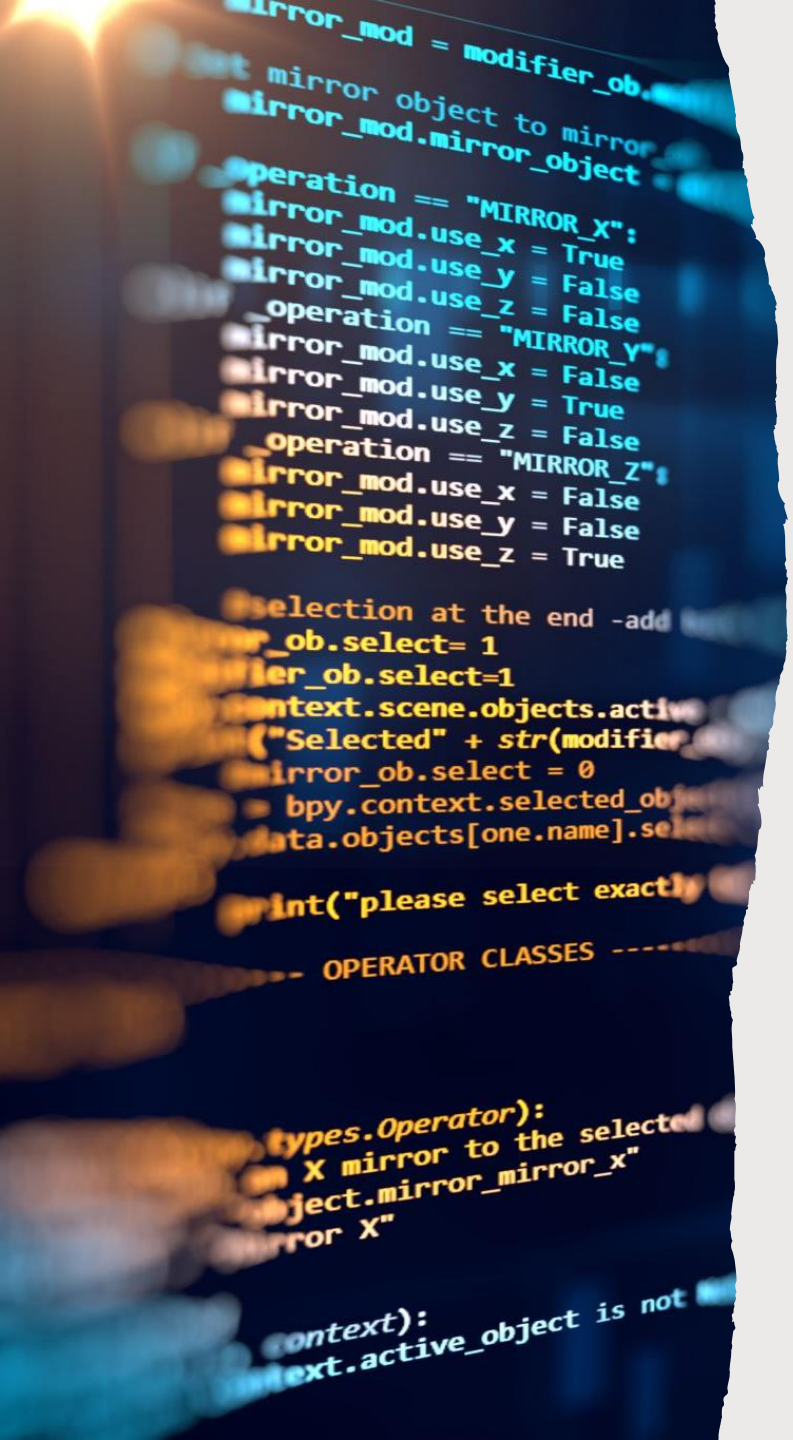
Module 0: Data Science Hacking Tools

CSCI1360: Foundations for Informatics and Data Analytics – Spring 2023

Objectives

- Use Command prompt\Terminal for OS basic actions
- Use Conda to install scientific tools
- Use Jupyter Notebook to write Python code
- Create GitHub account
- Use Git to interact with GitHub repositories



