This problem set will provide an opportunity for you to practice accessing computing resources on the OSCER server here at OU.

As with the previous proble set, you will submit this problem set by pushing the document to *your* (private) fork of the class repository. You will put this and all other problem sets in the path `/DScourseS22/ProblemSets/PS2/` and name the file `PS2_LastName.pdf`.

1. Make sure you are able to access OSCER.

2. If you have not already, clone your forked GitHub repository to your home directory on OSCER by following these steps:

   1. Make sure you are in your home directory by typing `cd ~`.

   2. Type `git clone https://github.com/[your-github-username]/DScourseS22.git`

   3. Check that everything works by typing `ls` and hitting enter. You should see a directory called "DScourseS22" and if you change to that directory, you should see an identical directory and file structure to what is on your GitHub private fork of the course repo.

3. Go to www.overleaf.com and create another .tex document, this time naming it `PS2_LastName.tex`. In it, write down a list (in a LaTeX `itemize` envrionment) that outlines the main tools of a data scientist (as discussed in class). Compile the PDF.

4. Compile your .tex file, download the PDF and .tex file, and transfer it to your cloned repository on OSCER using your SFTP client of choice, or via the `scp` command at the Terminal (in RStudio or elsewhere). If you're using FileZilla, you can set up your connection by clicking File → Site Manager → and then clicking on "New Site" and entering in "OSCER" as the name (instead of "New Site"), "dtn2.oscer.ou.edu" as the Host, "SFTP" as the protocol, and your username as the user. It's not a good idea to store passwords on your machine, so leave that field blank and make sure that under "logon type" you click "ask for password." If you are using something else besides FileZilla, there should be a similar "site manager" type of interface for you to enter in your credentials. Consult the help resources for your specific software.

   Note: you may also do the following. From the terminal in RStudio, navigate to a temporary directory on your computer (e.g. Downloads folder). Then type (all on one line) `scp -p PS2_LastName.tex`

   `username@schooner.oscer.ou.edu:/home/username/DScourseS22/ProblemSets/PS2/` (and the same for your `.pdf` file). Do **not** put these files in your fork on your personal laptop; otherwise git will detect a merge conflict and that will be a painful process to resolve.

5. Make sure that your .tex and .pdf files have the correct naming convention (see top of this problem set for directions) and are located in the correct directory. If the directory does not exist, create it using the `mkdir` command.

6. Update your OSCER git repository (in your OSCER home directory) by using the following commands:

   - `cd ~`
   - `cd DScourseS22/ProblemSets/PS2` (note: you may need to create the `PS2` directory if it doesn't exist)
   - `git add PS2_LastName.*`
   - `git commit -m 'Completed PS2'`
   - `git push origin master`

   Once you have done this, issue a `git pull` from the location of your other local git repository (e.g. on RStudio on your personal computer). Verify that the PS2 directory appears in the appropriate place in your other local repository.

7. Make sure that you have R version 4.0.2 installed on OSCER. To do this, type `nano ~/.bash_profile`. Type `module load R/4.0.2-foss-2020a` somewhere in the middle of that file. Then push Control+X and then type "yes" when it asks if you want to modify the file.

8. Now open R (on OSCER), verify that it is running the version mentioned above, and install the `xml2` package by typing `install.packages('xml2')` at the command prompt. (**NOTE: The package is "X-M-L 2," NOT "X-M 12"**) Answer 'yes' to install it in your home directory. Enter the number of any of the server mirrors that come up (I usually choose Texas). Building the package may take a couple of minutes. Once that has finished, repeat the process, but substitute `tidyverse` for `xml2`. Building `tidyverse` may take as much as 10-15 minutes. You don't need to watch it build, but be sure not to close your SSH terminal window or otherwise disconnect your session. When you are done with the build, exit R and log off of OSCER. You can always exit a program (or logout) on OSCER by typing Control+D.

9. Update your fork of the class repository on both your personal laptop as well as on OSCER, since I'm been continuously updating things (and as part of PS1, everyone in the class submitted pull requests). The command to do this is `git pull upstream master`. A step-by-step help for how to do this is located here. More simply, you may also just go to your fork on GitHub and click the button that says "Fetch upstream."