Write a Perl or Python script that takes as input a bacterial genome sequence in fasta format, searches it for ORFs (minimum size 150bp/50aa) then searches the NCBI swissprot database for proteins similar to each ORF. The output should be a table in comma separated format (CSV) suitable for opening in a spreadsheet program.
The table should have the following headers:

- Start: the start position of the ORF on the sequence
- End: the end position of the ORF on the sequence (note: for ORFs on the reverse strand, the end position index will be smaller than the start position)
- Strand: FORWARD or REVERSE
- BLAST hit: the accession number of the **top hit** matching the translated ORF in the **swissprot** database. If there is no similar sequence with an E-value equal or lower than 1 then this field should just contain "-" (no quotes)
- E-value: the e value for the top BLAST hit, provided it is equal or lower than 1. If it is >1 then the field should contain "-" (no quotes)

The lines in the CSV file should be sorted by start position. An example input and output are provided in Moodle. However, note that because the databases are constantly updated, the example output may not be exactly as your program will generate it.

- Your submission should be a single file named geneannot.pl or geneannot.py
- The script should be invoked at the command line using the command **`./geneannot.pl`** ***`filename`*****`.fasta`** (or **`./geneannot.py`** ***`filename`*****`.fasta`**) where ***filename*****.fasta** is the name of a nucleotide sequence file in fasta format containing a segment of a bacterial genome
- Your script should work with nucleotide sequences of length of up to 20,000 nucleotides (it can work with longer sequences if you want to but you should assume a test sequence of up to 20 kbps).
- Script should quit gracefully if the input sequence is invalid or not in fasta format.
- Script output should be in CSV format and go to standard output so it can be redirected if necessary. For example:
    - **`./geneannot.pl foo.fasta`** will display the CSV output on the screen.
    - **`./geneannot.pl foo.fasta > bar.csv`** will save the output to a file named bar.csv

- Suggested steps for the script:
    - Submit the input sequence to an ORF detection program of your choice (as long as it is installed in binftools)
    - For each ORF identified, its protein translation should be submitted to **BLASTP** at NCBI, searching the **swissprot** database and retrieving **the top hit** for each provided it has an **E-value lower than 1**. You will need to use the BLAST RESTful API at NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastDocs&DOC_TYPE=DeveloperInfo). For best results in the following step you may also want to use one of the alternative BLAST output formats rather than the default one.
    - Parse both the ORF detection and the BLAST outputs to extract the required information into a table in the specified format
    - Sort the ORFs by start position
    - Delete any intermediary files created by the script, keeping only the final CSV output (and any other files that were already there when you started the program).