**XSEDE Usage Flowchart**

1. Create an account at XSEDE portal <https://portal.xsede.org/>
2. Give instructor your XSEDE username (Fall 2019: in-class activity sign-up sheet; email)
3. Wait for XSEDE acknowledge after instructor added you to the project (resource allocation). Meanwhile read bridges user guide <https://portal.xsede.org/psc-bridges>
4. Change PSC password (see user guide) and (optional) set DUO MFA two factor authentication.
5. Upload files to XSEDE local directory using scp, sftp, rsync, e.g.

$scp -P 2222 input.txt lanyang@data.bridges.psc.edu://home/lanyang/

(note: XSEDE local directory has 10GB limit, for large files you may directly upload to pylon5)

1. Connect to bridges via ssh or putty and allocate processes (i.e. interact), e.g.

$ssh -p 2222 -l lanyang bridges.psc.edu

$interact -N 4 -t 00:30:00

1. Move files from XSEDE local directory to pylon5 (see note of step 5), e.g.

$scp -P 2222 input.txt lanyang@data.bridges.psc.edu/pylon5/<GroupName>/lanyang/

Note: our group name is **ac5fq2p**, find our group name using the projects command, i.e.

$projects

1. For convenience, you may put data and program in the same directory and also launch the software from that directory, e.g.

$cd /pylong5/ac5fq2p/lanyang

$module load spark

$pyspark #to run Spark program interactively

$spark-submit <program\_file> #to run with script

Notes:

1. Spark program could be run interactively similar to running Pandas program on Python notebook.
2. May view output similar to Pandas head(n) command, i.e. result.take(n), or save to a text file, e.g. result.saveAsTextFile(<filename>), or in scripting direct output to a file such as spark-submit submitme.py > output.txt