

Paper: Polaris

Paper: D3

Paper: Vega-Lite

Tamara Munzner

Department of Computer Science

University of British Columbia

CPSC 547, Information Visualization

Week 11: 13 November 2025

www.cs.ubc.ca/~tmm/courses/547-25

Plan for today

- peer review one direction
- break
- peer review other direction
- mini-lecture on systems papers

Paper: D3 System

Paper: D3

- paper types
 - design studies
 - technique/algorithm
 - evaluation
 - model/taxonomy
 - **system**

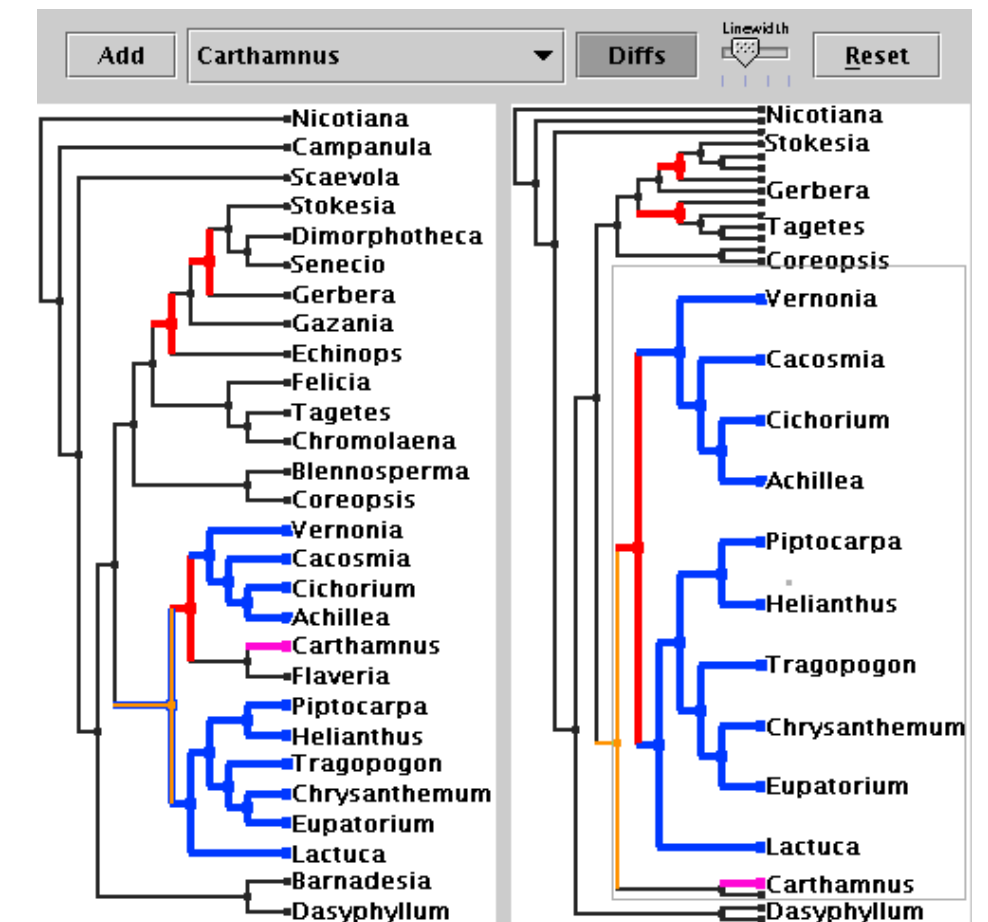
[D3: Data-Driven Documents. Bostock, Ogievetsky, Heer. IEEE Trans. Visualization & Comp. Graphics (Proc. InfoVis), 2011.]

Toolkits

- imperative: how
 - low-level rendering: Processing, OpenGL
 - parametrized visual objects: prefuse
 - also flare: prefuse for Flash
- declarative: what
 - Protoviz, D3, ggplot2
 - separation of specification from execution
- considerations
 - expressiveness
 - can I build it?
 - efficiency
 - how long will it take?
 - accessibility
 - do I know how?

