Intro to Fusion Tables

I am starting to get a grip on using fusion tables, so here is a small intro.

Eli Fisker

Document Shortcuts

Open Fusion Table

Making a graph

How to change what you want to compare

Adding a filter

Choosing what data to look at

Adding more filters

Getting an idea of the data distribution

Clever use of additional filters

Open Fusion Table

Link to the Fusion Table Omei made over the spreadsheet that Meechl send him:

Google Fusion Table

Here is the background forum post:

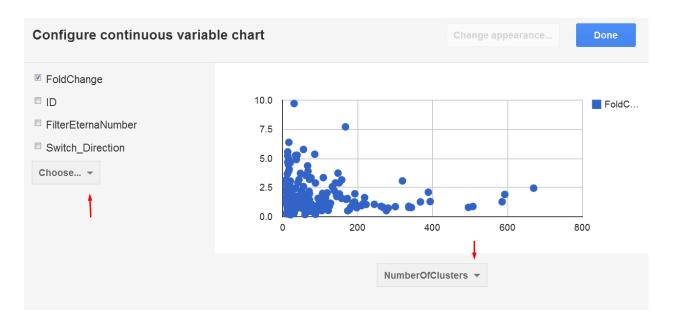
Additional descriptive data

Making a graph

First choose Chart

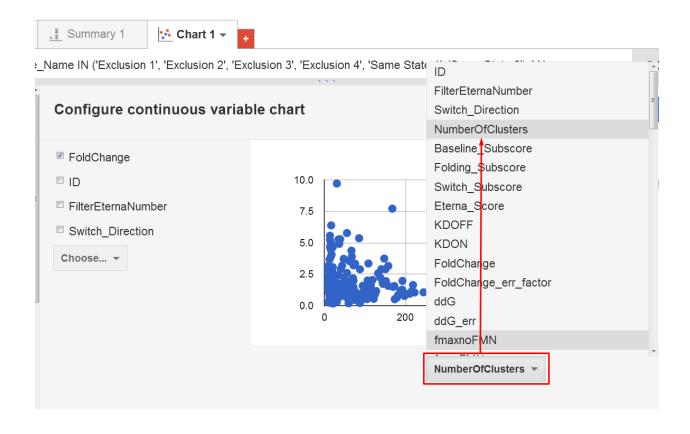


This will give you a graphs field. In this field - **Configure Continuous Variable Chart** - one can choose two of the data columns from the spreadsheet, to run against each other.



I choose to compare fold change against numbers of clusters. Just to see if there were any relation. I honestly had not expected to see any. But the designs that have a huge number of clusters, don't seems to be very good at switching. I have no idea why.

How to change what you want to compare

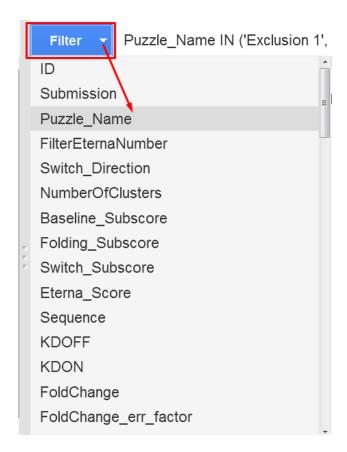


Adding a filter

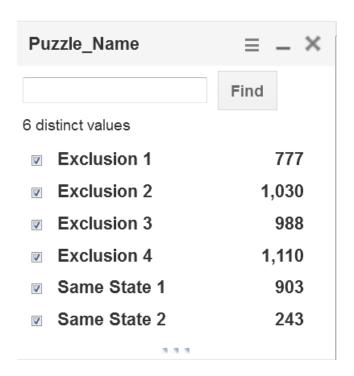
Choosing what data to look at

When playing with a big data sheet you may not want to watch all the data. You can limit what data you are looking at. I want to specify which labs I look at.

To do this you can add a filter:



Puzzle-Name will give you the name of the individual labs. This will give you a list where you can tick of which you wish to view.



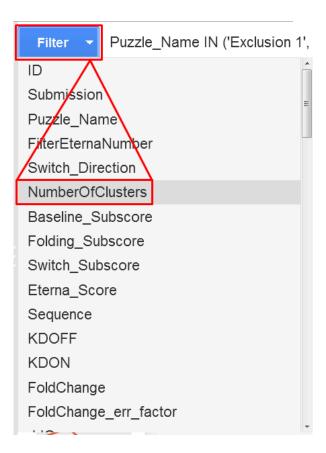
Now this menu after Filter shows what data I have selected.



