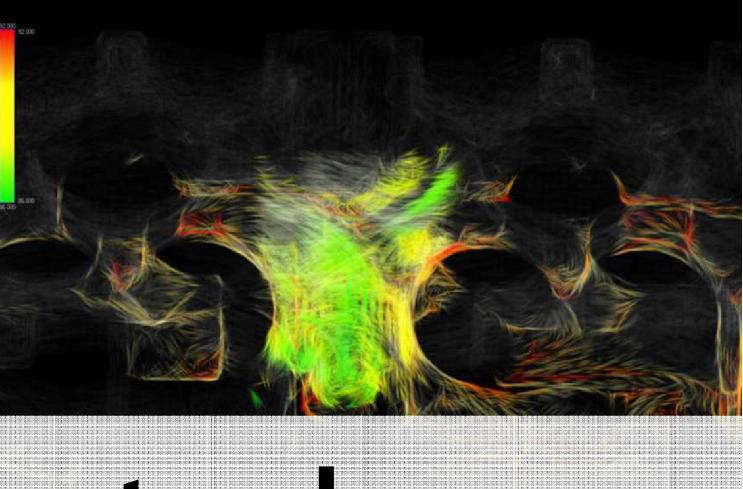
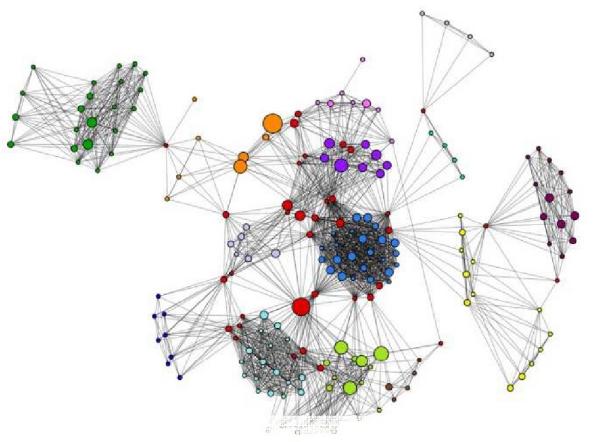


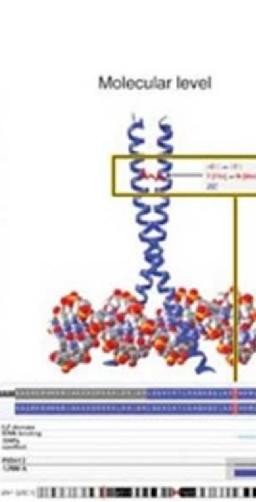
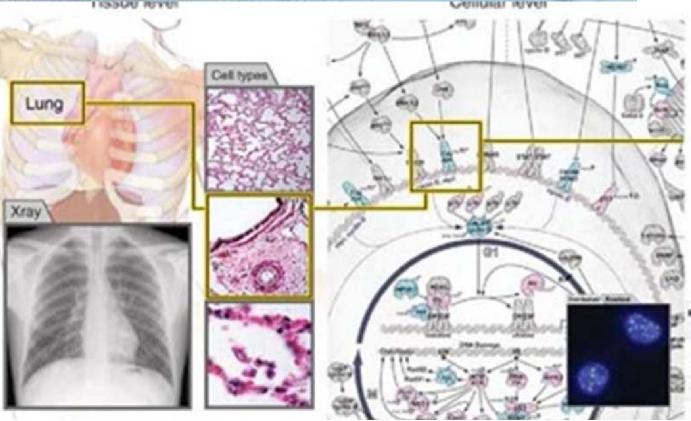
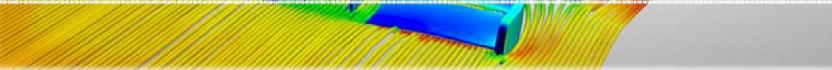
The background of the slide features a dynamic, abstract pattern of ink swirling in water. On the left side, a large, billowing cloud of teal and green ink dominates the space. To the right, a smaller, more concentrated cluster of blue, purple, and pink ink is visible, with a long, wispy trail of pink ink extending towards the bottom right corner.

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



+
°



1

Determine the
best visual



2

Balance the design



+

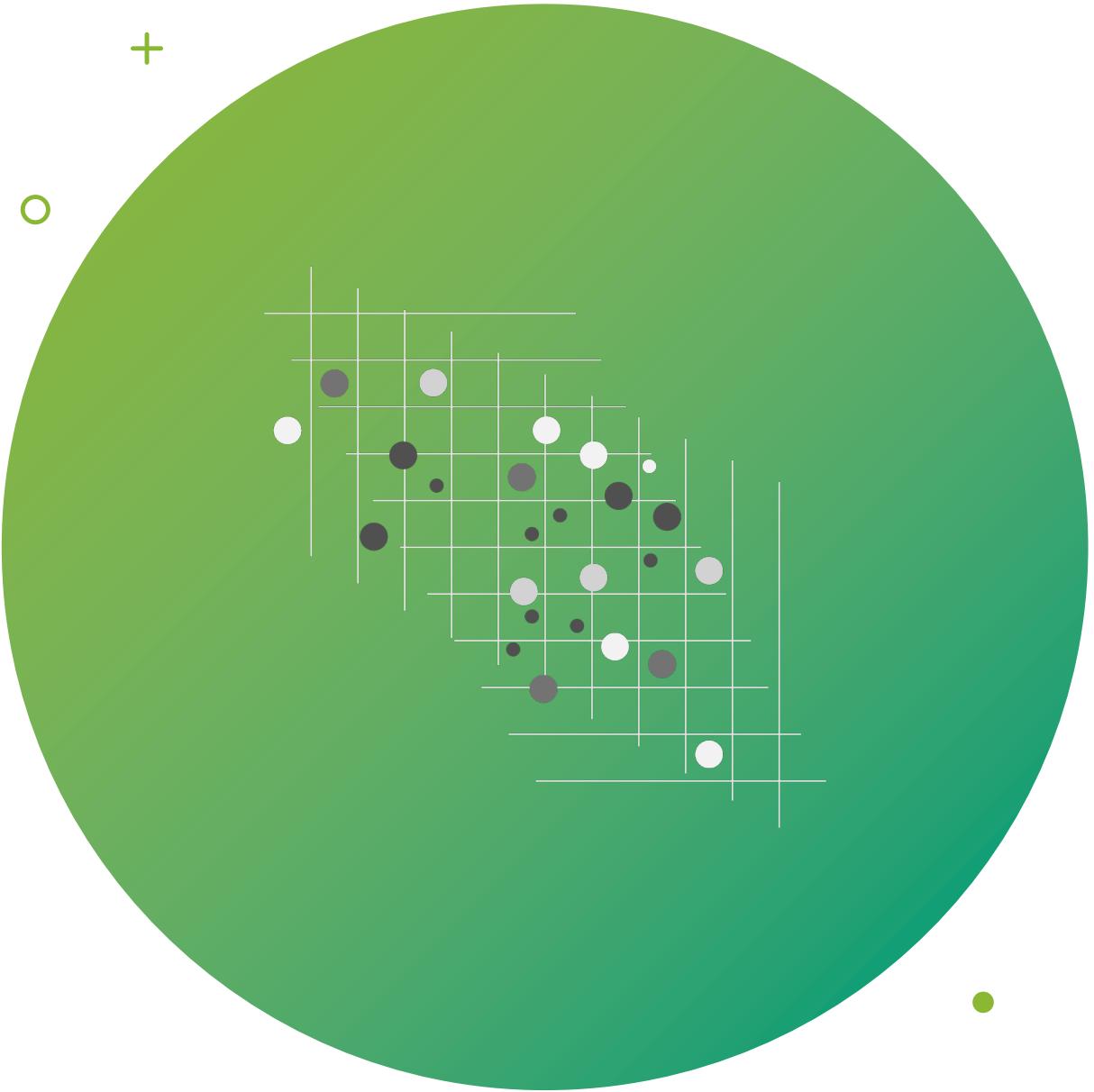
o

.



