# Peer-graded Assignment: Statistical Analysis

[**Instructions**](https://www.coursera.org/learn/genomic-data-science-project/peer/JlTaG/statistical-analysis)

The next step is to perform a statistical analysis to detect genes that are differentially expressed. To do this, set up the null and alternative hypotheses. Be sure to identify which covariates you are going to adjust for in your analysis. Then use a linear model to test for genes that are differentially expressed between age groups adjusting for an appropriate set of covariates. You can use one of several packages (DESeq, edgeR, limma+voom, or edge) to perform this analysis. Use a correction to identify genes that are differentially expressed after accounting for multiple testing. Make a plot of the fold change (or effect size) for age in each linear model versus the log10 p-value. This is called a [volcano plot](http://www.r-bloggers.com/using-volcano-plots-in-r-to-visualize-microarray-and-rna-seq-results/) which you can use to see if there are any major patterns among the observed differential expression.

## **Review criteria**

Upload a tab-delimited file with three columns - gene name, fold-change estimates, the p-value for that gene, and the adjusted p-value

* Does the file have gene names, fold-change estimates, p-values and adjusted p-values?
* Do the p-values seem appropriate?

Upload a PDF document describing your analysis and your results, including how many genes were differentially expressed at a given error rate, and a volcano plot of the results.

* Does the document have the required components (description of analysis and volcano plot)?
* Does the results seems justified.

**PROMPT 1**

Upload a tab-delimited file with three columns - gene name, fold-change estimates, the p-value for that gene, and the adjusted p-value

Differential Expression Table

[Differential Expression Table](https://coursera-assessments.s3.amazonaws.com/assessments/1632849816891/1cc8cdef-fdab-4235-e9ce-202f46bb1774/differential_Exp.txt)

log2 fold change (MLE): type fetal vs adult Wald test p-value: type fetal vs adult DataFrame with 21164 rows and 6 columns

**PROMPT 2**

Upload a PDF document describing your analysis and your results, including how many genes were differentially expressed at a given error rate, and a volcano plot of the results.

Description of Analysis and Volcano Plot

[Description of Analysis and Volcano Plot](https://coursera-assessments.s3.amazonaws.com/assessments/1632849832567/0ac3907d-9c43-43d9-981a-67a3e2df631c/Assignment%20W8_%20Statistical%20Analysis.pdf)