/hubwayPredict
Code Review Reflection and Synthesis

Feedback and Decisions

The primary feedback we received was how to structure our code. It was noted that we should store our data in memory, rather than parse through a text file every time we wanted to obtain the data. Pandas was suggested as a module that could help. Also, we were reminded that this project should have some object oriented elements to it, and it was suggested to us that we can represent our data with objects. We got a relatively late start on the project since we were not able to obtain data from the MBTA. We felt, and the code review participants agreed, that we need to start working in earnest on learning how to incorporate machine learning tools and actually implementing it.

We plan on looking into Pandas in the future, after we have our machine learning elements working. We plan on writing the code to represent our data as objects very soon, and we have already started reviewing machine learning concepts and have begun implementing them. We plan on researching machine learning tools and techniques that people have used for the Capital Bikeshare kaggle competition project, as it is very similar to what we are doing. We have decided that we would prefer not to simply use a black box method to predict Hubway use, but want to instead work on developing a more custom model. We may use some tools from Sci Kit Learn, but we plan on researching and understanding thoroughly any tool we use before implementing it.

We have already started narrowing down on what tools we want to focus on following the decisions we received in the code review on the machine learning part of the project. In particular, we are focusing on supervised machine learning techniques, as unsupervised learning appears to be the wrong type of machine learning to implement for several reasons: we know exactly what data we are feeding into the algorithm and unsupervised machine learning seems too difficult to understand and correctly implement in the time allotted for the project.

Review Process Reflection

We felt that the design review went well. We received a lot of constructive feedback that we can now act on. Overall, it would have been nice to get more feedback on the machine learning element of our project, though our audience didn't really know enough about the topic to give us enough detailed feedback in that area. One thing that we could have done to improve the design review is to provide our peers with pre-reading material on some of the machine learning techniques we were thinking of using. This would help them know more about certain machine learning techniques and put them in a position where they could actually provide feedback.

We did not adhere to our agenda as closely as we had in the previous design review. We fielded more clarification questions throughout and in turn, we solicited feedback throughout the presentation, instead of explaining all of our code and answering questions and getting feedback afterwards.

The structure of our design review was not as effective as we hoped it to be. We wanted more feedback about the machine learning decisions that we had to make, but our audience latched onto data processing. We probably should have front loaded our code review with machine learning questions (and provided pre-reading) in order to get more feedback about machine learning.