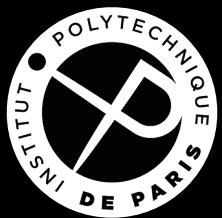


Tasks & Interaction

James EAGAN <james.eagan@telecom-paris.fr>



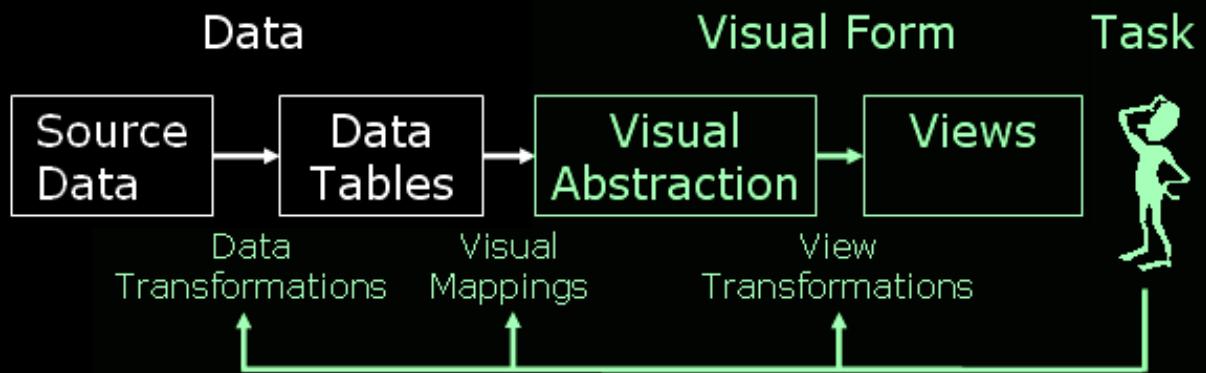
Includes slides adapted from John Stasko
(Georgia Tech), Petra Isenberg & Jean-Daniel
Fekete (INRIA), Nadia Boukhelifa (INRA),
Chris North (Virginia Tech), Tamara Munzner (UBC)



Updated: May 2020

1

InfoVis Pipeline



2

Outline

- How do people make sense of data?
- What do people do with a visualization system?
- What are ways people can interact with data visualizations?
- Shneiderman's mantra
- Dealing with Big Data

3

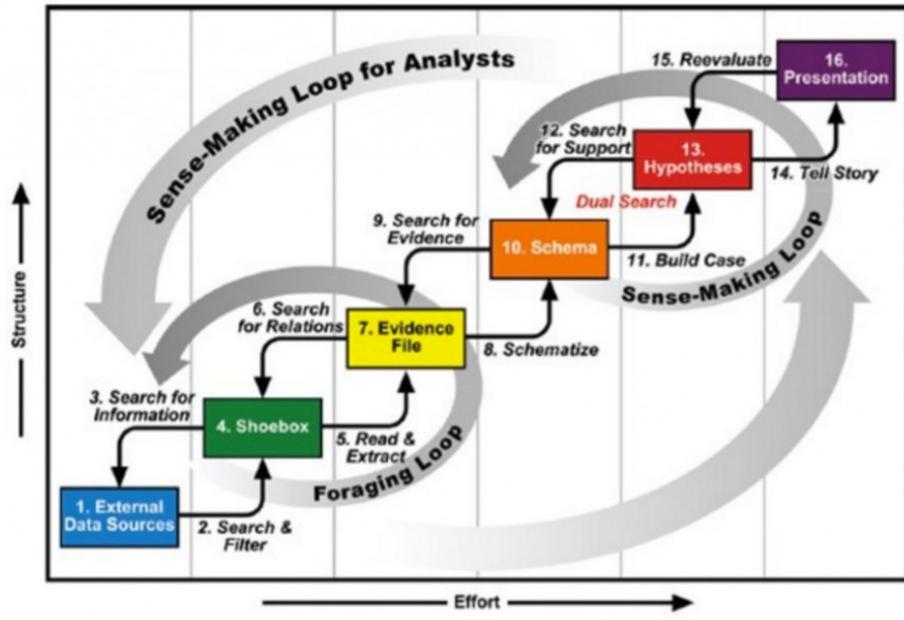
3

**How do people
make sense of data?**

4

4

Pirolli-Card Sensemaking Loop



Model: Pirolli, Peter, and Stuart Card. "The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis." *Proceedings of International Conference on Intelligence Analysis*. Vol. 5. McLean, VA: Mitre, 2005.

Image source: Thomas, James J. and Kristin A. Cook "Illuminating The Path" (2005): pp. 44

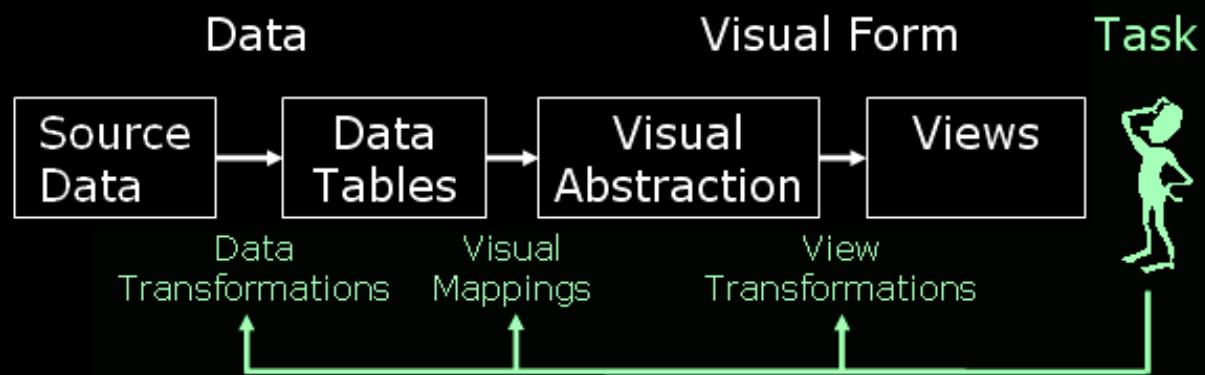
5

Sensemaking

- Rich domain of study
 - How do data workers deal with uncertainty?
[Boukhelifa, Perrin, Huron & Eagan, CHI '17]
 - How do data workers manage alternatives in sensemaking?
[Liu, Boukhelifa & Eagan, TVCG 26 (1), '20]
 - How do data scientists use computational notebooks?
[Kery & Myers, VL/HCC '17]
 - How do people perform exploratory programming?
[Rule, Tabard & Hollan, CHI '18]

6

InfoVis Pipeline



7

7

What kinds of questions do people ask of a visualization?

8

8

Low-level analytic tasks

- Retrieve value
- Filter
- Compute derived value
- Find extremum
- Sort
- Determine range
- Characterize distribution
- Find anomalies
- Cluster
- Correlate

[Amar, Eagan, & Stasko, InfoVis '05]

9

9

Retrieve value

For a specific data case, find its attributes

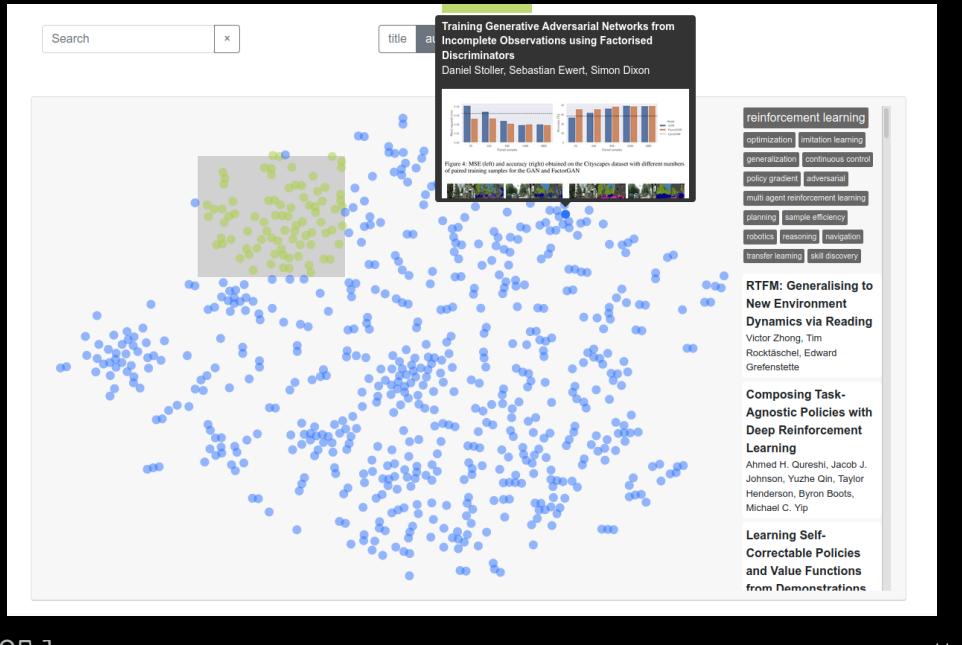
[Amar, Eagan, & Stasko, InfoVis '05]

10

10

Retrieve value

For a specific data case, find its attributes



[Amar, Eagan, & Stasko, InfoVis '05]

11

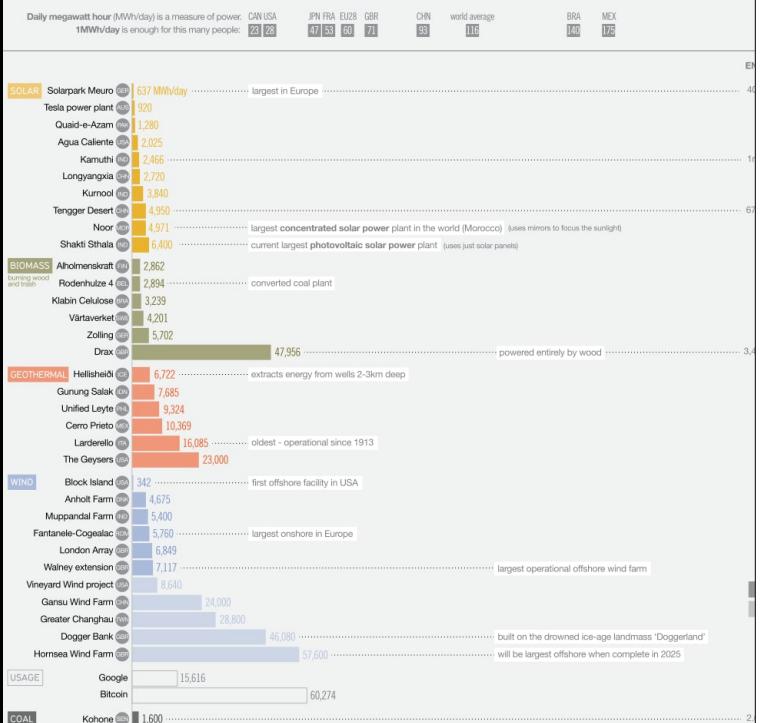
11

Retrieve value

For a specific data case, find its attributes

Mega-What?

