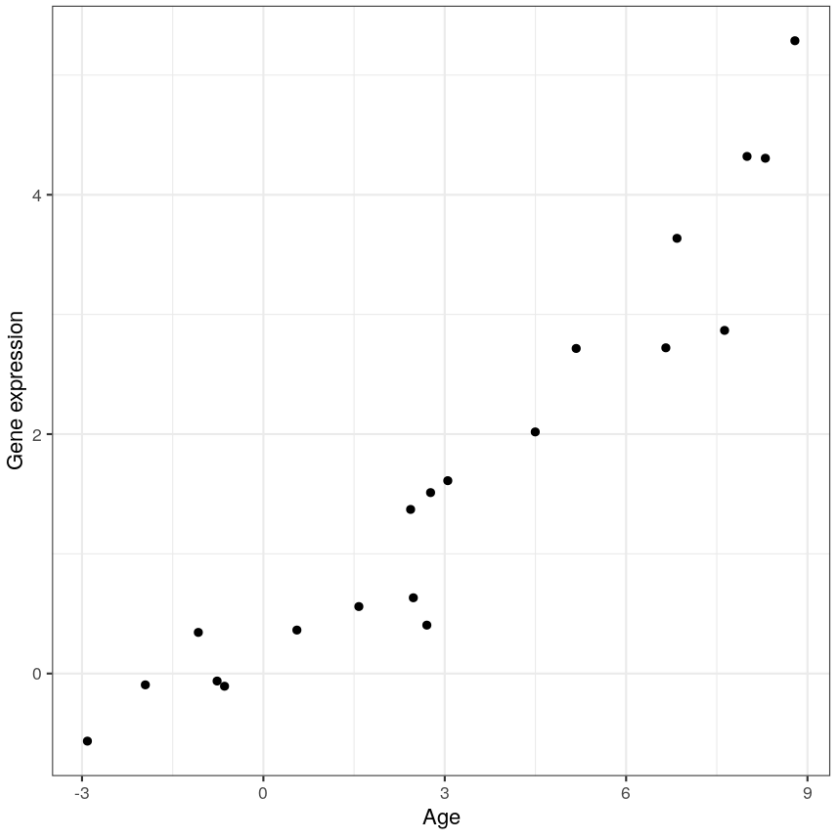
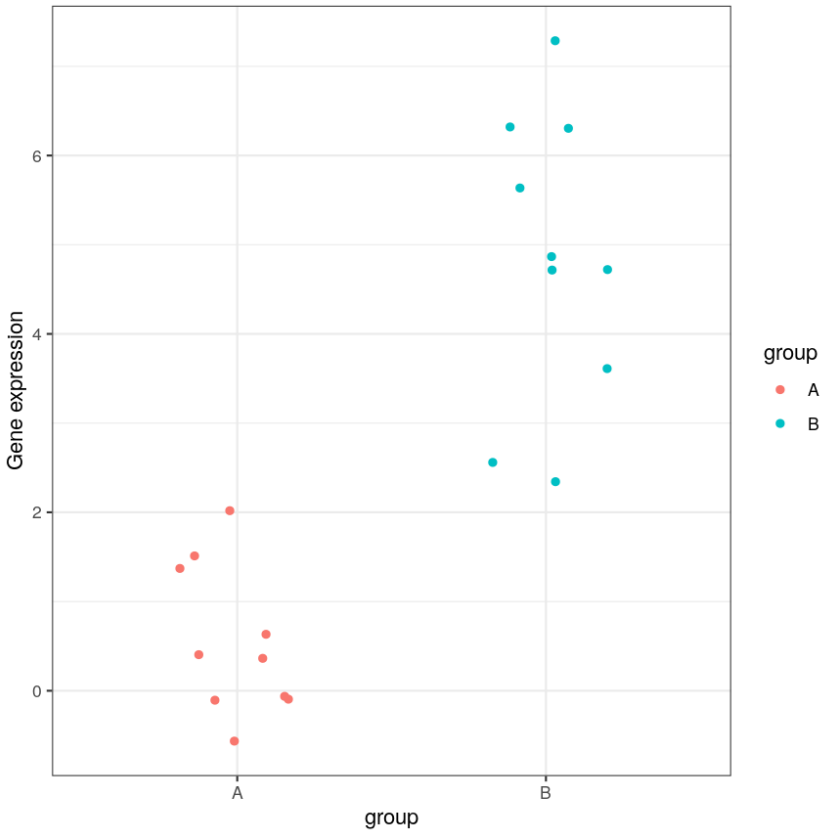
Regression with many features

