Lecture 7/8:
Design & Justification Exercises, Beyond R

Tamara Munzner
Department of Computer Science
University of British Columbia

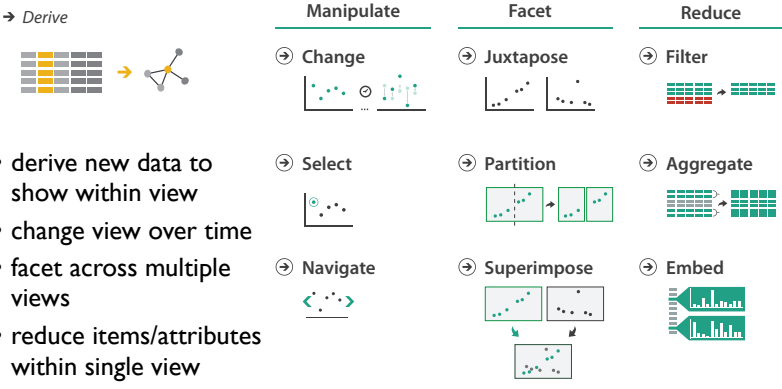*DSCI 532, Data Visualization 2*
Week 4, Jan 23 / Jan 25 2018

www.cs.ubc.ca/~tmm/courses/mds-viz2-17   @tamaramunzner

Viz theory

• block feedback: many people not seeing value of lecture material
• module covers **both** visualization tooling/code and visualization theory
  − lectures: teach theory (assessed with both viz and reasoning)
    • are you coding the right thing?
  − tutorials: teach tooling/code
    • how to code it
  − lab 1: 25% mechanics, 49% code, 21% theory, 5% writing
  − milestone 1: 5% mechanics, 65% theory, 30% writing
  − milestone 2: 15% mechanics, 45% code, 38% theory, 2% writing
  − milestone 3: 5+11=15% mechanics, 10% code, 75% theory
• today: in-class practice on theory to help you do well on milestone 3
  − bar is set considerably higher for milestone 3 than for milestones 1 & 2
    • now that more theory has been covered in class

2

How to handle complexity: 4 families of strategies

• derive new data to show within view
• change view over time
• facet across multiple views
• reduce items/attributes within single view

3

Scenario

• data: room occupancy rates
  − 1 room
  − occupancy measured every 5 min, duration 1 day
• task: characterize space usage pattern

• design
  • propose idioms (visual encoding, interaction)
  • justify idiom choice

4

Consider

• **what's the cardinality of the data?**
• **is a single static chart good enough?**
• **should you derive any useful additional data?**

5

Cardinality

• Marshall: 68 cities * 40 years * 4 crime types = 10,880
• Wine: 130K * 4 = 650,000
  − spatial (hierarchical), quantitative, categorical, free-form text

6

Scenario

• data: room occupancy rates
  − 20 rooms
  − measured every 5 min, duration 1 day
• task: compare space usage patterns between rooms

• design
  • propose idioms (visual encoding, interaction)
  • justify idiom choice

7

Consider

• what's the cardinality of the data?
• is a single static chart good enough?
• should you derive any useful additional data?

• **what are trade-offs between**
  − **filtering to see one chart at a time**
  − **showing all side by side with small multiples**
  − **superimposing all on top of each other**

8

Scenario

• data: room occupancy rates in building
  − 1 building: 200 rooms across 4 floors
  − measured every 5 min, duration 1 day
  − time series + floor plans
• task: characterize space usage patterns
  − trends, outliers

• design
  − propose & justify idioms

9

Consider

• what's the cardinality of the data?
• is a single static chart good enough?
• should you derive any useful additional data?
• what are trade-offs between
  − filtering to see one chart at a time
  − showing side by side with small multiples
  − superimposing on top of each other

• **multi-scale structure to exploit? aggregate, zoom, slice/dice, filter?**

10

Scenario

• data: room occupancy rates in building
  − 1 building: 200 rooms across 4 floors
  − measured every 5 min, duration 1 **year**
  − time series + floor plans **+ room sizes**
• task: characterize space usage patterns
  − trends, outliers

