

PHY 250L – Spring 2018

Python tutorial 2

Welcome back to the PHY 250L Python programming tutorial!

Things you should understand after week 2:

- plotting bivariate data
- histograms
- reading/writing files

Problems for 2.29.2018 The following problems should be completed and uploaded to Sakai by 09:45 on 2.29.2018. Each problem should correspond to its own python program (*i.e.*, each problem will correspond to a single file). The preferred names for the files are indicated in each problem.

1. collatz_len.py, 20 points

Though it is quite simple to state, the Collatz Conjecture is one of the most pernicious outstanding problems in number theory. For a number n (call it the *seed value*), a series of numbers is generated. Successive numbers in the series are generated by repeatedly applying the following rule:

- if n is even, the next number is $n/2$
- if n is odd, the next number is $3n + 1$

The series terminates if the number 1 is reached. For example, here is the Collatz series for seed value 7:

7, 22, 11, 34, 17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, 1 (1)

The Collatz Conjecture states that for any seed value in \mathbb{Z}^+ , the series eventually reaches 1. The problem with this statement is that it *seems* to be true, but it is not yet *proven*.

Write a **function**, `collatz(n)`, that computes the length of the Collatz series for input integer n . Your program should then apply this function to the following seed values and output the length of their Collatz series: 4, 8, 12, 13, 21, 45, 22.

2. collatz_plot.py, 30 points

Modify your code from the previous problem so that it computes the length of the Collatz series, $\mathcal{L}(n)$, for seed values n between 2 and 10^4 . Your program should plot $\mathcal{L}(n)$ versus n , and it should make a histogram of the values of $\mathcal{L}(n)$. If your program takes more than 20-ish seconds to run, something is wrong.

3. births_1.py, 30 points

On our Sakai resources page, you'll find a few datasets. Obtain either of the birth datasets. You will attempt to determine whether there is a time-dependence to the American birthrate.¹ Write a program that reads in the file, and plots the births per

¹ This might sound strange, but keep in mind that there is a huge time-dependence to the birth rate in the animal kingdom.

day versus the day of the year for a specific year (you choose the year). Note that the file gives the births on each *date*, but you want to convert these dates to the day of the year (*i.e.*, Jan 1 is the first day, Jan 2 is the second day, ...). There are a few ways to do this conversion, but the simplest is to note that the days are in order in the file; thus, you should be able to count as you read the lines and simply increment the day number.

Your program should determine (and print) the total number of births for the year, the maximum number of births per day, and the day on which this maximum occurred.

Don't forget axis labels and a title!