps to understand what is available.
- Although outside our scope today, MapBox also allows you to create your own custom background maps.

Click Undo until you get back to your full map of the U.S. The other map service that is available is WMS maps. Many of these can be found online, and Tableau works with them also.



Using Map Services





Using Map Services

- From the Maps menu, choose Background Maps > Map Services.
- On the Map Services dialog, click Add > WMS Servers.
- Let's try this one:
- <http://ows.terrestris.de/osm/service?>

This WMS map came with its own set of options. Click Map...Map Layers. Check the 4 checkboxes. This is probably not ideal for our context, but I wanted to show you the WMS option for your maps. You can find many other free WMS maps just by googling.



Using Map Services

