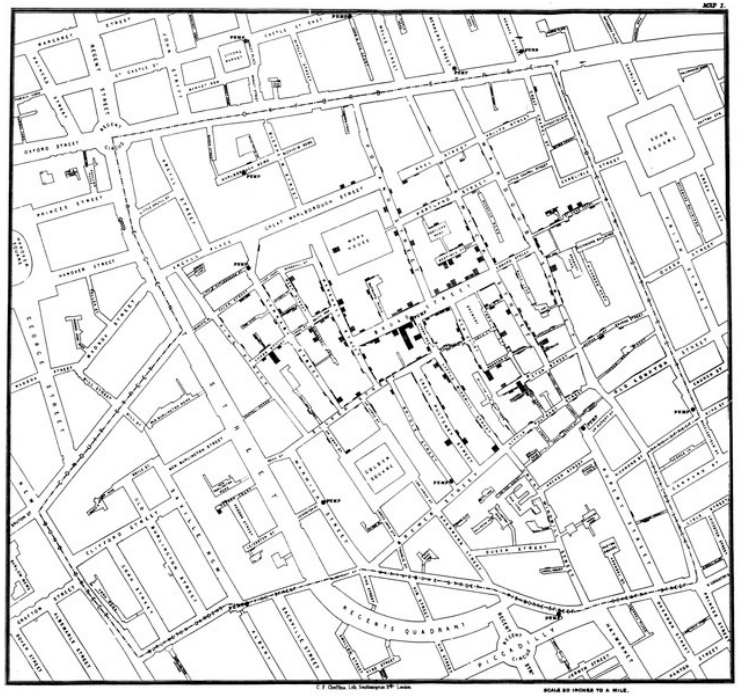
SECTION 1
WHAT IS DATA VISUALIZATION

PUBLIC HEALTH TOOL

John Snow's Map of the 1854 Broad Street Cholera Outbreak

Showed Cholera was being spread by water coming from a Broad Street Water Pump.

