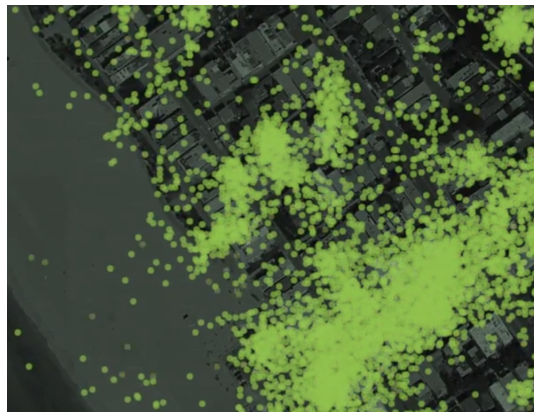


**CMPSC 301
Data Analytics
Fall 2020**

Lab 1 Assignment:
Researching and Reporting on Database Tools
Submit deliverables through your assignment GitHub repository.
Place report document writing/ directory

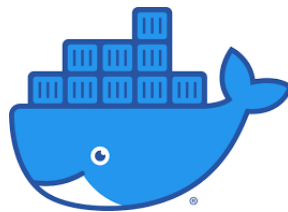


Objectives

To learn how to navigate the directories using a command line interface. To establish and configure the Docker container system and GitHub repositories that will be used in this course throughout the semester; learn basic commands for downloading course materials from the instructor and for submitting material to the instructor using `Git`. To find two (2) articles describing what could be a potentially unethical usage of data.

Part 1: Tools

In this part, there is no deliverable to submit, however it is necessary that you have these tools installed and working for the next steps of the class.



Docker

In our course, we will be using command-line tools that may be facilitated by in Docker container environment. If you have not already installed Docker, please visit the Docker *Getting Started* website (link: <https://www.docker.com/get-started>) to setup Docker on your machine. There is no deliverable for this installation in this lab, but it is assumed during class that you have correctly installed Docker on your machine.

For more information about installing and using Docker in our department, please watch Dr. Jumadinova's video, *Running and Testing Programs with Docker and GatorGrader* at the link; <https://www.youtube.com/watch?v=iceAgNEORCA>.



Git

Practicing software developers normally use a version control system to manage most of the artifacts produced during the phases of the software development life cycle. In this course, we will always use the **Git** distributed version control system to manage the files associated with our class, laboratory and practical sessions. In particular, we will securely communicate with the GitHub servers that will host all of our projects. In this laboratory assignment, we will perform all of the steps to configure the accounts on the departmental servers and the GitHub service. As you will be required to use **Git** in the remaining laboratory and practical assignments and during the class sessions, please be sure to keep a record of all of the steps that you complete and the challenges that you face. You are also responsible for working with a partner to ensure that each of you is able to successfully complete each of the steps outlined in this assignment. **Finally, please do not upload unnecessary files in your submission directory. These files include the daily lecture slides, data-sets given out in class or any other file set(s) which is/are NOT going to be considered for your grade.**

Part 2: GitHub Repositories

Create your Account on GitHub

It is assumed that you already have an account on GitHub but if you do not, then please visit github.com/ to create your account using your Allegheny email address. Follow the account generation links and be sure to add a current photo of yourself when completing your profile.

Git and MacOS

If you have never used **Git** on your MacOS machine, then simply going to the terminal prompt and typing,

```
$ git
```

will likely signal you to download Apple's development suite, **xCode**, to install the necessary packages to run **Git**.

Git and Windows

The **Git** software does not come pre-installed on a Windows machine; it must be downloaded and installed. Please visit *Git For Windows* (at link:<https://gitforwindows.org/>) to download and install this client. In addition, there are many pieces of information on this website to help you get started. If you are having trouble with your **Git** client on your Windows machine, please first perform online searches to research answers to your concerns. If you still have trouble after completing your online research, please see the course Technical Leader(s) or instructor who may be able to help.

Add Your SSH Keys?

SSH keys allow you to interact with GitHub using secure means. You may need to add your ssh keys to your GitHub account. Please visit <https://www.ssh.com/ssh/keygen/> to learn how to set-up your security keys. Let the instructor know if you have any questions with this task. For more information about SSH keys, please watch Professor Luman's video, *SSH Keys* at the link below.

https://www.youtube.com/watch?v=qEPjUGQFmzQ&list=PLsYZRXov75ZHSwWiCk0-jd1RcTuu_-zmD

Clone Your Assignment Repository

In this section, we will be using **Git** commands. It is suggested that the reader refer to online searches for help. For example, GitHub provides good documentation at the following link; <https://git.github.io/htmldocs/git.html>.

In many cases, you will be given a new repository containing assignment materials and you will save your files in this assignment repository as you continue to work on them. Copy and paste the assignment repository cloning command into your terminal to create your assignment repositories. Be sure to place your assignment repositories in a directory such as **cs301/** to keep your class materials organized by class.

Today's assignment repository can be found at the below link to a GitHub Classroom repository. Here you will work on your assignment and then push your work to the cloud where the instructor will be able to view your work for grading. Often, there will be files in your assignment repositories which you are to edit before you submit them by using the below commands for **git**.

<https://classroom.github.com/a/iewle-4K>

To use this link, please follow the steps below.

- Click on the link and accept the assignment
- Once the importing task has completed, click on the created assignment link which will take you to your newly created GitHub repository for this lab,
- Clone this repository (bearing your name) and work locally
- As you are working on your lab, you are to commit and push regularly. The commands are the following.

```
– git add -A
– git commit -m ‘Your notes about commit here’
– git push
```

Part 3: Database Tool Research

Deliverable: In our Data Analytics course, we discussed the article; (*Twelve Million Phones, One Dataset, Zero Privacy*, By Stuart A. Thompson and Charlie Warzel, at link <https://www.nytimes.com/interactive/2019/12/19/opinion/location-tracking-cell-phone.html>. In the article, location data, extracted from smartphones, was described as being used to track the whereabouts of seemingly any person who has a smartphone. We discussed some of the potential ethical disorders which could easily arise from a seemingly innocent form of analysis.

For your deliverable, you are asked to find two (2) articles which contain discussion of the analysis of data in an application that could be easily abused, or could converge on an unethical result. Your articles ought to be fairly current – within the last three years, and are from reputable sources of journalism. For both of your articles, you are asked to write a brief summery in your GitHub Classroom repository for this assignment.

You will be editing the `writing/report.md` file of your repository. Please copy and paste the text in the file to help organize each tool in your text.

- Your report should give the name of the tool, the Web site(s) and/or an article that introduces it. In addition, you are to provide a brief, but detailed, description of the features that the tool provides. Why is this tool important to databases? What does it do? Who uses it?
- Please use the Markdown file to prepare your report. For this, you will have to edit the file; `writing/report.md`. Be sure to add your name and the date to your work.

Required Deliverables

1. Please locate the Markdown file, `writing/report.md` in your cloned assignment directory. Edit this file using Markdown code to reflect your work. If you need help on this formatting, visit the below link for help.

- **Markdown Tidbits:**
<https://www.youtube.com/watch?v=cdJEUay5IyA>
 - **Markdown Cheatsheet:**
<https://github.com/adam-p/markdown-here/wiki/Markdown-Cheatsheet>
2. Do not forget to use the above `git` commands to push your work to the cloud for the instructor to grade your assignment. Please see the instructor if you have any questions about assignment submission.