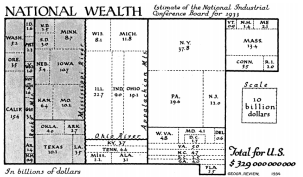
State of the Art in Cartograms

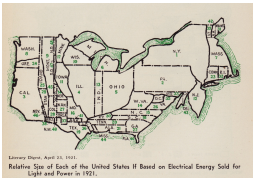
Sabrina Nusrat and Stephen Kobourov
University of Arizona

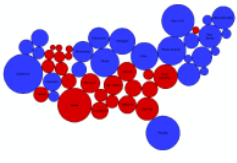
artist: Matt Cusick

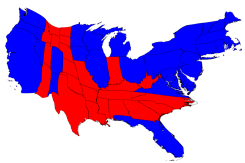
State of the Art in Cartograms

- Cartograms: combining statistics and geography for over 150 years!

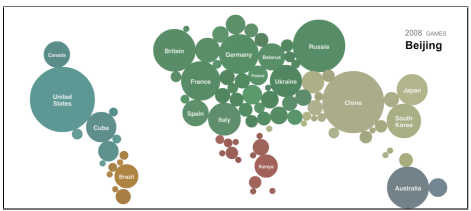


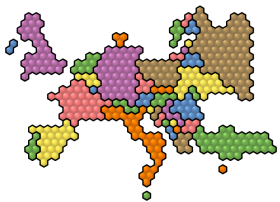


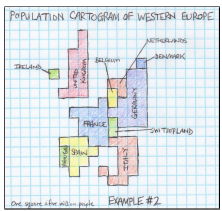




- Cartogram dimensions (statistics, geography, topology) and evaluations (quant, qual)







Cartograms: Definition

Definition

Cartograms, or value-by-area diagrams, are thematic maps, that combine statistical and geographical information, where areas of geographical regions (e.g., states) are scaled in proportion to some statistic (e.g., population)

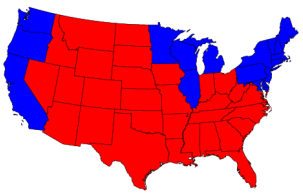
Cartograms: Difficult by Definition

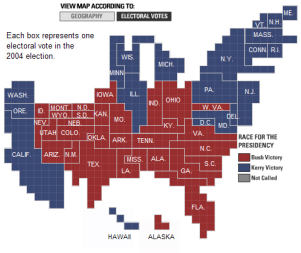
Geo-Statistic Visualization

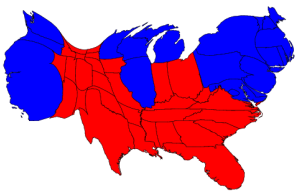
Cartograms attempt to *simultaneously* capture statistical and geographical information in order to provide insight into geo-statistical patterns.

Cartograms: Examples

- Red-state, blue-state US Presidential elections 2004
- Geographic map and two population cartograms: rectilinear and diffusion
- Seem to tell different stories!

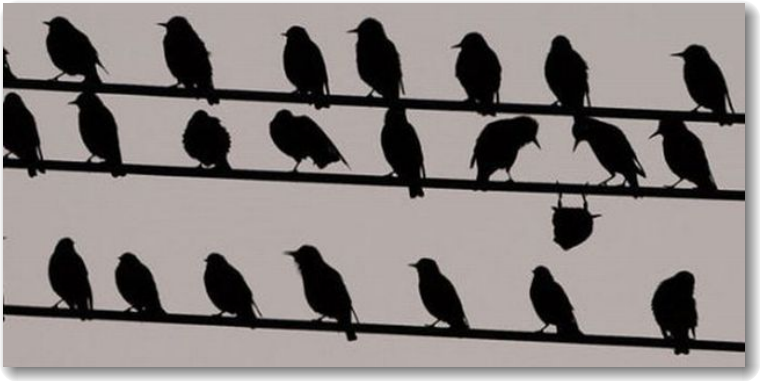






Cartogram Variants

- Many different variants over 150 years: rectangular, circular, diffusion,...
- Some more different that others



Cartogram Variants

- Popular: 506,000 google hits for “cartogram”

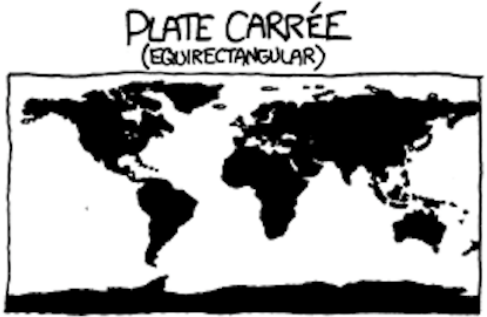
"Reality is Frequently Inaccurate"

Map Projections (xkcd)

Map Projections (xkcd)

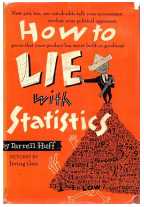
Map Projections (xkcd)

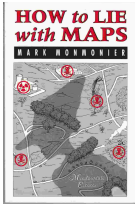
Map Projections (xkcd)

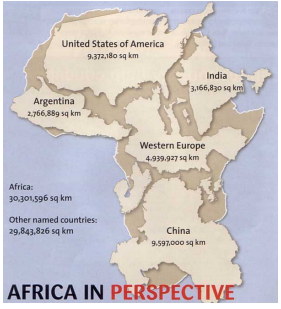


Deceptiveness

- How to lie with statistics
- How to lie with maps
- Statistics and maps: even more deception

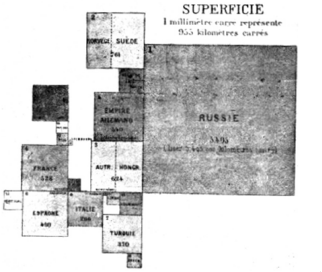






Cartogram History

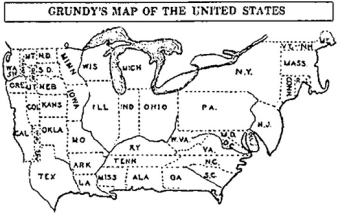
- Date back to the 19th century
- Many different variants from the beginning



Levasseur 1866

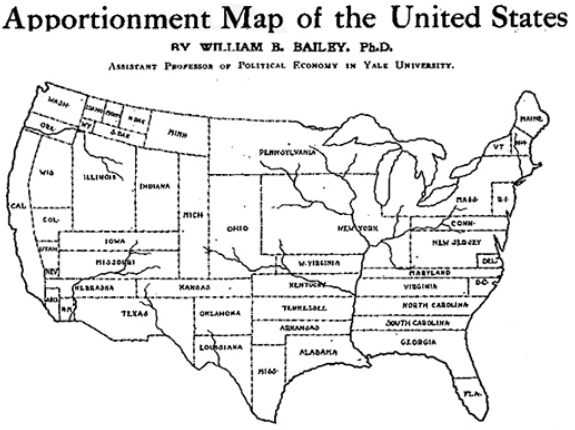
Fig. 24.5 Fragment of statistical diagrams in the Rand McNally World Atlas of 1897

Rand McNally 1897



Grundy 1929

A Cartogram from 1911



"The map shown on this page is drawn on the principle that the population is evenly distributed throughout the whole United States, and that the area of the States varies directly with their population."

