Evaluating Visualizations

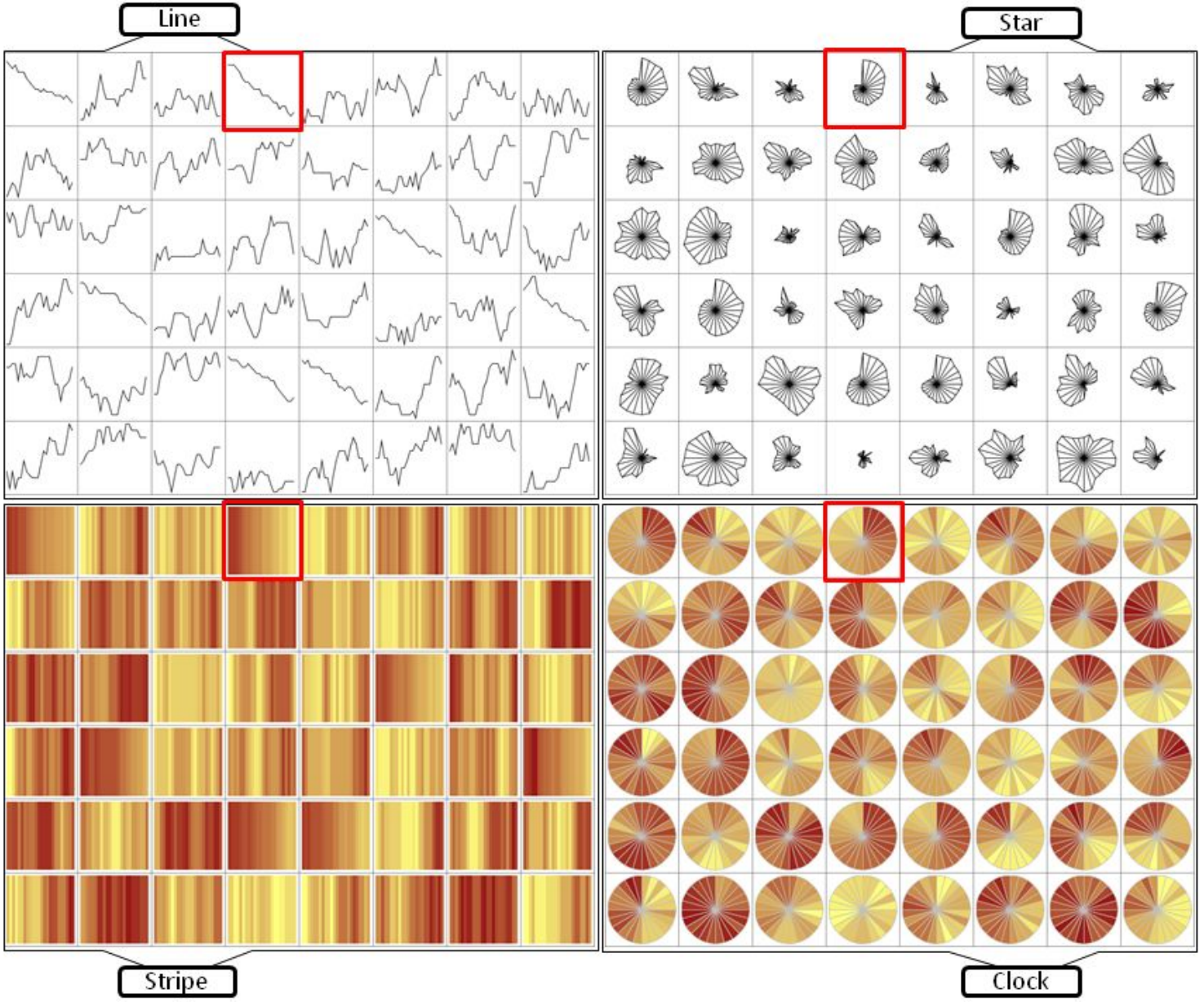
The Human Factor

Benjamin Bach

March 2019

Edinburgh University

bbach@inf.ed.ac.uk


Line
Star
Stripe
Clock

1st Rule of Evaluation:

Sit back, relax, evaluate.

2nd Rule of Evaluation:

Iterate.

Factors

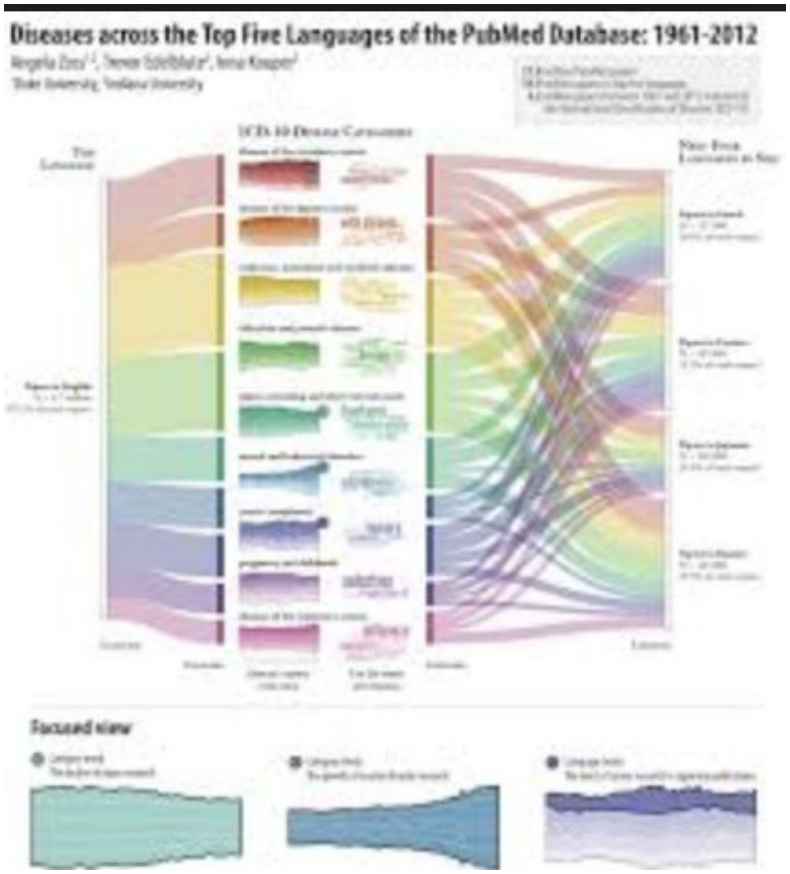
- Preknowledge
- Visual literacy, numeracy, data literacy
- Visualization literacy
- Perception
- Understanding
- Interest..

How do we know if a visualization is successful?

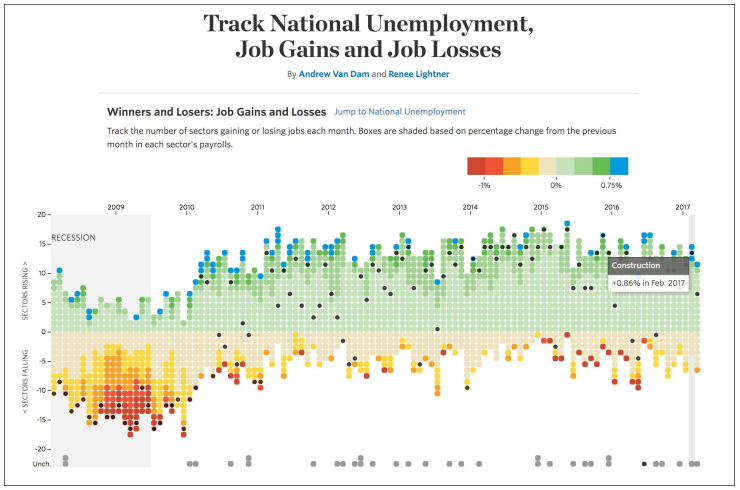
1. **Readability:** To correctly perceive a visualization's content

1. Readability

Resolution



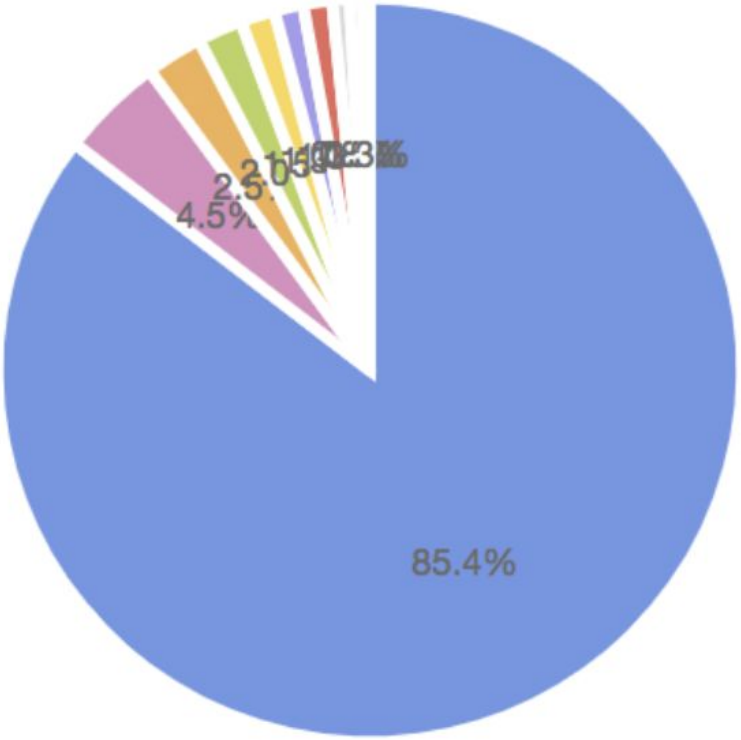
Size



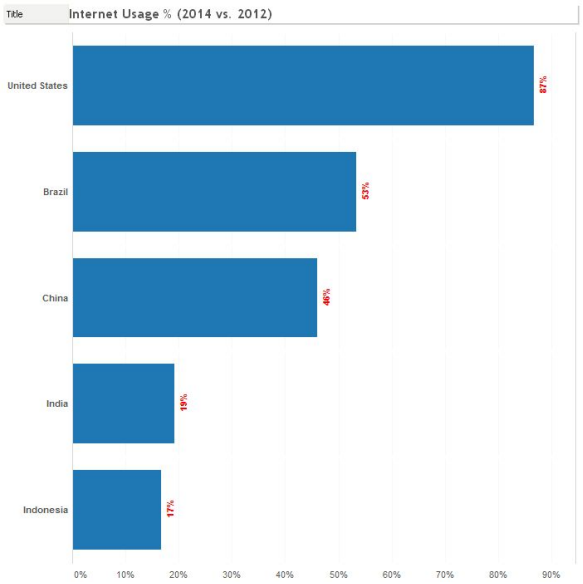
Lorem ipsum dolor sit amet, consectetur adipiscing elit. Pellentesque sollicitudin ipsum elit, et vestibulum nisl dictum et. Sed lobortis molestie felis. Praesent et ligula commodo magna fringilla egestas. Nam nec risus in magna facilisis sollicitudin nec et metus. Sed tincidunt dapibus lacus in viverra. Donec gravida finibus metus eget semper. In nec mauris