• design
  − propose & justify idioms

11

Consider

• what's the cardinality of the data?
• is a single static chart good enough?
• should you derive any useful additional data?
• what are trade-offs between
  − filtering to see one chart at a time
  − showing side by side with small multiples
  − superimposing on top of each other
• multi-scale structure to exploit? aggregate, zoom, slice/dice, filter?
• **can you normalize the data? should you - always vs on demand?**
• **how to handle multi-scale space and multi-scale time?**

12

Design Choices
(Additional Context)

13

Normalized vs Absolute

14

Idiom: **choropleth map**

• *use* given spatial data
  − when central task is understanding spatial relationships
• data
  − geographic geometry
  − table with 1 quant attribute per region
• encoding
  − use given geometry for area mark boundaries
  − sequential segmented colormap *[more later]*
  − (geographic heat map)
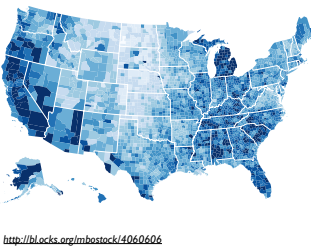
http://blocks.org/mbostock/4060606

15

Population maps trickiness

• beware!
• absolute/counts vs normalized/relative
  • population density vs per capita
• investigate with Ben Jones Tableau Public demo
  • http://public.tableau.com/profile/ben.jones#!/vizhome/PopVsFin/PopVsFin
    Are Maps of Financial Variables just Population Maps?
    • yes, unless you look at per capita (relative) numbers
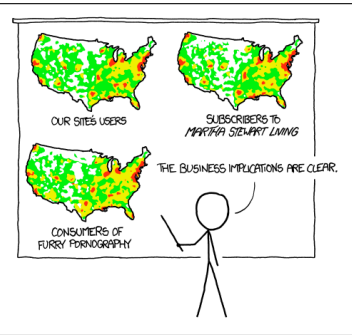
[https://xkcd.com/1138]

Idiom: Bayesian surprise maps

- use models of expectations to highlight surprising values
- confounds (population) and variance (sparsity)
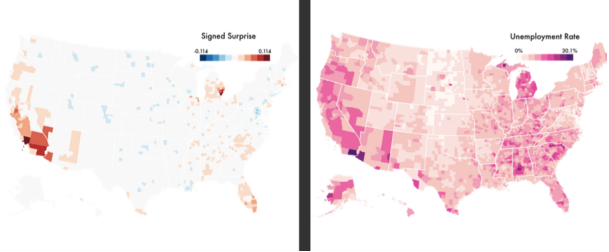
[Surprise! Bayesian Weighting for De-Biasing Thematic Maps. Correll and Heer. Proc InfoVis 2016]

https://medium.com/@uwdata/surprise-maps-showing-the-unexpected-e92b67398865   https://idl.cs.washington.edu/papers/surprise-maps/

17

Radial vs Rectilinear

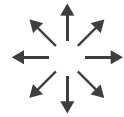18

Axis Orientation

→ Rectilinear   → Parallel   → Radial

19

## Idioms: **radial bar chart, star plot**

- radial bar chart
  - radial axes meet at central ring, line mark
- star plot
  - radial axes, meet at central point, line mark
- bar chart
  - rectilinear axes, aligned vertically

- accuracy
  - length unaligned with radial
    - less accurate than aligned with rectilinear

[Vismon: Facilitating Risk Assessment and Decision Making In Fisheries Management. Booshehrian, Möller, Peterman, and Munzner. Technical Report TR 2011-04, Simon Fraser University, School of Computing Science, 2011.]
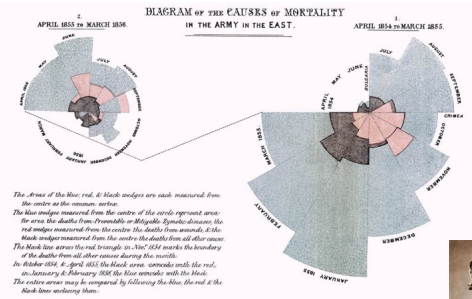
20

