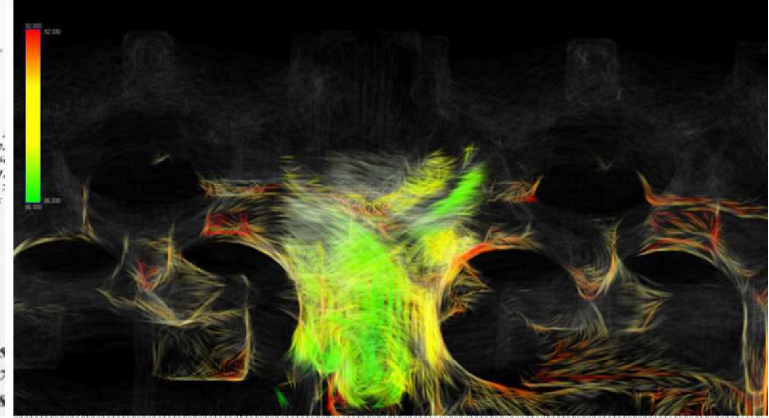
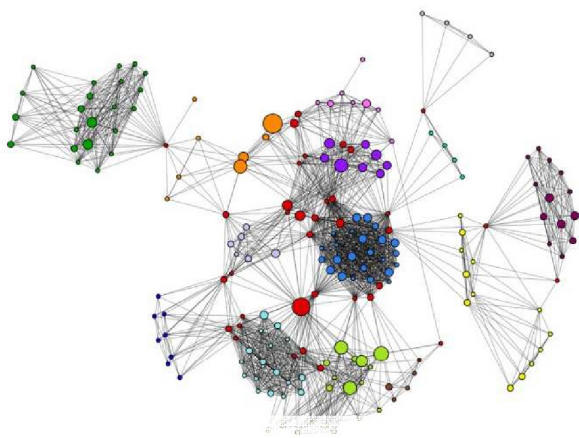


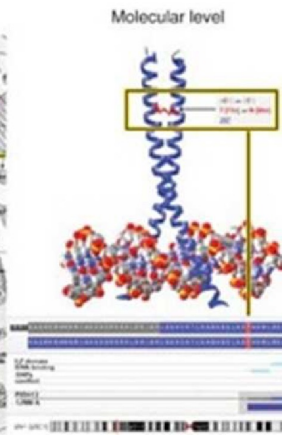
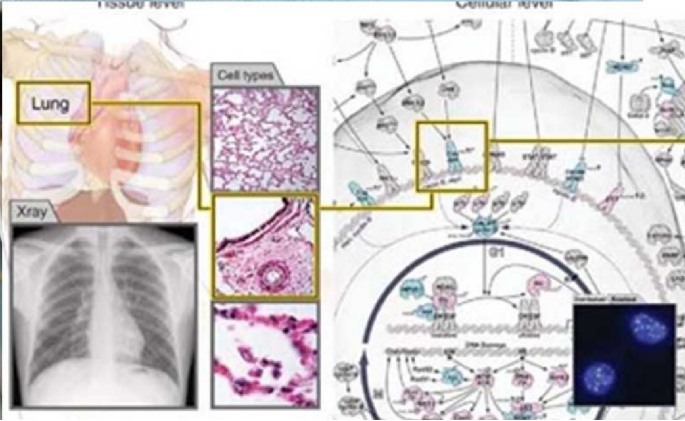
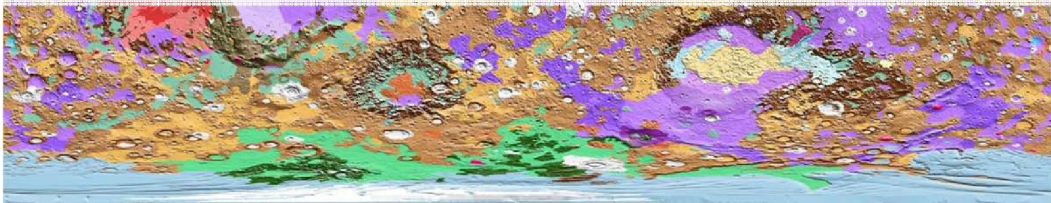
The background features a vertical green bar on the left side. To its right, there are dynamic, swirling ink-like splashes in shades of blue, purple, and pink against a white background. The text is overlaid on the green bar.

DATA VISUALIZATIONS TECHNIQUES AND TOOLS

**ASHISH
SENIOR RESEARCH DATA SCIENTIST
ROSEN CENTER OF ADVANCED
COMPUTING**



Data is generated everywhere and everyday



What is data visualization?



The graphical representation of information and data.



It is part art and part science. The challenge is to get the art right without getting the science wrong, and vice versa.



The goal of data visualization is to communicate data or information clearly and effectively to readers

+

- The human mind
- is very receptive to visual information.

That's why data visualization is a powerful tool for communication.

How many 3s can you count?

2487218401238740921659014760985609324
7209125629065098526590482758298568096
0986309584390564095878950374509284750
989475092984

- +
 - The human mind is very receptive to visual information.

That's why data visualization is a powerful tool for communication.

2487218401238740921659014760985609324
7209125629065098526590482758298568096
0986309584390564095878950374509284750
989475092984

KEY PRINCIPLES OF EFFECTIVE DATA VISUALIZATION



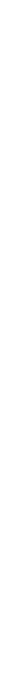
+

o



1

Determine the
best visual



2

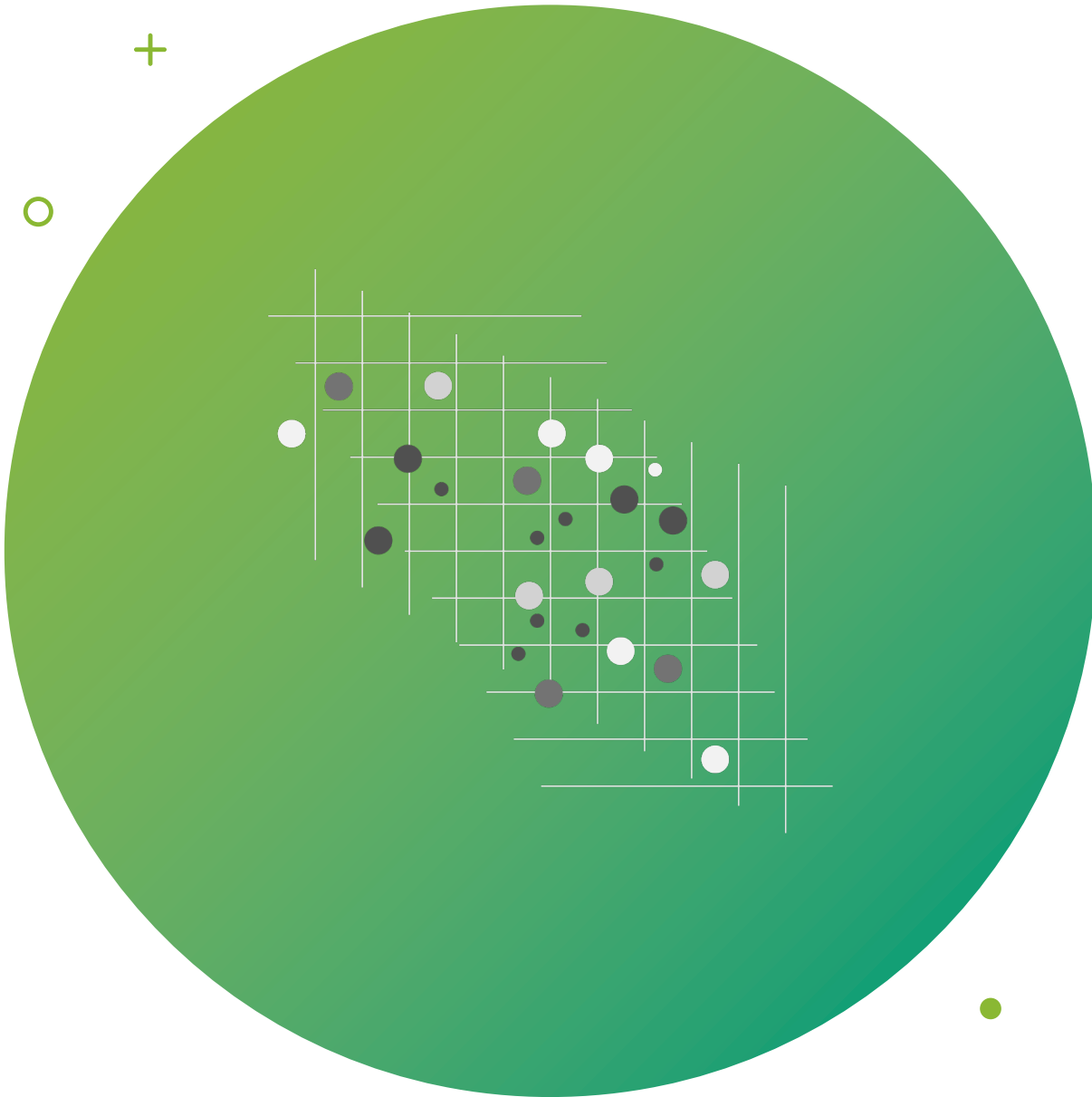
Balance the design





3

Focus on the key
areas



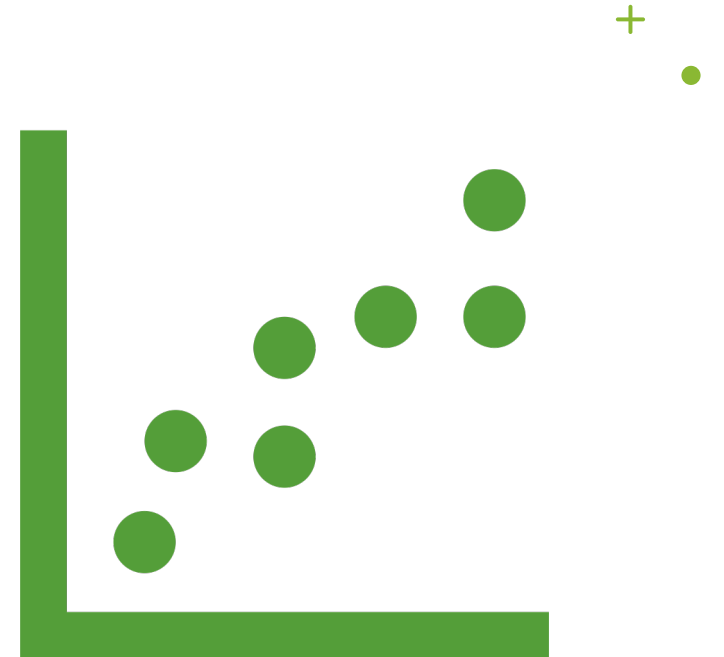
4

Keep it simple



5

Incorporate interactivity



o

+



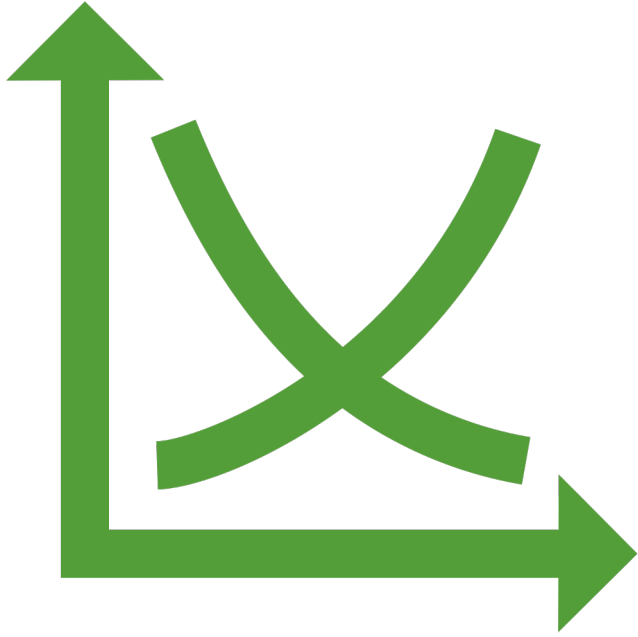
6

Use patterns



.





