

DAT16 SF: HOMEWORK 2 ASSIGNMENT

Assigned: Wednesday, August 5, 2015

Due: Monday, August 10, 2015, before class

Review Due: Wednesday, August 12, 2015, before class

The purpose of this homework is to review what we've learned about data sources, APIs, Pandas and Bokeh.

HOMEWORK QUESTIONS

DUE MONDAY:

1. Choose a public API from this list: <http://www.pythonapi.com/>, or use an API from Kimonolabs.
2. Define and state a question that you would like to answer using that data (a simple question)
3. Download a small dataset from the chosen API
4. Import the dataset to a Pandas data-frame in an Ipython notebook.
5. Explore the data:
 - a. Are there any missing data or NULL values? How could they be imputed?
 - b. How many features are there? Are they normalized?
6. Visualize the data with Bokeh to help answer the question
7. Export the visualization to a .png file
8. Upload the downloaded data, the Ipython notebook and the figure to a folder in homework in your personal repo

BONUS POINTS:

Choose a dataset from any of the lists on slide 33 of lecture 2 and repeat steps 2-6 above.

DUE WEDNESDAY:

1. Go to your new assigned review-buddy's repo
2. Read through your buddy's ipython notebook and make sure you understand what he/she is doing.
3. Open an issue in his/her repo and write comments on the things you don't understand and on the things you like in his/her code.
4. Quote the instructors in the comments so that we get notified about the open issue