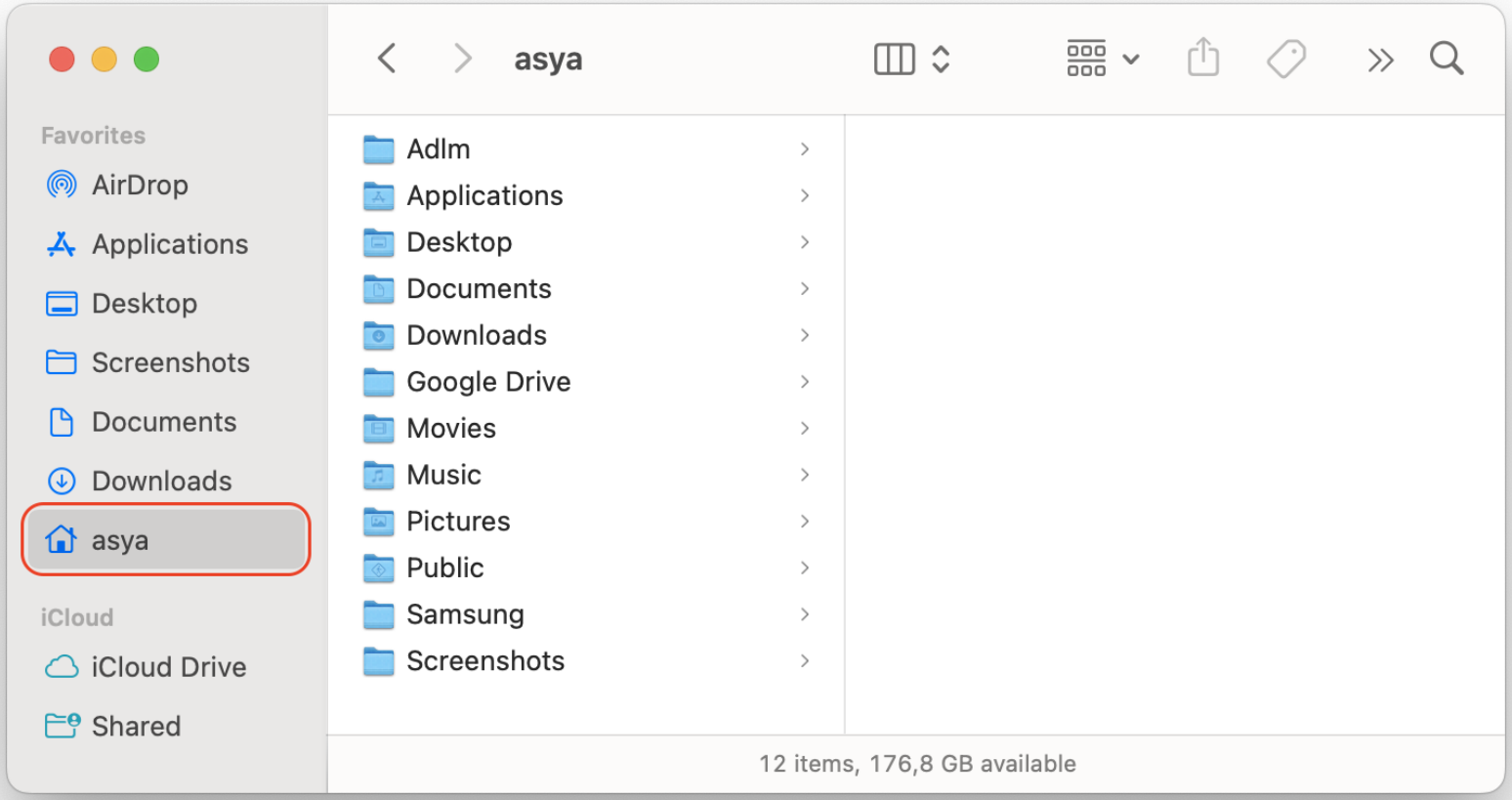
Command Prompt\Terminal

- Windows: Command Line Prompt or CMD
- Linux-based or Mac OS: Terminal
- Used to interact with the OS
  - To perform file \ folder actions
  - To start a program

