CS 4630/5630 Kresman Homework 6

Hydrology and Data Science: Watershed hydrologists and environmental scientists are vitally dependent on high-quality, real-world datasets that are generated through automated field sensors and harvested at national repositories.

**Problem**: Gather time series data from the US Geological Survey site NWISUV:**mySite** and do simple analyses. Use SOAP/SUD/DataFrame methods.

The from/to time periods and mySite are input, all one line, for example: 3/5/2019 4/5/2020 04193500. Create a data frame with a DateTime index and store data for the following (each variable’s data as a series). As well, add an extra series for water temperature in Fahrenheit.

00010: water temperature (Celsius)

00021: air temperature (degrees Fahrenheit)

00035: Wind speed (miles per hour)

00060: Discharge (cubic feet per second)

Print three tables (nice format):

1. For each series, the daily averages (since many values may be reported for the same day) - print just the tail() so the output is not too long.
2. For each series (across the entire time span), min, max, mean, and standard deviation
3. Rows for which any cell value is missing or 0 - just the tail() so the output is not too long

Plot all variables (except water temperature in Fahrenheit) in the same graph in a pretty format with nice legend. Use 100 cubic feet/sec as the unit for discharge, else graph may look squashed. The title of the graph is the name of the site (do not hard code the name).

**Notes**

* Must use DataFrame methods to the extent possible!
* Consider reviewing Units 10, 12, 14-15, 21
* If R or MATLAB is in your comfort zone, feel free to use it instead of Python😊 If so, document run, platform & OS instructions please!

**Canvas** submissions (do not zip): lastnameHw6.ipynb (or MATLAB/R code), a word doc of run snapshot