• + 7
○

Compare aspects



DATA VISUALIZATION TOOLS

- Tableau
- Power BI
- Matplotlib
- Plotly
- Bokeh
- Seaborn



DATA PREPROCESSING



TYPES OF DATA

QUANTITATIVE

DISCRETE

CONTINUOUS

CATEGORICAL

GUIDE TO CHART TYPES

BAR CHART



PIE CHART



LINE CHART



AREA CHART



SCATTER PLOT



BUBBLE CHART



HEAT MAP

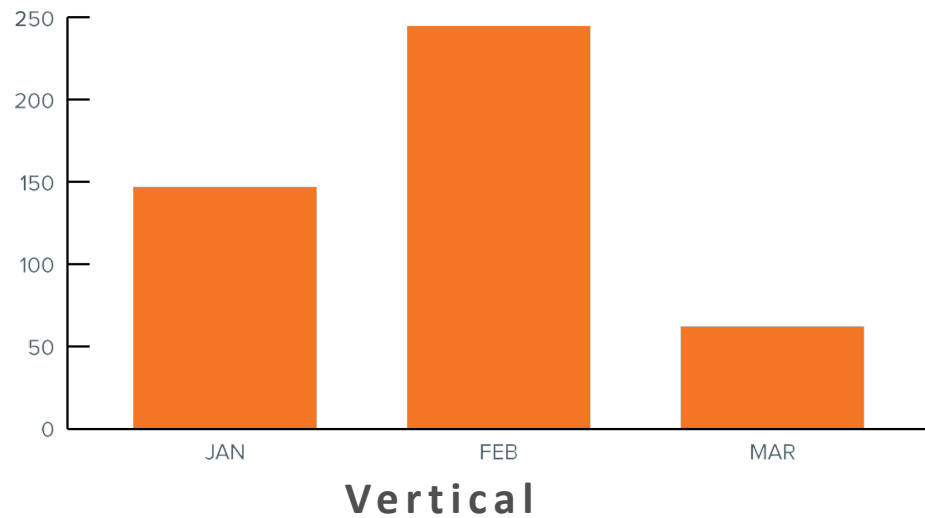


BAR CHART

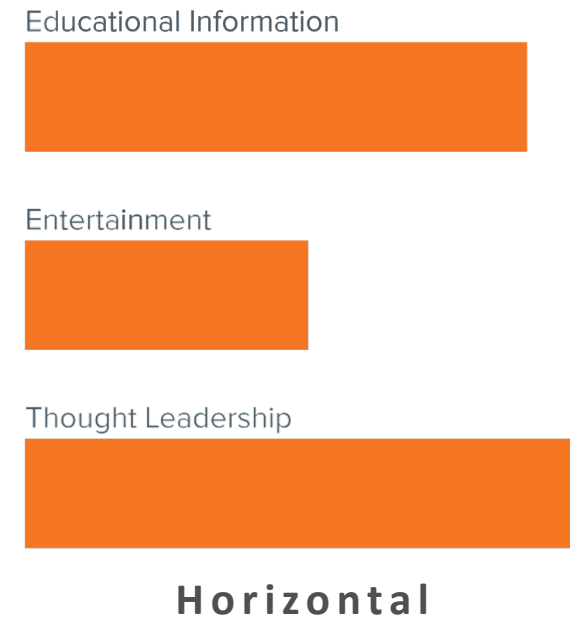


Variations of Bar Charts

PAGE VIEWS, BY MONTH



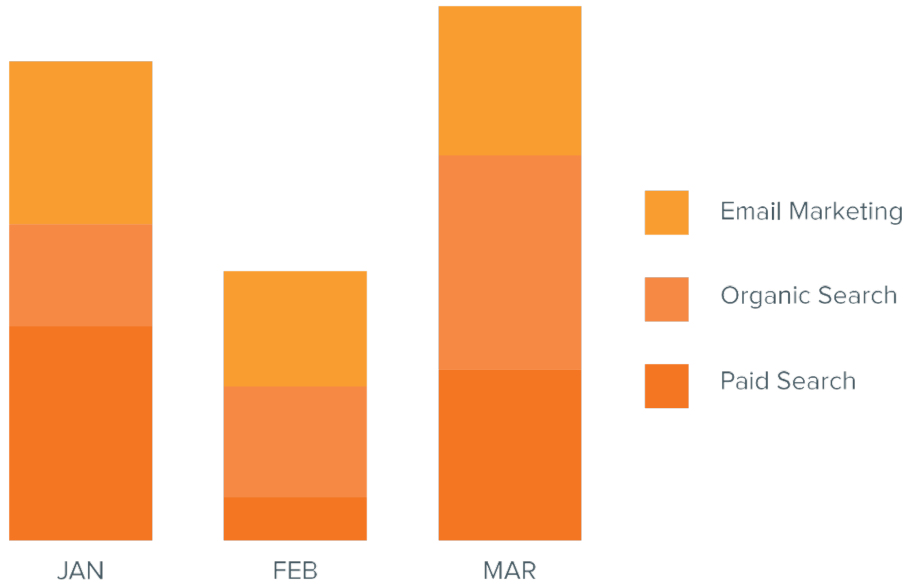
CONTENT PUBLISHED, BY CATEGORY



VARIATIONS OF BAR CHARTS (CONT.)

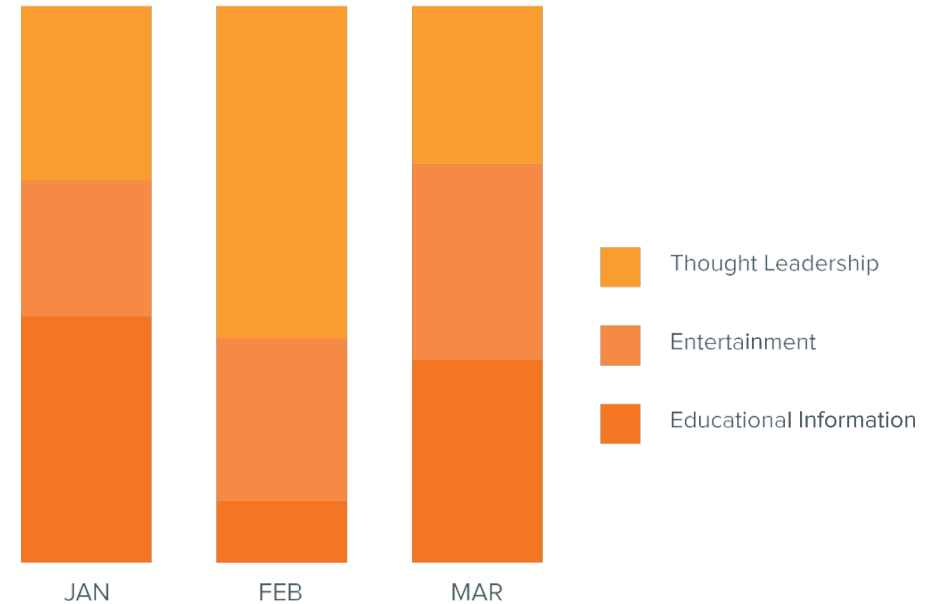


MONTHLY TRAFFIC, BY SOURCE



STACKED

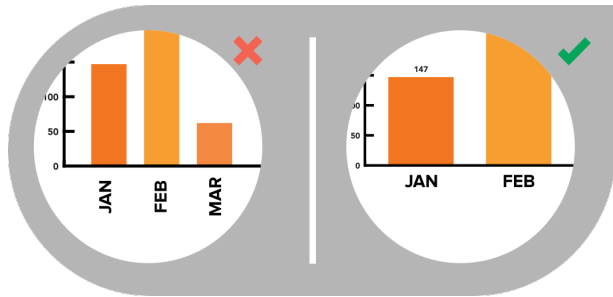
PERCENTAGE OF CONTENT PUBLISHED, BY MONTH