WebGL/OpenGL

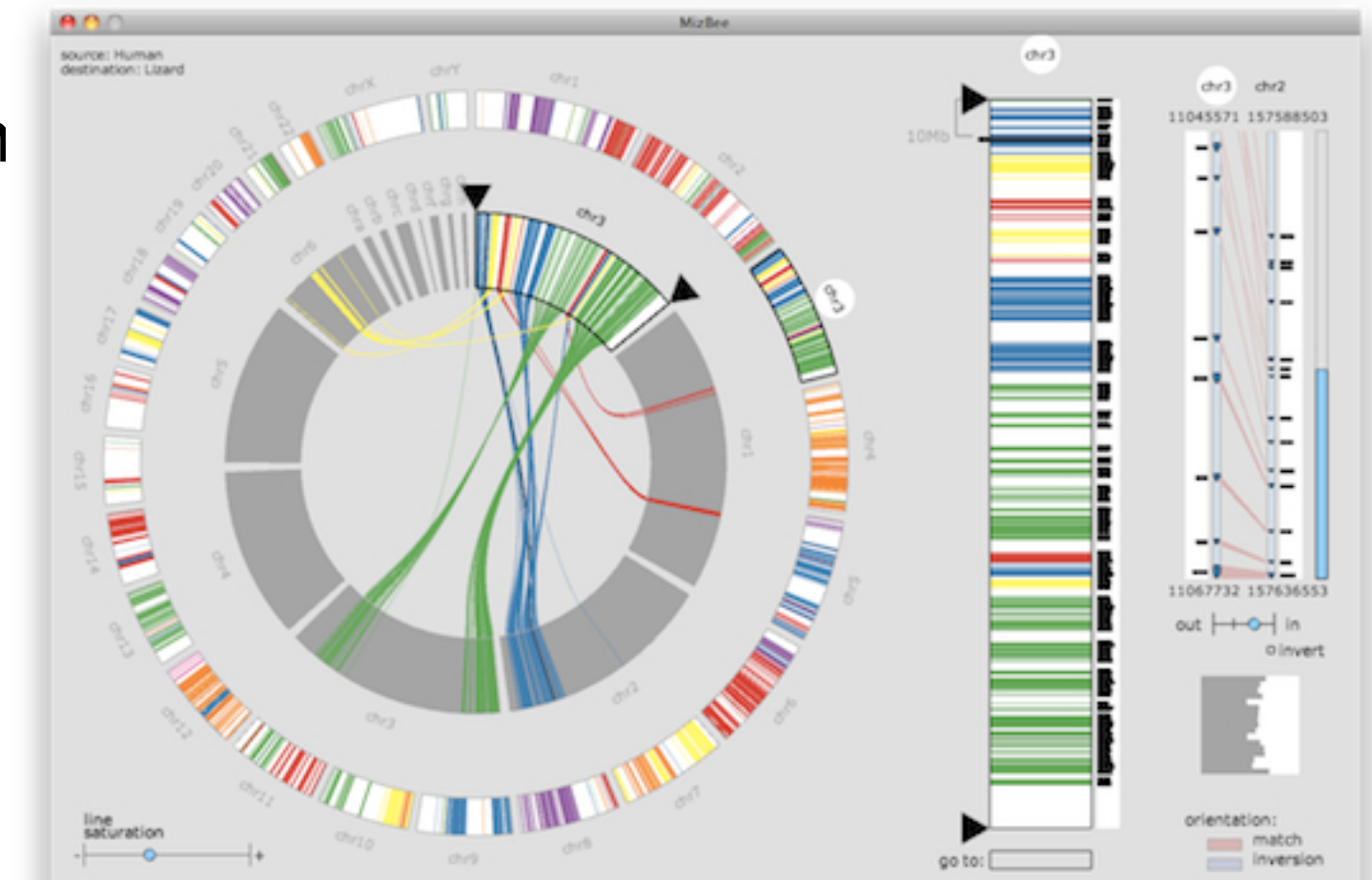
- graphics library
 - pros
 - power and flexibility, complete control for graphics
 - hardware acceleration
 - many language bindings: js, C, C++, Java (w/ JOGL)
 - cons
 - big learning curve if you don't know already
 - no vis support, must roll your own everything
 - example app: TreeJuxtaposer (OpenGL)



[Fig 5. Munzner et al. TreeJuxtaposer: Scalable Tree Comparison using Focus+Context with Guaranteed Visibility. Proc SIGGRAPH 2003, pp 453-462.]

Processing / p5.js

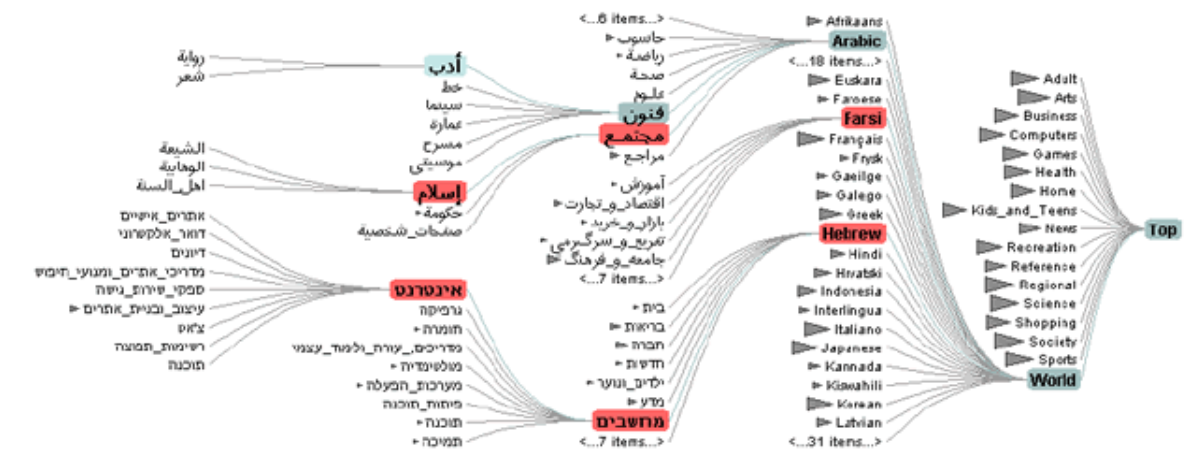
- layer on top of Java/OpenGL, Javascript/WebGL
- visualization esp. for artists/designers
- pros
 - great sandbox for rapid prototyping
 - huge user community, great documentation
- cons
 - poor widget library support
- example app: MizBee



[Fig 1. Meyer et al. MizBee: A Multiscale Synteny Browser. Proc. InfoVis 2009.]