3

Focus on the key
areas

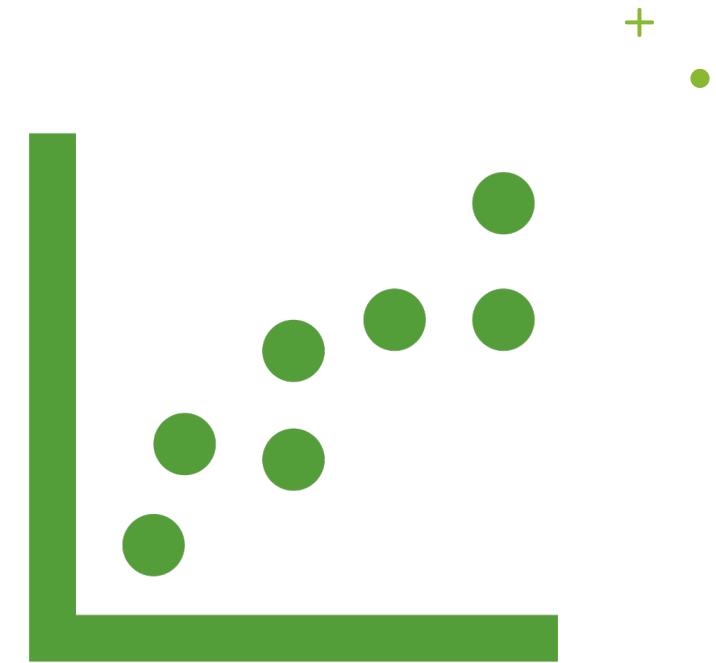


4

Keep it simple

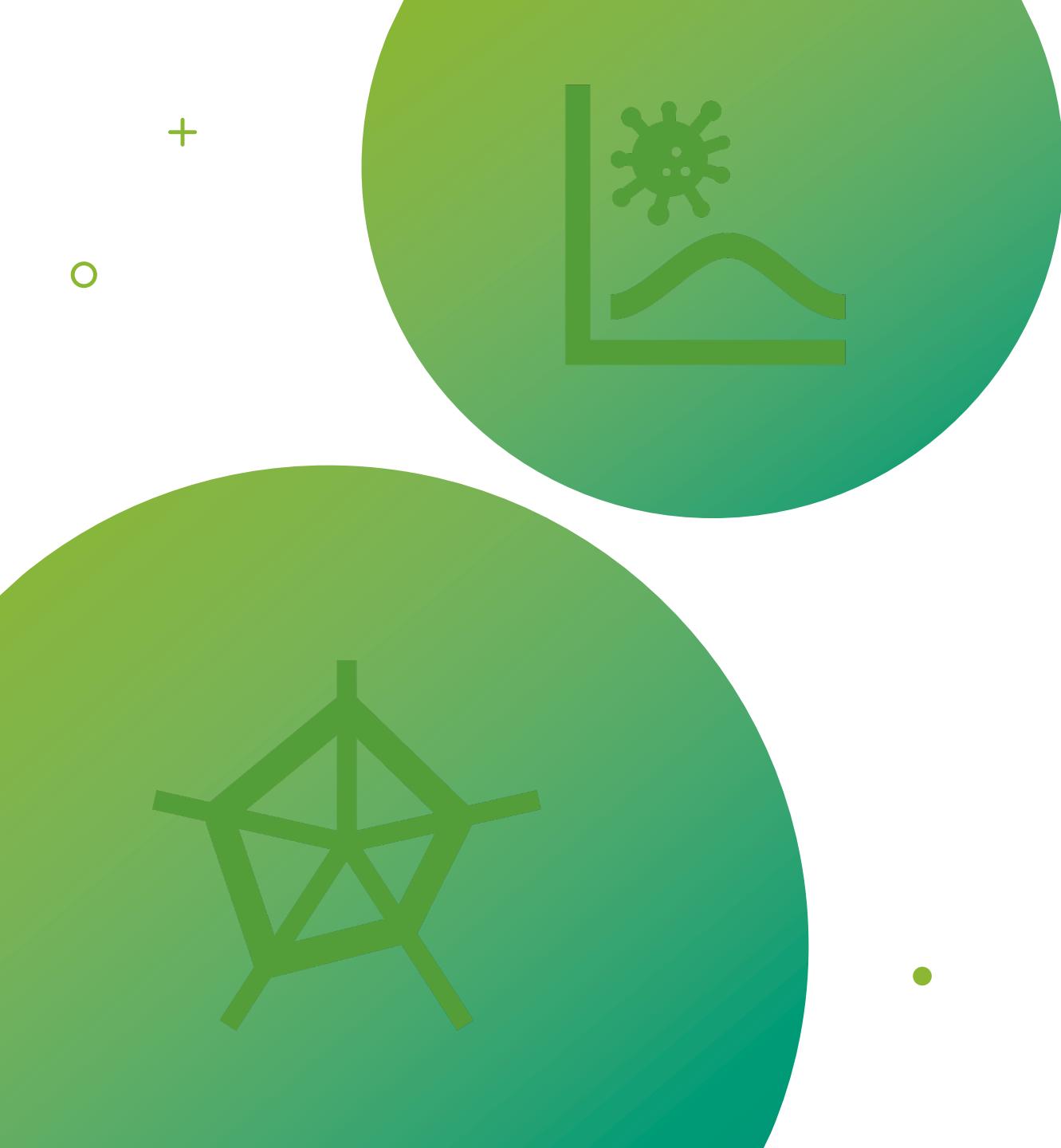
5

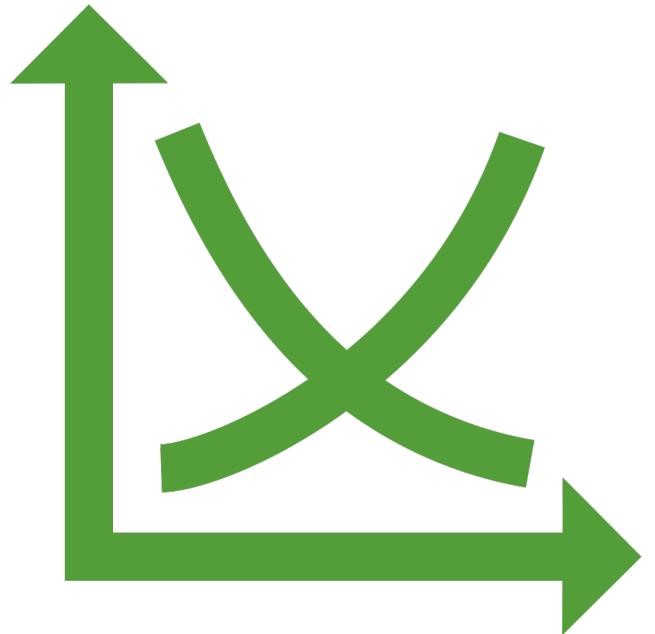
Incorporate interactivity



6

Use patterns





• +
○ 7

Compare aspects





DATA VISUALIZATION TOOLS

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

DATA PREPROCESSING



+

o

TYPES OF DATA

QUANTITATIVE

DISCRETE

CONTINUOUS

CATEGORICAL

GUIDE TO CHART TYPES

BAR CHART



PIE CHART



LINE CHART



AREA CHART



SCATTER PLOT



BUBBLE CHART



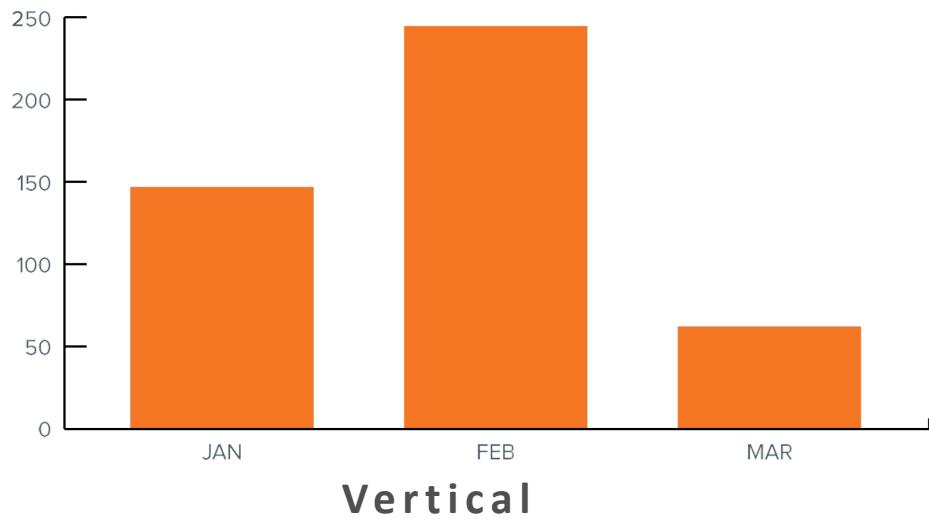
HEAT MAP



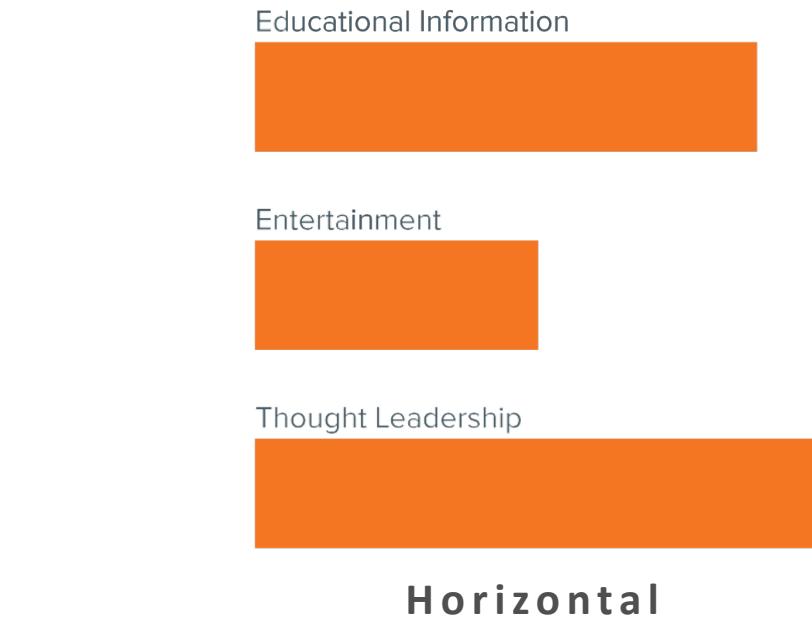
BAR CHART

Variations of Bar Charts

PAGE VIEWS, BY MONTH



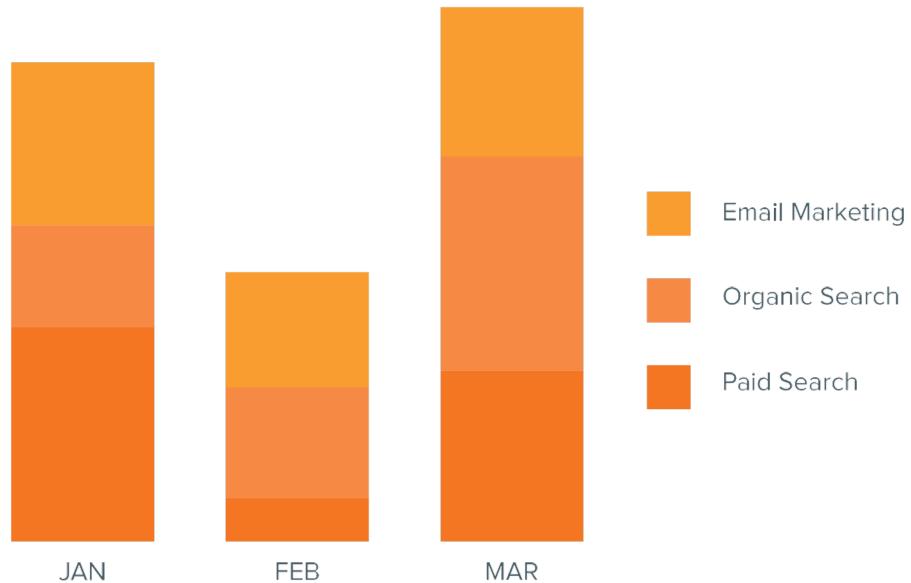
CONTENT PUBLISHED, BY CATEGORY



VARIATIONS OF BAR CHARTS (CONT.)