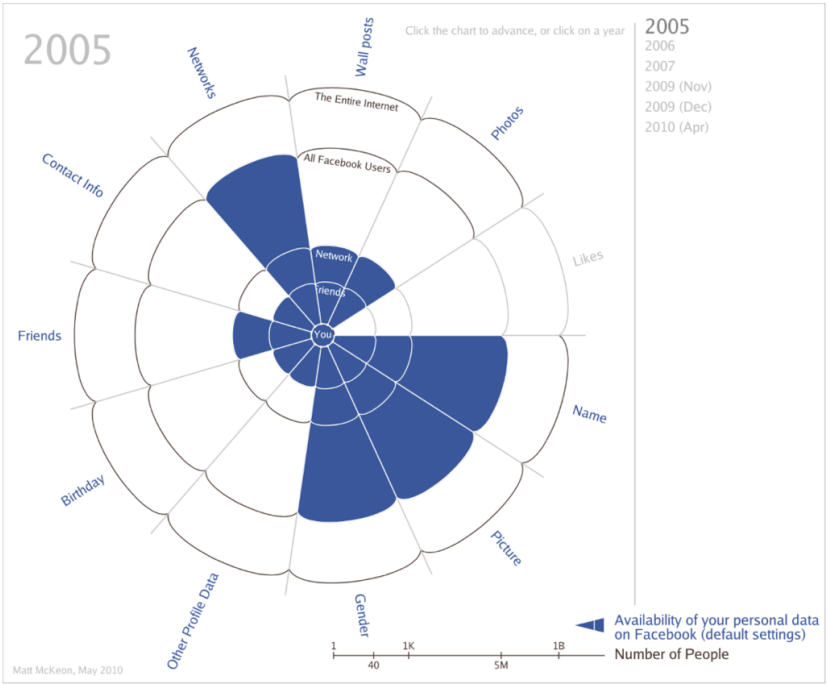


Source:
http://en.wikipedia.org/wiki/1854_Broad_Street_cholera_outbreak

PUBLIC / PRIVACY TOOL

How Public Your Personal Data is On Facebook By Default

Evolution of Facebook Default Privacy Settings



Source:
http://mattmckeon.com/facebook-privacy/

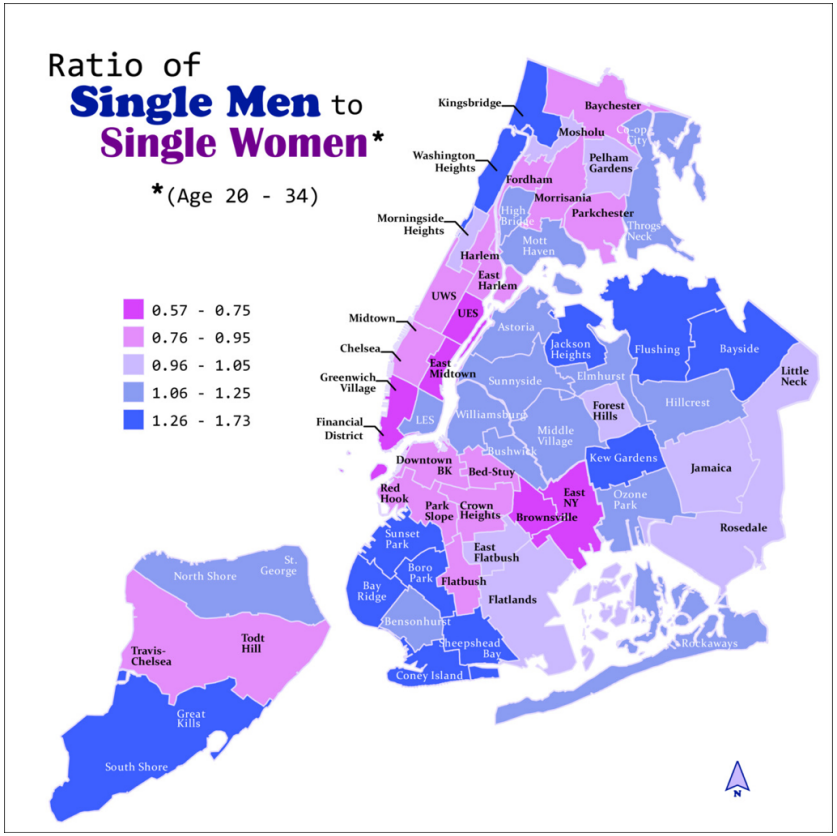
DATING TOOL

NYCEDC

Ratio of Single Men to Single Women (Age 20 – 34)

New York City's population is 53% female and 47% male. Using Census data, NYCEDC analyzed only the population who are never married singles between the ages of 20 and 34.

Source:
http://nycedc.tumblr.com/post/16175652438/ratio-of-single-men-to-single-women-in-nyc-new

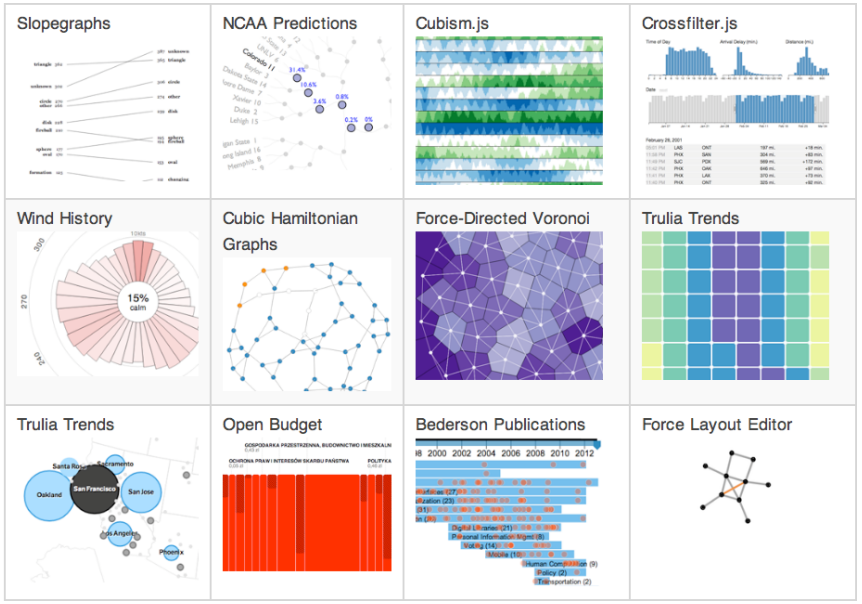


DATA VISUALIZATION IS A TOOL

Everything Generates Data

Visualizing this data leads to understanding.