1. Readability

Overlapping labels



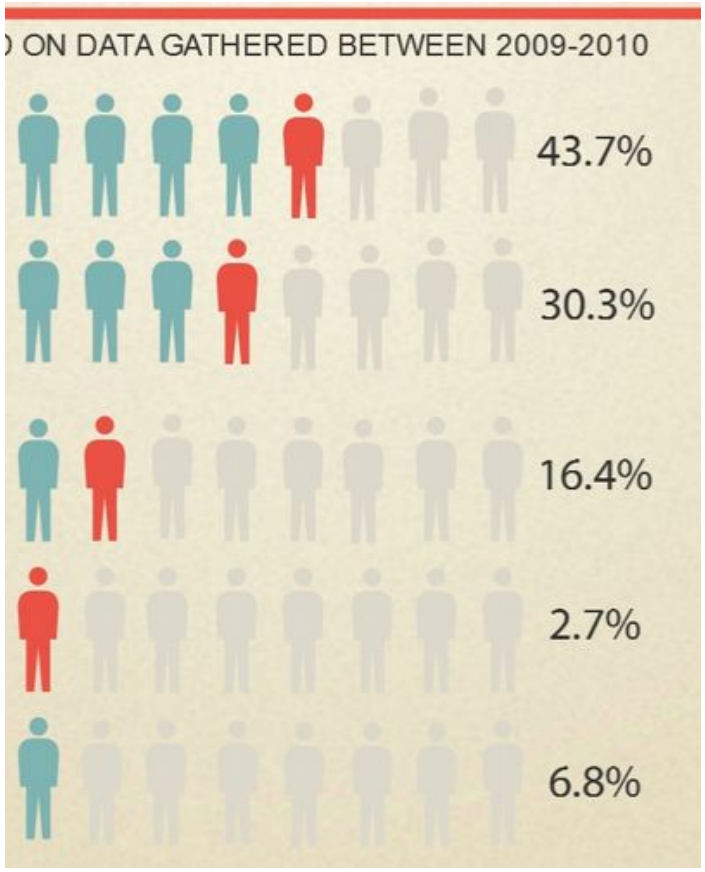
Label Size



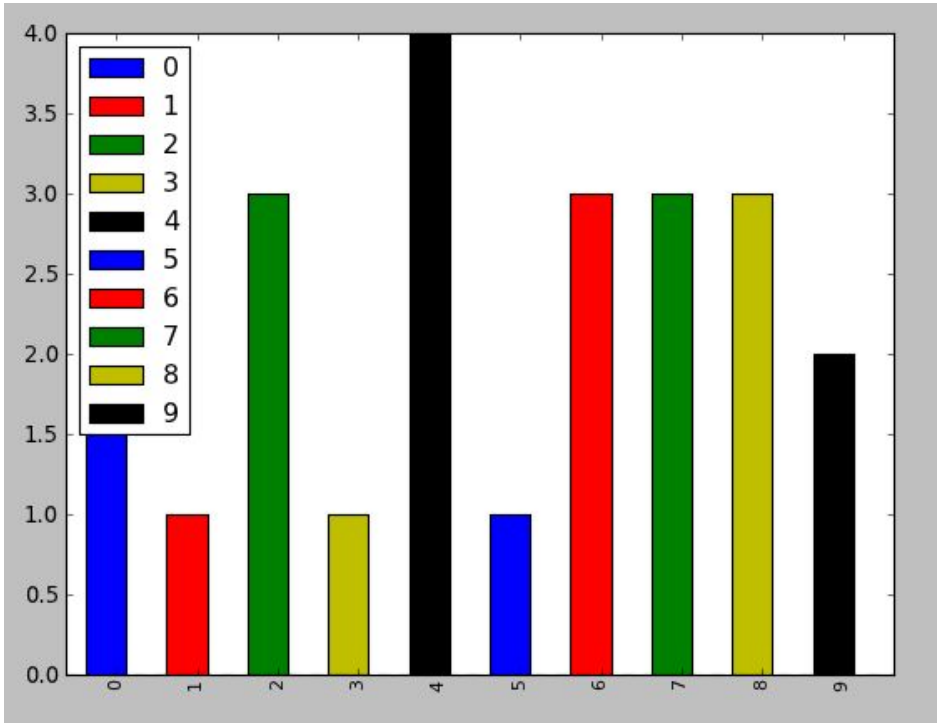
2. **Understandability:** Ability to support reasoning with the visualization.

