Séminaire Smart Governance

Data Visualisation: efficacité, bonnes pratiques et outils

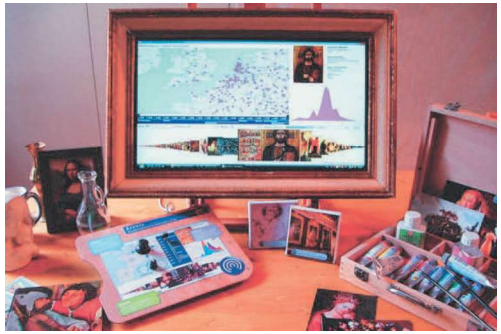
Bruno Dumas

UNIVERSITÉ DE NAMUR

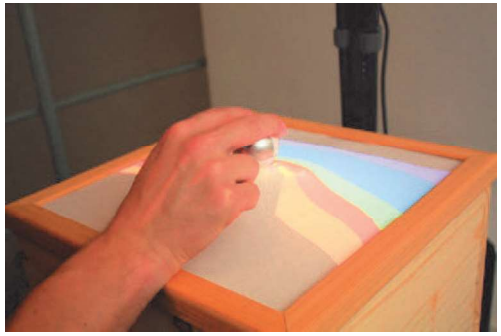
Bruno Dumas

- Associate professor @computer science faculty, University of Namur
- https://www.linkedin.com/in/brunodumas/

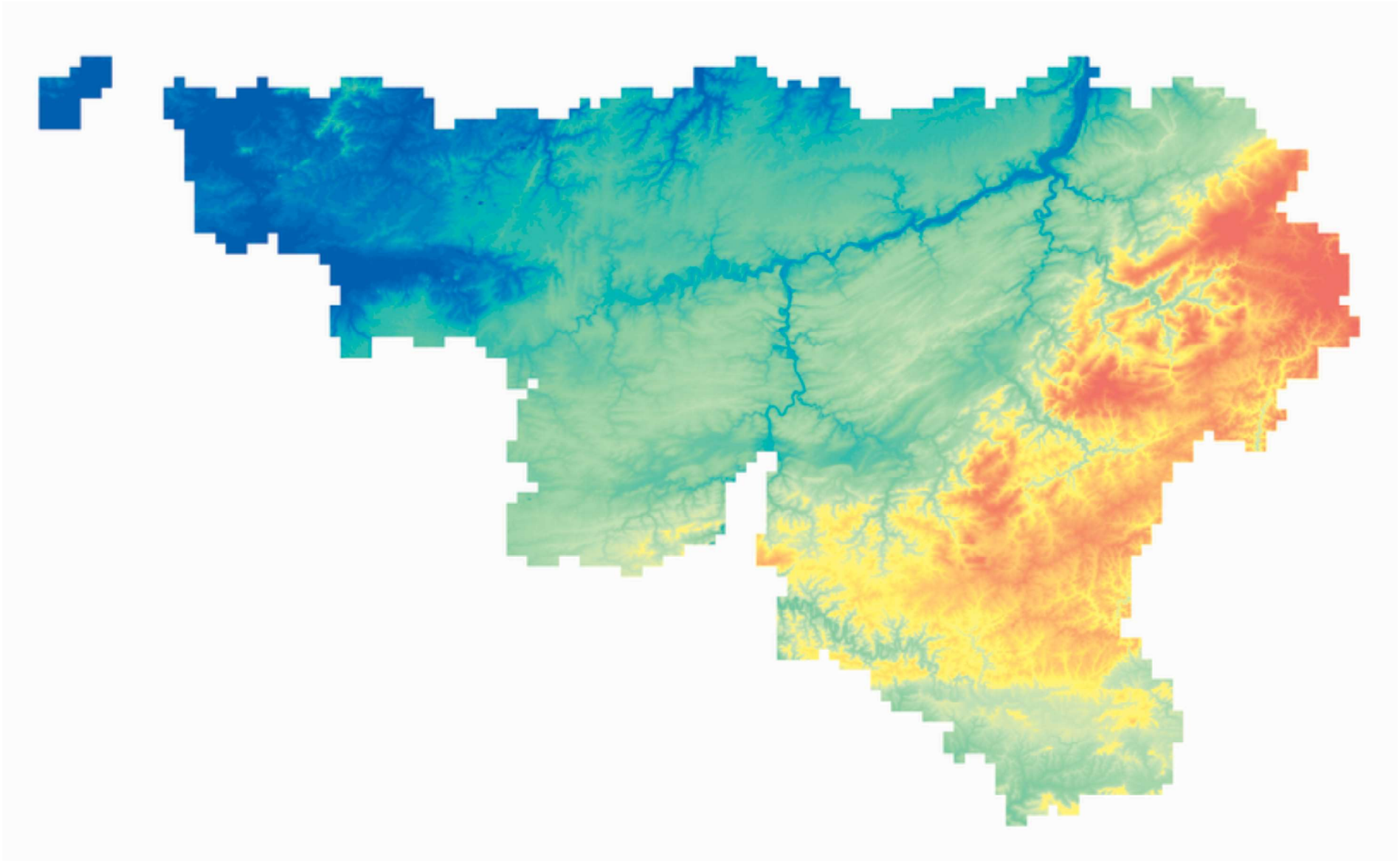
- Fields of Expertise
  - Multimodal interaction
  - Interface adaptation to user and context
  - Tangible user interfaces
  - Information visualisation
  - Augmented reality



Artvis tangible exploration tool



Infophys tangible visualisation



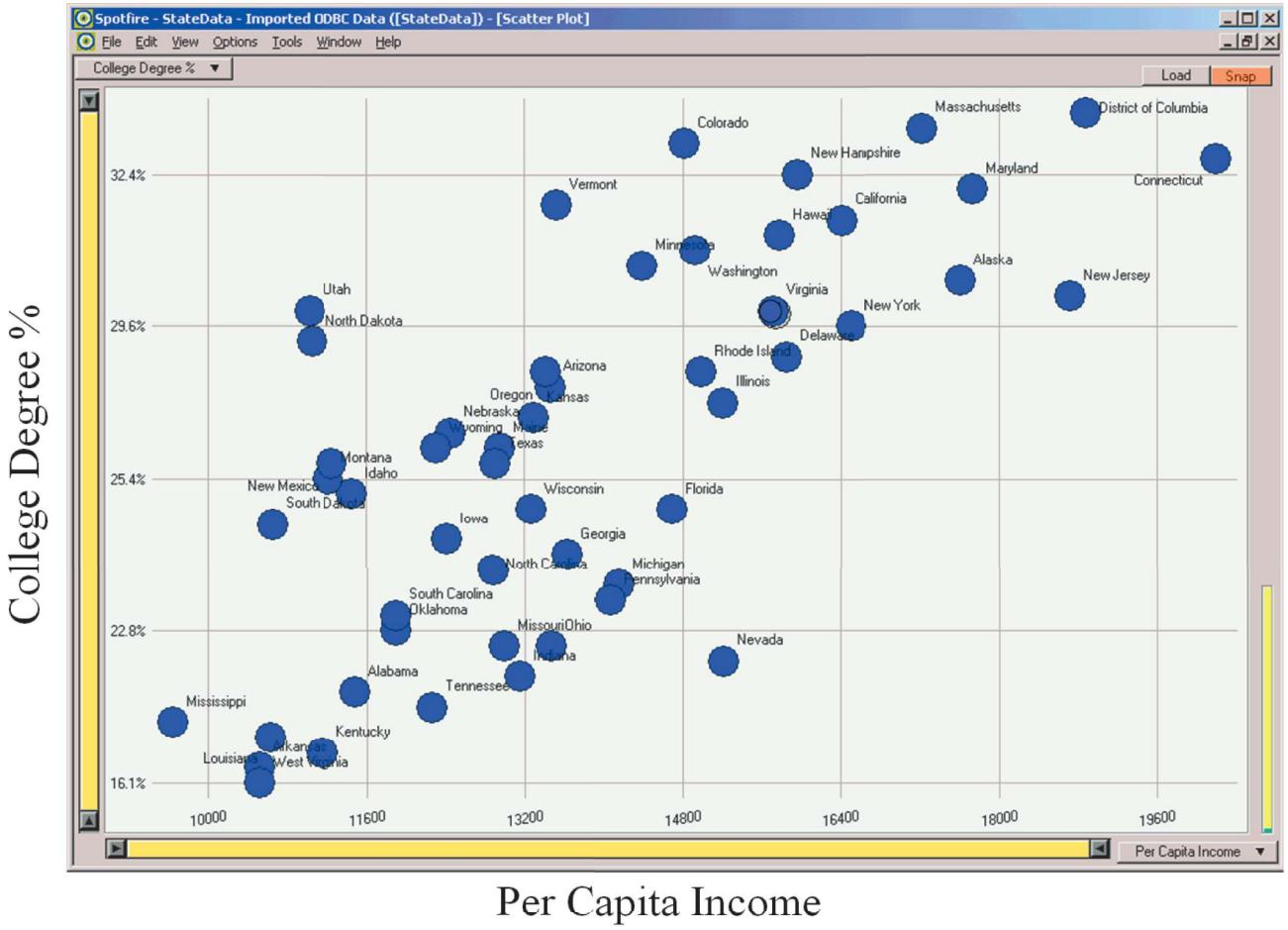
http://geoportail.wallonie.be/catalogue/6029e738-f828-438b-b10a-85e67f77af92.html

Human Vision

- Highest bandwidth sense
  - Fast, parallel
  - Pattern recognition
  - Extends memory and cognitive capacity
  - People think visually
- The Challenge
  - Transform the data into information (understanding, insight) thus making it useful to people
- Example:
  - Which state has the highest income?
  - Is there a relationship between income and education?

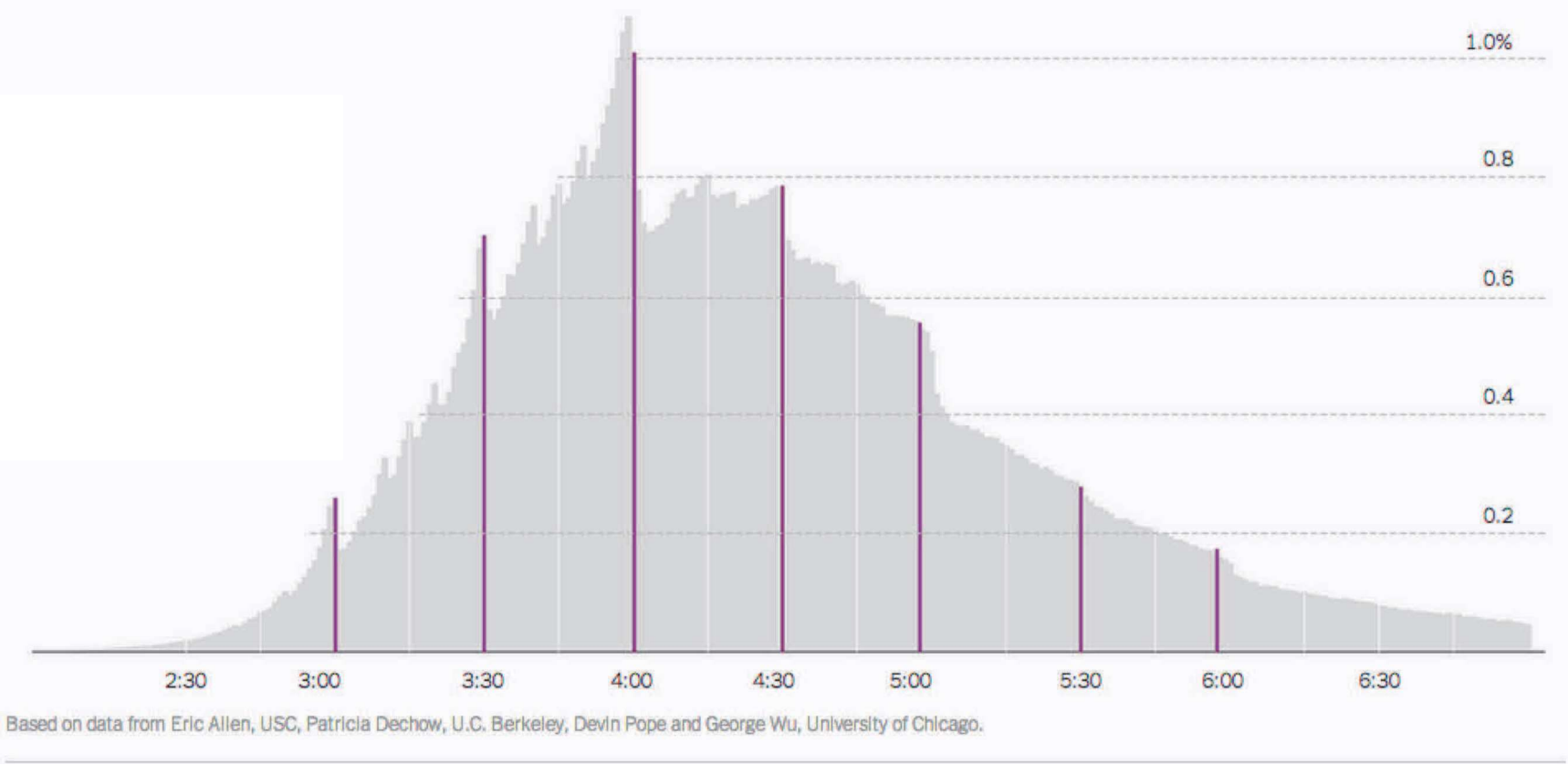
State	College Degree %	Per Capita Income
Alabama	20.6%	11486
Alaska	30.3%	17610
Arizona	27.1%	13461
Arkansas	17.0%	10520
California	31.3%	16409
Colorado	33.9%	14821
Connecticut	33.8%	20189
Delaware	27.9%	15854
District of Columbia	36.4%	18881
Florida	24.9%	14698
Georgia	24.3%	13631
Hawaii	31.2%	15770
Idaho	25.2%	11457
Illinois	26.8%	15201
Indiana	20.9%	13149
Iowa	24.5%	12422
Kansas	26.5%	13300
Kentucky	17.7%	11153
Louisiana	19.4%	10635
Maine	25.7%	12957
Maryland	31.7%	17730
Massachusetts	34.5%	17224
Michigan	24.1%	14154
Minnesota	30.4%	14389
Mississippi	19.9%	9648
Missouri	22.3%	12989
Montana	25.4%	11213
Nebraska	26.0%	12452
Nevada	21.5%	15214
New Hampshire	32.4%	15959
New Jersey	30.1%	18714
New Mexico	25.5%	11246
New York	29.6%	16501
North Carolina	24.2%	12885
North Dakota	28.1%	11051
Ohio	22.3%	13461
Oklahoma	22.8%	11893
Oregon	27.5%	13418
Pennsylvania	23.2%	14068
Rhode Island	27.5%	14981
South Carolina	23.0%	11897
South Dakota	24.6%	10661
Tennessee	20.1%	12255
Texas	25.5%	12904
Utah	30.0%	11029
Vermont	31.5%	13527
Virginia	30.0%	15713
Washington	30.9%	14923
West Virginia	16.1%	10520
Wisconsin	24.9%	13276
Wyoming	25.7%	12311