Important CMD\Terminal Commands

LINUX	WINDOWS	DESCRIPTION
ls	dir	List files and folders in the current working directory.
cd	cd	Change the current working directory.
cp	copy	Copy a file or folder to a new location.
mv	move / rename	Move or rename a file or folder..
mkdir	md	Make a new folder or subdirectory.
rm	del or rmdir	Delete a file or folder.
echo	echo	Send some text to the console screen.
cat	type	Show the contents of a text-based file
pwd	cd	Prints current path

Home Folder


asya
Favorites
AirDrop
Applications
Desktop
Screenshots
Documents
Downloads
asya
iCloud
iCloud Drive
Shared
Adlm
Applications
Desktop
Documents
Downloads
Google Drive
Movies
Music
Pictures
Public
Samsung
Screenshots
12 items, 176,8 GB available

Course Work Folder

LINUX	WINDOWS	DESCRIPTION
cd ~	cd c:\Users\<user_name>	Change the current working directory to Home
ls	dir	List files and folders in home directory
cd Desktop	cd Desktop	Changes to your Desktop folder
mkdir test_directory	md test_directory	Creates test_directory in Desktop
cd test_directory	cd test_directory	Change to test_directory
echo 'Hello World' > test.txt	echo 'Hello World' > test.txt	Creates test.txt
rm test.txt	del test.txt	Deletes test.txt
cd ..	cd ..	Moves to parent directory
rm test_directory -rf	rmdir test_directory	Deletes test_directory
mkdir CSCI1360	md CSCI1360	Creates CSCI1360

Miniconda Installation

- Link: https://docs.conda.io/en/latest/miniconda.html
- Mac OS
  - 64-bit bash package (.sh package)
  - Open terminal
  - Navigate to the package download folder
  - bash <package-name.sh>
- Windows
  - Download and install (Miniconda3 Windows 64-bit)
  - Use the default installation settings
  - Search (Anaconda Prompt)

Anaconda\Miniconda

- An open-source package management system and environment management system
- Regular coding environment
  - Packages depend on system libraries
  - Install packages with all the required steps by the vendor
  - One environment for all packages
- Conda environment
  - Isolated multiple environments
  - Packages depend on Conda libraries
  - One step package installation

Conda Basic Commands

Using environments	
Create a new environment named py35, install Python 3.5	`conda create --name py35 python=3.5`
Activate the new environment to use it	`WINDOWS: activate py35` `LINUX, macOS: source activate py35`
Get a list of all my environments, active environment is shown with *	`conda env list`
Make exact copy of an environment	`conda create --clone py35 --name py35-2`
List all packages and versions installed in active environment	`conda list`