Rectangular Cartograms

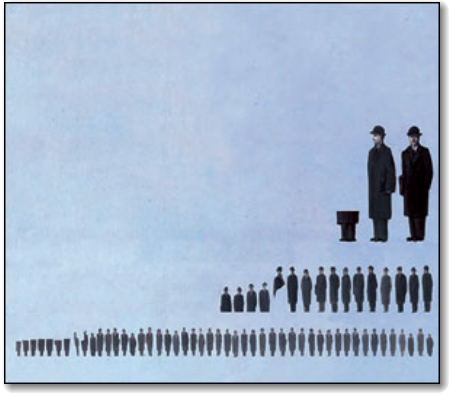
The desire for order:

René Magritte, Golconde, 1953

Rectangular Cartograms

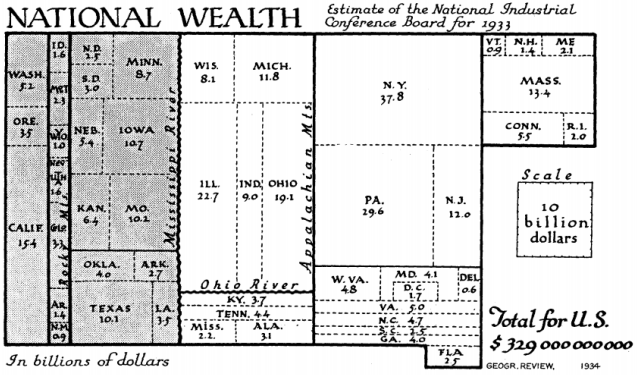
The desire for order:

René Magritte, Golconde, 1953



Ursus Wehrli, Tidying Up Art, 2008

Rectangular Cartograms: Raisz 1934



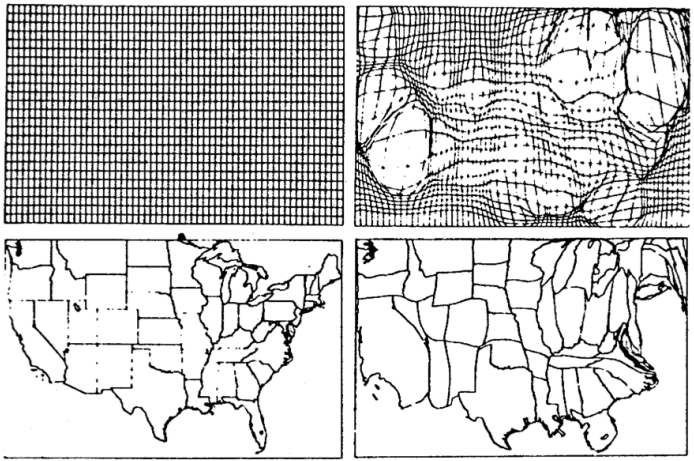
"The statistical cartogram is not a map. Although it has roughly the proportions of the country and retains as far as possible the relative locations of the various regions, the cartogram is purely a geometrical design to visualize certain statistical facts and to work out certain problems of distribution."

Making Cartograms

- Initially, more art than science (GM, 1938): geographers, economists, ...

Making Cartograms

- Initially, more art than science (GM, 1938): geographers, economists, ...
- Computer programs, e.g., Tobler's rubber sheet method, 1973

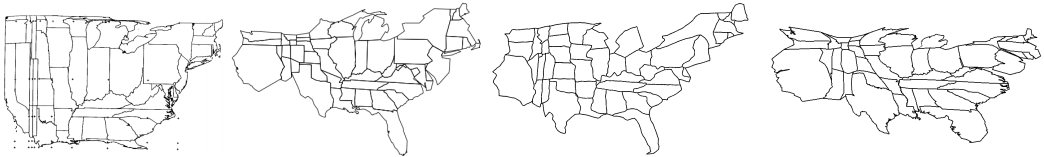


Making Cartograms

Cartogram-generation methods

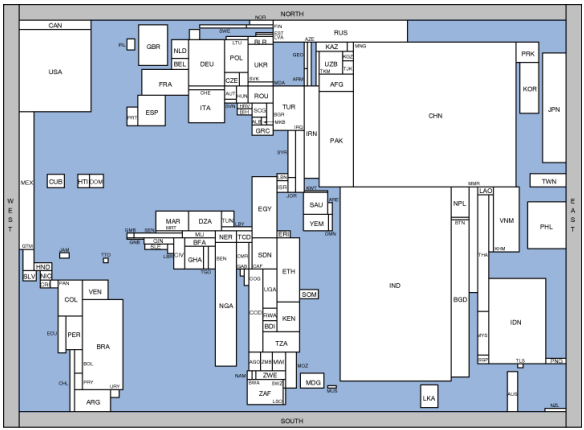
Given a geographic map (e.g., USA map) and some statistic (e.g., population), create a map that **simultaneously** captures the geography and the statistic.

- Looks easy
- Theory vs practice
- Every problem has a solution ...



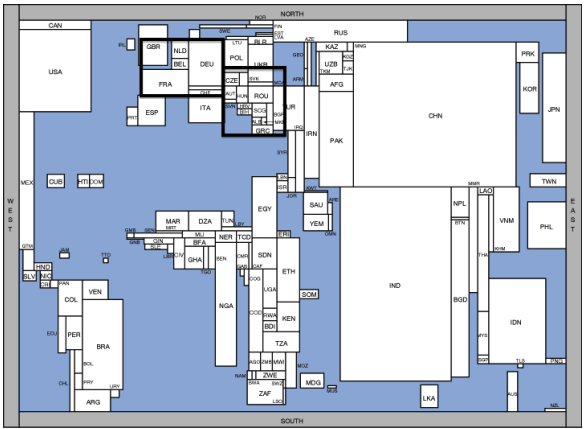
Rectangular Cartograms: Surely Easy

Consider this population cartogram of the world [Spe06]



Rectangular Cartograms: Surely Easy

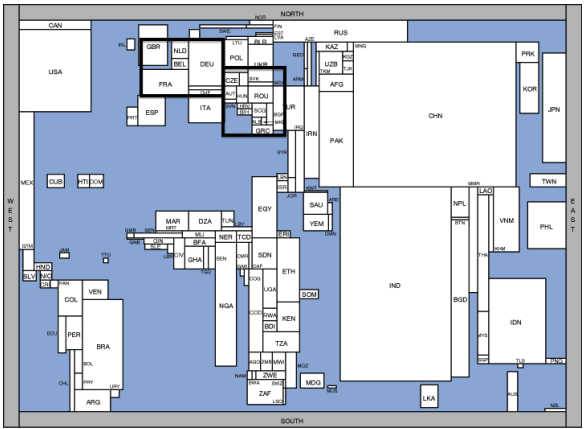
Consider this population cartogram of the world [Spe06]

