

2019 Homework Homework 1

1. Write a script called `download_count.sh` which does the following.
 - Download the data file `ftp://ftp.ncbi.nih.gov/blast/db/FASTA/vector.gz` from NCBI
 - Print out the count of the number of FASTA format sequences in this file (`https://en.wikipedia.org/wiki/FASTA_format`) - each record starts with a “>”
2. Write a script called `summary_exons.sh` which summarizes the total length of exons in the file `data/rice_random_exons.bed`
 - read in the file
 - use a loop structure to read each line
 - add up the values into a variable
3. Write a script called ‘strand_gene_count.sh’ to calculate the number of genes that are on the + and - strand in the file.
 - `https://fungidb.org/common/downloads/Current_Release/ScerevisiaeS288c/gff/data/FungiDB-45_ScerevisiaeS288c.gff`