

Data Visualization Principles: Interaction, Filtering, Aggregation

CSC444

Announcements

- Assignment 5 is due tonight
- Assignment 6 is posted
 - **Read this one early**
- Let's go over a solution for Assignment 4

What if there's too much data?

- Sometimes you can't present all the data in a single plot
- **Interaction:** let the user drive what aspect of the data is being displayed
- **Filtering:** Selectively hide some of the data points
- **Aggregation:** Show visual representations of subsets of the data

Focus+Context

When showing a limited view, try to hint at what is not being shown.

Demos: NYT Interactive charts

<http://www.nytimes.com/interactive/2014/06/05/upshot/how-the-recession-reshaped-the-economy-in-255-charts.html?abt=0002&abg=0>

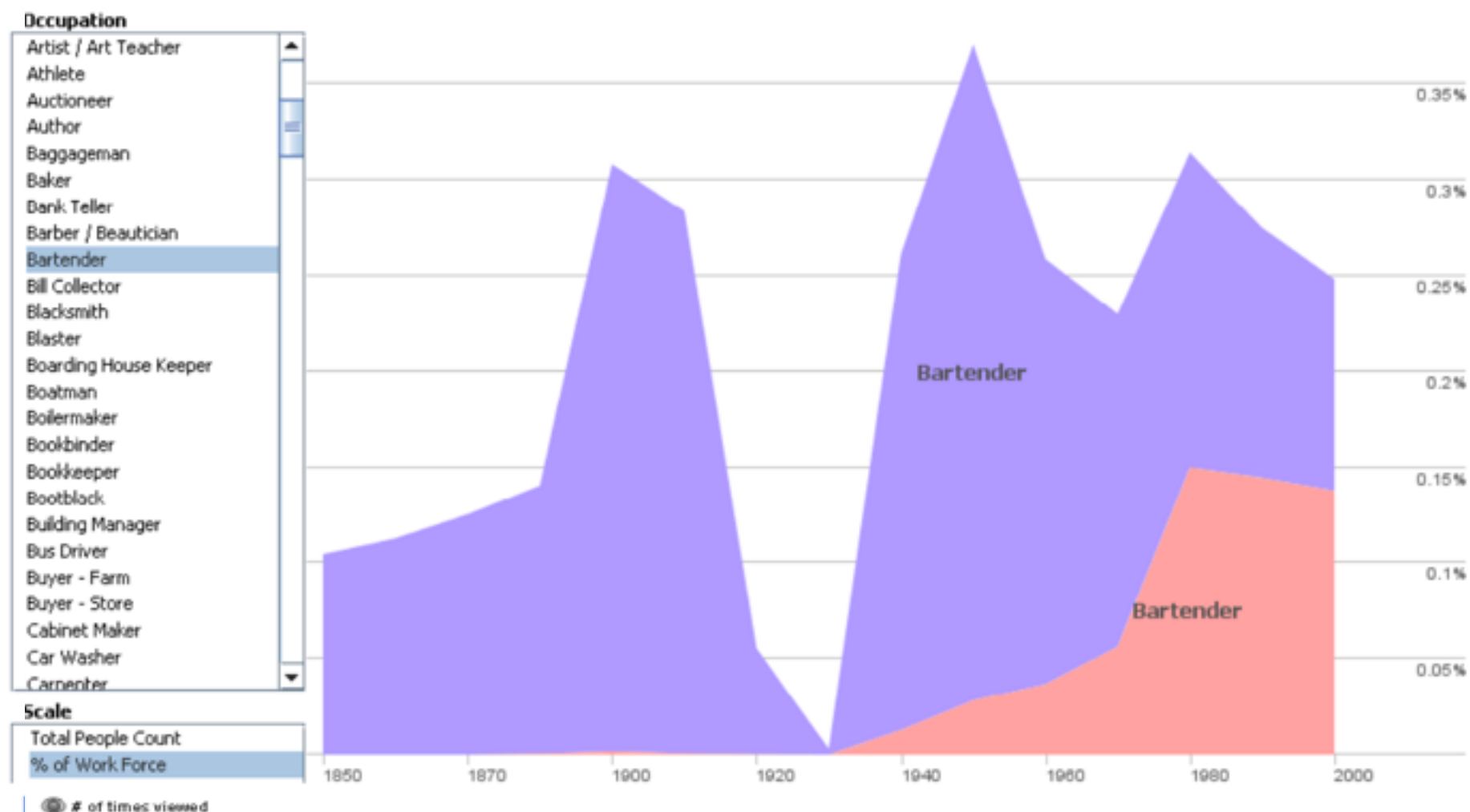
<http://www.nytimes.com/interactive/2014/09/19/nyregion/stop-and-frisk-map.html>

<http://www.nytimes.com/interactive/2014/upshot/buy-rent-calculator.html?abt=0002&abg=0>

INTERACTION

Fundamental idea

- Interpret the state of elements in the UI as a **clause** in a **query**. As UI changes, update data