Radial Orientation: Radar Plots
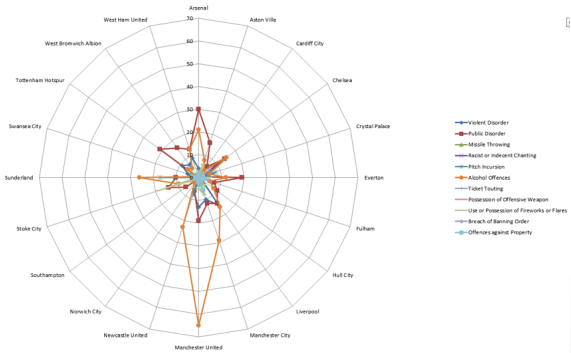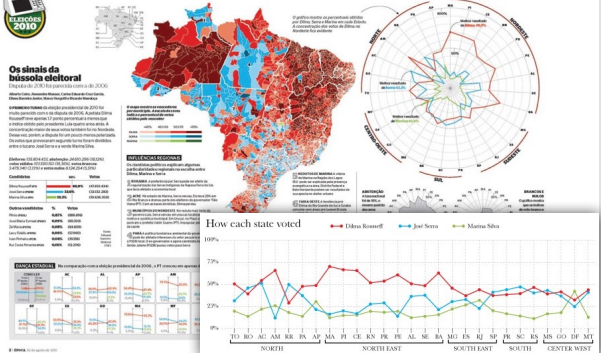

LIMITATION: Not good when categories aren't cyclic

[Slide courtesy of Ben Jones]

"Diagram of the causes of mortality in the army in the East" (1858)

[Slide courtesy of Ben Jones]
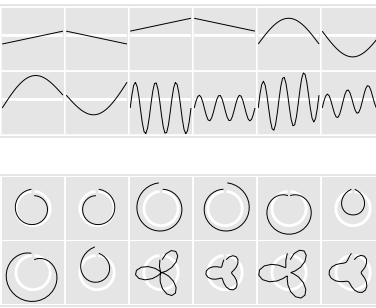

"Radar graphs: Avoid them (99.9% of the time)"

http://www.thefunctionalart.com/2012/11/radar-graphs-avoid-them-999-of-time.html

[Slide courtesy of Ben Jones]

Idiom: **glyphmaps**

- rectilinear good for linear vs nonlinear trends

- radial good for cyclic patterns

→ Axis Orientation
→ Rectilinear   → Parallel   → Radial

[Glyph-maps for Visually Exploring Temporal Patterns in Climate Data and Models. Wickham, Hofmann, Wickham, and Cook. Environmetrics 23:5 (2012), 382–393.]

24

Radial orientation

- perceptual limits
  - polar coordinate asymmetry
    - angles lower precision than lengths
    - frequently problematic
    - sometimes can be deliberately exploited!
      - for 2 attribs of very unequal importance
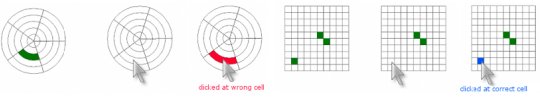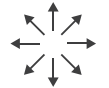
→ Axis Orientation
→ Rectilinear
→ Parallel
→ Radial

[Uncovering Strengths and Weaknesses of Radial Visualizations - an Empirical Approach. Diehl, Beck and Burch. IEEE TVCG (Proc. InfoVis) 16(6):935–942, 2010.]

25

Overview first, zoom and filter, details on demand

- influential mantra from Shneiderman

[The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. Shneiderman. Proc. IEEE Visual Languages, pp. 336–343, 1996.]

- overview = summary
  - microcosm of full vis design problem

→ Query
→ Identify   → Compare   → Summarise

26

Thursday

- Beyond R
  - Ana on broader landscape
  - Ana on direct comparison of Tableau to R
  - Vaden on python interactive tools
- Evaluations
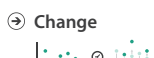- Further Design & Justification Exercises
- Next Steps