World's largest and notable energy sources

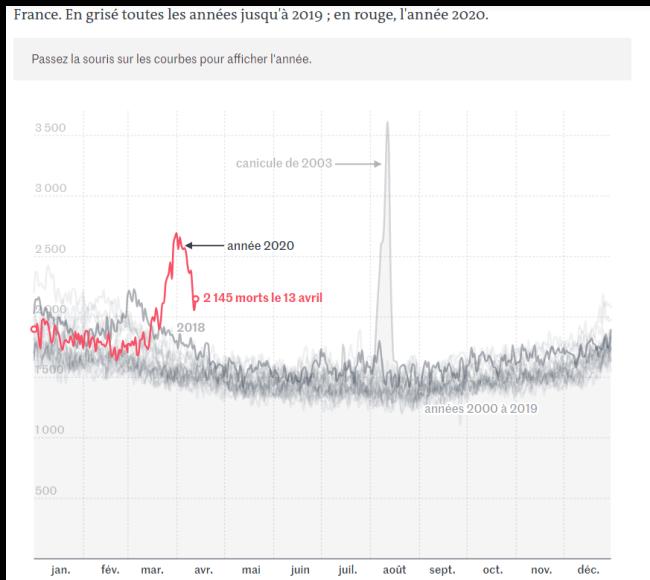


[Amar, Eagan, & Stasko, InfoVis '05]

12

Retrieve value

For a specific data case, find its attributes



[Amar, Eagan, & Stasko, InfoVis '05]

13

13

Filter

Show only the data cases matching a set of conditions based on attribute values

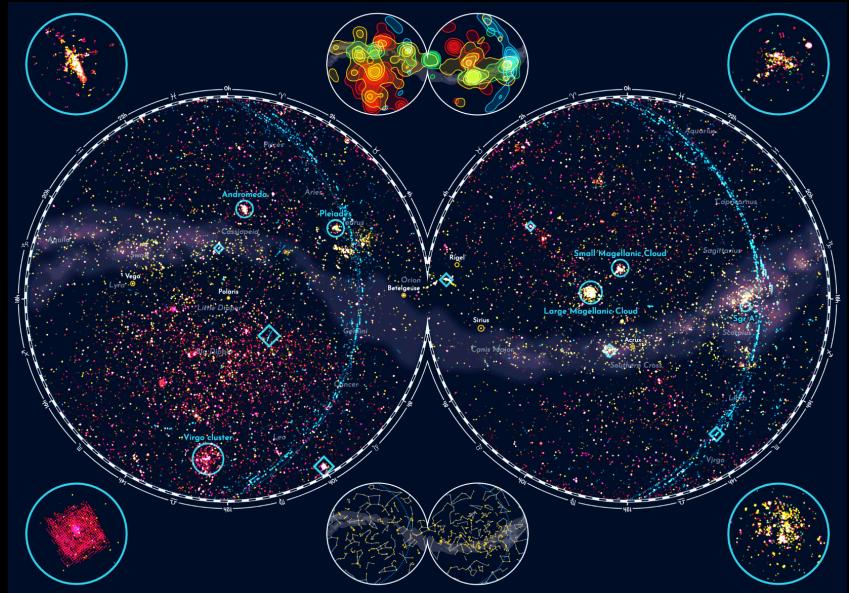
[Amar, Eagan, & Stasko, InfoVis '05]

14

14

Filter

Show only the data cases matching a set of conditions based on attribute values



[Amar, Eagan, & Stasko, InfoVis '05]

15

15

Filter

Show only the data cases matching a set of conditions based on attribute values

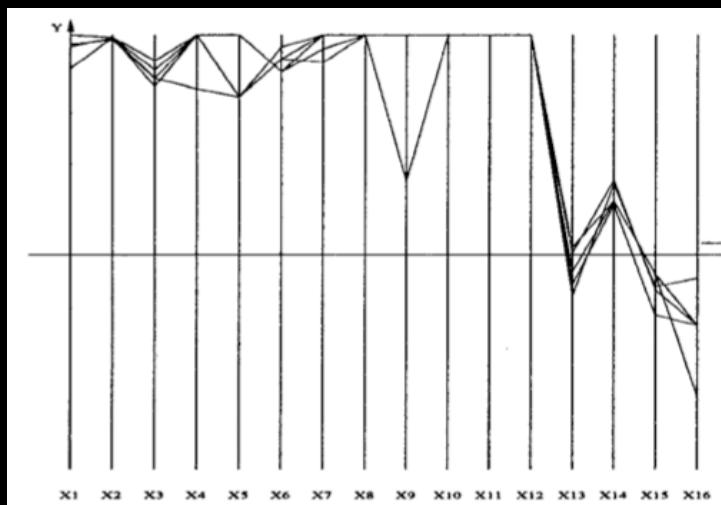


Figure 6: Batches with the highest Yields do not have the lowest defects in $X3$ and $X6$.

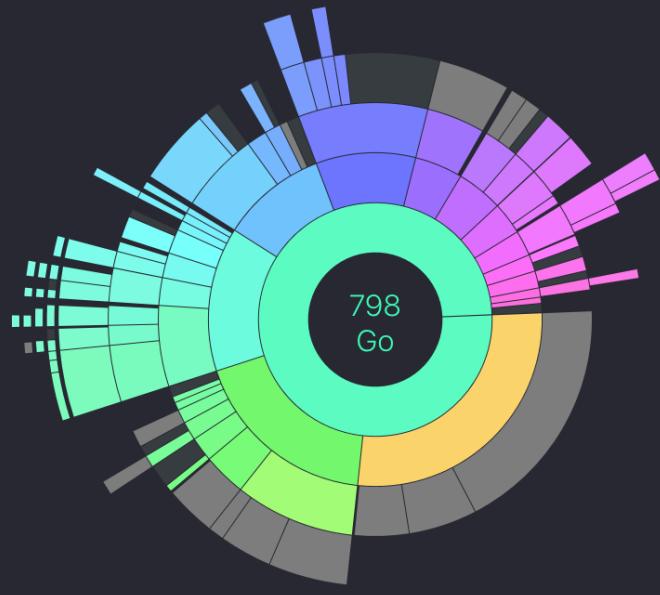
[Amar, Eagan, & Stasko, InfoVis '05]

16

16

Filter

Show only the data cases matching a set of conditions based on attribute values

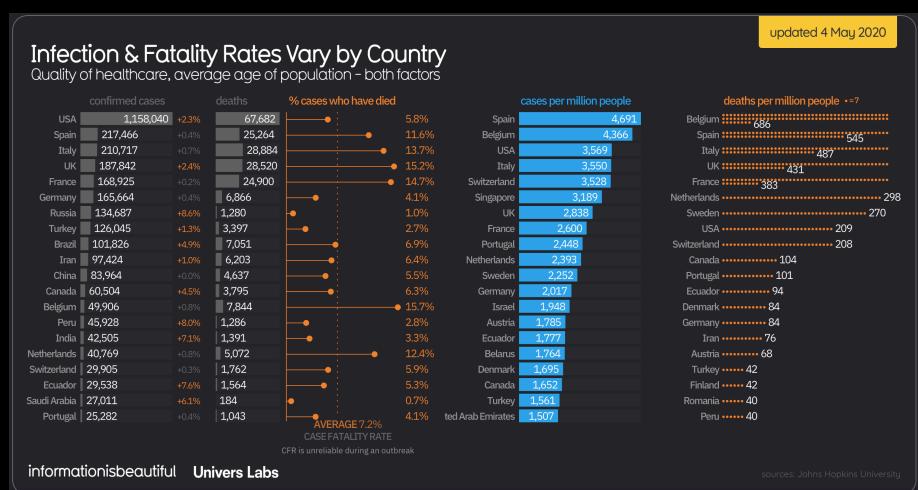


[Amar, Eagan, & Stasko, InfoVis '05]

17

Compute derived value

Compute an aggregate value for a set of data cases

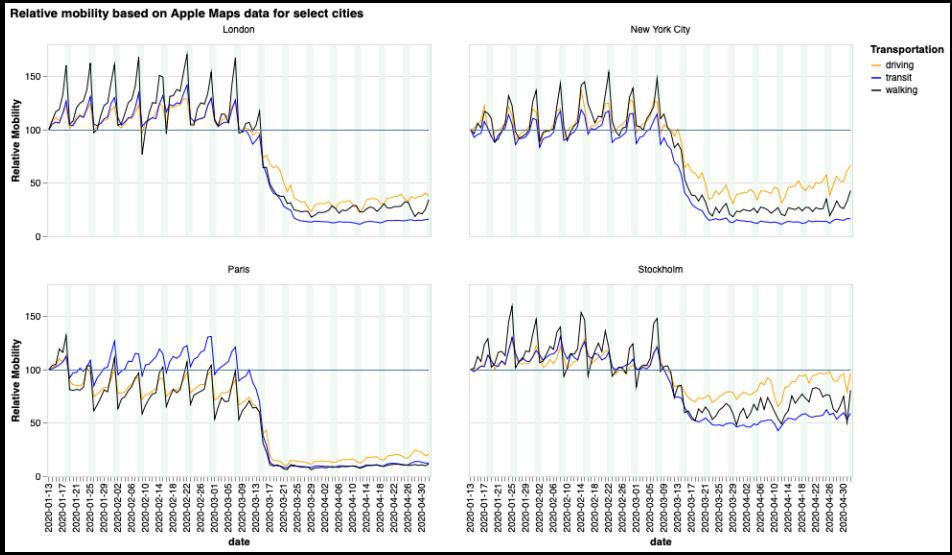


[Amar, Eagan, & Stasko, InfoVis '05]

18

Compute derived value

Compute an aggregate value for a set of data cases



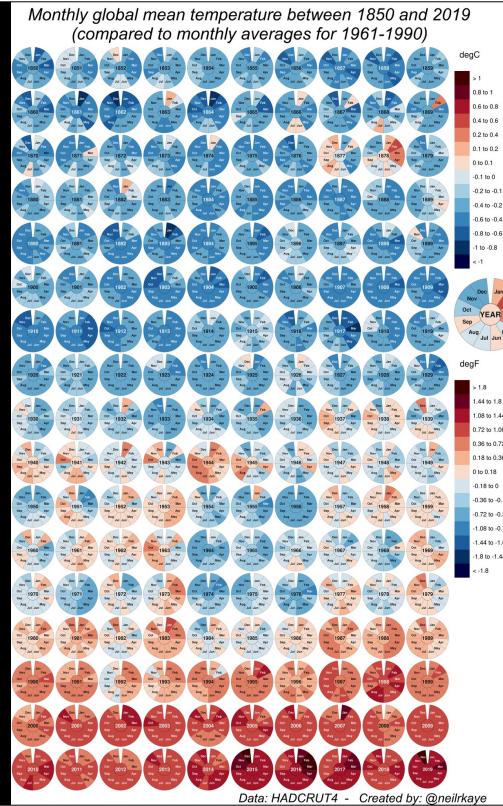
[Amar, Eagan, & Stasko, InfoVis '05]