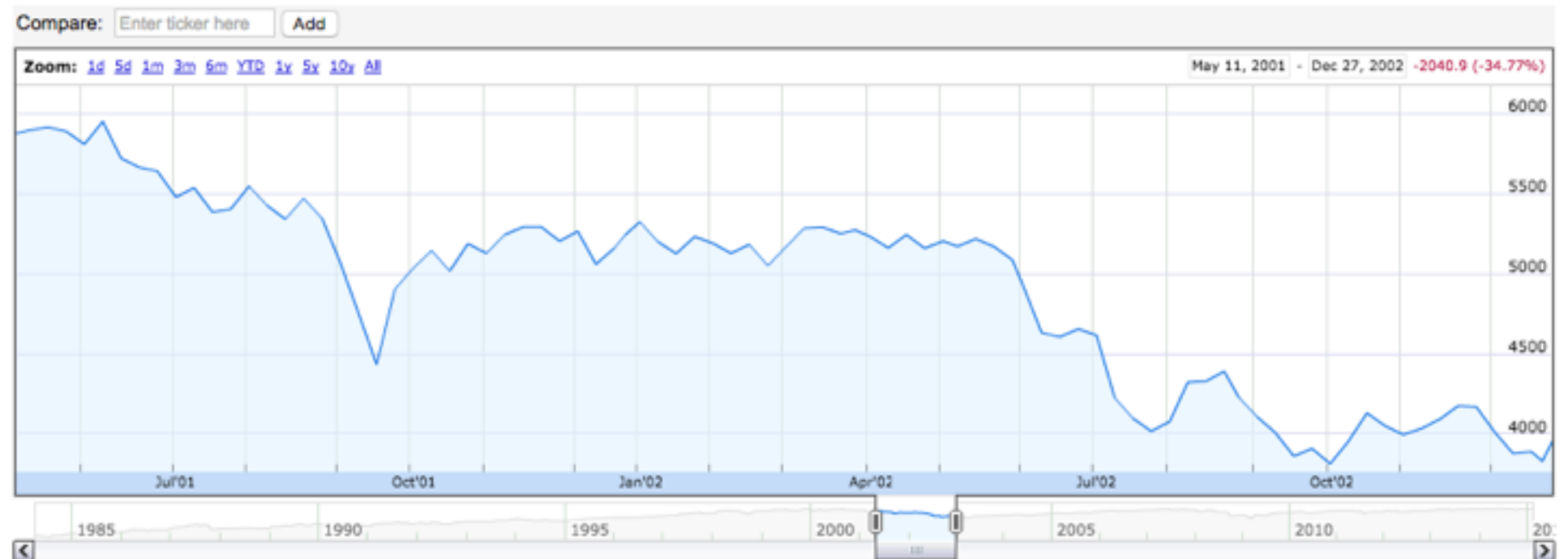
Willett et al., TVCG 2007 (*)