2. Understandability

Missing Legends

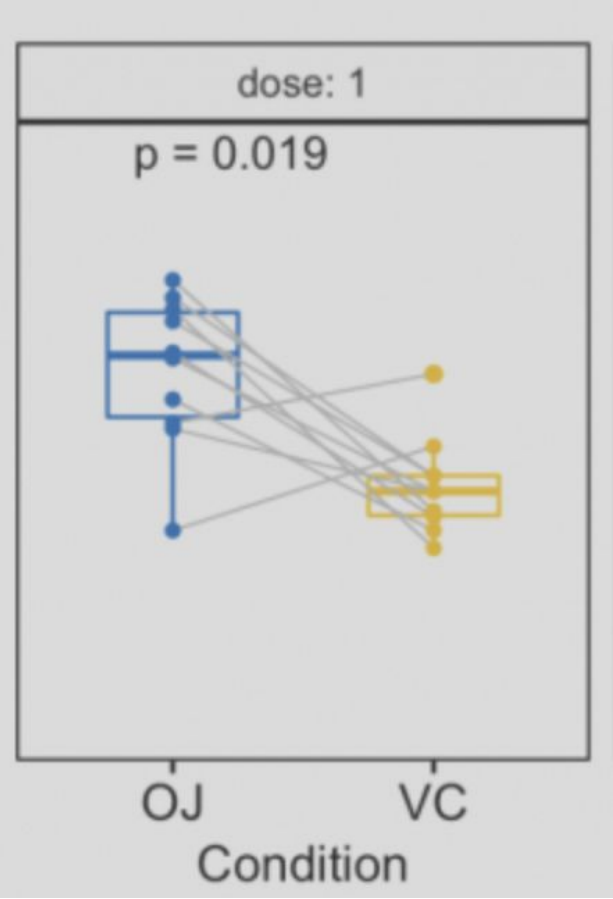


Meaningless encodings

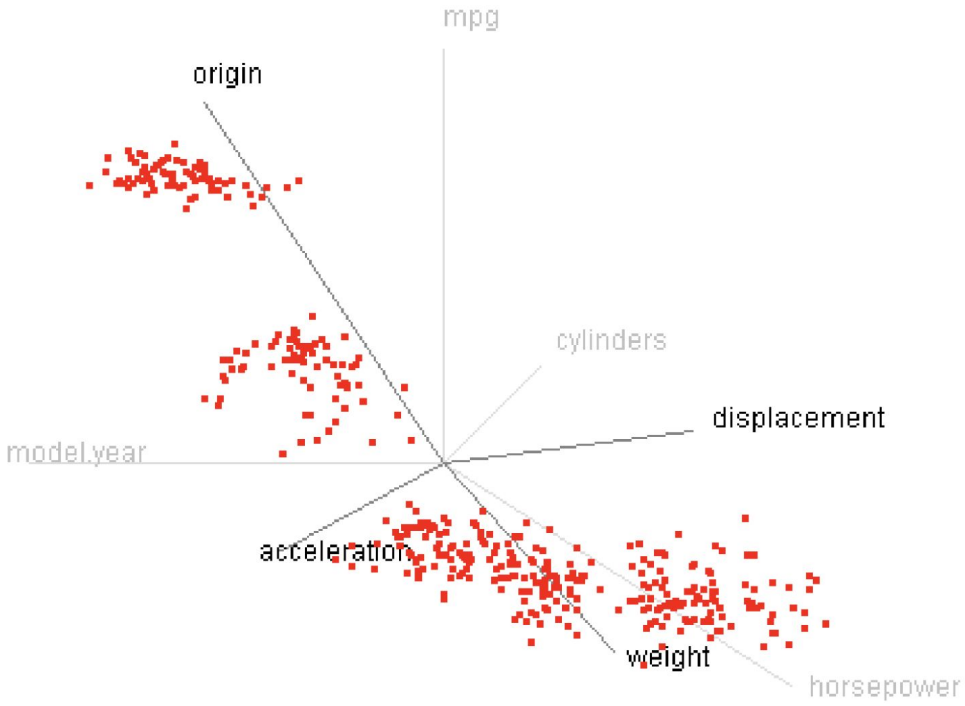


2. Understandability

Complicated / missing labels

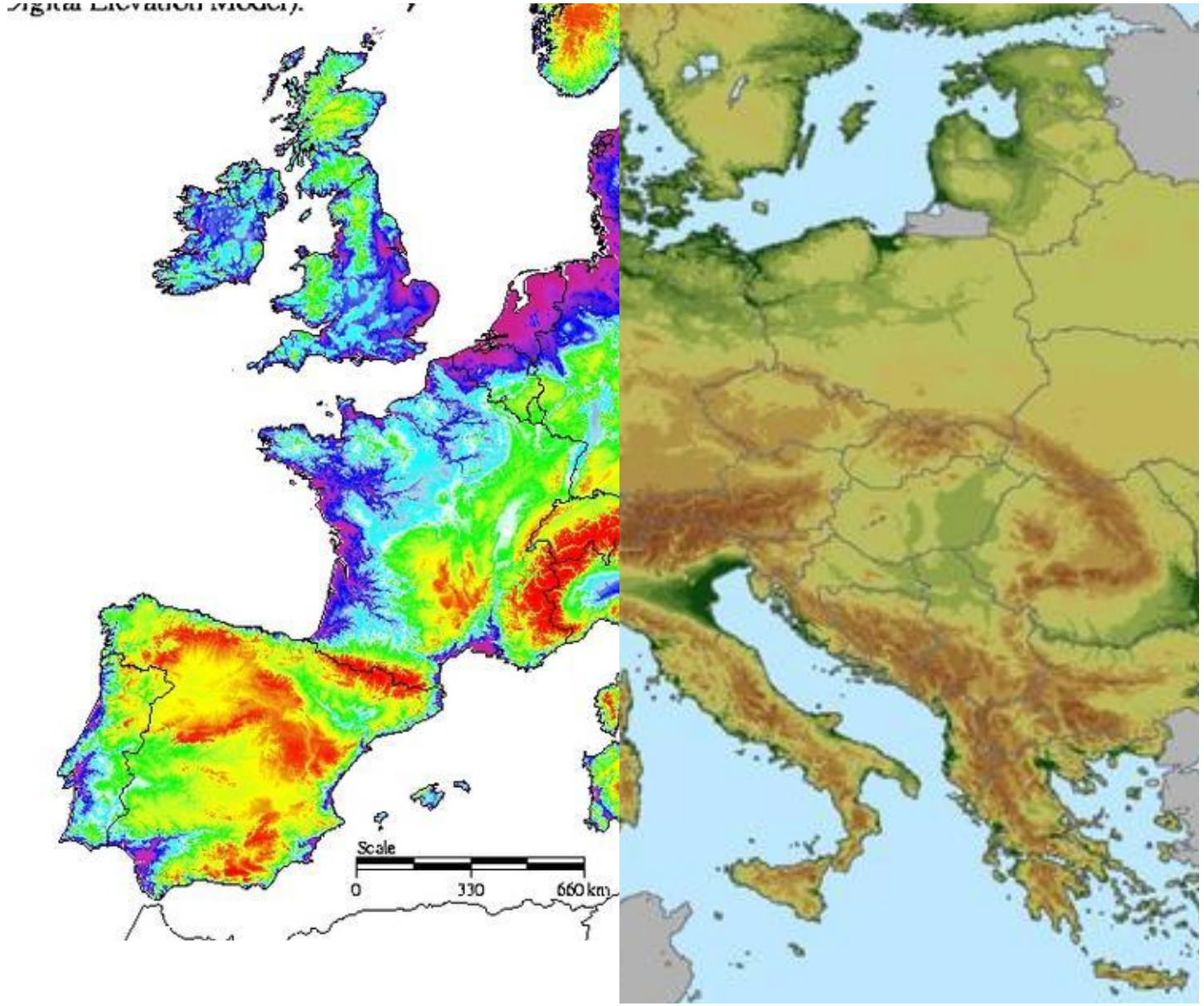


Unknown / unexplained visualizations



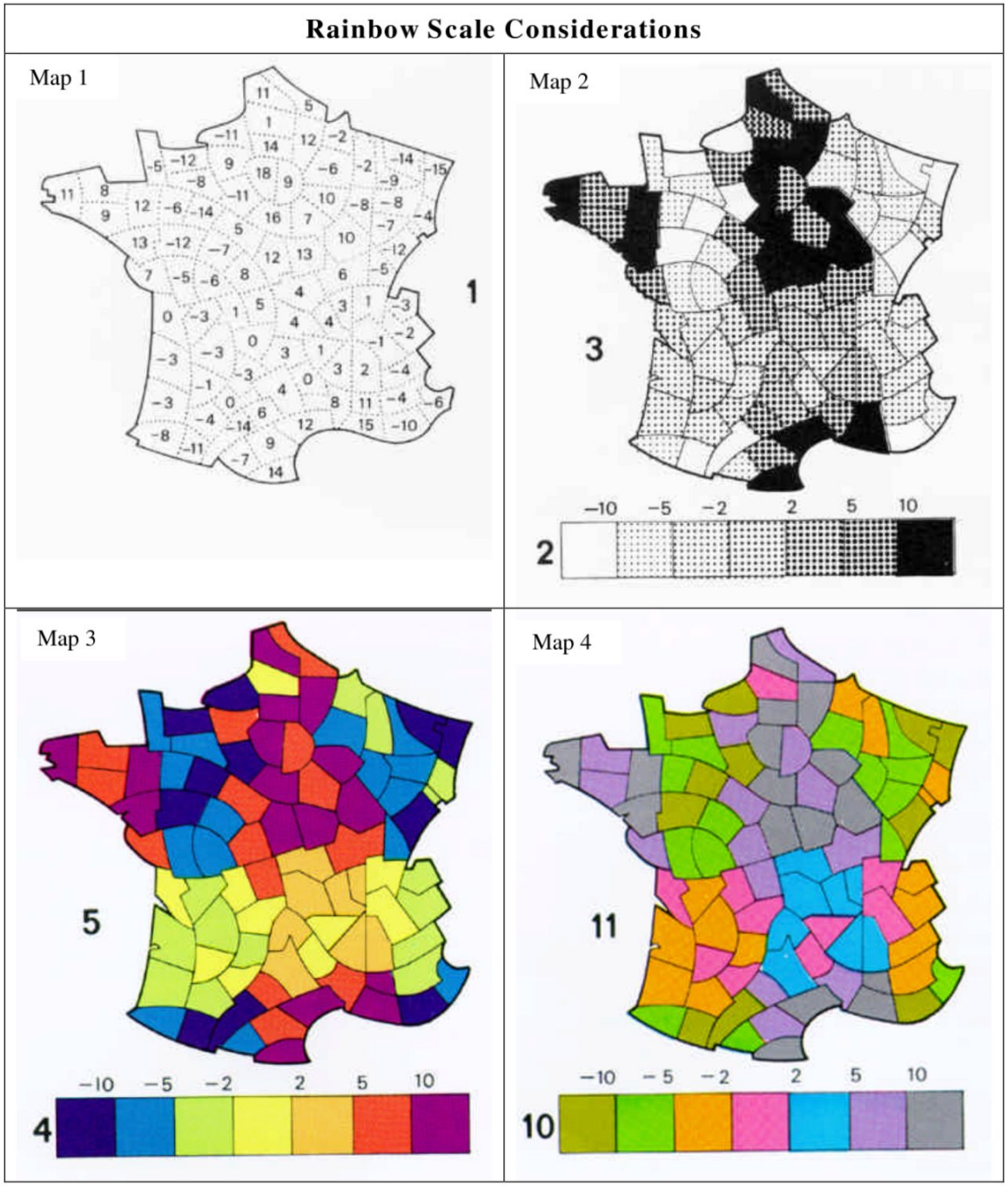
3. **Appropriateness:** Matches a given data set, task, message, context, audience…

