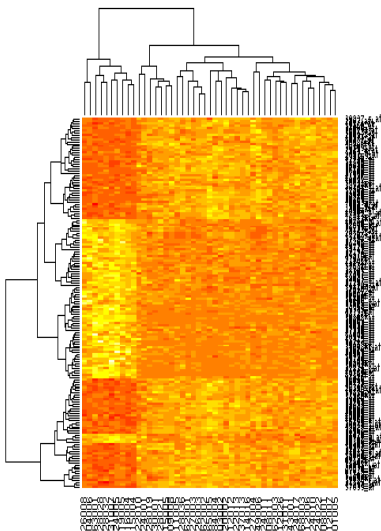


Stat 315c: Transposable Data Heatmaps

Art B. Owen

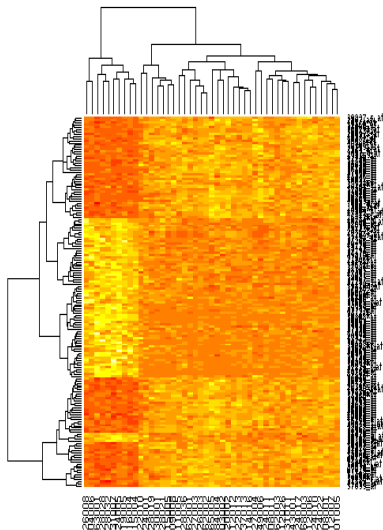
Stanford Statistics

Heatmaps



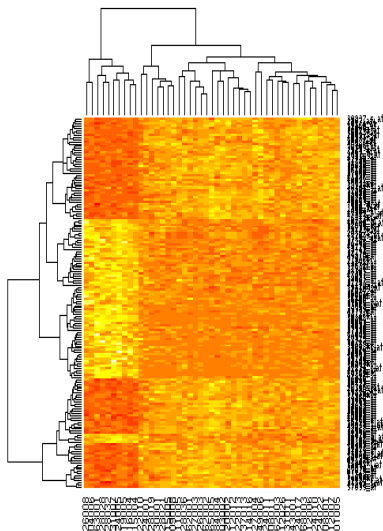
- Plot rows and cols
- Uses color for the value.

Heatmaps



- Plot rows and cols
- Uses color for the value.
- Ordering is crucial. Default is clustering. (more later)

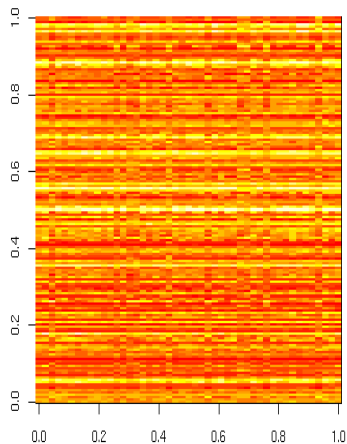
Heatmaps



- Plot rows and cols
- Uses color for the value.
- Ordering is crucial. Default is clustering. (more later)

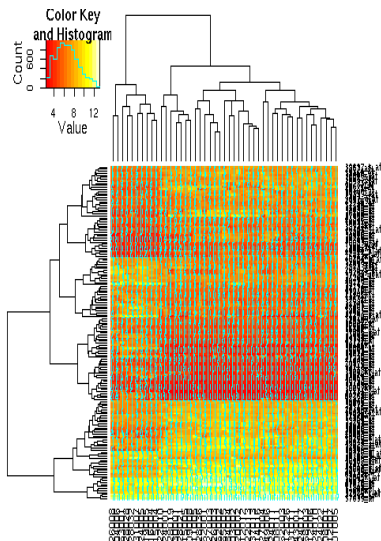
Rows are genes, columns are tissue samples. This is Peter Cock's reconstruction of a heatmap from Gentleman et al (2004). We see clear patterns.

Heatmaps



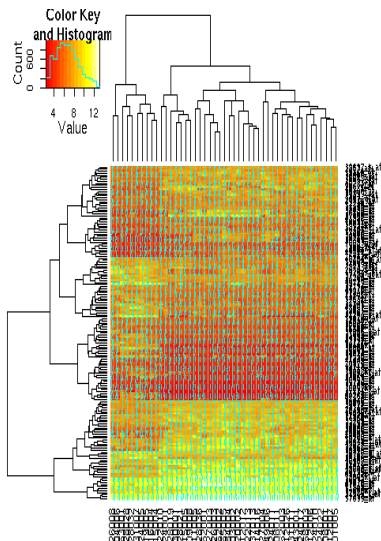
- Here are the data unordered.