Panning



<https://www.google.com/finance?q=INDEXFTSE>

Zooming



<https://www.google.com/finance?q=INDEXFTSE>

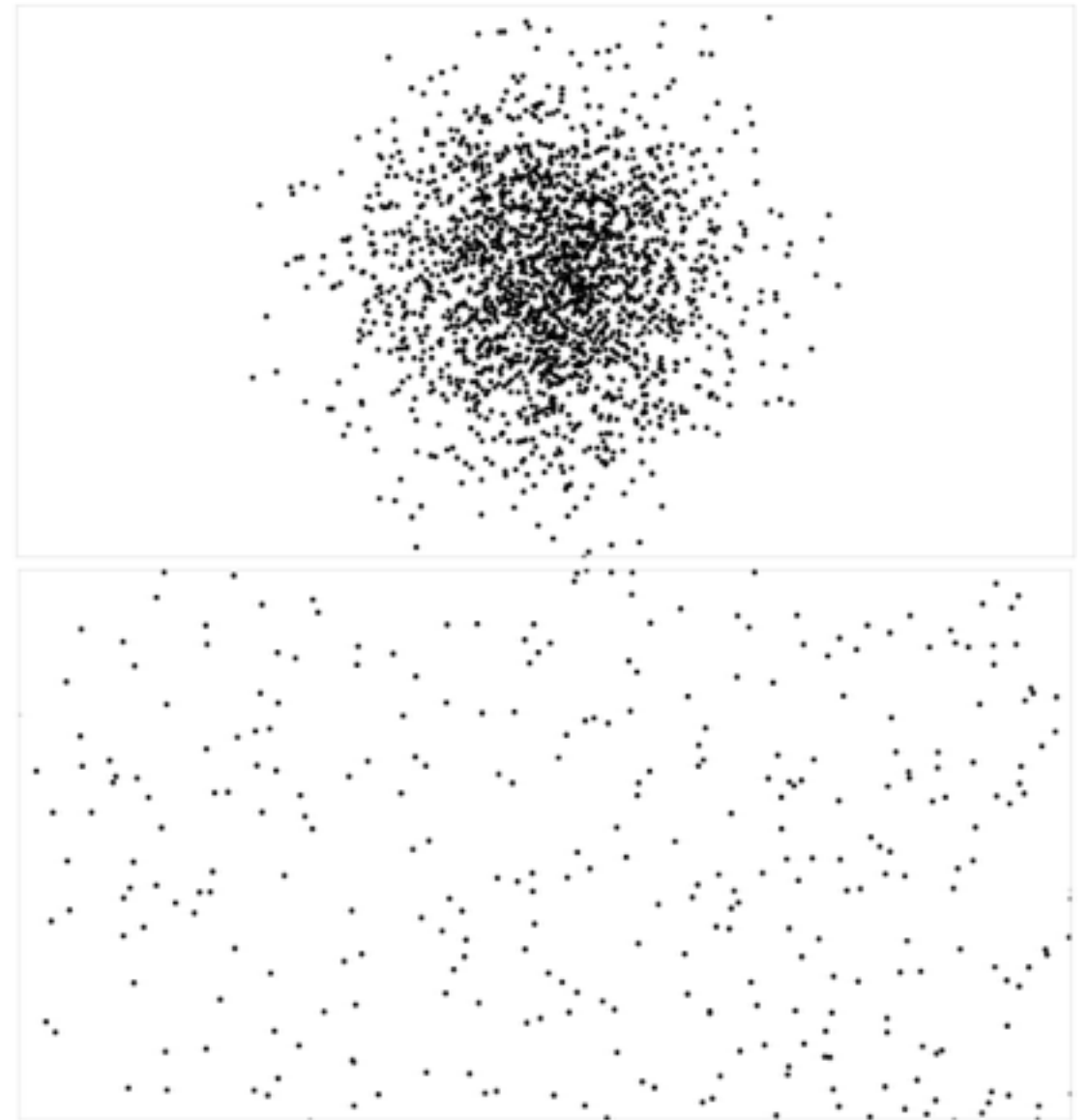
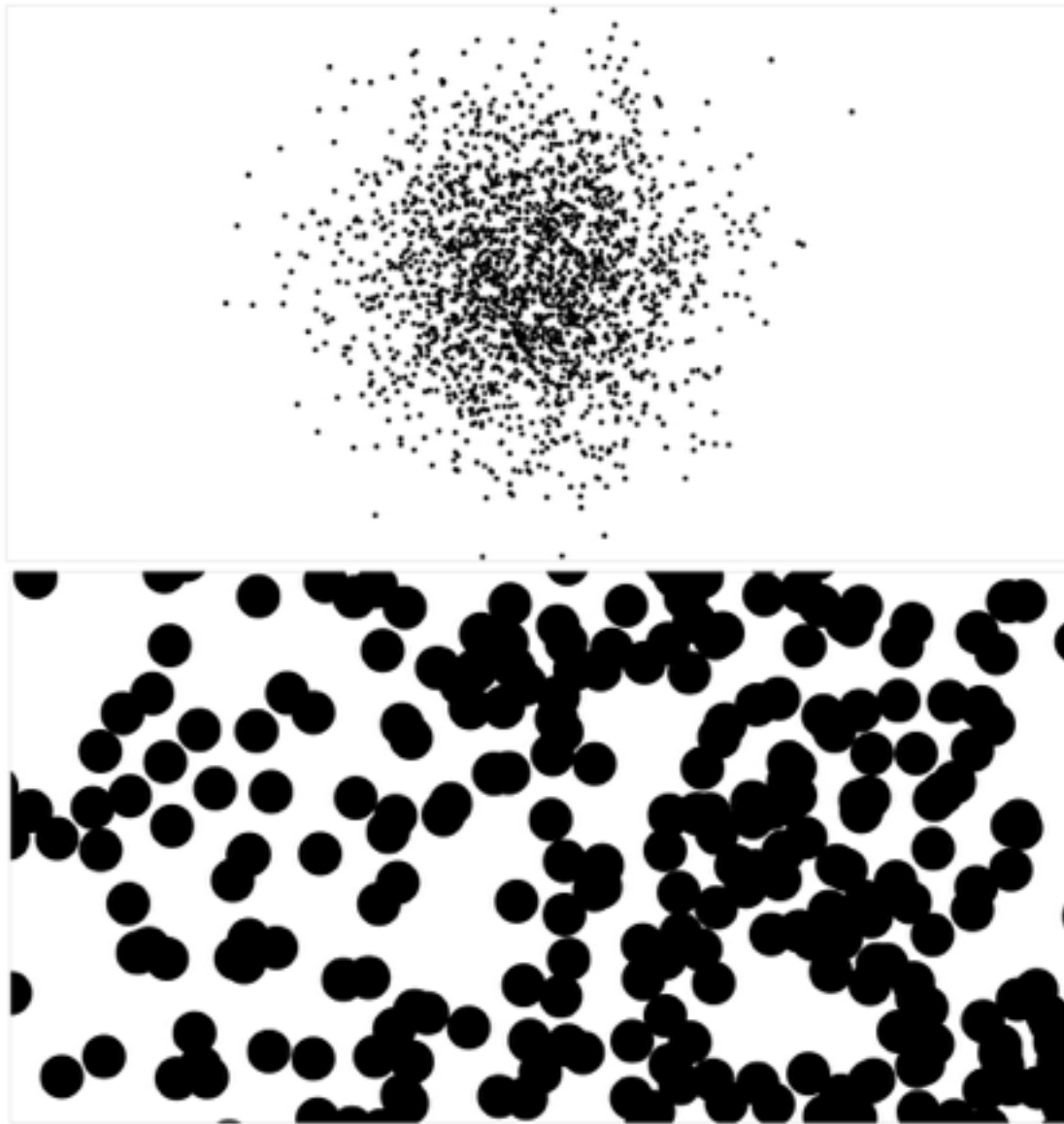
Focus+Context for Pan & Zoom



“Geometric” Zooming

vs.

“Semantic” Zooming



Smooth Zoom transitions (research highlight)

- What's the “best” way to go from one zoomed view to another?
- Differential equations to the rescue!

