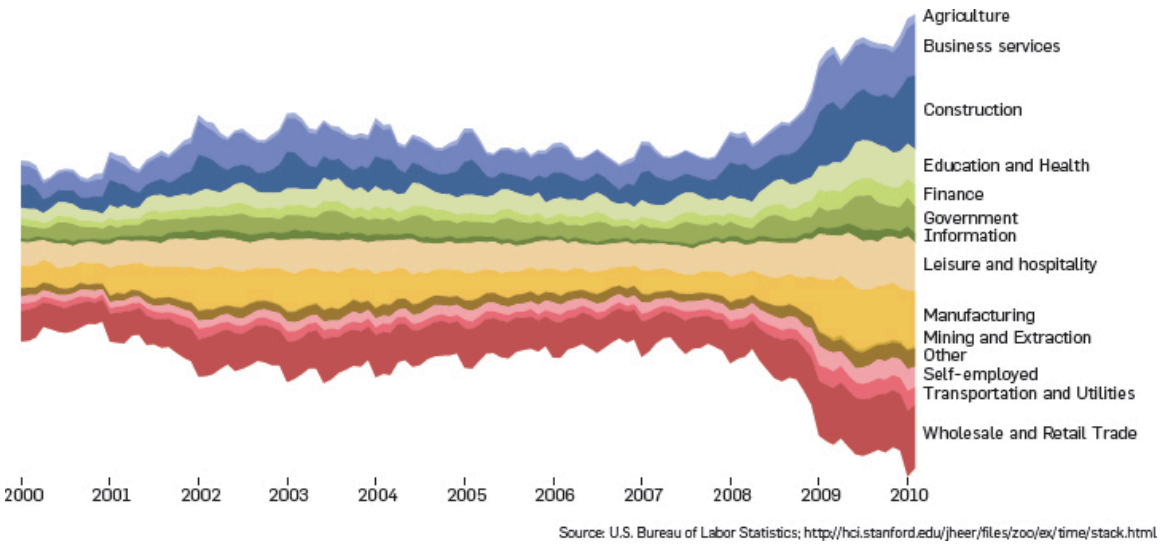
Nathaniel Burbank
March 28, 2006
High Dimensional Data Exploration

