**Managing Data on the Supercomputers**

**How much space is my group using?**

The groupquota command will show how much space a group has and how much is in use. The -a flag will show the breakdown for each group user. Use the -g flag to indicate which group is checked, if different from your default home group directory.

To see the usage of all individuals in the sadowsky group:

$ groupquota -a -g sadowsky

**How do I compress my files in my home directory?**

Studies in which data analysis has concluded (*e.g.*, once the paper has been submitted) but may require short-term reanalysis in the future can be compressed as tar balls (.tar.gz file format) to save space.

To create a tar ball:

$ tar -zcvf compressed\_files.tar.gz /path/to/data\_directory/

To unzip a tar ball in the current directory:

$ tar -zxvf compressed\_files.tar.gz ./

**Where can I put data from completed, published projects?**

Once work on a dataset is completely finished, the directory should be zipped into a tar ball, as described above. The tar ball can then be moved to second tier storage for permanent storage and will not count against the group’s data.

Second tier storage is navigated using the s3cmd command.

* Create a directory (called a *bucket*):  
  $ s3cmd mb s3://myfolder
* Add a file to your bucket:  
  $ s3cmd put s3://myfolder/compressed\_files.tar.gz
* Pull a file back down from your bucket:  
  $ s3cmd get s3://myfolder/compressed\_files.tar.gz
* To see the files in your bucket:  
  $ s3cmd ls s3://myfolder