Human Vision (cont.)

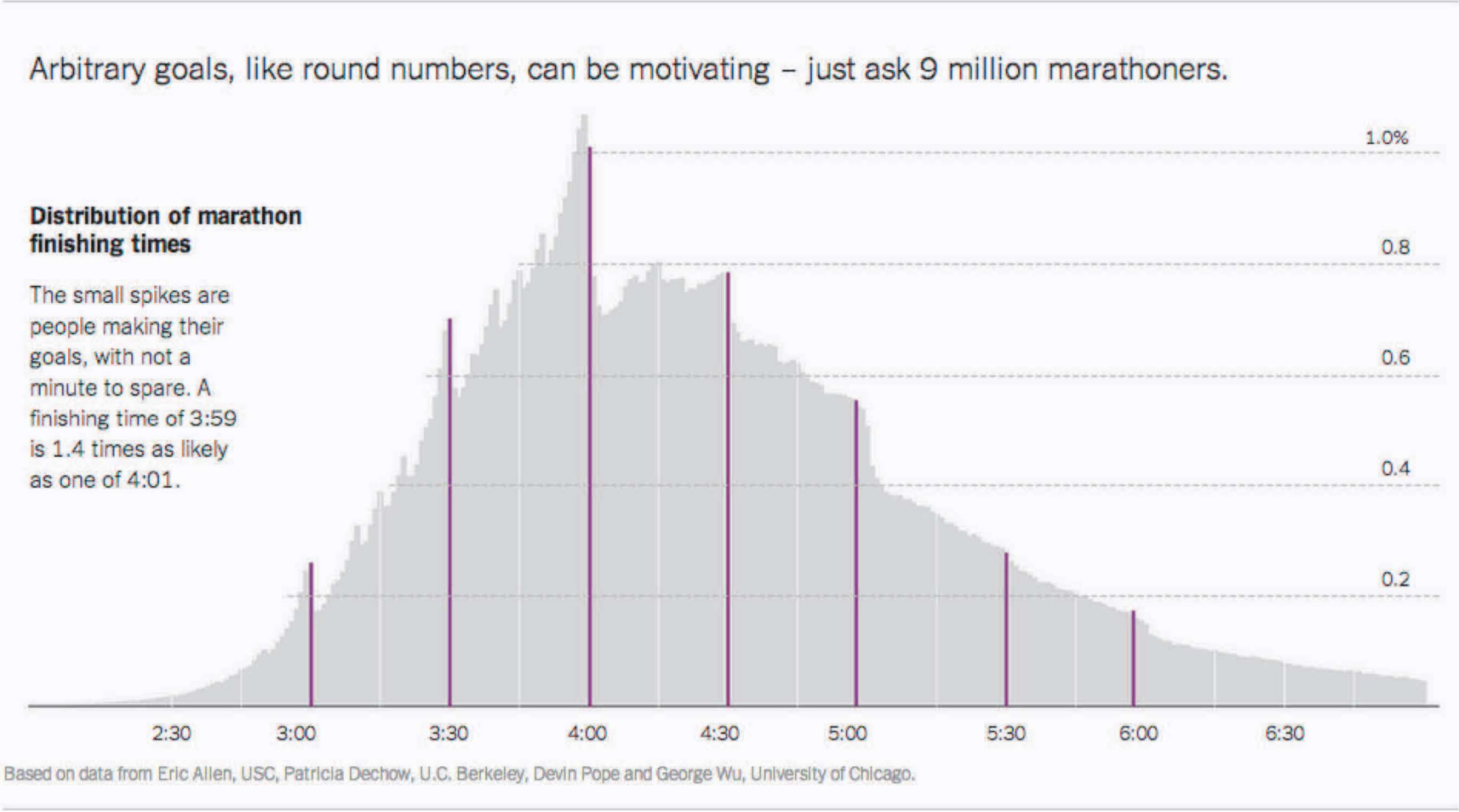


Use Case: A (Marathonian) Example

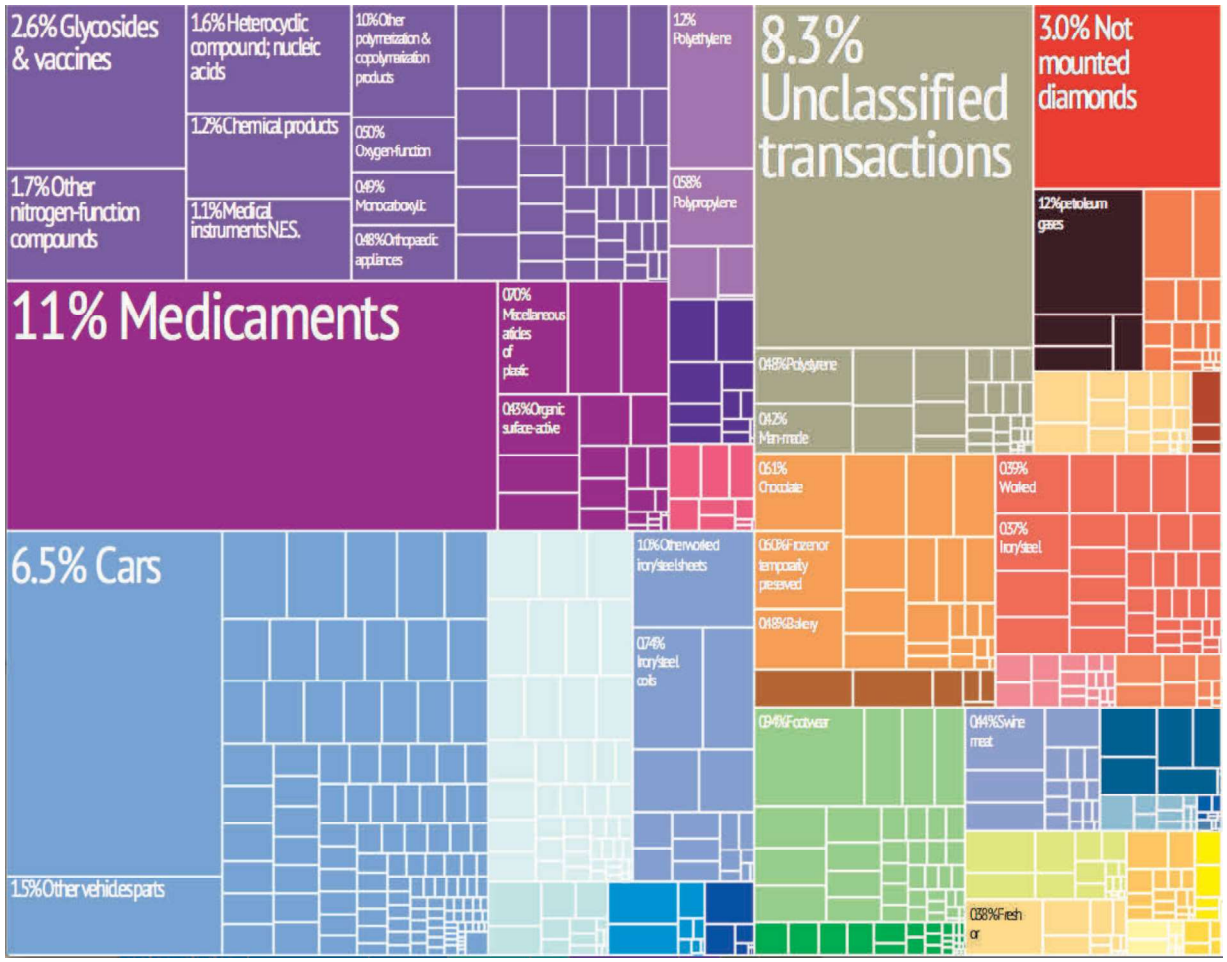
Use Case: A (Marathonian) Example



Use Case: A (Marathonian) Example

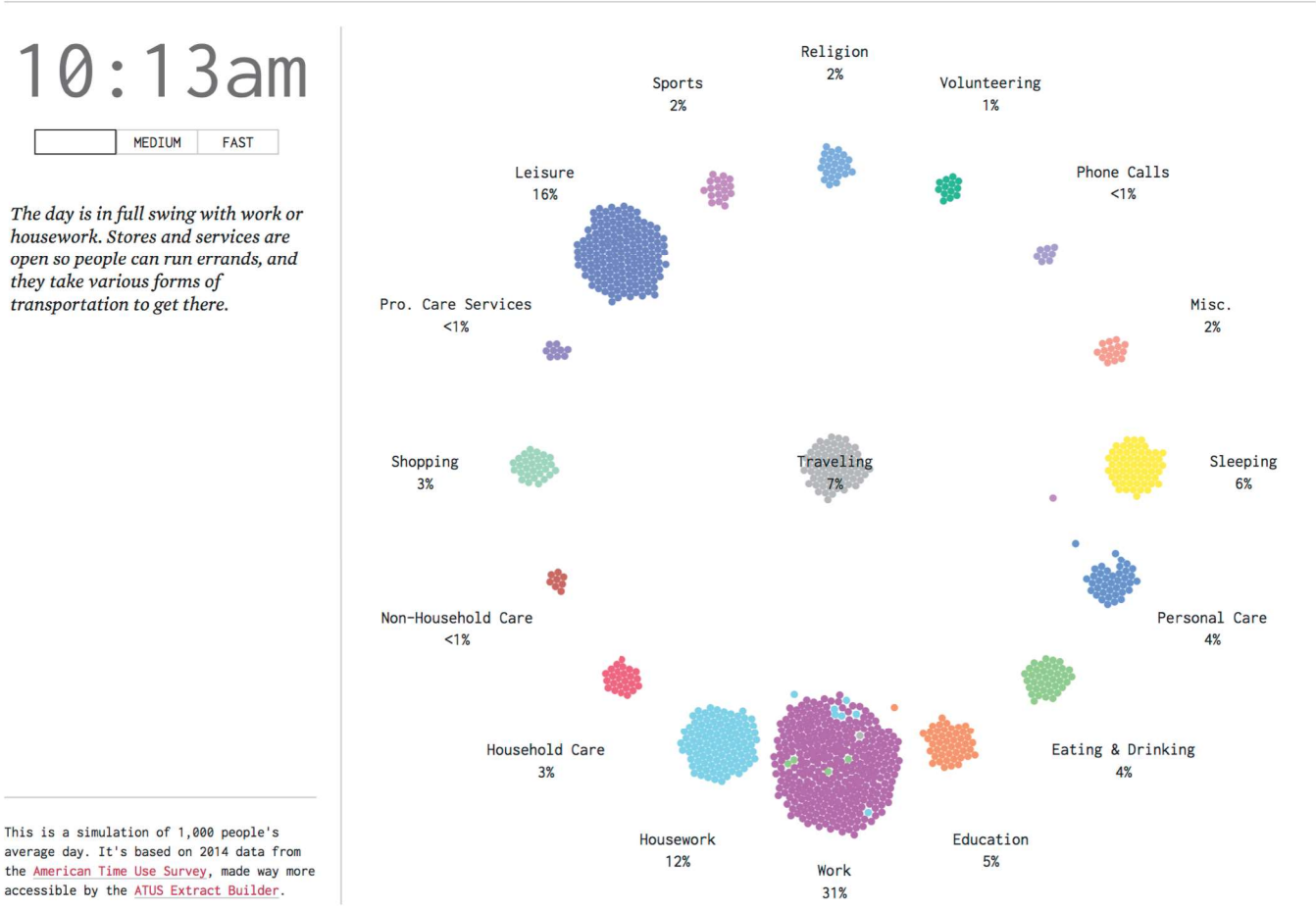


Use Case: 2009 Exports of Belgium



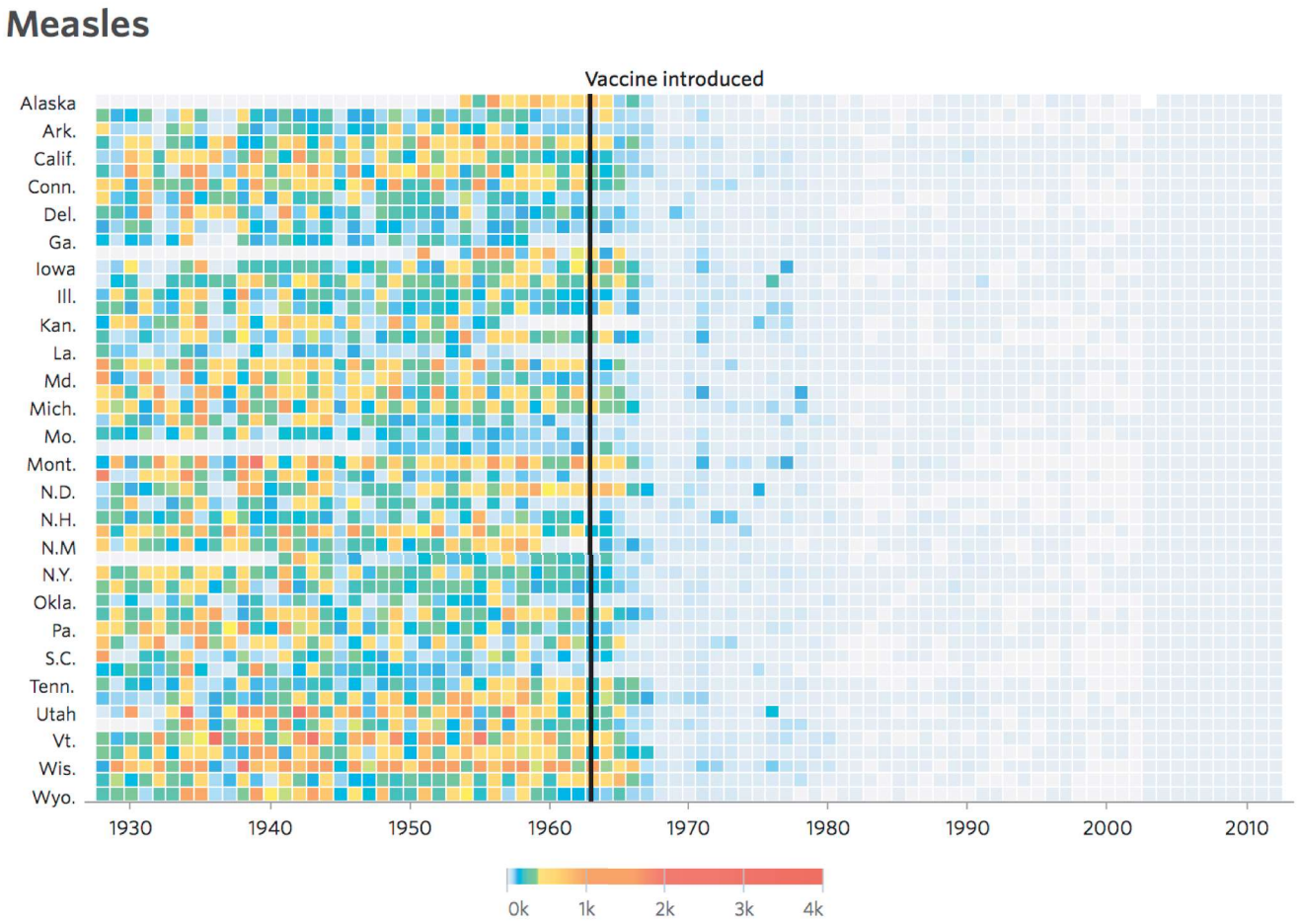
https://commons.wikimedia.org/wiki/File:Tree_map_export_2009_Belgium.jpeg

Use Case: A Day In the Life of Americans



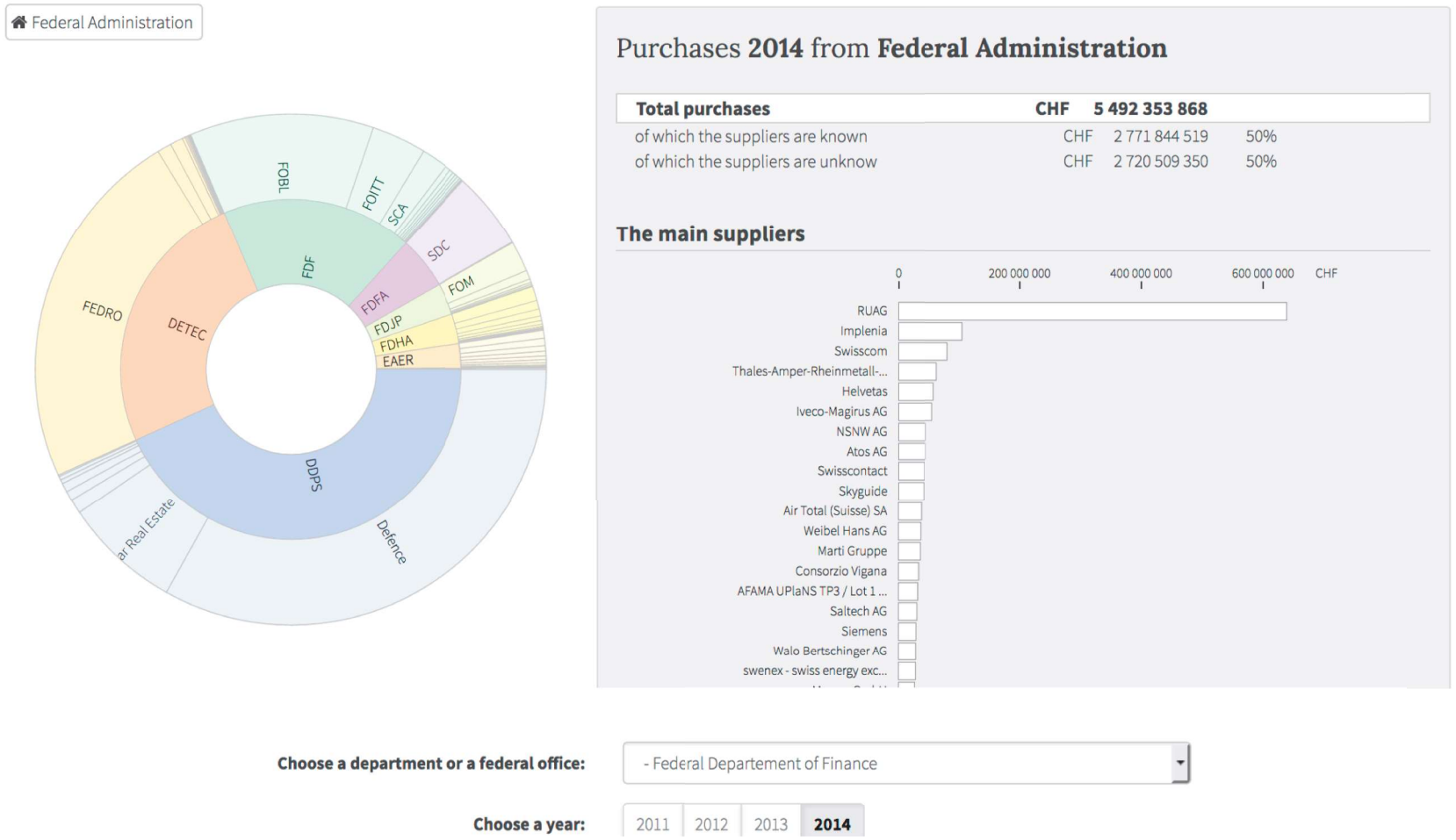
http://flowingdata.com/2015/12/15/a-day-in-the-life-of-americans/

Use Case: WSJ Infectious Diseases and Vaccines



http://graphics.wsj.com/infectious-diseases-and-vaccines/