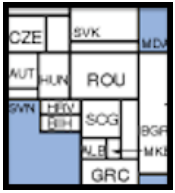


- missing neighbors: Serbia and Hungary

Rectangular Cartograms: Surely Easy

Consider this population cartogram of the world [Spe06]

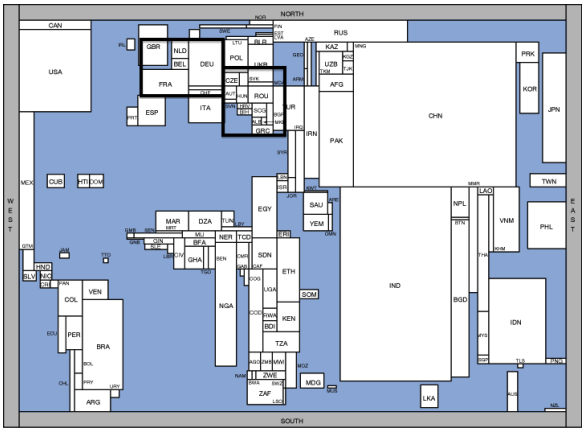


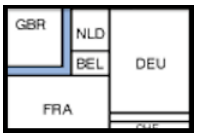


- missing neighbors: Serbia and Hungary
- extra neighbors: Moldova and Slovakia

Rectangular Cartograms: Surely Easy

Consider this population cartogram of the world [Spe06]

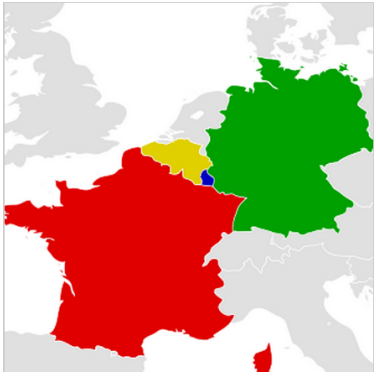




- missing neighbors: Serbia and Hungary
- extra neighbors: Moldova and Slovakia
- missing countries: Luxembourg

Rectangular Cartograms: Actually Impossible

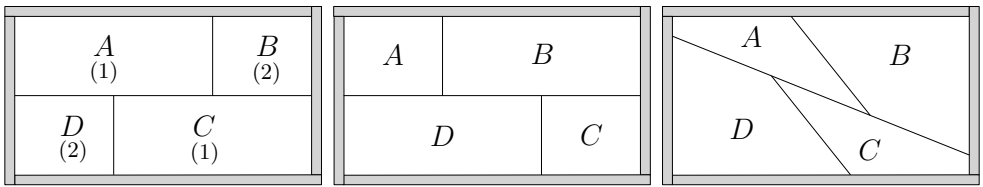
Consider these 4 countries



- It is not possible to represent them as rectangles and keep correct neighbors
- But it makes for a fun exercise

Cartograms: How Hard Can It Be?

Three cartograms for a map with an area assignment for four states A, B, C, D with desired areas 1, 2, 1, 2



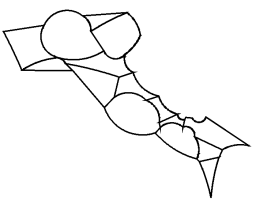
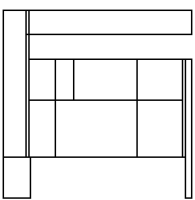
- the first one has **statistical errors**
- the second one has **topological errors**
- the third one has **geographical (shape) errors**

Worse yet, no cartogram exists without one of these errors!

Major Cartogram Design Dimensions

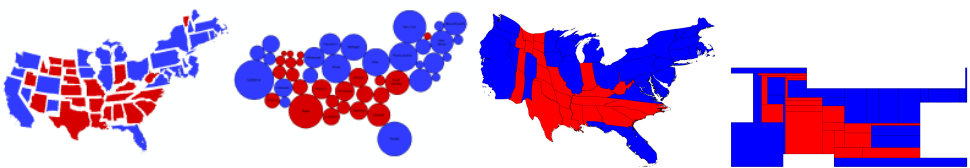
- Statistical accuracy: areas match given statistic
- Topology accuracy: neighbor relations preserved
- Geographical accuracy: region shapes and relative positions preserved



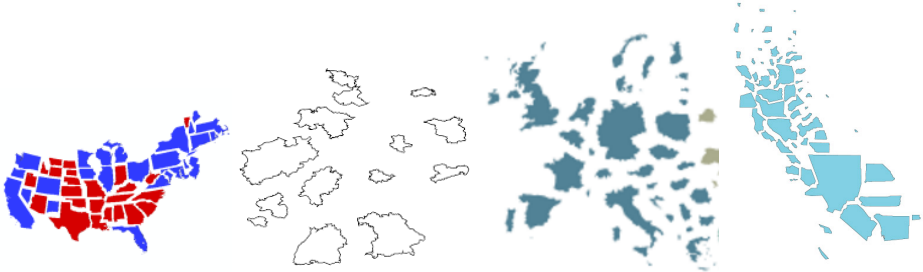
Major Cartogram Types

- Non-contiguous cartograms, e.g., Olson [O76]
- Dorling cartograms [Dor96]
- Contiguous cartograms, e.g., diffusion map [GN04]
- Rectangular cartograms, e.g., [vKS07]



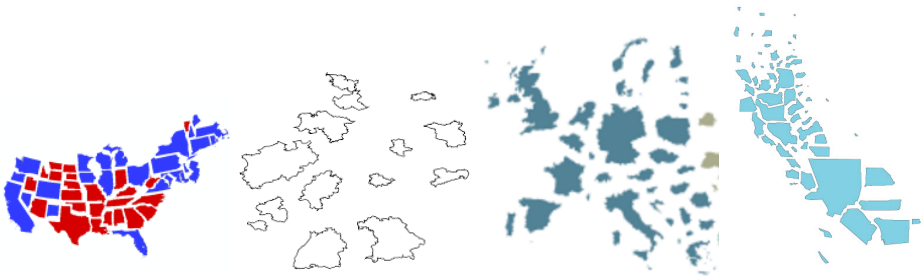
Non-Contiguous Cartograms

- Start with geographic map and scale down each region independently
- Some variants reposition the regions for better space utilization



Non-Contiguous Cartograms

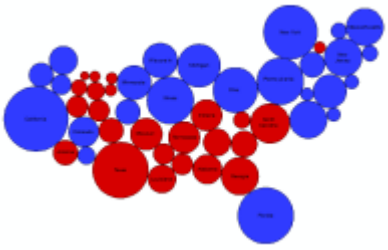
- Start with geographic map and scale down each region independently
- Some variants reposition the regions for better space utilization