van Wijk and Nuij, Infovis 2003
<http://bl.ocks.org/mbostock/3828981>

Research Highlight: smooth zoom transitions

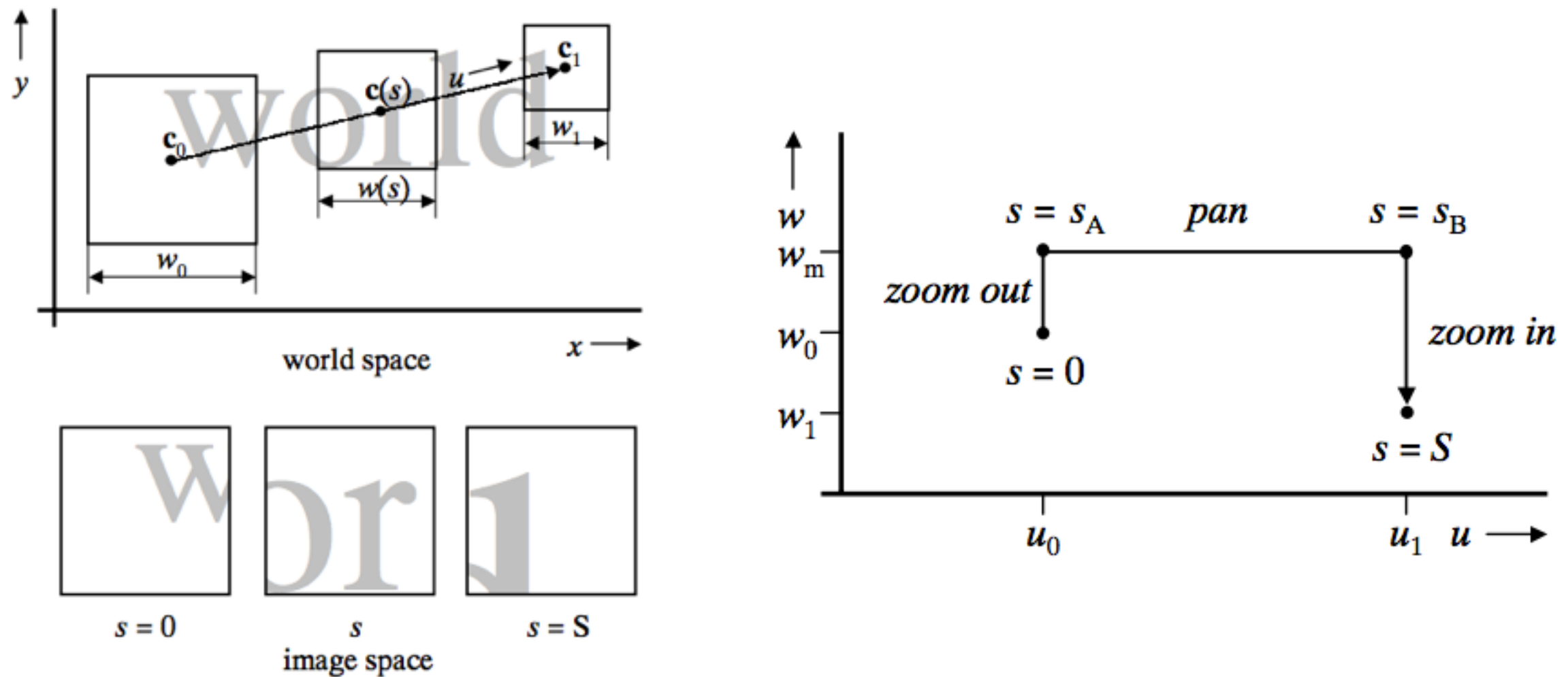
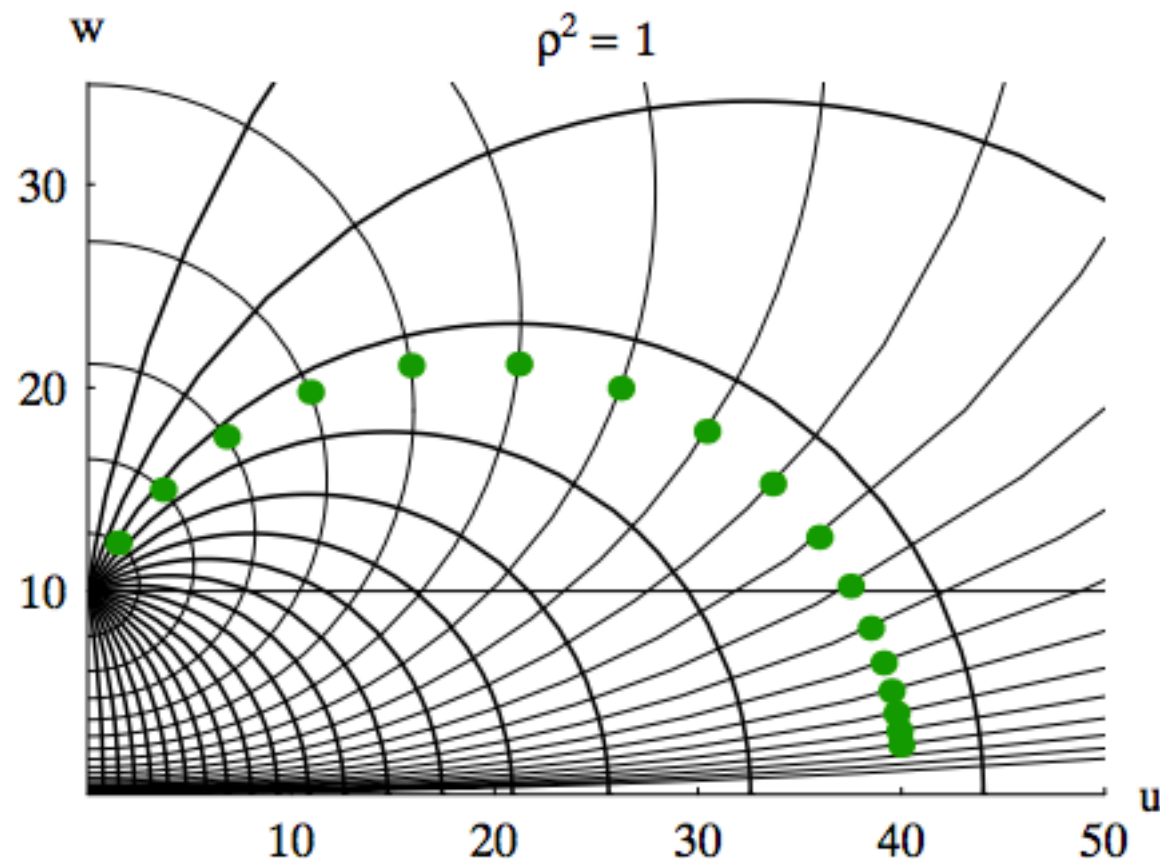


Figure 1: World space and image space

van Wijk and Nuij, Infovis 2003

<http://blocks.org/mbostock/3828981>

Research Highlight: smooth zoom transitions



Shortest paths in zoom
space!

3.3 Metric

We aim for a path that is smooth and efficient. Both require that we are able to measure the effect of changing c and w , as perceived by the viewer. Following and generalizing the approach of Igarashi and Hinckley [2000], we use the velocity of the moving image as a basis for measurements, i.e., we aim at a metric for the perceived average optic flow in the image window. To this end, we first consider the

...

(e.g. $E_u = \partial E / \partial u$). For our metric $E = \rho^2 / w^2$ and $G = 1 / \rho^2 w^2$, substitution gives

$$\ddot{u} - 2\dot{u}\dot{w}/w = 0, \text{ and}$$

$$\ddot{w} + \rho^4 \dot{u}^2 / w - \dot{w}^2 / w = 0.$$

(8)

van Wijk and Nuij, Infovis 2003

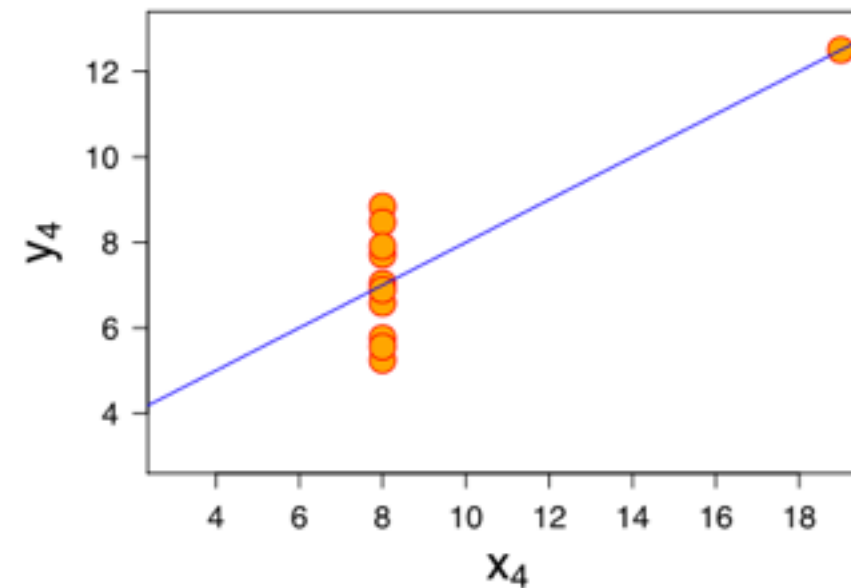
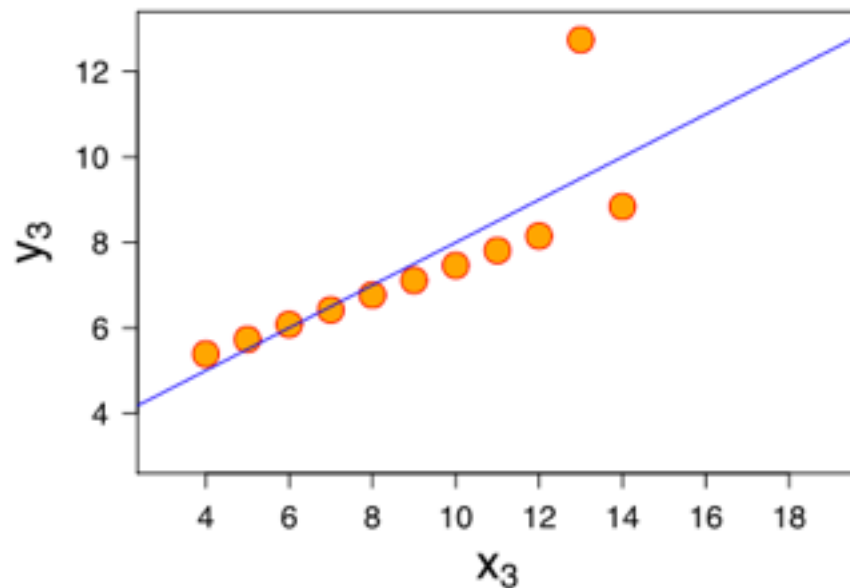
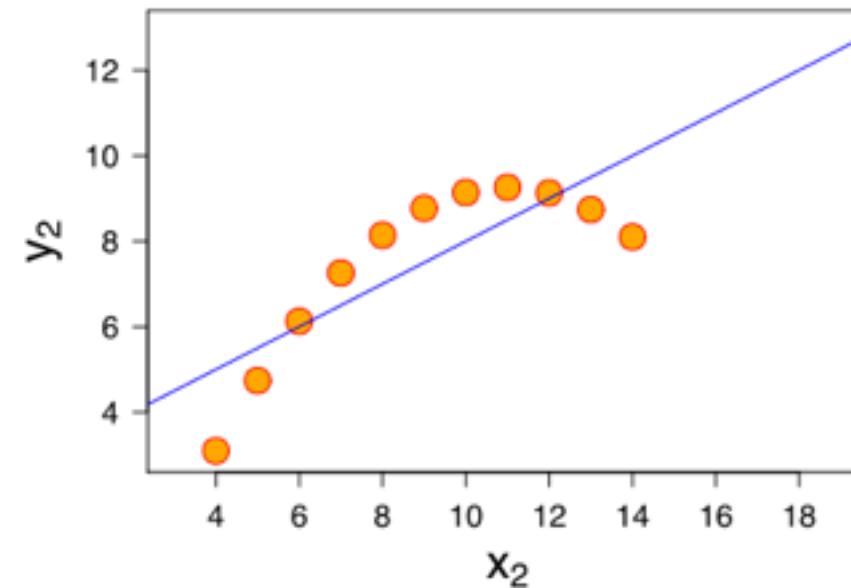
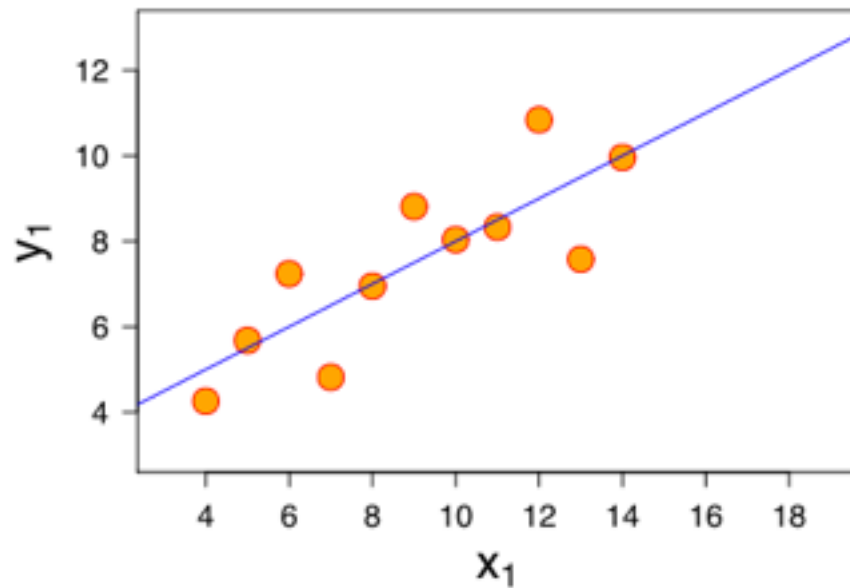
<http://blocks.org/mbostock/3828981>

FILTERING

Fundamental idea

- Choose a rule, hide elements that don't match that rule
- the more complex the rule, the better you will be able to find patterns in the data. **More focus**
- the more complex the rule, the less transparent it is, so user doesn't know what the filtering is doing. **Less context**

- Case in point: **do not hide outliers!**
- **Fancy outlier detection considered harmful**



Brushing, linked views

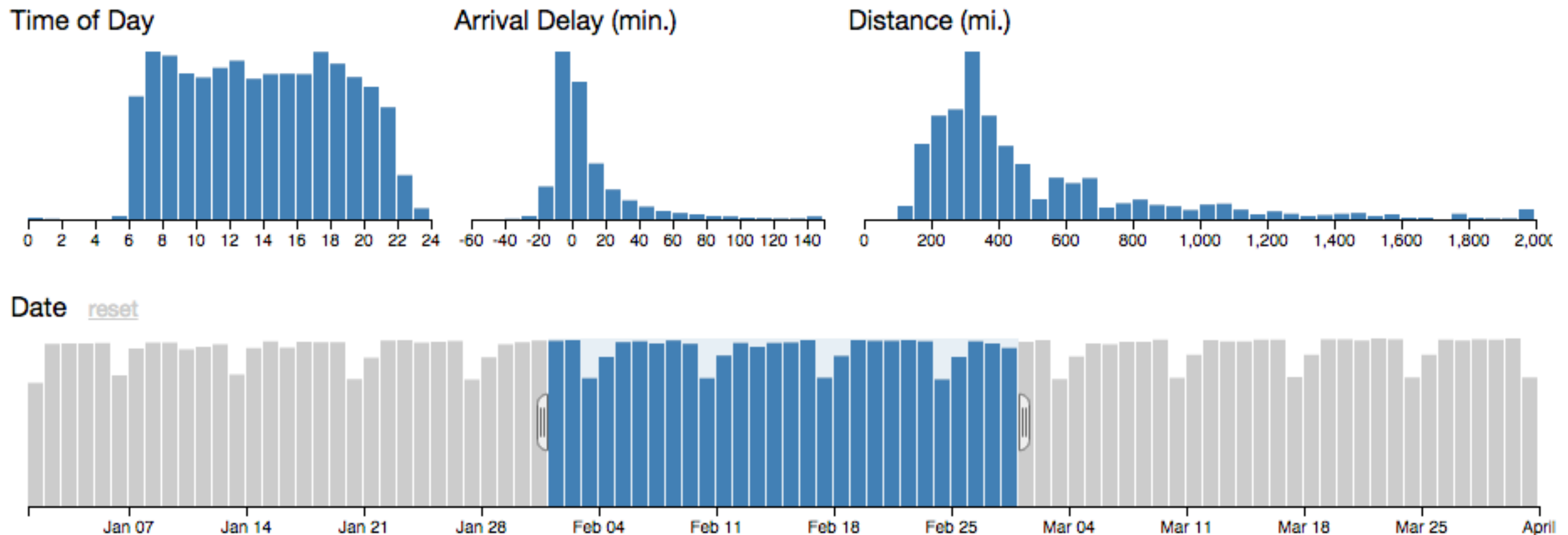
- **Filtering + Interaction**
- Show more than one view of the same data
- Users drag “brushes”: regions of each view, which are interpreted directly as queries
- No additional UI!