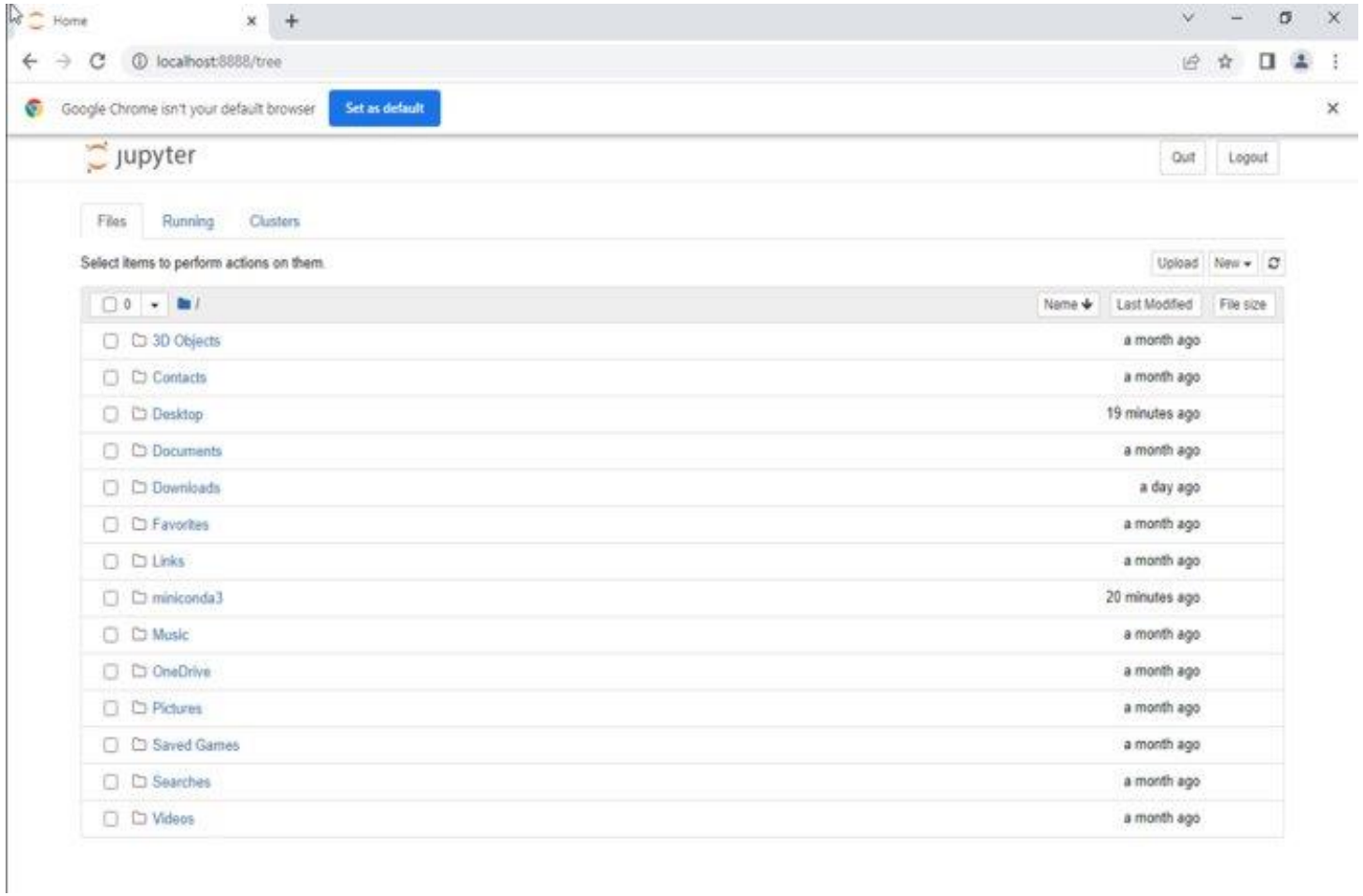
Installing and updating packages	
Install a new package (Jupyter Notebook) in the active environment	`conda install jupyter`
Run an installed package (Jupyter Notebook)	`jupyter-notebook`
Deactivate the current environment	`WINDOWS: deactivate` `macOS, LINUX: source deactivate`