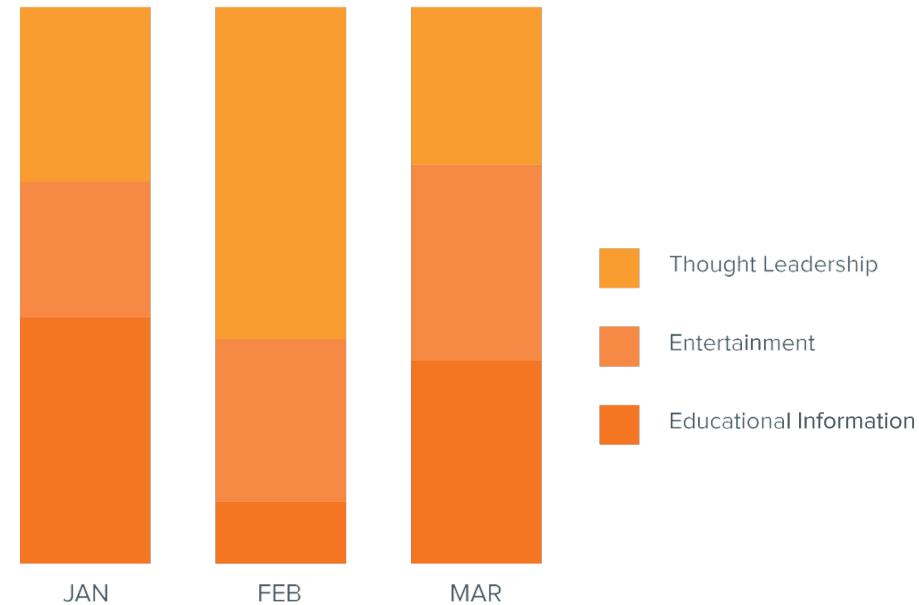


MONTHLY TRAFFIC, BY SOURCE



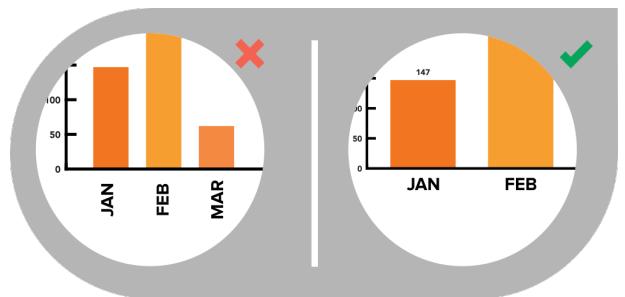
STACKED

PERCENTAGE OF CONTENT PUBLISHED, BY MONTH

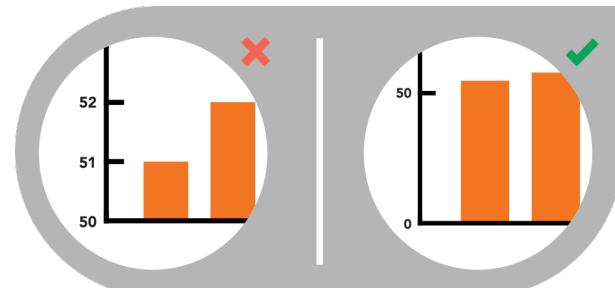


100% STACKED

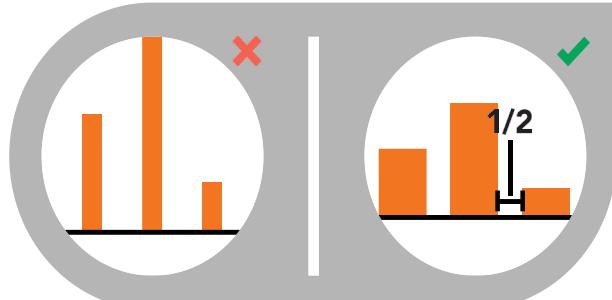
BAR CHART DESIGN BEST PRACTICES



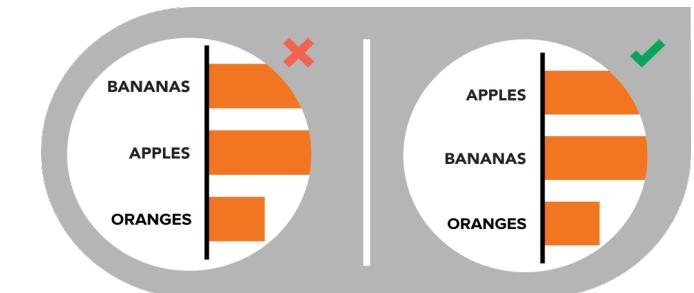
USE HORIZONTAL LABELS



START THE Y-AXIS VALUE AT 0



SPACE BARS APPROPRIATELY



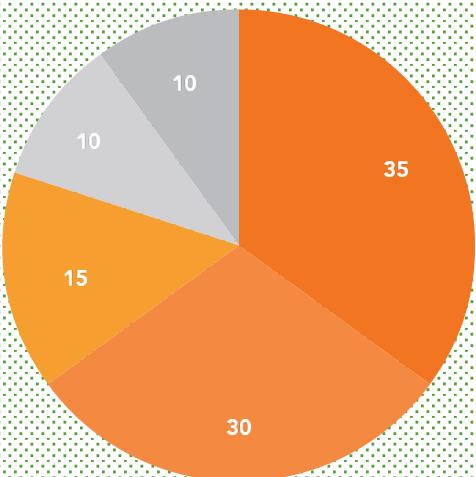
ORDER DATA APPROPRIATELY



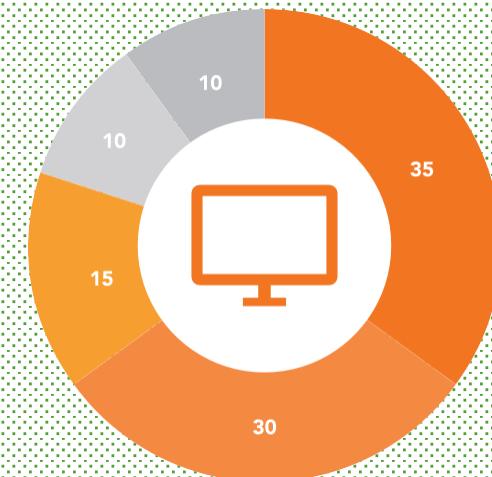
USE CONSISTENT COLORS

PIE CHART

VARIATIONS OF PIE CHARTS

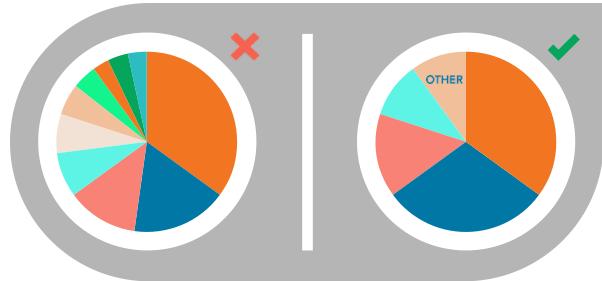


STANDARD



DONUT

PIE CHART DESIGN BEST PRACTICES



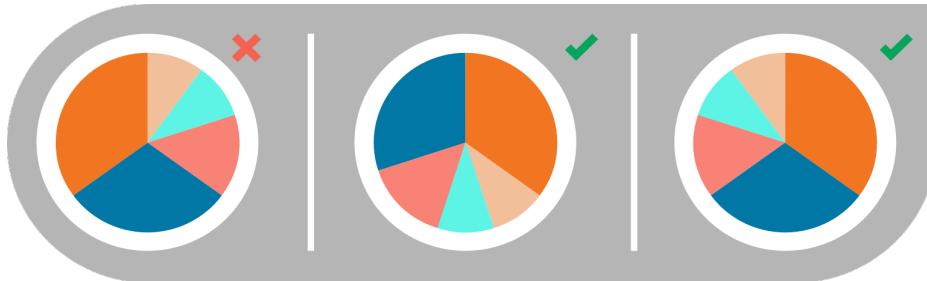
VISUALIZE NO MORE THAN
5 CATEGORIES PER CHART