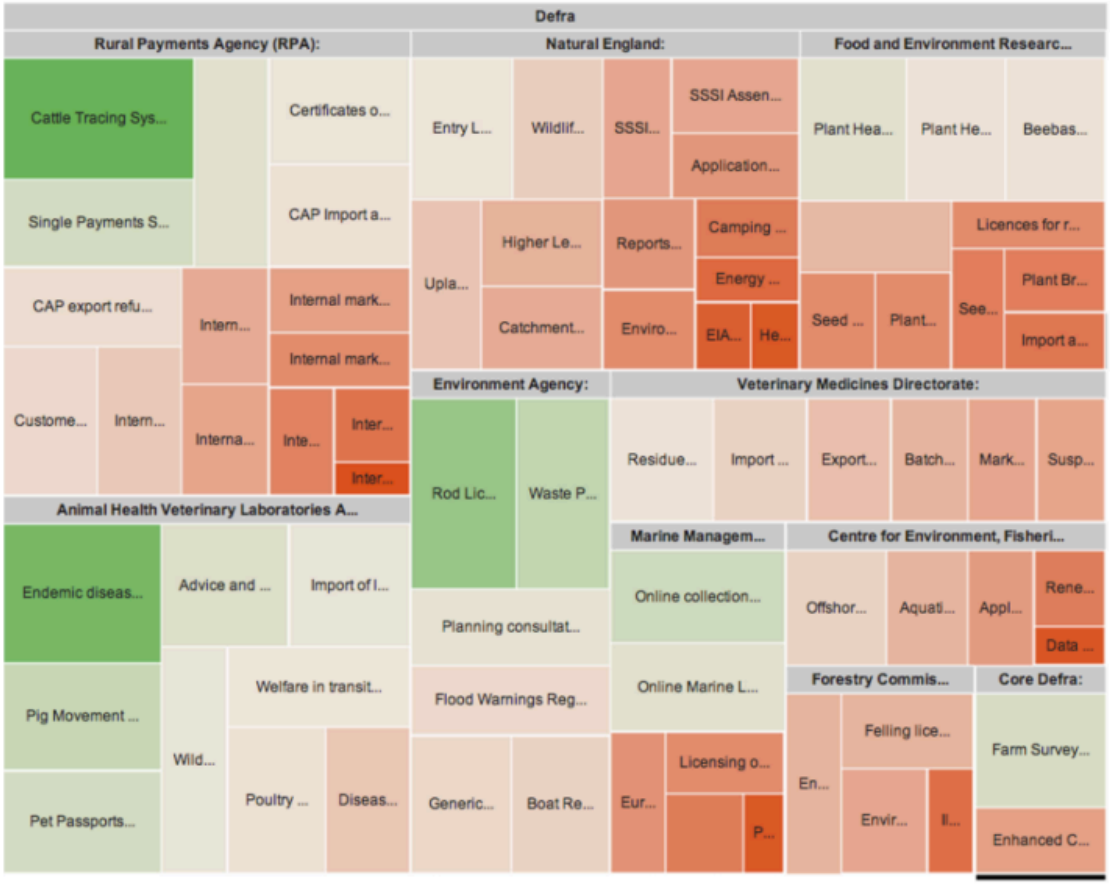
Example 1 – Stacked area chart



Source: http://deliveryimages.acm.org/10.1145/1750000/1743567/figs/f1b.jpg

A stacked area chart is probably the most standard way to display multidimensional time series data. While this is not "high-dimensional" per se, we will very likely be using a visualization like this in our project. A small but important variation of the stacked area chart above is the stacked percentage chart. Since our project concerns changes in proportions overtime, a natural (and again very standard way) to display our main dataset this would be via stacked area chart where the y-axis displays percentage for a given year rather than the raw values. This highlights relative changes in proportion that can be hard to see if there are also large aggregate changes taking place simultaneously.

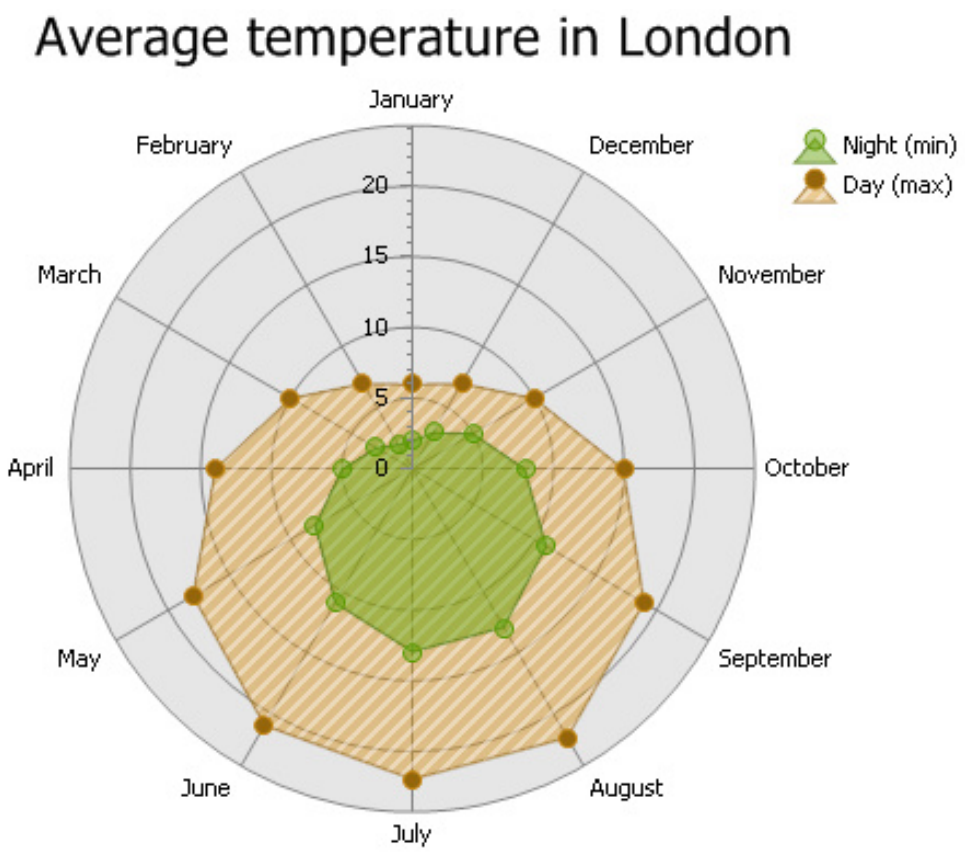
Example 2 – a Tree Map



Source : http://blog.ouseful.info/2012/08/02/pragmatic-visualisation-gds-transaction-data/

Another natural way to explore the types of hierarchical-data that we're working with would be a tree-map. This makes it easy to both see the larger categories while also displaying the detail of the varies sub-categories, allowing the user to use colors to display a third variable. A tree map could work well for us to display aggregate spending at a slice in time, and we could use animation to display the changes overtime.