Use Case: Purchases in 2014 from Swiss Federal Administration



http://enquete.lematindimanche.ch/interactif/achats/indexen.html?lang=en

Roles of Infovis

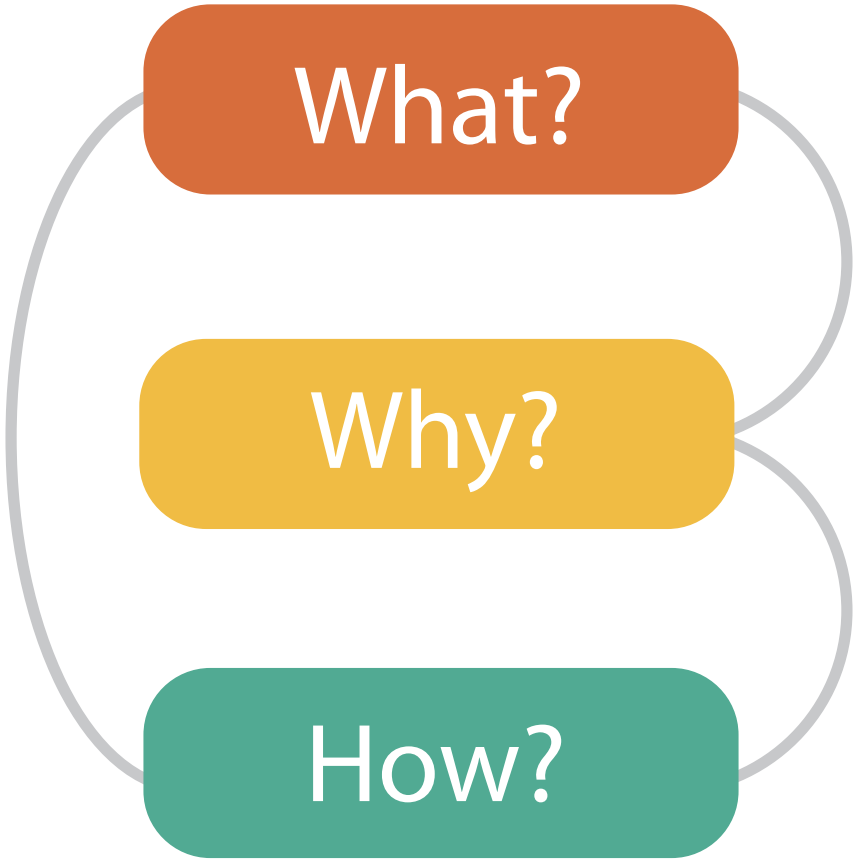
- Exploration
  - Prerequisite: domain knowledge
  - Outcome: new hypothesis
- Confirmation
  - Prerequisite: hypothesis
  - Outcome: confirmation/rejection (new hypothesis)
- Communication
  - Prerequisite: confirmed hypothesis
  - Outcome: clear visualisation
- Visual Analytics' Motto: *detect the expected and discover the unexpected*

Keim, Daniel A. "Information visualization and visual data mining." IEEE transactions on Visualization and Computer Graphics 8, no. 1 (2002): 1-8.