Create Conda Environment

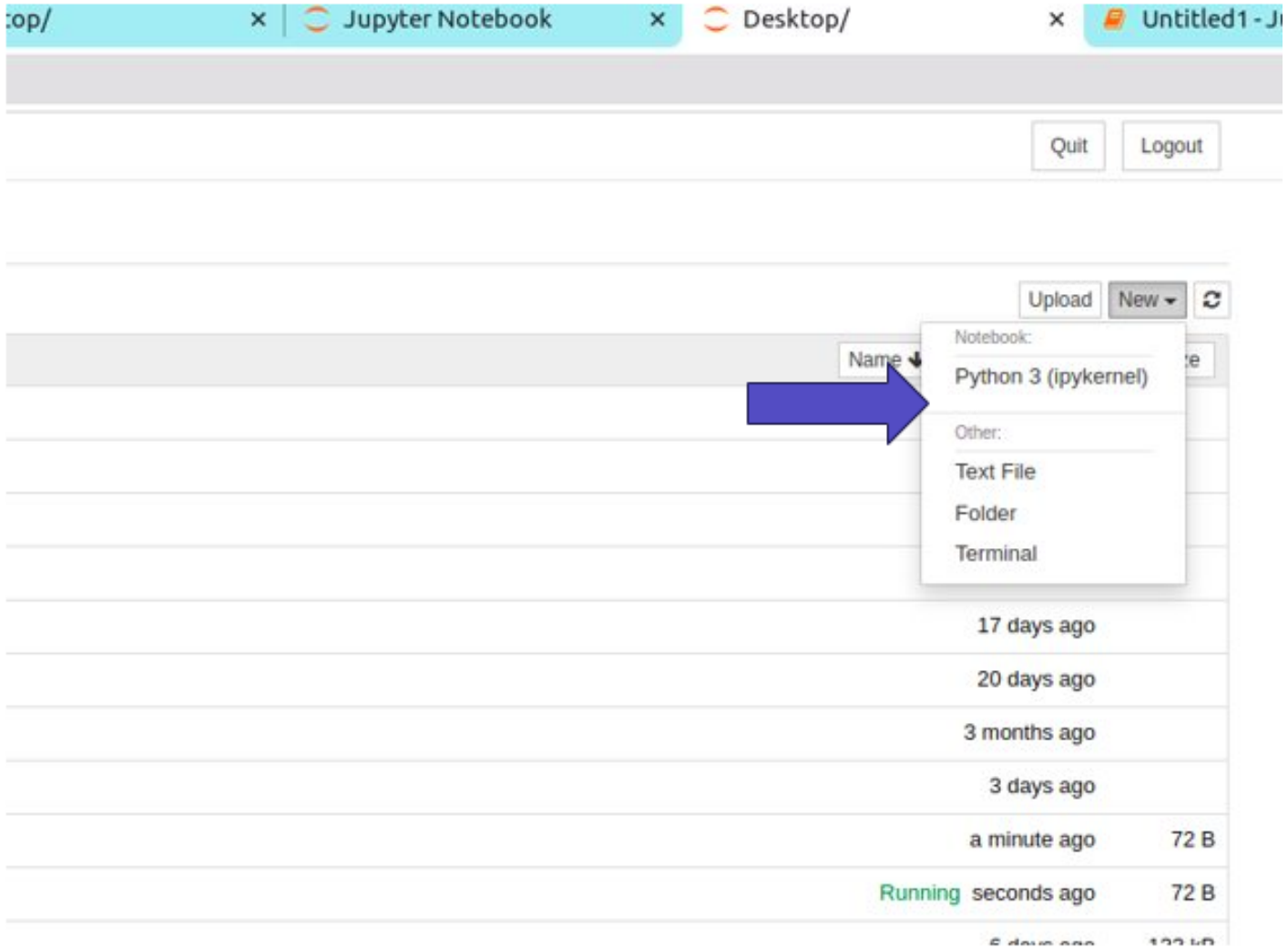
- `conda create --name csci1360 python=3.9`
- List the environments
- Activate the environment
- Packages to install
  - Notebook: `conda install notebook`
  - Git: `conda install git`

Jupyter Notebook

- Open-source web-based coding platform
- Features
  - Easy to use
  - Live code and visualizations output in cells
  - Enable explanatory text in markdown cells