prefuse

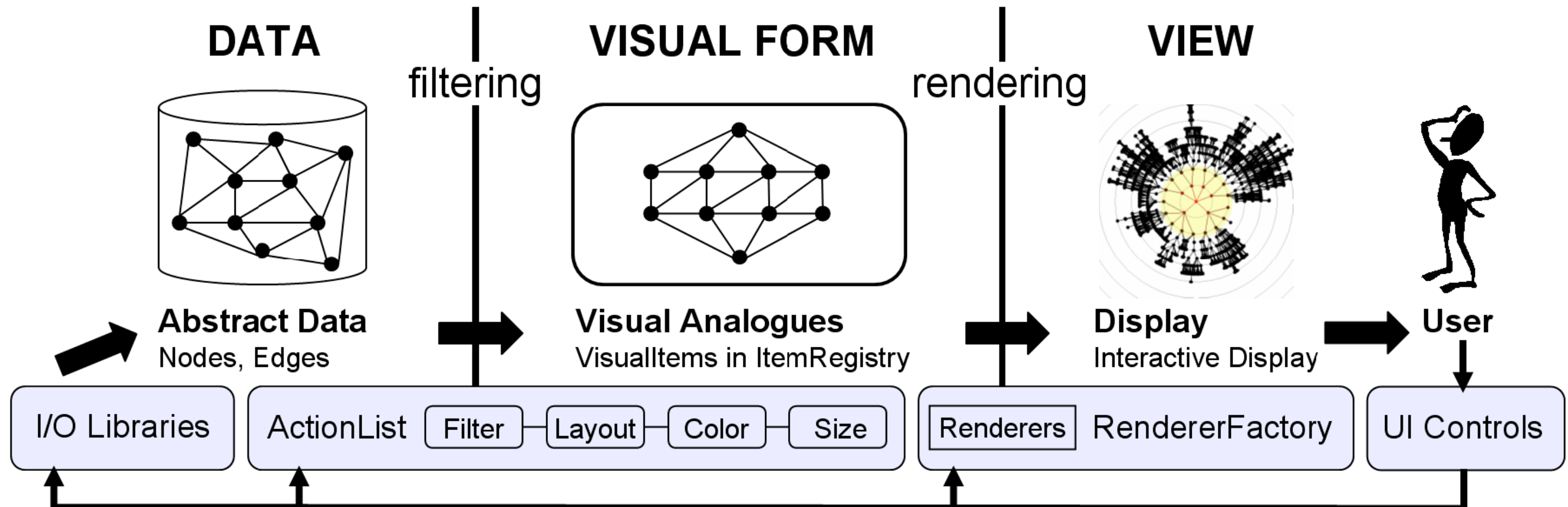
- infovis toolkit, in Java
- fine-grained building blocks for tailored visualizations
- pros
 - heavily used (previously)
 - very powerful abstractions
 - quickly implement most techniques covered so far
- cons
 - no longer active
 - nontrivial learning curve
- example app: DOI Trees Revisited



[DOITrees Revisited: Scalable, Space-Constrained Visualization of Hierarchical Data. Heer and Card. Proc. Advanced Visual Interfaces (AVI), pp. 421–424, 2004.]

prefuse

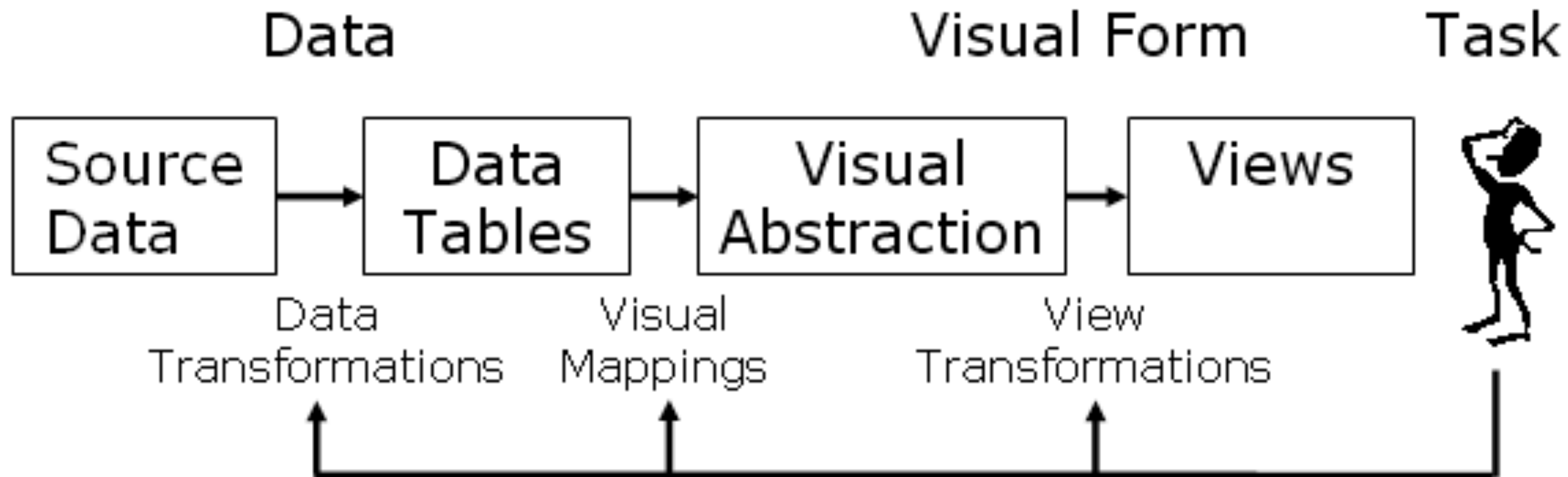
- separation: abstract data, visual form, view
 - data: tables, networks
 - visual form: layout, color, size, ...
 - view: multiple renderers



[Fig 2. Heer, Card, and Landay. Prefuse: A Toolkit for Interactive Information Visualization. Proc. CHI 2005, 421-430]

InfoVis Reference Model

- conceptual model underneath design of prefuse and many other toolkits
- heavily influenced much of infovis (including nested model)
 - aka infovis pipeline, data state model