3. Appropriate



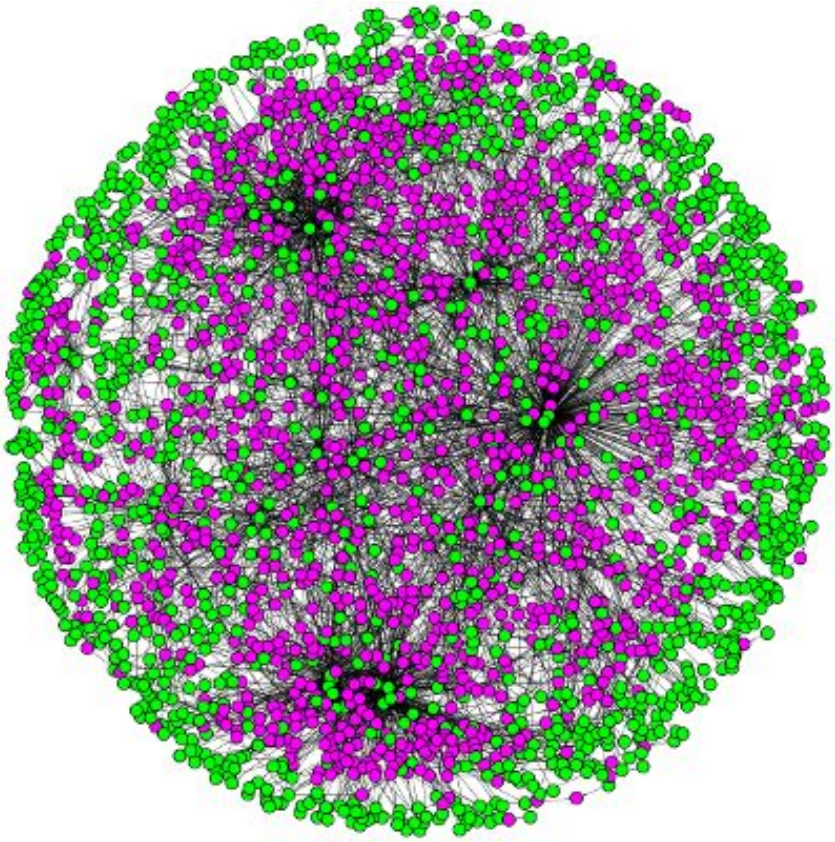
3. Appropriate

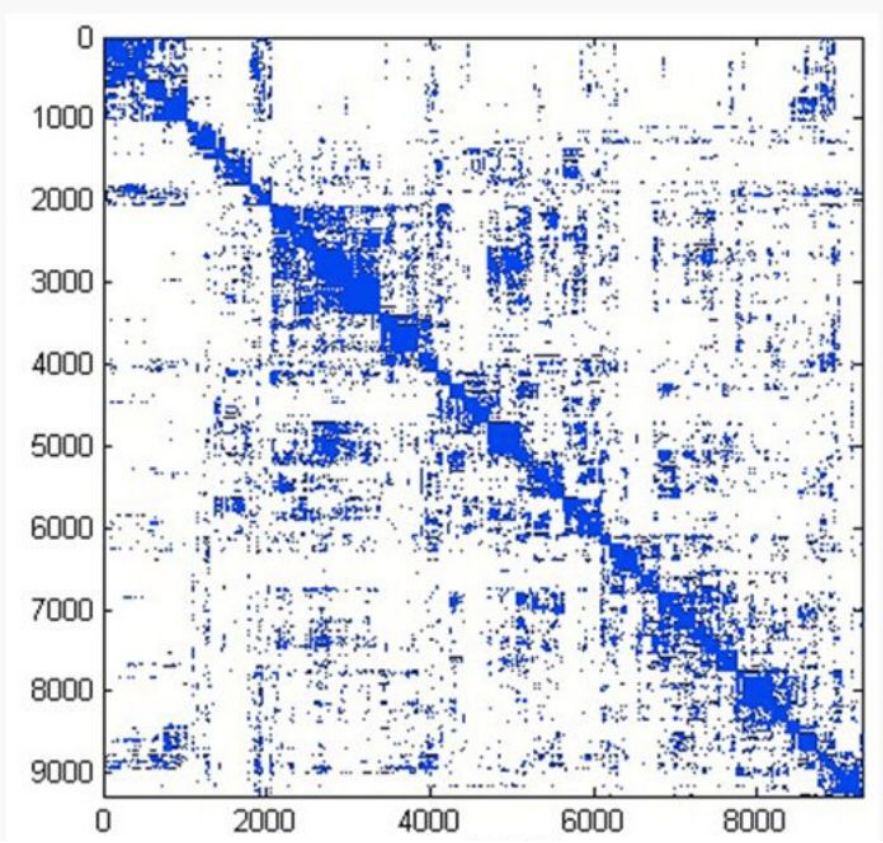
Visual Encoding



3. Appropriate

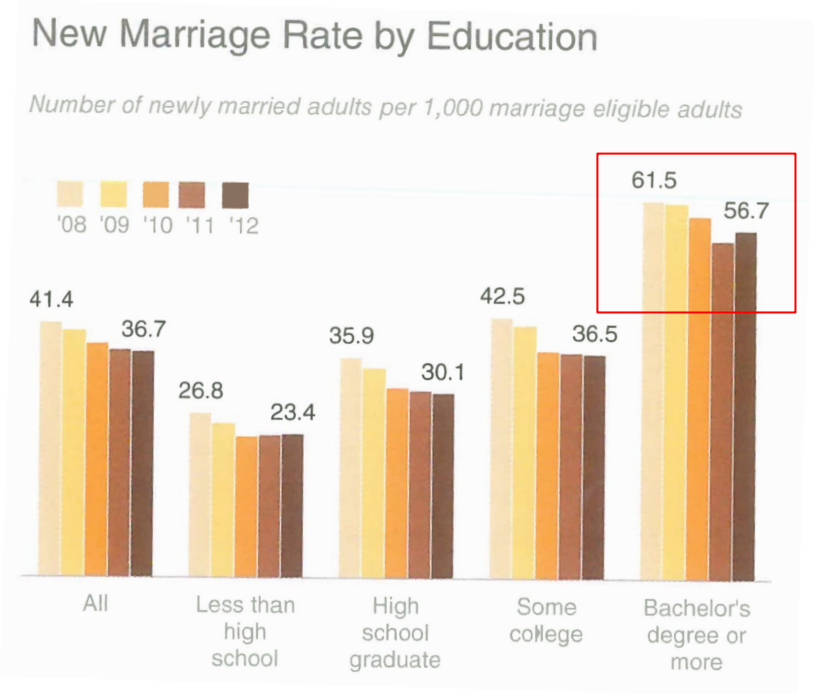
Appropriate for data

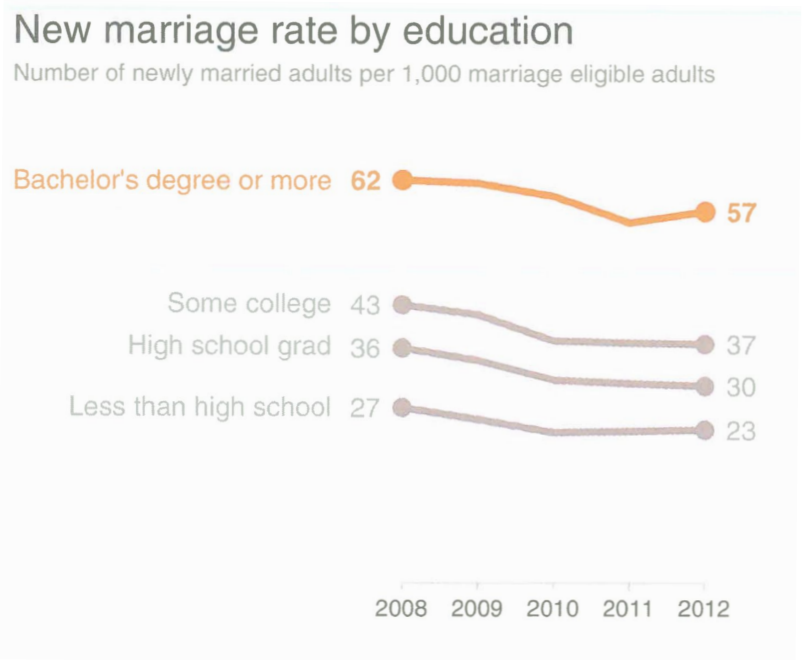




3. Appropriate

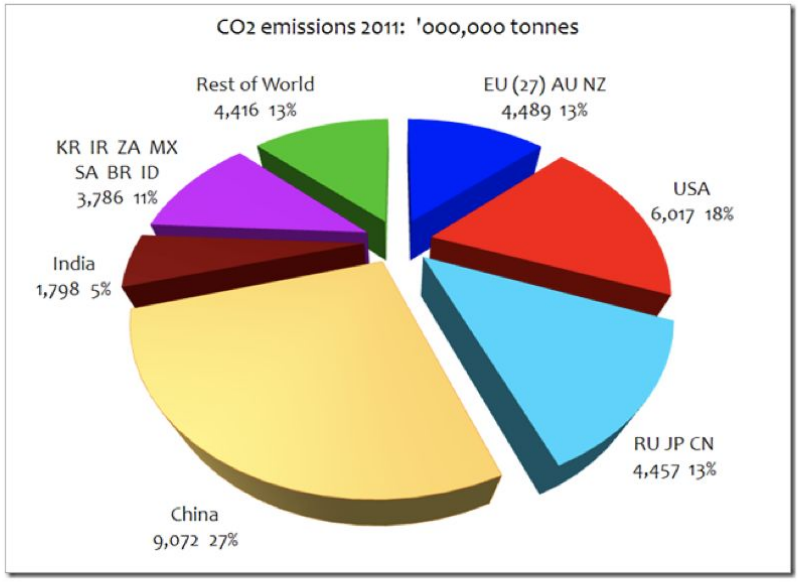
Appropriate for message

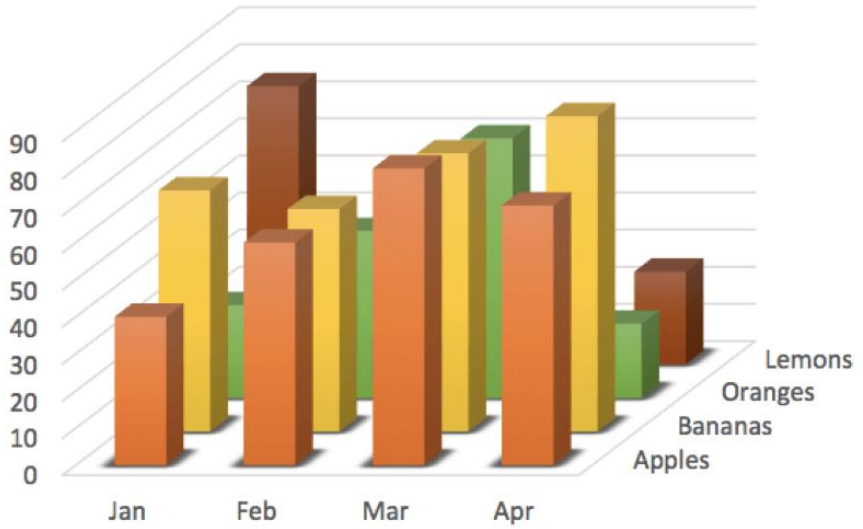


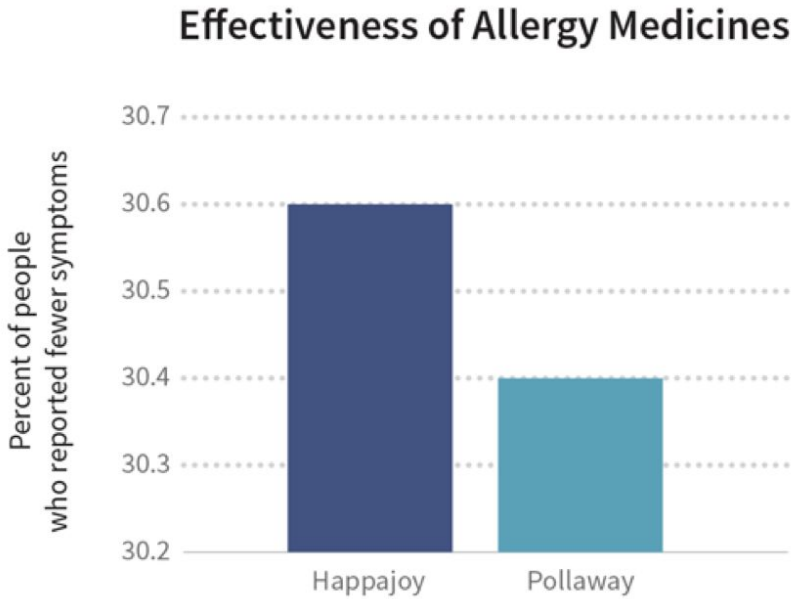


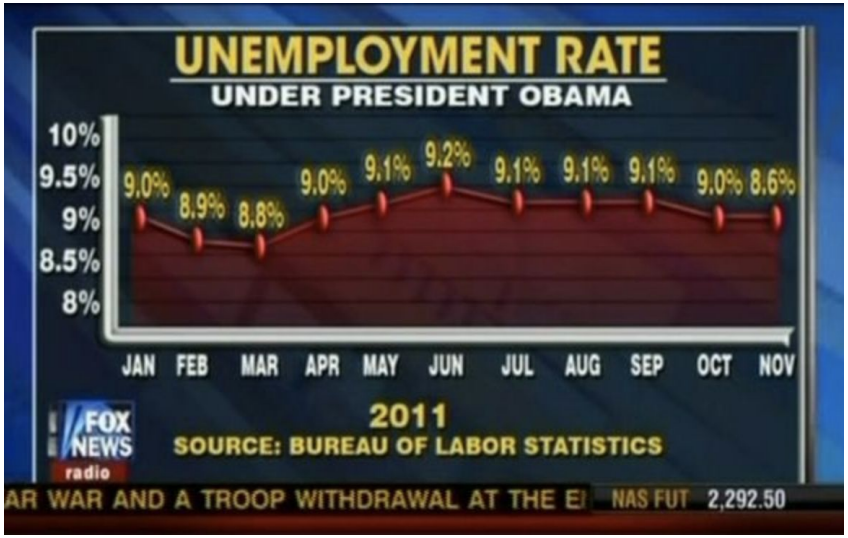
4. Truthful: Not deceive nor hide or distort.

4. Truthful


CO2 emissions 2011: '000,000 tonnes
Rest of World
4,416 13%
EU (27) AU NZ
4,489 13%
KR IR ZA MX
SA BR ID
3,786 11%
USA
6,017 18%
India
1,798 5%
RU JP CN
4,457 13%
China
9,072 27%


90
80
70
60
50
40
30
20
10
0
Jan
Feb
Mar
Apr
Lemons
Oranges
Bananas
Apples


Effectiveness of Allergy Medicines
Percent of people who reported fewer symptoms
30.7
30.6
30.5
30.4
30.3
30.2
Happajoy
Pollaway


UNEMPLOYMENT RATE
UNDER PRESIDENT OBAMA
10%
9.5%
9%
8.5%
8%
9.0%
8.9%
8.8%
9.0%
9.1%
9.2%
9.1%
9.1%
9.1%
9.0%
8.6%
JAN FEB MAR APR MAY JUN JUL AUG SEP OCT NOV
2011
SOURCE: BUREAU OF LABOR STATISTICS
