L&S 39F: Data science and the mind (Fall 2005)

Homework 4

November 17 (Due *beginning* of session on Tue, Dec 1)

All coding should be completed in a Python Jupyter Notebook (or Python). Please type your solutions into a *single* PDF that includes both figures and written explanations, with your own Python code enclosed in Appendix. Email this PDF to the course instructor with the title "L&S39F HW4 (YOUR NAME)". You should turn in your individual solutions, although collaborative discussions are encouraged (acknowledge your collaborator(s) in the write-up). Late work receives a deduction of 1 point per delayed hour.

In this assignment, you will explore the phenomenon of mental rotation. Specifically, you will perform linear regression and bootstrapped resampling on the data from Shepard and Metzler (1971). To do so, you will first need to take the following steps:

1) Download *Code* under *Hw4* in the NB column from the bcourses syllabus.
2) Copy and paste the content of "hw4_starter_code.txt" into a new Python notebook. You will now be in a position to explore the following questions.

Q1. Running the starter code, you should obtain two variables: a is an array of 9 rotated angles (20, 40, ..., 160) measured in degrees, and t is an array of 9 corresponding average reaction times. Linearly regress t against a. Report the values of the slope (regression coefficient) and intercept (offset) that you have obtained. [1 point]

Q2. Given the estimated slope and intercept, predict what the reaction time would be when the angle of rotation is 180 degrees. The actual observed reaction time is ~4.6 seconds - is your prediction close to this value? [1 point]

Q3. Perform bootstrap on the 9 paired points in a and t from Q1 by resampling these data 10,000 times. Each resampled set should draw 9 data points with replacement from the original pool of (a,t) pairs. For each set, estimate and record the corresponding regression coefficient by repeating the linear regression you performed in Q1 on the resampled data. Visualize these 10,000 coefficient values in a histogram with annotated axis labels. [1 point] Report the average coefficient value from the bootstrapped samples. [1 point] Explain why this bootstrapped average deviates from the actual slope observed in Q1. [1 point]