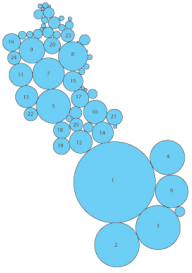


- high statistical accuracy
- mixed geographical accuracy
- low topological accuracy

Dorling Cartograms

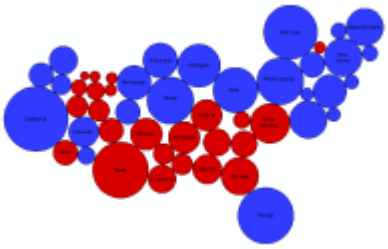
- Start with geographic map and replace each region by a circle
- Center each circle at the regional center and give area proportional to statistic
- Some variants reposition the regions for better space utilization

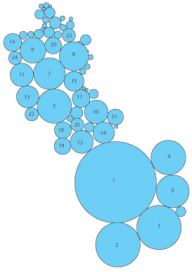




Dorling Cartograms

- Start with geographic map and replace each region by a circle
- Center each circle at the regional center and give area proportional to statistic
- Some variants reposition the regions for better space utilization





- high statistical accuracy
- low geographical accuracy
- low topological accuracy

Contiguous Cartograms

- Also known as continuous, deformation, diffusion cartograms
- The most popular algorithm of this type is the diffusion one [GN04]
- the original input map is projected onto a distorted grid
- or a region-to-region flow is modeled, until equal density is achieved

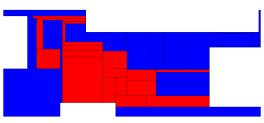
Contiguous Cartograms

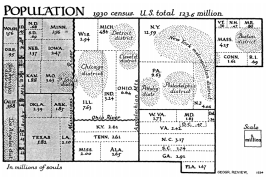
- Also known as continuous, deformation, diffusion cartograms
- The most popular algorithm of this type is the diffusion one [GN04]
- the original input map is projected onto a distorted grid
- or a region-to-region flow is modeled, until equal density is achieved

- high statistical accuracy
- variable geographical accuracy
- high topological accuracy

Rectangular Cartogram

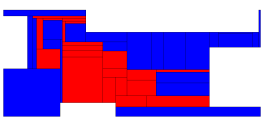
- Take the dual of the map and represent the resulting graph by rectangles
- Restricted to 3-connected graphs
- To make it work for all maps, L-shaped regions are required (Luxembourg)

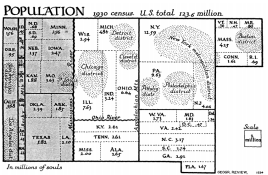




Rectangular Cartogram

- Take the dual of the map and represent the resulting graph by rectangles
- Restricted to 3-connected graphs
- To make it work for all maps, L-shaped regions are required (Luxembourg)

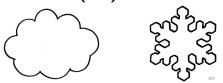




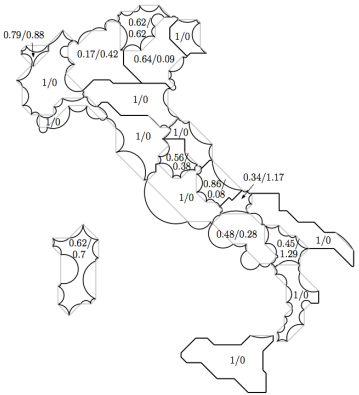
- variable statistical accuracy
- low geographical accuracy
- variable topological accuracy

Other Types: Circular-arc Cartograms

- Kämper et al., PacificVis'13
- shape shows relative change
- cloud (+) and snowflake (-)

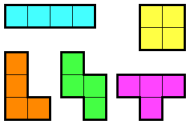


- low statistical accuracy
- high geographical accuracy
- high topological accuracy

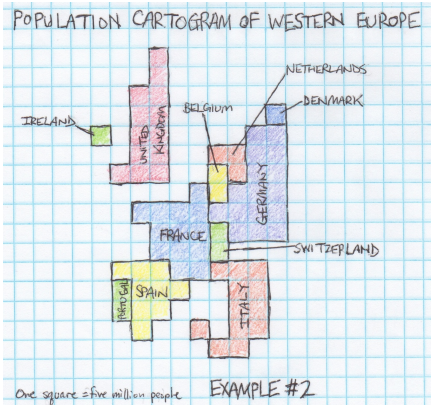


Other Types: Rectilinear Cartograms

- de Berg et al. ACM GIS'10
- more flexible than rectangular
- variants on shape complexity



- high statistical accuracy
- low geographical accuracy
- high topological accuracy

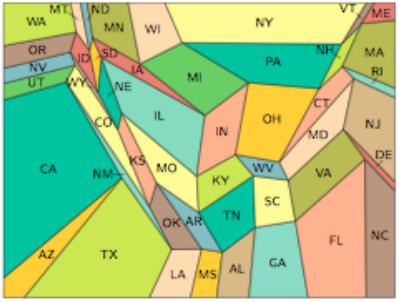


Other Types: Table Cartograms

- Evans et al. ESA'13
- first make a grid
- then modify cell areas

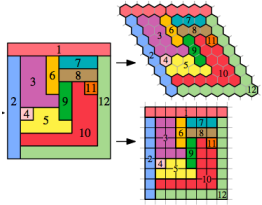
ID	ND	MN	IL	OH	VT	ME	RI
WA	MT	SD	WI	MI	NY	NH	MA
OR	WY	NE	IA	IN	PA	NJ	CT
NV	CO	KS	MO	KY	WV	MD	DE
CA	UT	OK	AR	AL	TN	NC	VA
AZ	NM	TX	LA	MS	FL	GA	SC

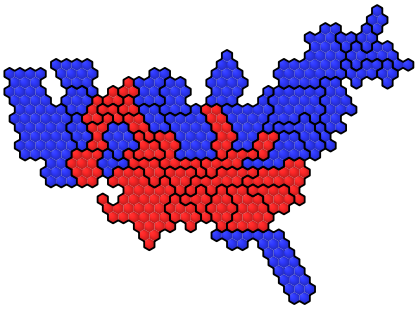
- high statistical accuracy
- low geographical accuracy
- low topological accuracy



Other Types: Mosaic Cartograms

- Cano et al., EuroVis'15
- tiles: squares, hexagons
- start with rectilinear

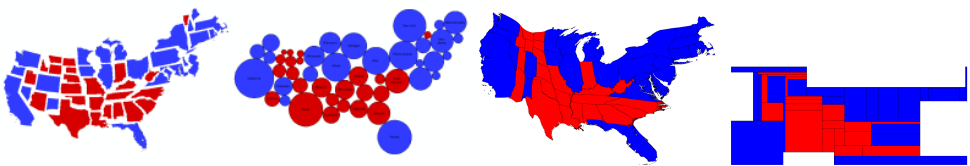




- variable statistical accuracy
- variable geographical accuracy
- variable topological accuracy

So, Which One is Best?