FOX NEWS radio
AR WAR AND A TROOP WITHDRAWAL AT THE E
NAS FUT 2,292.50

Successful visualizations

1. The visualization is **readable**
2. The visualization is **understandable**.
3. The visualization is **appropriate**.
4. The visualization is **truthful**.

How do we know if a visualization is successful?

Personas

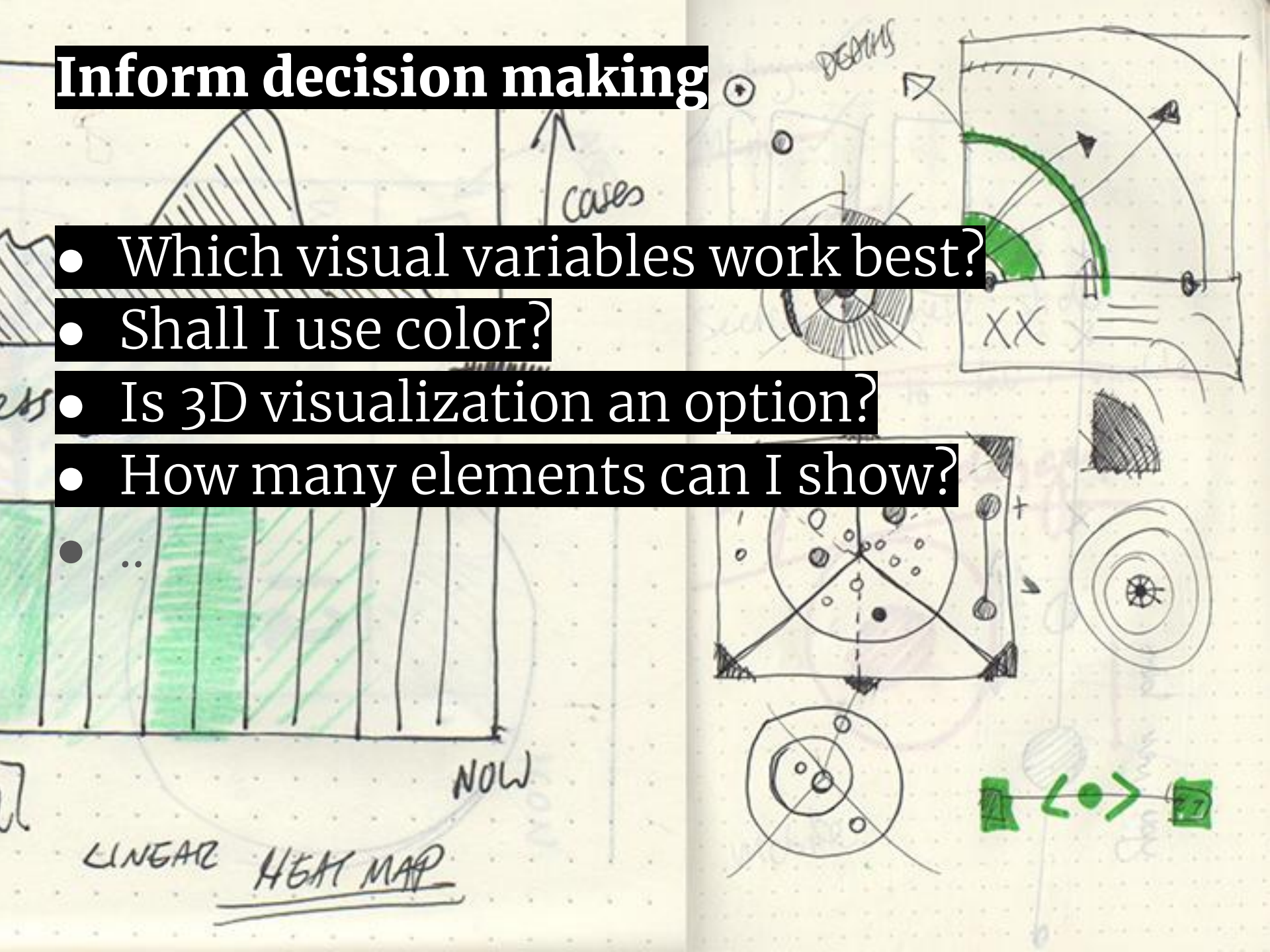
Personas

- Background & pre-knowledge
- Device & context
- Tasks & questions
- Visualization abilities (!)

1. Guidelines

Inform decision making

- Which visual variables work best?
- Shall I use color?
- Is 3D visualization an option?
- How many elements can I show?
- ..



Guideline:

1. The visualization is **readable**
2. The visualization is **understandable**.
3. The visualization is **appropriate**.
4. The visualization is **truthful**.

Guideline

"general rule, principle, or piece of advice"

Examples:

- *Overview first, zoom and filter, details on demand.*
- *Create the simplest graph that conveys the information you want to convey.*
- *Rainbow colormap considered harmful.*
- *Select meaningful axis ranges.*

