Essential commands

Command	Function
ls	List the contents of a directory
sort	Sort a file line by line
grep	Search inside files for pattern
awk	Write programs that process files line by line
sed	Transform a file line by line
cut	Cut out selected columns of a file
join	Join two files together based on columns
datamash	Multi-purpose tools to manipulate tab-delimited files
bedtools	Multi-purpose tool to manipulate BED files
samtools	Multi-purpose tool to manipulate SAM/BAM files

`$ man datamash`

NAME
       datamash - command-line calculations

SYNOPSIS
       datamash [OPTION] op [fld] [op fld ...]

DESCRIPTION
       Performs numeric/string operations on input from stdin.

       'op' is the operation to perform. If a primary operation is used, it must be listed first, optionally
       followed by other operations. 'fld' is the input field to use. 'fld' can be a number (1=first
       field), or a field name when using the **-H** or **--header-in** options. Multiple fields can be listed with
       a comma (e.g. 1,6,8). A range of fields can be listed with a dash (e.g. 2-8). Use colons for opera-
       tions which require a pair of fields (e.g. 'pcov 2:6').

   Primary operations:
               groupby, crosstab, transpose, reverse, check

   Line-Filtering operations:
               rmdup

   Per-Line operations:
               base64, debase64, md5, sha1, sha256, sha512, bin, strbin, round, floor, ceil, trunc, frac

   Numeric Grouping operations:
               sum, min, max, absmin, absmax, range

   Textual/Numeric Grouping operations:
               count, first, last, rand, unique, collapse, countunique

   Statistical Grouping operations:
               mean, trimmean, median, q1, q3, iqr, perc, mode, antimode, pstdev, sstdev, pvar, svar, mad,
               madraw, pskew, sskew, pkurt, skurt, dpo, jarque, scov, pcov, spearson, ppearson

   Grouping Options:
       -f, --full
               print entire input line before op results (default: print only the grouped keys)

       -g, --group=X[,Y,Z]
               group via fields X,[Y,Z]; equivalent to primary operation 'groupby'

       --header-in
               first input line is column headers

: