View Reduction Lecture Notes

Slide 4

Recall that Tableau makes a lot of use of different aggregation measures.

Slide 6

Figure 13.2(a) shows the full dataset. The interaction design features filtering, with immediate update of the visual display to filter out or add back items as sliders are moved and buttons are pressed. The visual encoding adapts to the number of items to display; the marks representing movies are automatically enlarged and labelled when enough of the dataset has been filtered away that there is enough room to do so, as in Figure 13.2(b). The system uses multiform overview–detail views, where clicking on any mark brings up a popup detail view with more information about that movie, as in Figure 13.2(c).

This uses a dynamic queries approach, where browsing using tightly coupled visual encoding and interaction is an alternative to searching by running queries, as for example with a database. All of the items in a database are shown at the start of a session to provide an overview, and direct manipulation of interface widgets replaces reformulating new queries. This approach is well suited to browsing, as opposed to a search where the user must formulate an explicit query that might return far too many items, or none at all.

Paper on FilmFinder: https://web.stanford.edu/class/archive/cs448b/cs448b.1166/cgi-bin/wiki/images/e/e2/FinalPaperCS448b.pdf

Rather dated video presentation https://www.youtube.com/watch?v=g9JadyYUyK8

Slide 8

Dataset of 215 attributes representing word counts and 298 points representing documents in a collection of medical abstracts. Applied here to star plots.

Slide 9

Tableau has a large number of aggregate functions: https://help.tableau.com/current/pro/desktop/en-us/calculations_calculatedfields_aggregate_create.htm

Slide 10

A short mathematical justification of the rule of thumb for the number of bins can be found on this page https://en.wikipedia.org/wiki/Histogram#:~:text=Variable%20bin%20widths,-Rather%20than%20choosing&text=When%20plotting%20the%20histogram%2C%20the,histogram%20approximate%20the%20density%20distribution.&text=%3B%20the%20coefficient%20of%202%20is,value%20from%20this%20broad%20optimum in the section labelled 'Remark'.
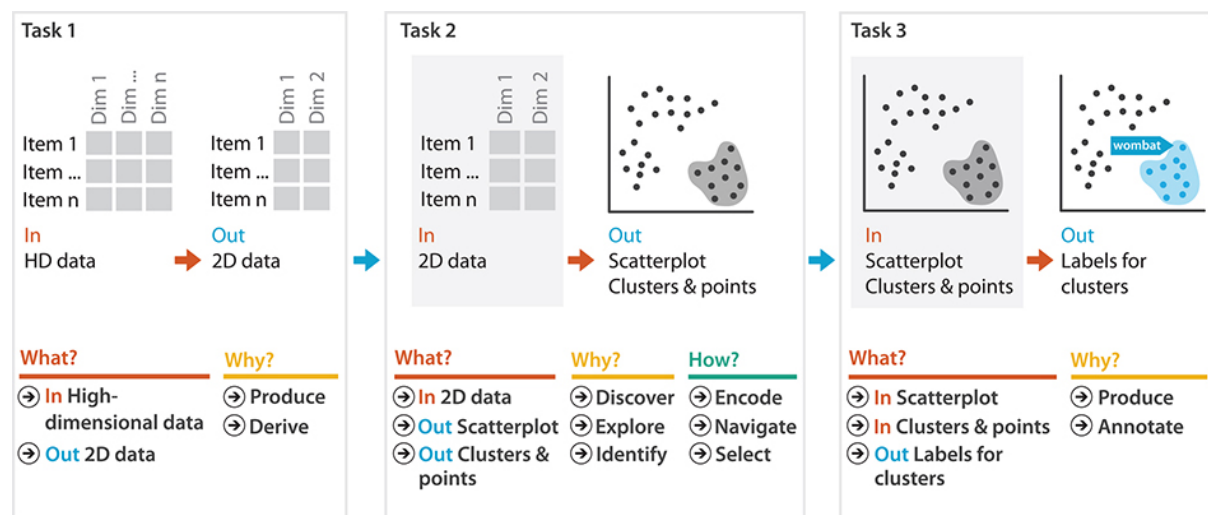
Slide 12

Tableau calls *fences*, *whiskers* (as in box-and-whisker plot) and uses them to display all points within 1.5 times the interquartile range (in other words, all points within 1.5 times the width of the adjoining box), or all points at the maximum extent of the data. Because these

plots are capturing information about distributions, it is best to turn off aggregation (recall that it is on by default in Tableau).

Slide 13

The level of detail displayed at a global level for the entire dataset can be interactively controlled by the user using a single slider. The parameter controlled by that slider is again a derived variable that varies the aggregate level of detail shown in a smooth and continuous way. The figure shows a dataset with eight attributes and 230,000 items at different levels of detail. Figure (a) is the highest-level overview showing the single top-level cluster, with very broad bands of green. Figure (b) is the midlevel view showing several clusters, where the extents of the tan cluster are clearly distinguishable from the now-smaller green one. Figure (c) is a more detailed view with dozens of clusters that have tighter bands; the proximity-based colouring mitigates the effect of occlusion.

Slide 16

What-why-how analysis of dimensionality reduction of document collection data.

Slide 19

The DOI function can be thought of as a continuous function that does not explicitly distinguish between focus items to show in detail, context items to aggregate, and completely elided items to filter out. Those interpretations are made by algorithms that use the function, often based on threshold values. These interest functions typically exploit knowledge about dataset structure, especially hierarchical relationships. For example, if a few subsections of a document were selected to be the foci, then a good context would be their enclosing sections.

The shaded triangles provide an aggregate representation showing the size of the elided subtrees. The context in which to show them is computed using tree traversal from the many focus nodes up toward their common ancestors and the tree root. In this case, distance is computed topologically based on hops through the tree, rather than geometrically through Euclidean space. The focus nodes can be chosen explicitly by clicking, or indirectly through searching.

Slide 20

Original paper https://research.tableau.com/sites/default/files/1993-ToolglassMagicLenses.pdf

Video of the original system https://www.youtube.com/watch?v=v7M3yw4Y71I Look at the screens in use back in 2008!

Slide 22

Two examples of a fisheye lens used with an online poker player dataset. The scatterplot in Figure (a) shows the percentage of time that a player goes to showdown (playing until people have to show all of their cards) versus the flop (playing until the point where three cards are placed face-up on the board). In the dense matrix view of Figure (b), blocks representing players are colour coded according to their winning rate, and a space-filling curve is used to lay out these blocks in order of a specific derived attribute; in this case, a particular betting strategy.

Slide 23

Hyperbolic geometry and Escher:
https://web.colby.edu/thegeometricviewpoint/2016/12/21/tessellations-of-the-hyperbolic-plane-and-m-c-escher/

Slide 24

The distortion idiom of hyperbolic geometry uses a single radial global focus with the interaction metaphor of hyperbolic translation. This approach exploits the mathematics of non-Euclidean geometry to elegantly accommodate structures such as trees that grow by an exponential factor, in contrast to standard Euclidean geometry where there is only a polynomial amount of space available for placing items.

An infinite non-Euclidean plane can be mapped to a finite Euclidean circle, and similarly an infinite non-Euclidean volume can be mapped to a finite sphere in Euclidean space. The interaction metaphor is hyperbolic translation, which corresponds to changing the focus point of the projection; the visual effect is changing which items are magnified at the centre, versus minimized at the periphery, for a global effect with similarities to using a fisheye lens that extends across the entire scene.