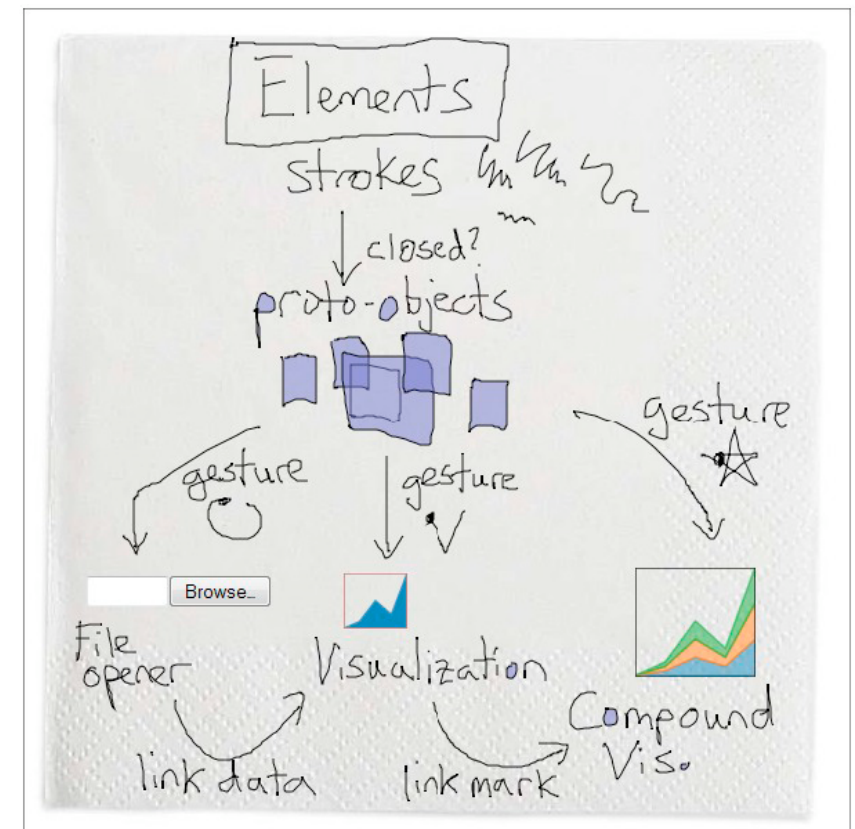
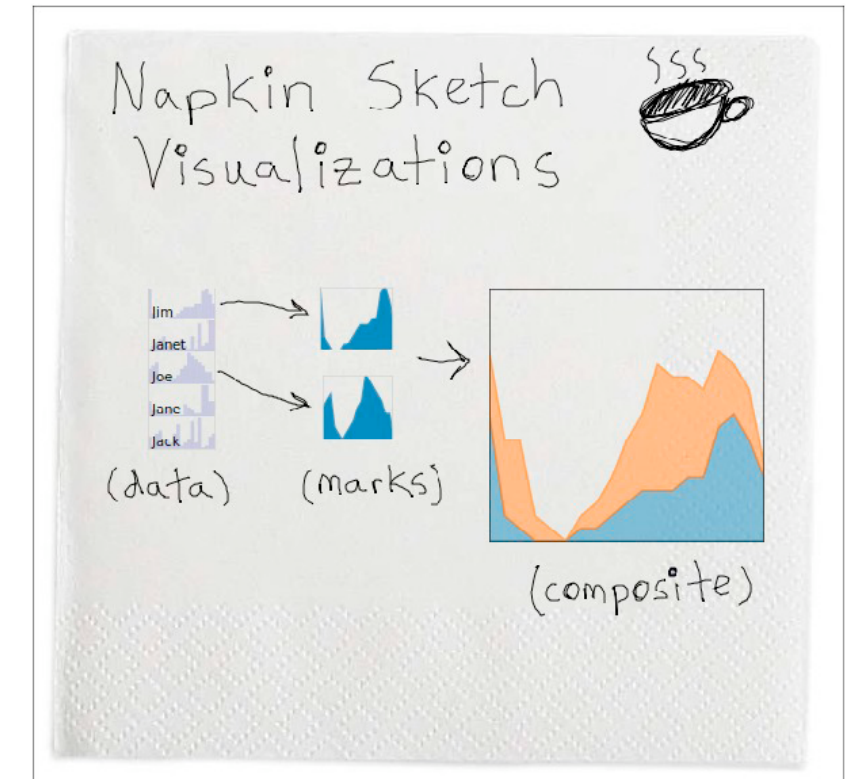
[Redrawn Fig 1.23. Card, Mackinlay, and Shneiderman. *Readings in Information Visualization: Using Vision To Think*, Chapter 1. Morgan Kaufmann, 1999.]

Declarative toolkits

- imperative tools/libraries
 - say exactly **how** to do it
 - familiar programming model
 - OpenGL, prefuse, ...
- declarative: other possibility
 - just say **what** to do
 - Protovis, D3

Protovis

- declarative infovis toolkit, in Javascript
 - also later Java version
- marks with inherited properties
- pros
 - runs in browser
 - matches mark/channel mental model
 - also much more: interaction, geospatial, trees,...
- cons
 - not all kinds of operations supported
- example app: NapkinVis (2009 course project)



Protovis Validation

- wide set of old/new app examples
 - expressiveness, effectiveness, scalability
 - accessibility
- analysis with cognitive dimensions of notation
 - closeness of mapping, hidden dependencies
 - role-expressiveness visibility, consistency
 - viscosity, diffuseness, abstraction
 - hard mental operations

[Cognitive dimensions of notations. Green (1989). In A. Sutcliffe and L. Macaulay (Eds.) People and Computers V. Cambridge, UK: Cambridge University Press, pp 443-460.]

D3

- declarative infovis toolkit, in Javascript
- Protovis meets Document Object Model
- pros
 - seamless interoperability with Web
 - explicit transforms of scene with dependency info
 - massive user community, many thirdparty apps/libraries on top of it, lots of docs
- cons
 - even more different from traditional programming model
- example apps: many