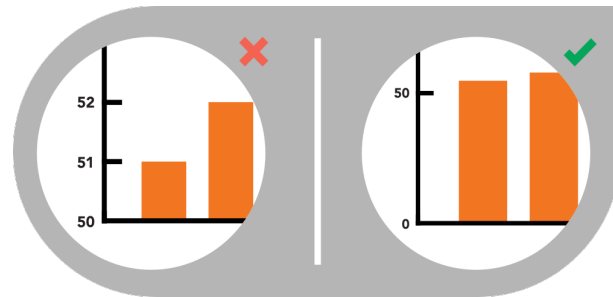
100% STACKED

BAR CHART

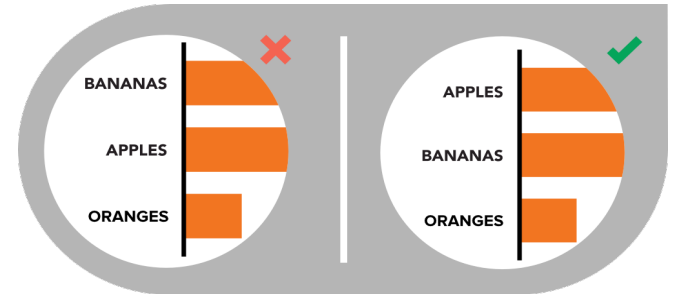
DESIGN BEST PRACTICES



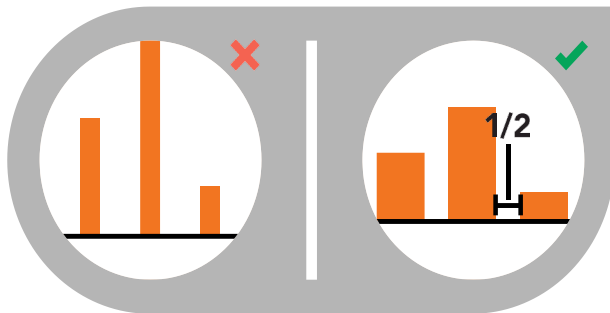
USE HORIZONTAL LABELS



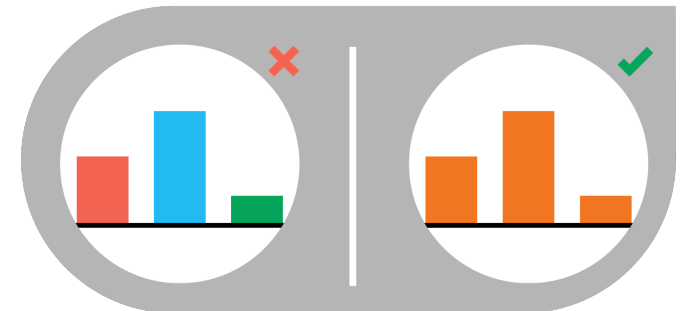
START THE Y-AXIS VALUE AT 0



ORDER DATA APPROPRIATELY



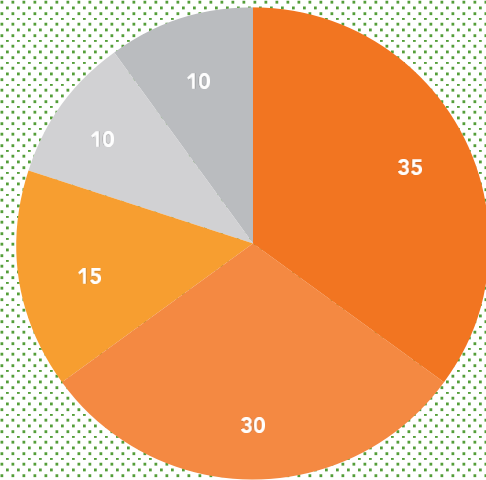
SPACE BARS APPROPRIATELY



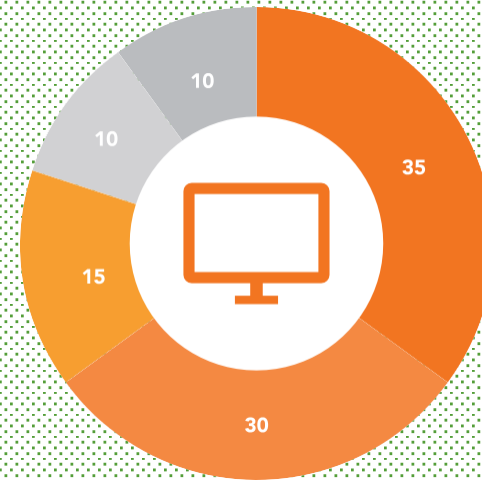
USE CONSISTENT COLORS

PIE CHART

VARIATIONS OF PIE CHARTS



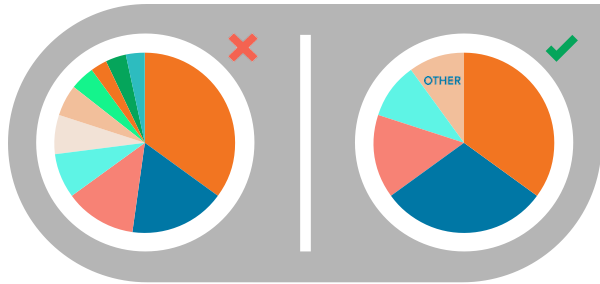
STANDARD



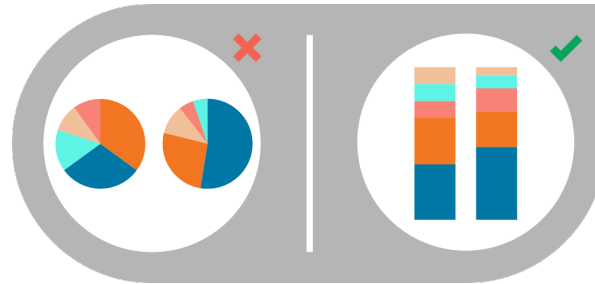
DONUT

PIE CHART

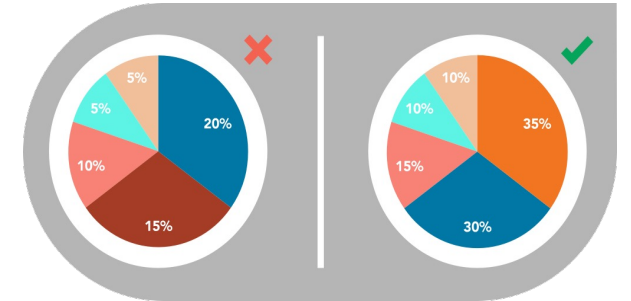
DESIGN BEST PRACTICES



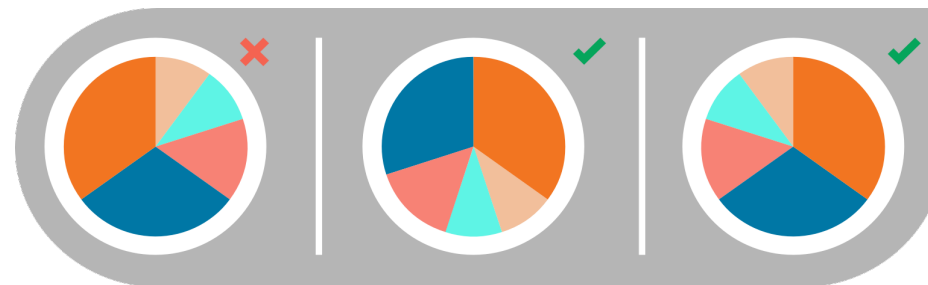
VISUALIZE NO MORE THAN
5 CATEGORIES PER CHART



