

# Algorithms and Data Structures for Data Science

## lab\_fundamentals

CS 277

January 19, 2024

Brad Solomon



UNIVERSITY OF  
**ILLINOIS**  
URBANA - CHAMPAIGN

Department of Computer Science

# Learning Objectives

Discuss lab structure and acceptable groupwork policies

Review the basics of variable type, conditionals and loops

Review the Jupyter notebook format for autograded assessments

Practice identifying sub-problems, coding functions based on I/O, and testing code

# Setting up your own machine

Lab assignments will (hopefully) all be runnable on Prairielearn

Mini-projects may involve larger datasets or runtimes than PL can handle

You must have **Python3** with **Numpy**, **Matplotlib**, and **Pandas** installed:

<https://courses.grainger.illinois.edu/cs277/resources/own-machine/>

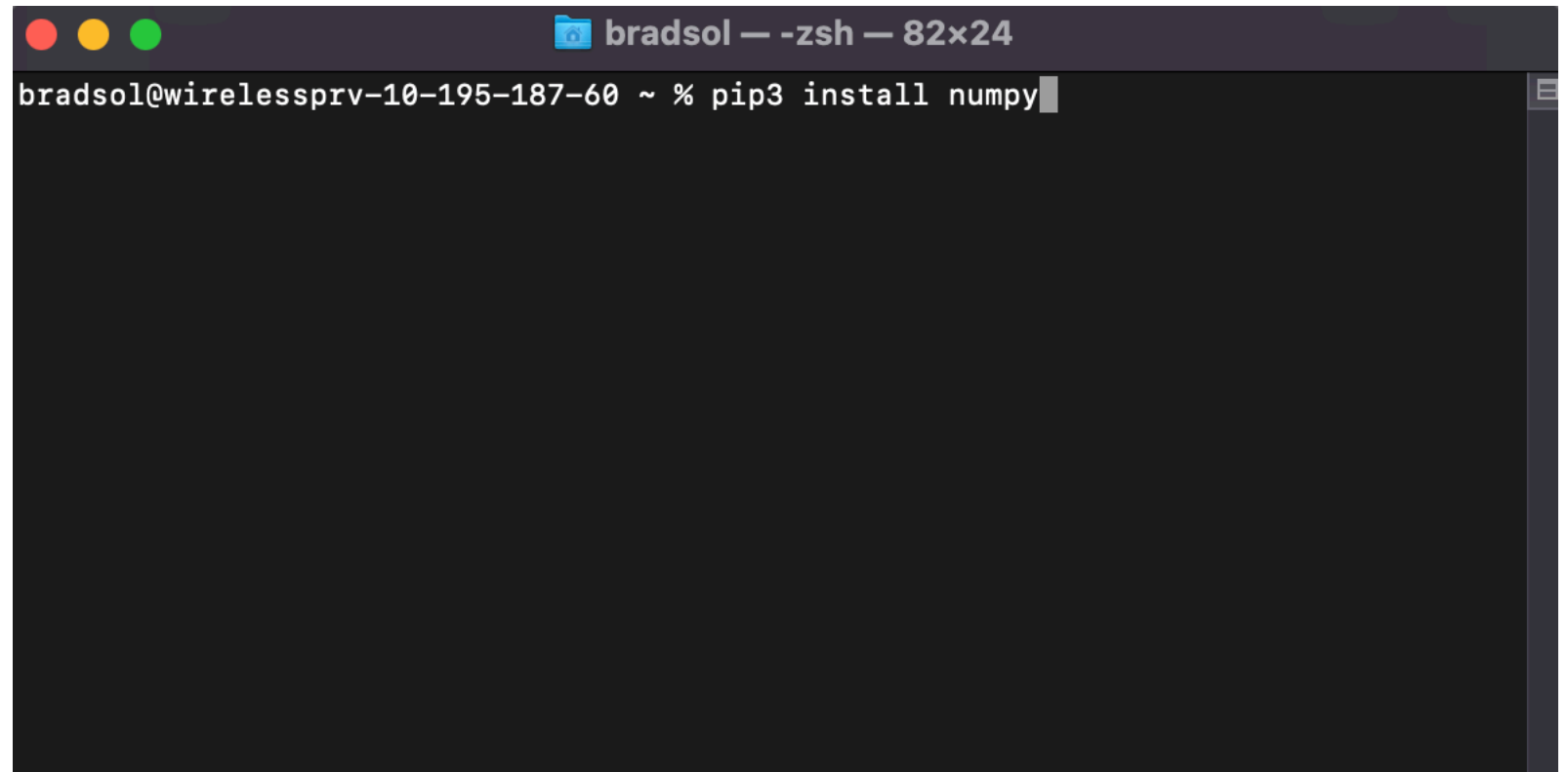
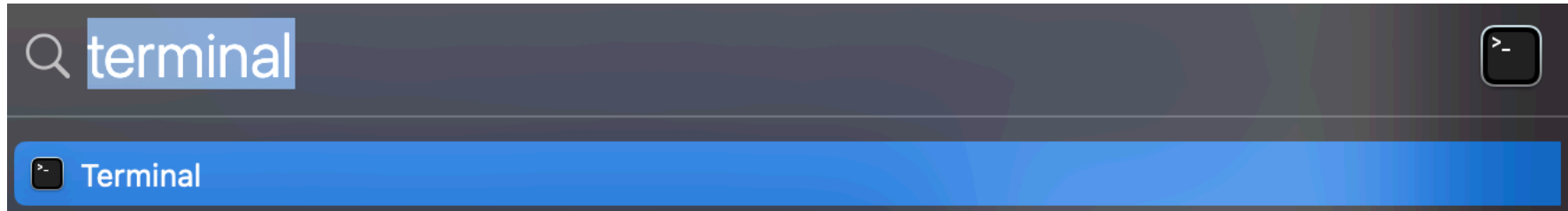
# Step 1: Installing Python