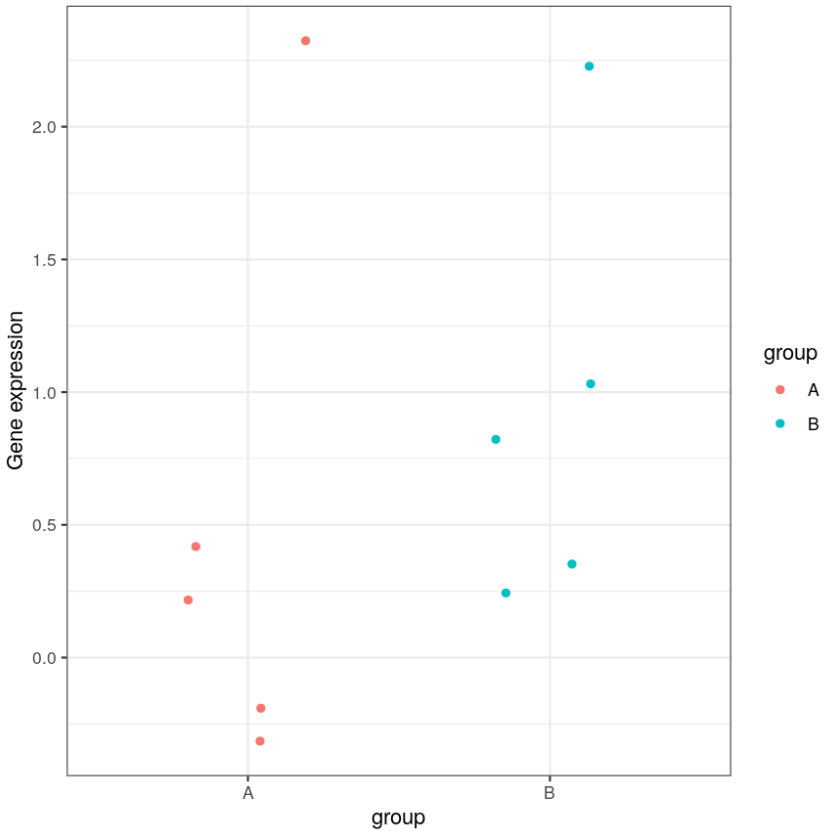
A scatter plot of age and a feature of interest.



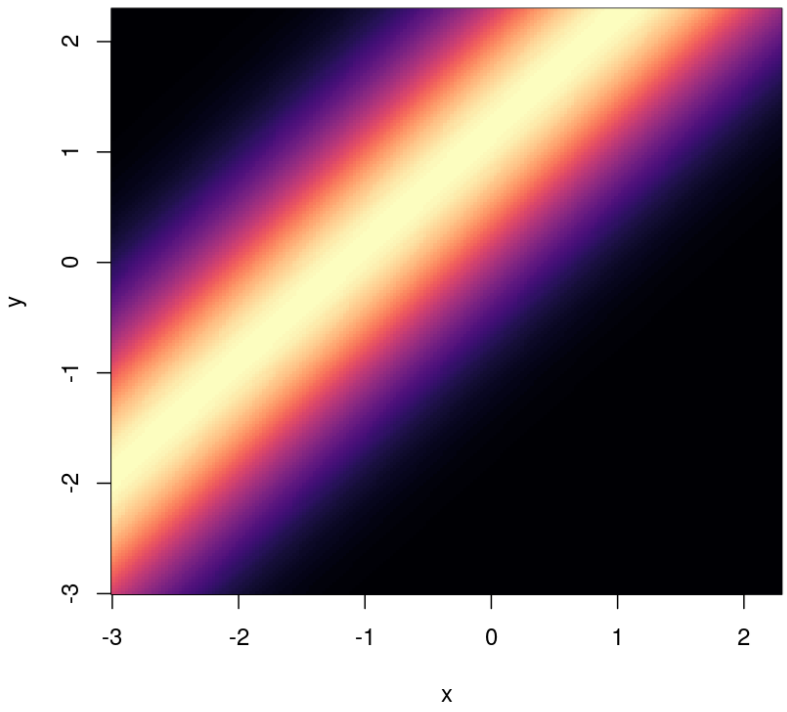
A scatter plot of a grouping and a feature of interest.



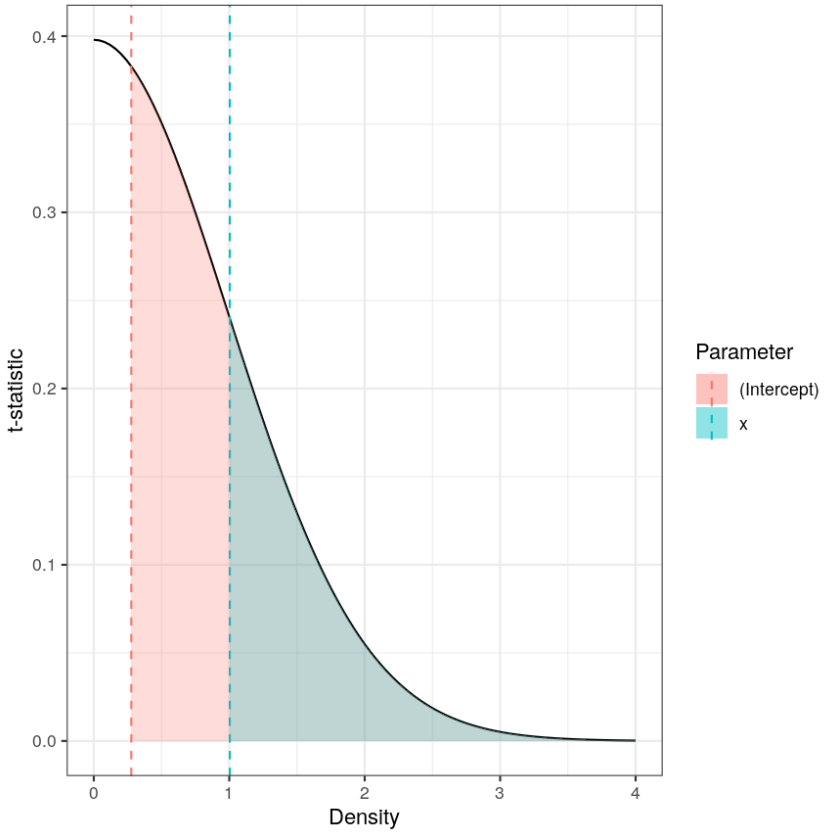
A scatter plot of a grouping and a feature of interest.