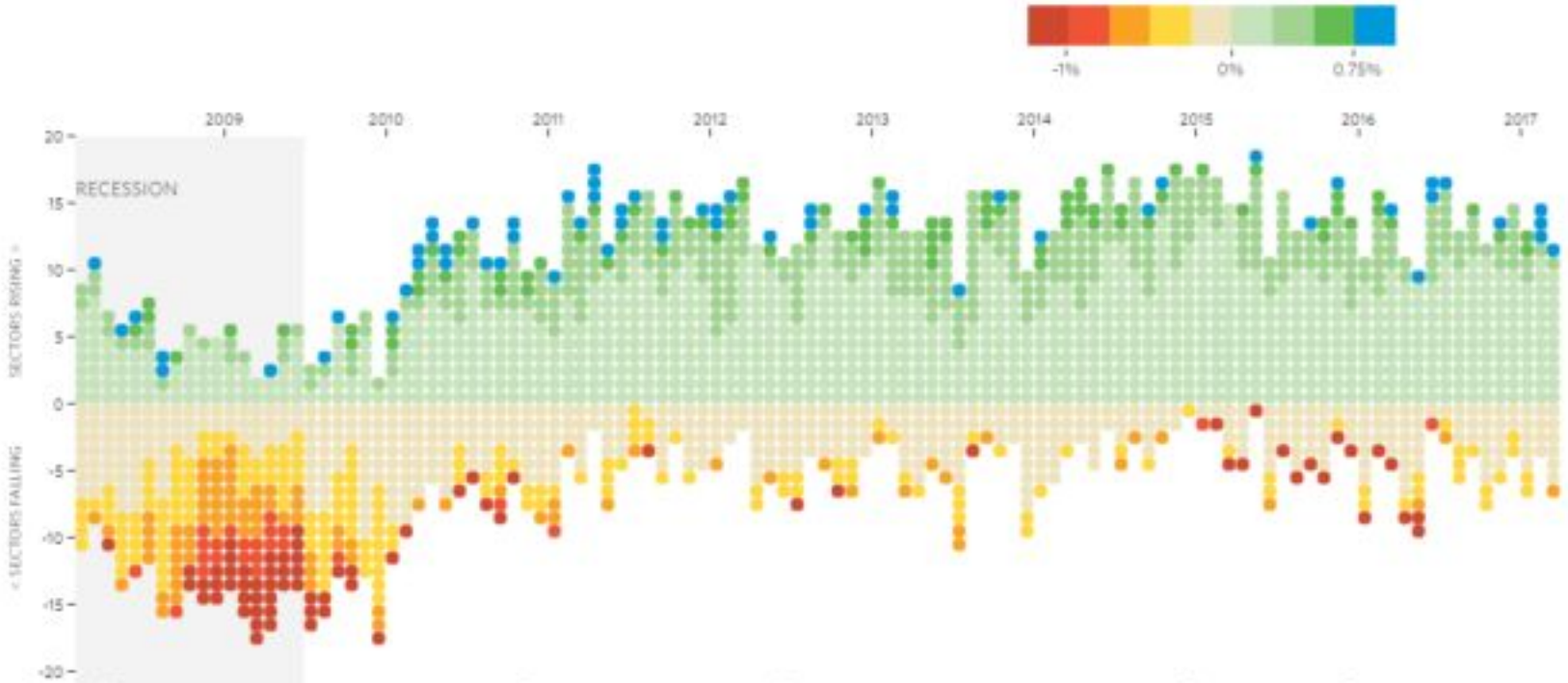
Where to find guidelines

- Text books (see course blog)
- Papers (see course blog)
- VisGuides (https://visguides.dbvis.de)

2. Case Study

Test and argue about a design

- Design is ready
- Does it tell me what I want it to tell?
- Can it convince?

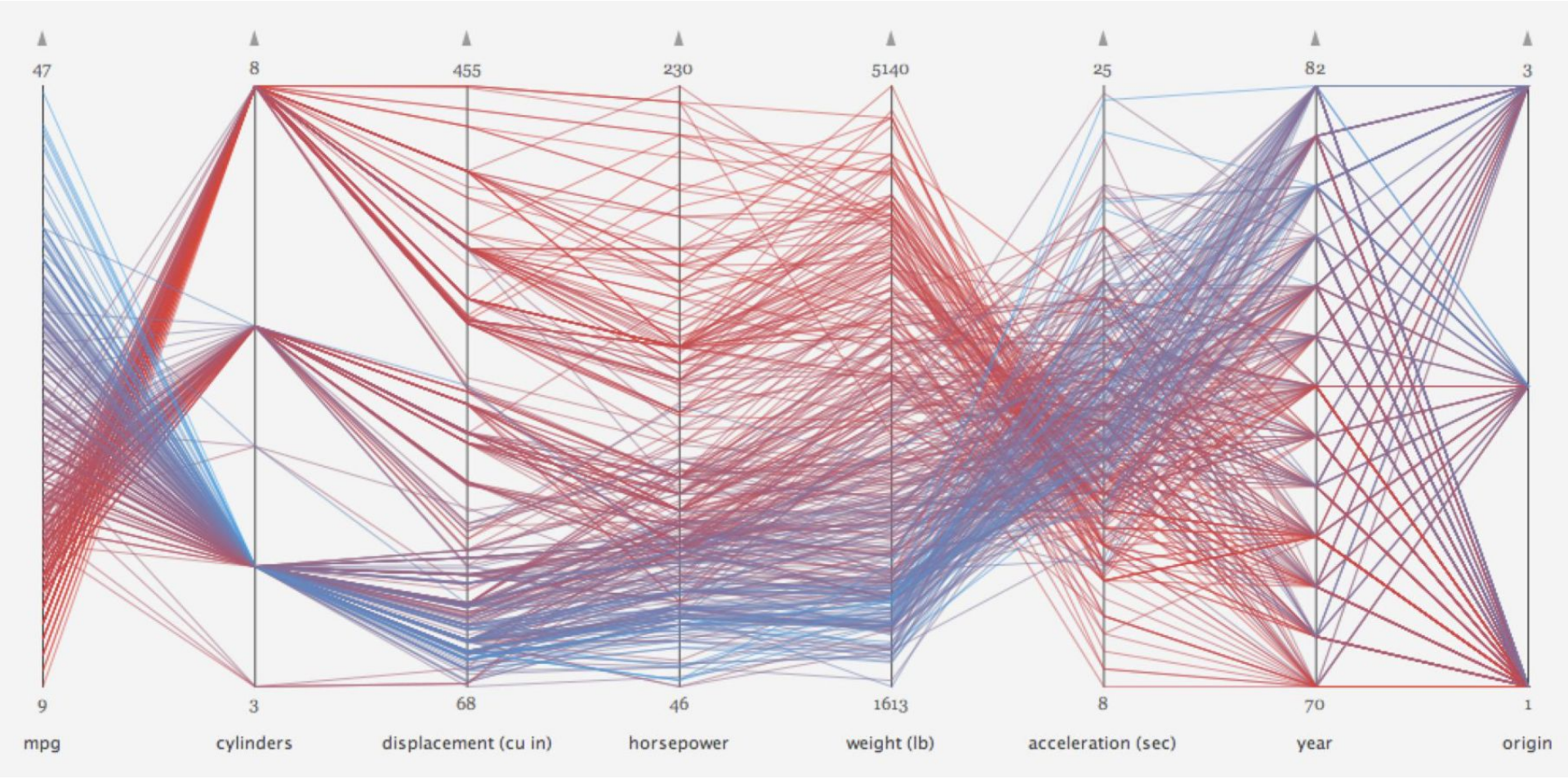


2. Case study

Explaining a visualization using example use cases

1. Explain context (data, users, motivation)
2. Explain tasks
3. Introduce visualization
4. Show how visualization solves task(s)

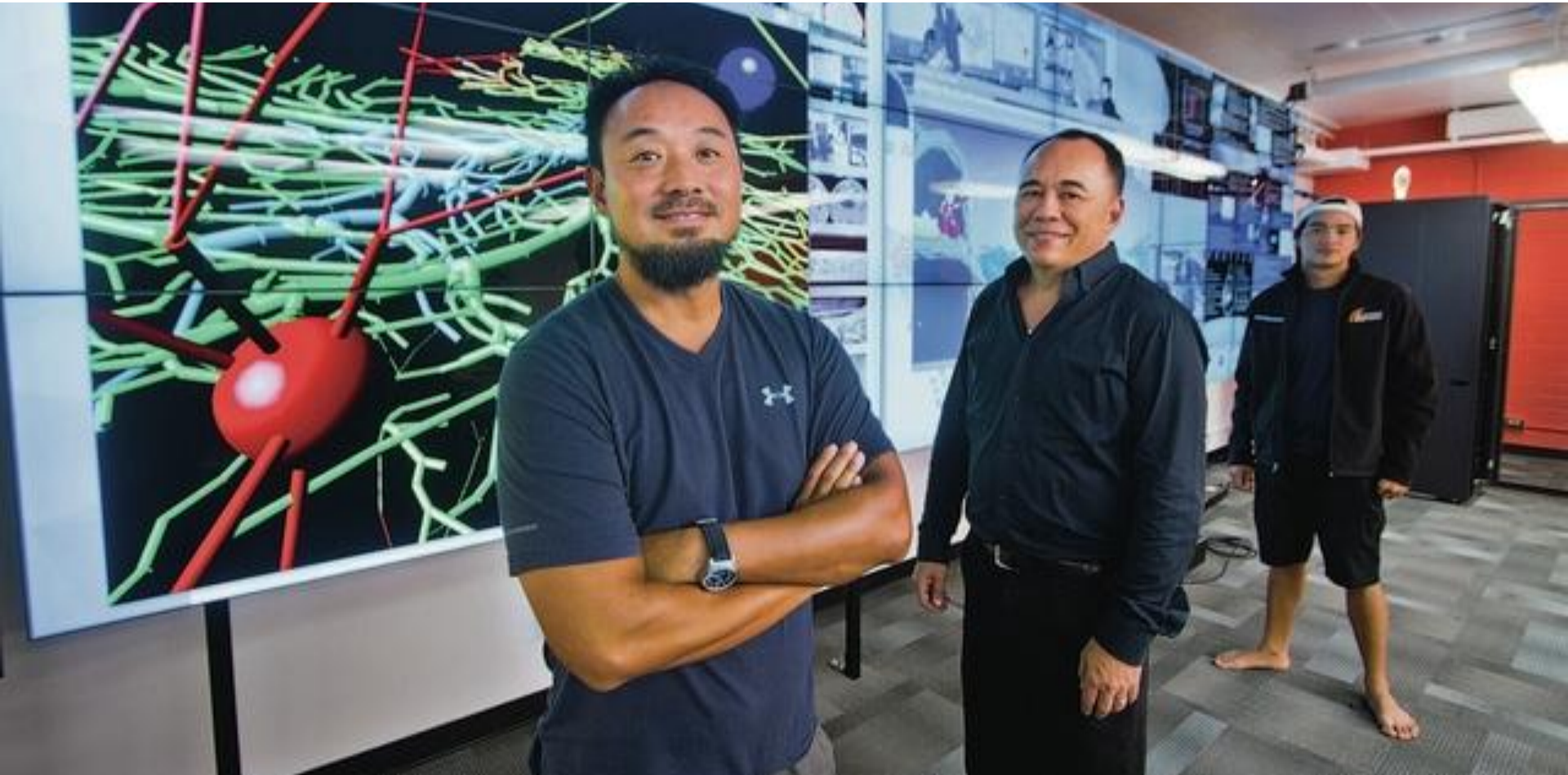
2. Case Study


47
8
455
230
5140
25
82
3
9
3
68
46
1613
8
70
1
mpg
cylinders
displacement (cu in)
horsepower
weight (lb)
acceleration (sec)
year
origin

3. Expert Feedback

External advice

1. Find an expert
2. Introduce your visualization & context
3. Ask specific questions



Example Question

- Can you see that X is larger than Y?
- Can you spot the four groups I am mentioning?
- Do you think the color encoding is appropriate?
- Do you think the visualization is appropriate
- ...

