###1. Without using a text editor examine the contents of the file HapMap.fas.txt

**###- How many lines does this file contain?**

Using command: wc -l HapMap.fas.txt, the lines are counted for you

There are 378672 lines in this file

**###- How many characters? (Hint: check out the options of wc)**

Using command: wc -m HapMap.fas.txt the characters are counted to be 15503584

**###- What is the first line of this file? (Hint: read the man page of head)**

Using command: head -1 HapMap.fas.txt you get >TP1\_query\_64 as an output

**###- What are the last 3 lines? (Hint: read the man page of tail)**

Using command: tail -n 3 HapMap.fas.txt you get

CGATTTTTTTTTTTTATCAAAAATGTTATTTTTCTTATTTTGCCGCTGTCTTAAATGGCCCGCT

>TP95869\_hit\_64

CGATTTTTTTTTTTTATCAAAAATGTTATTTTTCTTATTTTGCCGCTGTCTTAAATGGCCCTCT

As an output

**###- How many sequences are in the file? (Hint: use grep)**

Using command: grep ^[CTAG] HapMap.fast.txt | wc -l will provide an output of 189336

The ^ states “find the beginning of a line that starts with [CTAG] in the Hap Map file. The | is a pipe in a command that redirects the previous output of the grep command into further process of the word count “wc -l” command which states how many lines were produced from the previous command.

**###Print all lines that contain a phone number with an extension (the letter x or X followed by four digits).**

Used command: (\d\d\d.\d\d\d.\d\d\d\d\s\w+) and Replace: \1

Then I hit extract to pull it onto another document

**###Print all lines that begin with three digits followed by a blank using the repetition specifier ({})**

Using command: (^\d{3}\s) followed by Replace: \1 and extracted

Saying: start at the beginning of a line with a 3 digits \d followed bye \s

**###Print all lines containing a vowel (a, e, i, o, or u) followed by a single character followed by the same vowel again. Thus, it will find “eve” or “adam” but not “vera”. Hint: \( and \).**

After a lot of trying, the only command I could come up with was: grep \([aeiou]\).\1 grep.txt

However, this command did not bring up anything and I am not sure why

**###Print all lines that do not begin with a capital S.**

Grep ^[^S] grep.txt

Start at the beginning of a line with a line that does not being with S (brackets)

**###Print all lines that contain CA in either uppercase or lowercase.**

command: Grep [Cc][Aa] grep.txt

**###Print all lines that contain an email address (they have an @ in them), preceded by the line number.**

command: Grep -n @ grep.txt

-n signifies that you want the line number in there

**###Print all lines that do not contain the word Sep. (including the period).**

Using command: grep -v Sep\. grep.txt

-v meaning selected lines that do NOT match this pattern

**###Print all lines that contain the word de as a whole word**.

Using command: grep de grep.txt receiving output: Don Henderson

Daren de Heras