- Sports
- Commerce
- Weather
- Real Estate
- Publications
- Social Media
- Etc.....



Source:
https://github.com/mbostock/d3/wiki/Gallery

DATA VIZ GUIDES THINKING

Data in Columns & Rows

Audience has to think hard

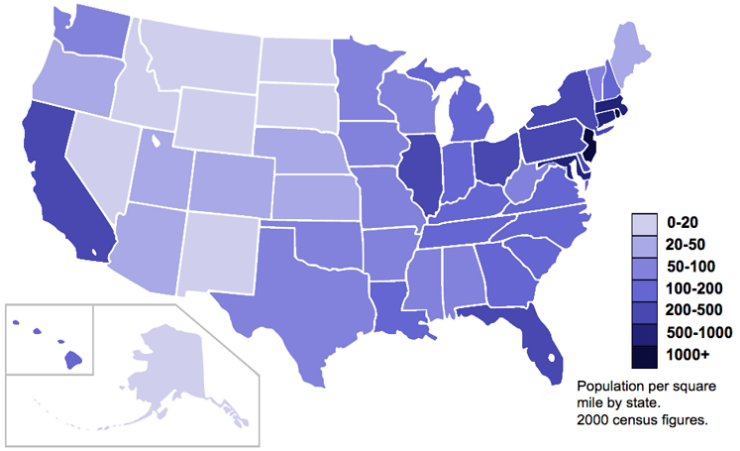
A39 | Maryland

	A	B	C	D	E	F	G	H
20	Alaska	D-3	R-3	R-3	R-3	R-3	R-3	R-3
21	Arizona	R-5	R-5	R-6	R-6	R-6	R-7	R-7
22	Arkansas	D-6	(\1)	R-6	D-6	R-6	R-6	R-6
23	California	D-40	R-40	R-45	R-45	R-45	R-47	R-47
24	Colorado	D-6	R-6	R-7	R-7	R-7	R-8	R-8
25	Connecticut	D-8	D-8	R-8	R-8	R-8	R-8	R-8
26	Delaware	D-3	R-3	R-3	D-3	R-3	R-3	R-3
27	District of Columbia	D-3	D-3	D-3	D-3	D-3	D-3	D-3
28	Florida	D-14	R-14	R-17	D-17	R-17	R-21	R-21
29	Georgia	R-12	(\1)	R-12	D-12	D-12	R-12	R-12
30	Hawaii	D-4	D-4	R-4	D-4	D-4	R-4	D-4
31	Idaho	D-4	R-4	R-4	R-4	R-4	R-4	R-4
32	Illinois	D-26	R-26	R-26	R-26	R-26	R-24	R-24
33	Indiana	D-13	R-13	R-13	R-13	R-13	R-12	R-12
34	Iowa	D-9	R-9	R-8	R-8	R-8	R-8	D-8
35	Kansas	D-7	R-7	R-7	R-7	R-7	R-7	R-7
36	Kentucky	D-9	R-9	R-9	D-9	R-9	R-9	R-9
37	Louisiana	R-10	(\1)	R-10	D-10	R-10	R-10	R-10
38	Maine	D-4	D-4	R-4	R-4	R-4	R-4	R-4
39	Maryland	D-10	D-10	R-10	D-10	D-10	R-10	R-10
40	Massachusetts	D-14	D-14	D-14	D-14	R-14	R-13	D-13
41	Michigan	D-21	D-21	R-21	R-21	R-21	R-20	R-20
42	Minnesota	D-10	D-10	R-10	D-10	D-10	D-10	D-10
43	Mississippi	R-7	(\1)	R-7	D-7	R-7	R-7	R-7
44	Missouri	D-12	R-12	R-12	D-12	R-12	R-11	R-11
45	Montana	D-4	R-4	R-4	R-4	R-4	R-4	R-4
46	Nebraska	D-5	R-5	R-5	R-5	R-5	R-5	R-5
47	Nevada	D-3	R-3	R-3	R-3	R-3	R-4	R-4
48	New Hampshire	D-4	R-4	R-4	R-4	R-4	R-4	R-4
49	New Jersey	D-17	R-17	R-17	R-17	R-17	R-16	R-16
50	New Mexico	D-4	R-4	R-4	R-4	R-4	R-5	R-5