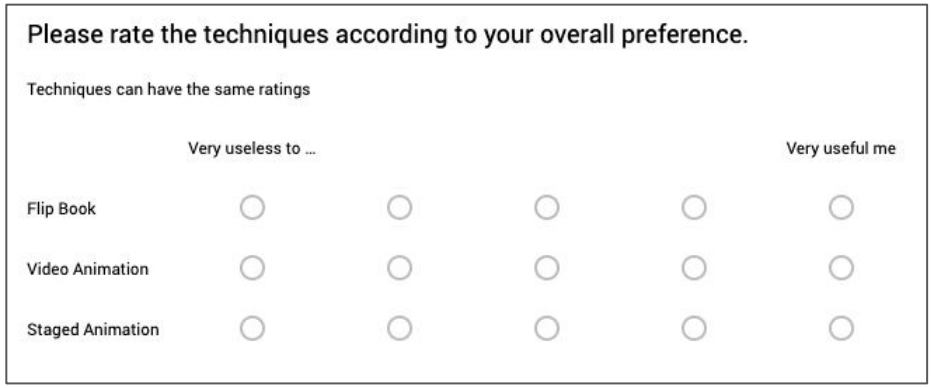
4. User Feedback

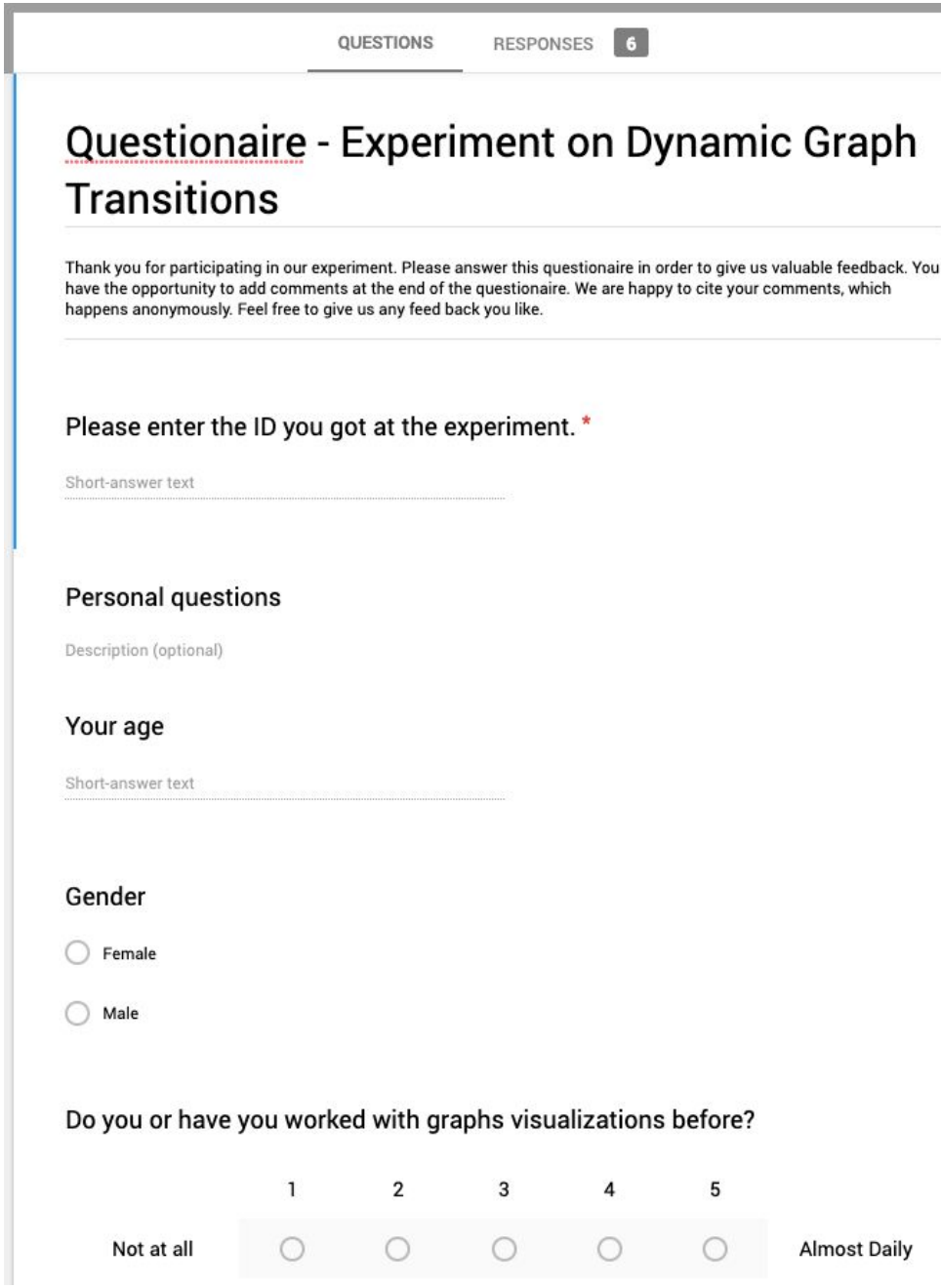
User feedback

1. Find end users
2. Demos, interviews, questionnaires, survey,

Questionnaires: Google forms

- Demographics
- Preferences
- Rankings

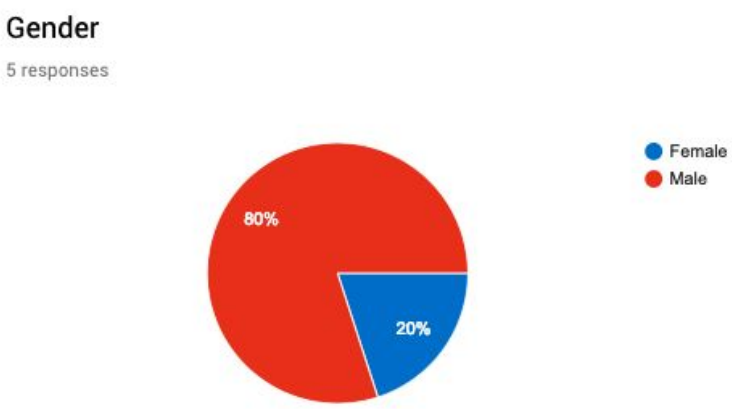


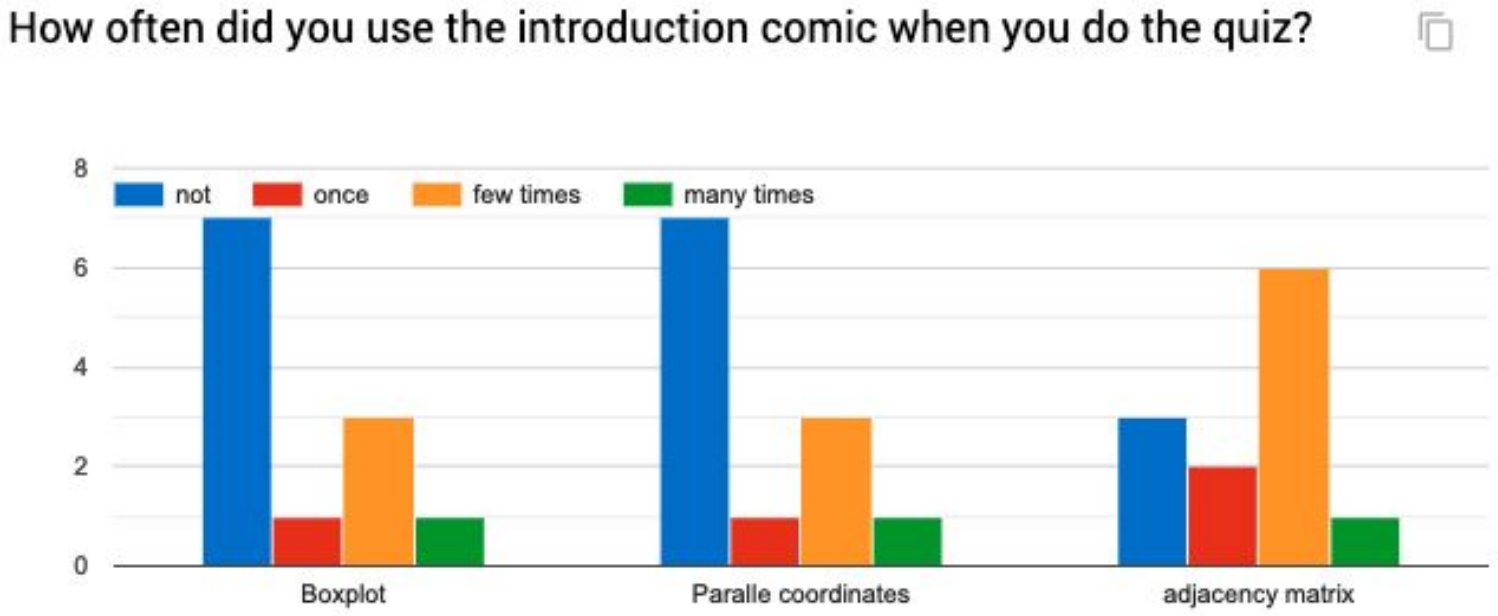


Questionnaires:
Google forms

- Pose very specific questions: "*how hard was it to understand X.*"
- Likert scales ~5 to 7 points
- Ask for positive and negative feedback
- Leave space for open comments