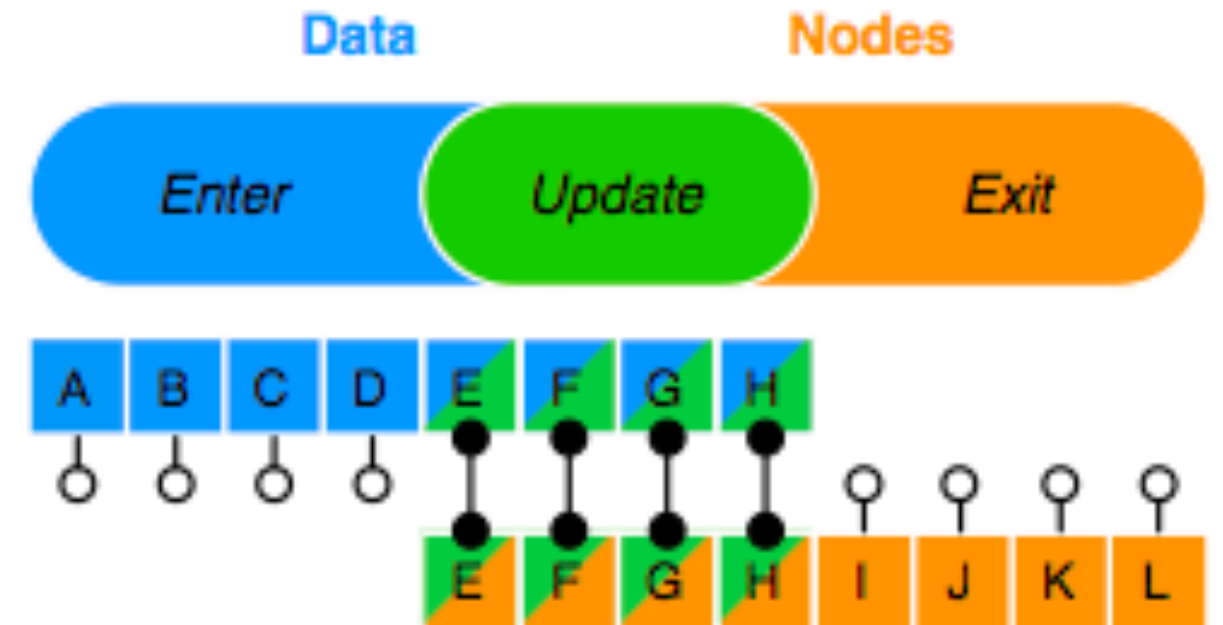
D3

- objectives
 - compatibility
 - debugging
 - performance
- related work typology
 - document transformers
 - graphics libraries
 - infovis systems
 - general note: all related work sections are a mini-taxonomy/typology!

[D3: Data-Driven Documents. Bostock, Ogievetsky, Heer. IEEE Trans. Visualization & Comp. Graphics (Proc. InfoVis), 2011.]

D3 capabilities

- query-driven selection
 - selection: filtered set of elements queries from the current doc
 - also partitioning/grouping!
 - operators act on selections to modify content
 - instantaneous or via animated transitions with attribute/style interpolators
 - event handlers for interaction
- data binding to scenegraph elements
 - data joins bind input data to elements
 - enter, update, exit subselections
 - sticky: available for subsequent re-selection
 - sort, filter



[D3: Data-Driven Documents. Bostock, Ogievetsky, Heer. *IEEE Trans. Visualization & Comp. Graphics (Proc. InfoVis)*, 2011.]

D3 Features

- document transformation as atomic operation
 - scene changes vs representation of scenes themselves
- immediate property evaluation semantics
 - avoid confusing consequences of delayed evaluation
- validation
 - performance benchmarks
 - page loads, frame rate
 - accessibility
 - (adoption)
 - everybody has voted with their feet by now!

