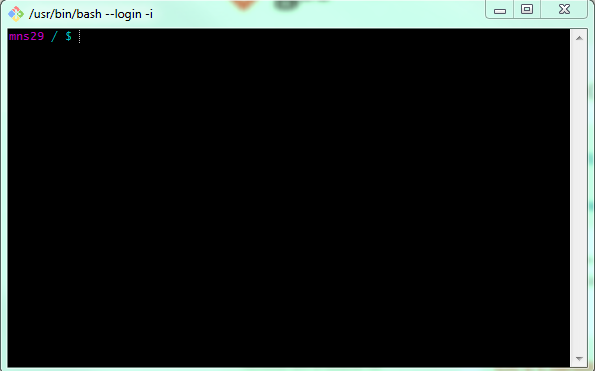
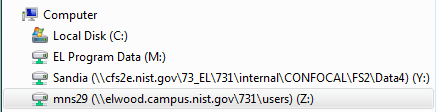
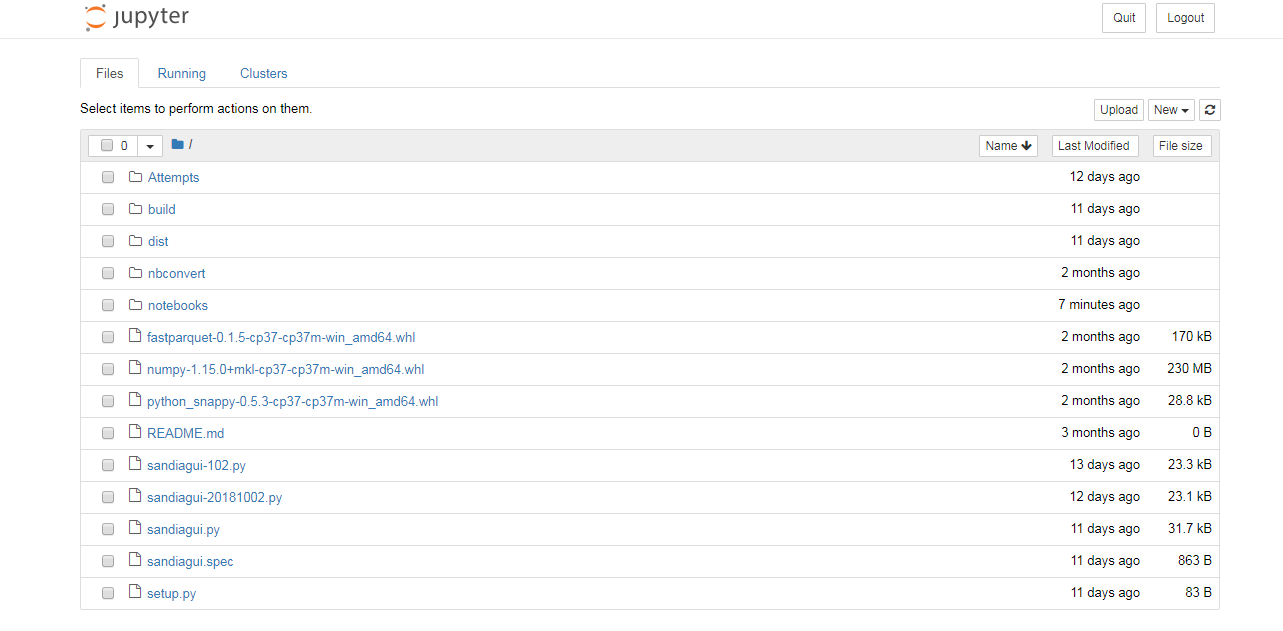
1. Download **Git Bash**
   1. **Link:** <https://git-scm.com/downloads>
2. Download Python (3.7)
   1. **Link:** <https://www.python.org/downloads/>
3. **Open** Git Bash
   1. 
   2. **Type** these commands
      1. **cd** ~ or **cd** z
      2. **mkdir** *foldername*  (ONLY THE FIRST TIME)
      3. **cd** *foldername*
      4. **ls** (Displays all contents in the folder)
         1. **cd**: goes into specified folder
         2. **mkdir**: makes a new directory
         3. **ls**: views all files/folders in the directory
4. **Go** to your folder directory and **Click** the one with your username. The new folder should be listed there.



1. Put the given **Piplock** and **Piplock.lock** file and **parquet\_processing.ipynb** file in that folder from the folder directory.
2. Go to the **folder** in Git Bash.
   1. **cd** *foldername*
3. **Type** these commands:
   1. pip install pipenv
   2. pipenv update
      1. Wait for the dependencies to be locked, it takes around 5 – 10 minutes.
      2. If there are errors, please contact Manika Sachdeva ([manika.sachdeva@nist.gov](mailto:manika.sachdeva@nist.gov))
   3. pipenv run jupyter nbextension enable --py widgetsnbextension
   4. pipenv run jupyter notebook
      1. A jupyter notebook will open on your browser
4. **Click** on the parquet\_processing.ipynb file
   1. **Wait** 5 minutes the Kernel is ready
      1. 
   2. **Comment out** the DATA# variable by using “#” before the line and it will be grayed out.
      1. 
   3. **Type** the new variable
      1. DATA# = SANDIA / “Name of Sandia Data Folder” (remember the quotes)
   4. **Change the sorted function variable to the DATA#**.
      1. 
      2. **For Dry and Humid!**
   5. In the **DRY\_PARQUET** and **HUMID\_PARQUET** variables change the file names to the dates of the current data set.
      1. 
   6. **Press the Run button** for each block of code
      1. 
      2. Make sure that **the first three blocks** have the numbers
      3. 
      4. Once you get to **blocks 4 and 5** it should have **[\*]** because that means it is still running.
      5. It takes **10 minutes to run blocks 4 and 5**, it will display the time once it is generated
      6. 
   7. **Check the Parquet Files Folder** for the new generated parquet files
      1. **Link:** [\\cfs2e.nist.gov\73\_EL\731\internal\CONFOCAL\FS2\Data4\Sandia\Parquet Files](file:///\\cfs2e.nist.gov\73_EL\731\internal\CONFOCAL\FS2\Data4\Sandia\Parquet%20Files)
5. You can close the Git Bash and the Jupyter Notebook and repeat the process for new data sets.