27

Evaluations

28

Viz theory

- block feedback: many people not seeing value of lecture material
- module covers **both** visualization tooling/code and visualization theory
  - lectures: teach theory (assessed with both viz and reasoning)
    - are you coding the right thing?
  - tutorials: teach tooling/code
    - how to code it
  - lab 1: 25% mechanics, 49% code, 21% theory, 5% writing
  - milestone 1: 5% mechanics, 65% theory, 30% writing
  - milestone 2: 15% mechanics, 45% code, 38% theory, 2% writing
  - milestone 3: 5+11=15% mechanics, 10% code, 75% theory
- today: in-class practice on theory to help you do well on milestone 3
  - bar is set considerably higher for milestone 3 than for milestones 1 & 2
    - now that more theory has been covered in class

29

How to handle complexity: 4 families of strategies

→ Derive

	Manipulate	Facet	Reduce
	→ Change	→ Juxtapose	→ Filter
	→ Select	→ Partition	→ Aggregate
	→ Navigate	→ Superimpose	→ Embed

- derive new data to show within view
- change view over time
- facet across multiple views
- reduce items/attributes within single view

30

Scenarios last time

- 1 room, occupancy every 5 min over 1 day
- 20 rooms, occupancy every 5 min over 1 day
- 200 rooms across 4 floors, occupancy every 5 min over 1 day, floor plans
- 200 rooms, 4 floors, occupancy every 5 min over 1 year, floor plans, room sizes

31

Scenario

- data: currency exchange rates
  - 30 countries (each against CAD)
  - measured every 5 min, duration 5 years
  - time series + country names + continent names (+ map shapefiles) + country populations
- task: find groups of similarly-performing currencies

- design
  - propose & justify idioms

32

Scenario

- data: CPU usage across many machines
  - 100 machines, belonging to 20 companies
  - measured every 5 min, duration 1 month
  - time series + company name + company location (country)
- task: capacity planning for machine room

- design
  - propose & justify idioms

Scenario

- data: many metrics across many machines
  - 100 machines, belonging to 20 companies
  - 4 metrics measured every 5 min, duration 1 month
    - CPU, memory, disk I/O, network traffic
  - time series + company name + company sector (finance/tech/entertainment/other)
- task: forensic analysis to determine possible causes of crashes

- design
  - propose & justify idioms

Consider

- what's the cardinality of the data?
- is a single static chart good enough?
- should you derive any useful additional data?
- what are trade-offs between
  - filtering to see one chart at a time
  - showing side by side with small multiples
  - superimposing on top of each other
- multi-scale structure to exploit? aggregate, zoom, slice/dice, filter?
- can you normalize the data? should you - always vs on demand?
- how to handle multi-scale space and multi-scale time?
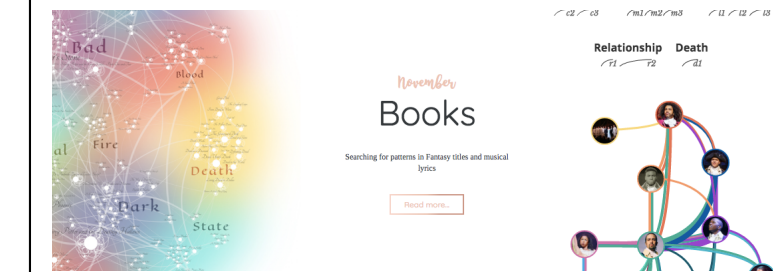- **is spatial information germane or extraneous?**

Next Steps
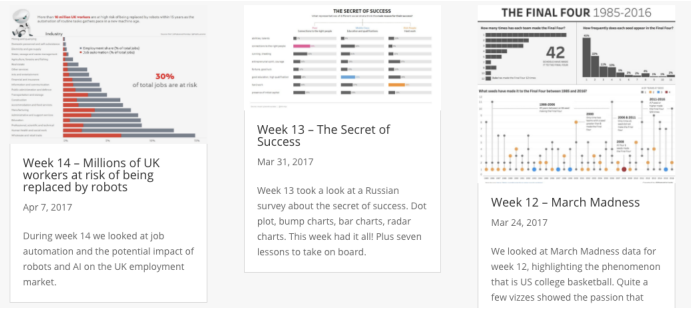
Visual Design Process In Depth: **Dear Data**

http://www.dear-data.com/by-week/

Visual Design Process In Depth: **Data Sketches**

http://www.datasketch.es/

Redesign En Masse: **Makeover Mondays**

http://www.makeovermonday.co.uk/blog/