DON'T USE MULTIPLE PIE
CHARTS FOR COMPARISON



MAKE SURE ALL DATA
ADDS UP TO 100%



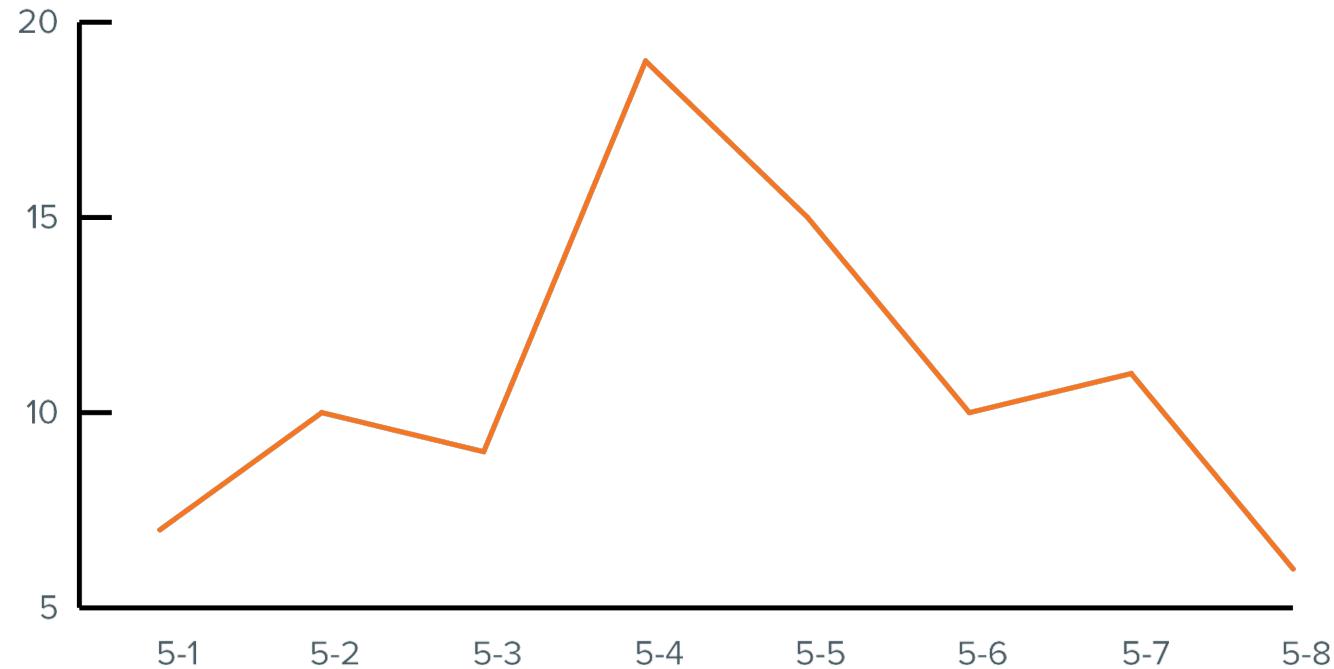
ORDER
SLICES
CORRECTLY

OPTION 1

OPTION 2

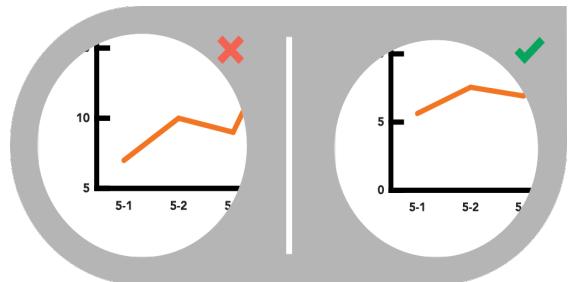
LINE CHART

DIRECT MARKETING VIEWS, BY DATE

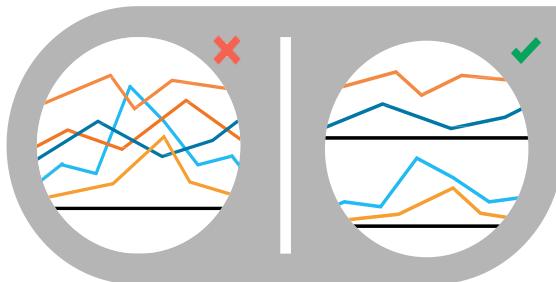


LINE CHART

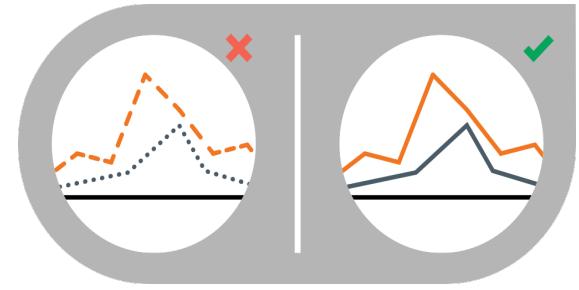
DESIGN BEST PRACTICES



INCLUDE A ZERO BASELINE IF POSSIBLE



DON'T PLOT MORE THAN 4 LINES



USE SOLID LINES ONLY



LABEL THE LINES DIRECTLY

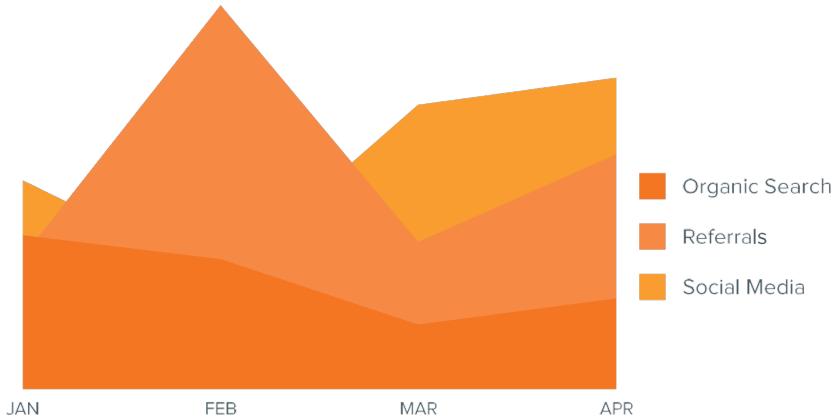


USE THE RIGHT HEIGHT

AREA CHART

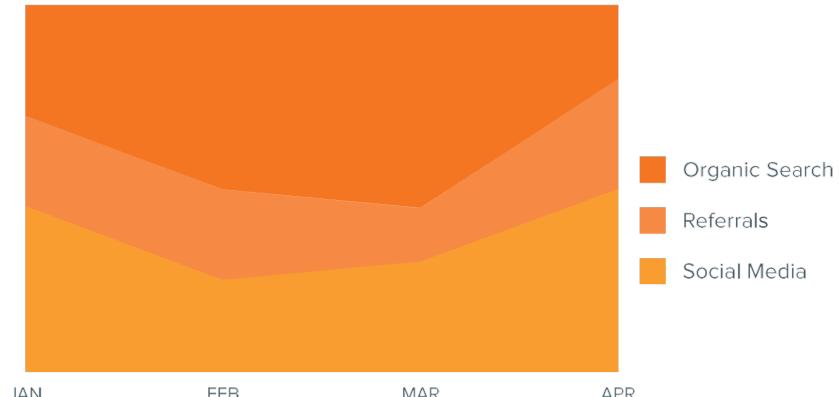
VARIATIONS OF AREA CHARTS

NEW CONTACTS, BY SOURCE



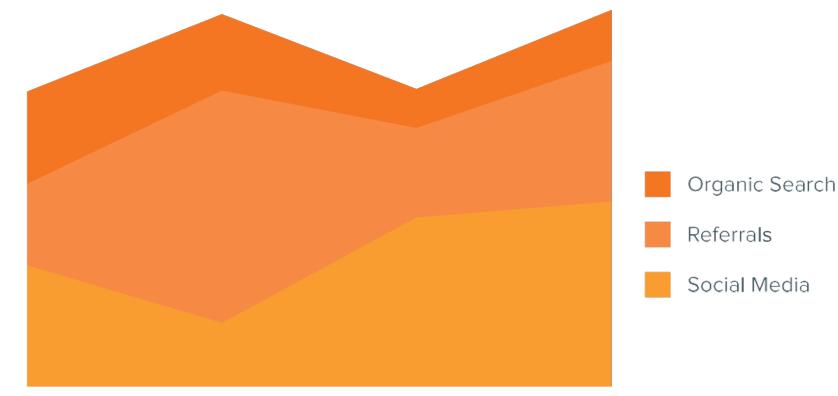
AREA CHART

NEW CONTACTS, BY SOURCE



100% STACKED AREA

NEW CONTACTS, BY SOURCE

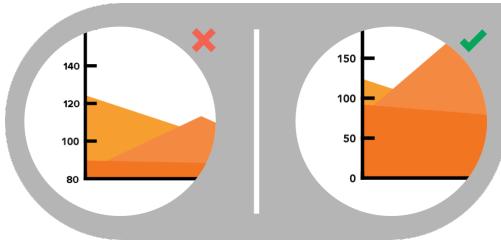


STACKED AREA

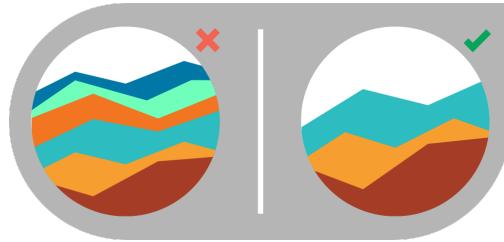