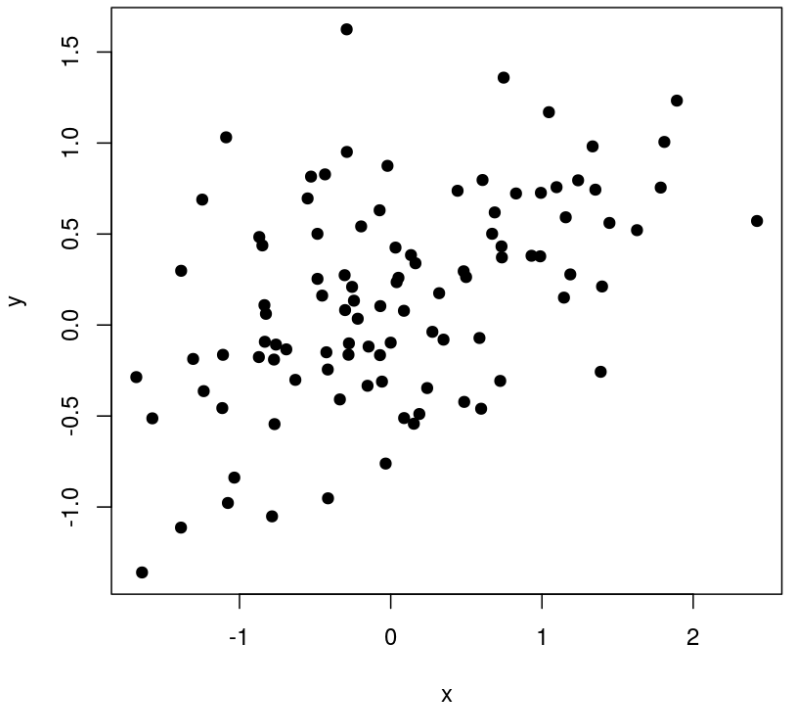
Regression is like a normal distribution with varying mean.



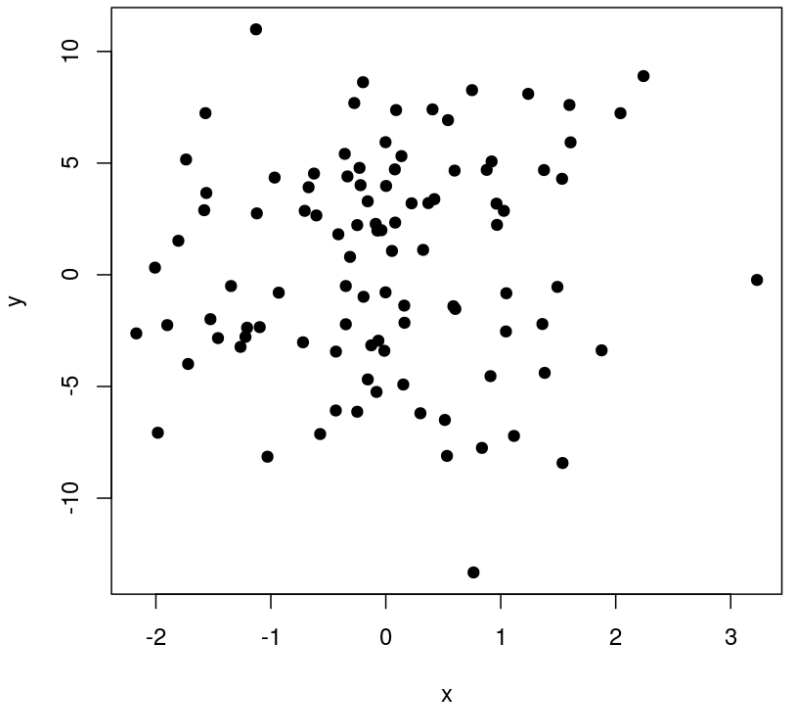
The p-value for a regression coefficient represents how often it'd be observed under the null.