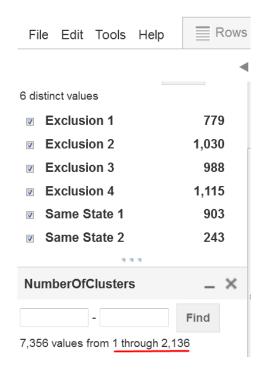
Adding more filters

However I still haven't asked the questions I want to ask the data. I wish to only see the data for designs that have more than 10 clusters, since data that has very few clusters are not very certain.

So I add a filter more. This time I pick **NumbersOfClusters**.



Now a new filters version pops up in the left side of the window.



It tells me the range of clusters there are. From 1 to 2136. And how many of the data points in the spreadsheet that has a cluster number attached to it. (7356)

So now I just limit start number to 10 and choose all the rest of the cluster numbers and click find and the graph will update.

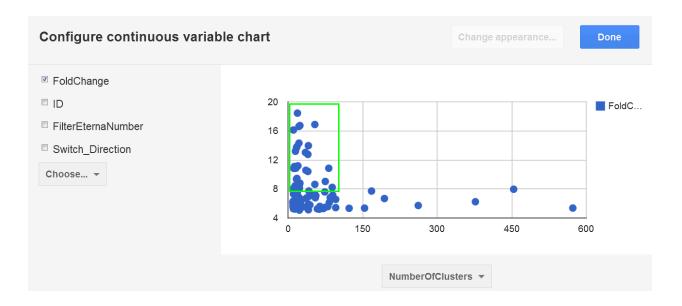


Now I want to limit the data also to the better scoring ones. So I choose to filter with **Eterna_Score** and sets it to count data with folding score 80% and up.



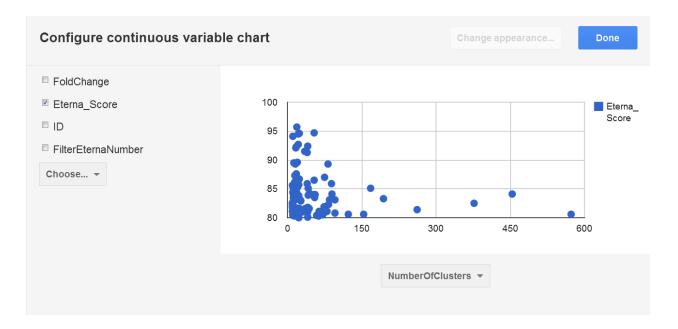
8

Most of the data that has large folding shift, which means a switch is happening, have a cluster size of below 150 clusters.

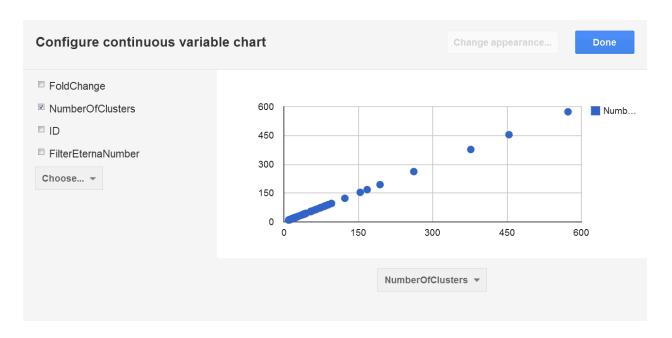


Getting an idea of the data distribution

There seems to be no relation between FoldChange and numbers of clusters. When I change FoldChange to show EternaScore versus FoldChange, I get a similar image. There simply are few designs that have really many clusters.



This becomes clear when I run **NumberOfCluster** against **NumberOfCluster**. Most of the designs have cluster size of below 150. Mystery solved. :)



Clever use of additional filters

For adding of more filters that will help you look at the best data, check out Omei's analysis of Maslow designs. Its practically a guide to Fusion tables.

Variations on the Maslow design, Round 2

Also jandersonlee has made a really fine guide on how to create fusion tables and how to find close to identical designs.

Creating a Nearby Solutions Fusion Table

Written 13 April 2015, last update 17 April