Interactivity

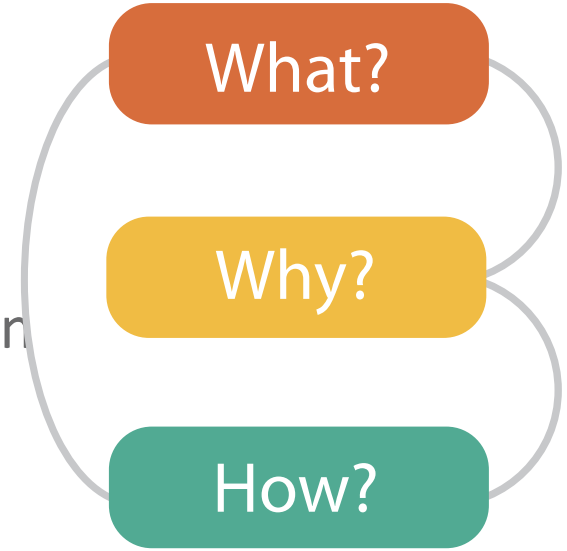
- Interactivity is crucial for handling complexity
  - Typically when data is large enough…
- Interactivity involves
  - Changing representations
  - Zooming
  - Filtering
  - Getting details
  - Summarizing
  - Combining views
  - …

A Three-Part Analysis Framework

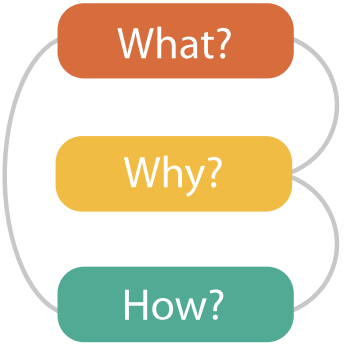


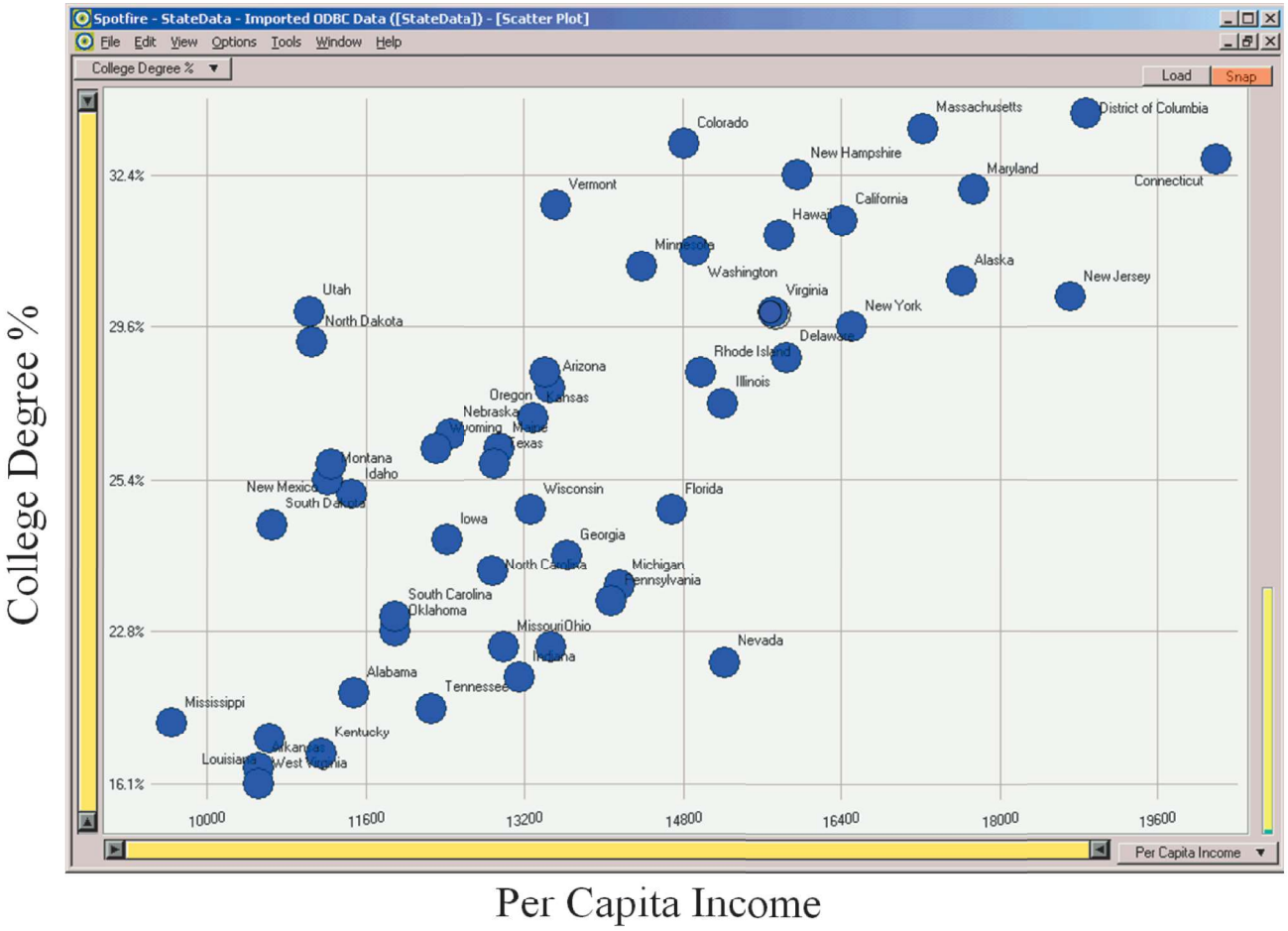
A Three-Part Analysis Framework

- WHAT data the user sees
  - Aka the data
- WHY the user intends to use a vis tool
  - Aka the task
- HOW the visual encoding and interaction idiom are constructed
  - Aka the visual idiom and interaction



A Simple Example:





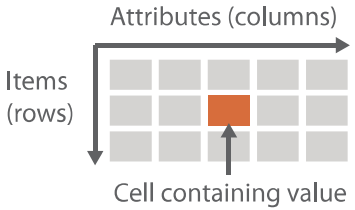
1. Data

- Datasets
  - Data types
  - Data and dataset types
  - Dataset types
  - Dataset availability
- Attributes
  - Attribute types
  - Ordering Direction

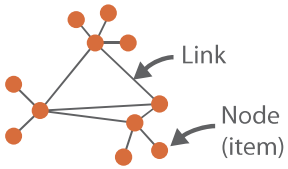
Datasets: Dataset Types

Dataset Types

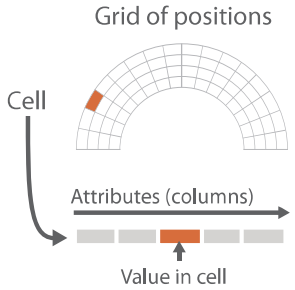
→ Tables



→ Networks

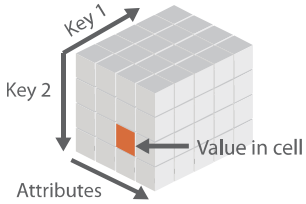


→ Fields (Continuous)

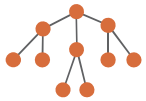


→ Geometry (Spatial)

