CSSS 569 · Visualizing Data

#### GALLERY 2: MAPS AS VISUAL DISPLAYS OF INFORMATION

Christopher Adolph

**Department of Political Science** 

and

Center for Statistics and the Social Sciences

University of Washington, Seattle



CENTER for STATISTICS and the SOCIAL SCIENCES

## Gallery

Today's lecture starts with Napoleon's March

In honor of Minard's graphic, our opening gallery surveys modern data maps

## Maps as VDQIs

Three main ways to use maps in VDQIs:

- 1. Match data to space: choropleth maps
- 2. Match space to data: cartograms
- 3. Create a space for data: e.g., treemaps

#### data $\rightarrow$ space: choropleth maps

Code regions of a map with non-spatial data

Advantage: Connects abstract variables to familiar cartography

Disadvantage: Often conflates spatial and non-spatial variation

All variables directly linked to space, not to each other

But space may be a distraction or a confounder

Actual social relationships may be independent of space but not uniformly spatially distributed

## French kissing map



Not as exciting as it sounds

Preferred number of kisses to cheeks in greeting, per survey data

Source: Gilles Debunne, combiendebises.free.fr in Jacobs, *Strange Maps* 

## South Carolina BBQ sauce preferences



Source: Kovacik & Winberyy, *South Carolina: A Geography* in Jacobs, *Strange Maps* Unclear how assessed. Danger of 50% cutoffs hiding ambiguous results in choropleths

## **Didactic maps**



Source: *Superinteresante* magazine, Sao Paulo, in Jacobs, *Strange Maps* GDP of US states explained through international analogs

## **Didactic maps**



Source: *Superinteresante* magazine, Sao Paulo, in Jacobs, *Strange Maps* GDP of African states explained through international analogs

## **Counties voting 65%+ for Clinton over Obama**



Source:

65% vs plurality: The cutoff for coding a region matters

Data from 2008 primaries, but before WV & KY. Map suggests a model & forecast

## 2008 Dem Primary: Winner by County



Source: Wikipedia, "Results of the 2008 Democratic Party presidential primaries"

## 2008 General: Southern Voting



Source: Allen Gathman, cstl-csm.semo.edu/gathman/cottonvote.htm

#### **1860 Cotton Production**



Source: Allen Gathman, cstl-csm.semo.edu/gathman/cottonvote.htm

## 1860 Cotton Production vs. 2008 Presidential Vote



Source: strangemaps.wordpress.com

## 1990 Black Population & 2004 Population Density



Source: Allen Gathman, cstl-csm.semo.edu/gathman/cottonvote.htm Maps hide models

The right model:

Geography  $\rightarrow$  cotton regions  $\rightarrow$  slaves  $\rightarrow$  black voters  $\rightarrow$  Democratic strongholds

Nothing actually surprising here

# 2008 in a Map: Appalachia & Cotton Regions

... while the rest of the nation shifted Democratic.



Source: New York Times

#### Map as reference

Most of the time, maps are a tool for look-up, not gestalt evalutation Interactivity helps with lookup, and can establish a sense of scale & context [Ben Fry zip code]

http://benfry.com/zipdecode/

We'll talk more about interactive graphics as the course progresses, but here's a great example published this week on segregation

[Tomas Monarrez segregation maps]

https://www.vox.com/2018/1/8/16822374/school-segregation-gerrymander-map

A seminal interactive map

[NYT interactive map of the US Census]

http://projects.nytimes.com/census/2010/map

## An iconic map: 2004 Red-blue America



# **Richer version: County level**



#### $Map \rightarrow data: cartograms$

As a political scientist, I look at these maps and see an urban-rural split.

I know not to infer overall electoral support from areas on this map

Many people, though, saw Republican dominance

How do we make this a useful tool for assessing electorate size?

**Cartograms** distort geography to match some variable (e.g., population) while retaining it's distinctive shape

Successful cartograms exist in the eye of the beholder: do you see the old map *and* the variable, or just one?