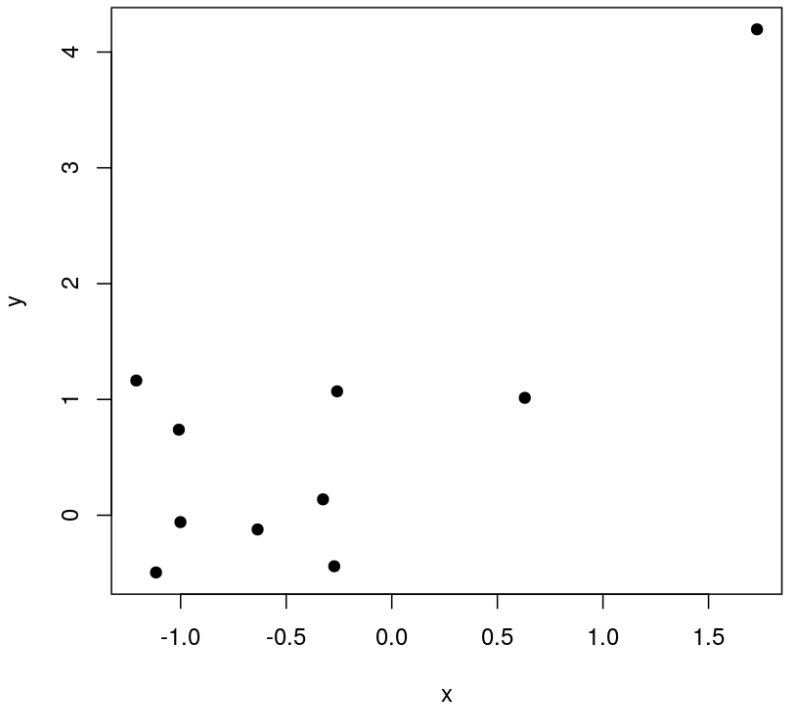
Relationships can be significant with small noise and small effects.



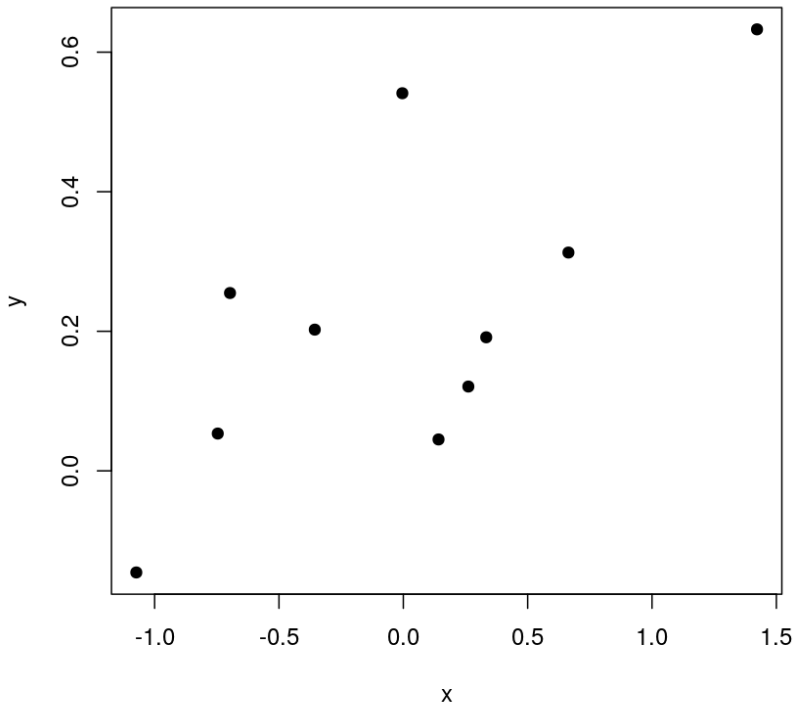
Relationships can be non-significant with large noise and large effects.



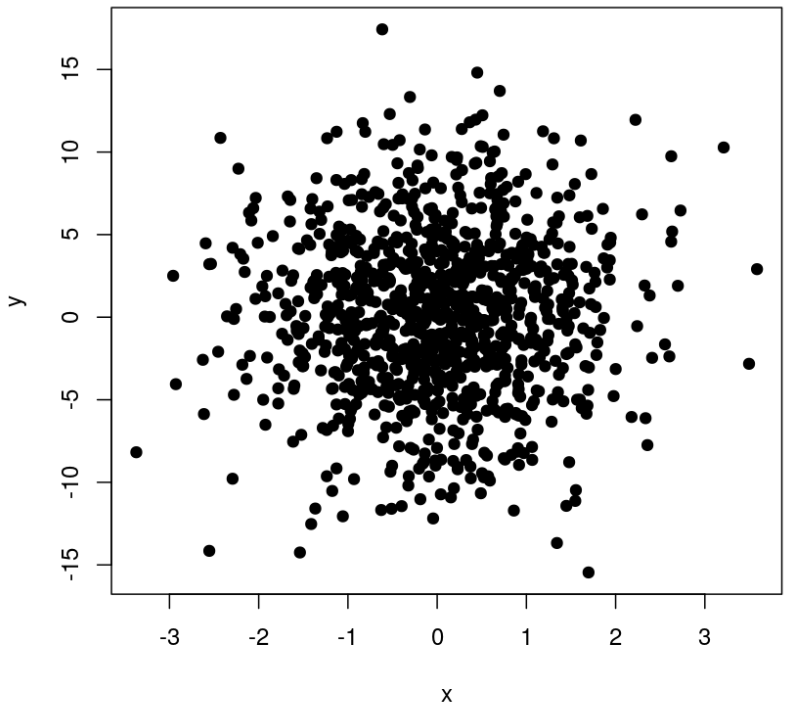
With few points, relationships can be non-significant with large noise and large effects.



