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

Bioimaging
Molecular Dynamics
Sequence Analysis
Drug Discovery
Cheminformatics
Systems Modeling
Data Analysis
Protein Dynamics
Protein Structure

Proteomics
Genomics
Biomedical Informatics
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.

“Programming”

**Definition of PROGRAMMING**

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

**Programming**

Computing
algorithm, and its
target code,
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-readable 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.
“Python”
Python

Designed to be easy to learn
Full featured, powerful language
Free - Costs nothing and open-source
Ideal for *scripting*
Popular
Introduction to Bioinformatics
Programming in Python

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

<table>
<thead>
<tr>
<th>Department of Computational Biology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction to Bioinformatics</td>
</tr>
<tr>
<td>Programming in Python</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Translations</th>
<th>License</th>
<th>Programming Language</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bio-Informatics Software</td>
<td></td>
<td>Java (892)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Python (330)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Perl (315)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C++ (300)</td>
<td></td>
</tr>
<tr>
<td>Molecular Science Software</td>
<td></td>
<td>Java (109)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Python (75)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C++ (69)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C (61)</td>
<td></td>
</tr>
<tr>
<td>Chemistry Software</td>
<td></td>
<td>Java (155)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C++ (102)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Python (100)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C (57)</td>
<td></td>
</tr>
<tr>
<td>Physics Software</td>
<td></td>
<td>C++ (311)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Python (174)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C (173)</td>
<td></td>
</tr>
</tbody>
</table>
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/
COVID-19 Hospitalizations by Vaccination Status

Hospitalization rate for not fully vaccinated residents is **32 times higher** than fully vaccinated residents.

<table>
<thead>
<tr>
<th>Vaccination Status</th>
<th>Hospitalization Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Fully Vaccinated*</td>
<td>1.60</td>
</tr>
<tr>
<td>Fully Vaccinated **</td>
<td>0.05</td>
</tr>
<tr>
<td>County Overall Rate</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Dates: 8/1-8/7/2021

Figure 3. 7-Day Rolling Average COVID-19 Hospitalization Rate by Vaccination Status in San Diego County Residents since March 1, 2021

Data for this period may be incomplete.
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
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

Fall 2021 Next Generation Sequencing Workshops

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

High throughput sequencing has brought abundant sequence data along with a wealth of new "-omics" protocols, and this explosion of data is 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)
2. This week-long course requires that participants either have a Pittnet user (for on-campus) or logged in by VPN (page 2 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