Module 4

This week we will be looking at Enterococcus levels in the Hudson River, using data from the organization Riverkeeper (http://www.riverkeeper.org/).

Background: Enterococcus is a fecal indicating bacteria that lives in the intestines of humans and other warm-blooded animals. Enterococcus (“ Entero”) counts are useful as a water quality indicator due to their abundance in human sewage, correlation with many human pathogens and low abundance in sewage free environments. The United States Environmental Protection Agency (EPA) reports Entero counts as colonies (or cells) per 100 ml of water.

Riverkeeper has based its assessment of acceptable water quality on the 2012 Federal Recreational Water Quality Criteria from the US EPA. Unacceptable water is based on an illness rate of 32 per 1000 swimmers.

The federal standard for unacceptable water quality is a single sample value of greater than 110 Enterococcus/100 mL, or five or more samples with a geometric mean (a weighted average) greater than 30 Enterococcus/100 mL.

Data: I have provided the data on our github page, in the folder https://github.com/charleyferrari/CUNY_DATA608/tree/master/lecture4/Data. I have not cleaned it – you need to do so.

Each question should be a separate dash app. A single app.py for each will be sufficient.

Questions 1:

You’re a civic hacker and kayak enthusiast who just came across this dataset. You’d like to create an app that recommends launch sites to users. Ideally an app like this will use live data to give current recommendations, but you’re still in the testing phase. Create a prototype that allows a user to pick a date, and will give its recommendations for that particular date.

Think about your recommendations . You’re given federal guidelines above, but you may still need to make some assumptions about which sites to recommend. Consider your audience. Users will appreciate some information explaining why a particular site is flagged as unsafe, but they’re not scientists.

Question 2:

This time you are building an app for scientists. You’re a public health researcher analyzing this data. You would like to know if there’s a relationship between the amount of rain and water quality. Create an exploratory app that allows other researchers to pick different sites and compare this relationship.