DON'T USE MULTIPLE PIE
CHARTS FOR COMPARISON



MAKE SURE ALL DATA
ADDS UP TO 100%



ORDER
SLICES
CORRECTLY

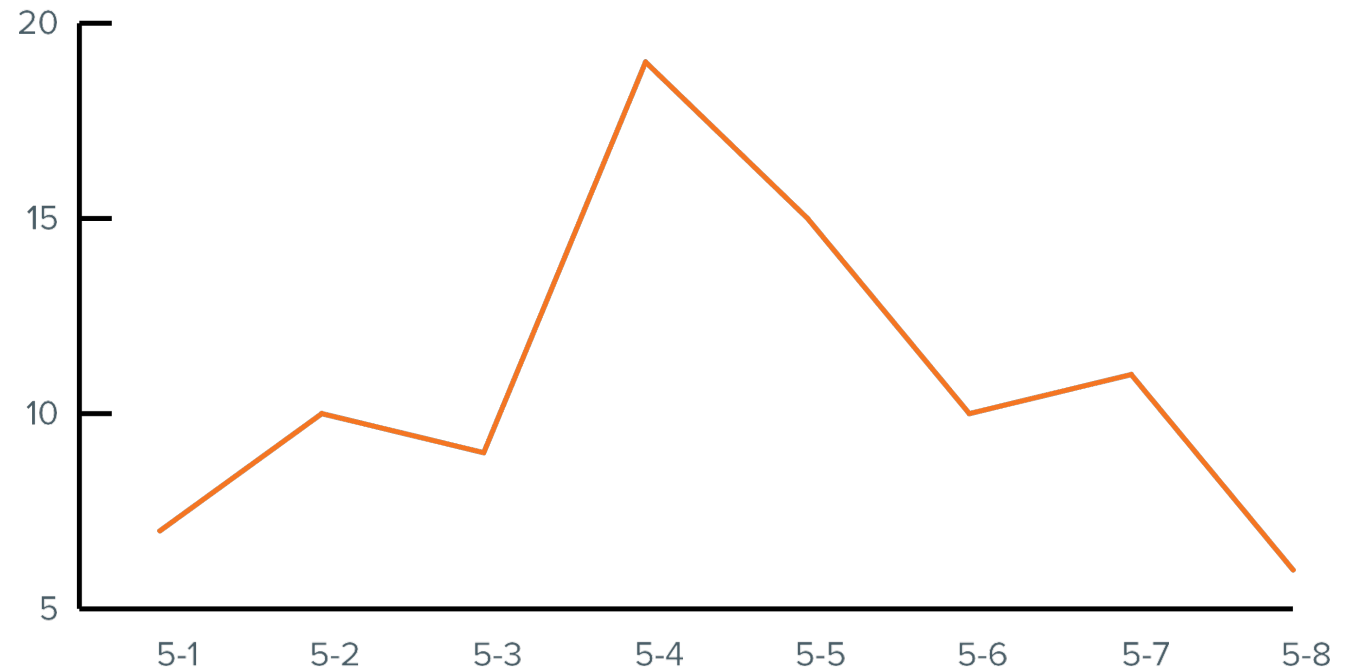
OPTION 1

OPTION 2

LINE CHART

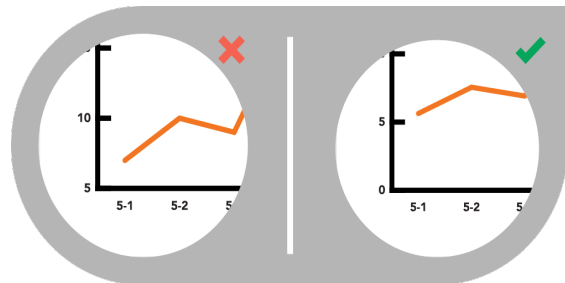


DIRECT MARKETING VIEWS, BY DATE

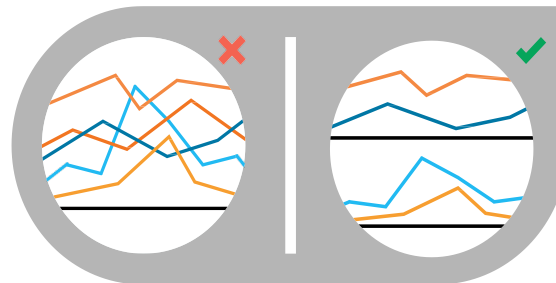


LINE CHART

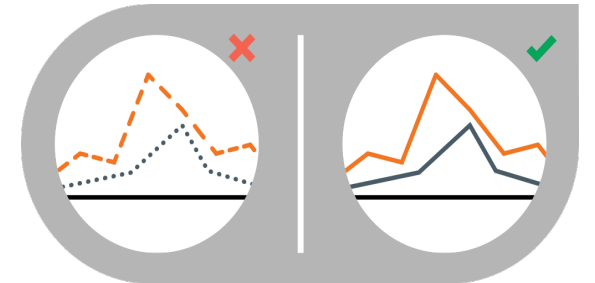
DESIGN BEST PRACTICES



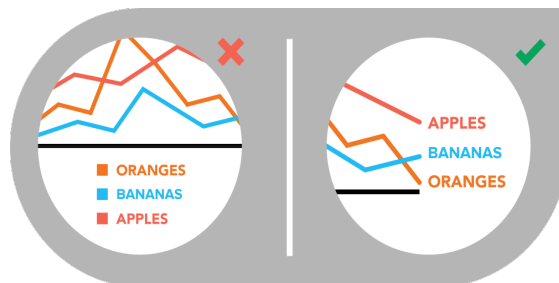
INCLUDE A ZERO BASELINE IF POSSIBLE



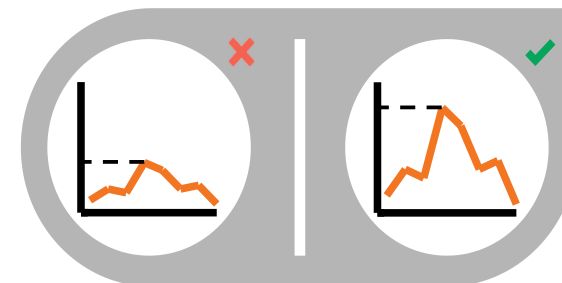
DON'T PLOT MORE THAN 4 LINES



USE SOLID LINES ONLY



LABEL THE LINES DIRECTLY

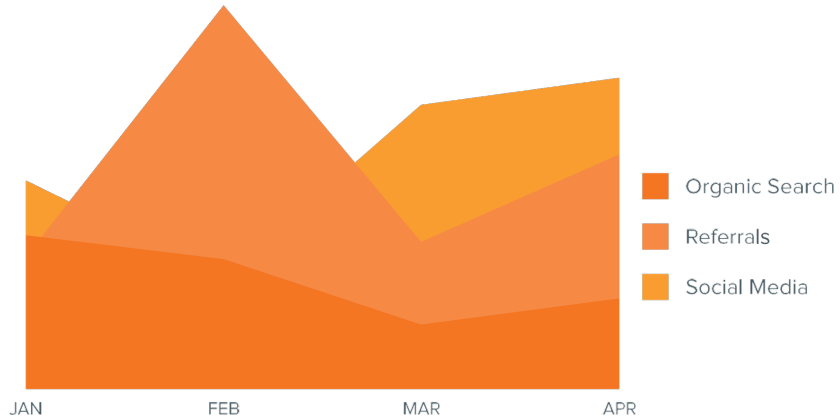


USE THE RIGHT HEIGHT

AREA CHART

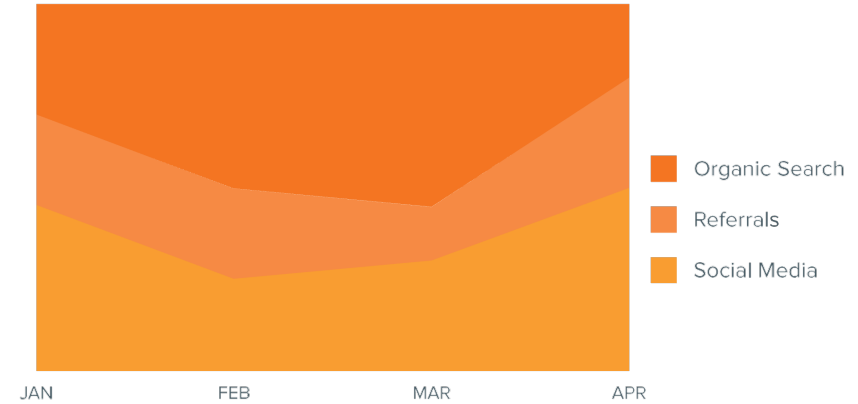
VARIATIONS OF AREA CHARTS

NEW CONTACTS, BY SOURCE



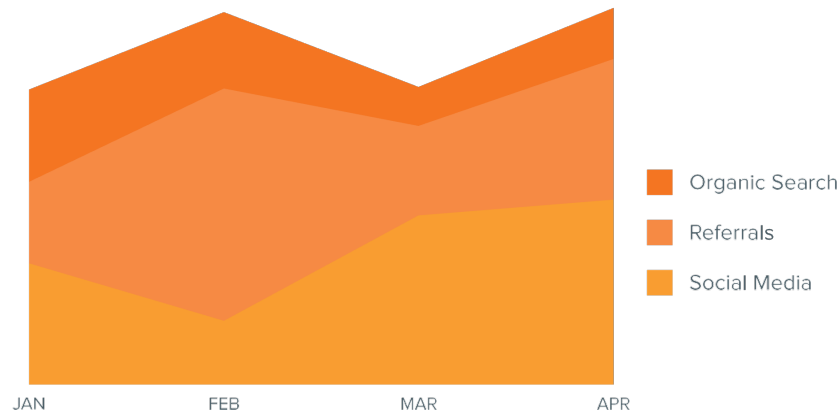
AREA CHART

NEW CONTACTS, BY SOURCE



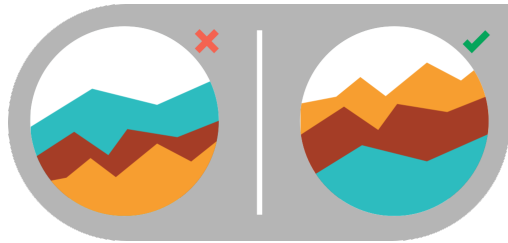
100% STACKED AREA

NEW CONTACTS, BY SOURCE

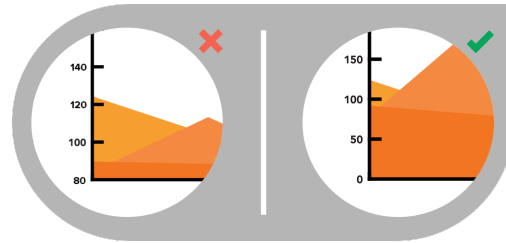


STACKED AREA

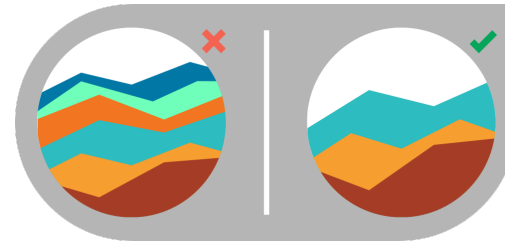
AREA CHART DESIGN BEST PRACTICES



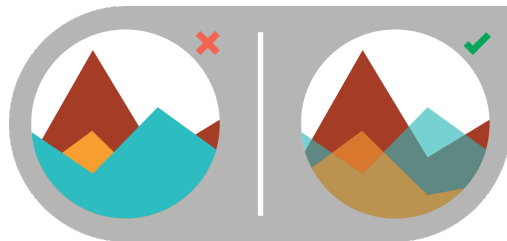
MAKE IT EASY TO READ



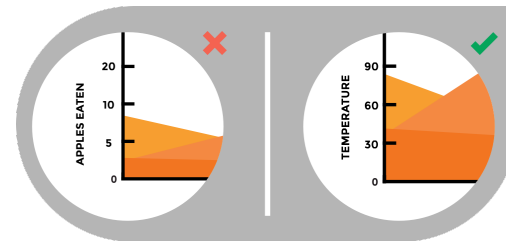
START Y-AXIS VALUE AT 0



DON'T DISPLAY MORE THAN
4 DATA CATEGORIES



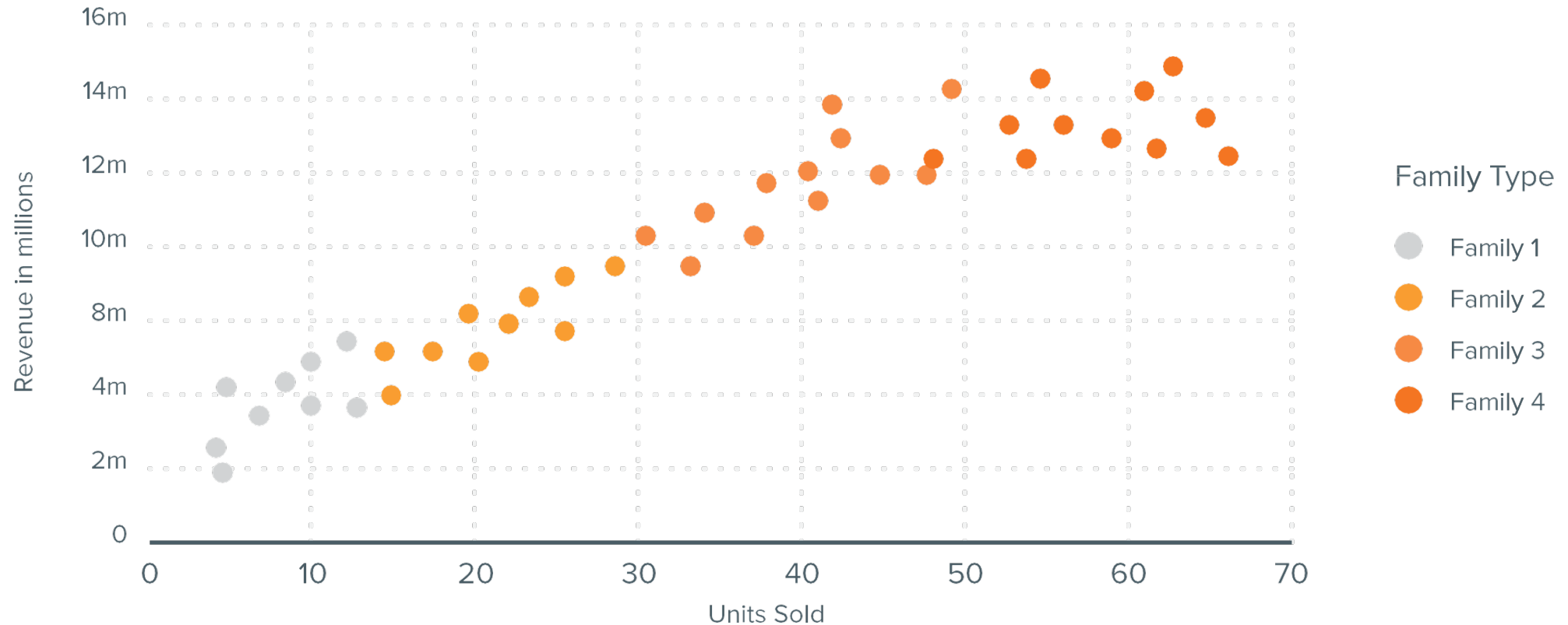
USE TRANSPARENT COLORS



DON'T USE AREA CHARTS TO
DISPLAY DISCRETE DATA