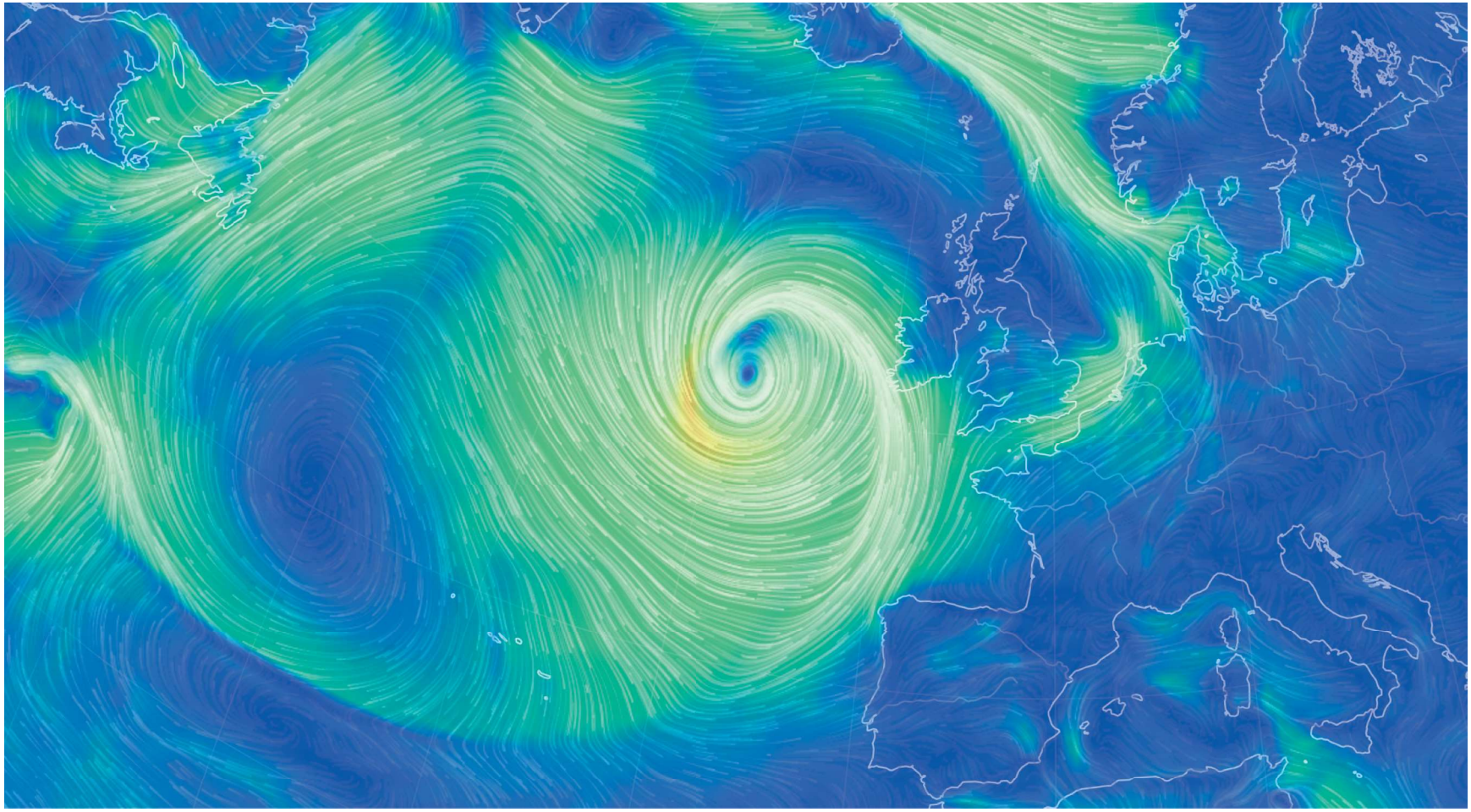
→ *Multidimensional Table*



→ *Trees*



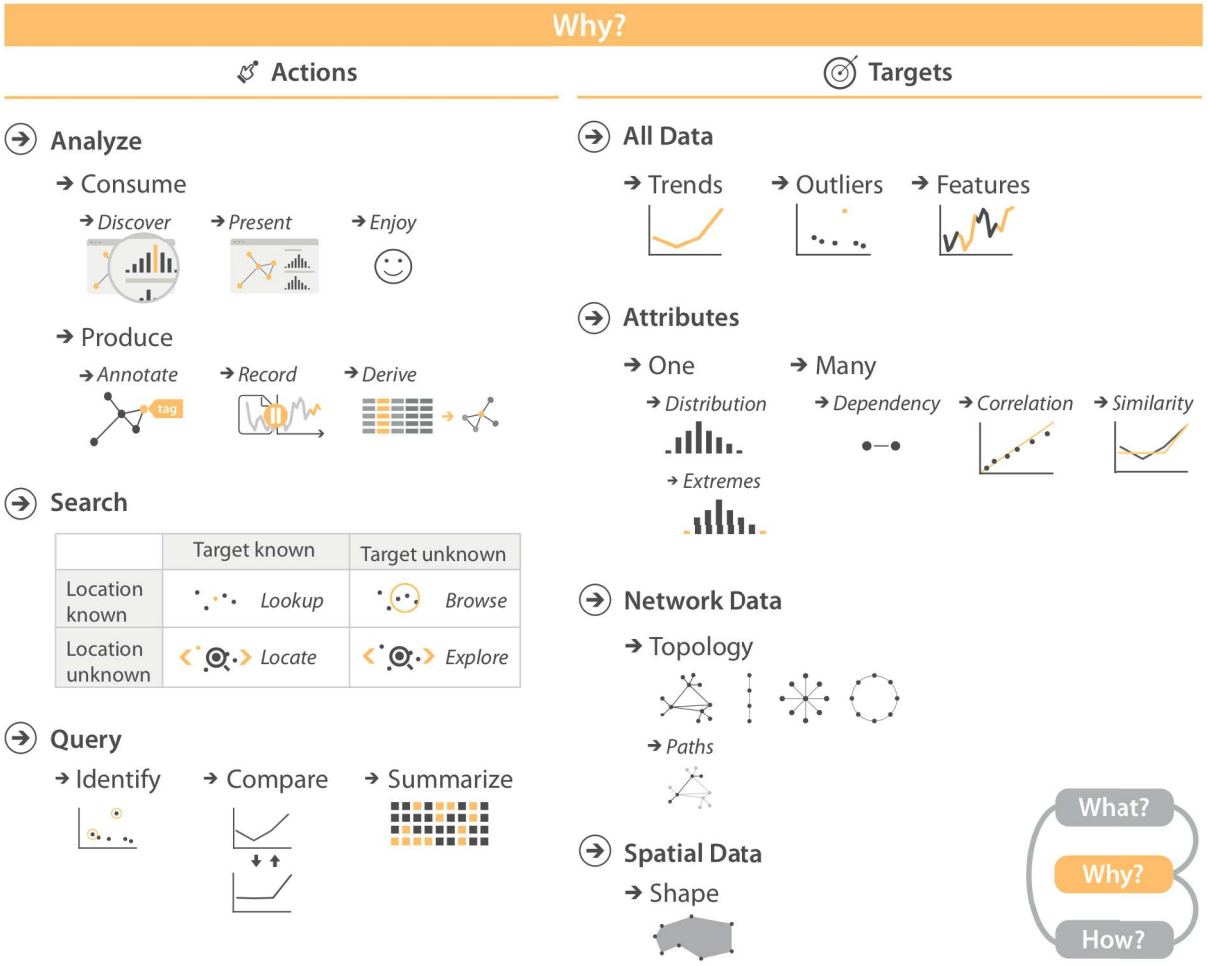
An Example



https://earth.nullschool.net/#current/wind/surface/level/orthographic=-0.61,47.21,1137/loc=26.765,35.377

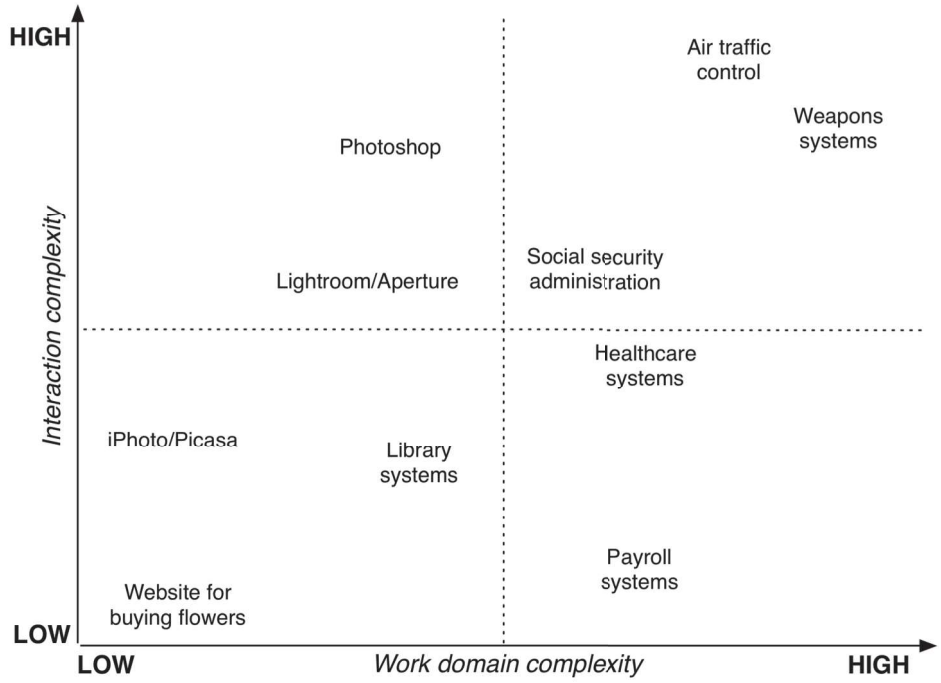
2. Task Abstraction

- Why analyse?
- Who?
- Actions (verbs):
  - Analyse
  - Produce
  - Search
  - Query
- Targets (nouns):
  - All data
  - Attributes
  - Network data
  - Spatial data



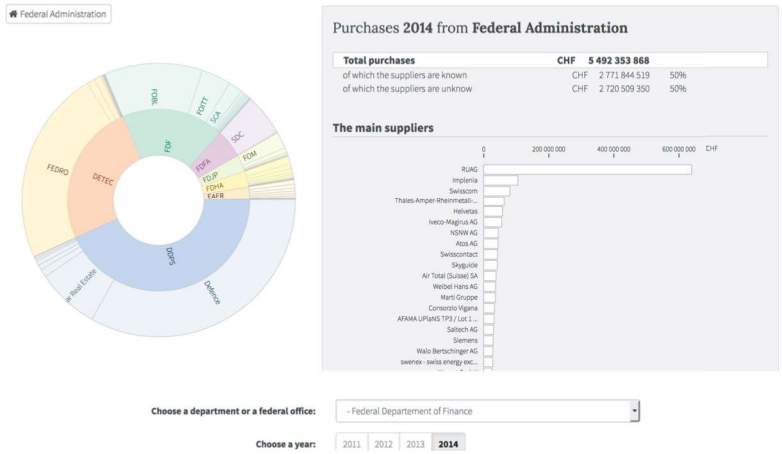
Who Is the Audience?

- Domain expert? Average web user? Namur resident? You?

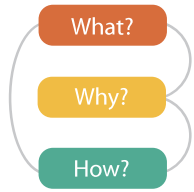


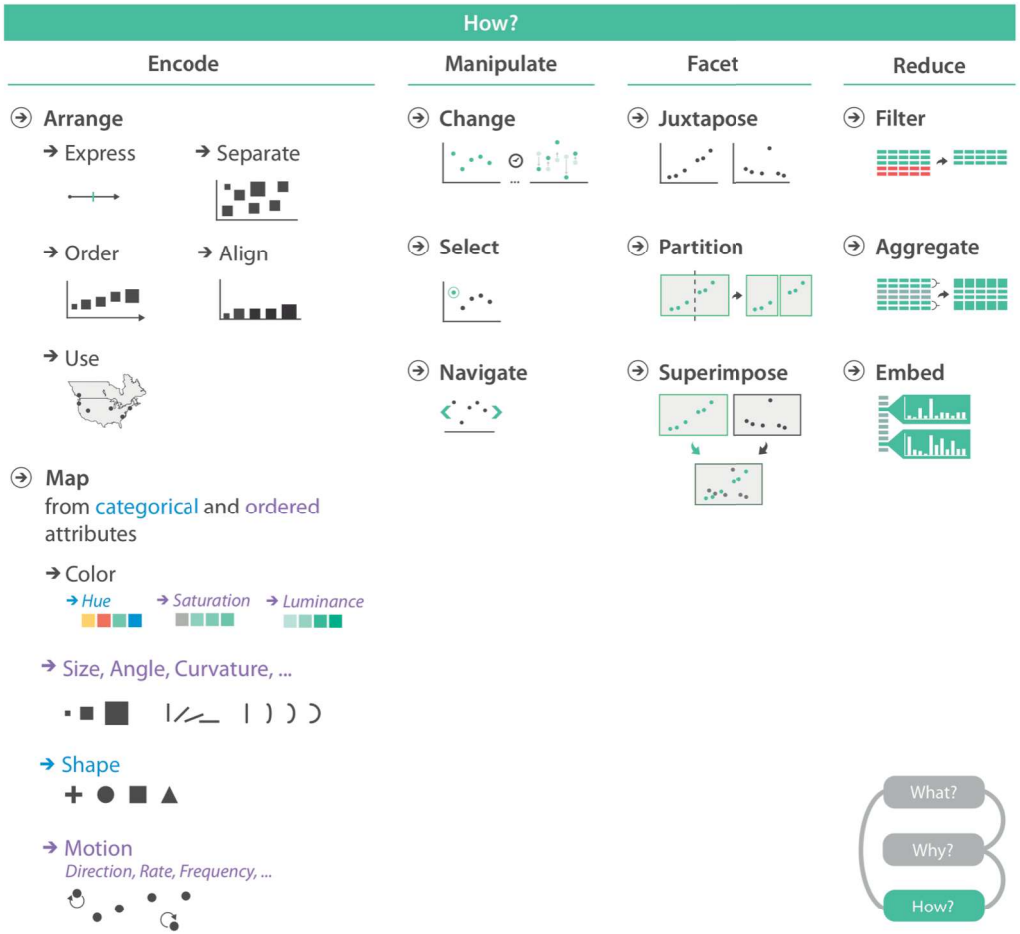
Who Makes the Choices?

- Part of the analysis: who will be the main user of the visualisation?
- Corollary of that question: who refines the design?
  - Filters data?
  - Selects representation?
  - Focuses on a specific part of the data?
  - Offline choices (designer) vs. Online choices (user)
  - … Vs. both

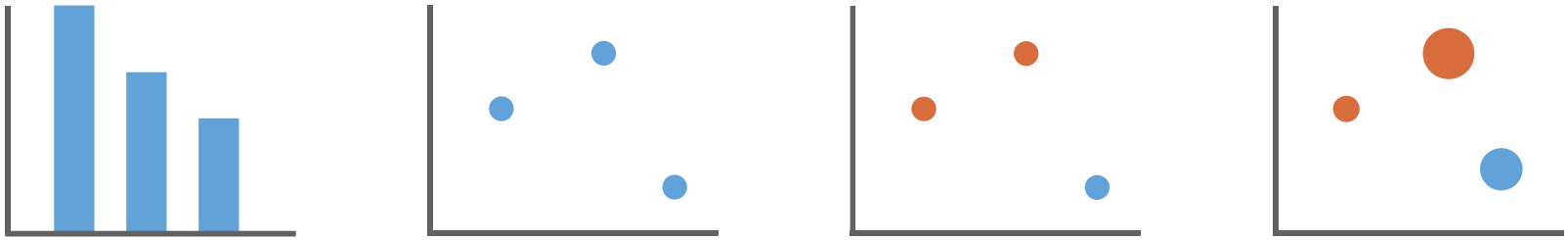


How?...


What?
Why?
How?


How?
Encode
Manipulate
Facet
Reduce
Arrange
Express
Separate
Order
Align
Use
Map
from categorical and ordered attributes
Color
Hue
Saturation
Luminance
Size, Angle, Curvature, ...
Shape
Motion
Direction, Rate, Frequency, ...
Change
Select
Navigate
Juxtapose
Partition
Superimpose
Filter
Aggregate
Embed
What?
Why?
How?

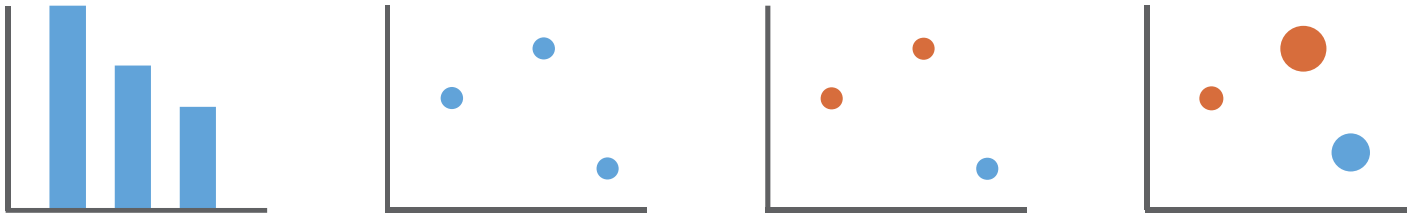
An Introductory Example



How many encoded attributes?