Source:
http://www.census.gov

Data Visualization

Audience has to think less hard



Source:
http://commons.wikimedia.org/wiki/File:US_2000_census_population_density_map_by_state.svg

"RAW DATA" IS HARD TO READ

Anscombe's quartet

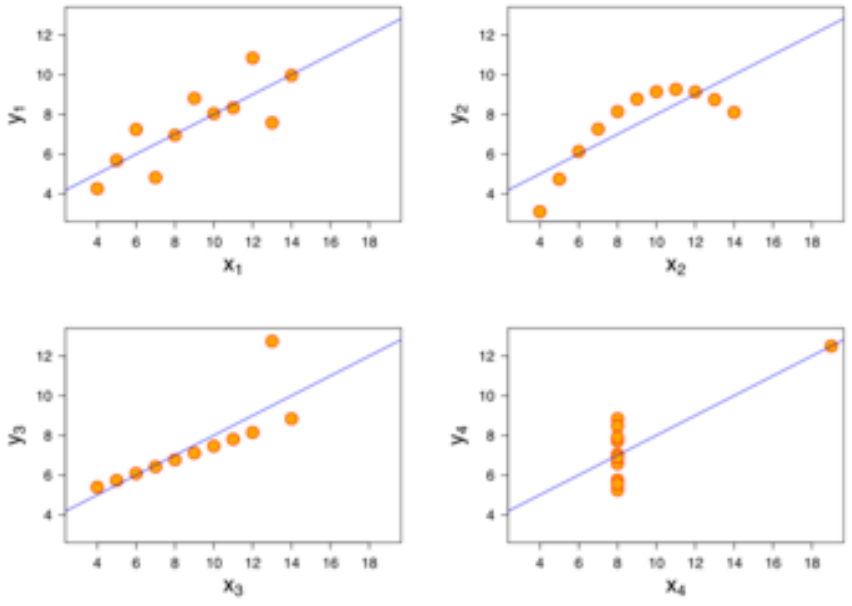
I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

- Mean of x: 9
- Variance of x: 11
- Mean of y: 7.50
- Variance of y: 4.12
- Correlation between x and y: 0.816
- Linear regression line for each case:
  $y = 3.00 + 0.500 * x$

DATA VIZ DRIVES INTUITION OF DATA

Anscombe's quartet

I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

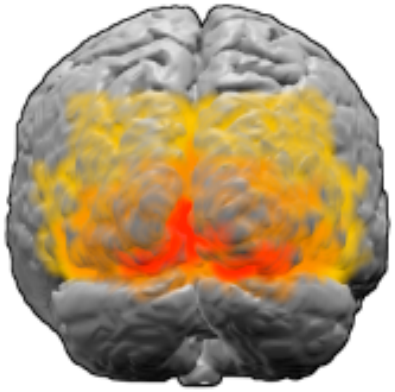


DATA VIZ USES VISUAL CORTEX

66% of stimuli reaching the brain are visual (Zaltman 1996)

50% of brain devoted to processing visual images (Bates & Cleese 2001)

80% of learning is visually based (American Optometric Assoc. 1991)



Eye Image Source:
http://www.flickr.com/photos/orangeacid/234358923/

Brain Image Source:
http://en.wikipedia.org/wiki/File:Brodmann_areas_17_18_19.png

DATA VIZ USES BRAIN'S SHORTCUTS

720349656089226535931140790070322302
076958689027429003358787115045223998
424533087922668417382319480046553364
246202505406711172160430997890121737
608183566145635519888049583302306957
749597705315240714467203496560892265
359311407900703223020769586890274290
033587871150452239984245330879226684
173823194800465533642462025054067111
721604309978901217376081835661456355
5202642463355640084913283

720349656089226535931140790070322302
076958689027429003358787115045223998
424533087922668417382319480046553364
246202505406711172160430997890121737
608183566145635519888049583302306957
749597705315240714467203496560892265
359311407900703223020769586890274290
033587871150452239984245330879226684
173823194800465533642462025054067111
721604309978901217376081835661456355
5202642463355640084913283