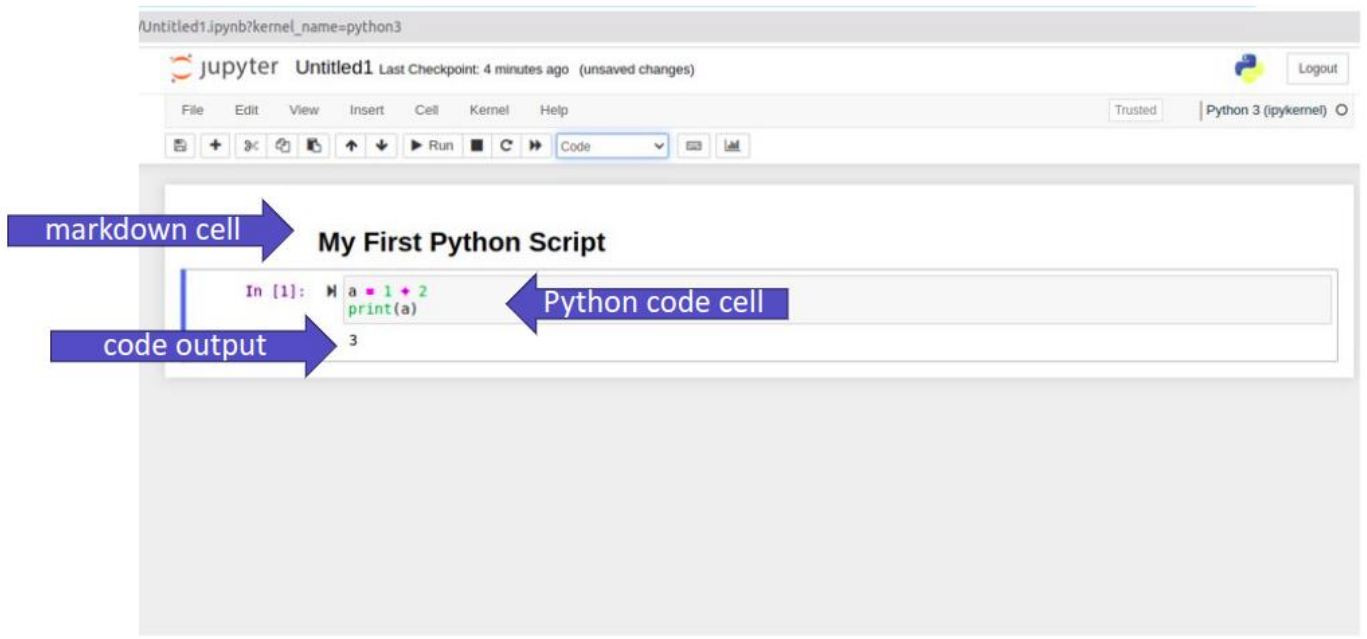


Create a Notebook



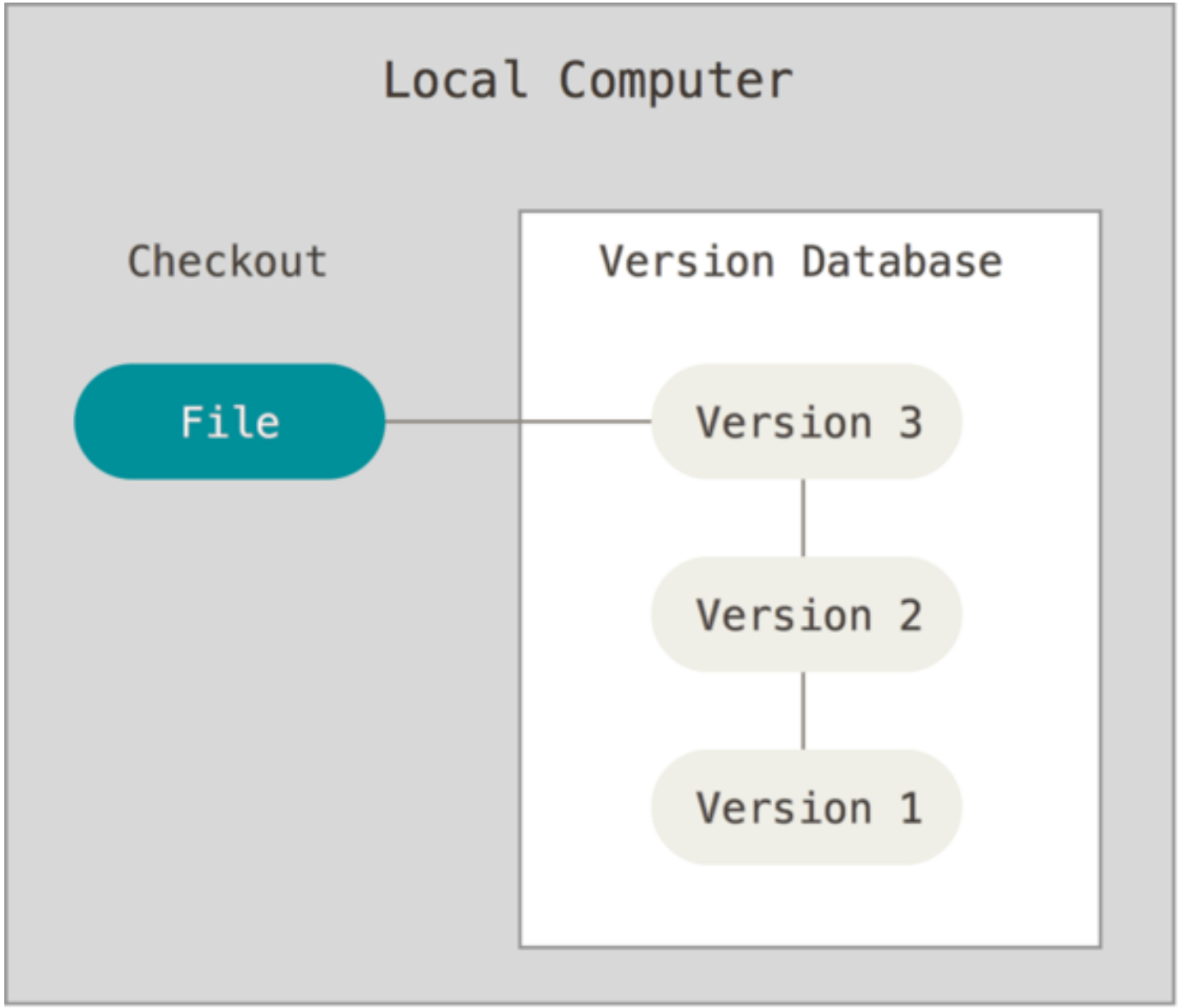
My First Python Script

- Click the Run button to execute the code.



Git

- Version control is a system that records changes to a file or set of files over time so that you can recall specific versions later
- Usages
  - Local environment
  - Collaboration environment (beyond the scope of this course)



Git Basics

- `git clone <url>`
  - To clone an online repository to your local environment
  - To include in your local Git version control system
- `git add <file>`
  - Start tracking a file
- `git commit –m "my commit note"`
  - Commits changes, new version of your code will be stored in Git database
  - Changes after last `add` command won't be considered
  - Unless: `git commit -a –m "my commit note"`

Git Basics

- `git status`
  - to determine which files are in which state
- `git push`
  - Updates the online repository with your local changes
  - Submits your assignment solution
- `git config --global user.email "you@example.com"`
  - to set your GitHub account for push commands
- `git config --global user.name "Your Name"`
  - to set your name for push commands

GitHub

- GitHub is a hosting service for software development and version control using Git
- Cloud based Git repository
- Web interface and commands interface through Git
- A repository contains all of your project's files and each file's revision history.

Create Your GitHub Account

Use your UGA email for this account
Generate private token to be used when submitting your assignment solutions
settings
Developer settings
Personal access tokens
Tokens
Generate new token (Classic)
Set Expiration
click "repo" checkbox
Generate token
Keep token file private in your machine

GitHub Classroom

- A platform that lets teachers and students interact in programming based courses
- https://github.com/UGA-CSCI1360
- Assignment and Lectures repositories
- Steps to work on an assignment
  - Get your assignment repository created through the assignment link
  - Clone the repository to your local course folder
  - Add your solution to the notebook template
  - Commit and push your changes using Git

Demo

```
cd [PATH to CSCI1360]
# create assignment-0 directory
cd assignment-0
git clone [assignment link]
jupyter-notebook
git add *.ipynb
git commit -m "my commit note"
git push
# Enter GitHub username and token string
```