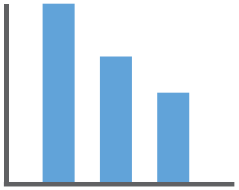
An Introductory Example

- In this example…
  - 1) Two attributes (X categorical/Y quantitative)
  - 2) Two attributes (X quantitative/Y quantitative)
  - 3) Three attributes (X quantitative/Y quantitative/colour)
  - 4) Four attributes (X quantitative/Y quantitative/colour/size)
- Each attribute can convey one dimension…
- … Or you can introduce redundancy between attributes…
- … Some channels are more efficient for certain attributes…
  - E.g. bars in bar chart convey a sense of size better than scatterplots

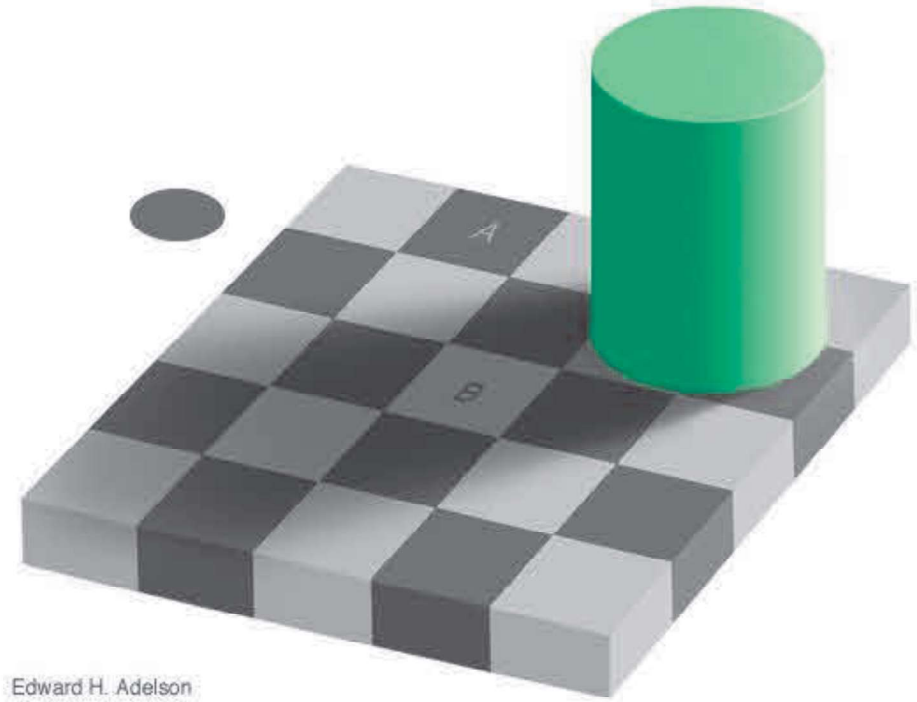


Channel Types

- Two main types of channels: identity and magnitude
  - Identity: tells us information about *what* something is or *where* it is
  - Magnitude: tells us *how much* of something there is
- Identity example: shapes, hue, motion pattern…
- Magnitude example: line length, luminance, size…
- Linked to how the human perceptual system works
- -> All channels are not equal!

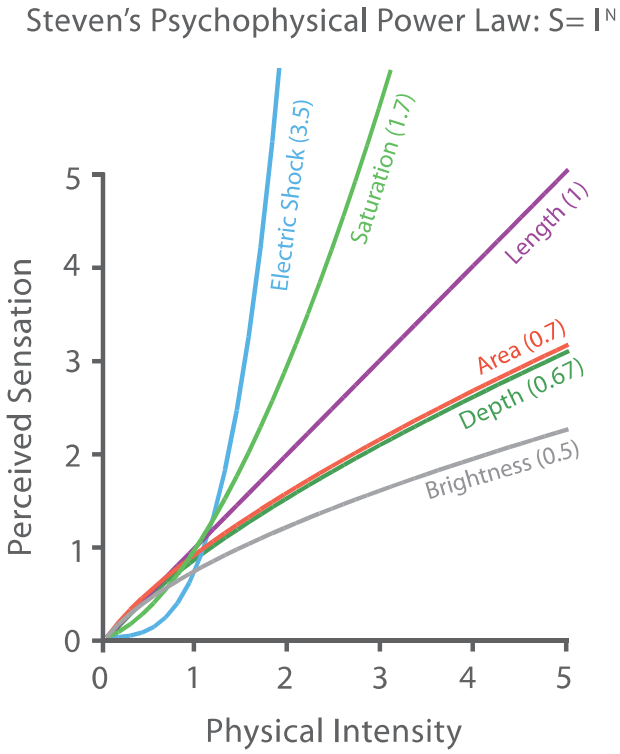


Break: A Small Brightness Test…



Channel Effectiveness: Accuracy

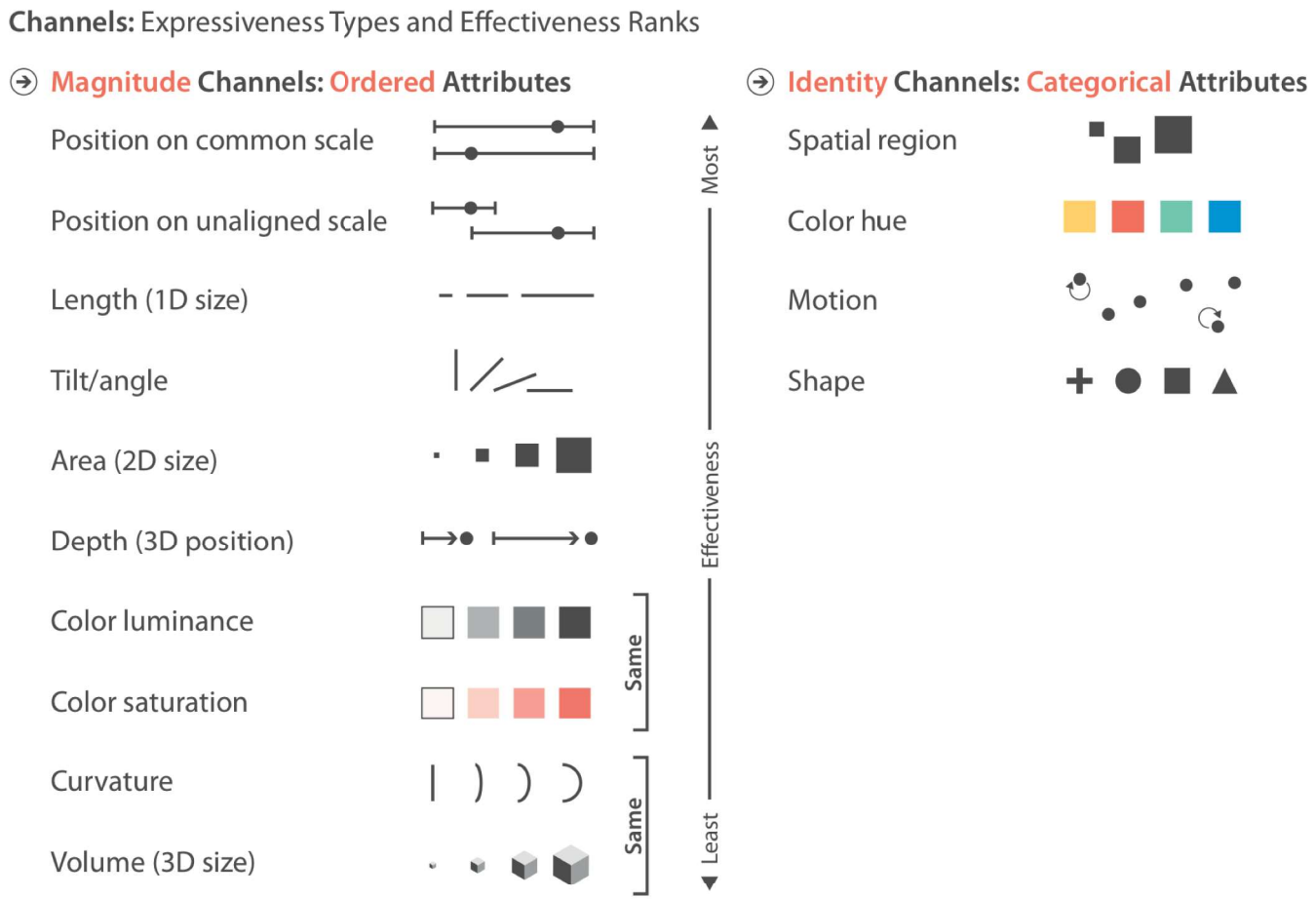
- Accuracy to assess channel effectiveness w.r.t. perception
- Demonstrated as exponential law



S: perceived sensation
I: physical intensity
N: dependent of channel

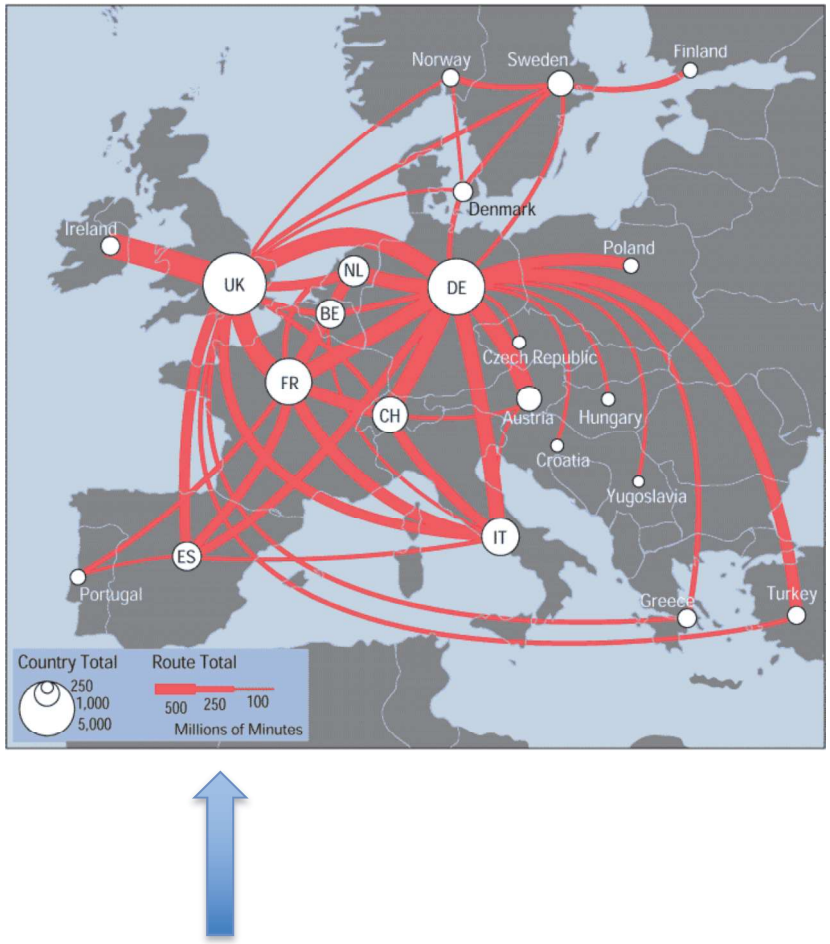
Stevens, S. S. (2017). Psychophysics: Introduction to its perceptual, neural and social prospects. Routledge.

Channel Effectiveness: Ranking of channels



Channel Effectiveness: Discriminability

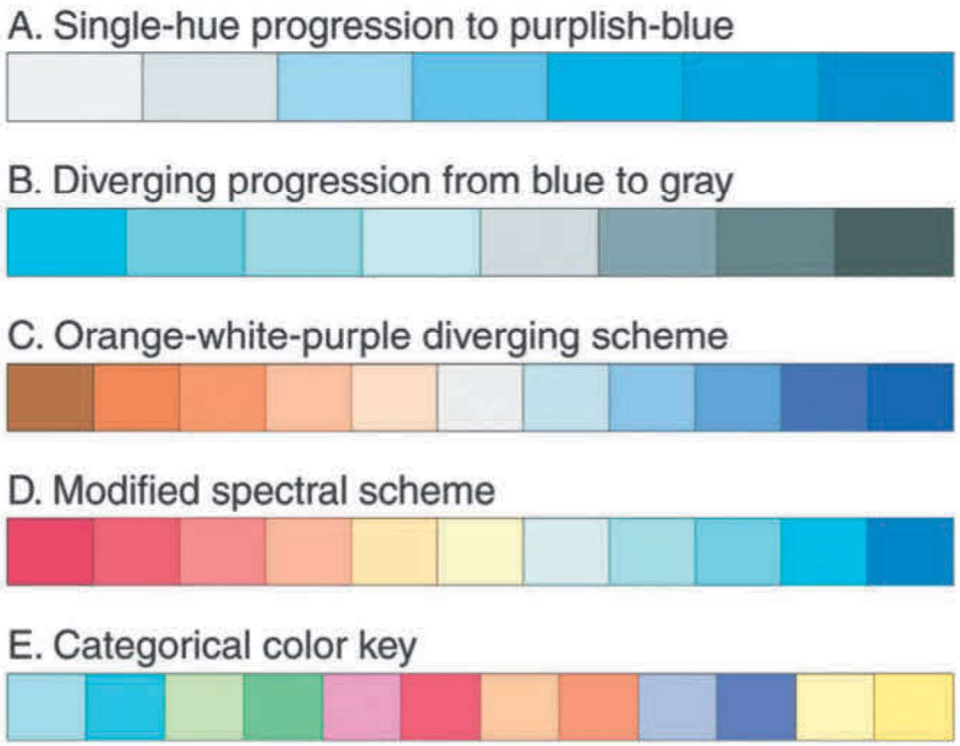
- Discrimanibility: are the differences between items perceptible to the user?
- Concept of bin = distinguishable level within a visual channel
  - Exemple with line width: how many levels can you spot on this image?