19

19

Find extremum

Find data cases at either extreme of the data set for a given attribute



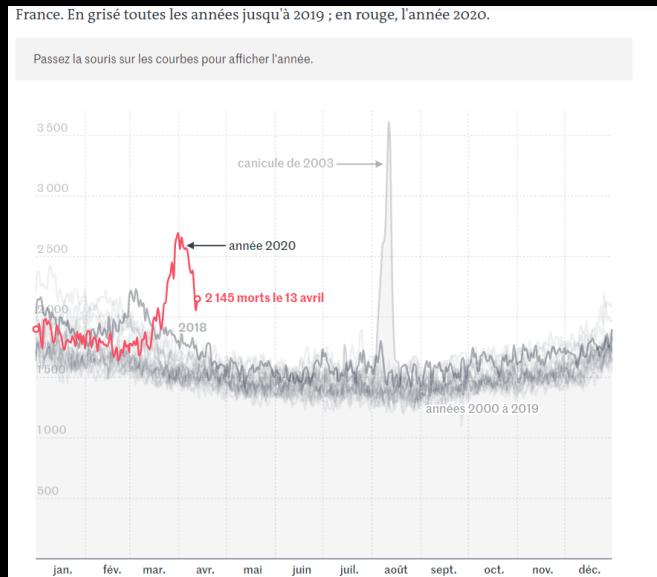
[Amar, Eagan, & Stasko, InfoVis '05]

20

20

Find extremum

Find data cases at either extreme of the data set for a given attribute



[Amar, Eagan, & Stasko, InfoVis '05]

21

21

Sort

Rank data cases according to some ordinal metric

[Amar, Eagan, & Stasko, InfoVis '05]

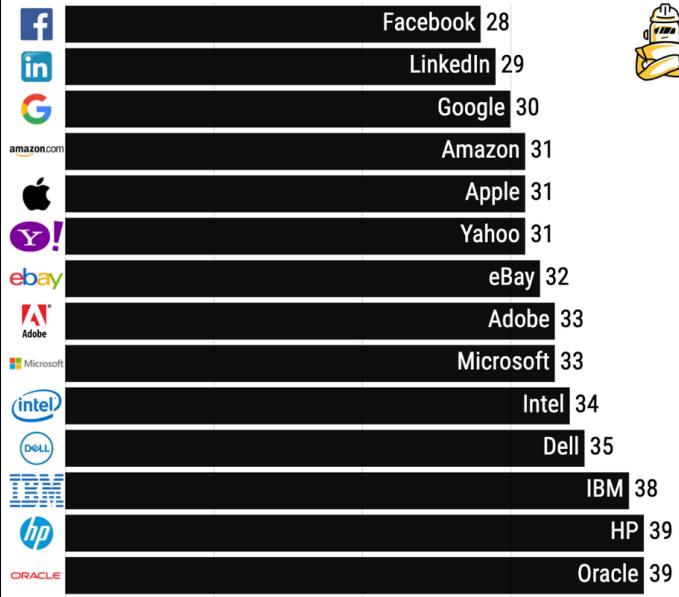
22

22

Sort

Rank data cases according to some ordinal metric

Median Employee Age Of Top Tech Firms



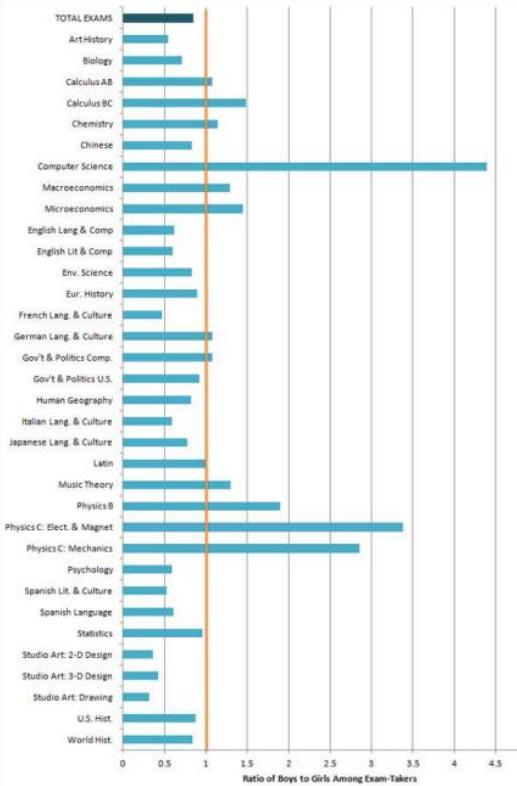
@WORLDWIDE_ENGINEERING <https://t.me/worldwideengineering>

[Amar, Eagan, & Stasko, InfoVis '05]

23

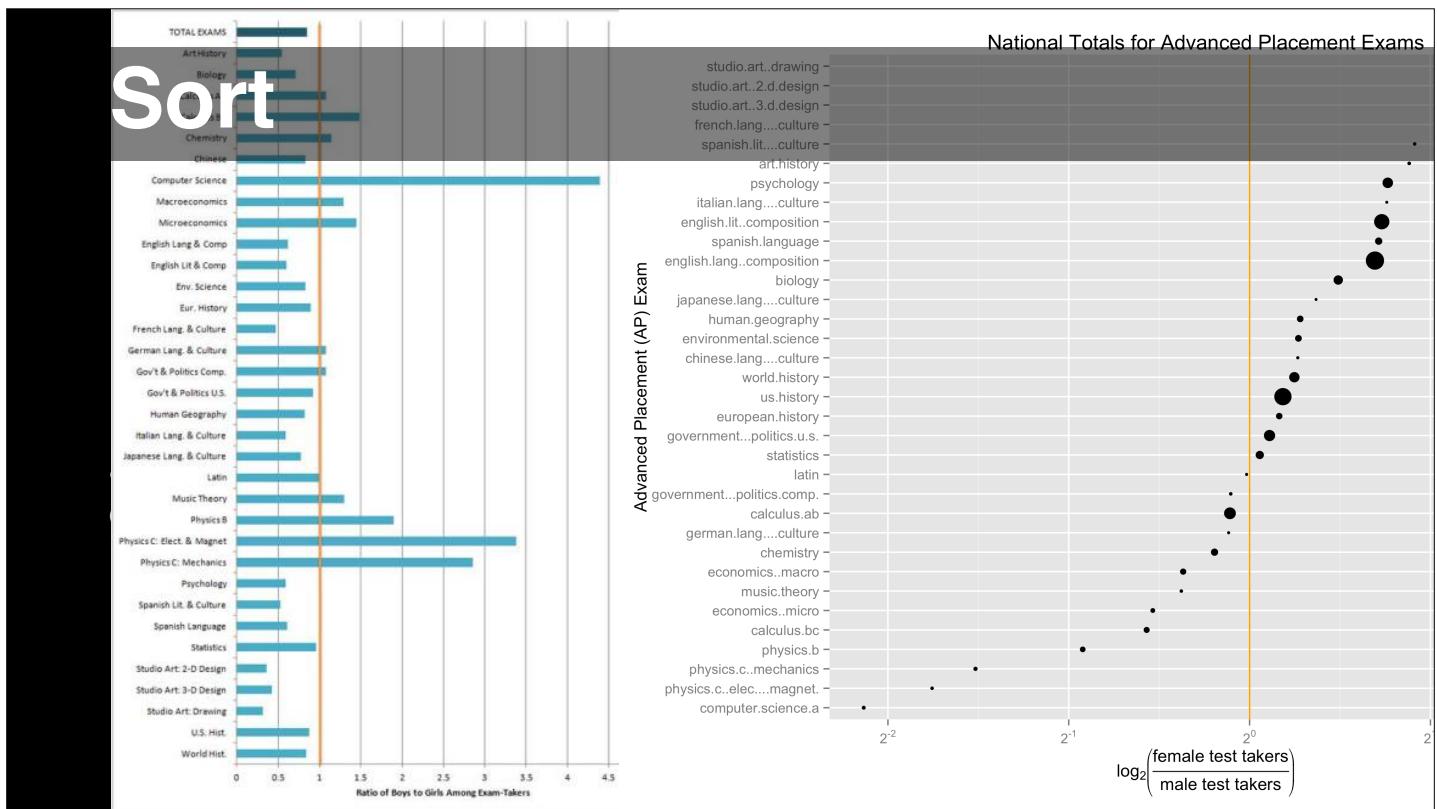
Sort

Rank data cases according to some ordinal metric



[Amar, Eagan, & Stasko, InfoVis '05]

24



25

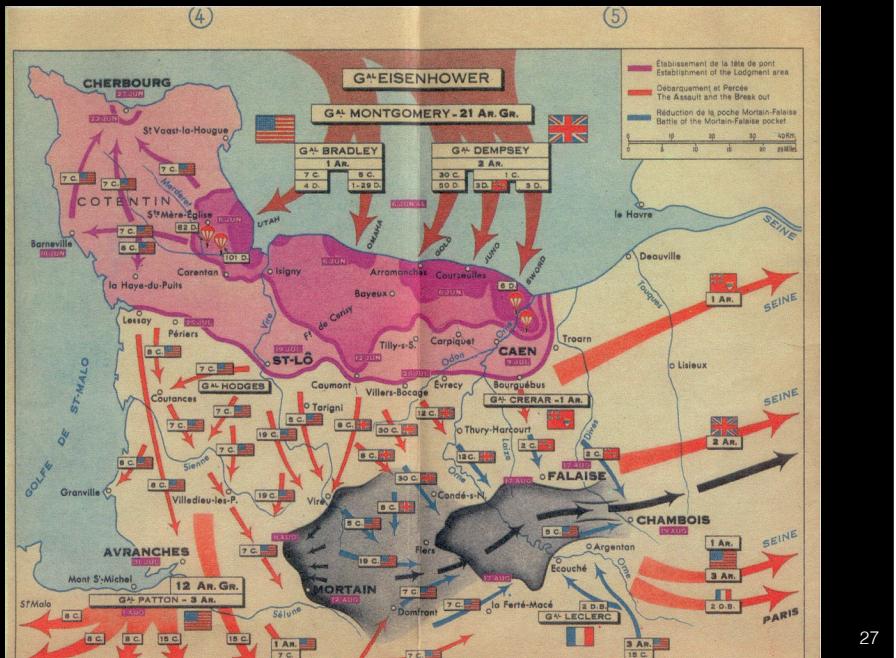
Determine range

Identify the span of values for a given set of data cases and attributes

Determine range

Identify the span of values for a given set of data cases and attributes

[Amar, Eagan, & Stasko, InfoVis '05]



27

Determine range

Identify the span of values for a given set of data cases and attributes



[Amar, Eagan, & Stasko, InfoVis '05]



28

Characterize distribution

For a given set of cases and attributes of interest, characterize the distribution of that value over the set

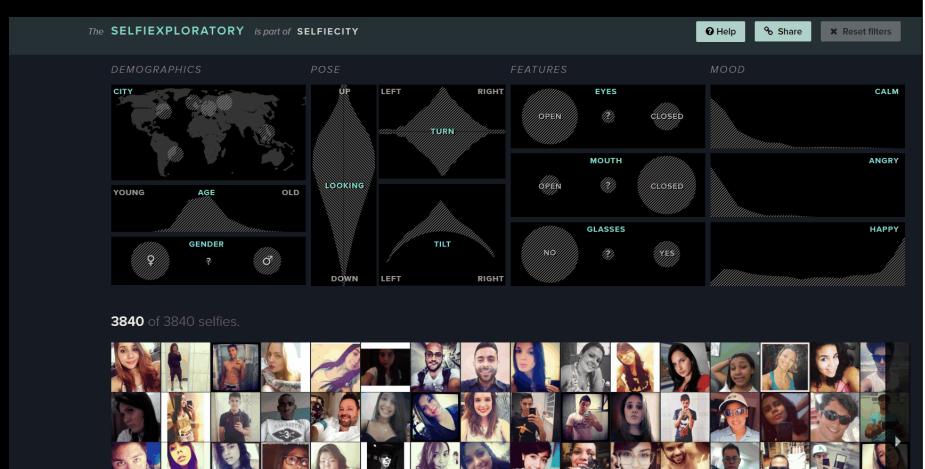
[Amar, Eagan, & Stasko, InfoVis '05]