BRAIN'S PRE-ATTENTIVE PROPERTIES

COLOR HUE

ORIENTATION

TEXTURE

POSITION & ALIGNMENT

COLOR BRIGHTNESS

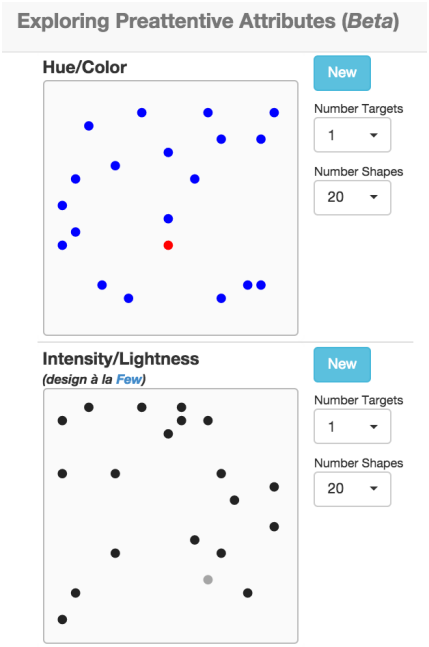
COLOR SATURATION

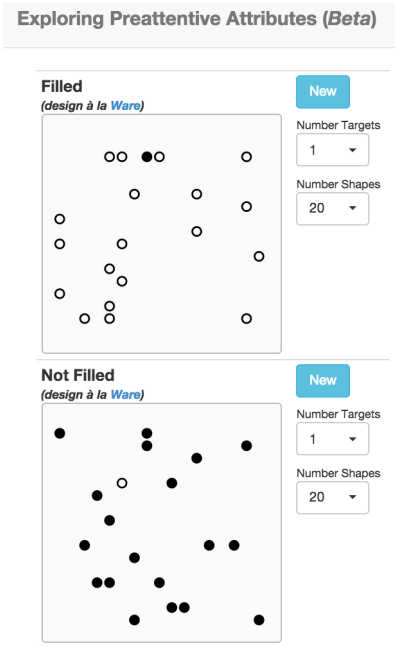
SIZE

SHAPE

D3 EXAMPLES OF PRE-ATTENTIVE PROPERTIES

http://learnforeverlearn.com/preattentive/





DATA VISUALIZATION IS A...

- Tool
- That helps guide thinking
- When trying to “understand” / “develop intuition about” data
- That uses the Visual Cortex
- To take advantage of the brain’s shortcuts
- And pre-attentive properties
- To achieve a goal.

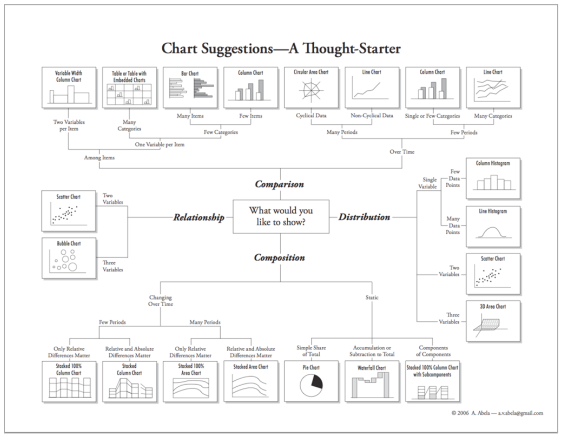
SEVEN MAIN TYPES OF DATA VIZ GOALS

Visual Encoding Depends on your goal -

- Time / Evolution
- Drill down
- Zoom out
- Contrast
- Intersections
- Factors
- Outliers