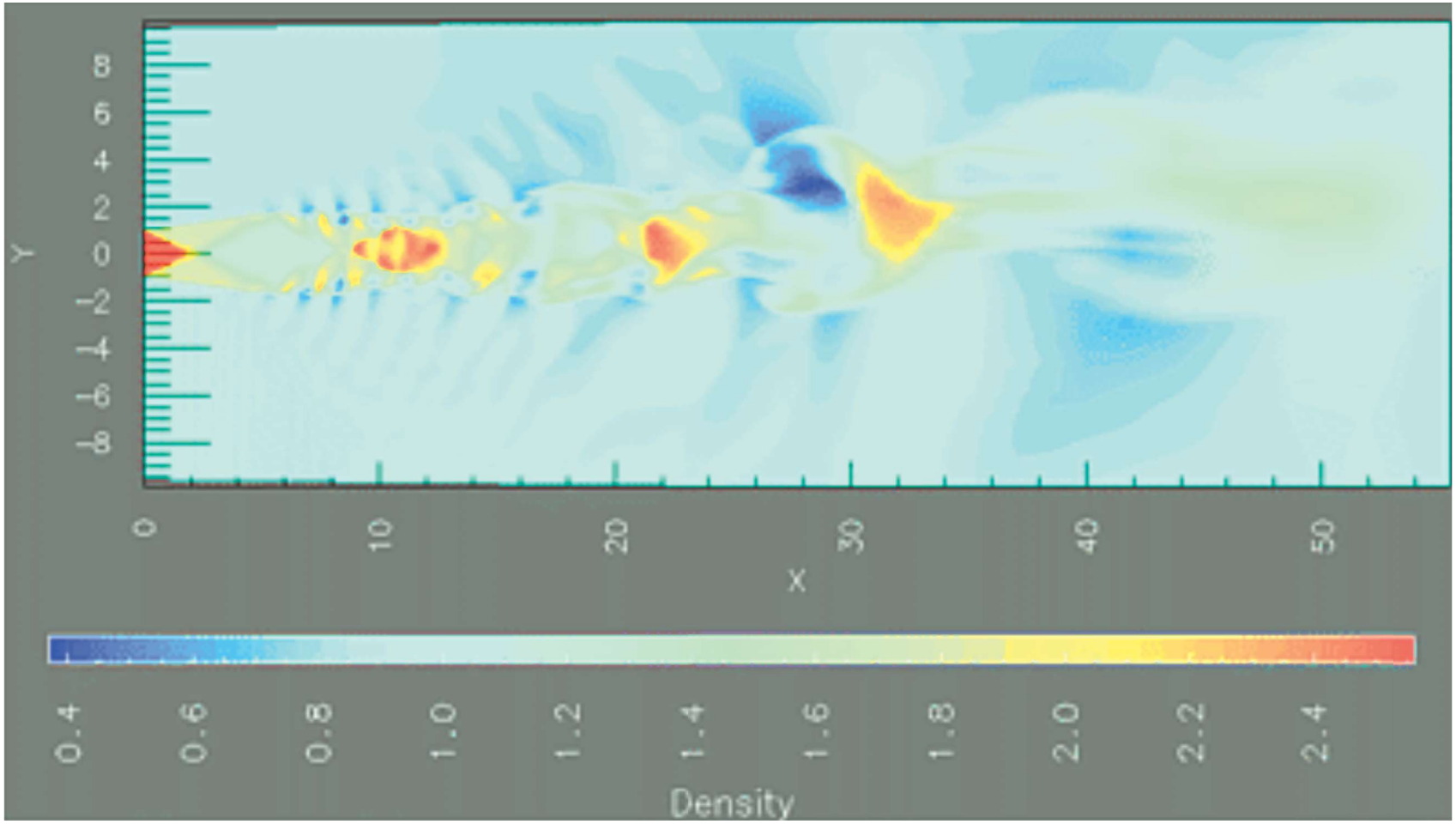
https://mappa.mundi.net/maps/maps_014/telegeography.html

Colourmaps

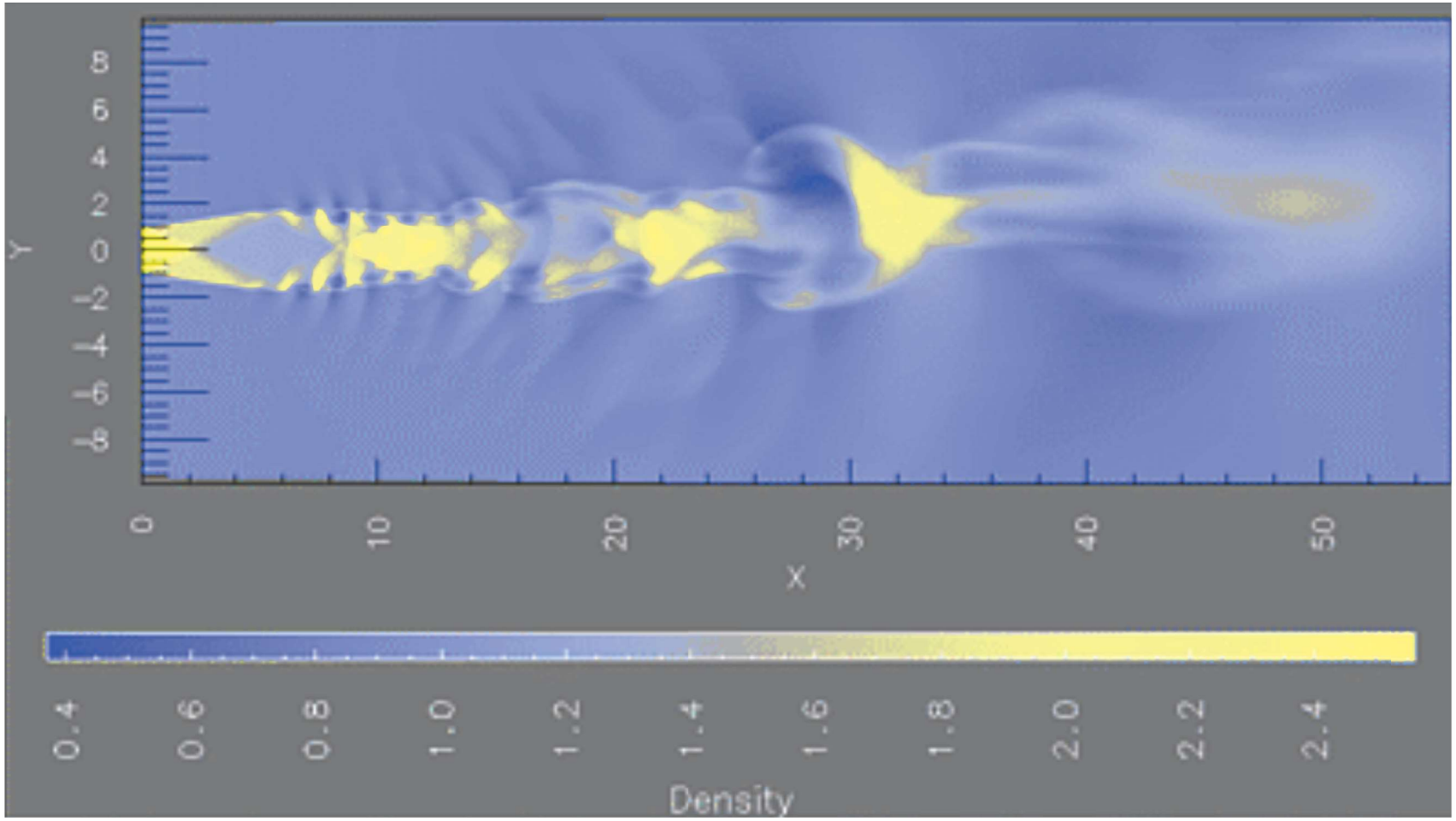
- Colourmap = mapping between colours and data values
  - ~ visual encoding with colour