29

29

Characterize distribution

For a given set of cases and attributes of interest, characterize the distribution of that value over the set



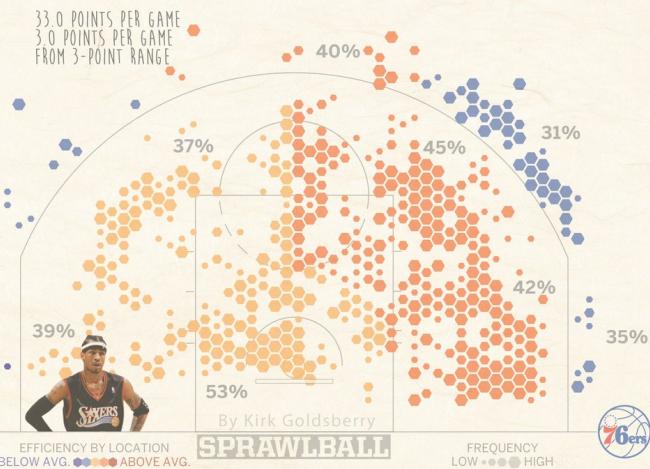
[Amar, Eagan, & Stasko, InfoVis '05]

30

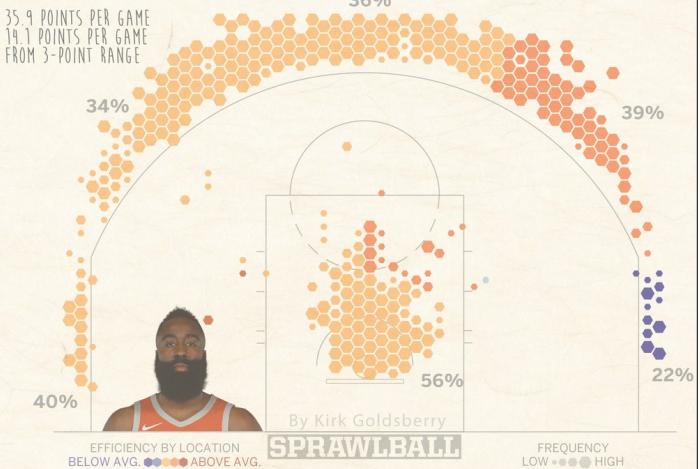
30

Characterize distribution

ALLEN IVERSON, 2005-06



JAMES HARDEN, 2018-19

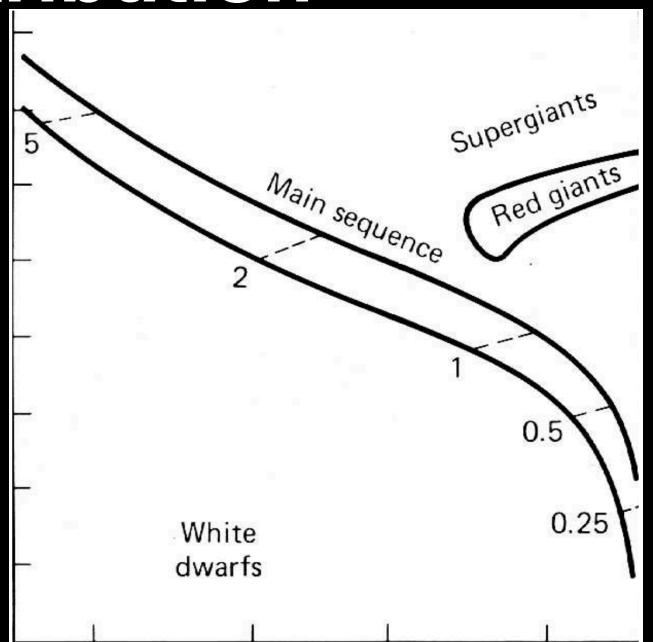
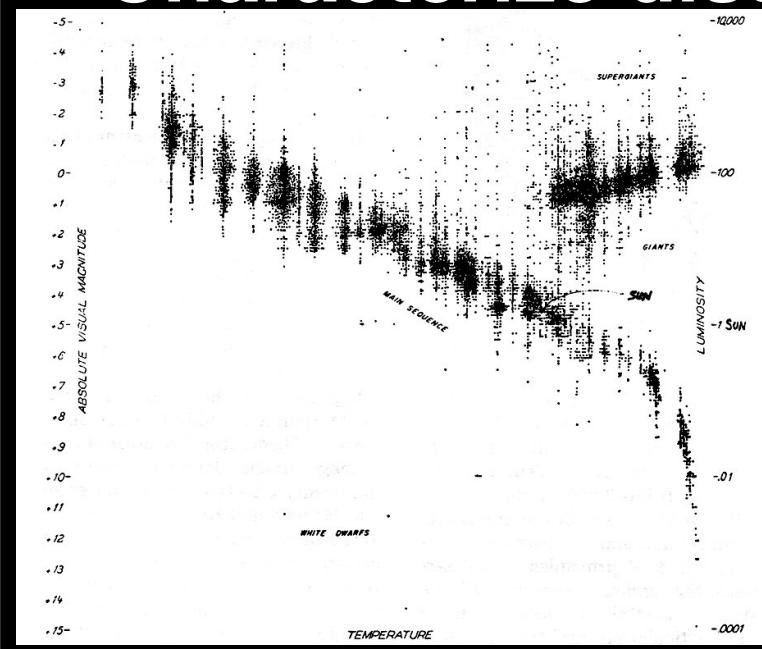


[Amar, Eagan, & Stasko, InfoVis '05]

31

31

Characterize distribution



[Amar, Eagan, & Stasko, InfoVis '05]

32

32

Find anomalies

Identify any anomalies in a set of data cases with respect to a given relationship or expectation (e.g. statistical outliers)

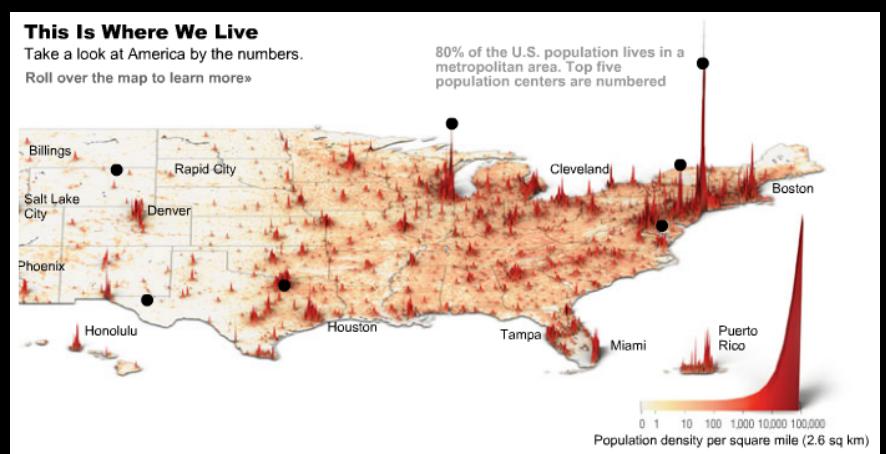
[Amar, Eagan, & Stasko, InfoVis '05]

33

33

Find anomalies

Identify any anomalies in a set of data cases with respect to a given relationship or expectation (e.g. statistical outliers)



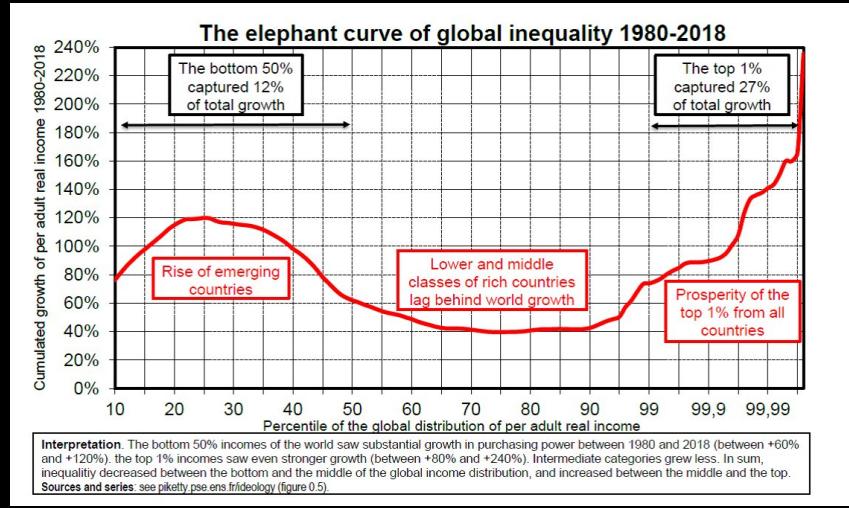
[Amar, Eagan, & Stasko, InfoVis '05]

34

34

Find anomalies

Identify any anomalies in a set of data cases with respect to a given relationship or expectation (e.g. statistical outliers)



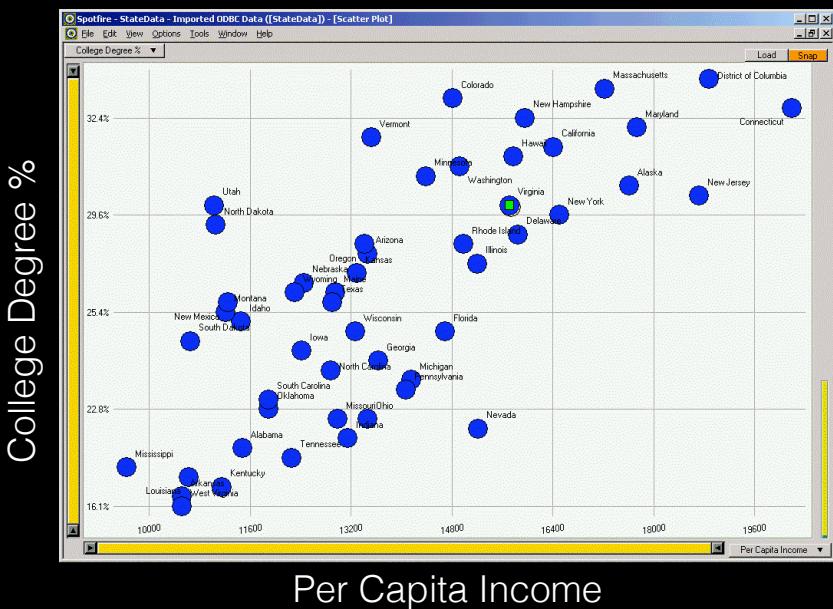
[Amar, Eagan, & Stasko, InfoVis '05]

35

35

Find anomalies

Identify any anomalies in a set of data cases with respect to a given relationship or expectation (e.g. statistical outliers)



[Amar, Eagan, & Stasko, InfoVis '05]

