Data Analytics
CS301
Tools for Working with Data

Week 1: 26th August
Fall 2021
Oliver BONHAM-CARTER
To install for this class

- **ClassDocs** – all class material
- **Git** - to work with GitHub
- **Atom** – an editor
- **Docker** – run programs in a virtual environment (container) on your computer
- **RStudio** – Used for programming in R

Two ways to install this:
- Locally
- Using Docker

Notes to install each software are below
ClassDocs: All Class Materials

- We will be using GitHub to manage all class material. The links below are used to pull over your to classDocs repository to get slides and labs.

  - **HTTP based repository pull**: works in absence of installed ssh keys.
    - `git clone https://github.com/Allegheny-Computer-Science-301-F2021/classDocs.git`
  
  - **SSH based repository pull**: uses installed ssh keys.
    - `git clone git@github.com:Allegheny-Computer-Science-301-F2021/classDocs.git`
Installing Git

- **MacOS**: go to your *Terminal*, type in “git” and if not installed, MacOS will offer to install the free *Xcode* software development suit from Apple that contains git.

- **Ubuntu**: Git may already be installed. If not, use the command, `sudo apt install git` to install git. You will need your password.
  - Good ref: [https://www.digitalocean.com/community/tutorials/how-to-install-git-on-ubuntu-20-04](https://www.digitalocean.com/community/tutorials/how-to-install-git-on-ubuntu-20-04)

- **Windows**: Git does not come with the Windows OS and so it must be installed. Please visit [https://gitforwindows.org/](https://gitforwindows.org/) to install and learn more.
Git and Your Class Repositories

- **PULL** your classDocs before class (cloud data sent to you).

  ```shell
  git pull
  ```

- **PUSH** assignment repos to submit homework (your data sent to the cloud)

  ```shell
  git add -A
  git commit -m "My commit mesg"
  git push
  ```
The Atom Editor: Suggested for Programming

- We will be programming and Atom facilitates this task
- If you do not already have it, please download it from: https://atom.io/
Docker for Running Software

- A container in which to run programs in isolation.
- Please be sure that you machine will work with the regular Docker, **not** Docker ToolBox.
- Verify: www.cs.allegheny.edu/canirundocker

Yes!

Check the [docker docs](https://docs.docker.com) for more information about the Linux system requirements and installation procedure.

No / Maybe

- Windows: Purchase a Windows Enterprise activation key
- Dual boot: Linux and Windows
- Use another computer

All Set!
Get Started With Docker

- Running and Testing Programs with Docker and GatorGrader (Dr. Jumadinova):
  - https://www.youtube.com/watch?v=iceAgNEORCA

- Main site
  - https://www.docker.com/

- Downloads
  - https://www.docker.com/get-started

- Tutorial
  - https://www.docker.com/101-tutorial
Learning About Docker

- Play-with-Docker
  - https://www.docker.com/play-with-docker

- Once Docker has been installed, you can play with it.

- First, build a work container:
  - docker run -dp 80:80 docker/getting-started

- Then, to learn more use your browser to go to the url:
  - http://localhost/
Please Install Your Software

- We will be using Git and GitHub. Please setup your account **by next class** at https://github.com/ and also download a Git client software from https://git-scm.com/downloads (All OS’s) or https://gitforwindows.org/ (Windows only).

- We will also be using the Atom editor to write code. Please download and install your editor from https://atom.io/

- For some labs, we may be using Docker. Please download and install your Docker Desktop installation (note: not the Docker ToolBox) from https://www.docker.com/. Help: https://hub.docker.com/

- If necessary, please help each other to install this software.

**Links to download sites are above!**
R programming: A Local Install

Install directly on your machine

RStudio Desktop 1.4.1717 – Release Notes

1. Install R. RStudio requires R 3.0.1+.

2. Download RStudio Desktop. Recommended for your system:

   **DOWNLOAD RSTUDIO FOR MAC**
   1.4.1717 | 203.06MB

   Requires macOS 10.14+ (64-bit)

   https://www.rstudio.com/products/rstudio/download/#download
A Local Install of rStudio

- You must first install R and then rStudio
  - The R programming language
    - https://cran.rstudio.com/
  - Rstudio
    - https://rstudio.com/products/rstudio/download/

If you install these, you may not need to use Docker containers for your R programming.
RStudio With Docker

FYI: Using containers

version 2.1.0.5 (40693)
channel stable
Docker Alternative of: R Programming at Bash

- Build and run container:
  - `docker run -ti --rm r-base`

- Linux, Mac; Build, mount local drive and run container:
  - `sudo docker run -ti --rm -v "$PWD":/home/docker -w /home/docker -u docker r-base`

- Windows; Build, mount local drive and run container:
  - `docker run -ti --rm -v /home/docker -w /home/docker -u docker r-base`

Note: the directory where you run this becomes your local directory in the container.
Docker Container Setup: rStudio

Note: the directory where you run this becomes your local directory in the container.

- **Username**: rstudio
  - **Password**: letmein

- **Linux, Mac; Build, mount local drive and run container**:  
  `docker run --rm -p 8787:8787 -e PASSWORD=letmein rocker/rstudio`

- **Windows; Build, mount local drive and run container**:  
  - `docker run --rm -e PASSWORD=letmein -p 8787:8787 -v "PWD":/home/rstudio/ rocker/verse`

- **Browser**:  
  - **URL**: Use Browser address: http://localhost:8787/
Programming in R

R and RStudio

To run:
Type "R" at terminal

To run:
Find its icon or type rstudio at terminal
R by Jdoodle

- https://www.jdoodle.com/execute-r-online

```r
x <- 10
y <- 25
z <- sum(x, y)
cat("x + y = ", z)

x + y = 35
```