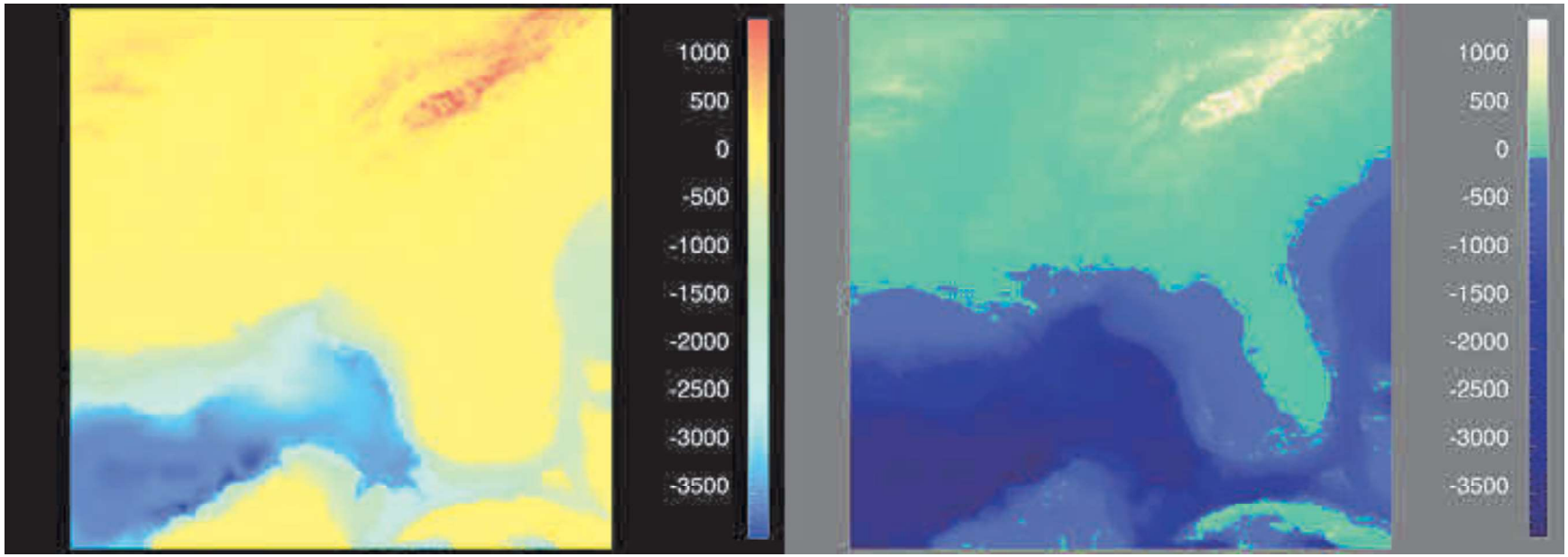
Ordered Colourmap Example



Alternative Colourmap for the Example Before

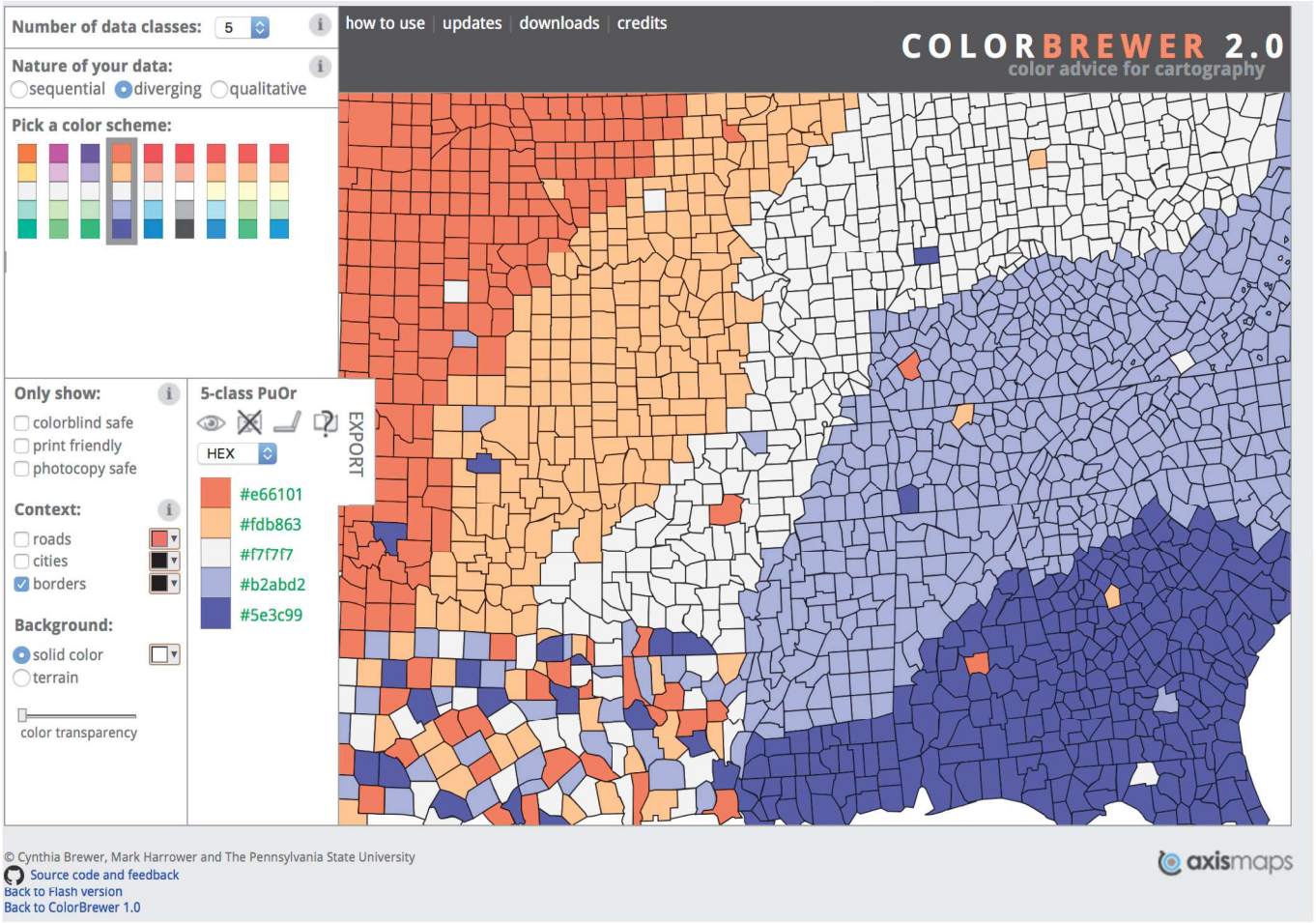


Ordered Colourmaps Example – Rainbow vs. Carefully Designed Colourmap



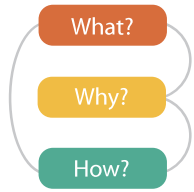
https://www.research.ibm.com/people/l/lloydt/color/color.HTM

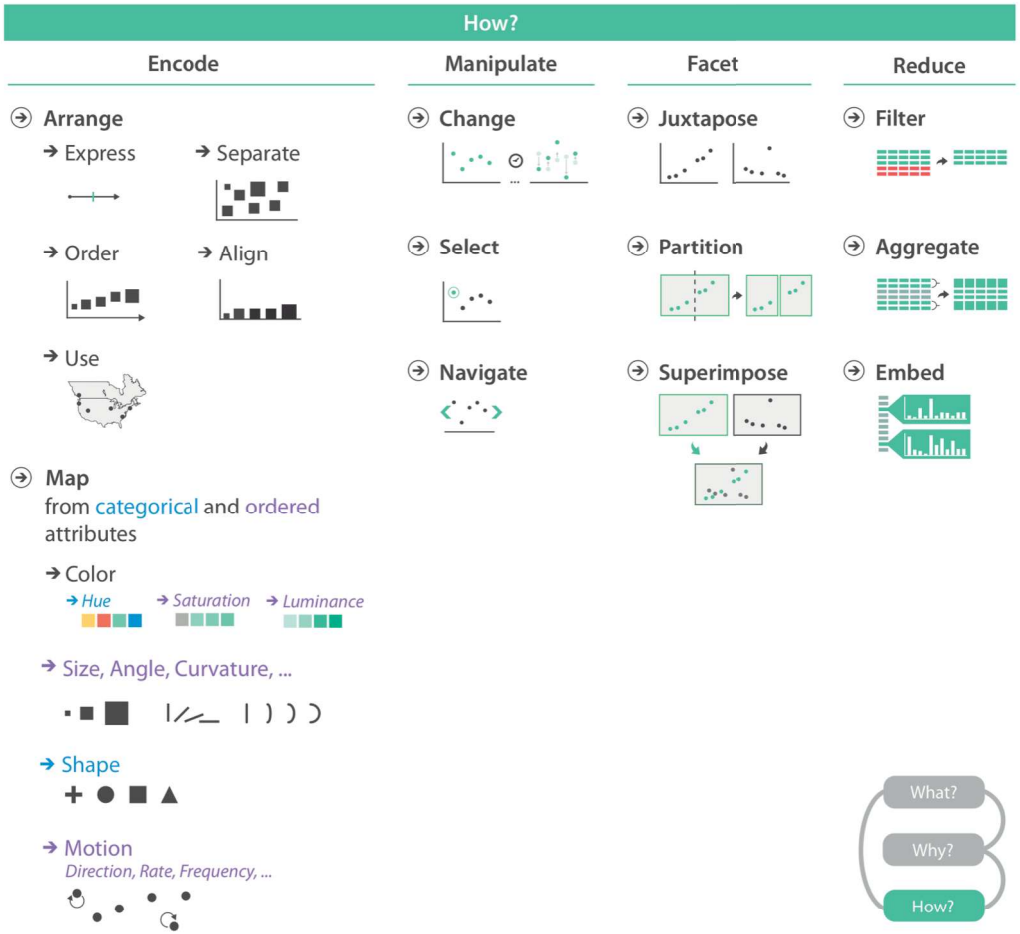
Categorical Colourmaps: Colorbrewer



http://colorbrewer2.org/

How?...


What?
Why?
How?


How?
Encode
Manipulate
Facet
Reduce
Arrange
Express
Separate
Order
Align
Use
Map
from categorical and ordered attributes
Color
Hue
Saturation
Luminance
Size, Angle, Curvature, ...
Shape
Motion
Direction, Rate, Frequency, ...
Change
Select
Navigate
Juxtapose
Partition
Superimpose
Filter
Aggregate
Embed
What?
Why?
How?

Examples of Visualisation Techniques

https://github.com/d3/d3/wiki/Gallery

Tools

- Two main approaches:
  - Dashboard-like tools
  - Visualisation libraries

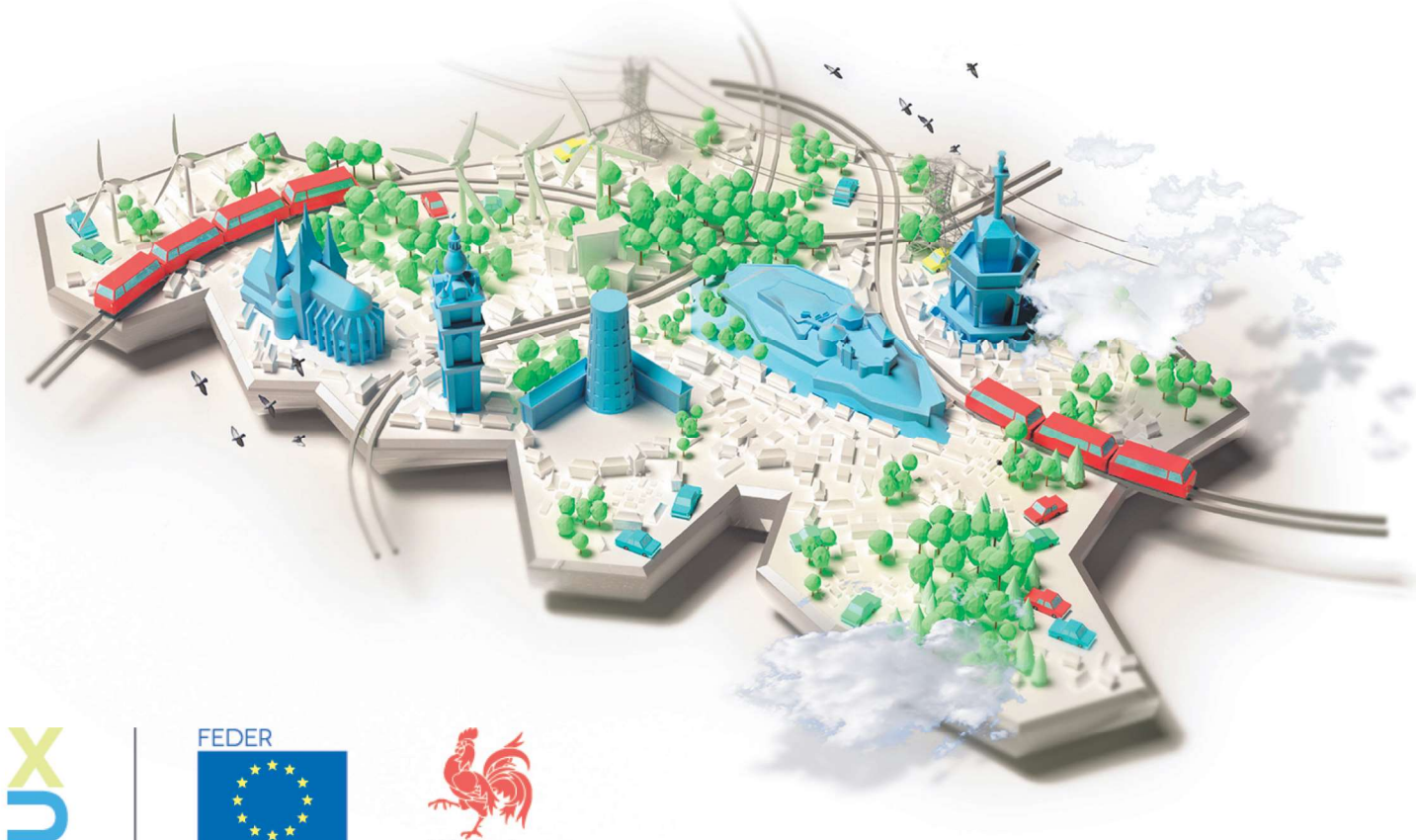

```
<!DOCTYPE html>
<meta charset="utf-8">
<style>

.axis--x path {
  display: none;
}

.line {
  fill: none;
  stroke: steelblue;
  stroke-width: 1.5px;
}

</style>
<svg width="960" height="500"></svg>
<script src="//d3js.org/d3.v4.min.js"></script>
<script>

var svg = d3.select("svg"),
    margin = {top: 20, right: 80, bottom: 30, left: 50},
    width = svg.attr("width") - margin.left - margin.right,
    height = svg.attr("height") - margin.top - margin.bottom,
    g = svg.append("g").attr("transform", "translate(" + margin.left + "," + margin.top + ")");

var parseTime = d3.timeParse("%Y%m%d");

var x = d3.scaleTime().range([0, width]),
    y = d3.scaleLinear().range([height, 0]),
    z = d3.scaleOrdinal(d3.schemeCategory10);

var line = d3.line()
    .curve(d3.curveBasis)
    .x(function(d) { return x(d.date); })
    .y(function(d) { return y(d.temperature); });

d3.tsv("data.tsv", type, function(error, data) {
  if (error) throw error;

  var cities = data.columns.slice(1).map(function(id) {
    return {
```

Tools: Starting Easy...

Tools: Some Dashboards

Tools: Some Libraries

FEDER Wal-E-Cities



www.enmieux.be

FEDER
UNION EUROPEENNE

Wallonie

LE FONDS EUROPÉEN DE DÉVELOPPEMENT RÉGIONAL
ET LA WALLONIE INVESTISSENT DANS VOTRE AVENIR

Sources & Reference

- Tamara Munzner. Visualization Analysis and Design. A K Peters Visualization Series, CRC Press, 2014.