36

36

Cluster

Within a set of cases, identify any clusters of similar attribute values

[Amar, Eagan, & Stasko, InfoVis '05]

37

37



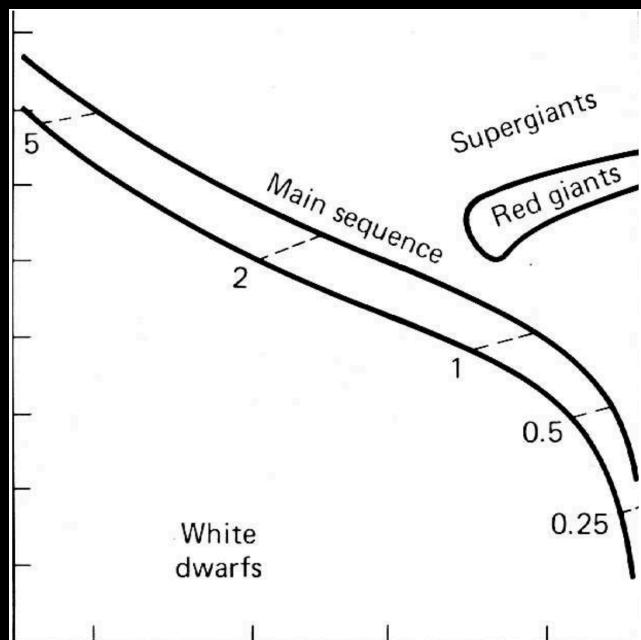
[Amar, Eagan, & Stasko, InfoVis '05]

38

38

Cluster

Within a set of cases, identify any clusters of similar attribute values



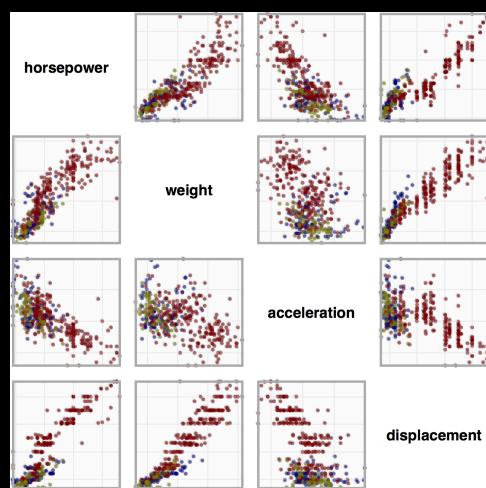
[Amar, Eagan, & Stasko, InfoVis '05]

39

39

Correlate

For a given set of cases and two attributes, determine useful relationships between the values of those attributes



[Amar, Eagan, & Stasko, InfoVis '05]

40

40

Low-level analytic tasks

- Retrieve value
- Filter
- Compute derived value
- Find extremum
- Sort
- Determine range
- Characterize distribution
- Find anomalies
- Cluster
- Correlate

[Amar, Eagan, & Stasko, InfoVis '05]

41

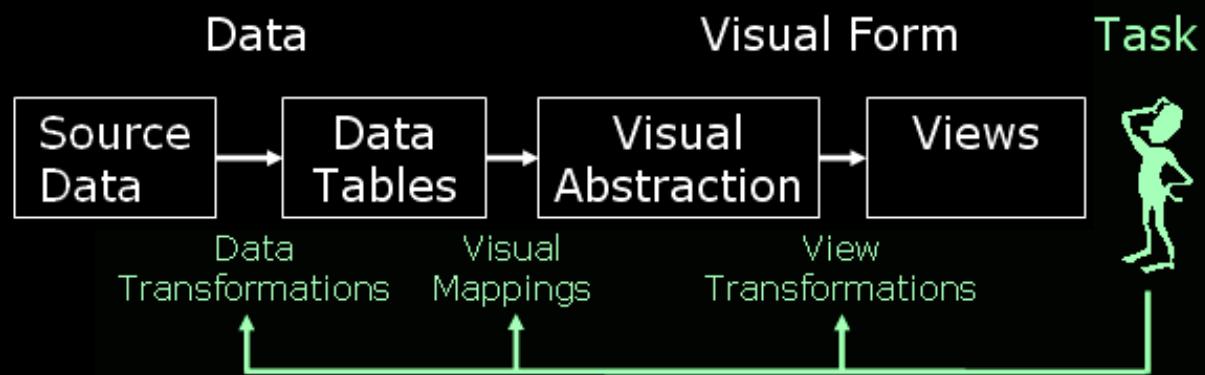
41

**How do people
interact with visualizations?**

42

42

InfoVis Pipeline



43

43

Kinds of interaction

Select: mark something as interesting

Explore: show something else

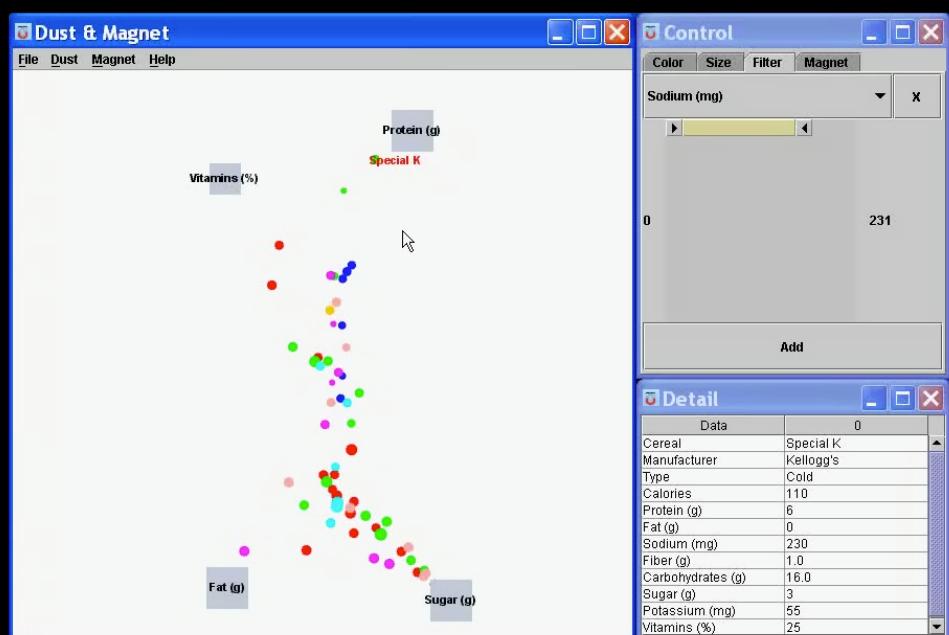
Reconfigure: show a different arrangement

Encode: show a different representation

Abstract/Elaborate: show less or more detail

Filter: show something conditionally

Connect: show related items

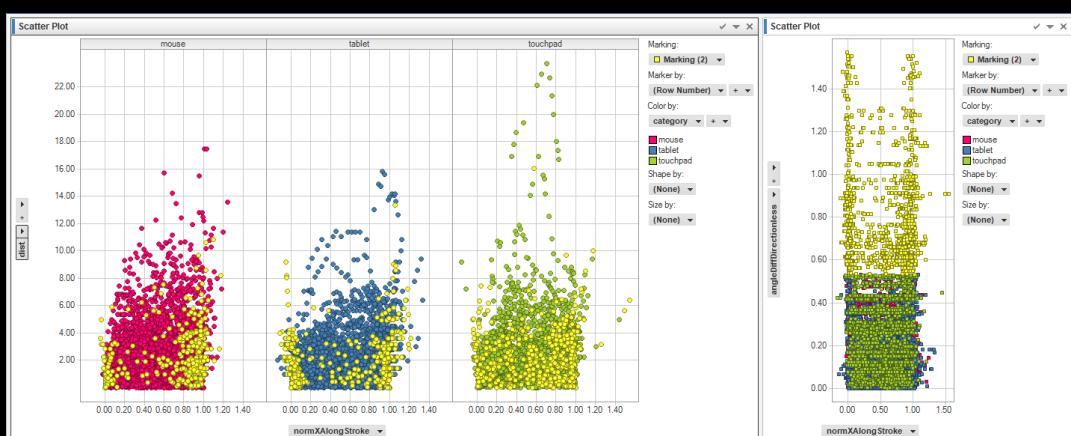


Select

[Yi et al., InfoVis '05]

45

45

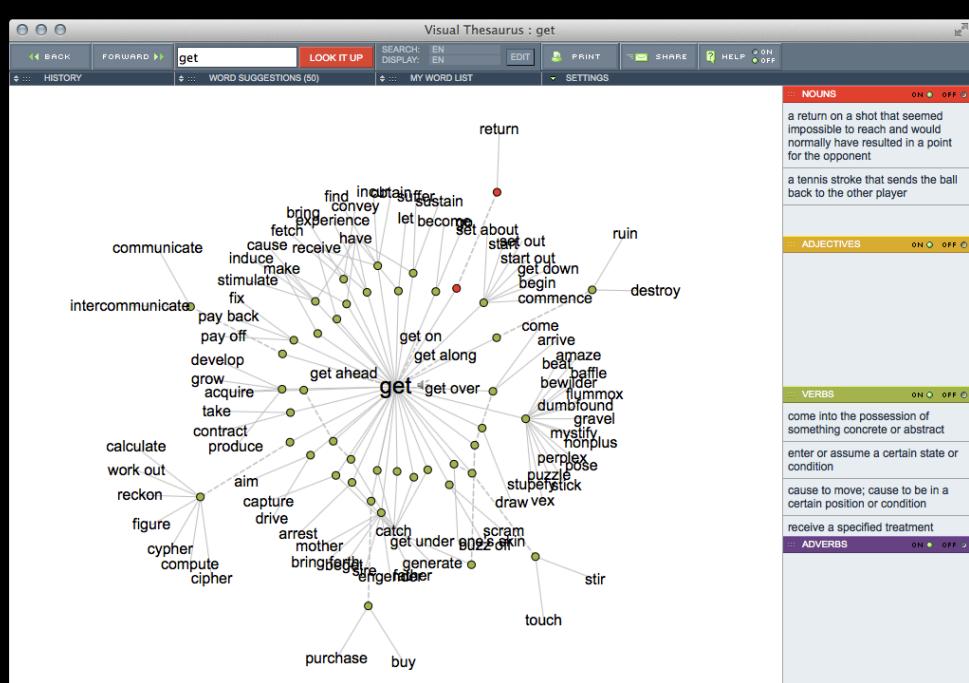


Select

[Ahlberg, SIGMOD Record '96]