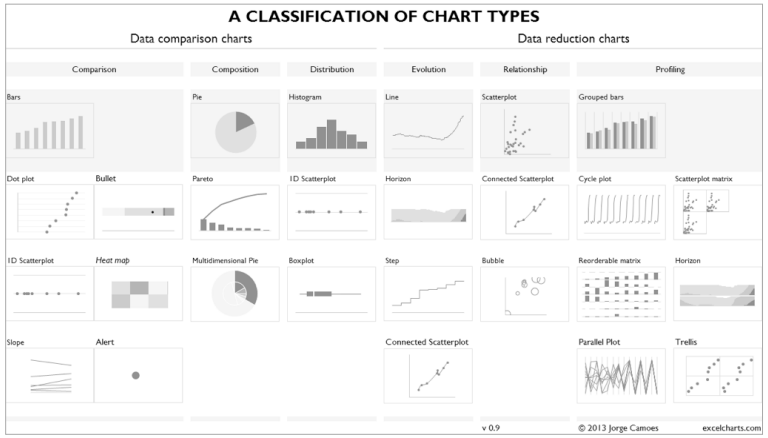
Source:
https://twitter.com/dataremixed

VISUALLY ENCODING YOUR GOAL



- Comparison
- Distribution
- Composition
- Relationship

Source:
© A. Abela - a.v.abela@gmail.com



Comparison
- Comparison
- Composition
- Distribution

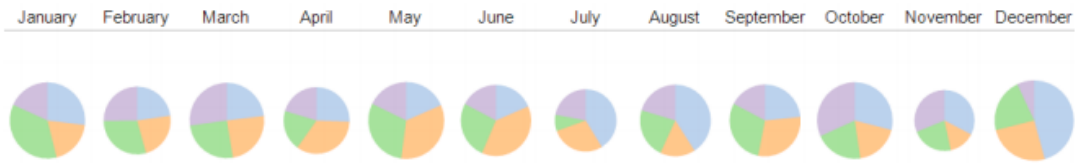
Reduction
- Evolution
- Relationship
- Profiling

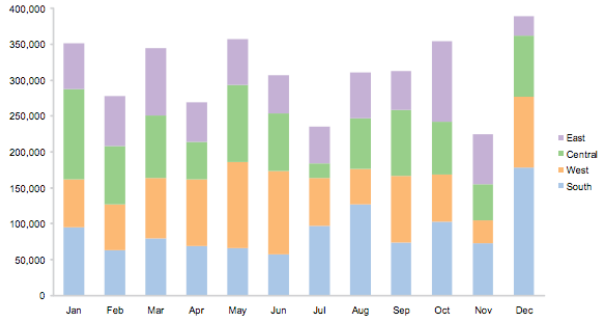
Source:
excelcharts.com

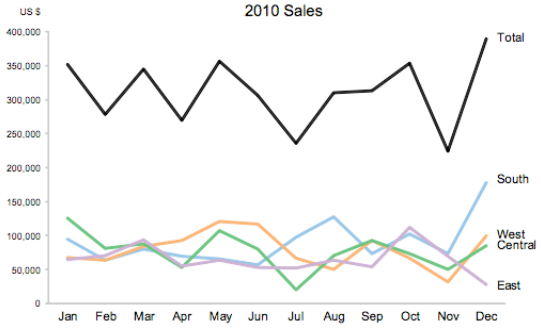
VISUAL ENCODING EFFECTIVENESS

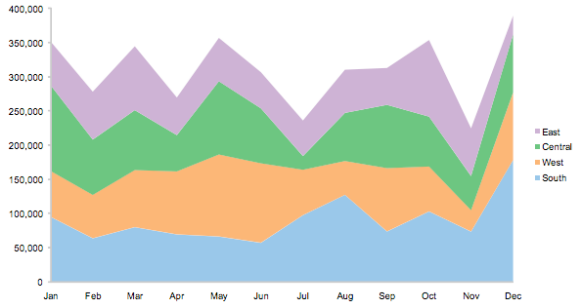
Depends on you goal

12 Months of Sales
- By Region
- By Month









Source:
Enrico Bertini, Assistant Professor at NYU-Poly (@filwd)

PURPOSE, DATA, AND AUDIENCE

Seth Godin

Three questions to ask your marketing team

Q1 - Who are you trying to reach?

Q2 - Why do they decide to support us?

Q3 - What do you need in order to make this happen more often?