- Huge variety of cartogram types and subtypes
- Question: which type should we use?
- Answer: it depends...
- Quantitative evaluations, task taxonomy, qualitative evaluations



Quantitative Cartogram Evaluation

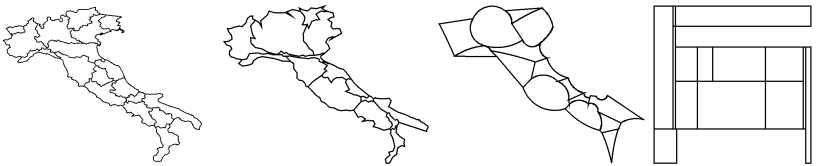
- Many ad hoc quantitative measures in the literature
- A set of **seven standard** measures, Alam et al., EuroVis'15

Criteria	Measure & Notation	
Statistical Accuracy	Average Cartographic Error	ε
	Maximum Cartographic Error	ξ
Topological Accuracy	Adjacency Error	τ
Geographical Accuracy	Angular Orientation Error	θ
	Hamming Distance	δ
	Average Aspect Ratio	α
Complexity	Polygonal Complexity	η

Statistical Accuracy

- How well is the statistic captured by the areas of the regions?
- Individual cartographic error (for each region): $|o(v) - w(v)|$
  - $o(v)$: obtained area in cartogram
  - $w(v)$: required area
- Average and maximum cartographic error in the [0,1] range (for the entire map)

$$\varepsilon = \frac{1}{|V|} \sum_{v \in V} \frac{|o(v) - w(v)|}{max\{o(v), w(v)\}} \qquad \xi = \max_{v \in V} \frac{|o(v) - w(v)|}{max\{o(v), w(v)\}}$$



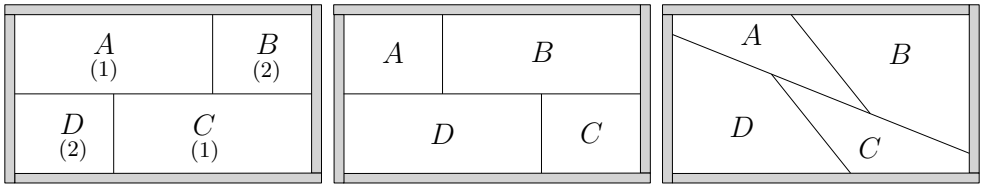
$\xi = 0.25$ $\xi = 0.70$ $\xi = 0.0$

Topological Accuracy

- How well are adjacencies b/n pairs of neighbors preserved?
- Topological error in the [0,1] range:

$$\tau = 1 - \frac{|E_c \cap E_m|}{|E_c \cup E_m|}$$

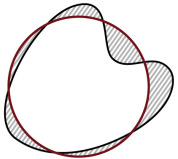
  - E_m: adjacencies between countries in the map
  - E_c: adjacencies between countries in the cartogram



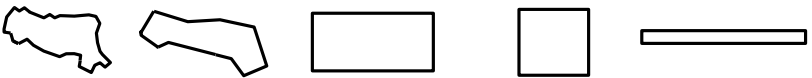
$\tau = 0$ $\tau = 0.33$ $\tau = 0$

Geographical Accuracy: Shapes

- How well are country shapes preserved?
- Shape distortion measure:
  - *Hamming distance*, δ
  - aka symmetric difference

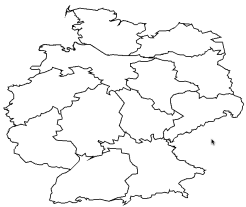


- Definition:
  - pair of polygons are superimposed
  - measure fraction of area that is in exactly one of the polygons
  - normalize polygon areas to unit area (scale invariance)
  - over all possible translations (translation invariance)
  - use the one that gives the smallest error

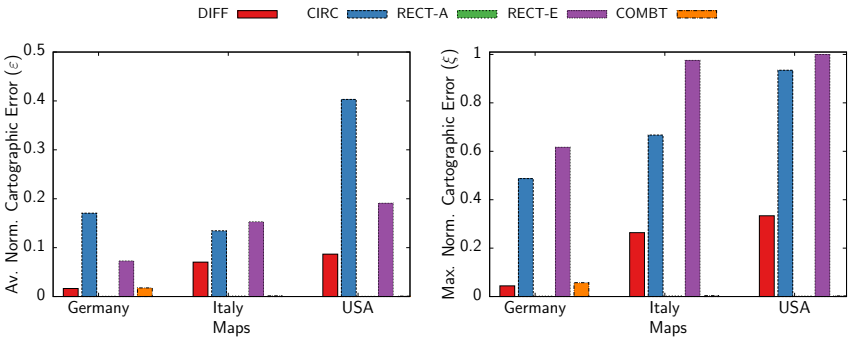


Quantitative Experiment

- Compare 5 different types of cartograms:
  - Diffusion [GN04]
  - Circular-arc [KKN13]
  - Rectilinear-A [BSV12]
  - Rectilinear-E [BSV12]
  - Combinatorial T-shape [ABF13]
- Experimental data:
  - 3 countries (USA, Germany, Italy)
  - 2 statistics (GDP, population)

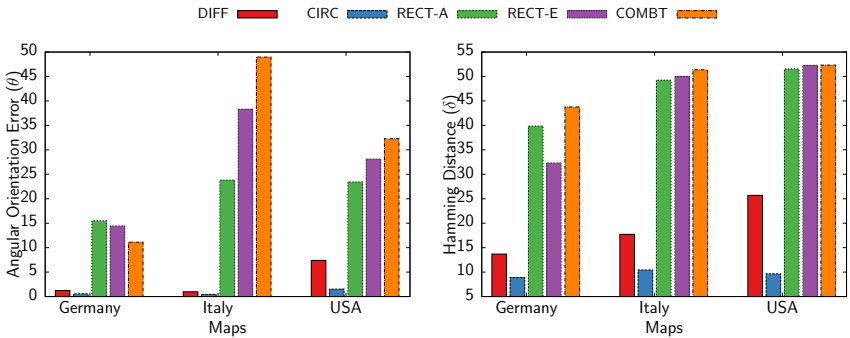


Results: Statistical Accuracy



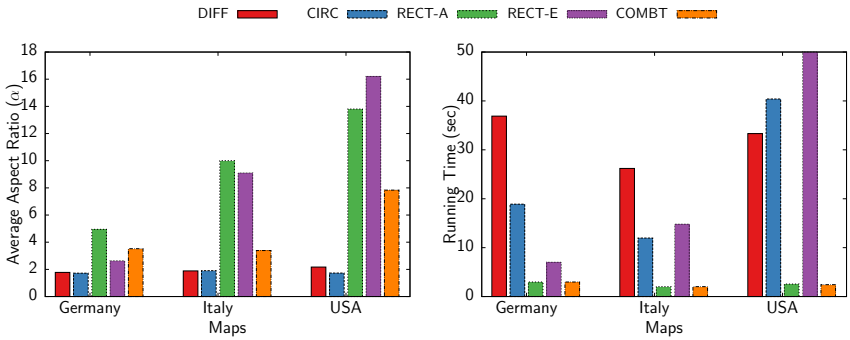
- Rectilinear and Dorling cartograms have low cartographic errors
- But they have high topological errors