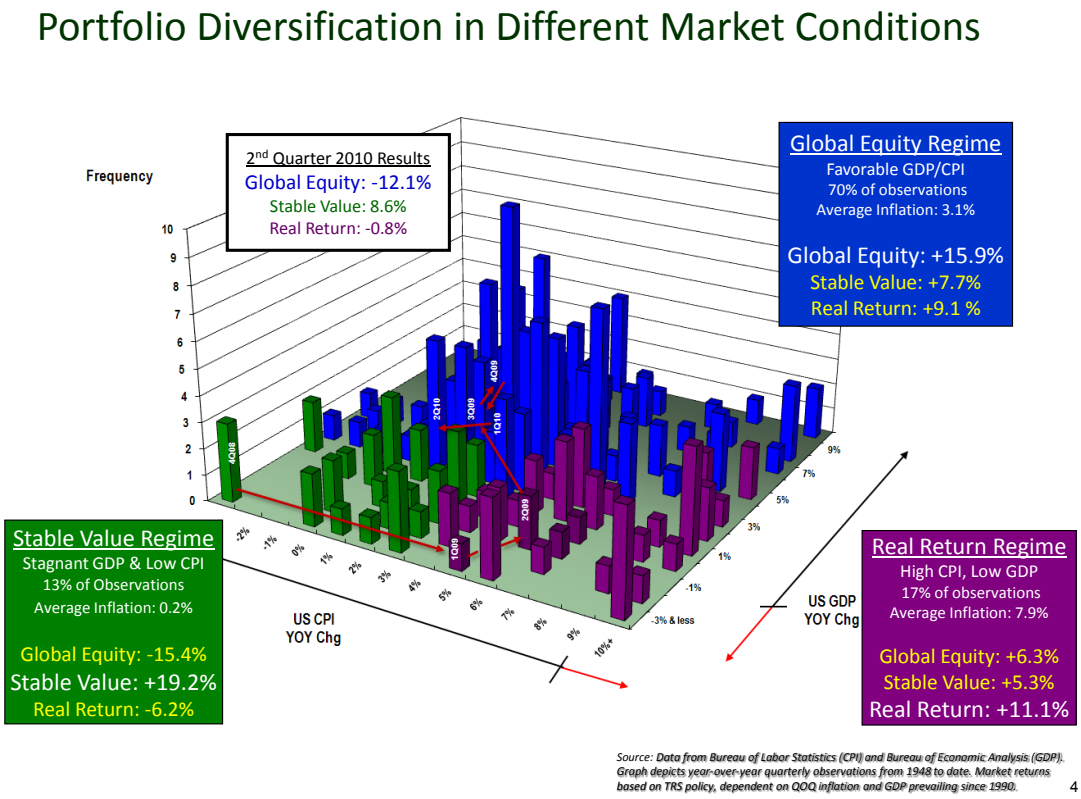
Example 3 – Polar area chart.



Source: https://apandre.files.wordpress.com/2011/03/2101rad.jpg

Polar area charts are a great way to visualize repeating data, such as the seasonal trends that take place over the course a of a year. We might be able to use a visualization like this in our project to display annual trends in consumer spending – i.e. holiday shopping bumps, etc.

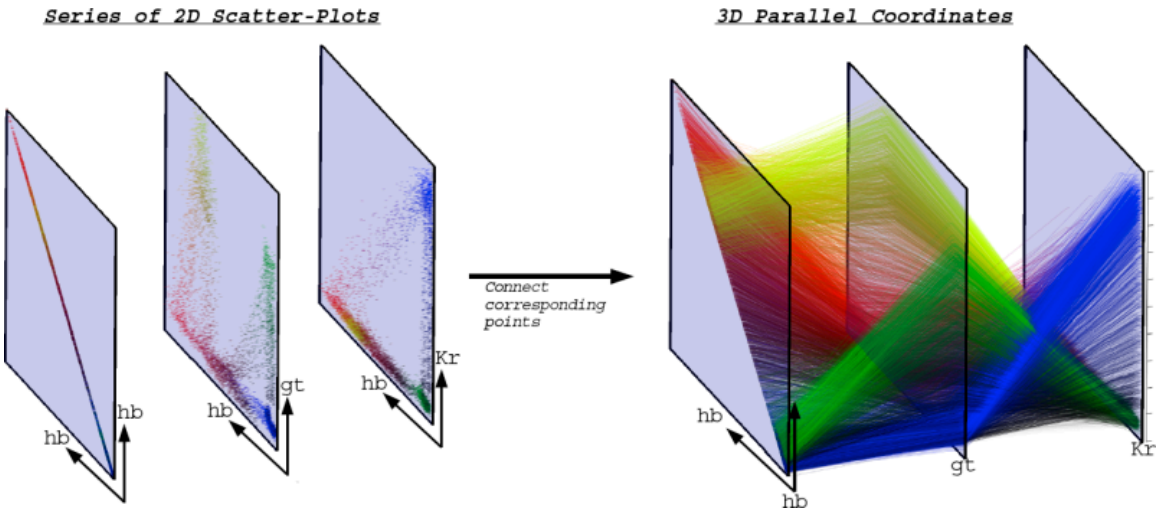
Example 4 – 3D Bar chart



Source: https://fortunedotcom.files.wordpress.com/2012/02/texas_way_v2.pdf

Perhaps the most standard way to try and visualize high-dimensional data is to expand into three (or more) dimensions. While these graphs can look interesting, I don't particularly like these because they are so hard to read. However, if the points or bars nicely group into distinct areas of 3d space (as they do above) it can work.