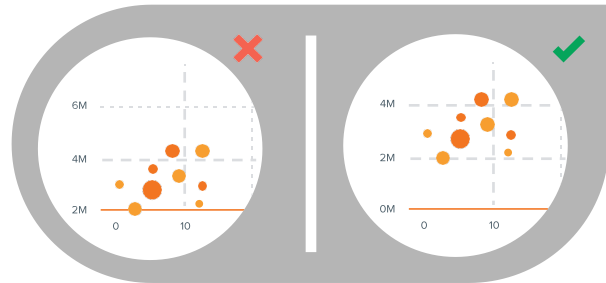
SCATTER PLOT

REVENUE, BY PRODUCT FAMILY

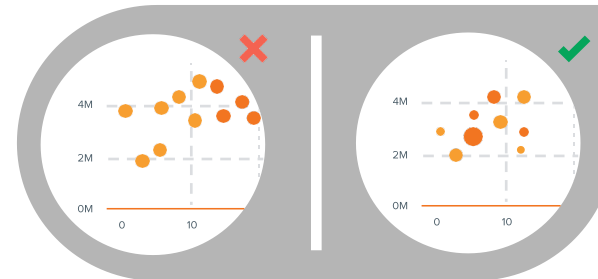


SCATTER PLOT

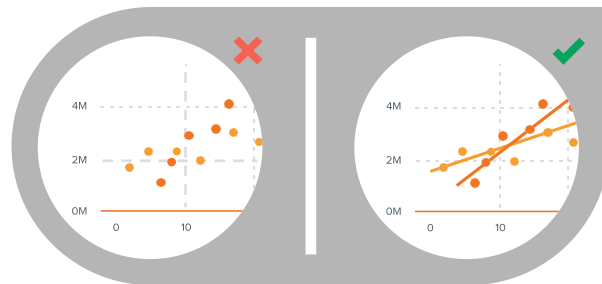
DESIGN BEST PRACTICES



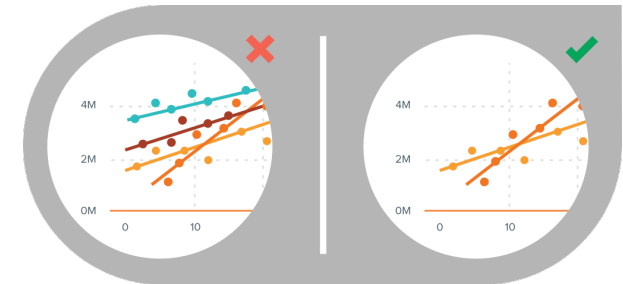
START Y-AXIS VALUE AT 0



INCLUDE MORE VARIABLES



USE TREND LINES

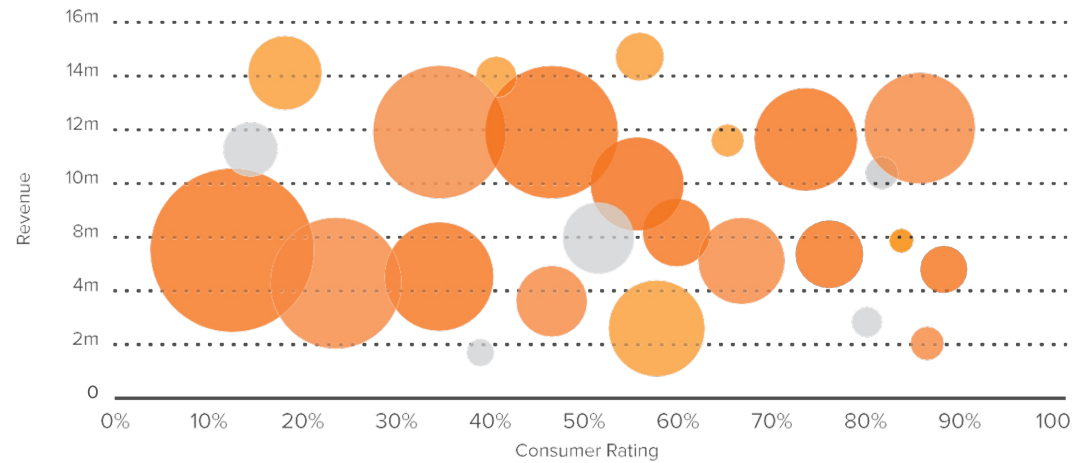


DON'T COMPARE MORE THAN 2 TREND LINES

BUBBLE CHART

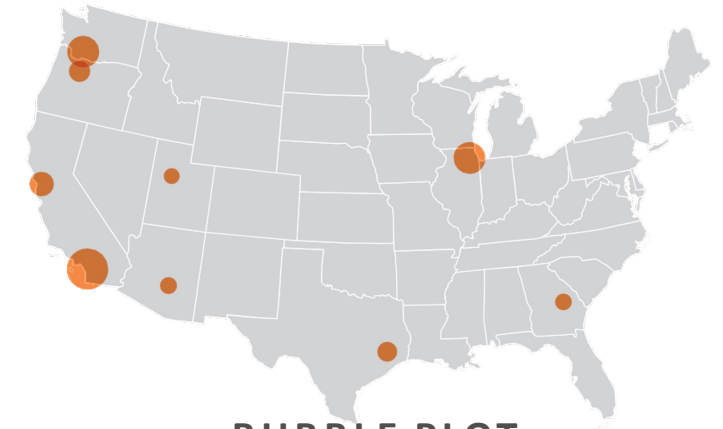
VARIATIONS OF BUBBLE CHARTS

REVENUE VS. RATING



BUBBLE PLOT

BIGGEST SALES INCREASE

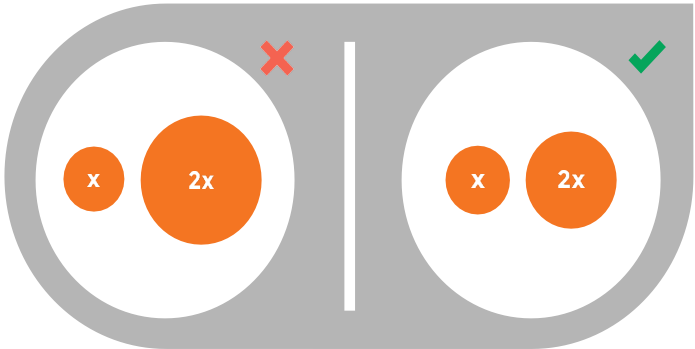


BUBBLE PLOT

BUBBLE CHART DESIGN BEST PRACTICES



MAKE SURE LABELS ARE VISIBLE



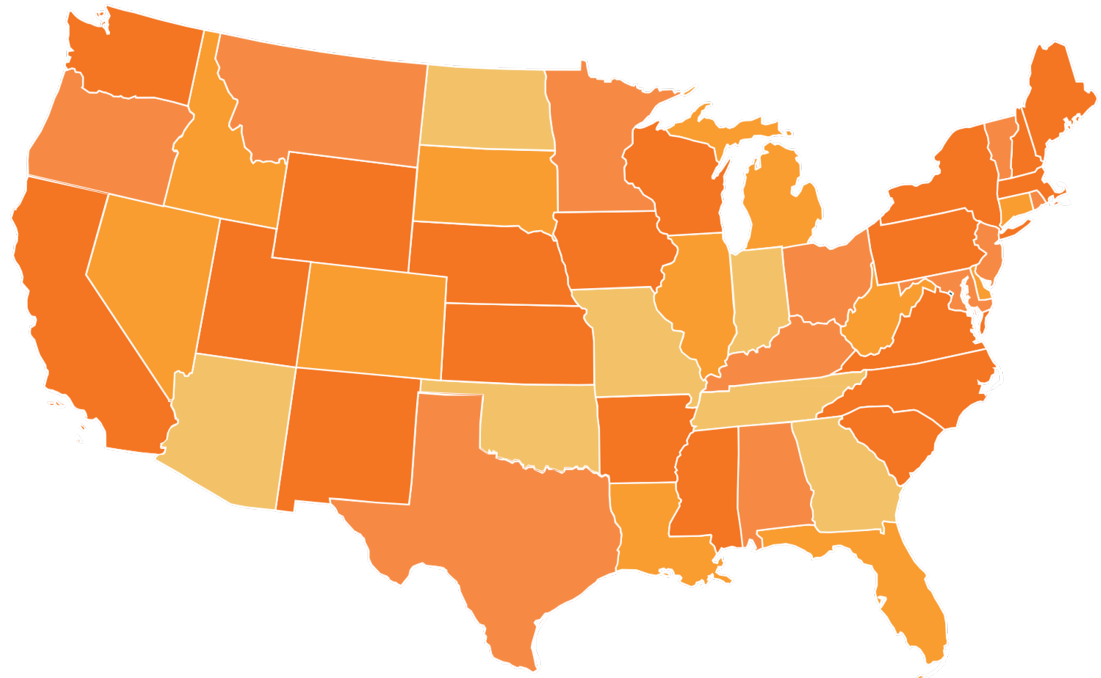
SIZE BUBBLES APPROPRIATELY



DON'T USE ODD SHAPES

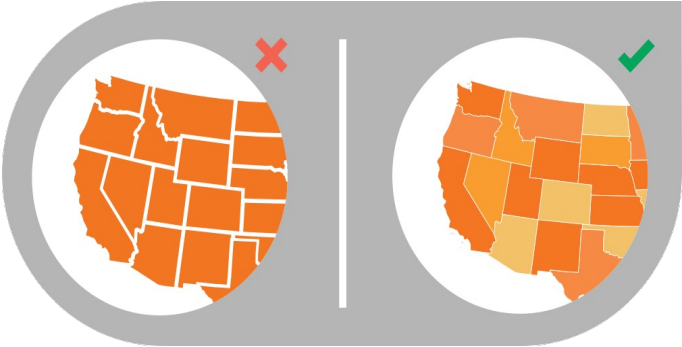
HEAT MAP

+
●
○
STATES WITH NEW SERVICE CONTRACTS

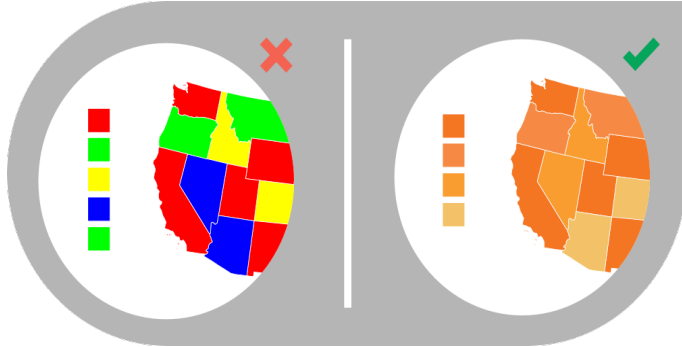


● 75-76 ● 77-78 ● 79-80 ● 81+

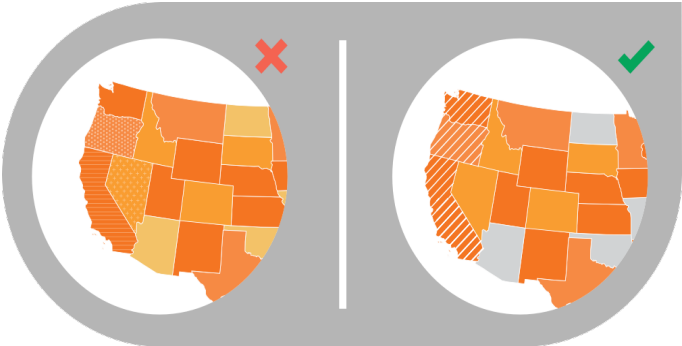
HEAT MAP DESIGN BEST PRACTICES



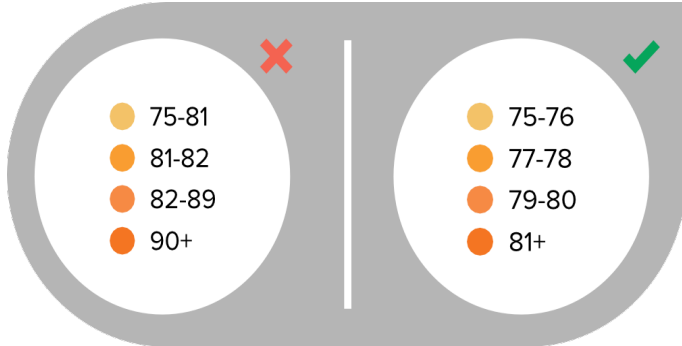
USE A SIMPLE MAP OUTLINE



SELECT COLORS APPROPRIATELY



USE PATTERNS SPARINGLY



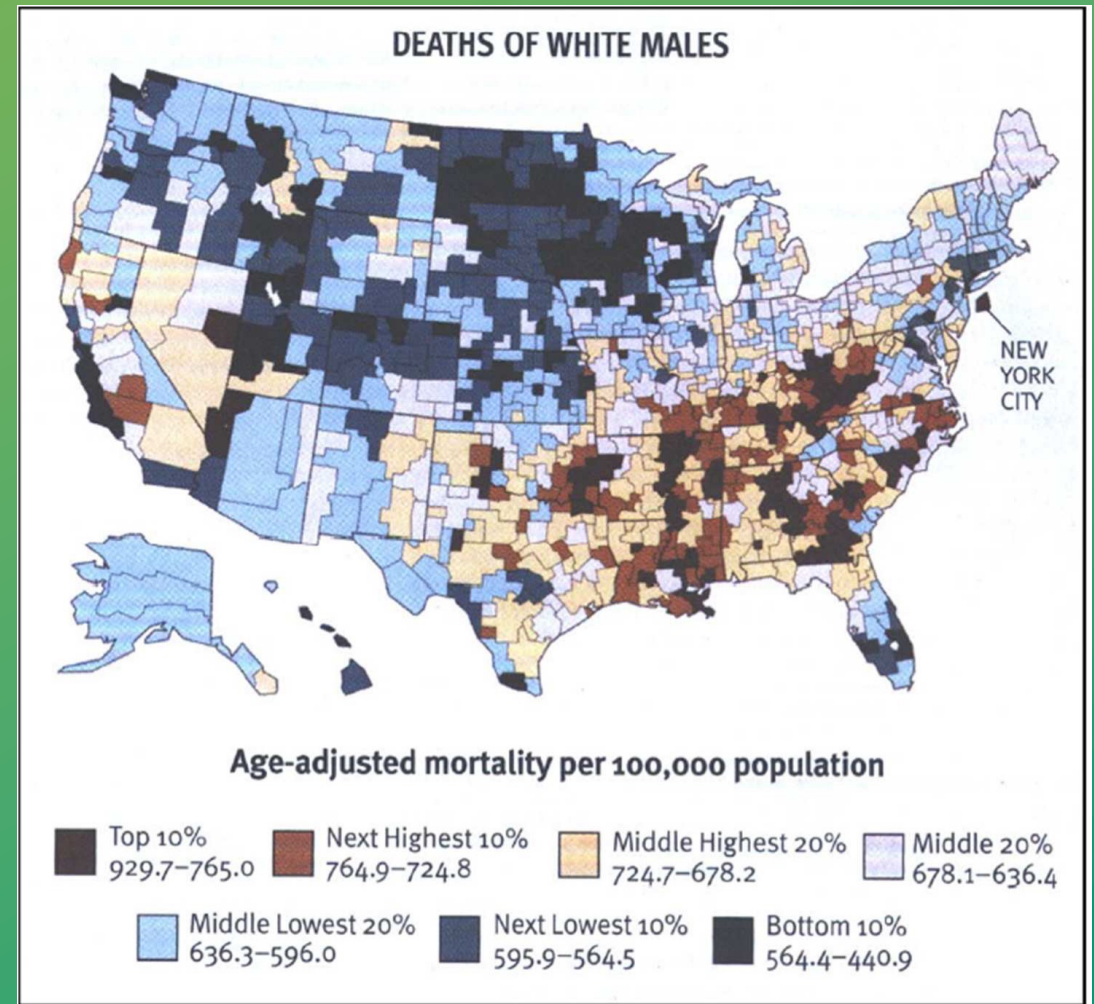
CHOOSE APPROPRIATE DATA RANGES

+
○ ●

USE COLOR WISELY



WHAT IS WRONG WITH THIS COLOR SCALE



**DO NOT ATTEMPT TO FIGHT
PRE-ESTABLISHED
COLOR MEANINGS**

+

•