Results: Geographical Accuracy



- Diffusion and circular-arc cartograms have good geographical accuracy
- But they have high statistical errors

Results: Topological Accuracy



- Diffusion, circular-arc and rectilinear cartograms have high topological accuracy
- But they are computationally expensive

Summary of Quantitative Evaluation

Cartogram type	Topological accuracy	$O(1)$-Complexity	Statistical accuracy	Geographical accuracy	Run time
Diffusion	Y	N	Y	V	S
Circular-arc	Y	N	N	H	S
Rectilinear	N	Y	Y	L	F
Dorling	N	Y	Y	L	F

- Green is good: Yes (Y), Fast (F), High (H)
- Red is bad: No (N), Slow (S), Low (L)
- Yellow is Medium (M) or Variable (V)

Cartogram Task Taxonomy

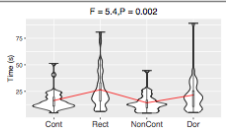
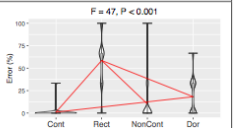
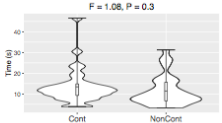
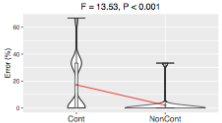
Tasks categorized by goals, means, characteristics and cardinality, based on the questions why, how, what, and where [NK15].

	Goals			Means				Characteristics		Cardinality		
	Query	Search	Extract	Map Relation	Data Relation	Navigation	Derive	Low Level	High-Level	Single	Multiple	All
Recognize	√	×	×	√	×	×	×	√	×	√	×	×
Detect Change	√	×	×	√	×	×	×	√	×	√	×	×
Compare	√	×	×	×	√	×	×	√	×	×	√	×
Find top-k	×	√	×	×	√	×	×	×	√	×	×	√
Filter	×	√	×	×	√	×	×	×	√	×	×	√
Cluster	×	√	×	×	√	×	×	×	√	×	×	√
Locate	×	√	×	×	×	√	×	√	×	×	×	√
Find Adjacency	×	√	×	×	×	√	×	×	√	×	√	×
Summarize	×	×	√	×	×	×	√	×	√	×	×	√
Identify	×	×	√	×	×	×	√	√	×	√	×	×

- Detect change: *Given a population cartogram of the USA, determine if the state of California has grown or shrunk.*
- Summarize: *Given a red-blue election results cartogram, determine whether it was a close election, or a "landslide win."*

Task-based Evaluations

- First cartogram effectiveness tests [Dent 1975]:
  "attitudes point out that these (value-by-area) cartograms are thought to be confusing and difficult to read; at the same time they appear interesting, generalized, innovative, unusual, and having - as opposed to lacking - style"
- Formal task-based evaluation of 4 major types, Nusrat et al., [NAK15]

		Input	Question	Time (s)	Error %
H1	*Locate*	An original undistorted map is given and a state is highlighted in red. A cartogram of the map is shown.	Locate this state in the cartogram.		
H2	*Recognize*	A state from the map of a country, and shapes of three states from the cartogram of that country are shown.	Find out which cartogram state corresponds to the state from the original map.		

Evaluation Results

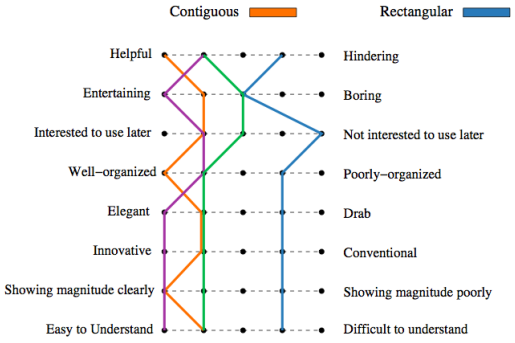
- Alam et al. [ANK'15] task-based evaluation, analyzing time and accuracy
- 33 participants, 4 different cartogram types, covering spectrum of tasks

Algorithm	locate	recognize	compare	detect change	find adjacency	summarize
Contiguous	H	V	H	M	H	V
Non-contig.	H	H	L	H	L	H
Rectangular	L	L	L	L	M	M
Dorling	L	L	H	M	L	H

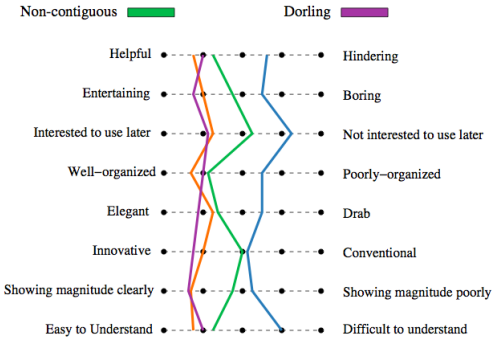
- Green is good: High (H)
- Red is bad: Low (L)
- Yellow is Medium (M) or Variable (V)

Subjective Preferences

- Recall Dent's comment: *"cartograms ... appear interesting, generalized, innovative, unusual, and having - as opposed to lacking - style"*
- Alam et al. [ANK'15] subjective preferences (using Likert 1-5 scale)
- same 33 participants after working with 4 different cartogram types



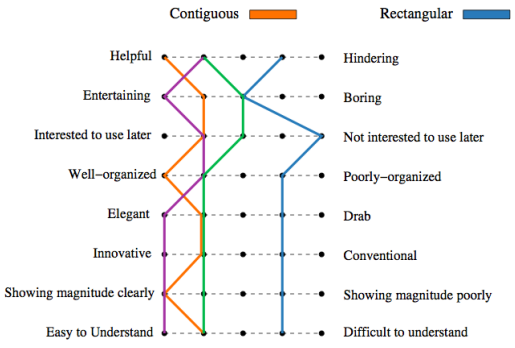
mode



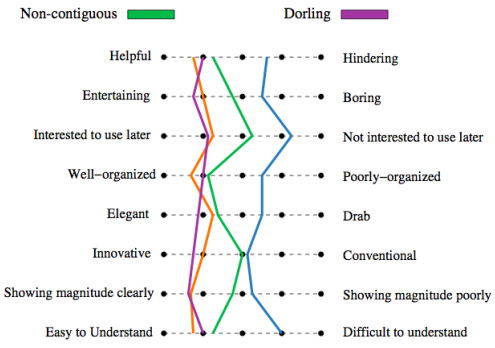
mean

So, Which One is Best?

