



# Introduction

## David Koes

8/31/21



# Hybrid Instruction

In-person (Murdoch 814)

Follow all University guidelines

Building access

Masks required

Zoom

Attend **synchronously**

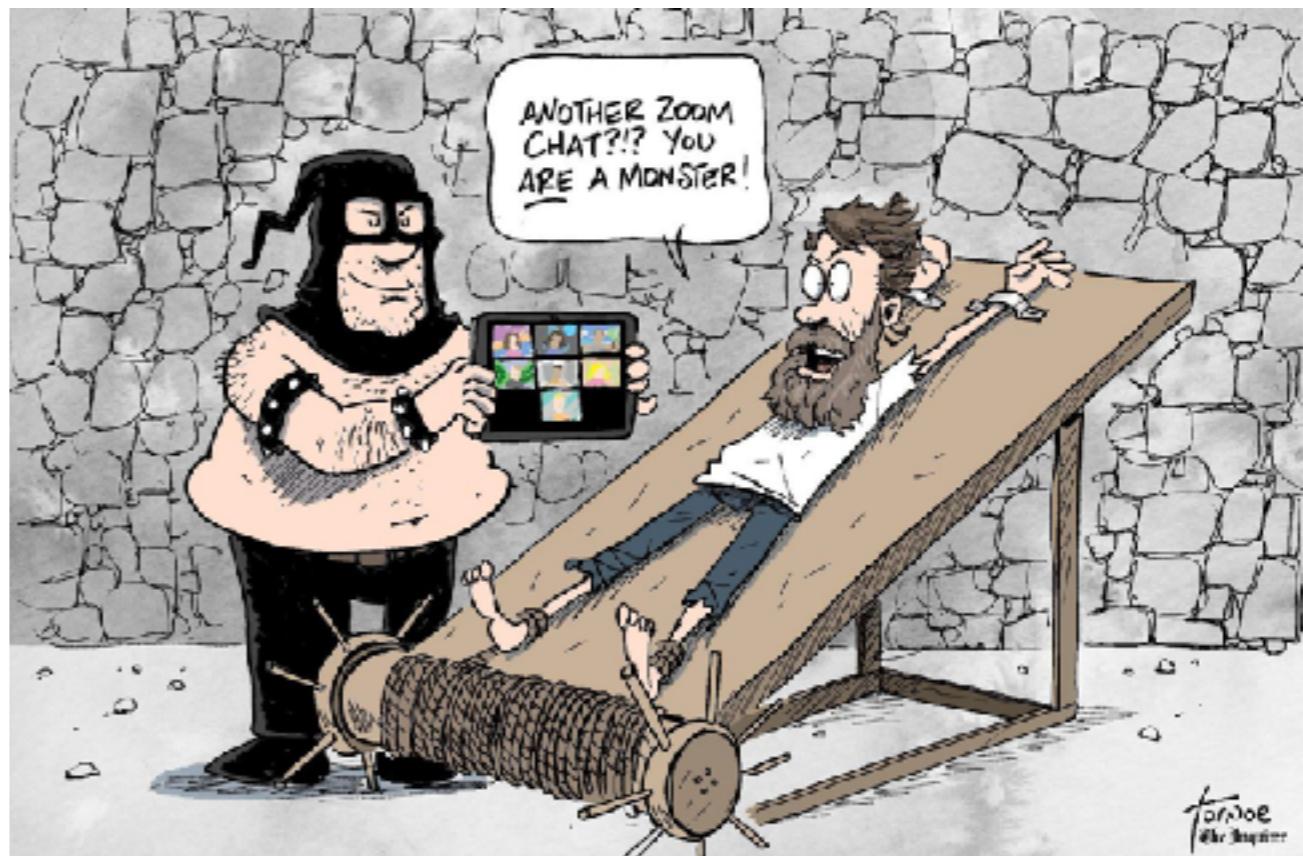
Lecture recordings are for reference





# Zoom Etiquette

- Video on is preferred (but not required)
- Stay on Mute
- Use Chat to ask/respond to questions
  - "Raise Hand" if I don't notice chat or want to share screen/speak on Zoom
- Let me know if something isn't working