An old artistic technique, now easy to generate: diffusion cartograms (see Gastner & Newman, 2004 *Proc Nat Acad Sci*, 101:20, 7499-7504, and http://www-personal.umich.edu/mejn/election/)

## State-level population cartogram



# **County-level population cartogram**



#### **Red-blue population cartogram**

Now we see an even mix of Rs & Ds. (Dense islands of the latter in a thin ocean of the former)

Criticism: map still paints areas in binary fashion

Most counties are internal divided (e.g., 55-45 Dem split)

A color scale helps

But the scaling itself is arbitrary

This map shows the same data as the last, but "looks" very different

# Linear Red-purple-blue population cartogram



# **Categorical Red-purple-blue population cartogram**



## Cartograms don't always reflect population



**Fig. 5.** The distribution of news stories by state in the United States. (*a*) Albers conic projection. (*b*) Cartogram in which the sizes of states are proportional to the frequency of their appearance in news stories. States are the same shades in *a* and *b*.

#### And some think cartograms aren't the best solution. . .



Randall Monroe

http://xkcd.com/1939

## Map as analytical tool



Disease clusters. Is (lung) cancer in New York state clustered?

## Map as analytical tool



Or is this an artifact of population clustering?

## Maps of data

Treemaps have ambiguous cartography

Just show data in a useful spatial grouping

Treemap of market status
http://www.smartmoney.com/marketmap

Treemap of news stories http://www.marumushi.com/apps/newsmap/newsmap.cfm

Treemap of average color of webpages containing English nouns http://loop.aiga.org/resources/loop/loop9/colorproject



+ about + permalink	+ Select all count	RIES AUSTRAL AUSTR	XIA CANADA	FRANCE D	EUTSCI	INDIA	italia	NEW ZE	ESPANA	U.K. 🗹 U.S.
In pictu	Russia and Ukraine settle gas dispute		US: Accusation by former Syrian official warrants investigation		New cri cast sha over Palestin poll	ses adow iian	Worker Bonus es Sought At Flvi	SUV's no safer for kids A new ruler in	US Retailers' Sales Risa 2.9% in the Fisal Week of December S S CBS, Viacon trading boosts both firms' shares S	
US MINE			Moscow's proposal to Tehran still stands		Indonesian landslide buries village, 200 feared dead			Dubai Dow Jones CEO to Step Down in	percent of a program of a progr	
trageuy			Fraud probe delays final results of Iraq election	Gummen b road to Ga crossing	block Bolivia's president- elect meets Châvez, vows leftist campaign		Dollar slips as Fed hints it may halt rates rises	Roe woes for caviar lovers	Alde Badari Satur, a working Satur, a working Nintando reports strong DS sales	
Suspected N Arson Fire in C	orthern alifornia	Update 26: B 13th Victim F	Body of Found	Sharon to Transfer Power Temporarily	Egpyt assures UNHCR that Sudanese won't be deported	S.Lanka to probe youth deaths over execution report Arrest may help crack USs chooleuth	Roe woes for caviar lovers	Dueling claims to inspiration for Narnia story	Lohan treated after asthma attack	Phoenix returns to Johnny Cash's prison stage Pyan Searc family murdered set ablaze Pyan stage Searc family family murdered set ablaze set ablaze
Contained Clean Up		at German Rink		Australia plane crash Britons named	France to call off state of emergen cy today	Police Pres Ada Chierto Estada: Fajmut la Faca e Reage et Cherges	Sharnerd Sauth Court scientifit courts of weatwork for eggs - TV	Dick Clark's Rockin'	Ancho rs' array on ABC	Deal With Est Version of Version
Politicians take Cover as Lobby ist Updat Bush directive shocks local		Penn State	Slaton, Mountain eers make	Al Davis' revisionis of the Rai e	is' Hingis startii nist history to look Raiders ominous		starting s	Chair Eve ID as science is now	Oh, yes you cai	SUV's no safer for kids
prepares to talk	ents Downed flier: <sup>2</sup> Snowmobilens Killed in Colo. Avalanche Avalanche	star hurt	statement for Big East	New coach promises big	Passan: The Power of the Fans	Marquette's Novak outshines all of UConn's stars	STEIN LINE EDITION: An exiled King's ratum	history Microsoft Rov to Patch ent Windows sec	ers Mot ar ola	U Study says statists fail in lower cancer risk degalizes medical marijua
Iraq Notebook Teenager activist?	ntzlligen Grei kej Lane. e Panel chemistrichen beiten 27 chemistrichen 27 c	Football Notes: Clarett posts	MIKE DUFFY'S TV TODAY:	for Sonics	Wilson signs one -year deal with	Billups inspires Pistons to victory over Magic	Rockets break out in big way	Flaw Next yea Week Mai	ron unv s iRau	dio Study: Cheerleading- related injuries on
reneges on talking about trip	Mana Anadh and hard and gang warning Solder Weighand in bag	bond on 2 robbery charges	Longhorns, Trojans promise grea action in	NHL Roundup: Crosby leads Penguins to victory	Astros Cheisea kity to peace offering at Highbury	No help for Henrik Elias' return ignites Devils	hrushin harpina The State The State The State Has an and the state of the state of the state of the state	Inside Nodel More Intel	Ante Ante Ante Ante Ante Ante Ante Ante	All Cires All Cires
Thursday January 5, 2006 (			+ SELECT A	LL CATEGO	RIES		LAYOL	IT: 🔳 SQU	arified 📃 standard	
ARCHIVED SRT SUN MON TUE WED YE	ST. TODRY HOW	LESS THAT	n 10 minutes ago	WORLD	- HATION	P BUS	INESS 🔽 1	TECHHOLI 🔽 SF	ORTS	ENTERTAI 🛩 HEALTH
00:00		More That	n 10 minutes ago							