- There is no clear winner
- Advantages and disadvantages for all
- Subjectively, contiguous and Dorling cartograms are preferred



mode



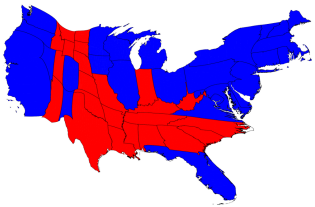
mean

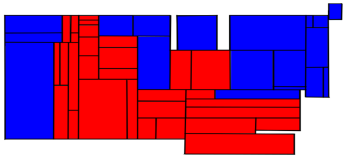
Cartogram Criticism: Difficult to Interpret

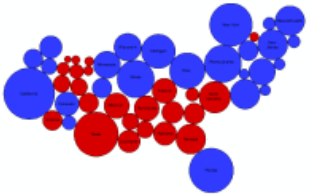
- *"A frequent criticism of cartograms is that even cartograms based upon the same variable for the same areas of a country can look very different" [Dor96]*
- *"These maps are open to potential criticism when it comes to their informative value... the variation of the depicted topic within the territorial borders is not taken into consideration." [HDR09]*
- *"can be hard to interpret without additional information to help the user locate towns and cities." [FBC00]*

Cartogram Criticism: Area Perception

- Humans perceive length with minimal bias, but underestimate differences in area [Ste57]
- The accuracy of area judgment for rectangles and circles is similar, but both are worse than length judgments [HB10]
- There is a stronger correlation between actual area and apparent area for irregular polygons than for circles [Teg65]

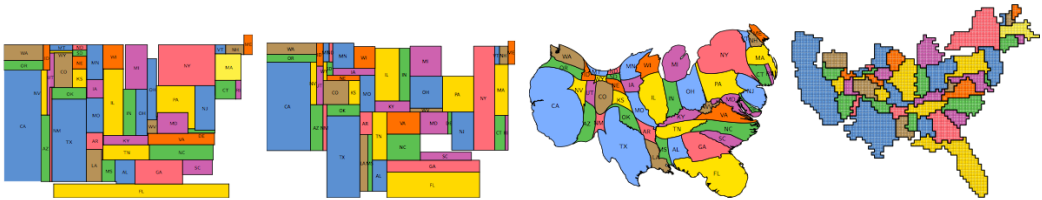






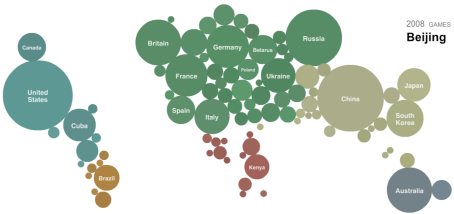
Cartogram Criticism: Shape Distortion

- *"Topology preservation at the expense of shape: even if I know what a county looks like on a normal map, I'm going to have a hard time identifying it here... The bottom line is that many – perhaps even most – cartograms are essentially used for shock value..." [Woo08]*
- *"A cartogram cannot work if people cannot recognize the geography. It no longer surprises/shocks/intrigues if we can't figure out where anything is and how much larger/smaller a place is than we expect. There's definitely a balance that needs to be struck. Good cartograms are still uncommon, I think." [Duf13]*



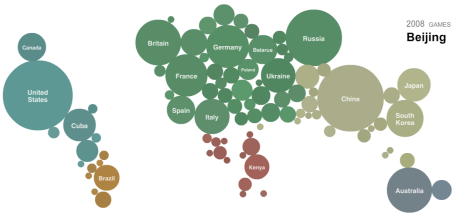
Cartogram Applications

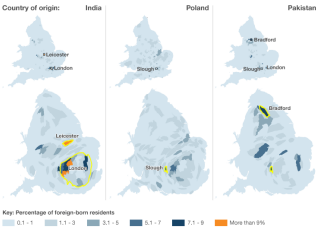
- Olympic games (NYT)



Cartogram Applications

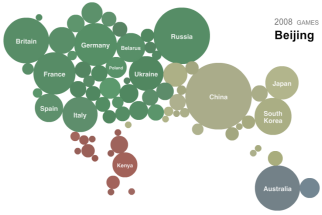
- Olympic games (NYT)
- Social structure (Guar)

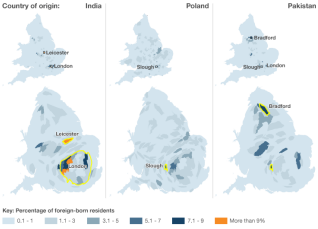


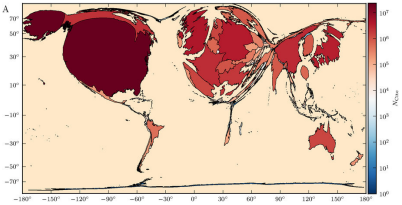


Cartogram Applications

- Olympic games (NYT)
- Social structure (Guar)
- Collaboration networks

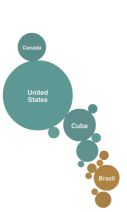


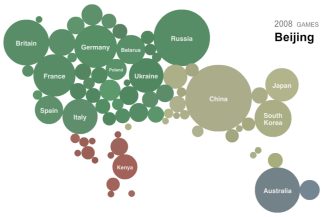


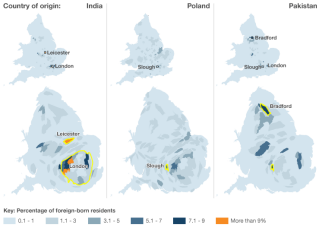


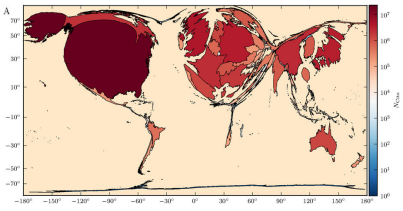
Cartogram Applications

- Olympic games (NYT)
- Social structure (Guar)
- Collaboration networks
- TED talks



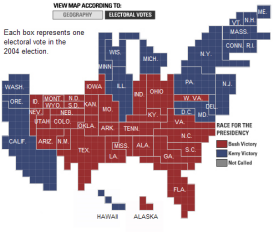






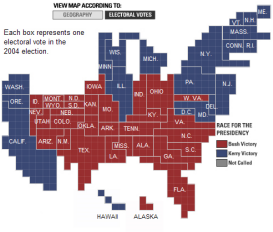
Election Cartograms

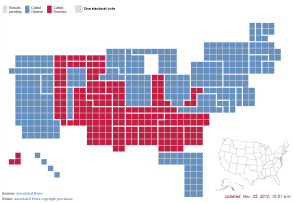
- US President (NYT)



Election Cartograms

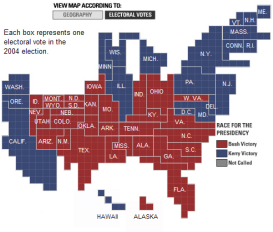
- US President (NYT)
- US President (LAT)

