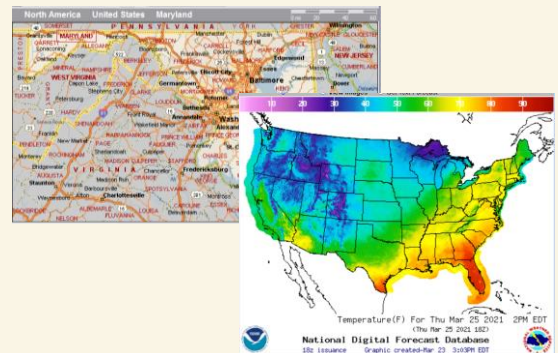


From Covid-19 to Election Results: Understanding Data Maps

Linda W. Pickle, Ph.D.

1

“... but I already understand maps!”



2



Maps are
Everywhere!

Coronavirus Maps,
P. 1 of Washington Post,
Sunday, March 14, 2021

3

Why map data?

- Many things vary by location: demographics, disease rates, behaviors, election results,...
- Summary statistics mask these geographic differences
- Technical advances now allows anyone to map data, not just specialists using proprietary software
- More data now available with geographic identifiers
- Examples of disease causes discovered by mapping

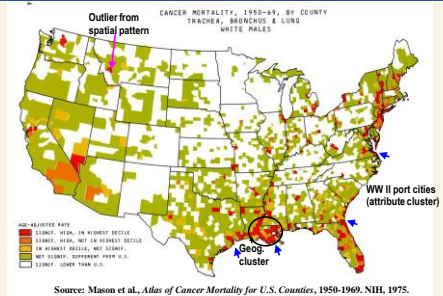
4

John Snow's Cholera Map, London, 1854



5

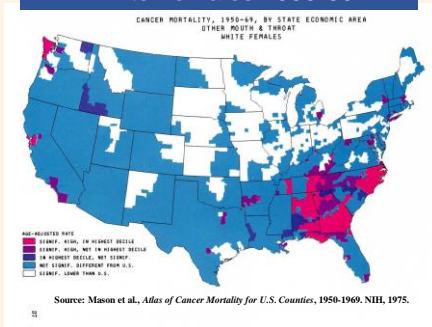
Findings from 1st NCI Cancer Atlas, 1975
Lung cancer among white males, 1950-69



Source: Mason et al., *Atlas of Cancer Mortality for U.S. Counties, 1950-1969*. NIH, 1975.

6

Oral Cancer Mortality Rates White Females 1950-69



7

Outline: What do we need to know to correctly interpret data maps?

- What is being mapped?
- What geography is represented & how?
- How are data represented?
- Cognitive biases, limitations in map reading
- How best to compare maps
- A critique of some election results maps

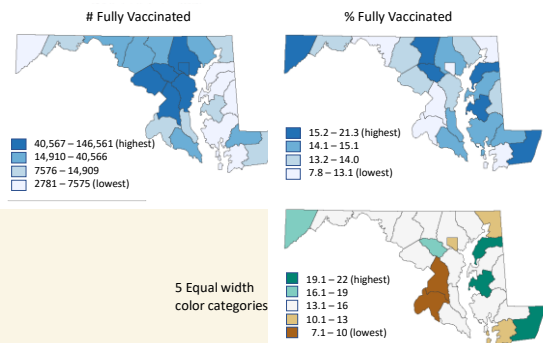
8

What is being mapped?

- Counts? Percents? Rates (crude or adjusted)?
- Are values categorized? If so, how?
 - Quantiles: an equal # places in each ranked group
 - Equal interval: each group has equal width of values

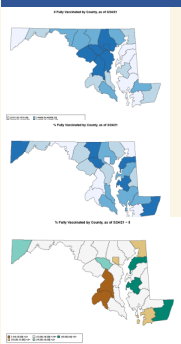
9

vs. % Fully vaccinated in MD Counties, 3/24/21 Quartiles – 6 counties in each of 4 color categories



10

What question is best to ask of each design?



- Counts
 - Resource allocation, e.g., # vaccine doses
 - Disease burden, e.g., # cancer cases
- Quantile (equal # places in each category)
 - Rank order of places more important than the actual value
- Percents (and rates)
 - Account for population differences
 - Relative value more important than actual count
- Equal width categories
 - When actual values more important than rank order of places
- Diverging (2 hues) color scheme
 - When there is a middle (referent) value of interest for comparison, e.g., a US rate

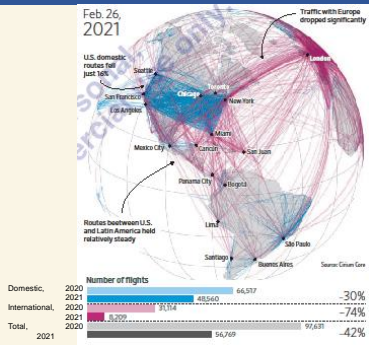
11

How is geography represented?