AREA CHART DESIGN BEST PRACTICES



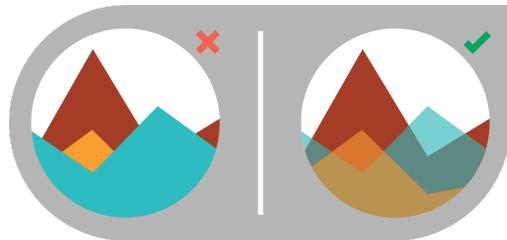
MAKE IT EASY TO READ



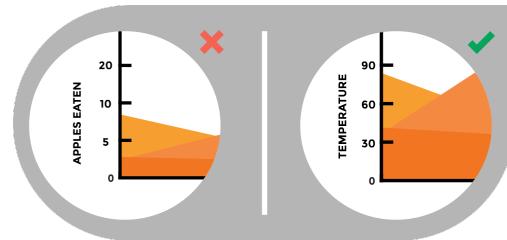
START Y-AXIS VALUE AT 0



DON'T DISPLAY MORE THAN 4 DATA CATEGORIES



USE TRANSPARENT COLORS

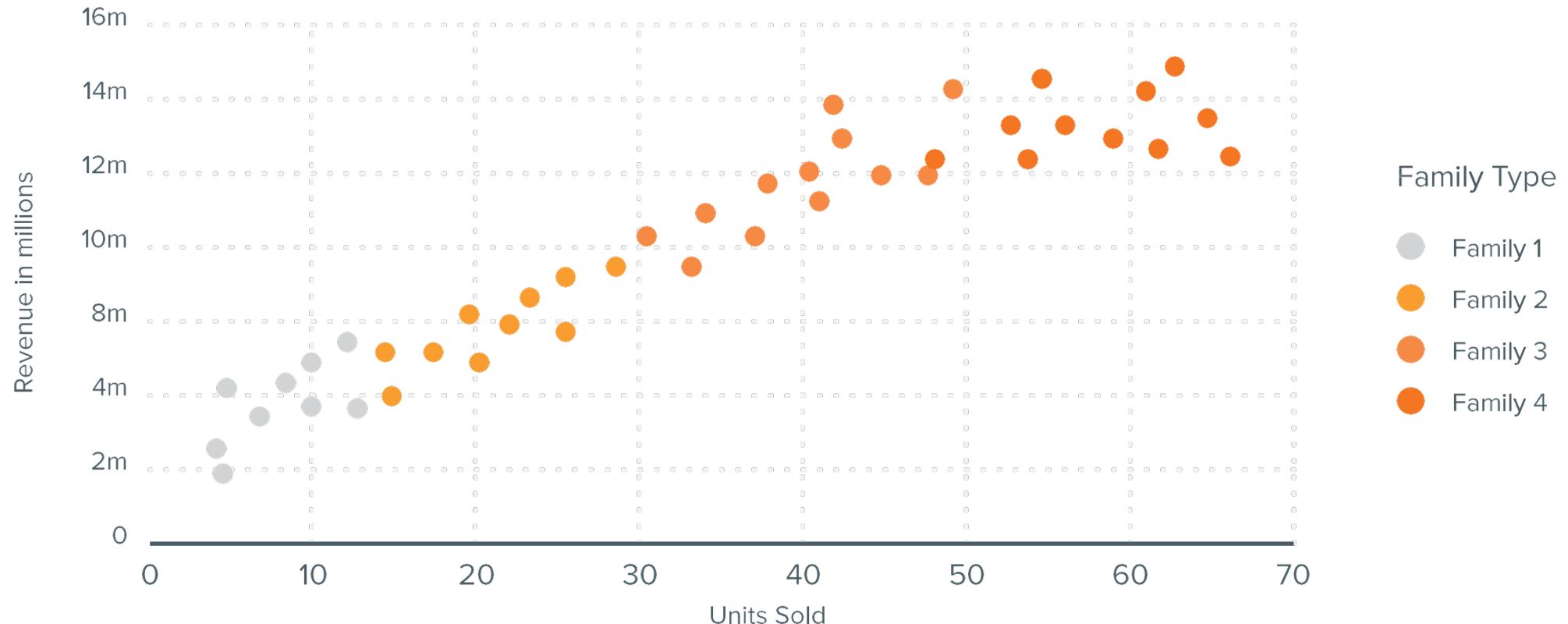


DON'T USE AREA CHARTS TO
DISPLAY DISCRETE DATA



SCATTER PLOT

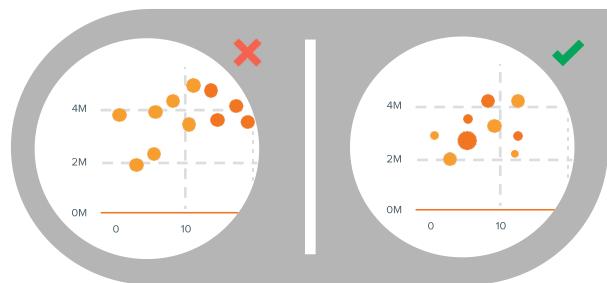
REVENUE, BY PRODUCT FAMILY



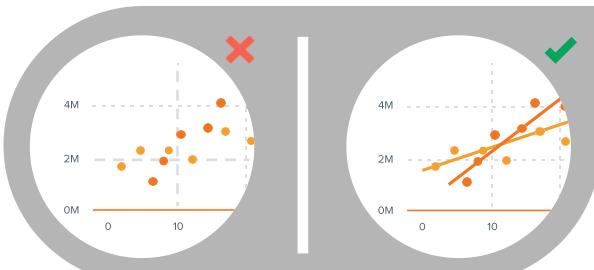
SCATTER PLOT DESIGN BEST PRACTICES



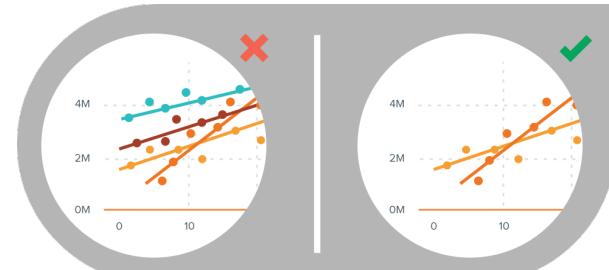
START Y-AXIS VALUE AT 0



INCLUDE MORE VARIABLES



USE TREND LINES

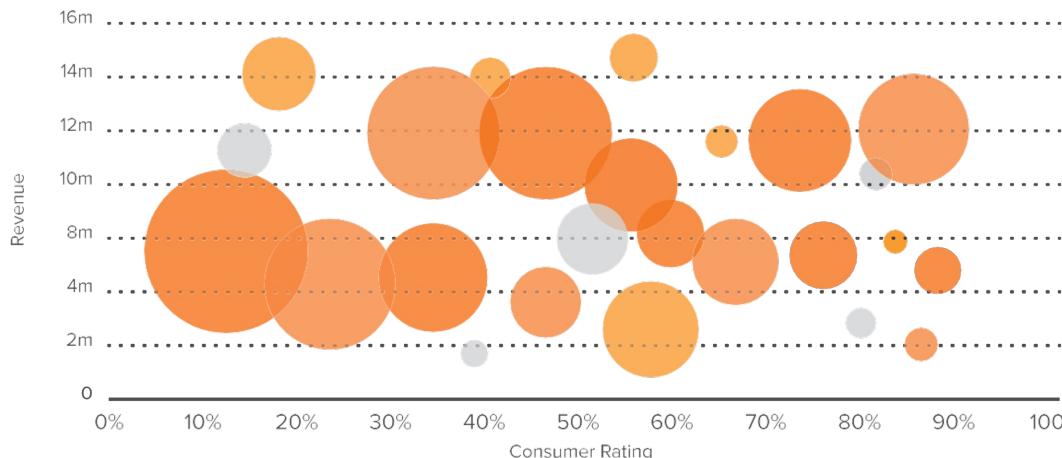


DON'T COMPARE MORE THAN
2 TREND LINES

BUBBLE CHART

VARIATIONS OF BUBBLE CHARTS

REVENUE VS. RATING

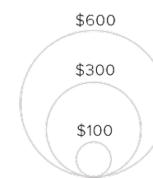


BUBBLE PLOT

Product Type

- Product 1
- Product 2
- Product 3
- Product 4

Production Cost