Use the official site to download: <https://www.python.org/downloads/>



# Step 2: Installing Packages

Pip3 comes standard with Python now so installing packages is easy!



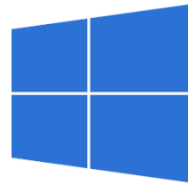
# Step 3: Install VSCode

You can use your own code editor but this is recommended:

<https://code.visualstudio.com/download>

## Download Visual Studio Code

Free and built on open source. Integrated Git, debugging and extensions.



↓ Windows

Windows 8, 10, 11

User Installer	x64	x86	Arm64
System Installer	x64	x86	Arm64
.zip	x64	x86	Arm64
CLI	x64	x86	Arm64



↓ .deb

Debian, Ubuntu

↓ .rpm

Red Hat, Fedora, SUSE

.deb	x64	Arm32	Arm64
.rpm	x64	Arm32	Arm64
.tar.gz	x64	Arm32	Arm64
Snap	Snap Store		
CLI	x64	Arm32	Arm64



↓ Mac

macOS 10.11+

.zip	Intel chip	Apple silicon	Universal
CLI	Intel chip	Apple silicon	

# How to work through coding assignments

## **1. Make sure you understand the problem**

What is the **input** and **output** of the problem?

Can you break the problem down into parts?

Do any of the sub-problems build off each other?

## **2. Solve (and test) each part one at a time**

What should the output be given an input?

Are there any edge cases you are missing?

# getTotalTime()

Given **HH:MM:SS** format, I want to know the exact difference between start and stop times in minutes. How would we approach this problem?



# getSmallestEven()

Given three integers, return the smallest even number.

# electricBill()

Given one integer, return total energy bill according to conditional logic.

# oddCountByIncrement()

Given start, stop, and increment integers, return all odd values in **range**.

# sumUntilGreater()

Given start and stop integers, sum numbers from start until total is greater than stop.

# Lab Tips

First couple of assessments will help establish background knowledge

Give feedback (anonymously or not) about class pacing