BST281: Genomic Data Manipulation, Spring 2018

Monday 01: Course overview

Convey the conceptual understanding and technical tools to do high-quality quantitative biology research.

<http://huttenhower.sph.harvard.edu/bst281>

## Location

Lecture: MW 3:45-5:15 Kresge 200

Lab: R 3:45-5:15 FXB G03 (mandatory during problem set assignments)

## Instructors and teaching assistant

Curtis Huttenhower, chuttenh@hsph.harvard.edu Office hours: W 10:00-11:00, SPH1 413

Eric Franzosa, franzosa@hsph.harvard.edu Office hours: W 10:00-11:00, SPH1 413

Mike MacArthur, macarthur@g.harvard.edu Office hours: TBD

## Course structure

6 biweekly problem during the first 3/4 of the course.

 Assigned during class on Monday, due by end-of-day (midnight) the following week's Friday (~2 weeks).

Group midterm journal club, plus final project (including 2-4 page individual writeups).

## Collaboration policy

Assignments and code may be discussed but not copied; all problem sets must be submitted individually and will be checked for duplications.

Questions but not specific code should be posted publicly on the course discussion board; please email questions about specific code privately to TAs.

# Python

Simple, efficient language with many tools already available for biological data handling.

 Think of it as a computational pipette: not complicated, but flexible and crucial for accurate experiments!

Codecademy: resource for interactive learning.

Anaconda: software allowing Python command line execution.

Jupyter: graphical interface to interactive Python snippets, visualization, and documentation.

# Textbooks

Bioinformatics overview: Pevsner, Chapter 1 p3-15

Computing overview: Haddock and Dunn, Chapter 0 p1-6

Python setup: Chapter 1 p9-15, Chapter 4 p47-55, 59-62

# Literature

[A systems approach to mapping DNA damage response pathways. Workman et al, Science 2006.](https://www.ncbi.nlm.nih.gov/pubmed/16709784)

[Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. Cimermancic et al, Cell 2014.](https://www.ncbi.nlm.nih.gov/pubmed/25036635)