46

46

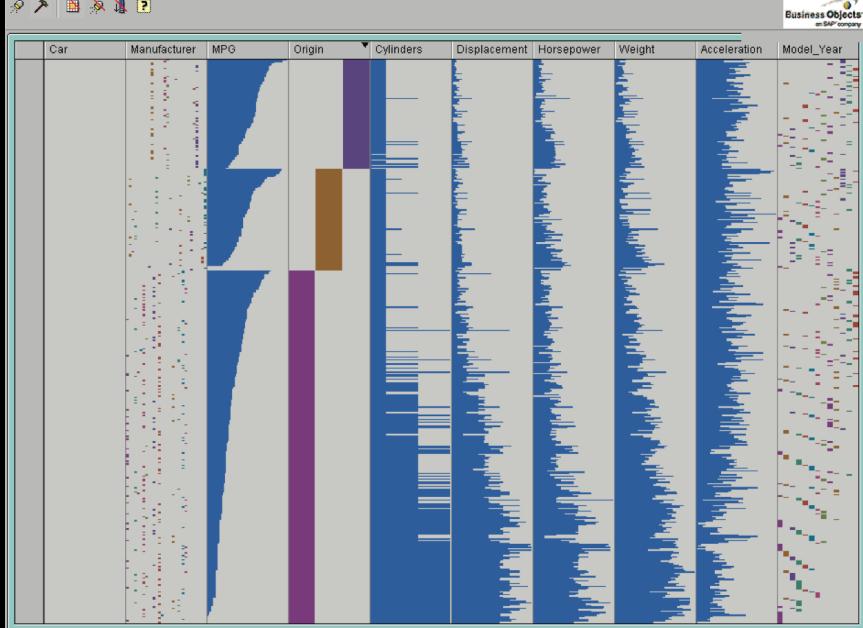


Explore

[www.visualthesaurus.com]

47

47

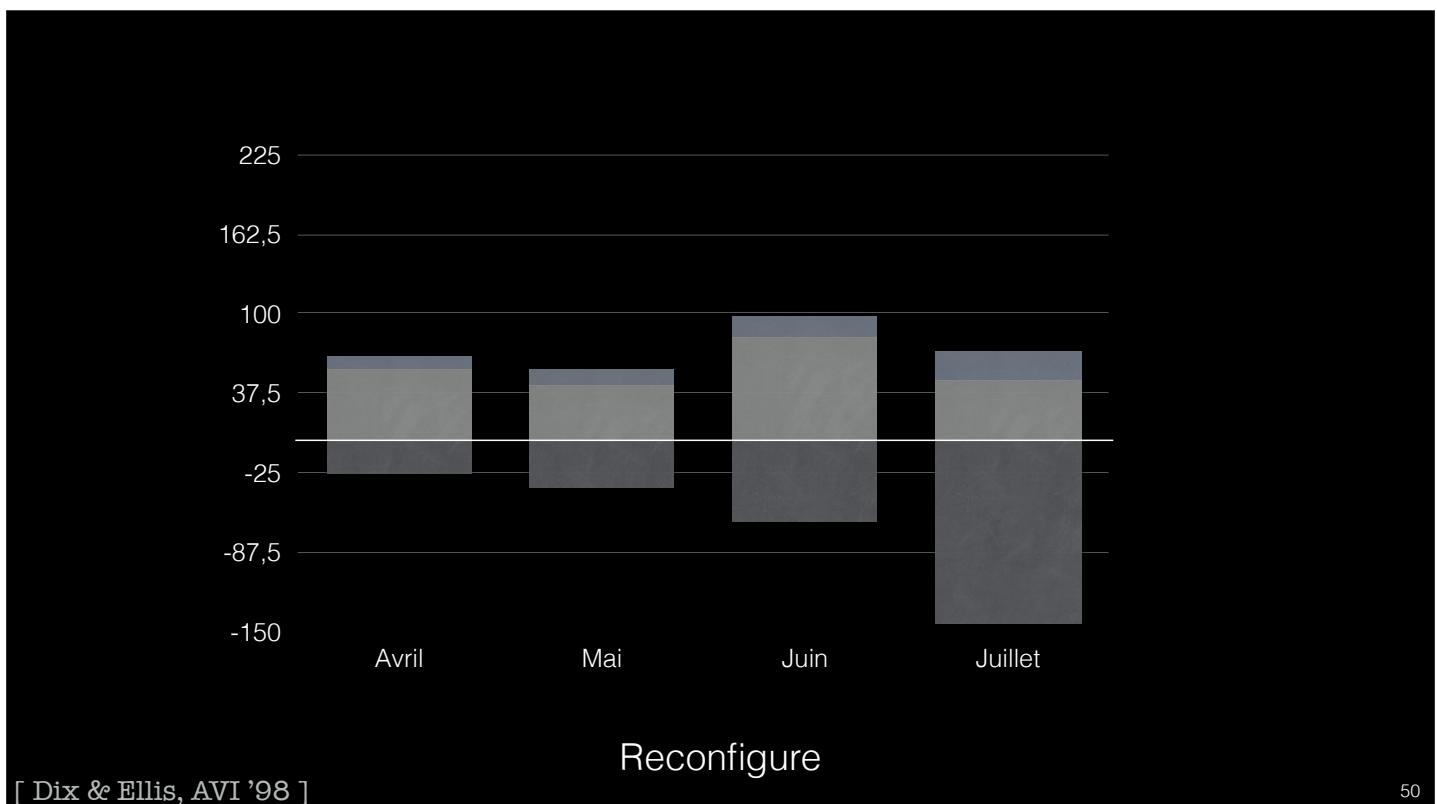
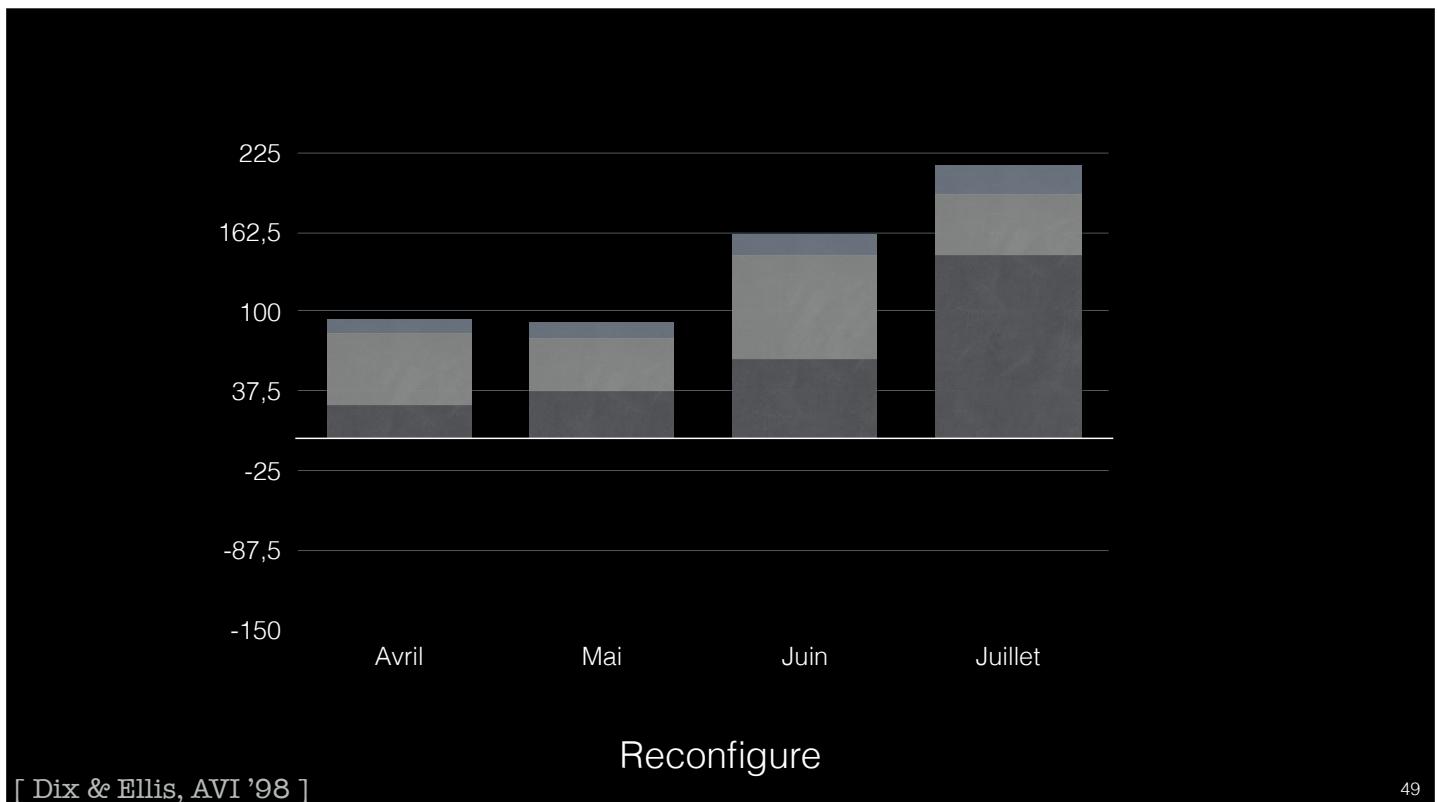


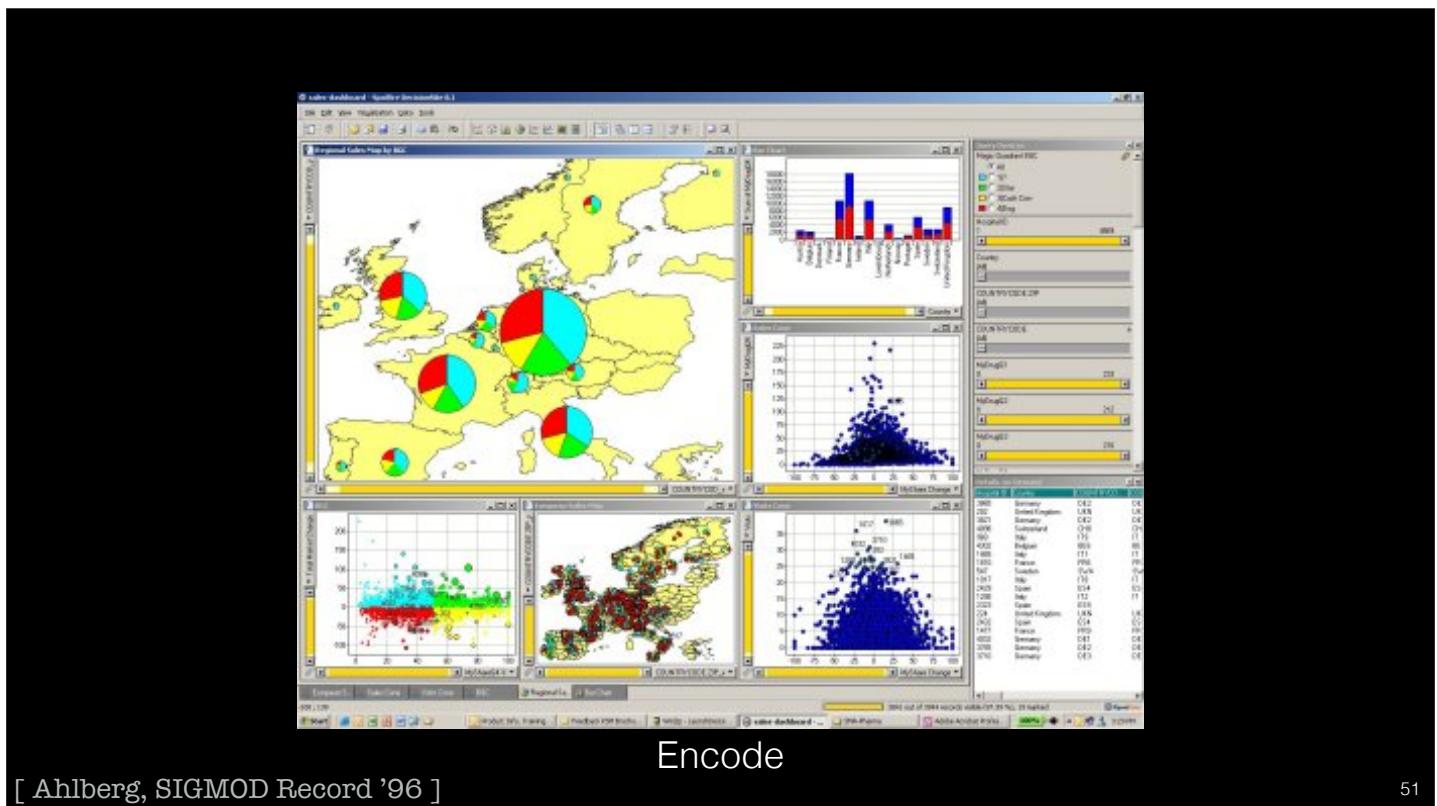
Reconfigure

[Rao & Card, CHI '94]