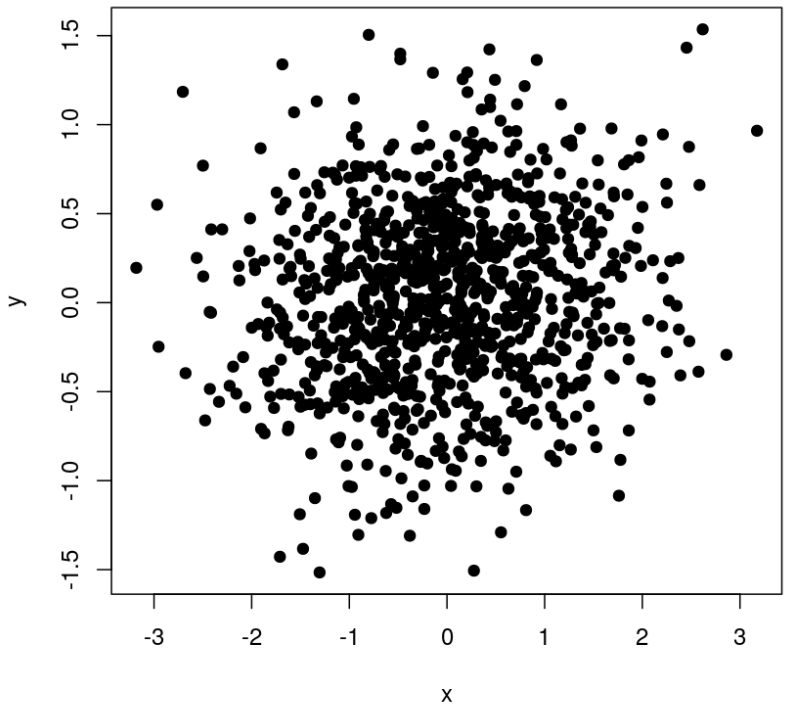
With few points, relationships can be significant with small noise and small effects.



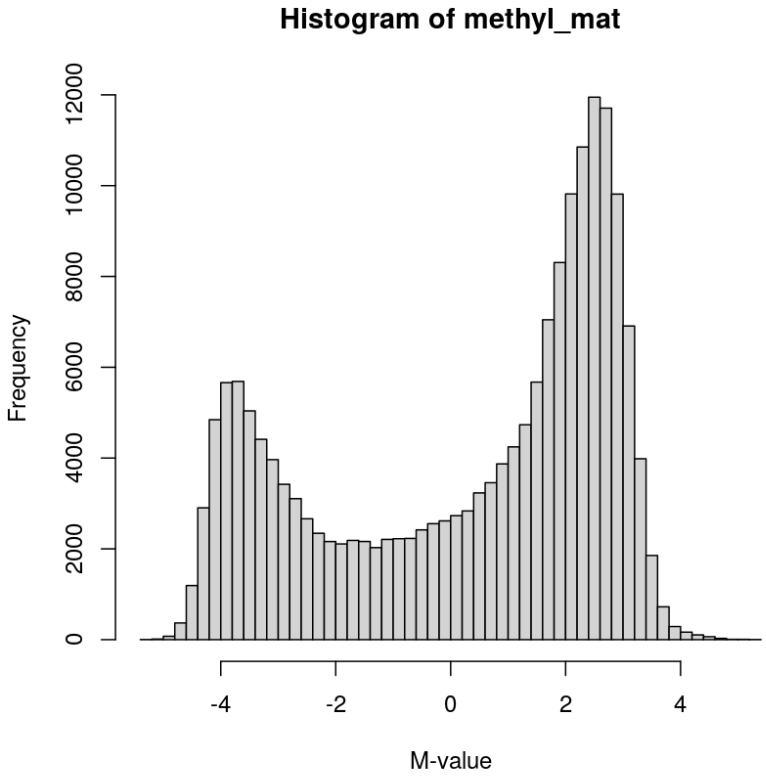
With many points, relationships can be significant with large noise and small effects.



