1. How many sequences are in lab2_nucleotide.fna? How many sequences are in lab2_protein.faa? Give the command and the number of sequences for each fasta file.

grep -c "^>" lab2_nucleotide.fna
Outputs 4

grep -c "^>" lab2_protein.faa
Outputs 11

2. Give the sed command to replace "lab 2" with your name in lab2file1.txt.

sed 's/lab 2/Laasya/g' lab2file1.txt > Laasyafile1.txt

3. Give the sed command to delete lines containing the word "this" in lab2file1.txt.

sed '/this/d' lab2file1.txt

4. Give the sed command to convert lab2_nucleotide.fna to lowercase.

sed 's/.*/\L&/' lab2_nucleotide.fna

5. Give the sed command to convert lab2file1.txt to uppercase.
sed 's/.*/\U&/' lab2file1.txt

6. Give the grep command that counts all occurrences of substrings "AT" and "GC" in lab2_nucleotide.fna. How many are there?

grep -o "AT" lab2_nucleotide.fna | wc -l   outputs 30
grep -o "GC" lab2_nucleotide.fna | wc -l   outputs 68

7. Give the sed command that deletes lines with the last word being "text" in lab2file1.txt.

sed '/text$/d' lab2file1.txt

8. Give the grep command to count the number of lines in lab2_protein.faa that start with "M". How many are there?

grep -c "^M" lab2_protein.faa   outputs 7

9. Give the grep command to print the lines and line numbers in lab2_protein.faa that DO NOT start with "M".

grep -n '^[^M]' lab2_protein.faa

10. Give the grep command to print out lines in lab2_protein.faa that contain "T__A", where the empty spaces can be any single character.

grep 'T..A' lab2_protein.faa

BONUS 1 (1 pt): Give an awk command to display Sequence 4 through 6 (including headers) in lab2_protein.faa.

awk '/^>/{seq++} seq>=4 && seq<=6' lab2_protein.faa

BONUS 2 (1 pt): Give an awk command to find the longest line in lab2_protein.faa. What is the length?

awk '{if(length > max) max = length} END {print max}' lab2_protein.faa