# “Bioinformatics”

## Bioinformatics, Computational, and Systems Biology

**Bioinformatics** bio-in-fər-mä-tiks bio-in-fər-mä-tiks bio-in-fər-mä-tiks  
*construction* \bī-ō-in-fär-'ma-tiks\

**Definition of BIOINFORMATICS** ..... [Cite!](#) [g+1](#) [Like](#)

**Bioimaging** : the collection, classification, storage, and analysis of biochemical and biological information using computers especially as applied to molecular genetics and genomics

— bio·in·for·mat·ic [adjective](#)

Molecular  
Dynamics

Systems Modeling

Protein Dynamics

Sequence Analysis

Drug Discovery

Cheminformatics

Data Analysis

Proteomics

Genomics

Biomedical  
Informatics

Protein Structure



# “Programming”

Computer programming

pro·gram·ming *noun* \-mĕng\

## Definition of PROGRAMMING

Cite! G+1 Like

- 1** : the planning, scheduling, or performing of a program
- 2** **a** : the process of instructing or learning by means of an instructional program
- b** : the process of preparing an instructional program

implementation of the build system and management of derived artefacts such as machine code of computer programs. The algorithm is often only represented in human-parseable form and reasoned about using logic. Source code is written in one or more programming languages (such as C++, C#, Java, Python, Smalltalk, etc.). The purpose of programming is to find a sequence of instructions that will automate performing a specific task or solve a given problem. The process of programming thus often requires expertise in many different subjects, including knowledge of the application domain, specialized algorithms and formal logic.

*There is an on-going debate on the extent to which the writing of programs is an art form, a craft, or an engineering discipline.*

Computing  
s,  
algorithm,  
and its  
a target  
code,

“Python”





