Module 1: Shell Workshop

# Goal

Going back to basics. The computer “Terminal” has been the access point for serious computer scientists for decades. But why do you need to know about this?

Reason #1: We are going to be working this semester with tools in the Hadoop ecosystem (including Pig, Hive, and Spark). These tools can be integrated with SAS (as well as R and Python), but fundamentally work at a more basic level: the Unix shell.

Reason #2: We will be working on the Boston University cluster in this class, a remote resource that will help us scale analyses on large datasets. The best way to work on this resource is through a terminal and the language we want to speak is the Unix shell.

Today’s workshop will cover the basics of interacting with your system via the Unix shell.

# Instructions

Set up and log into your BU cluster account: <https://scc-ondemand2.bu.edu/>

Follow along with the in class live-coding based on this guide: <http://swcarpentry.github.io/shell-novice/>

**Turn In: DUE January 31, 2021**

For each of the commands we use (see the tutorial webpages if you can’t remember) record the following

1. Name of the command (e.g., cd, cp, ls, cat)
2. A plain language explanation of what the command does in your own words.
3. An example of code using this command.