**Overview:**

Clone the nvsl training GitHub repo.

Notice where you are downloading it on the filesystem.

Change to the data directory.

List the files the data directory.

How many files and folders are in the directory?

How many FASTQ file are there?

Make a new directory and mv all the FASTQ files to this directory.

We will return later to the SRA-Tool kit.

**Getting Started**

Download Anaconda and install

Make a vsnp3 environment

`conda create -n vsnp3 -c bioconda -c conda-forge vsnp3`

Activate vsnp3

`which vsnp3\_assembly.py`

echo $PATH  
  
make a new file using `touch`

`touch myfile.txt`

Open this file in a GUI

**Command-line basics**

Unzip FASTQ: `pigz`

Look at the first 10 lines of the FASTQ file: `head`

Count the number of line in the FASTQ file: `grep -c`

Find the error:

pwd; dir=`pwd`; cd~; cd $dir; pwd # find the error

for \*\_R1\*; echo $i; done

for i in \*\_R1\*fastq; count=`grep -c 'GTGTAA' `; echo "$count in $i"; done

**Compute**

Who here has the least computer memory?

If you have a Mac, do you have Intel or M chip?

What are our current download speeds?

What type of computing environments will each of us work with?

**Edit text**

Add a line to your profile that echos what profile has been added. Test by opening a new terminal window.

**Package manager**

Follow training doc.

**FASTQ and FASTA basics**

See the SRA Tool kit installation in Overview.

**vSNP**

Follow training doc.

**kSNP**

Follow training doc.

**Kraken**

Follow training doc.

**AMRFinderPlus**

Follow training doc.