48

48





51

51



52

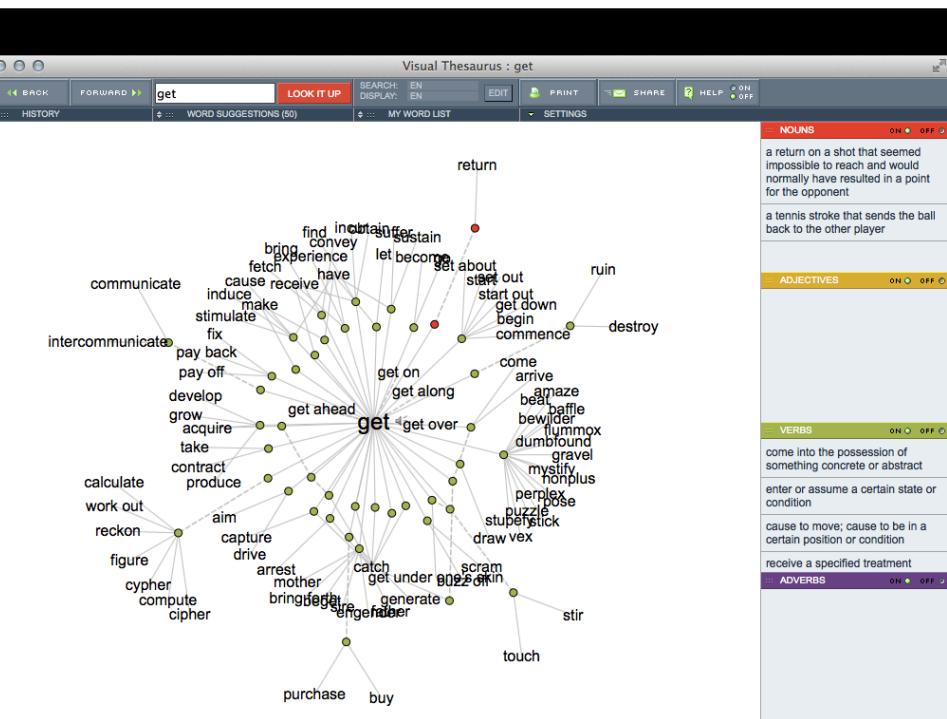


Filter

[Ahlberg et al., CHI '92]

53

53



Connect

[www.visualthesaurus.com]

54

54

Kinds of interaction

Select: mark something as interesting

Explore: show something else

Reconfigure: show a different arrangement

Encode: show a different representation

Abstract/Elaborate: show less or more detail

Filter: show something conditionally

Connect: show related items

[Yi et al., TVCG '07]

55

**How do we
put this all together?**

56

Shneiderman's mantra

- Overview first, zoom & filter, details on demand
- Overview first, zoom & filter, details on demand
- Overview first, zoom & filter, details on demand

57

57

Putting it all together

- *Problem: Finding a home*

58

58

Dynamic queries

- Query language
- **Select** house-address
From sf-realty-db
Where price >= 200,000 **and**
 price <= 400,000 **and**
 bathrooms >= 3 **and**
 garage == 2 **and**
 bedrooms >= 4

59

59

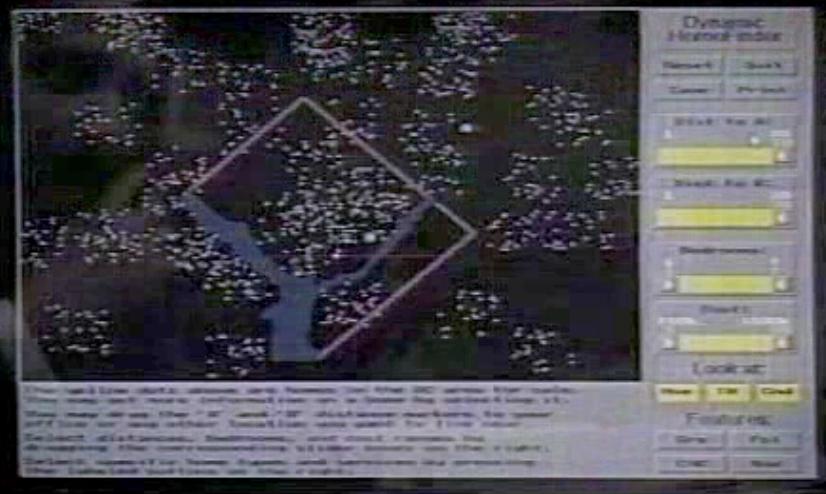
Dynamic queries

- Query language
- **Select** house-address
From sf-realty-db
Where price >= 200,000 **and**
 price <= 400,000 **and**
 bathrooms >= 3 **and**
 garage == 2 **and**
 bedrooms >= 4

60

60

Dynamic queries



[Ahlberg & Shneiderman, '92]

61

61

Response time

- 0.1 sec — *animation, visual continuity*
- 1 sec — *system response, conversation break*
- 10 sec — *cognitive response*

62

62

Case Study: Twitter Log Data

63

63

Twitter logs

- Suppose you have logs of user events on Twitter

Component	Description	Example
client	client application	web, iphone, android
page	page or functional grouping	home, profile, who_to_follow
section	tab or stream on a page	home, mentions, retweets, searches
component	component, object, or objects	search_box, tweet
element	UI element within the component	button, avatar
action	actual user or application action	impression, click, hover

Field	Description
event_name	event name
user_id	user id
ip	user's IP address
timestamp	timestamp
event_details	event details

Terabytes per day!

web:home:mentions:stream:avatar:profile_click

65

65

web:home:mentions:stream:avatar:profile_click

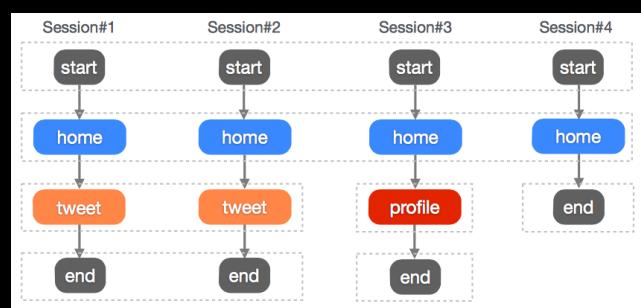
66

66

Overview

67

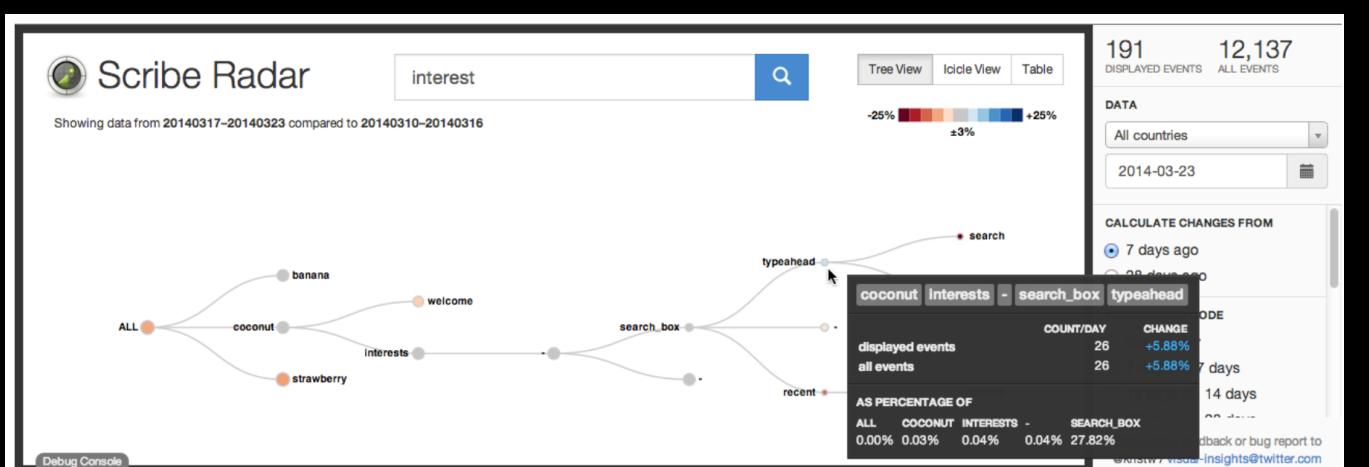
67



Task: See changes in patterns

69

69

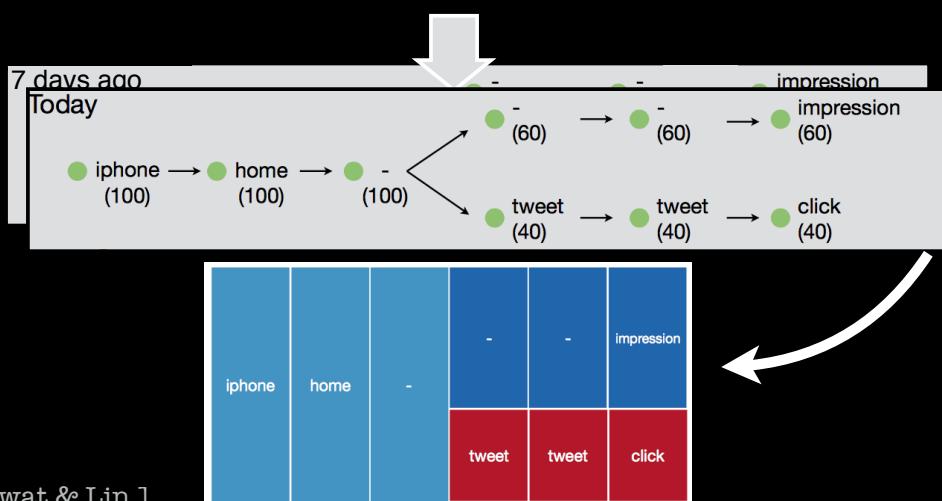


[Wongsuphasawat & Lin, VAST '14]

70

70

client	page	section	component	element	action	count
iphone	home	-	-	-	impression	60
iphone	home	-	tweet	tweet	click	40



[Wongsuphasawat & Lin]