○

+

○

•

Examples of Pre-Established Color Meanings

Red

Stop
Off
Dangerous
Hot
High stress
Oxygen
Shallow
Money loss

Green

On
Plants
Carbon
Moving
Money

Blue

Cool
Safe
Deep
Nitrogen

+

•

○

USE GOOD CONTRAST, AS THE
HUMAN EYE IS
GOOD AT
IDENTIFYING THE
DIFFERENCE



Color Alone Doesn't Cut It

**I sure hope that my
life does not depend
on being able to read
this quickly and
accurately!**

Luminance Contrast is Crucial

**I would prefer that
my life depend on
being able to read *this*
quickly and
accurately!**

Use good contrast

	Black	Black	Black	Black	Black	Black	Black	Black
White		White	White	White	White	White	White	White
Red	Red		Red	Red	Red	Red	Red	Red
Yellow	Yellow	Yellow		Yellow	Yellow	Yellow	Yellow	Yellow
Green	Green	Green	Green		Green	Green	Green	Green
Blue	Blue	Blue	Blue	Blue		Blue	Blue	Blue

OTHER RULES

It is not possible to list all the useful rules. They come with a lot of experience!

- Limit the total number of colors if viewers are to discern information quickly.
- Be aware that our perception of color changes with:
 - 1) surrounding color
 - 2) how close two objects are
 - 3) how long you have been staring at the color
 - 4) sudden changes in the color intensity.



Beware of Color Pollution



Just because you have millions of
colors to choose from

doesn't mean you must use them all ...