# Python

Designed to be easy to learn

Full featured, powerful language

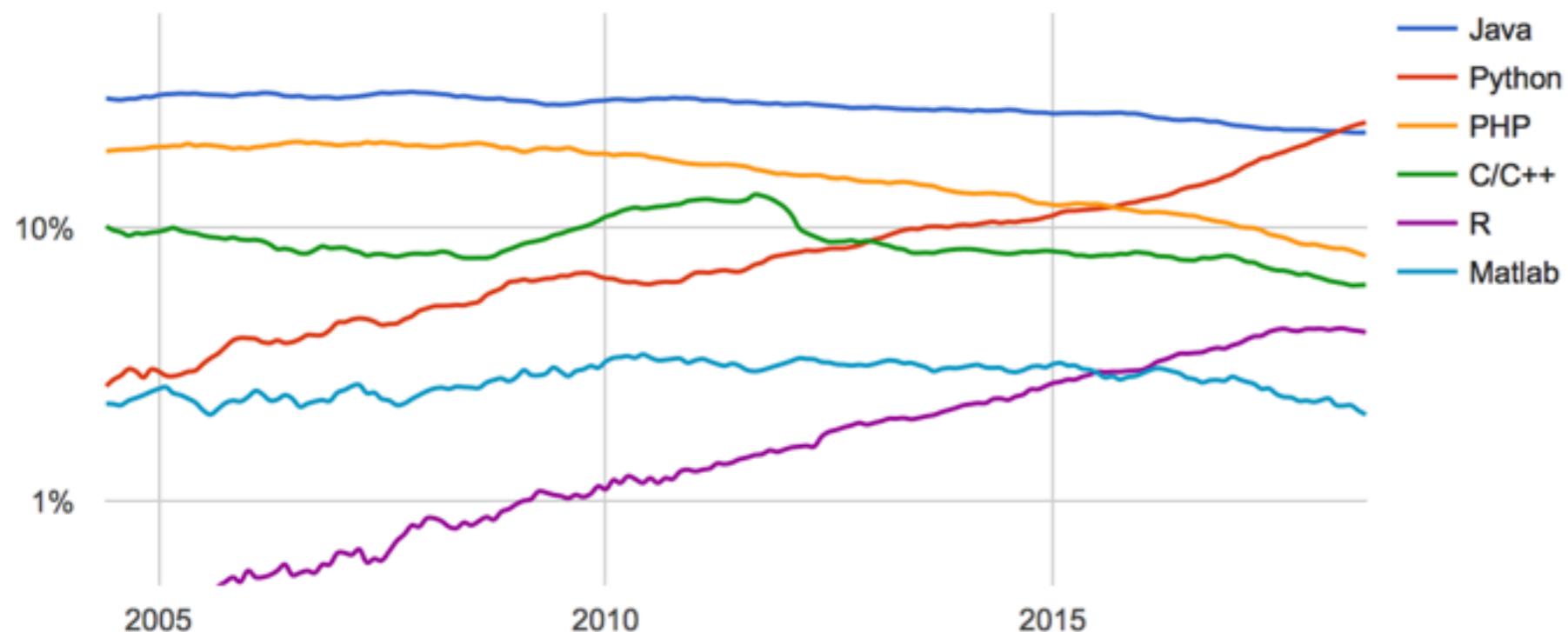
Free - Costs nothing and open-source

Ideal for *scripting*

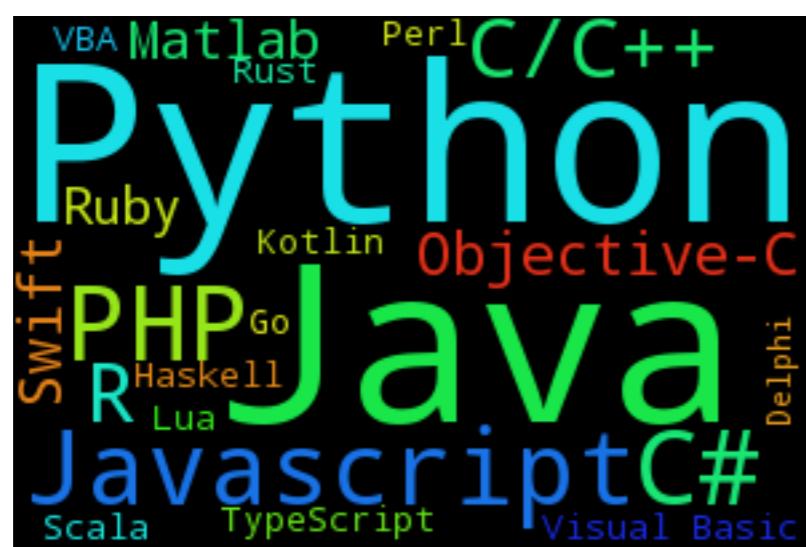
Popular



### PYPL PopularityY of Programming Language



<http://pypl.github.io/PYPL.html>



Language Rank	Types	Spectrum Ranking
1. Python	🌐💻⚙️	100.0
2. C++	💻⚙️	98.4
3. C	💻⚙️	98.2
4. Java	🌐💻	97.5
5. C#	🌐💻	89.8
6. PHP	🌐	85.4
7. R	💻	83.3
8. JavaScript	🌐💻	82.8
9. Go	🌐💻	76.7
10. Assembly	⚙️	74.5

<https://spectrum.ieee.org/at-work/innovation/the-2018-top-programming-languages>



Translations ▾ License ▾ Programming Language ▾ Status

**Bio-Informatics Software**

Java (892)

Python (330)

Perl (315)

C++ (300)

Hot topics in Bio-Informatics Software

S

Translations ▾ License ▾ Programming Language ▾ Status

**Molecular Science Software**

Java (109)

Python (75)

C++ (69)

C (61)

Hot topics in Molecular Science

S

Translations ▾ License ▾ Programming Language ▾ Status

**Chemistry Software**

Java (155)

C++ (102)

Python (100)

C (57)

Hot topics in Chemistry

S

Translations ▾ License ▾ Programming Language ▾ Status

**Physics Software**

C++ (311)

C++ (311)

Python (174)

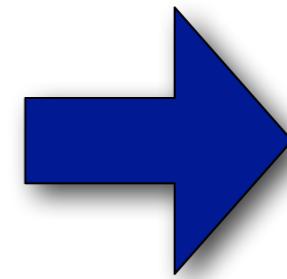
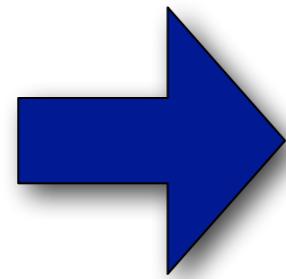
C (173)