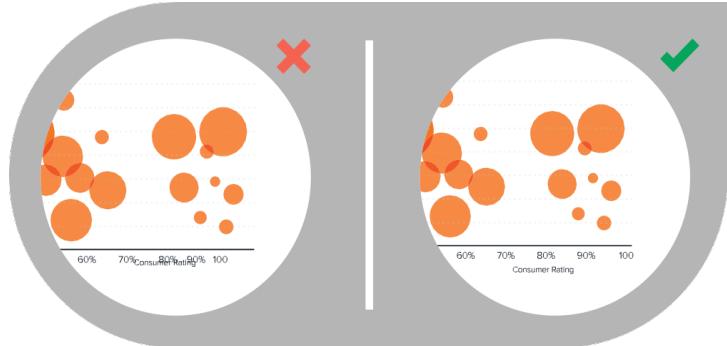


BIGGEST SALES INCREASE

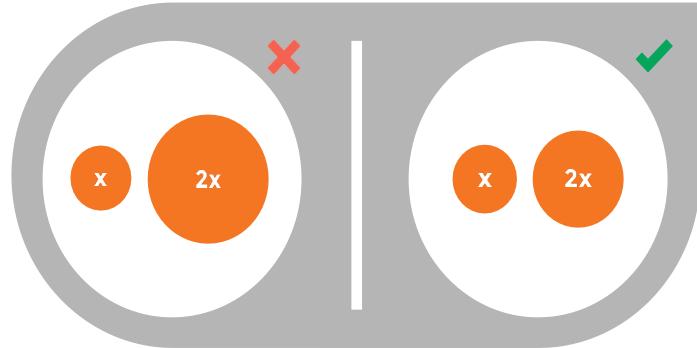


BUBBLE PLOT

BUBBLE CHART DESIGN BEST PRACTICES



MAKE SURE LABELS ARE VISIBLE



SIZE BUBBLES APPROPRIATELY

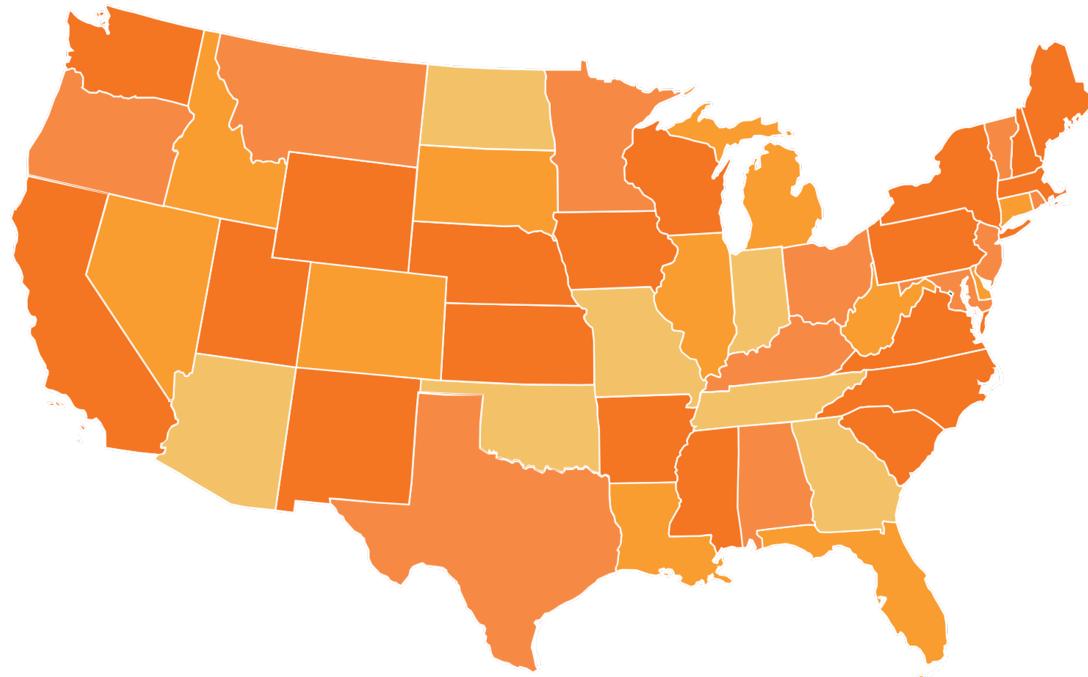


DON'T USE ODD SHAPES

HEAT MAP

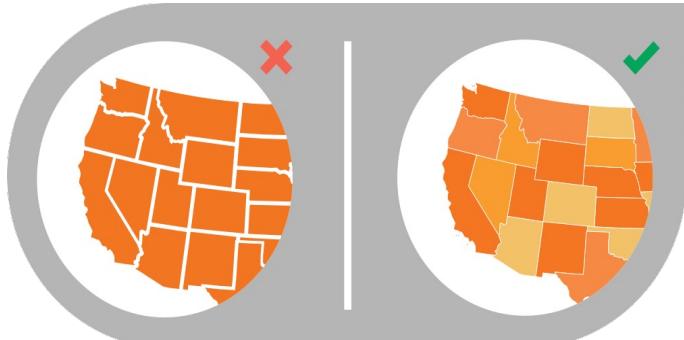


STATES WITH NEW SERVICE CONTRACTS

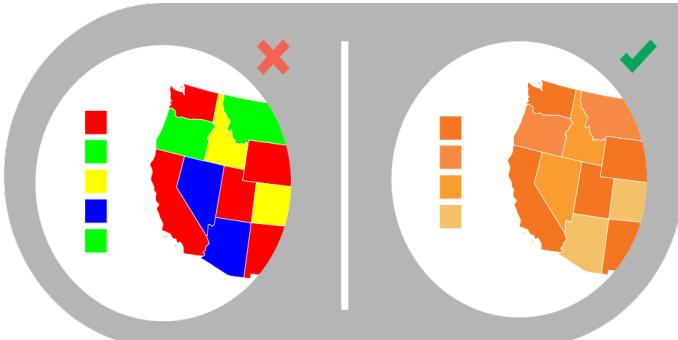


● 75-76 ● 77-78 ● 79-80 ● 81+

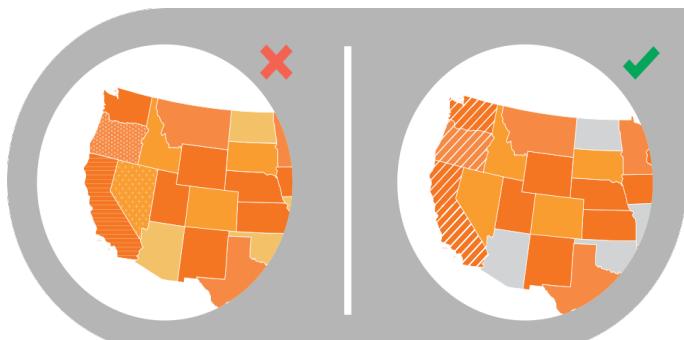
HEAT MAP DESIGN BEST PRACTICES



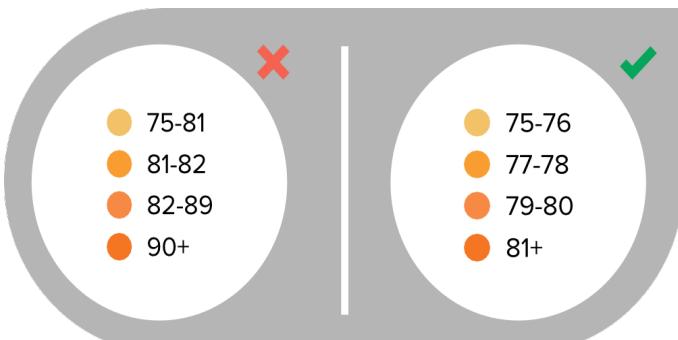
USE A SIMPLE MAP OUTLINE



SELECT COLORS APPROPRIATELY



USE PATTERNS SPARINGLY



CHOOSE APPROPRIATE DATA RANGES



+

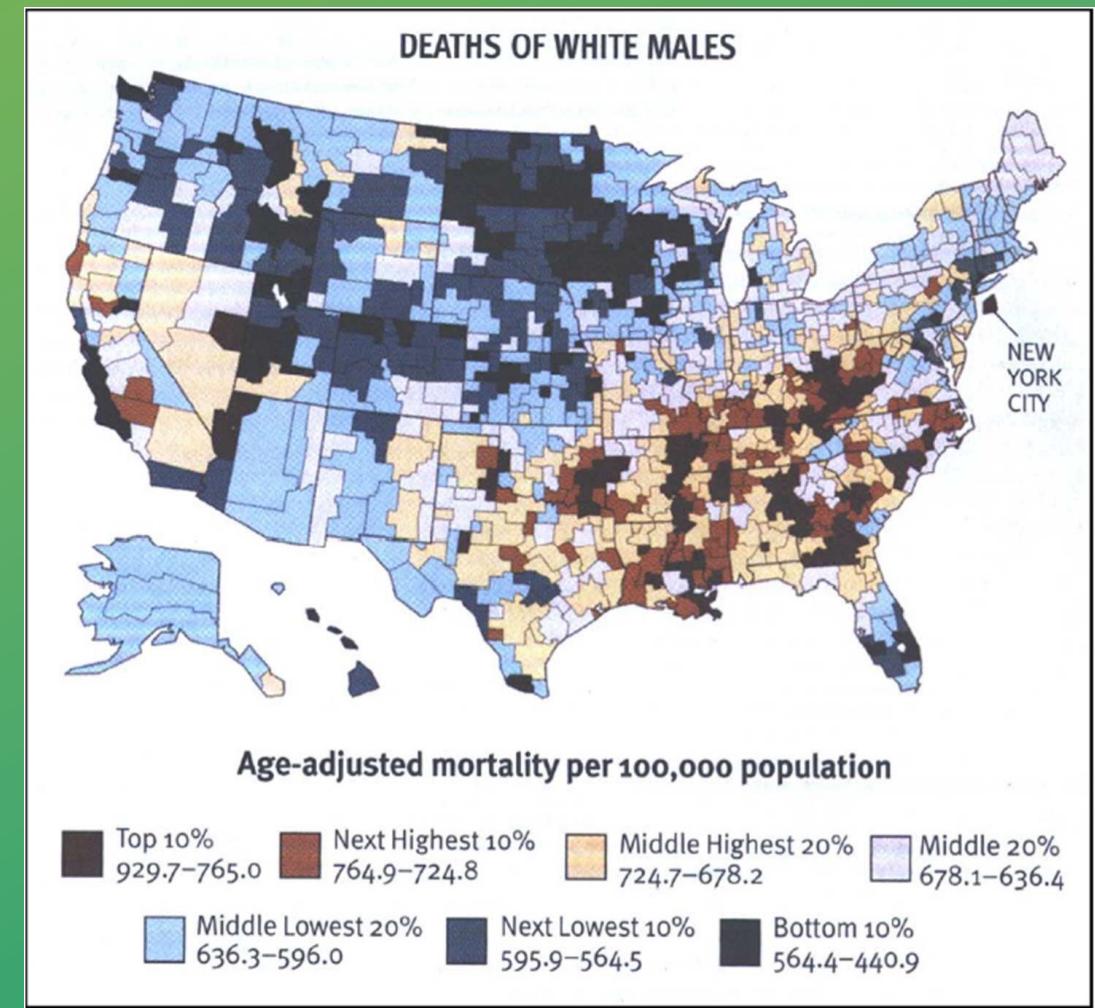
•

○

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
- o



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



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.