- What geographic areas are on the map?
- Geographic units need to match data units
- Projection – representing 3D image on 2D map
- Cartograms – a deliberate distortion

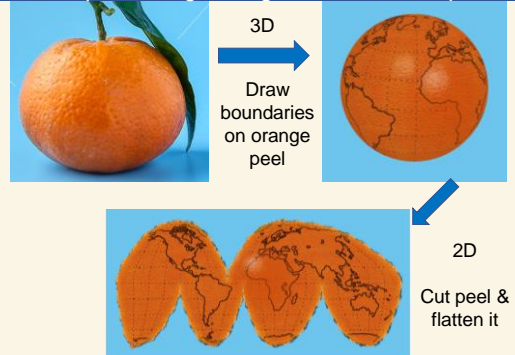
12

Show Data in a 3D Image (like a globe)



13

Map Projections Explained: Representing a 3D globe on a 2D map



Source: John Nelson, <https://www.esri.com/arcgis-blog/products/arcgis-pro/education/earth-peel/>

14

2 Common Projections



Mercator projection distorts areas away from equator, preserves direction & shape. Note: Greenland (white) looks to be same size as Africa, but Africa is 14 times the area of Greenland.



Albers projection distorts shape, preserves areas. Note: Greenland (arrow) is now much smaller than Africa.

15

Which projection is this?



16

What projection is best for U.S. maps?

Area-preserving (Albers)

Not area-preserving (lat-long)



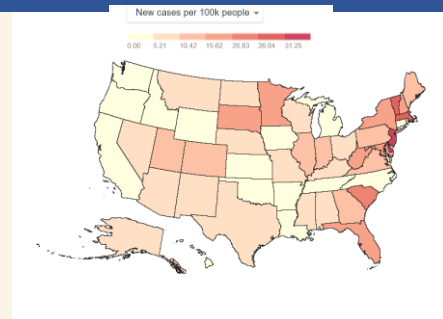
Blue



Green

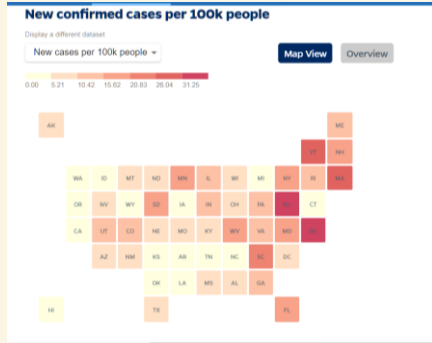
17

New daily coronavirus cases per 100k pop, 3/22/21, JHU Coronavirus Center



18

Cartogram of new daily coronavirus cases per 100k pop, 3/22/21, JHU Coronavirus Ctr.



19

How are data represented?

- Symbols vs colors – pros & cons
- Circles are better because they eliminate area bias
- But
 - We don't judge circle area as well as length, so can be fooled about relative sizes of circles
 - Sometimes difficult to find a range of circle sizes that can represent very large values AND show various small values
 - With many areas to represent, circles can overlap, making them hard to read
- Color doesn't have these problems, but visual bias by area size can fool the eye (e.g., NYC vs MT)

20

Representing values by circles or colors

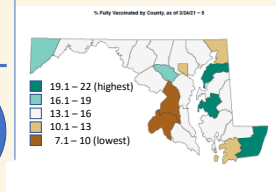
Circle's AREA is proportional to value

Diameter of B is 2 times that of A



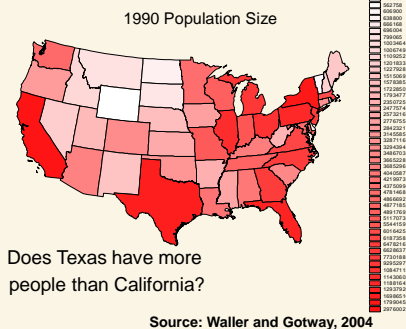
But Area of B is 4 times that of A

Match color on map to legend to read value



21

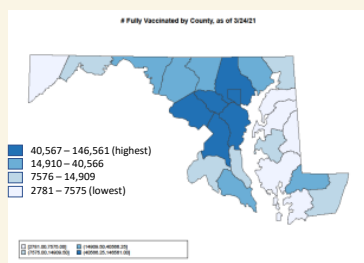
Are values categorized or not? Example of uncategorized color-coding