Heatmaps



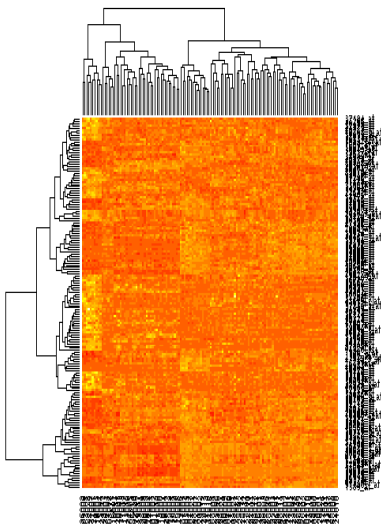
- This is heatmap.2 in R
- Awkward spacings.

Heatmaps



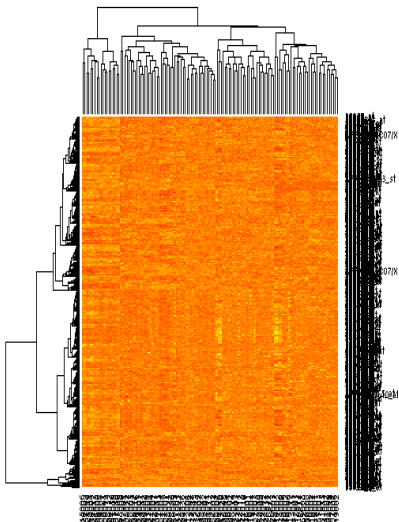
- This is heatmap.2 in R
- Awkward spacings.
- Has color key and histogram.

Heatmaps



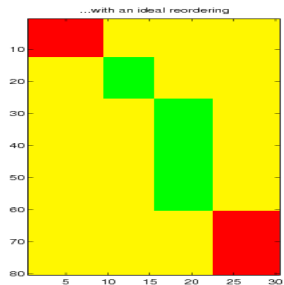
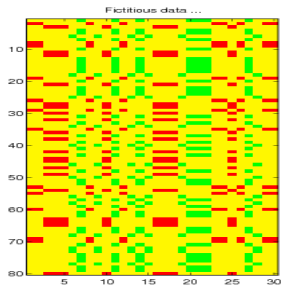
- Now all 128 samples, not just 47 differentially expressed
- Pattern is there, but harder to see

Heatmaps



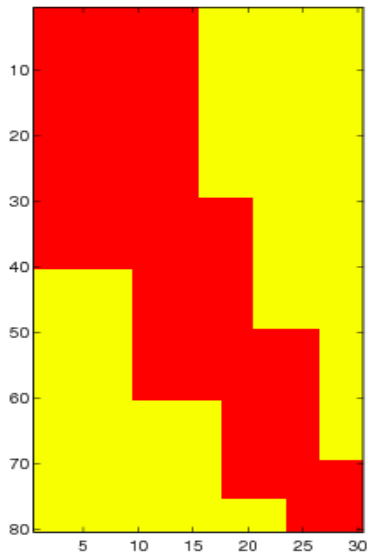
- Now try with 1000 randomly selected genes
- Heatmaps are good for zooming in on structure you've found some other way
- They scale awkwardly. Even genomic data sets are too large for “heatmap(x)” to be an automatically good image. Web sized data would be worse.

Idealized heatmaps

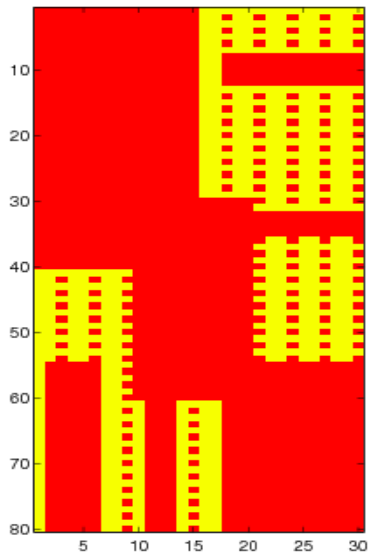


Overlaps happen & Some data can't be ordered

4 Overlapping blocks



4 Unorderable blocks



Show more slides

Wilkinson slides from the web
Slides from AGEMAP work