#### COLOR CODE: <u>Home | FAQ | Image gallery | Contact</u>

Show All	33096 English nouns arranged by m	eaning. Point and click on th	e mosaic.			
Arrangement: Meaning Color		BODY PART	ECOD	DISH	MINERAL	HAPPENING
M Show Labels All starting with:	CHANGE ACTIVITY	DRUG REGION				
	CONTENT	PLACENTAL	HERB	SHRUB	NATURAL OBJECT	DISEASE
	SYMPTOM	BIRD	FLOWER	TREE	STRUCTURE	
	MESSAGE MEASURE			WORKER	DEVICE	PROCESS
	QUALITY PROPERTY-					

## Treemaps an effective tool?

Do these displays work? Or are they just toys?

- Detailed lookup is hard
- What is the clustering based on?

Would Tufte like these plots?

#### Treemaps an effective tool?

Do these displays work? Or are they just toys?

- Detailed lookup is hard
- What is the clustering based on?

Would Tufte like these plots? Probably not

Wasted dimension; arbitrary arrangement; unclear what space actually means

Fun for data exploration, though.

But they can be used to make an argument, as well, by varying the underlying space

## Maps as an analytical tool



*Market peak* ... Sectoral composition of the Wilshire 5000 on March 24, 2000 (record high). Source: New York Times.

#### Maps as an analytical tool



... *and collapse.* Sectoral composition of the Wilshire 5000 on July 18, 2002. Source: New York Times.

The sudden appearance of an epidemic typically inspires rapt attention, panic and action. Once the crisis subsides, public attention wanes although the threat of contagion continues, especially among the world's poor. Compare our response to severe acute respiratory syndrome, or SARS, with the more familiar germs that plague us daily. Compare it to the dangers of smoking or getting in a car and heading out on the road. Every life is precious, but when you look at the numbers, SARS just isn't as formidable a threat as we've made it out to be. Its death rate is far lower than that for AIDS or malaria: coronaviruses. like the one believed to cause SARS. tend to be most active in the winter and early spring. In addition to taking the steps necessary to keep SARS at bay-watching out for new cases and isolating people who are contagious to others-we would do well to channel our energies into something more lasting: a permanent. integrated and accountable global public health system for the surveillance and prevention of the microbes that are certain to emerge in the future. Right now, worldwide accounting of disease is incomplete at best, hampered in large measure by sketchy reporting from developing countries. These gaps slowed our containment of SARS and allowed rumor to spread more rapidly than reliable information. When the facts are few, it's easy for fear to fill the vacuum. Howard Markel. professor of pediatrics and communicable diseases at the University of Michigan, is author of the forthcoming "When Germs Travel."



#### The Epidemic Scorecard

(Source: Howard Markel & Stephan Doyle, "The Epidemic Scorecard," *New York Times*, April 30, 2003.)

A good idea, horribly executed

Massive lie factor: compare the area of rectangles for AIDS and TB

Or SARS and anything: SARS still far too big (My first reaction: "Wow! SARS causes  $\sim 10\%$  as many deaths as AIDS!")