

# L09: Logistics

Eric Franzosa (franzosa@hsph.harvard.edu)

<http://franzosa.net/bst273>

# Overview

- GSU strike
- Final Project overview
- Regular Expressions (Jupyter Notebook)

# GSU Strike

# GSU Strike officially began yesterday

- BST 273 lectures and assessments will continue
- We may be down TA support during the strike
  - = Reduced office hours
    - mine will proceed as usual, and you can also email me
  - = Longer wait for HW feedback
- This should not affect your grade for the course
  - HW4 due date will be extended to next Monday
  - I will post HW4 hints in proportion to missed office hours
  - Similar adjustments to Final Project if strike continues into next week
- Email me if you have concerns that are specific to your situation
  - E.g. if you need your final grade on time for early graduation

# Final Project

# Final Project

- Officially begins next Monday and runs through the end of the course
- You will choose between
  - Default Final Project
  - Custom Final Project
- Today's overview is to help you make that decision
- If doing a custom final project, you must send me a proposal this week
  - Details in subsequent slides
- **I will post a more detailed Final Project description on Canvas next Monday**
  - Along with sample files for the Default Final Project

# Default Final Project: Document Comparison



- Assessing similarity between documents is a real-world computing task
  - Assigning unknown works to authors
  - Detecting plagiarism
  - Making book recommendations
- You will write a script which...
  - Takes two documents (text files) as input
  - Breaks up and counts the words in each document
  - Compares the word counts using a given formula (Bray-Curtis distance)
  - Prints the documents' similarity to the screen
- You will be given
  - An English-language description of a recommended solution
  - Sample documents
  - A grid of expected similarities between those documents

# Custom Final Project

If you are considering a Custom Final Project, email me with your idea before the end of this week so we can refine the scope ASAP.

- You will propose a script which...
  - Takes some sort of input data as a file
  - Analyzes that data
  - Makes a report about the analysis (to the screen or a separate file)
- Appropriate scope...
  - I expect custom projects to be similar in scope to HW4
    - i.e. about 3 functions and some support code
  - I will approve a script that does a bit less than that if it uses a new module
    - i.e. one that we did not specifically cover in class
- You must provide
  - All default hand-in materials (see next slide)
  - A sample input file
  - A sample command, output, and means to verify that the output is correct



# What you'll hand in (both Final Project options)

- The Python script you wrote
- A custom input file for your script (if applicable)
- A sample command and output from your script
- A separate text document containing answers to free-response questions
  - General feedback about the course
  - Specific questions about your approach to solving the Final Project task
    - A place to get partial credit if you aren't able to make a functioning script

Questions?

Transition to  
L09 lecture material