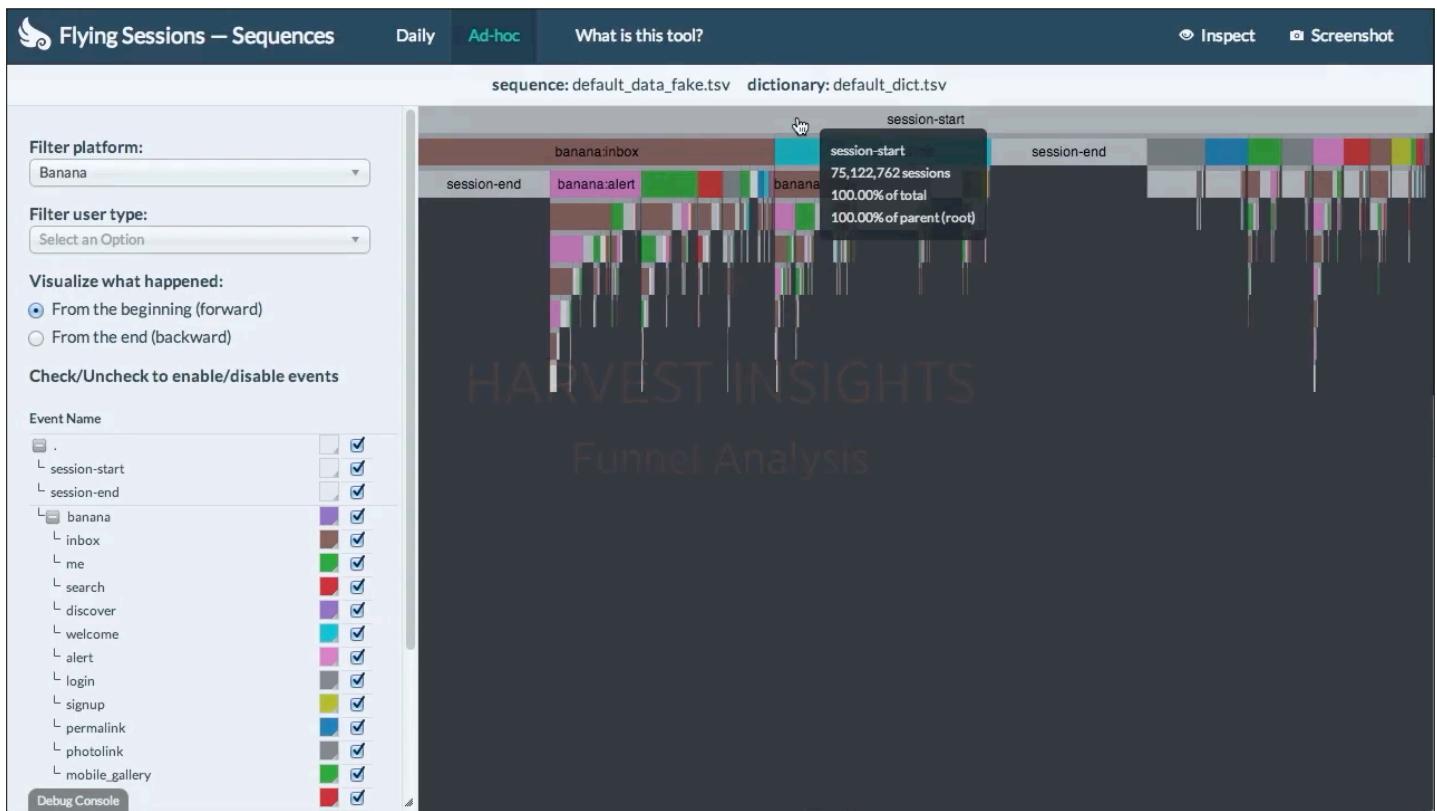
71

71

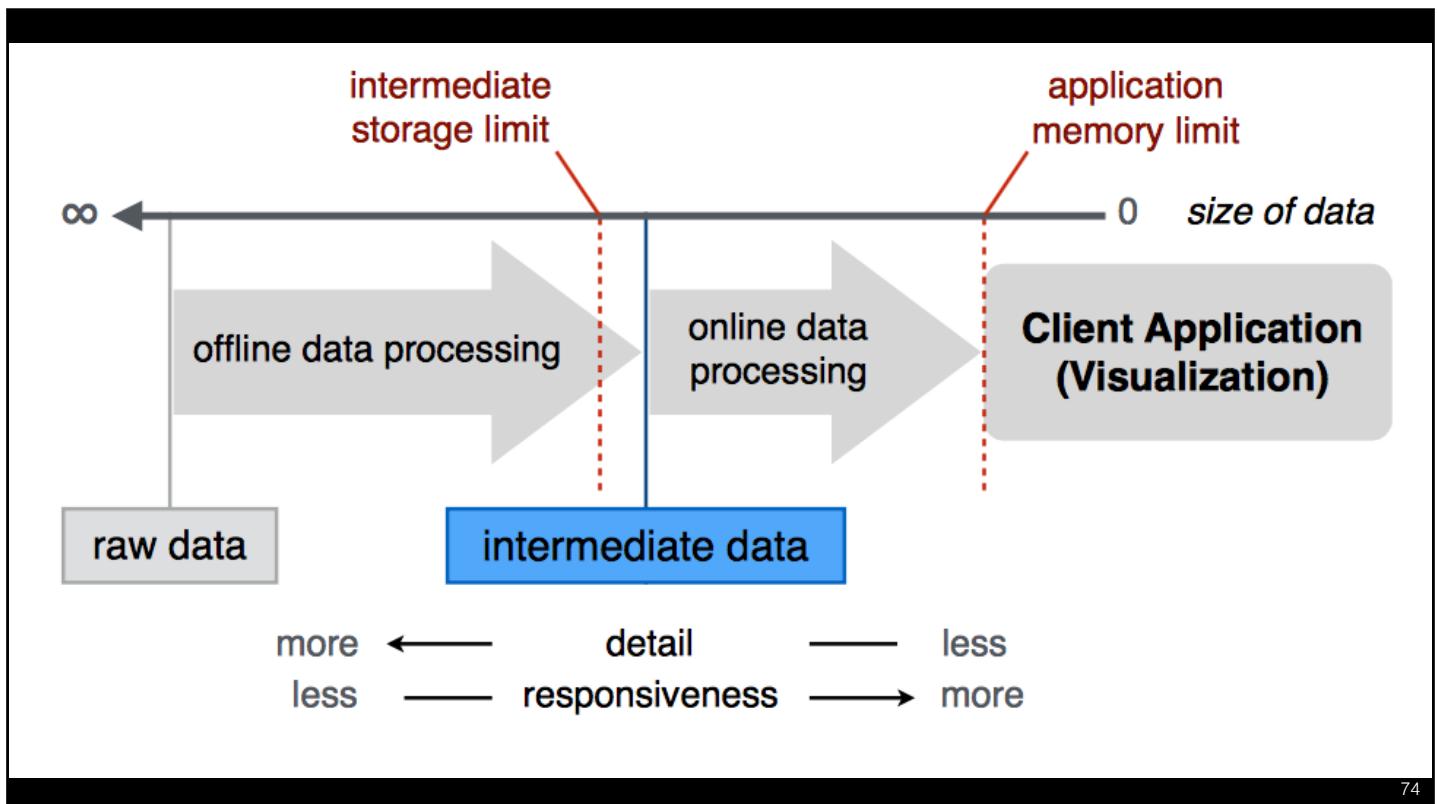
Maintaining Interactivity

72

72



73



74

Aggregation

75

75

web:home:mentionCons:stream:available:profile_click

76

76

ABCDEDEF → ABCDEF

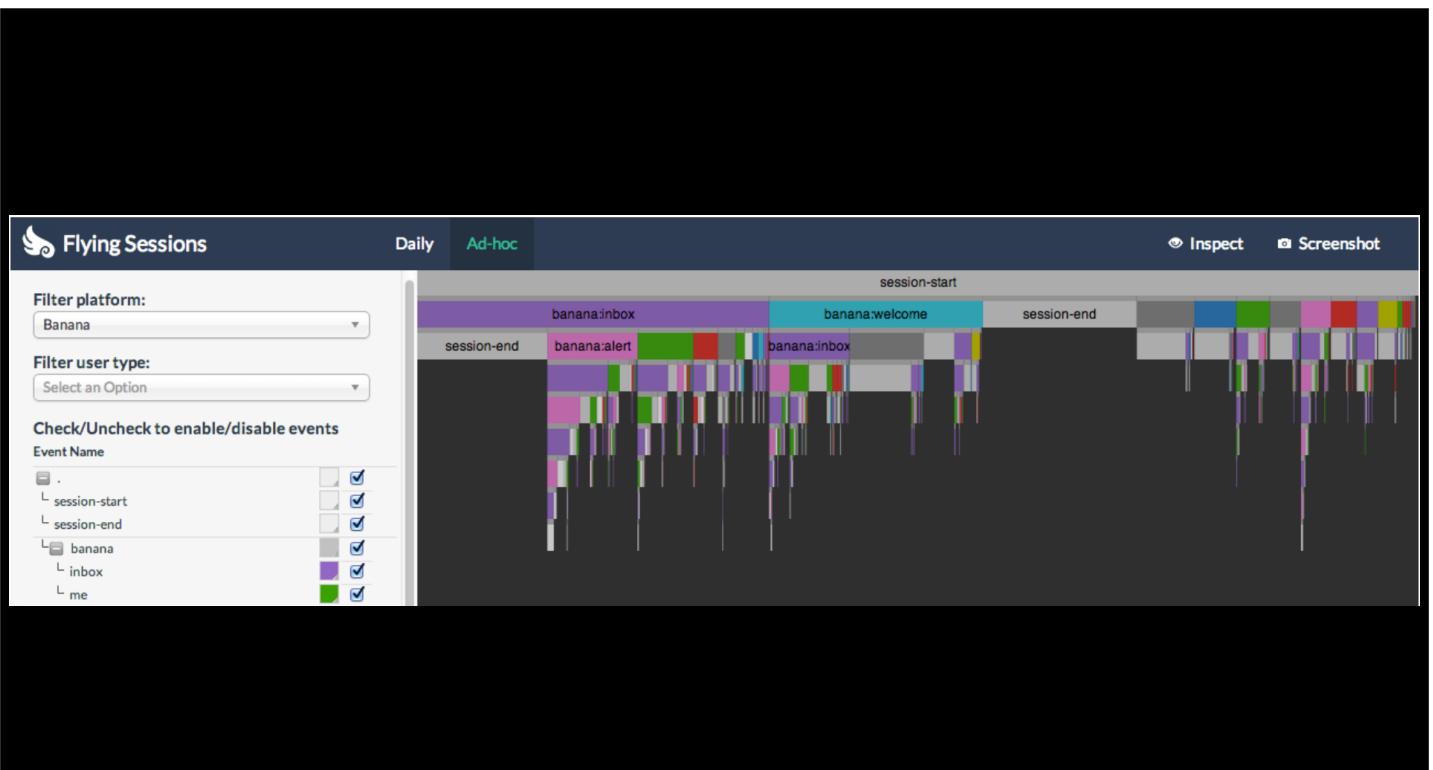
ABC (2000)
ABCD (80)
ABCE (20)
ABCDF (1)
ABCDG (1)

ABC (2000)
ABCD (80)
ABCE (20)
ABCDx (2)

ABC (2000)
ABCx (102)

77

77



[Wongsuphasawat & Lin, VAST '14]

78

78