Source: Waller and Gotway, 2004

22

Better: Categorized color ranges



23

Cognitive research into map reading

- Questions asked of a map
 - Reading a value for 1 area
 - Identifying patterns on a map
 - Comparison of patterns on multiple maps
- Map elements studied:
 - General style
 - Legend
 - Categorization of rates
 - Color
 - Indicating unreliability

Sources: Pickle, *Spatial & Spatio-temporal Epidemiology*, 2009
Carr & Pickle, *Visualizing Data Patterns with Micromaps*, CRC Press, 2010.

24

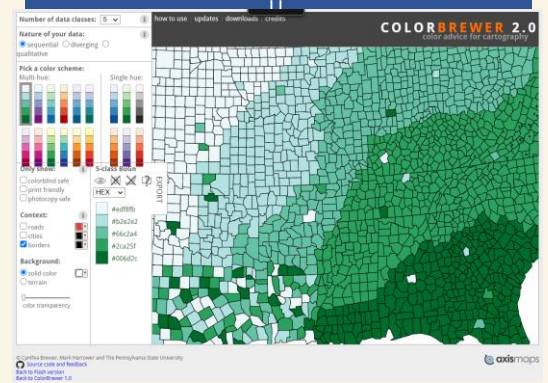
Results of Color Research

- Very distinct colors best for rate readout (Hastie 1995)
- Color gradient best for pattern recognition (Lewandowsky 1995): For low to high values, use light to dark OR cool to warm
- Diverging (double-ended) color scheme combines 2 distinct hues + color gradient; good for highs & lows
- Excellent source of info on color choices: ColorBrewer

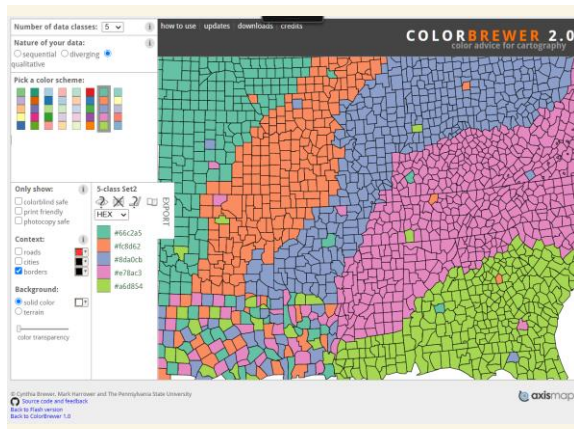
Source: Brewer CA, MacEachren AM, Pickle LW: Mapping mortality: Evaluating color schemes for choropleth maps. Annals of the Amer Assoc of Geographers, 1997.

25

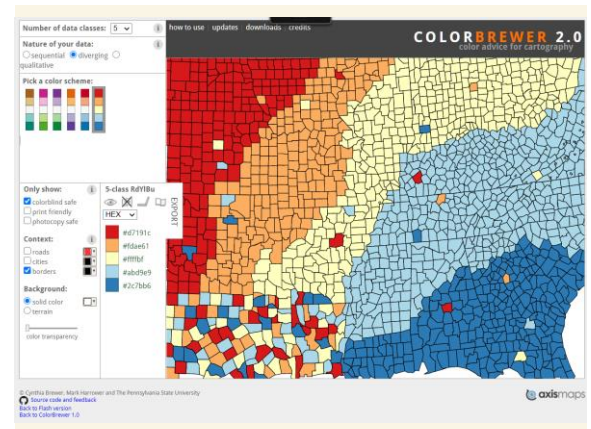
<http://colorbrewer2.org> by Brewer & Harrower, Penn State