+
•
○

INTERACTIVE VISUALIZATION WITH BOKEH



**BOKEH: DATA
VISUALIZATION LIBRARY
IN PYTHON THAT
PROVIDES INTERACTIVE
CHARTS AND PLOTS.**



+
o

•

DATA FEATURES OF BOKEH

FLEXIBILITY

PRODUCTIVITY

INTERACTIVITY

POWERFUL

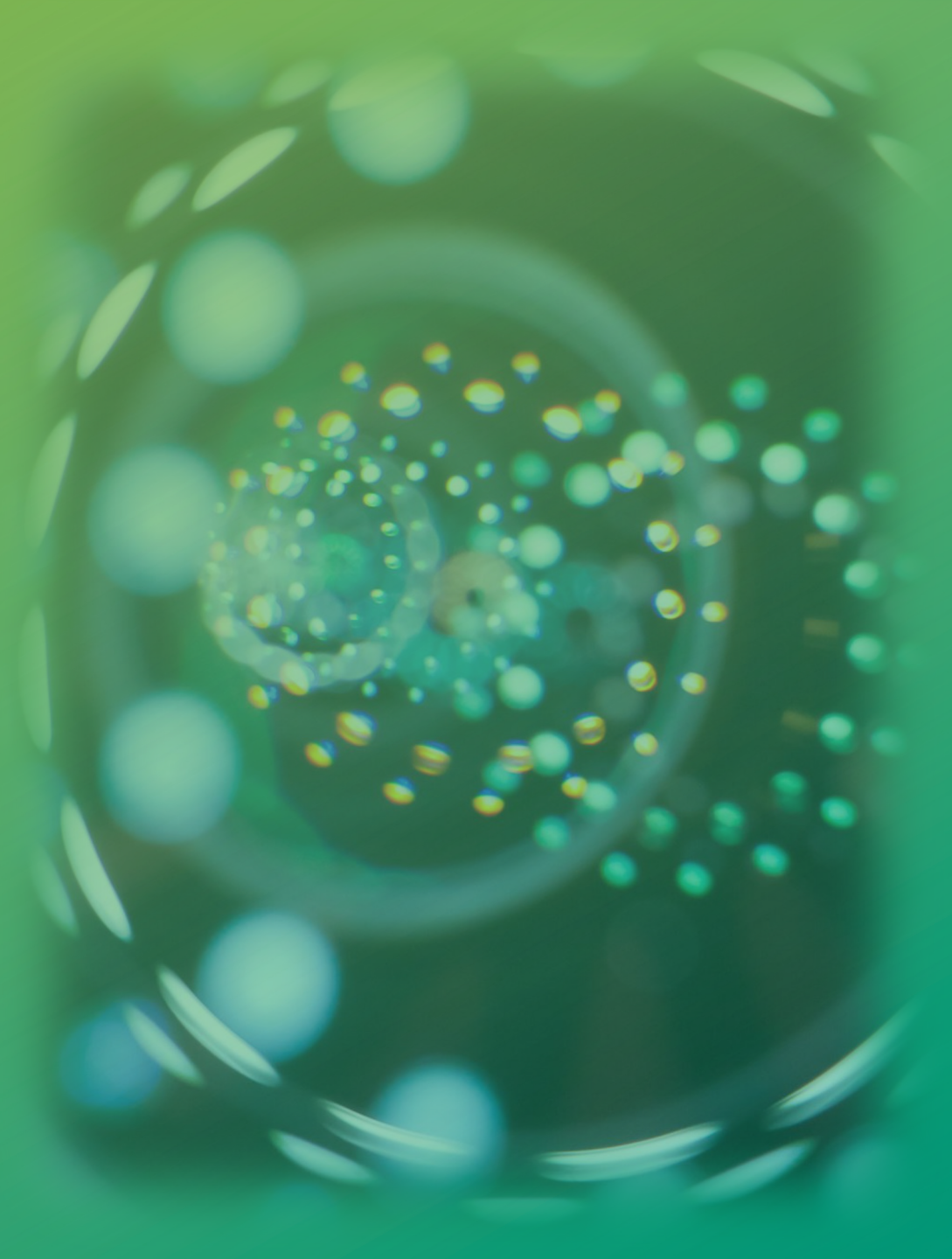
SHAREABLE

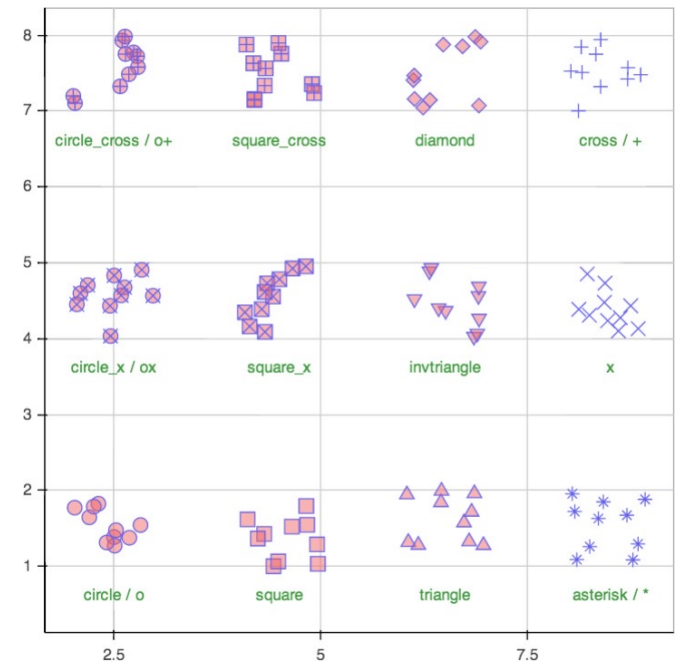
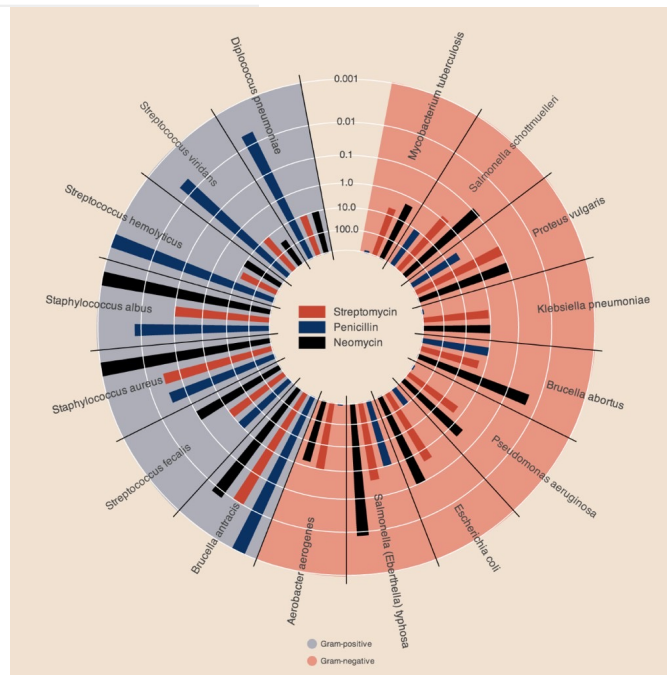
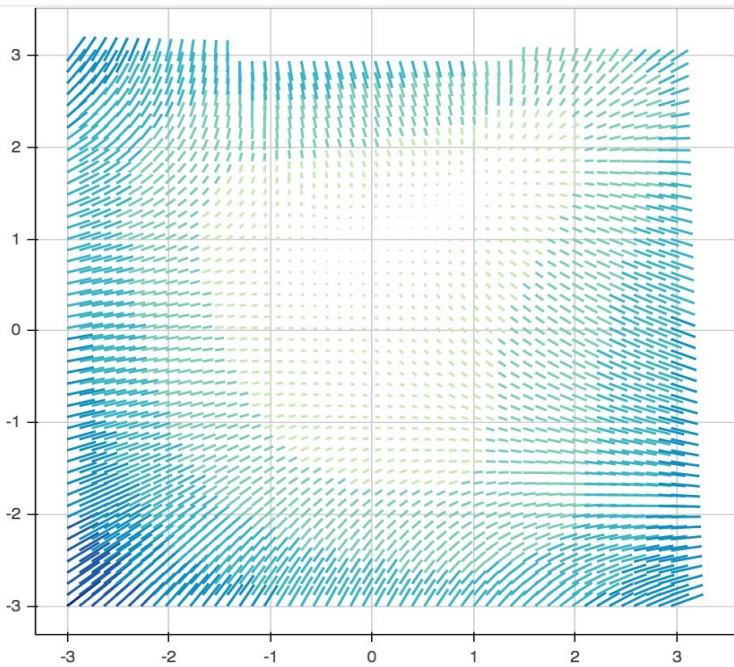
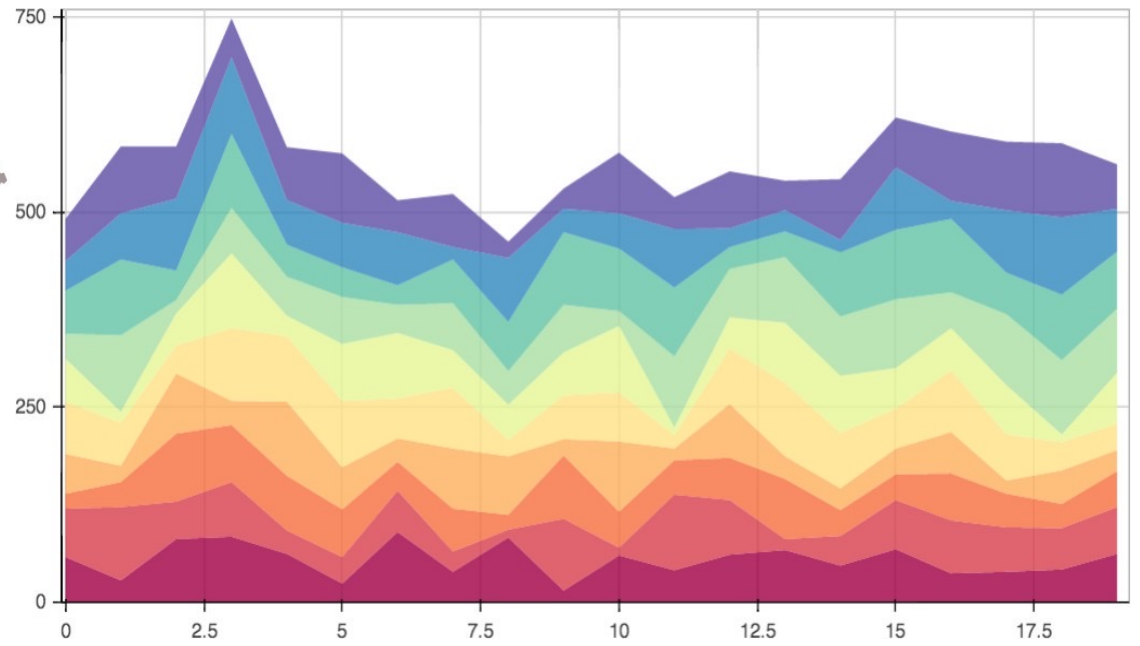
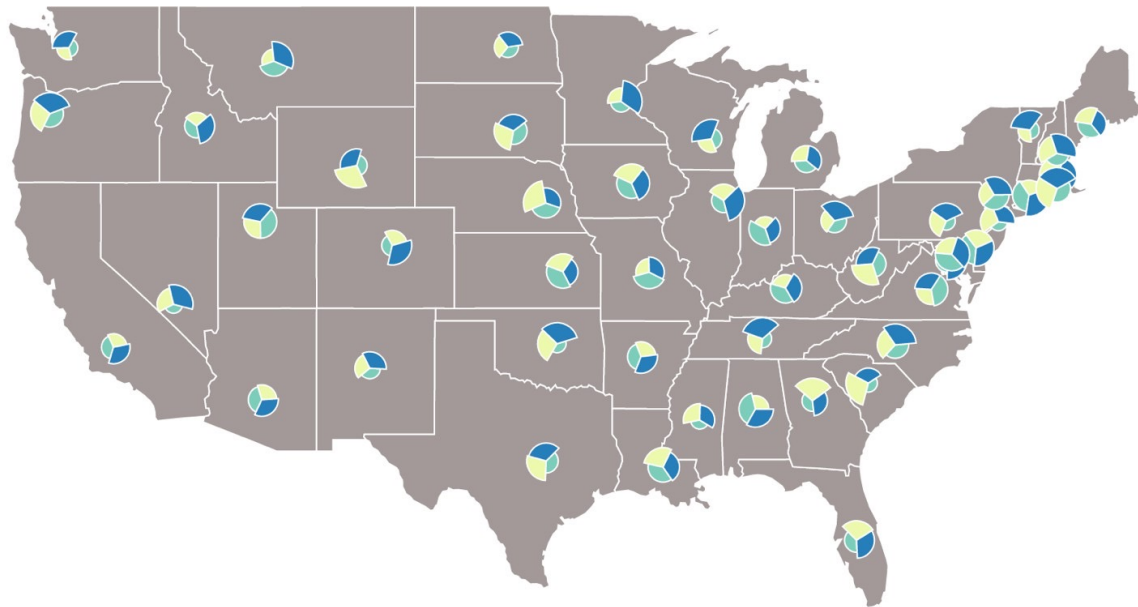
OPEN SOURCE