DATA FEATURES OF BOKEH

FLEXIBILITY

PRODUCTIVITY

INTERACTIVITY

POWERFUL

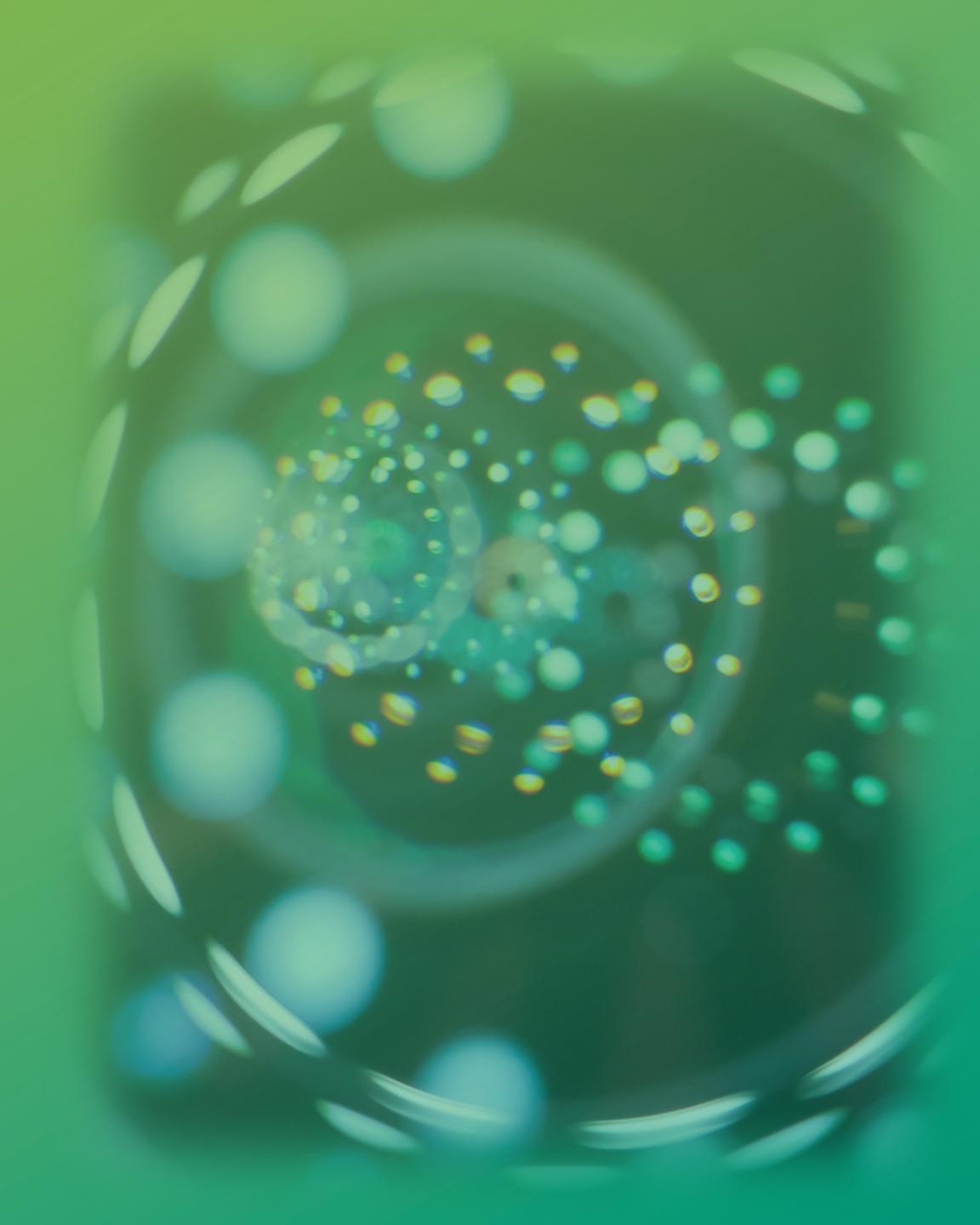
SHAREABLE

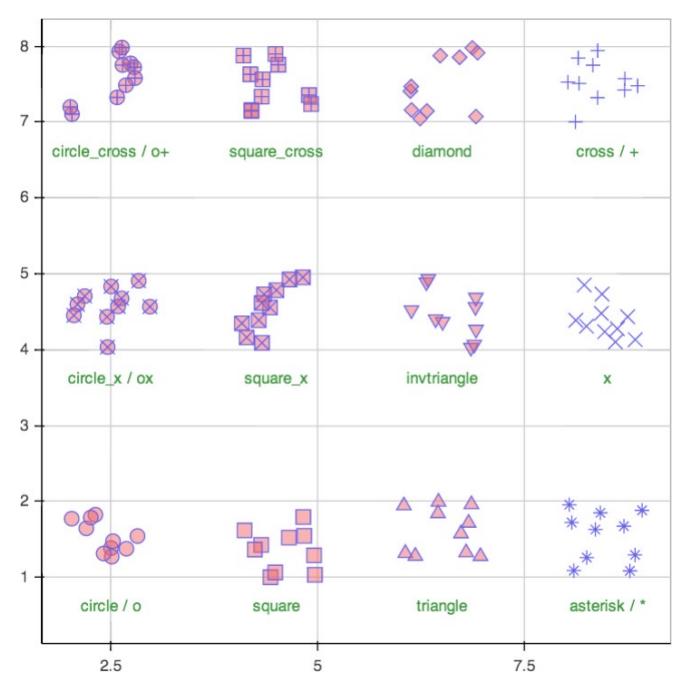
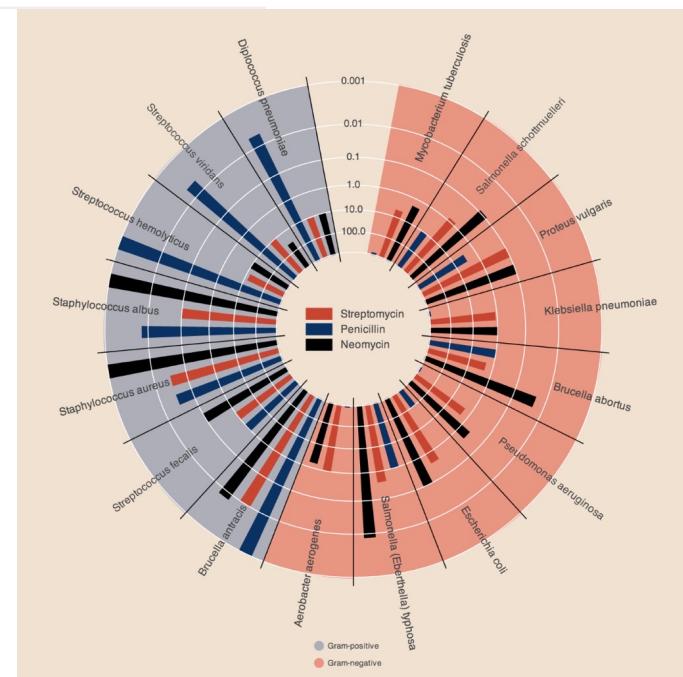
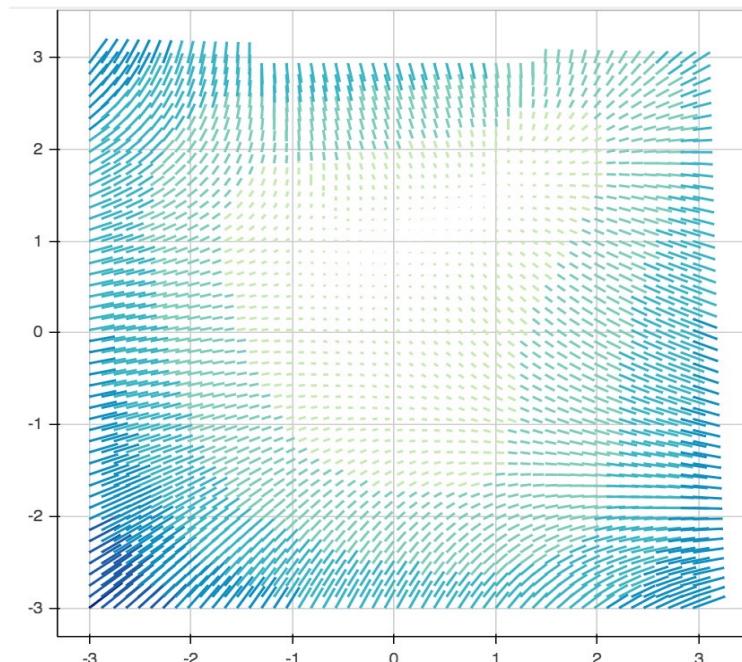
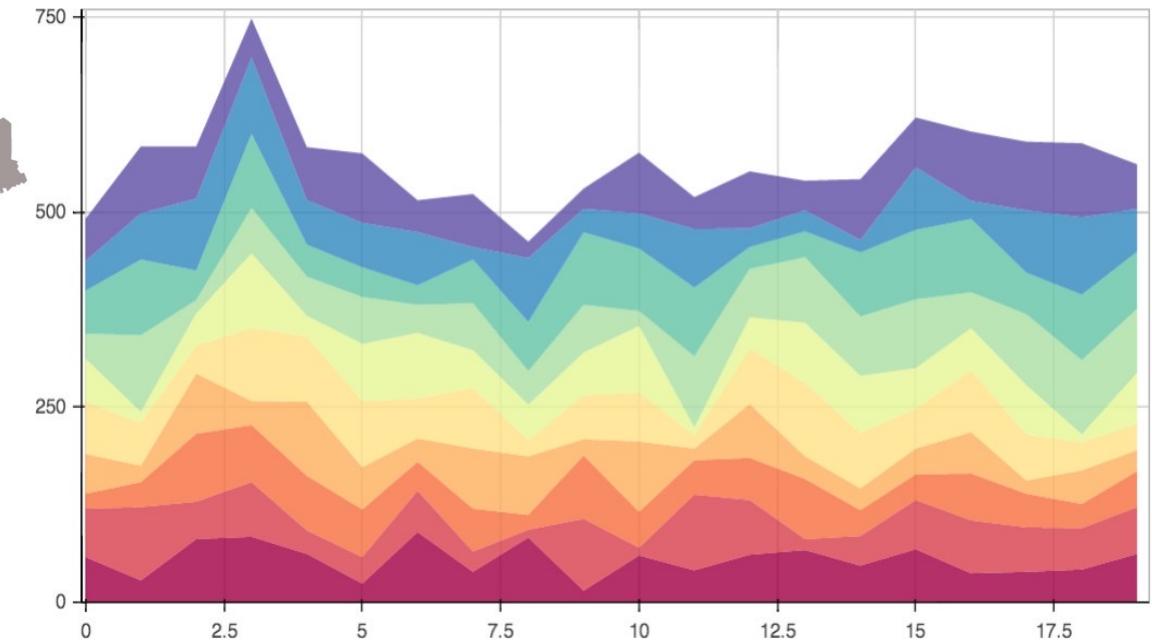
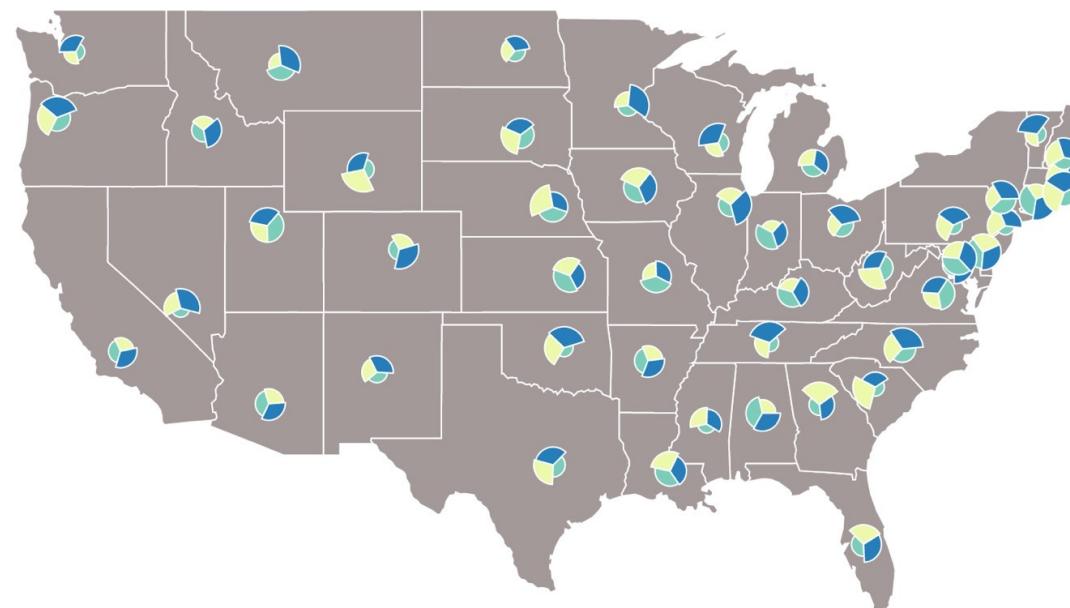
OPEN SOURCE