Hot topics in Physics Software

S

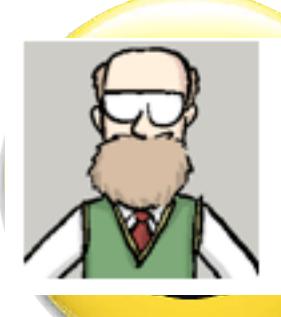
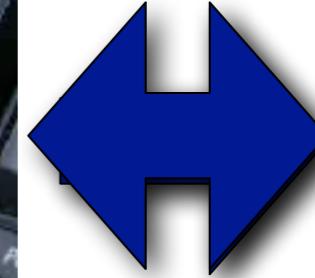
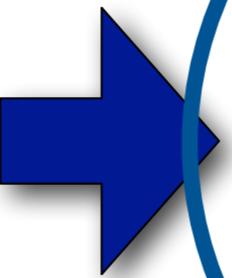
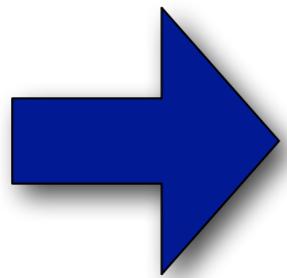
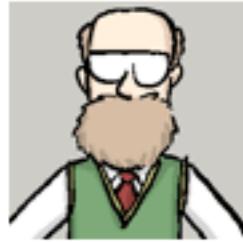


# Course Goals



“Analyze the data”

“Do it again”



“Analyze the data”

“Do it again”



# Course Goals

Gain experience programming

Learn Python

Survey computational methods

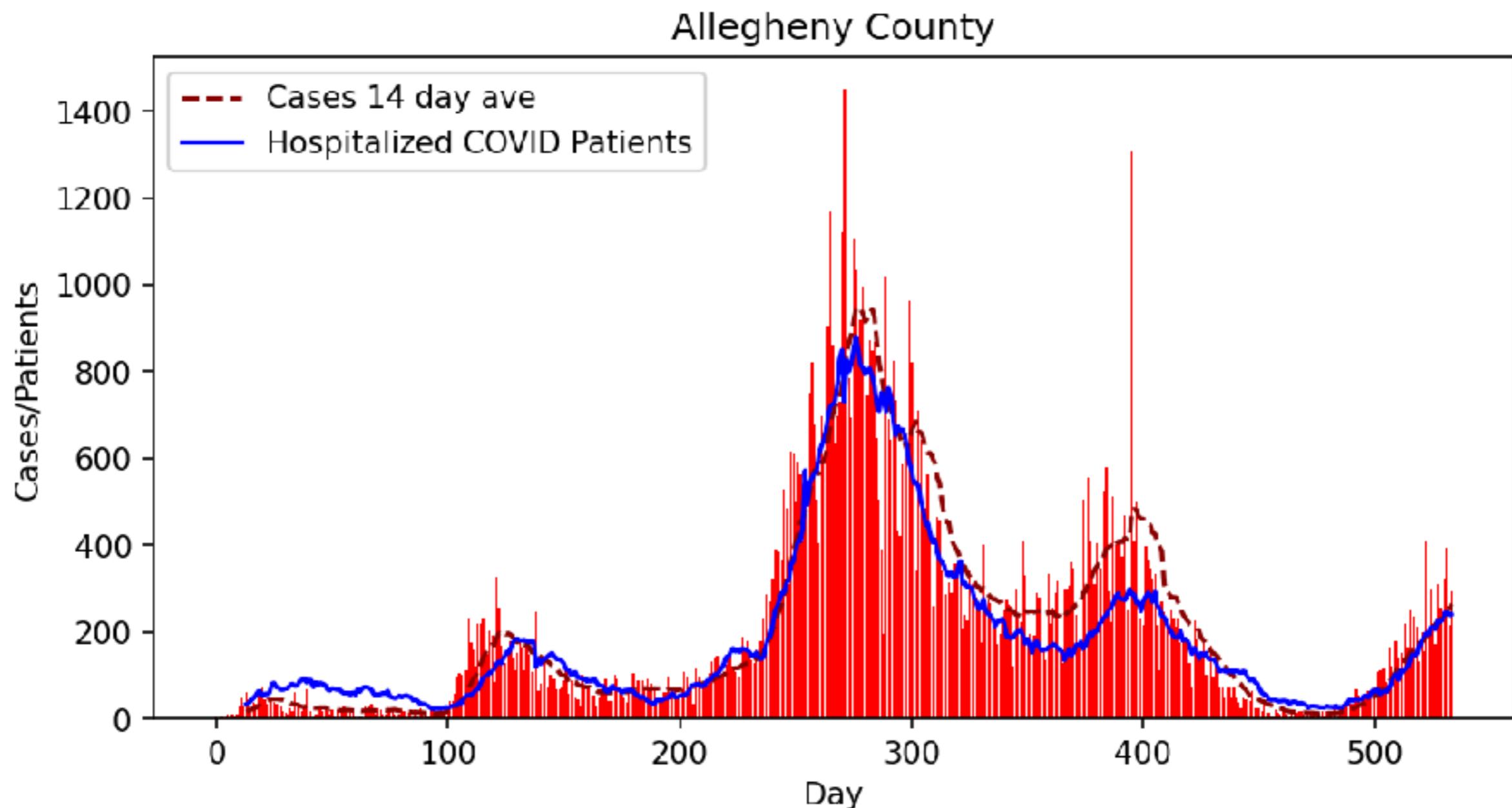
*Improve skills to be a more productive  
and successful researcher*