26



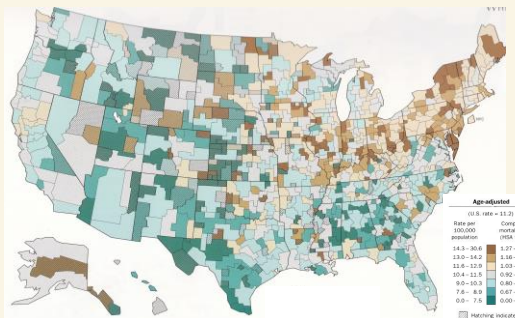
27



28

White Female Colon Cancer Mortality, 1988-92

Broad patterns: Rates are high in NE & Great Lakes regions

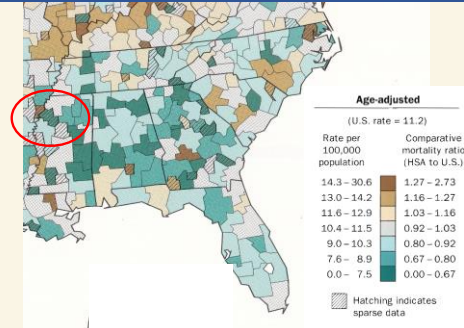


Source: Pickle et al., NCHS, 1996

29

White Female Colon Cancer Mortality, 1988-92

Unreliable local rates are hatched over color.



Source: Pickle et al., NCHS, 1996

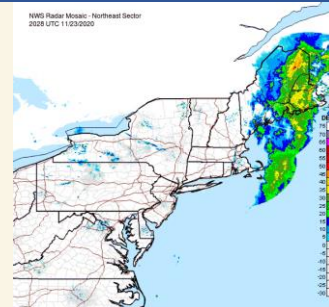
30

Comparing Maps in Series

- Methods for map series display:
 - Animation – good for smoothly changing patterns
 - Side-by-side maps ("small multiples")
 - Map the differences between pairs of maps
- Accuracy of comparing maps significantly better when the same method of representation is used for each map (Brewer & Pickle 2002)

31

NWS Radar Map of Northeast US 11/23/20



Source: https://radar.weather.gov/Conus/Loop/northeast_loop.gif, accessed 11/23/20, 5pm

32

The Coronavirus Pandemic: Month where each county reached its peak # cases.

Source: Tierney L, Meko T. The Coronavirus Pandemic: More than 250,000 lives lost. Wash Post, p. A18, Nov 22, 2020.

33

March



34

April



35

May



36

June



37

July



38

August



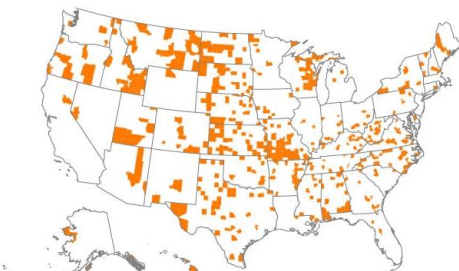
39

September



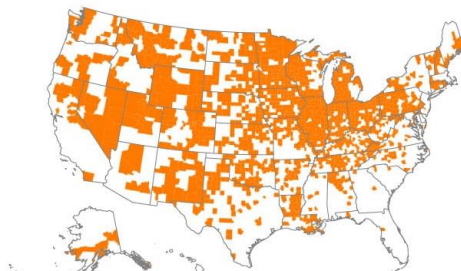
40

October



41

November (1-17)



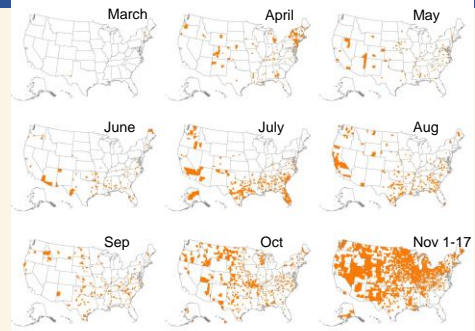
42

Impact of Attention & Memory on Animation

- Attention span: can attend to only 3-4 items at once
- In order to see the change on the map, we must focus our attention on relevant part of image
- Types of blindness relevant to animation:
 - Attentional blindness: looking at wrong part of image, e.g., attention drawn away by movement, saturated colors or large objects
 - Change blindness: 1st image mentally erased before 2nd image is seen, so differences not seen

43

Wash Post, 11/22/20, p.A18, The Coronavirus Pandemic; When each county reached its peak # of cases.