+

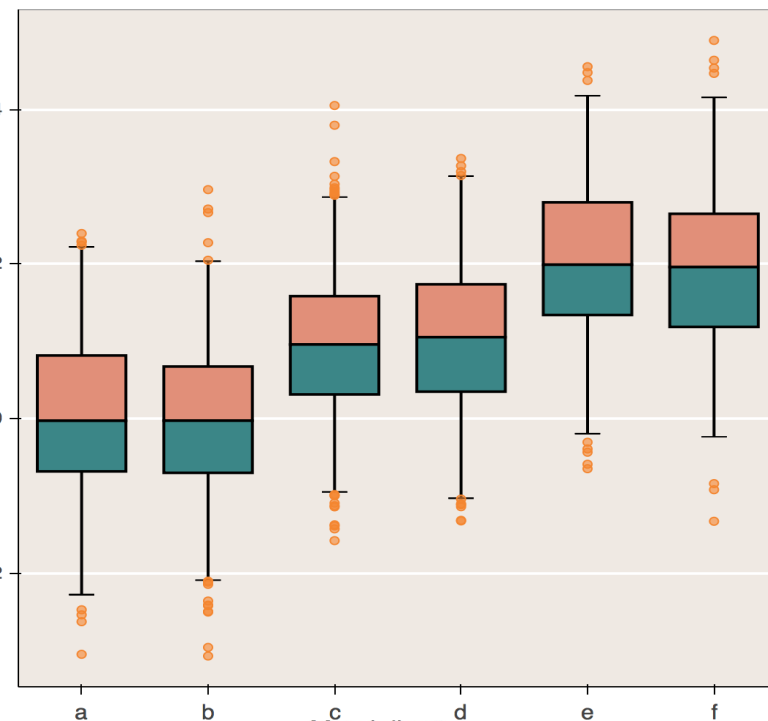
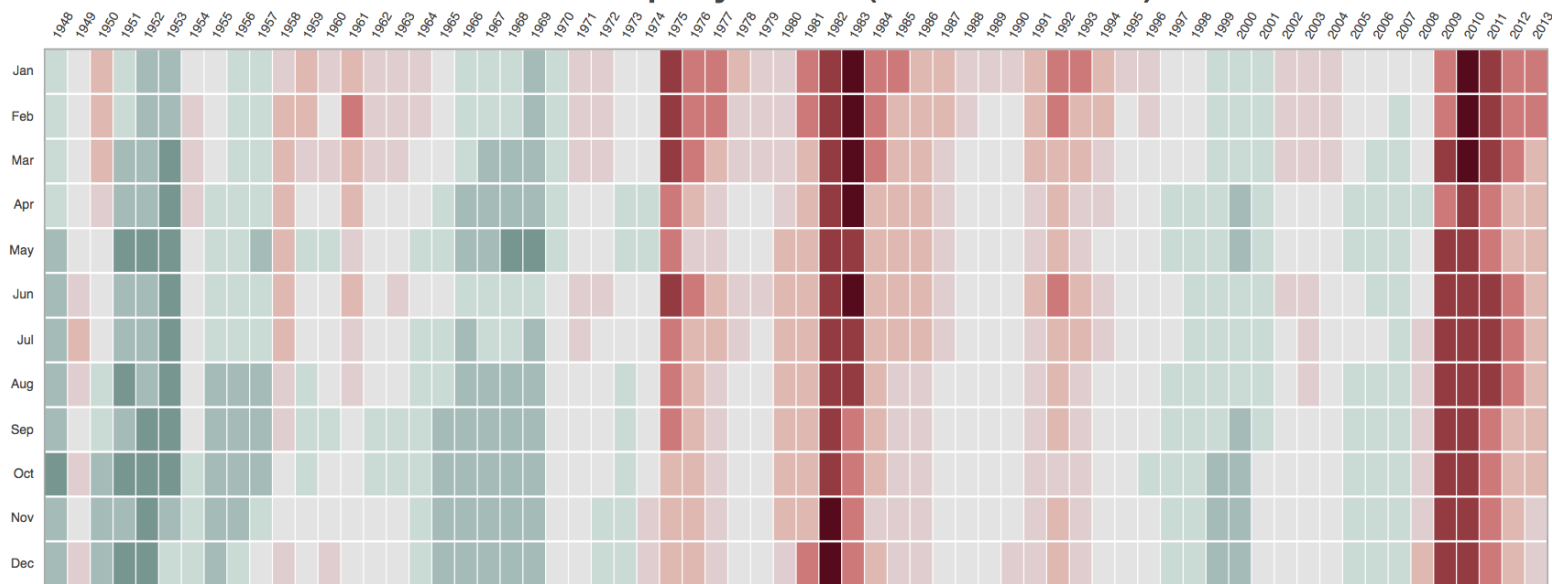
•

○





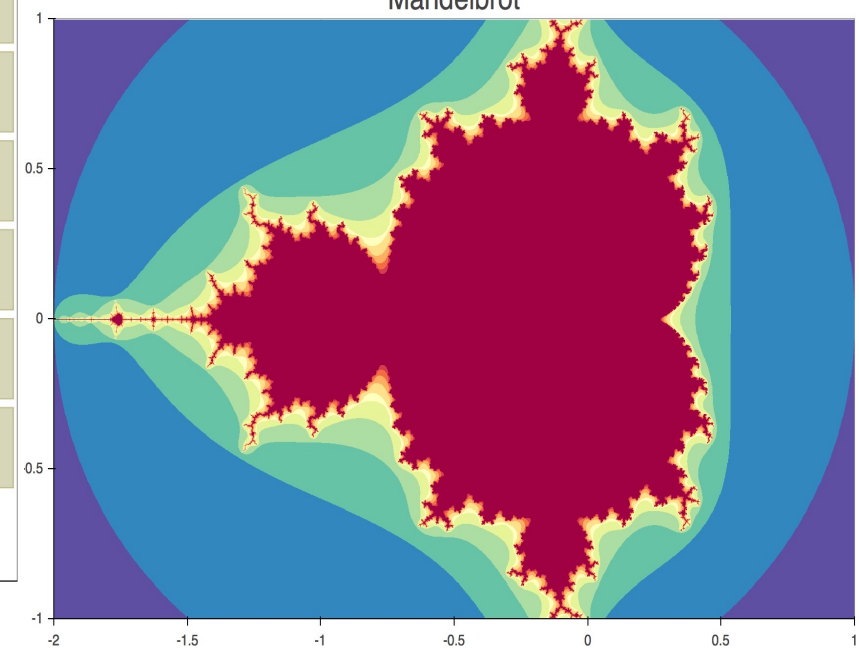
US Unemployment (1948 - 2013)



Periodic Table

1 H Hydrogen 1.00794(4)																	2 He Helium 4.002602(2)
3 Li Lithium 6.941(2)	4 Be Beryllium 9.012182(3)											5 B Boron 10.811(7)	6 C Carbon 12.0107(8)	7 N Nitrogen 14.0067(2)	8 O Oxygen 15.9994(3)	9 F Fluorine 18.9984032(5)	10 Ne Neon 20.1797(6)
11 Na Sodium 22.98976928(2)	12 Mg Magnesium 24.3050(6)											13 Al Aluminum 26.9815386(6)	14 Si Silicon 28.0855(3)	15 P Phosphorus 30.973762(2)	16 S Sulfur 32.065(5)	17 Cl Chlorine 35.453(2)	18 Ar Argon 39.948(1)
19 K Potassium 39.0983(1)	20 Ca Calcium 40.078(4)	21 Sc Scandium 44.955912(6)	22 Ti Titanium 47.867(1)	23 V Vanadium 50.9415(1)	24 Cr Chromium 51.9961(6)	25 Mn Manganese 54.938045(5)	26 Fe Iron 55.845(2)	27 Co Cobalt 58.933195(5)	28 Ni Nickel 58.6934(4)	29 Cu Copper 63.546(3)	30 Zn Zinc 65.38(2)	31 Ga Gallium 69.723(1)	32 Ge Germanium 72.64(1)	33 As Arsenic 74.92160(2)	34 Se Selenium 78.96(3)	35 Br Bromine 79.904(1)	36 Kr Krypton 83.798(2)
37 Rb Rubidium 85.4678(3)	38 Sr Strontium 87.62(1)	39 Y Yttrium 88.90585(2)	40 Zr Zirconium 91.224(2)	41 Nb Niobium 92.90638(2)	42 Mo Molybdenum 95.96(2)	43 Tc Technetium [98]	44 Ru Ruthenium 101.07(2)	45 Rh Rhodium 102.90550(2)	46 Pd Palladium 106.42(1)	47 Ag Silver 107.8682(2)	48 Cd Cadmium 112.411(8)	49 In Indium 114.818(3)	50 Sn Tin 118.710(7)	51 Sb Antimony 121.760(1)	52 Te Tellurium 127.60(3)	53 I Iodine 126.90447(3)	54 Xe Xenon 131.293(6)
55 Cs Cesium 132.9054519(2)	56 Ba Barium 137.327(7)	71 Lu Lutetium 174.9668(1)	72 Hf Hafnium 178.49(2)	73 Ta Tantalum 180.94788(2)	74 W Tungsten 183.84(1)	75 Re Rhenium 186.207(1)	76 Os Osmium 190.23(3)	77 Ir Iridium 192.217(3)	78 Pt Platinum 195.084(9)	79 Au Gold 196.966569(4)	80 Hg Mercury 200.59(2)	81 Tl Thallium 204.3833(2)	82 Pb Lead 207.2(1)	83 Bi Bismuth 208.98040(1)	84 Po Polonium [209]	85 At Astatine [210]	86 Rn Radon [222]
87 Fr Francium [223]	88 Ra Radium [226]	103 Lr Lawrencium [262]	104 Rf Rutherfordium [267]	105 Db Dubnium [268]	106 Sg Seaborgium [271]	107 Bh Bohrium [272]	108 Hs Hassium [270]	109 Mt Meitnerium [276]	110 Ds Darmstadtium [281]	111 Rg Roentgenium [280]	112 Cn Copernicium [285]	113 Uut Ununtrium [284]	114 Uuq Ununquadium [289]	115 Uup Ununpentium [288]	116 Uuh Ununhexium [293]	117 Uus Ununseptium NaN	

Mandelbrot



+

•

○

THANK YOU

Questions