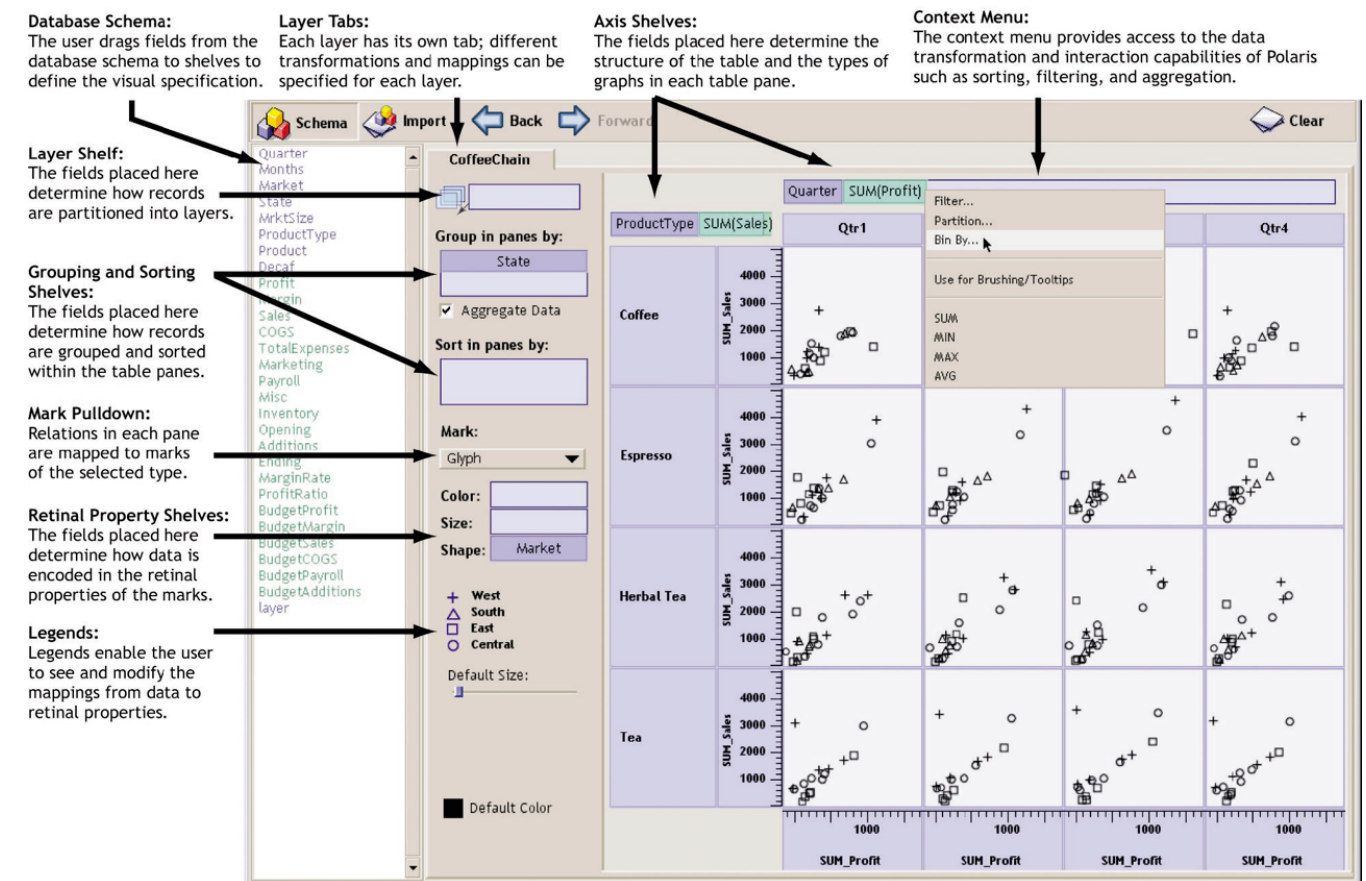
Paper: Polaris/Tableau System

Polaris

A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases

Chris Stolte, Diane Tang, Pat Hanrahan

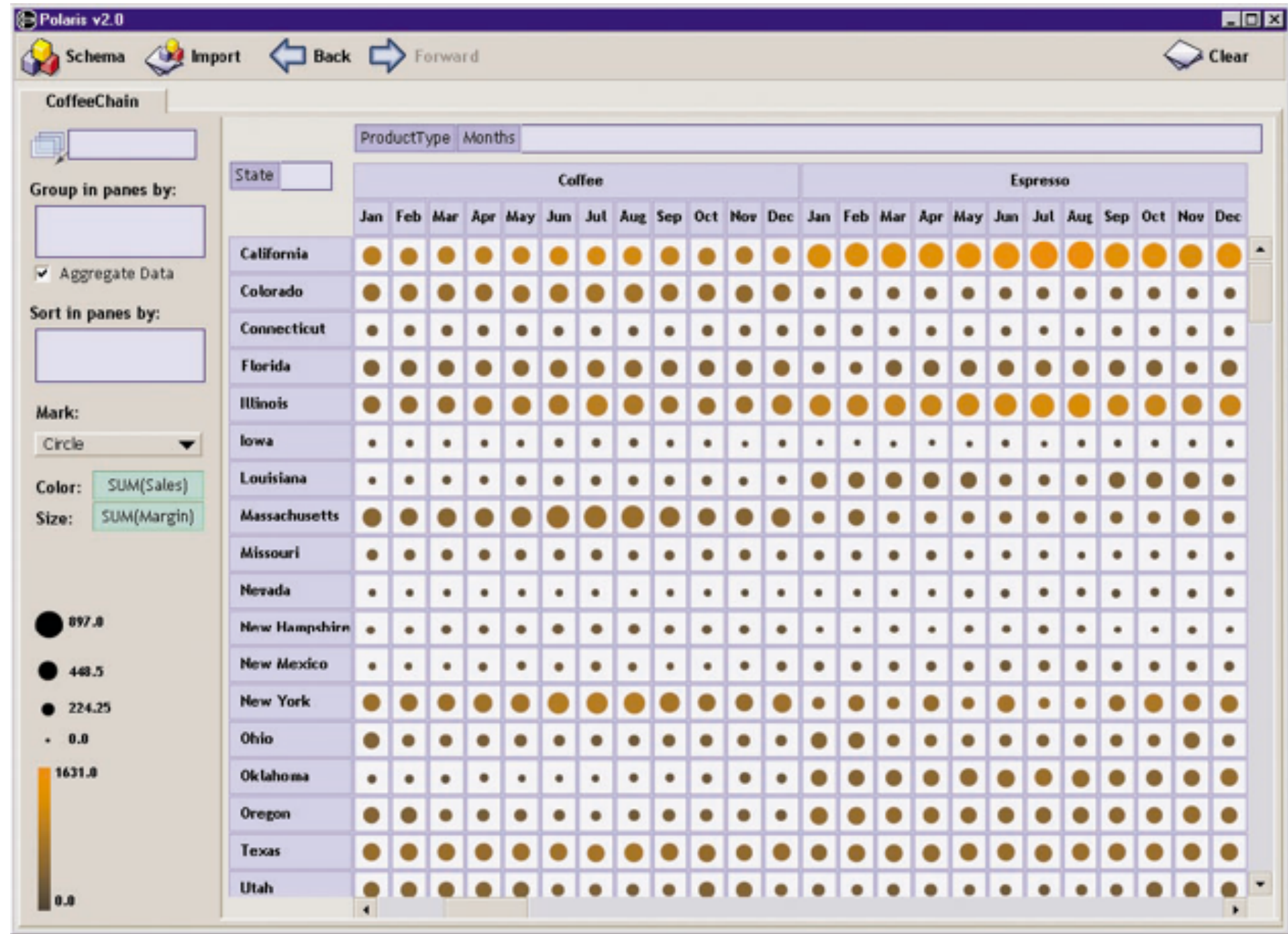
<http://www.graphics.stanford.edu/projects/polaris/>



Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases.
Stolte, Tang and Hanrahan, *IEEE TVCG* 8(1):52-65 2002

Polaris: Stolte, Tang, and Hanrahan

- infovis spreadsheet
 - table cells have graphical elements, not just numbers
 - wide range of channels and marks
- example
 - marks: circles
 - color channel: saturation
 - size channel: area
 - partition: state x product:month
 - ord x ord

