# Peer-graded Assignment: Get Feature Counts

# [**Instructions**](https://www.coursera.org/learn/genomic-data-science-project/peer/jTAos/get-feature-counts)

Now that you have performed alignment and quality control, the next step is to calculate the abundance of every gene in every sample. This count is an approximation to the expression level for the gene. It is calculated by comparing the aligned reads to pre-specified positions in the genome that correspond to known genes. To calculate these counts you need a file that tells you the positions of the genes. You will need the “gtf file” describing the locations of the genes for the corresponding genome you aligned to during the alignment step. You can then use one of many types of software for obtaining the gene counts (featureCounts, HTseq, etc). The result should be a table that is formatted with one gene per row and one sample per column. The count for that gene in that sample is in the corresponding cell.

**Review criteria**

Upload the gene expression table in a tab-delimited text file

* Is there a tab delimited file with the gene expression counts for each sample?
* Do the counts appear to be appropriate ?(you may have to download the file and make some plots)

**PROMPT1**

Upload the gene expression table in a tab-delimited text file

Expression Data

[Expression Data](https://coursera-assessments.s3.amazonaws.com/assessments/1631397972954/dd53f52e-7983-4207-b761-92ec3ba93906/fetaure_counts.txt)

All samples were aligned to b37 hg19 human reference genome using Hisat2 in command line. Gene counts were obtained using featureCounts from subread package in command line. Fetal samples: SRR1554537, SRR1554538, SRR1554541, SRR1554566, SRR1554567, SRR1554568 Adult samples: SRR1554535, SRR1554536, SRR1554539, SRR1554556, SRR1554561, SRR1554534