Questionnaires: Google forms



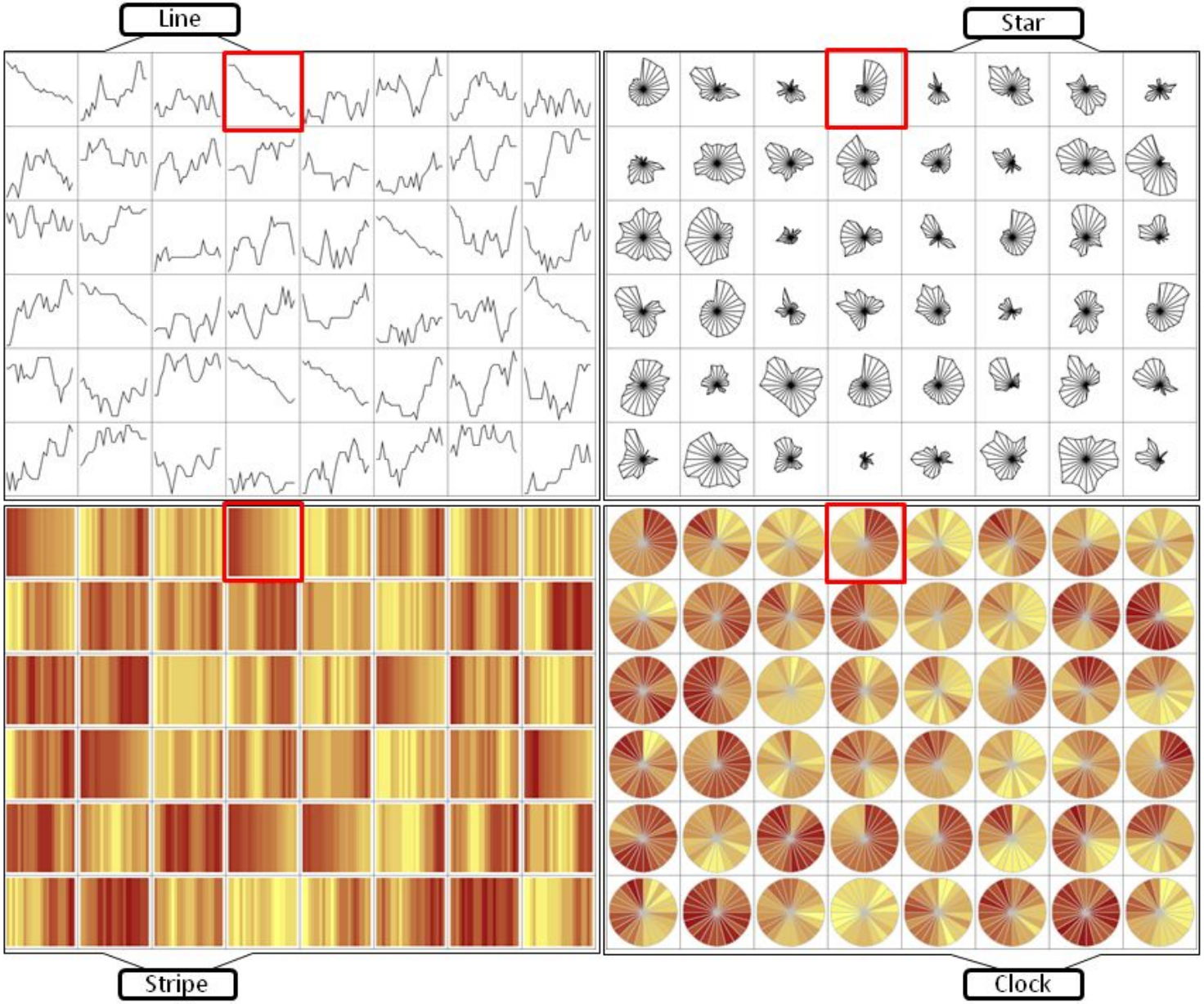


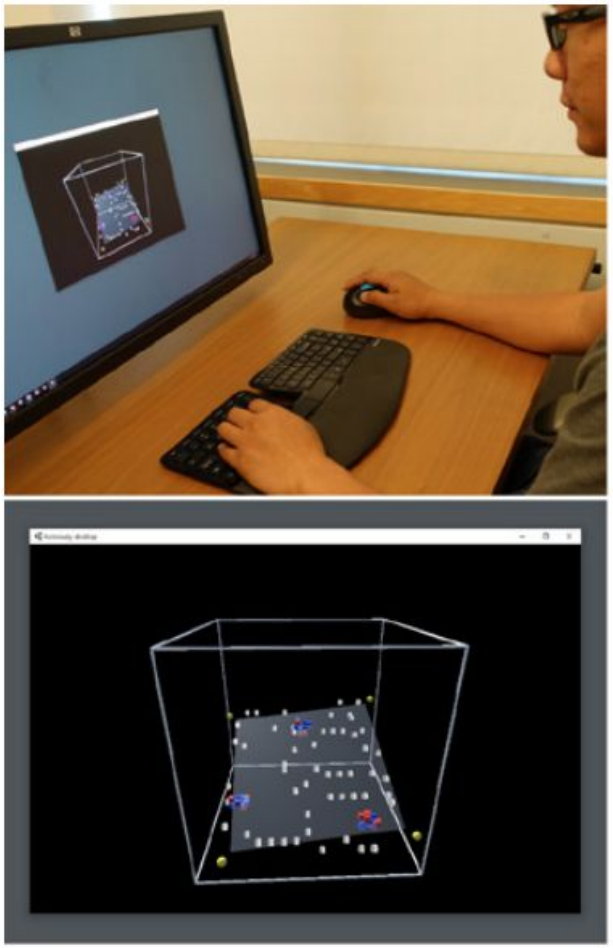
5. Observations

Observations: Think-aloud protocol

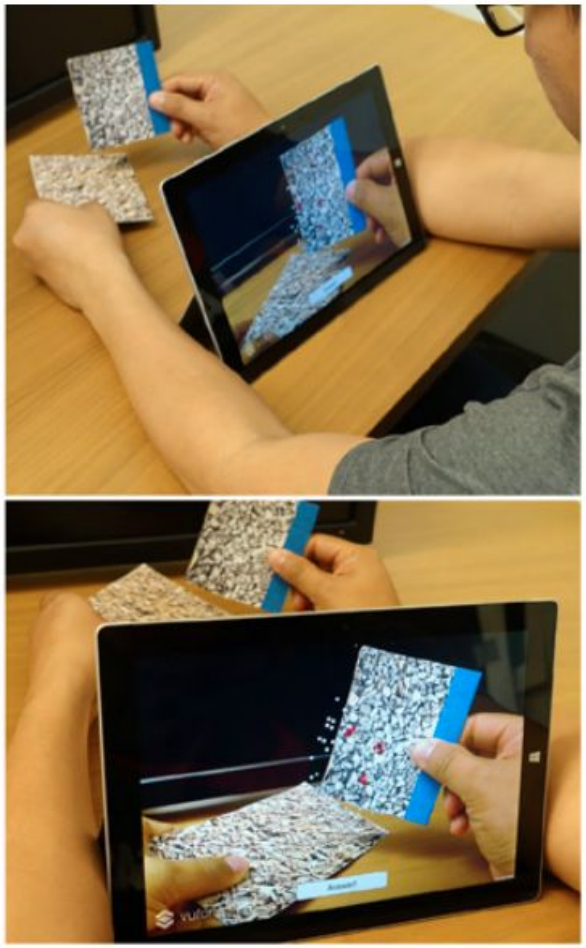
1. Give precise instructions / tasks:
   a. Which country has the highest value X.
   b. Which countries have lower values than X.
2. Let participants annotate the visualization
3. **Talk while thinking**
4. Record feedback or make notes
5. *Do not disturb*

6. Measuring User Performance

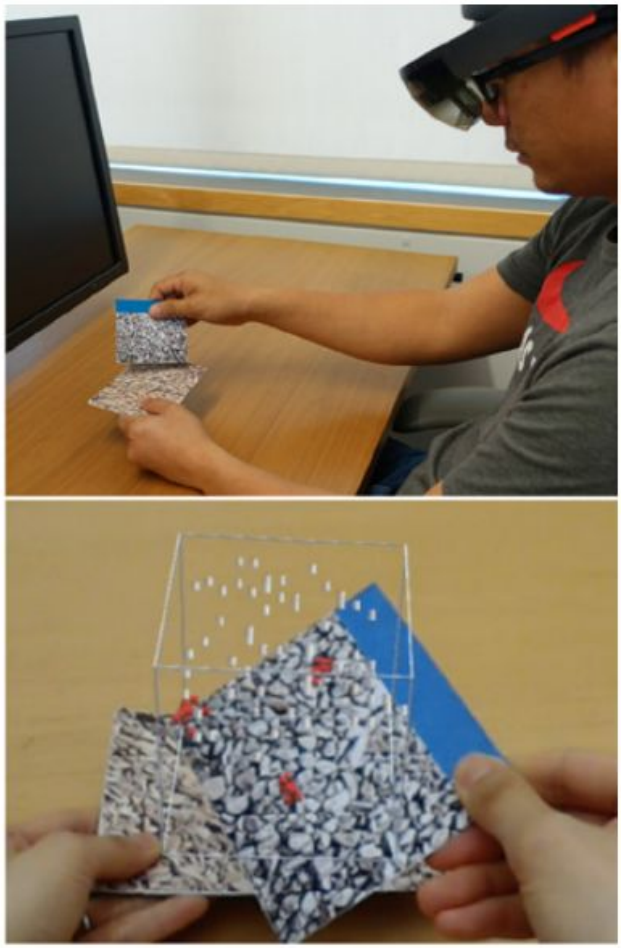

Line
Star
Stripe
Clock



Desktop



Tablet AR



Immersive AR

5. Measuring User Performance

1. Measuring:
    a. Task-completion time
    b. Task error rate
2. Reporting statistics
    a. Mean, median, mode
    b. Distributions & outliers
    c. Differences
3. > combine with subjective feedback

Evaluation methods

1. Guidelines
2. Case studies
3. Expert feedback
4. User feedback
5. Observations
6. Measuring User Performance

Preparation

- **Idea**: sketches, prototype, examples, etc.
- **Task analysis:** What do you want your visualization to support?
- **Evaluation questions:** What are you unclear about?
- **Evaluation technique:** experts, users, guidelines, etc.