

**MAST 467/667: Introduction to Polar Oceanography (Fall 2021)**  
**Andreas Muenchow (muenchow@udel.edu)**

**Sept.-22, 2021**

**Workshop/Homework-2:** Automating work flow via Command Line Interface

**Data:** Ocean Melts Greenland (OMG) at <https://omg.jpl.nasa.gov/portal/>

**Introduction.** For our second workshop we will continue to probe Arctic Ocean Water Profiling around Greenland as done by a large NASA program call OMG.

Together we will explore shell processing for automated data downloads and extraction of location data using the Windows (or macos) command line with tools like gawk (or awk)

**Goal.** Acquire command line skills to automate data download and processing.

**Expectations.** I expect us all to work within a collaborative environment. Connect with other students in this class to identify which tasks you can and which tasks you cannot do. I do NOT expect that all students can solve all aspects of this assignment. I expect that every person makes an honest attempt and is able to identify and verbalize aspects the she/he/they are unable to conduct without help of the instructor. Substantial help from the instructor is an expectation. Learn by doing.

**Assignment.** Determine the number of data files within your assigned sector of Greenland.

1. Download manually the executable for gawk (Windows users only) from <https://sourceforge.net/projects/gnu-awk-for-windows/> and place it into a directory that includes your computer's executables. On my Windows machine I put it in c:\Users\Andreas\AppData\Local\Microsoft\WindowsApps that is also part of the "path" command.

[move files]

2. Open a command line (Windows: search for "cmd" app) or terminal (Mac: search for "terminal" app); I posted a few useful web-sites at <http://muenchow.cms.udel.edu/classes/Arctic/links.html>

Try typing a few commands such as "echo Hello World" or "dir" or "cd"  
Make a directory by typing "mkdir MAST467"  
Go into this directory by typing "cd MAST467"  
Check the files in this directory by typing "dir"  
Check, if you have an application curl by typing "curl -help |more"

[files and directories]

3. Open a new file called "First.cmd", add a command like "echo Hello World", save it, and execute it from the command line by typing "First.cmd"

*[edit text, create/execute files/scripts]*

4. Make a data directory under the MAST467 directory by typing "mkdir Data" and make a code directory under MAST467 by typing "mkdir Code" and go into the Code directory by typing "cd Code" where ideally your future codes.

*[managing files and work space]*

5. In your "Code directory" create a new file that will automatically download a file from NASA's web-site and search this file for its latitude and longitude. Your new ".cmd" file will need to include the following 4 lines to download data [exclude the "LineX>" string that indicates just what I mean by 4 separate lines].

```
Line1>set input=OMG_20180822145112_CH12_294.edf
Line2>set ID=d09ca830-e9e9-11e9-bd33-c3a895d253bd
Line3>curl -X GET -o %input% https://omg.jpl.nasa.gov/cas-product-omg/data?refIndex=0&productID=%ID%&uid
Line4>EXIT /B 0
```

Can you explain what each of these lines does?

What is the difference between "input" and "%input%" in the script?

*[shell scripting with variables]*

6. Add another line to the script that will extract the latitude and longitude from this file. You will need awk/gawk for this that you can invoke with gawk "\$1 ~ \"Lat\" {print}" %input%

*[executing gawk command]*

7. Add a line to the above script to move the downloaded file with a different and shorter name into the Data directory.

*[automatic workflow and data management]*

8. How could you generalize, modify, or replace this script to perhaps download ALL files that reside in OMG data folders or directories on the web.

[Hint: Have a look at the hyperlinked cURL files at

[https://omg.jpl.nasa.gov/portal/browse/products/urn:omg:OMG\\_Ocean\\_AXCTD\\_L1](https://omg.jpl.nasa.gov/portal/browse/products/urn:omg:OMG_Ocean_AXCTD_L1)