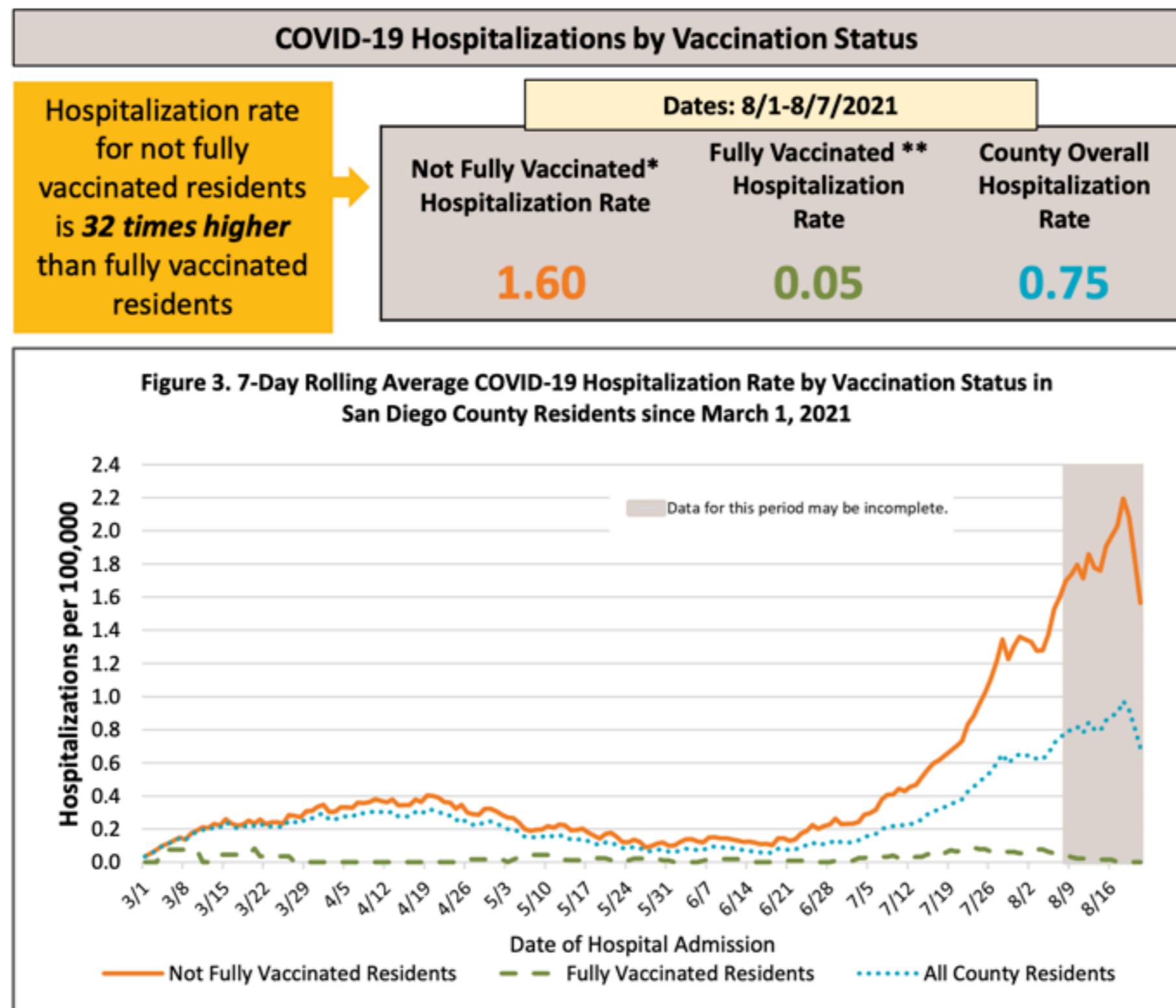


# Coarse Goal 2021

Don't get sick!