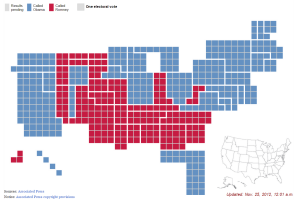


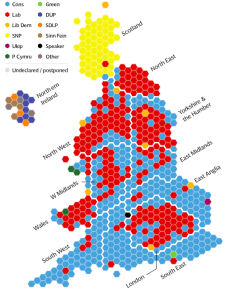


Election Cartograms

- US President (NYT)
- US President (LAT)
- UK Elections (Tel)



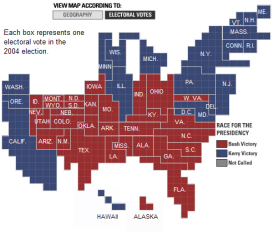


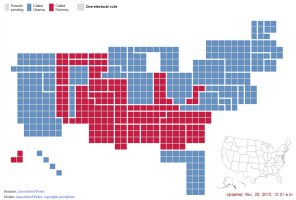


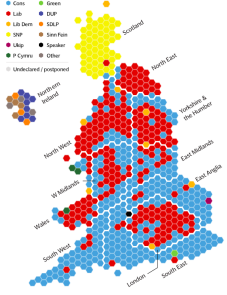
Stephen Kobourov, University of Arizona

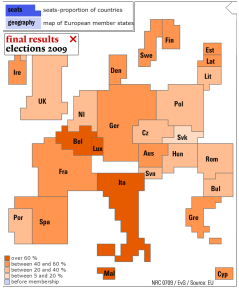
Election Cartograms

- US President (NYT)
- US President (LAT)
- UK Elections (Tel)
- EU Elections (NRC)



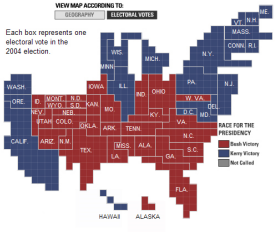


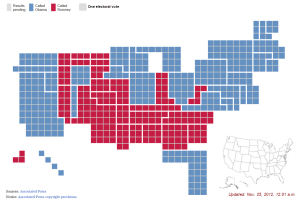


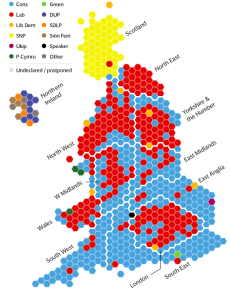


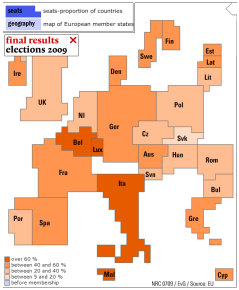
Election Cartograms

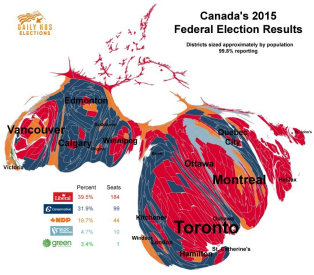
- US President (NYT)
- US President (LAT)
- UK Elections (Tel)
- EU Elections (NRC)
- Canada Elections (DK)







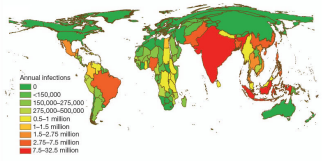




Stephen Kobourov, University of Arizona

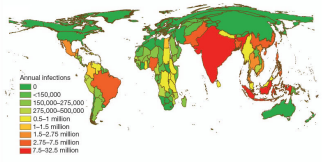
Epidemiology Cartograms

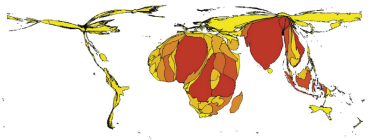
- Dengue



Epidemiology Cartograms

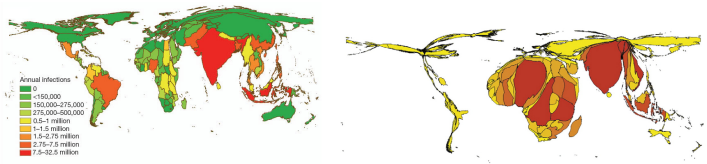
- Dengue
- Malaria

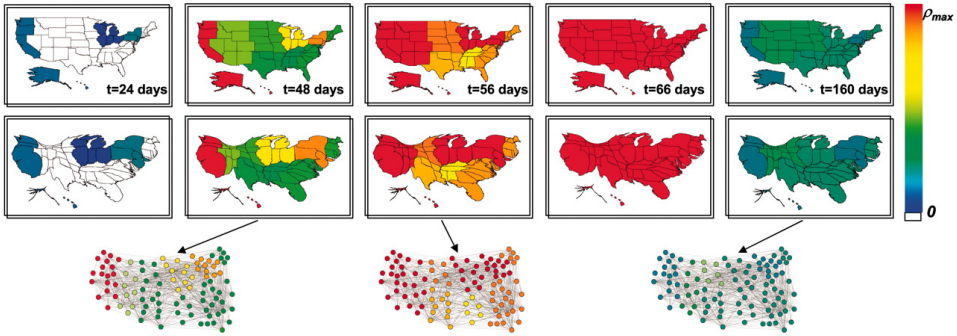




Epidemiology Cartograms

- Dengue
- Malaria
- Disease Spread

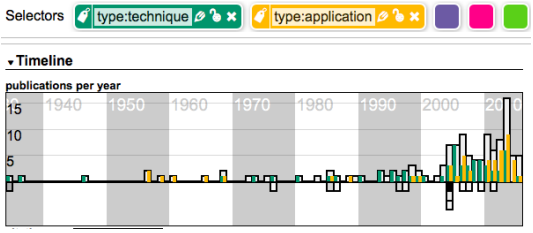




Bibliographic Analysis

- Web-based literature browser, SurVis [BKW16]
- Populated with the usual suspects (InfoVis, EuroVis, ...)

Paper	Year	Cartogram type	Citations
Gastner et al. [GN04]	2004	Contiguous	411
Dorling [Dor96]	1996	Dorling	161
Dougenik et al. [DCN85]	1985	Contiguous	122
Raisz [Rai34]	1934	Rectangular	119
Keim et al. [KNP04]	2004	Contiguous	99
van Kreveld et al. [vKS07]	2007	Rectangular	99



Future Work

- Addressing cartogram limitations
- Interaction
- Dynamic cartograms
- Appropriate labeling
- Evaluation of utility
- Multivariate cartograms
- Uncertainty in cartograms
- Engagement and enjoyment
- Memorability and recall

Conclusions

William Playfair, Statistical Breviary, 1801

Making an appeal to the eye when proportion and magnitude are concerned is the best and readiest method of conveying a distinct idea.

- 150 years of cartograms and still active and growing research area!
- Excellent early surveys by Kocmoud '97 and Tobler '04
- This paper @ `http://arxiv.org/abs/1605.08485`

artist: Matt Cusick