<http://bl.ocks.org/mbostock/4063663>

AGGREGATION

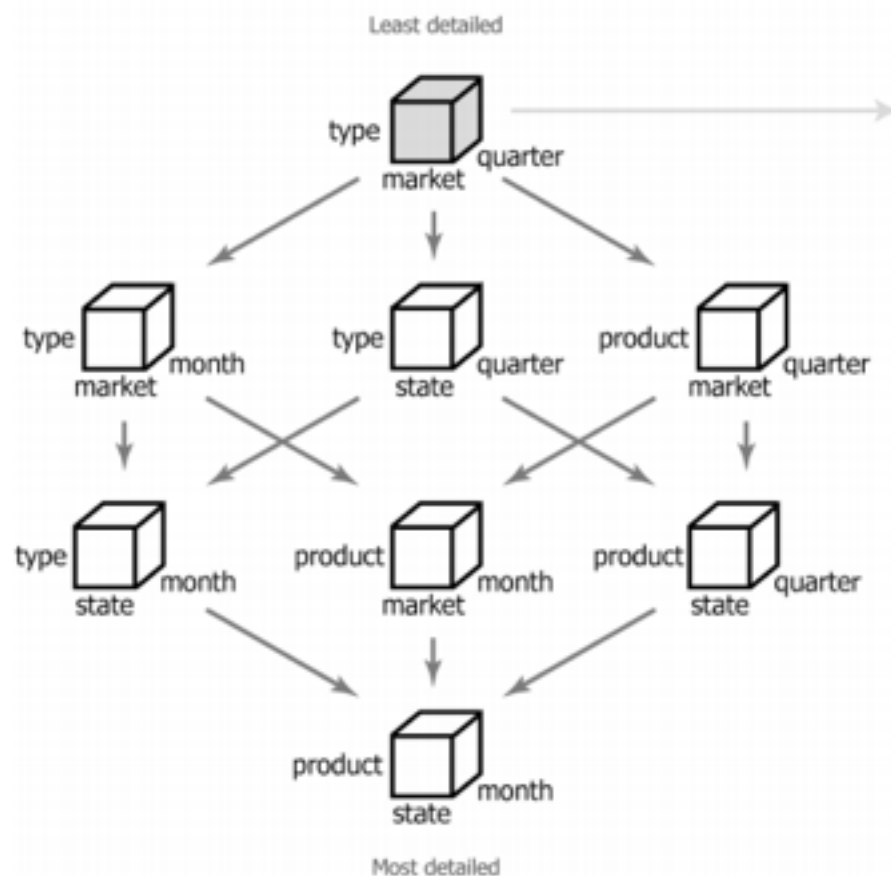
Fundamental idea

- If there's too much data, replace individual data points with representation of **subsets**

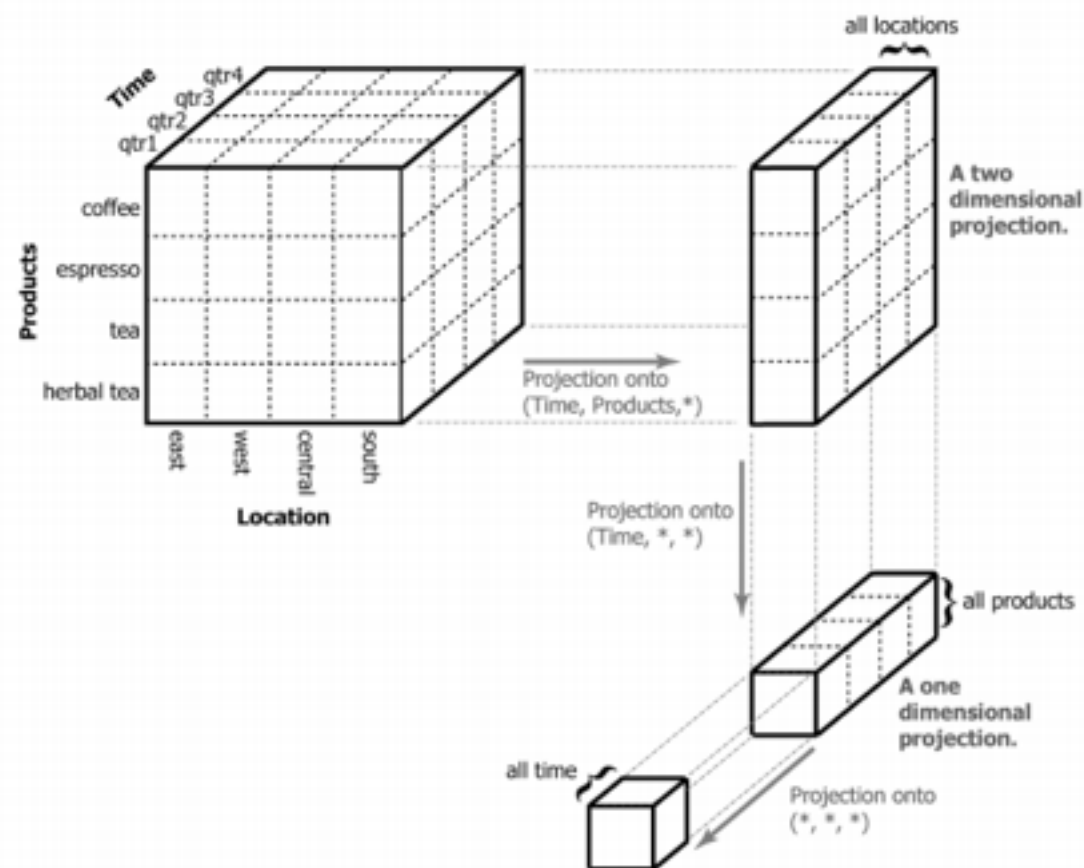


Data Cubes: aggregate by collapsing attributes

(a) The lattice of data cubes



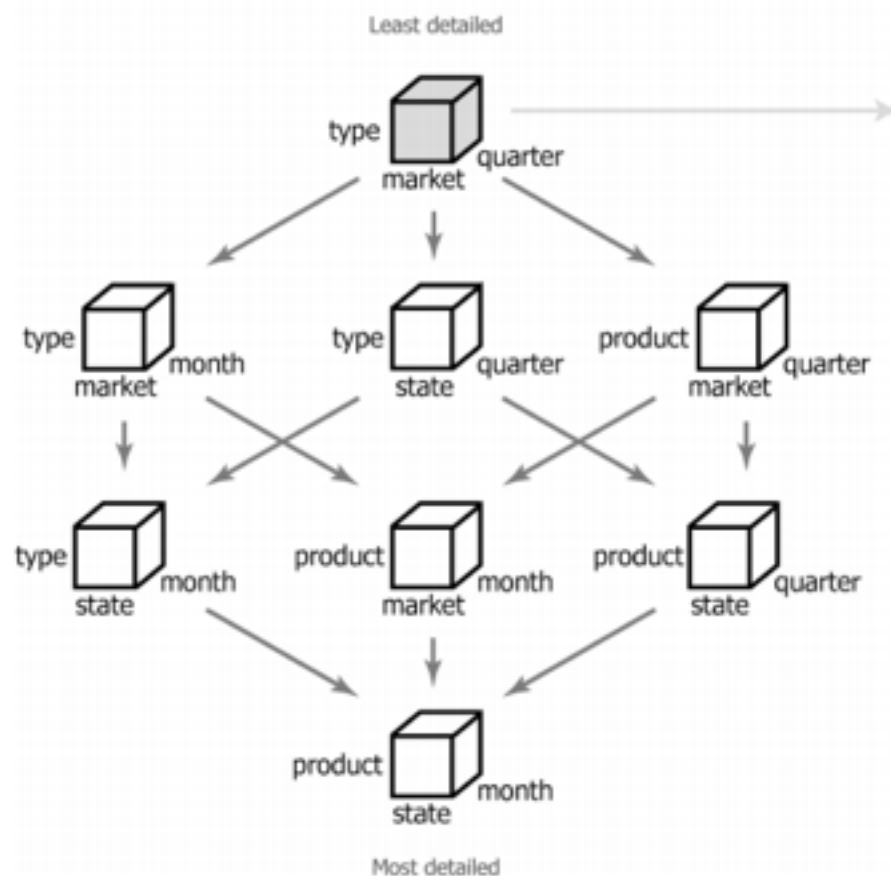
(b) Projecting a three dimensional data cube