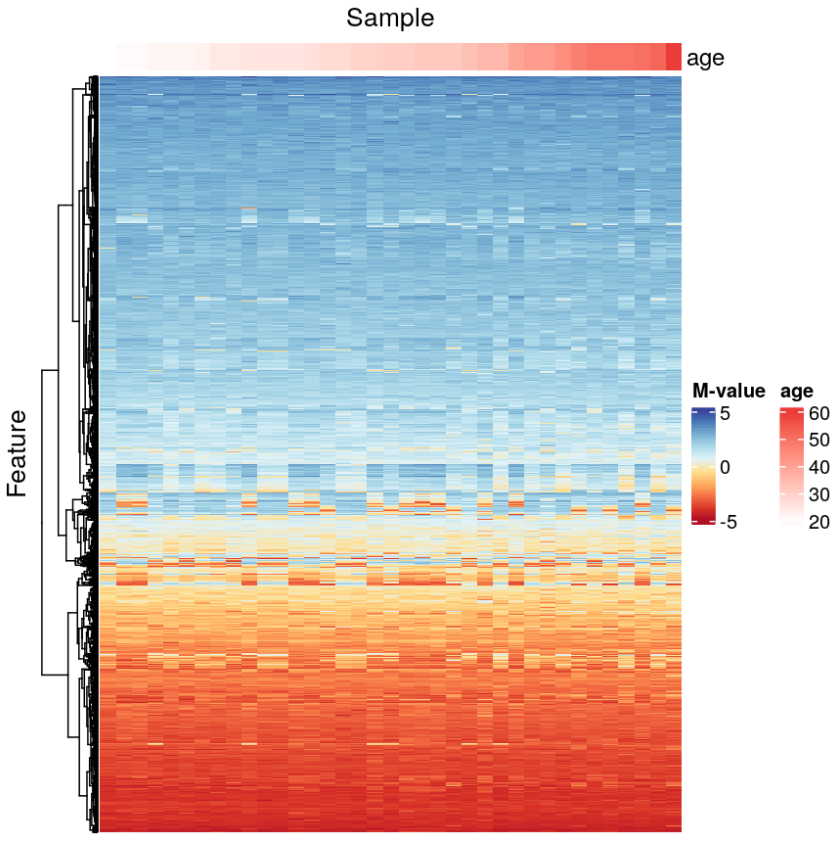
With many points, relationships can be significant with small noise and tiny effects.



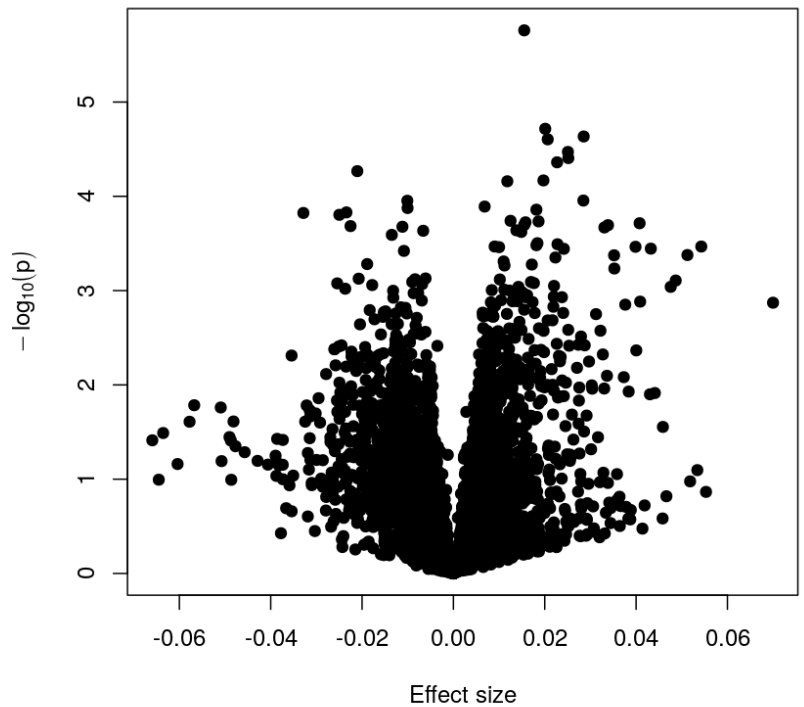
Methylation levels are generally bimodally distributed.



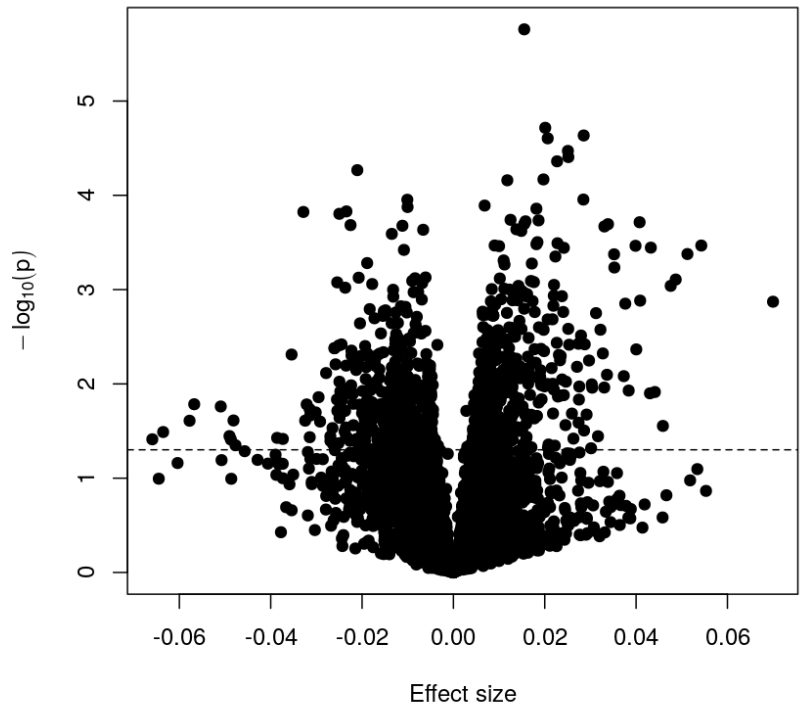
Visualising the data as a heatmap, it's clear that there's too many models to fit 'by hand'.