<https://www.coronavirus.pitt.edu/>







# Logistics

12 Programming Assignments

Due midnight on Tuesday

Autograded - submit until it works

1 day late - 90% credit

2 days late - 50% credit

**>3 days late - 0% credit**

Late penalty only applied to *additional* points

Each assignment worth ~7%

Final Project (create an assignment)

Final Grades

A: >93%

B: >85%



# Logistics

## Communication over Slack

<http://mscbio2025.slack.com>

The screenshot shows the Slack web interface for the '#general' channel of the 'mscbio2025' workspace. The sidebar on the left lists various workspace settings and team members. The main area displays four images: a classroom setup, a hallway, a street view, and a person standing near an '8 TH FLOOR' sign. A message from user 'dkoes' at 8:50 PM reads: 'Please bring a laptop to class.' A message input field at the bottom allows users to send messages to the channel.



# Getting Help

No TA this year :-(

General questions

Ask in #general

Use threaded conversations

Ask after class in classroom

One-on-one help

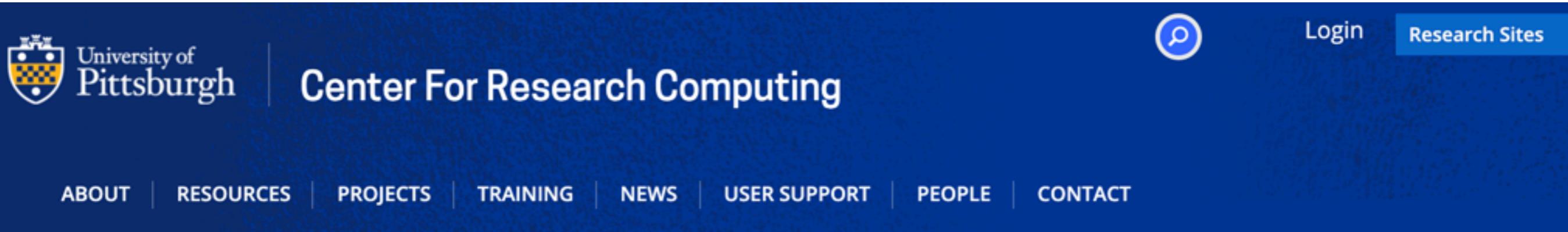
Come to my office (748) after class

Direct message on slack



# Extra Training

<https://crc.pitt.edu/training/fall-2021-next-generation-sequencing-workshops>



The screenshot shows the University of Pittsburgh Center for Research Computing website. At the top left is the University of Pittsburgh logo and name. In the center is the text "Center For Research Computing". At the top right are links for "Login" and "Research Sites", and a search icon. Below the header is a navigation bar with links for "ABOUT", "RESOURCES", "PROJECTS", "TRAINING", "NEWS", "USER SUPPORT", "PEOPLE", and "CONTACT".

## Fall 2021 Next Generation Sequencing Workshops

These workshops were supported in part by the University of Pittsburgh seed project titled "University of Pittsburgh Computational Gen Training Program".

High throughput sequencing has brought abundant sequence data along with a wealth of new "-omics" protocols, and this explosion of as bewildering as it is exciting. Our multi-day hands-on workshops give researchers the research, open-sourced tools to plan and execute successful bioinformatics and genomics experiments. These workshops, taught by experienced Bioinformatics core faculty, cover both the theoretical and practical aspects of a wide range of NGS data, using the [HTC cluster](#).

These workshop have hands-on components that require the following requirements be set up before a workshop begins.

1. Participants should have an account on the HTC cluster, which is the cluster we will use for demonstration purposes. (page 1 of [this documentation](#))



# Academic Honesty

Do your own work

Do **not** share or look at other students' code

Do discuss concepts and problem solving strategies



# Website

<http://mscbio2025.net>

Commandline Basics

Laptop setup

Change Account Password on [python.mscbio2025.net](http://python.mscbio2025.net)