+

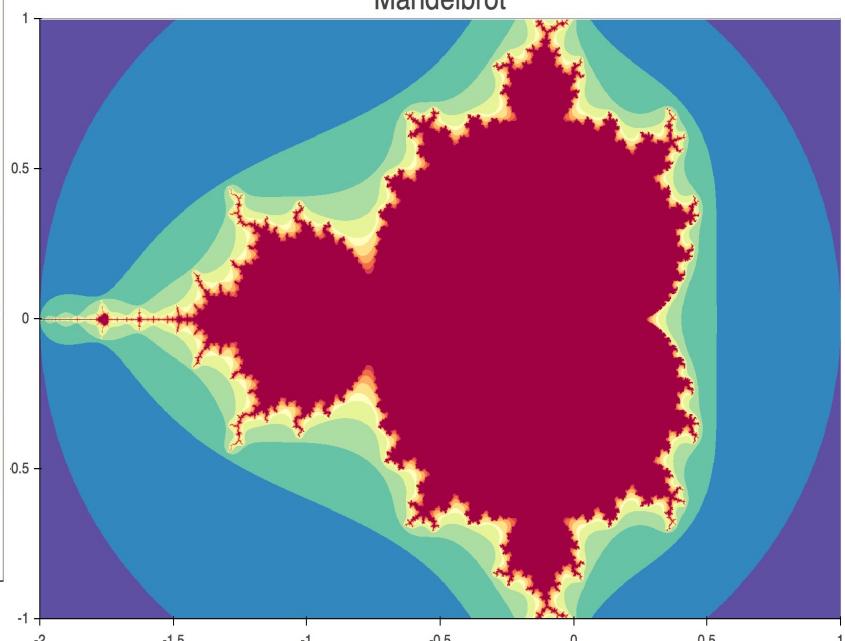
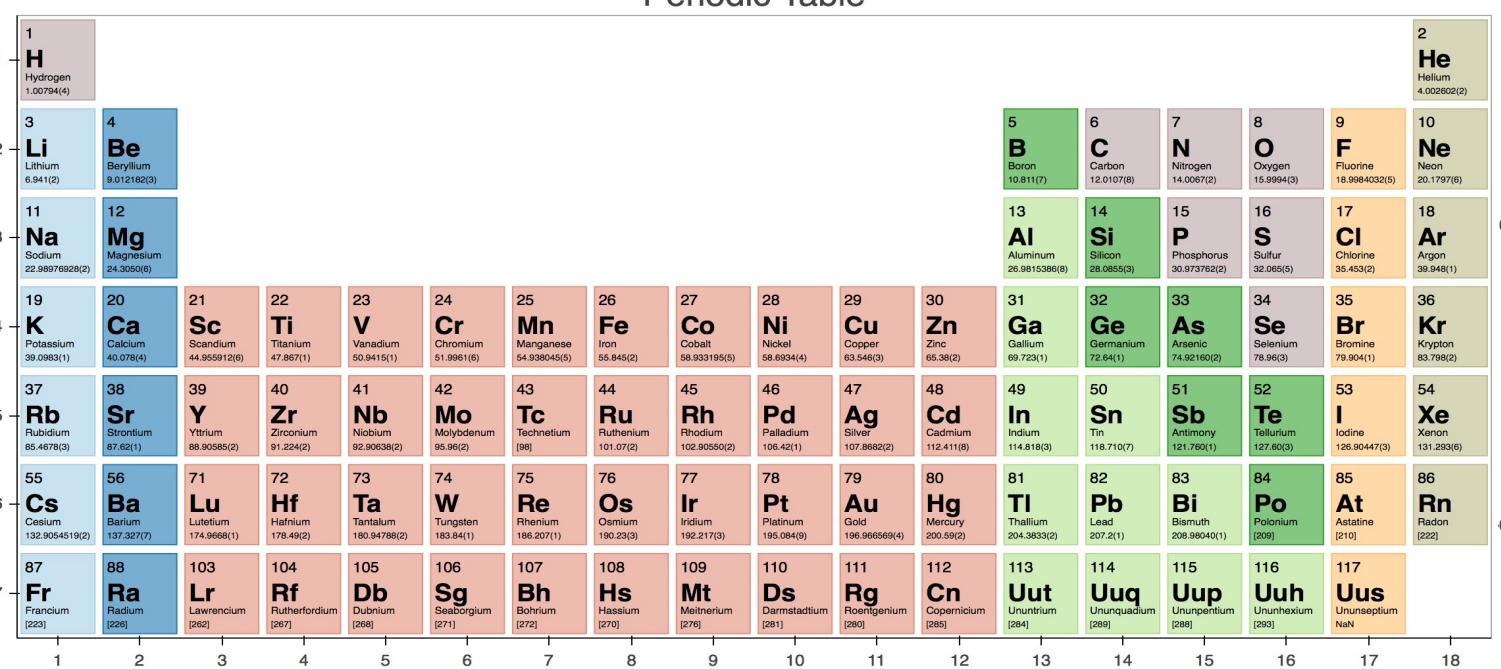
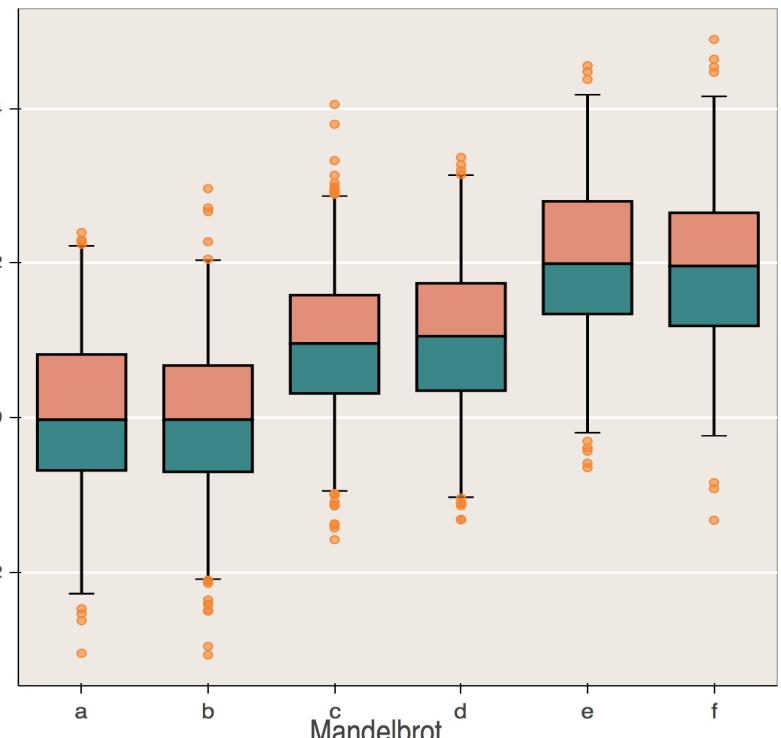
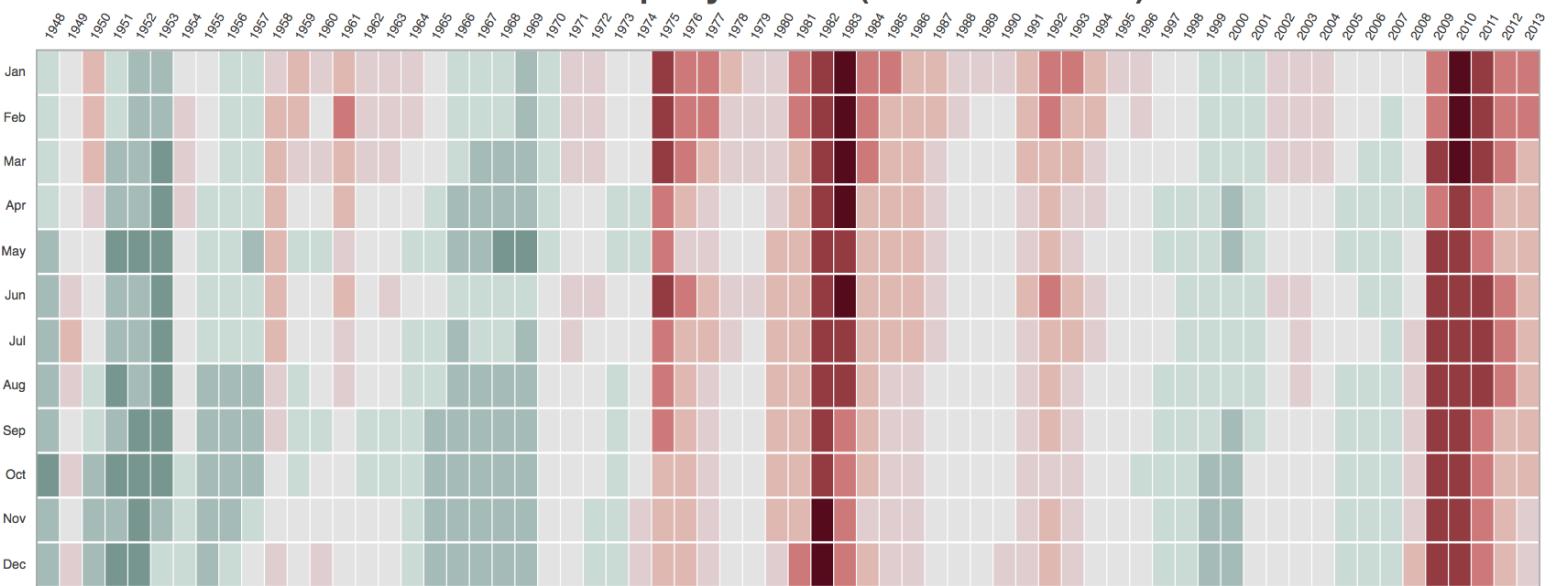
•

○





US Unemployment (1948 - 2013)



+
•
○

THANK YOU

Questions