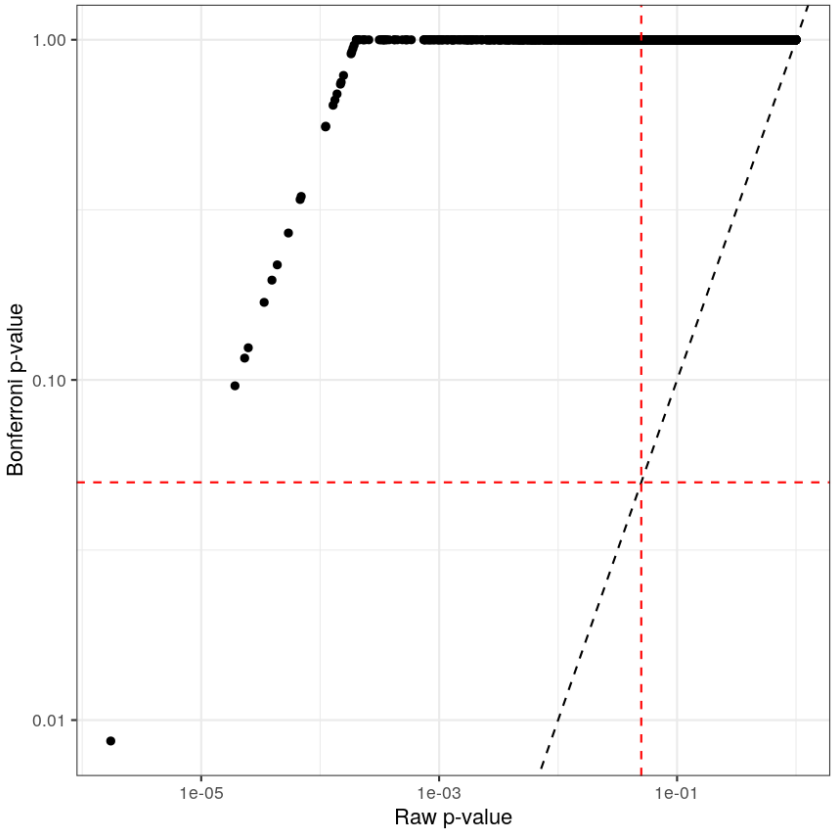
Plotting significance against effect size, it's clear that the two are related (but not 1-1).



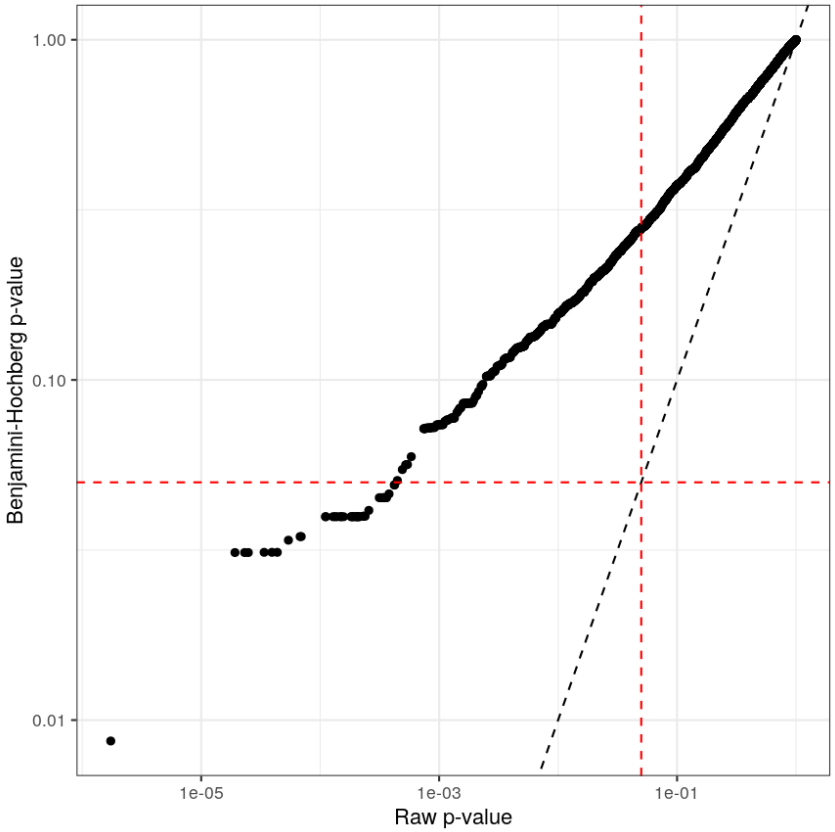
Plotting p-values against effect sizes for a randomised outcome shows we still observe ‘significant’ results.