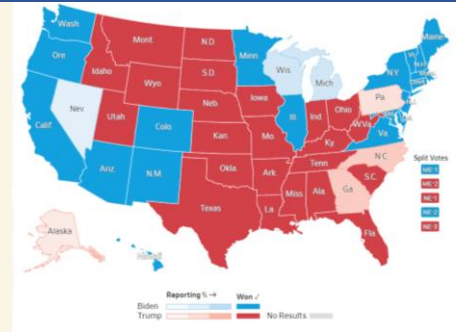
44

A Critique of Election Maps Published by the New York Times, Wall Street Journal, & the Washington Post

all data as of Nov 4, 2020

45

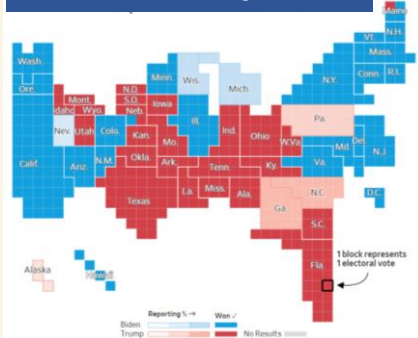
Winner of Popular Vote by State (11/4/20)



Source: Wall Street Journal online, downloaded 11/4/20, 2pm

46

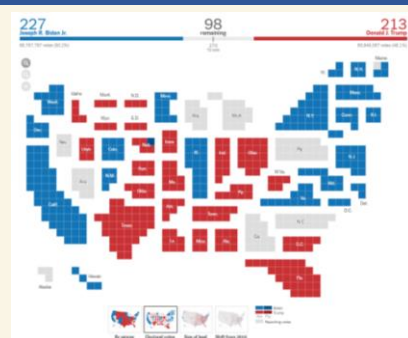
Electoral Vote Cartogram, WSJ



Source: Wall Street Journal online, downloaded 11/4/20, 2pm

47

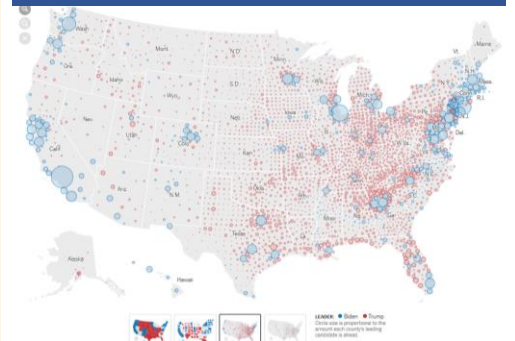
Electoral Vote Cartogram, NYTimes



Source: New York Times online, downloaded 11/4/20, 2pm

48

Popular Vote Winners by County, 11/4/20



Source: New York Times online, downloaded 11/4/20, 2pm

49

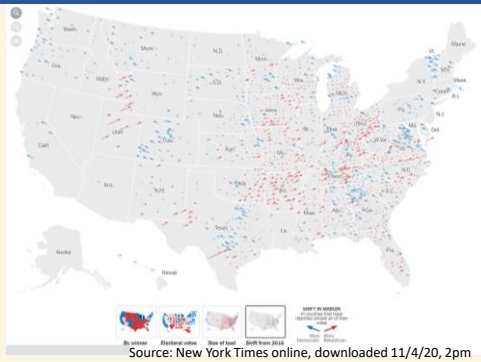
Popular Vote Winners by County, 11/4/20

LEADER: ● Biden ● Trump
Circle size is proportional to the amount each county's leading candidate is ahead.

Source: New York Times online, downloaded 11/4/20, 2pm

50

A Difference Map: 2016 vs. 2020



Source: New York Times online, downloaded 11/4/20, 2pm

51

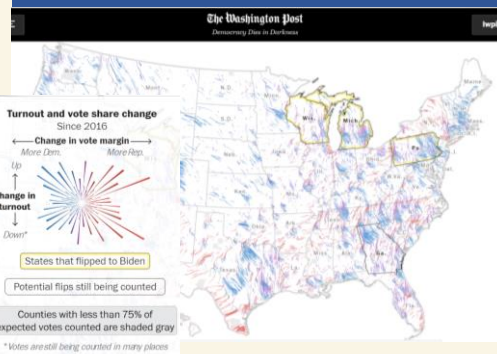
A Difference Map: 2016 vs. 2020 by County

SHIFT IN MARGIN
In counties that have reported almost all of their votes

More Democratic (blue arrow)
More Republican (red arrow)

52

Wash. Post County Difference Map



53

Recap: What do we need to know to correctly interpret data maps?

- What is being mapped?
- What geography is represented & how?
- How are data represented?
- Cognitive biases, limitations in map reading
- Further reading
 - Brewer, *Designing Better Maps: A Guide for GIS Users*, ESRI Press, 2005.
 - Carr & Pickle, *Visualizing Data Patterns with Micromaps*, CRC Press, 2010.
 - Monmonier, *How to Lie with Maps*, 3rd ed., 2018.
 - Howard Wainer – any publications on visualization

54