Friday, 09/15
Activities to be completed before coming to class

PART 1

Submitting Assignments

If you check your CS account, each of you has now a new folder there: **dav** [~~cs234~~]. The folder contains subfolders named week01, week02, etc. The purpose of the folders is to upload in them work that was assigned during that week. Generally, it is recommended that you submit that work by the end of the successive week (Friday before class). For example, work assigned during Week 1 is expected by Friday of Week 2, and so on.

However, to be consistent with the course policy of learning at your own pace, you may submit work later (hopefully not in a way that prevents you from doing new work).

You can use Cyberduck, Fetch, WinSCP to transfer files from your laptop to the cs234 folder (in the corresponding subfolder).

Work assigned on Week 1 that is expected on Week 2

1. Create a first notebook using Python to calculate descriptive statistics. [Use Markdown to add titles and text notes, etc.]
2. Tidy data activity for food data: either typed text files (stored as PDFs) or scanned copies (or photos) of your handwritten notes. If they are not very legible, please type them.
3. Visualization notebook (with your two charts task at the end)
4. Pandas notebook annotated with your notes in Markdown (don't use comments in cells, they are not readable).

For notebooks submit both the IPYNB file and a HTML file which you can get through File | Download As | HTML.

Collaboration Policy: Feel free to create a study group and do these tasks together with your peers. However, everyone should work on their computer and only discuss issues in English, without sharing code. The best way to learn is to do the tasks yourself.

PART 2

Running Python Script to extract timestamps from photos

Some of you tried to run the Python script, but weren't successful. In many cases this had to do with the way you transferred files from your phone to the laptop. If you used Facebook or some other social media platform, they usually strip all the information about how and when the photo was taken to preserve a user's privacy. If you looked at the Python code, it's calling the method **`_getexif`**. EXIF stands for Exchangeable Image File Format (https://en.wikipedia.org/wiki/Exif). If EXIF data is removed (Facebook routinely does that), then our script produces None values for the photos.

My suggestion is to try one of the following:

1. Use an USB cable to transfer your files from your phone to your laptop.
2. Upload your photos to Google Drive or another Cloud service that doesn't remove EXIF.

There is no need to do this for all your photos at this point. **3-4 photos suffice**. We want to get a sense of how timestamps of photos from different students are formatted, so that we can write a script to automatically convert them to variables we are interested.

Run the script and send the output file via email to Eni, preferably by 11am on Friday.

PART 3 - Bayes

Review the class activity sheet about the Oreo Cookies and Bayes' Theorem. The problem was modified from Allen Downey, who has a blog post with more such problems, which I encourage you to challenge yourself to solve.
http://allendowney.blogspot.com/2011/10/all-your-bayes-are-belong-to-us.html

Once you've done the review, on a piece of paper (or you can type it), solve the Diagnosis problem at the end of the activity sheet. First define all events involved (test is positive, etc.), write the general Bayes formula using these events, and then perform the calculations and find the correct answer.

PART 4 - Food Data variables

We'll create a spreadsheet (we'll give you one with columns and data validation) that you'll fill with information from your food data collection. Below are the proposed variables and values for them. Because we want to aggregate data from all students, we need to make sure we'll be using the same variables and values for them. If you have suggestions for the values of the variables below (especially the time intervals for the meal names, or missing beverage names or locations), **please drop me an email before class.** Other variables are welcome too, as long as they aren't too unique.

MealName From 6-9: Breakfast From 9-12: Morning Snack From 12-3: Lunch From 3-6: Afternoon Snack From 6-9: Dinner From 9-12: Night Snack	**ContainsBeverage** True False **IsVegetarianMeal** True False	**BeverageHasCaffeine** True False **CaffeineAmount** ?? **EatingWithFriends** True False
BeverageType Water Tea Coffee Soda Milk ?	**Location** Dining Hall Res Hall Academic Building Off Campus ?	**ContainsVeggieOrFruit** True False

Recommended Reading related to the Food Survey: [10-15 minutes]
http://fivethirtyeight.com/features/you-cant-trust-what-you-read-about-nutrition/
Don't miss the short video at the end of the article!