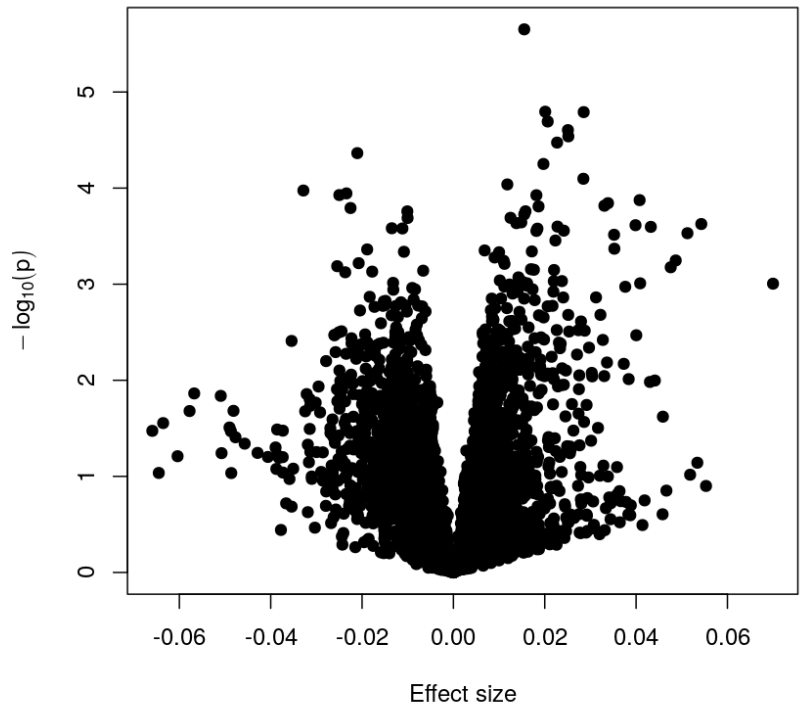
Bonferroni correction often produces very large p-values, especially with low sample sizes.



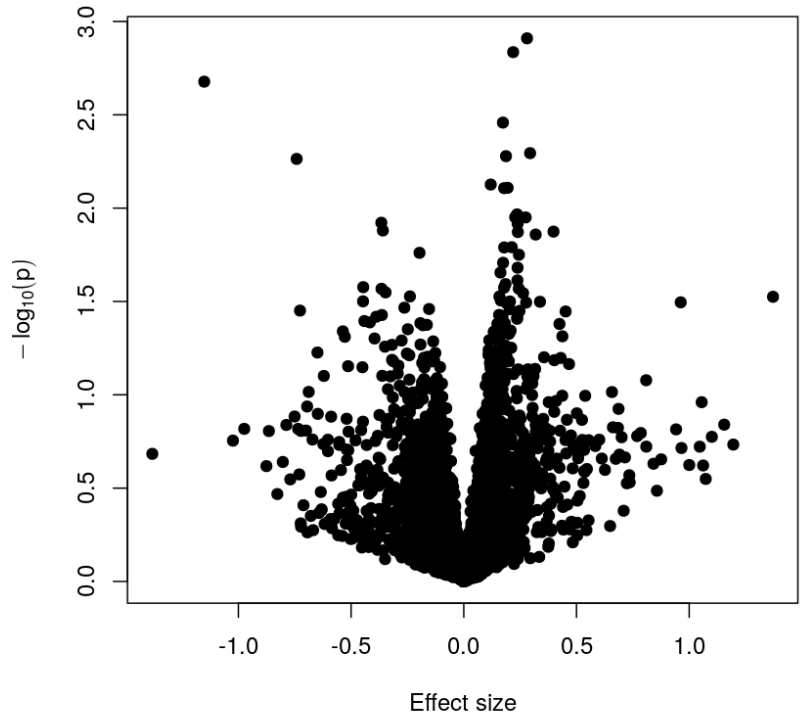
Benjamini-Hochberg correction is less conservative than Bonferroni



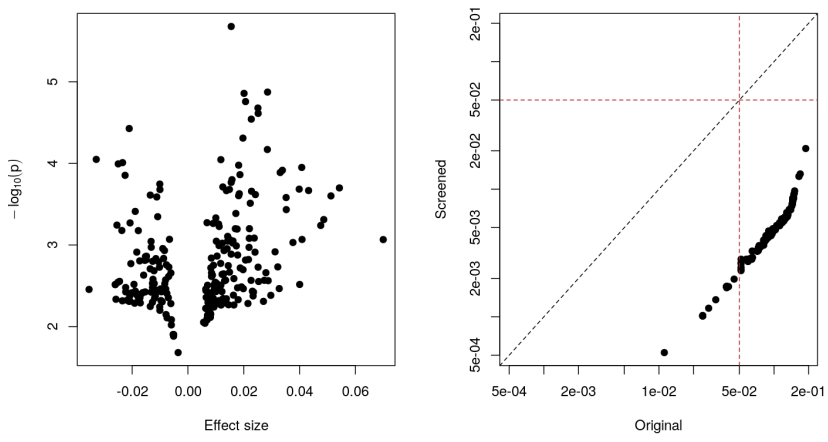
Plotting p-values against effect sizes using limma; the results are similar to a standard linear model.



A plot of significance against effect size for a regression of smoking against methylation.



Caption


$-\log_{10}(p)$
2
3
4
5
-0.02
0.00
0.02
0.04
0.06
Effect size
Screened
5e-04
2e-03
5e-03
2e-02
5e-02
2e-01
5e-04
2e-03
1e-02
5e-02
2e-01
Original