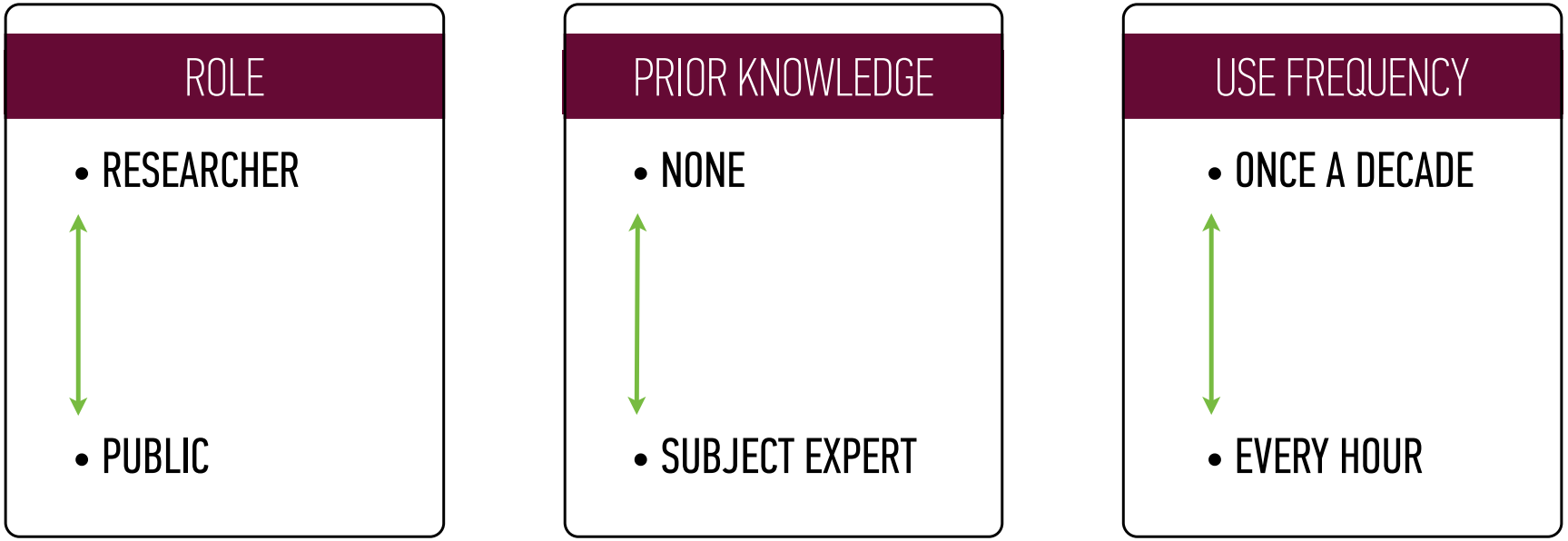
Three questions to ask of your data visualization

Q1 - Who are you trying to reach? (Audience)

Q2 - Why do they decide to support us? (Purpose)

Q3 - What do you need in order to make this happen? (Data)

NEED TO KNOW – AUDIENCE



NEED TO KNOW – HOW IT IS VIEWED

PRINT	WEB	VIDEO	PRESENTATION
• BLACK AND WHITE? • SOME COLOR? • ALL COLOR?	• INTERACTIVE? • NON-INTERACTIVE?	• NEWS SEGMENT? • COMMERCIAL? • SHOW?	• GUIDED? • UNGUIDED?

NEED TO KNOW – AVAILABLE DATA

PRIMARY	SECONDARY	GENERATED
• YOU COLLECT IT • YOU OWN IT • NOBODY ELSE HAS IT	• OTHERS COLLECT IT • OTHERS OWN IT • EVERYONE HAS IT	• FROM PRIMARY • FROM SECONDARY • FROM COMBINATION

NEED TO KNOW – GOAL / PURPOSE / WIN

HYPOTHESIS	GOAL	PARAMETERS
WHAT ARE WE TRYING TO SHOW?	HOW DO WE KNOW IF WE ACHIEVED IT?	WHAT ARE THE BOUNDARIES?

AUDIENCE VIEWS DATA VIZ, THEN WHAT

Seth Godin

Source:
http://www.sethgodin.com/

- What are you trying to tell me?
- What do you want me to do now?

DATA VISUALIZATION IS A...

- Tool
- That helps guide thinking
- When trying to “understand” / “develop intuition about” data
- That uses the Visual Cortex
- To take advantage of the brain’s shortcuts (pre-attentive properties)
- So that you / your audience can achieve a goal / purpose
- Without thinking too much