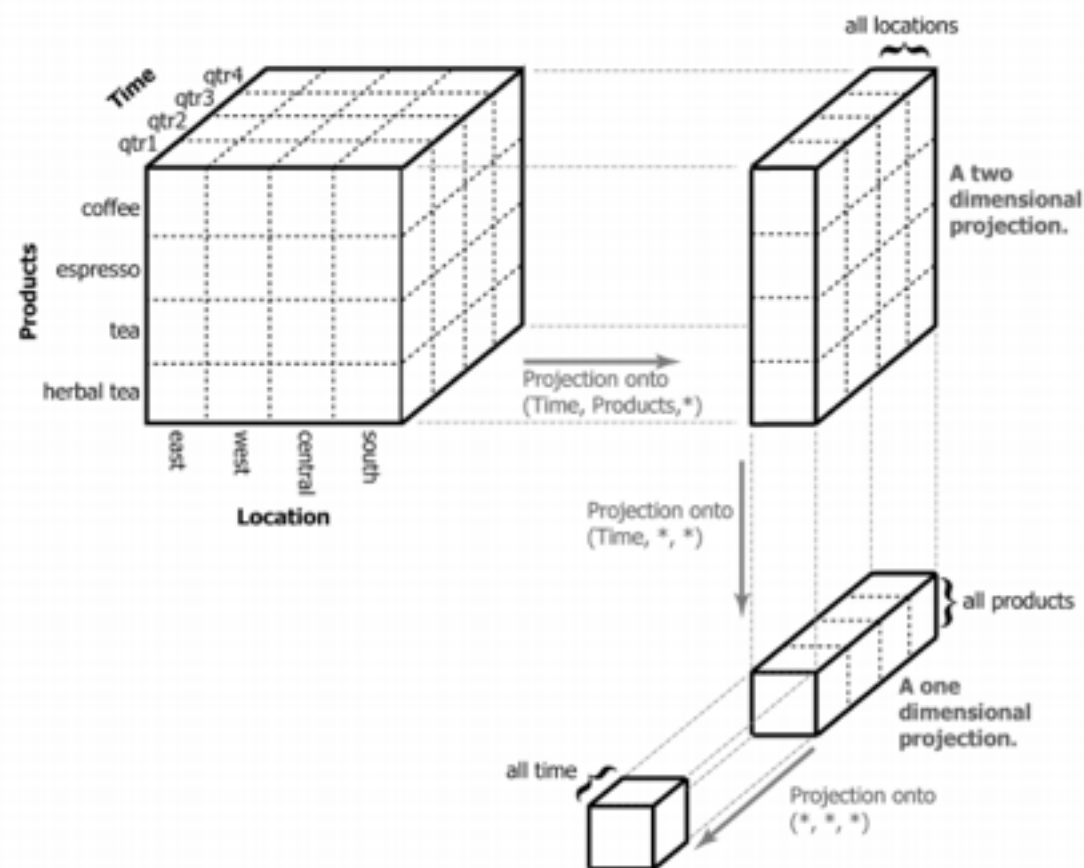
Multiscale Visualization using Data Cubes,
Stolte et al., Infovis 2002

Data Cubes: aggregate by collapsing attributes

(a) The lattice of data cubes



(b) Projecting a three dimensional data cube



Multiscale Visualization using Data Cubes,
Stolte et al., Infovis 2002

Data Cubes: aggregate by collapsing attributes

- recent: data cubes specifically designed for vis:
 - Bostock et al.'s Crossfilter (<http://square.github.io/crossfilter/>)
 - Liu et al.'s Immens (<http://vis.stanford.edu/papers/immens>)
 - Lins et al.'s Nanocubes (<http://nanocubes.net/>)
- **Filtering + Aggregation + Interaction**

Scented widgets (Willett et al., 2007)

- If UI is necessary, summarize data on UI overlay
- **Filtering + Aggregation + Interaction**



Research Questions

- “Torture your data enough, and it’ll tell you anything”, Ronald Coase
 - (<http://tylervigen.com/>)
- Statistics has tools to mitigate this problem
- Interaction is much less well-studied!

Shneiderman's “Visual information seeking mantra”

**Overview first,
zoom and filter,
then details-on-demand**

Demos

<http://www.nytimes.com/interactive/dining/new-york-health-department-restaurant-ratings-map.html>

<http://square.github.io/crossfilter/>

<http://cscheid.net/static/mlb-hall-of-fame-voting/>

Overview first:

Before all else, show a “high-level” view, possibly through appropriate aggregation

Zoom and Filter:

Use interaction to create
user-specified views

Details on Demand:

Individual points or attributes
should be available, but only
as requested