grep

grep 'string' input.txt

grep -c 'string' file
  count ocurrences of string

grep -f file1 file
  find patterns listed in file1

grep -v 'string' file
  select non-matching lines

grep -l 'string' /home/genomics/*
  list files containing *string*

grep -L 'string' /home/genomics/*
  list files not containing *string*

grep -n 'string' file
  print line numbers of matches

grep -o 'string' file
  only prints *string*, not the line

grep -x 'string' file
  match entire line

grep -e 'regex' file
  use regular expressions

sed

sed 's/string/newstring/[flag]' input.txt

sed 's/stringA/stringB/g' file
  replace all ocurrences in file

sed 's/stringA/stringB/2' file
  replace two ocurrences in file

sed 's/stringA/stringB/p' file
  print only substituted line

sed 's@/home/genomics/@/home/@'
Substitution character can be any
character: / | ^ @ !

sed 's/stringA/>&/' file
  & gets the matched pattern
  prints ">stringA"

sed 's/hello\(stringA\)/\1/' file
\(stringA\) groups stringA
\1   refers to stringA
\2   would refer to the next group

awk

awk condition {action}

awk BEGIN {action} condition {action}
  first action is executed only once at the start.

awk condition {action} END {action}
  last action is executed only once at the end.

Fields: $0: entire row
         $1, $2, ... column 1, column 2

Variables:
NR: Number of records
NF: Number of fields
FS: Field separator
RS: Record separator
FILENAME: name of file

Functions:
print(x)   sum col.:
length(x)   sum+=$2
rand()
sqrt(x)
sub(x,y)

#evomics2023