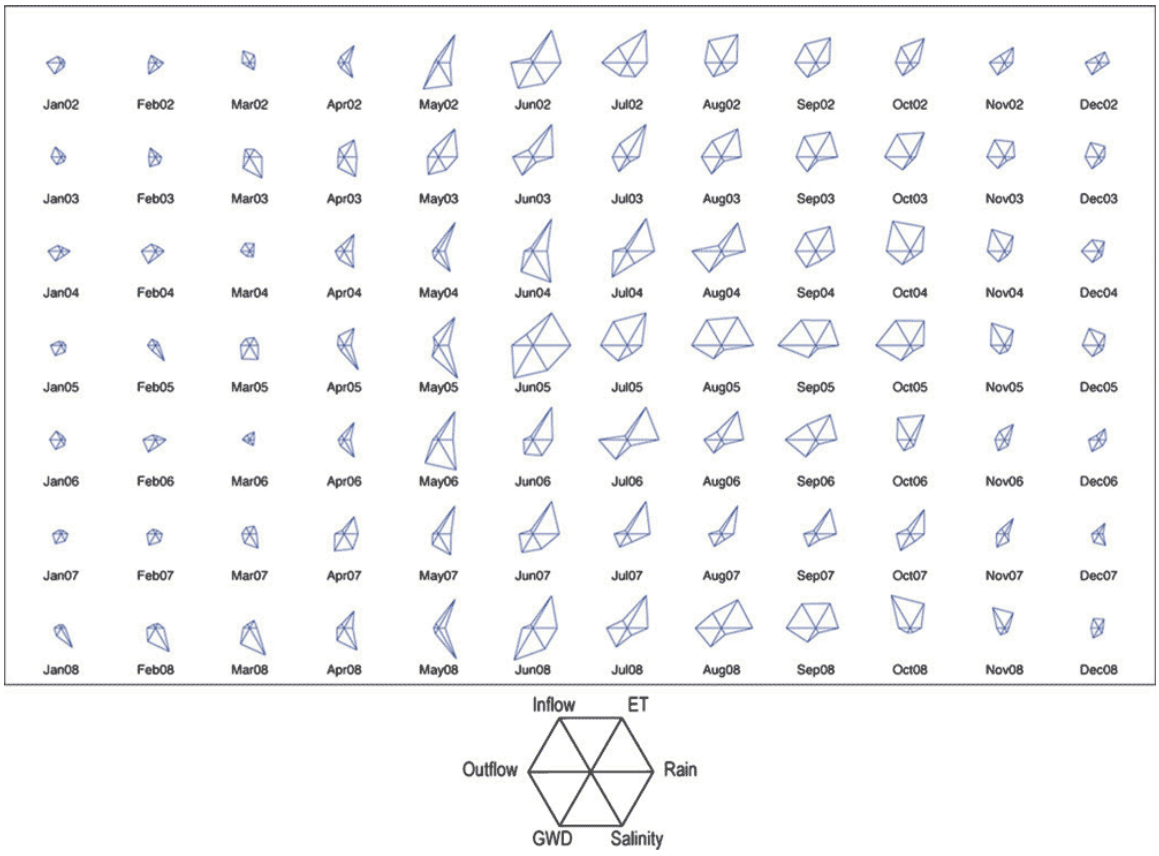
Example 5 – 3d Parallel coordinates



Source: http://bdtnp.lbl.gov/Fly-Net/content/bid/pcx/ParallelCoordinates/3DParallelCoordinates_Illustration.png

Parallel coordinates are another way to display high dimensional data, and show how a set of data points are connected over a third dimension, such as time or ordinal categories. I find these visualizes a bit hard to make sense of, but if effective colors are used (above) patterns are easy to pick out.

Example 6: Glyphplot starplot matrix overtime



Source: https://qph.is.quoracdn.net/main-qimg-288521e3f8a2a8ba0aaabf5d90cf33b2?convert_to_webp=true

I'll admit I was entirely unfamiliar with this graph type before completing this exercise, which is a matrix of star plots. A star plot represents each observation as a "star" whose ith spoke is proportional in length to the ith coordinate of that observation, and as we can see above, six different variables are represented, which is a lot! I'm not sure if we'll be able to build anything like this, but it is interesting.