Using removelowcov.py and removehighcov.py

These scripts are used to filter out records from fasta files based on their coverage values. The first script, removelowcov.py , will remove entries *below* the threshold. The second script, removehighcsv.py , does the opposite, sorting out values that are *above* the specified threshold.

Dependancies

In order to use this script, you need to have an installation of python 3, as well as the sys and biopython libraries. The sys library is installed by default. In order to install biopython, use

```
$ pip install biopython
```

> 🔓 If installation is unsuccessful, try using sudo to install your package as root.

Syntax

From the command line, run the python script, and include your fasta file as the first argument. Specify your threshold as the second argument. You do not need to redirect the output into a file, the script will automatically create the two files it needs.

```
$ python removelowcov.py input.fa threshold
$ python removehighcov.py input.fa threshold
```

> ❗ **Remember, the threshold is inclusive in both scripts.** If you call the removelowcov.py program with a threshold of 3, then entries with coverage 3 will be included in the clean.fa output file. Entries with a threshold of 2.99 will be put into the lowcov.fa output file.

> ⚠️ **Make sure you are using Python 3.** If you are using Python 2, then the program will not work!

How it works