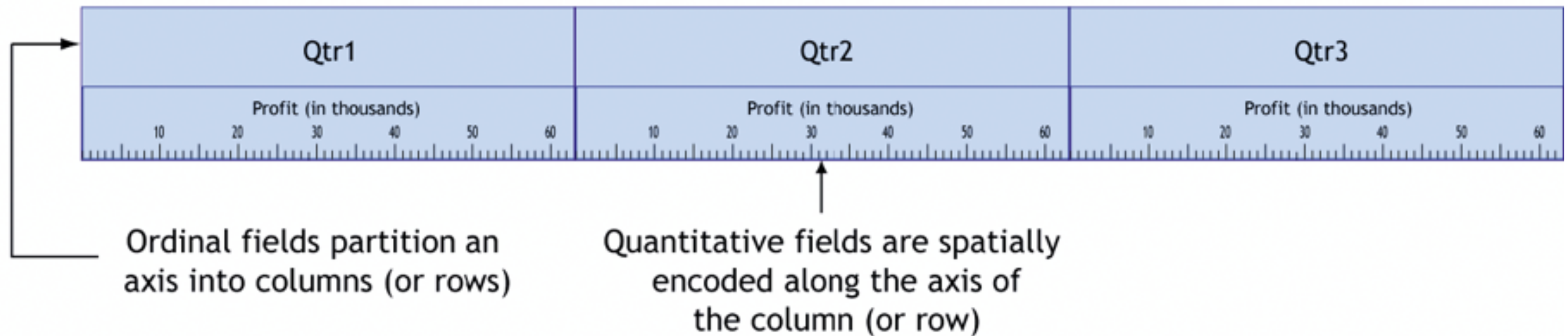


[Fig 3a. Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases. Stolte, Tang and Hanrahan, IEEE TVCG 8(1):52-65 2002.]

Table Algebra :: Interactive Interface

- drag and drop actions map to formal language underneath
 - partitioning using shelves
 - different results for ord vs quant

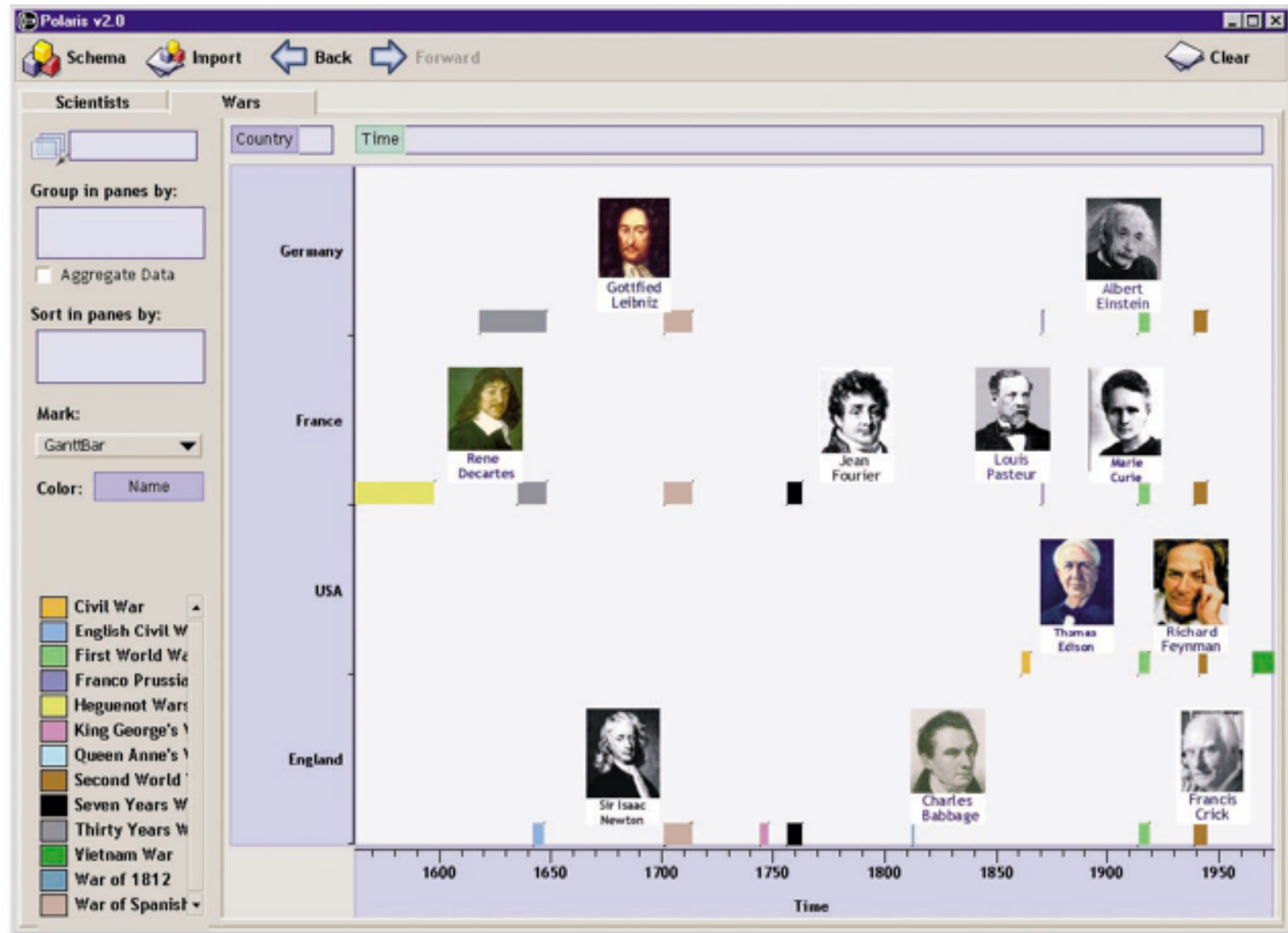
$O \times Q = \text{Quarter} \times \text{Profit} = \{(\text{Qtr1}, \text{Profit}), (\text{Qtr2}, \text{Profit}), (\text{Qtr3}, \text{Profit}), (\text{Qtr4}, \text{Profit})\}$:



[Fig 2. Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases. Stolte, Tang and Hanrahan, IEEE TVCG 8(1):52-65 2002.]

