# Week 7 in CS 234

# What is due on Tuesday, 10/17/17?

**Time allocation: budget for Tue and Fri about 3h/each to complete the tasks for that day.**

1. Every team needs to put the Wikipedia project description (one or two research questions) in this common Google Doc, before class. Most teams have already sent me an email and I've given them feedback on what to focus. The earlier you do this step, the sooner you can start working on the project.
2. The members of each team will collaboratively do the following tasks (decide among yourselves who will do what for Tue, then make sure to complete all of them by next Tuesday).
   a. Student A works on completing the Notebook Time, Datetime, Dateutil, which shows how to perform operations with date and time.
   b. Student B works on completing the Notebook Wikipedia Revisions.
   c. Student C works on identifying research papers for the group, puts them in a shared folder for the group and summarizes the paper that it's most relevant to their research questions. This summary needs to help the project: what were the research questions, how did the authors translate them into concrete quantities that could be measured; how did they collect the data; what transformations did they do; what techniques did they use; how did they analyze the results, etc.
3. Each student creates a subfolder "wikipedia" in their "cs234" subfolder of "public_html" folder in their CS account. In this subfolder, you create a blog.html page where you'll be "blogging" about your work in this project. For example, if you are Student A of your group, you'll write an entry that might look like this:
   **10/17/17**
   **Completed Datetime operations notebook**
   It took me about 1.5 hours to complete the notebook, since I had never heard of any of these modules. The biggest takeaways from the notebook are:
   - Something important one
   - Something important two

You should remember to write in this blog page, every time that you complete a task related to the project, or you work on something of your own initiative on the project. Finally, put a link to this blog page from your CS234 portfolio page, where you have your link to the eating habits page. The two other students will write similar summaries.

# What is due on Friday, 10/20/17?

1. The members of the team will rotate the tasks to complete one of the subtasks in Point 2 from Tuesday. In Task 2.c where student C has already compiled a list of papers, the task is to summarize another paper that might be relevant to what you're doing.
2. The team members work together to create their first Wikipedia project exploration notebook. The student(s) who did the Wikipedia Revision task should be easily able to produce a list of all edit timestamps for a popular page that their project has identified (probably as a CSV file). Working

together with the two other students, the team will create different histograms:

   a. Distribution of edits over the years (that the page has existed)
   b. Distribution of edits for the year with most edits (the months of the year)
   c. Distribution of edits for the month with most edits (the days of the month)

The datetime module will make it easy to group counts by years, months, days, etc.
A distribution like 2.c will allow you to see whether there are particular days that trigger edits
(external events to Wikipedia). Upload this notebook in your project01 folder in dav/drop (the
IPYNB file), but link the HTML page of the notebook from your blog page (and blog about it).

**Challenge yourself**

Teams who want to challenge themselves can decide to additionally create a heatmap (I've
created the one below with Plotly, you can substitute having edits instead of emails per day):

Heatmap: emails per day [2015-2016]
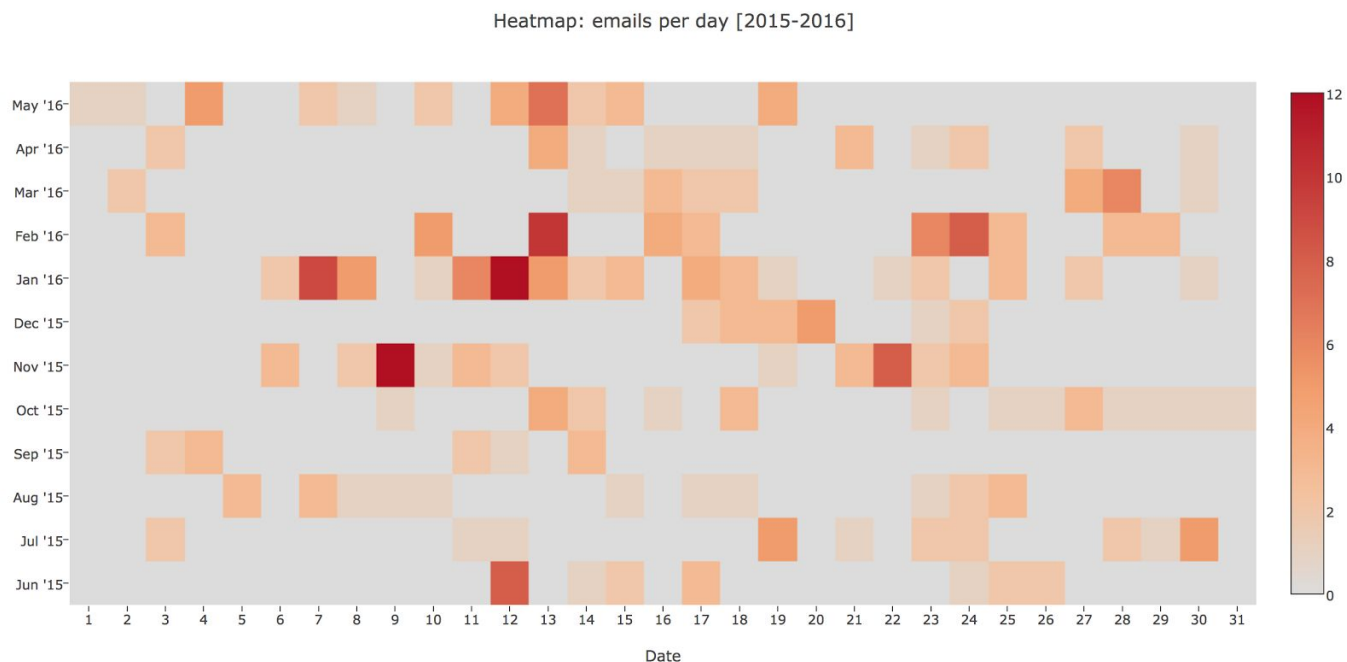


Thinking ahead of the communication of the results, you can think of a Dash app that has a
dropdown menu that lets one select a particular year and update the heatmap, or have two
heatmaps that are being compared, side by side, etc.

Another fun visualization challenge is the use of a Dash app with an adjustable histogram (Rami
& Sophie explained one in class) that lets you see the edits over all the time period, but also lets
you zoom in to a particular year or a particular month. The wind speed histogram they showed
cannot be adopted as it is, but you may think of pairing it with other widgets to achieve the effect
you want.

**Have fun working on the project!**