Polaris

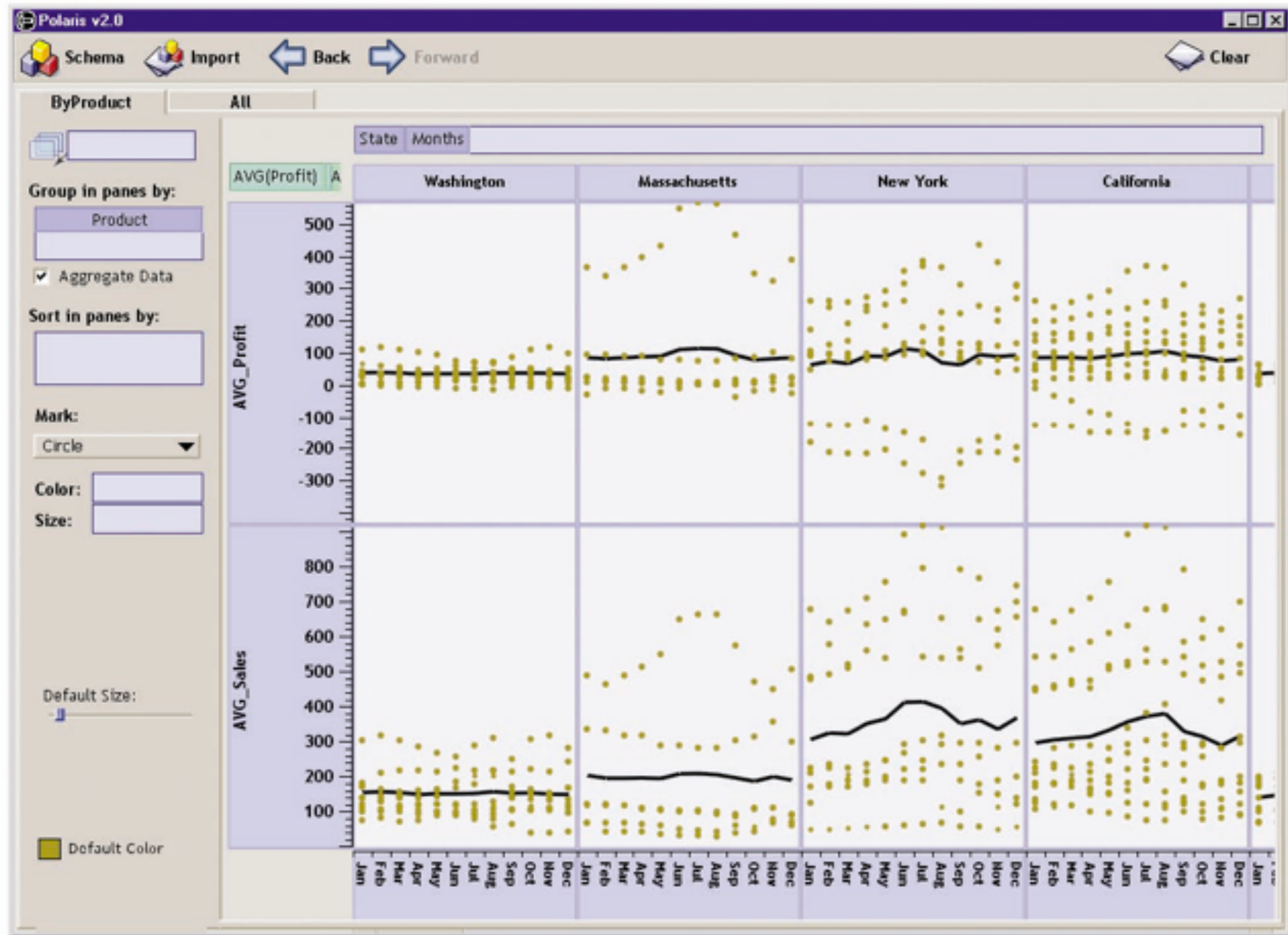
- example
 - marks: Gantt chart bars
 - color channels: nominal / categorical
 - spatial position channels: country x year
 - ord x quant



[Fig 3b. Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases. Stolte, Tang and Hanrahan, IEEE TVCG 8(1):52-65 2002.]

Polaris

- example
 - views: scatterplots
 - marks: points
 - spatial position channels:
profit x month
 - quant x (2 ord)



[Fig 3d Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases. Stolte, Tang and Hanrahan, IEEE TVCG 8(1):52-65 2002.]

Terminology I: Now and Upcoming

- Marks and Channels
 - retinal variables/properties: *visual channels*
 - mark: *mark*
- Data Abstraction
 - column or field: *attribute*
 - nominal: *categorical*
 - ordinal: *ordered*
 - quantitative: *quantitative*
 - row or record: *item*
 - dimension / independent / ordinal: *key attribute*
 - all ordinal fields treated as dimensions in Polaris
 - measure / dependent : *value attribute*
 - all quantitative fields treated as measures in Polaris

Terminology II: Upcoming

- Data Abstraction
 - deriving data
- Map Color and Other Channels
 - hue: *hue*
 - value: *saturation*
 - brightness: *luminance*
- Manipulate View
 - sorting
- Facet Into Multiple Views
 - pane: *view*
 - partitioning
 - brushing: *linked highlighting*
- Reduce Items and Attributes
 - aggregation, filtering

Polaris: Pre and post

- influences
 - Bertin's Semiology of Graphics book (1967 / 1998)
 - Wilkinson's Grammar of Graphics book (1999 / 2005)
 - Mackinlay's APT paper/system (1986)
 - Cleveland's Visualizing Data book (1993)
- Stolte and Hanrahan commercialized as Stanford spinoff Tableau Software
 - major success story in vis, \$2B IPO in 2013
 - Mackinlay joined in 2004, Wilkinson joined in 2014
- Tableau use in this course
 - very useful for analysis projects
 - possible sandbox for experimentation when starting programming projects
 - you can request free student license, good for one year
 - <http://www.tableau.com/academic/students>