**Part 1:**

* ~~We will need to load several csv files into Python. They contain tabulated data and we will need to run queries on them, search for specific data, sort them, etc. Do you know what kind of data structure is the best choice? Can you figure out what is the best kind of data structure to use? While these operations are very easy to do in MATLAB or R, we need to find the most efficient way of doing so in Python. Ideally, we want to avoid "for loops" and use vectorization methods. Examples of tasks we'll have to perform are:~~
  + ~~ex1: find all the rows in the table for which the ID column is a certain value.~~
  + ~~ex2: add the first two columns together, and put the results in a new column.~~
  + ~~ex3: find all the rows in which the X column is larger than the Y column.~~
* ~~Given the latitudes and longitudes of many points, plot all the points on a map~~
* ~~Is there any way to project (as in project data onto a map) the latitude and longitude using Lambert Conformal Projection (https://matplotlib.org/basemap/users/lcc.html)? Also, how can we plot county and state borders for a certain state?~~

**Part 2:**

* ~~Plot the location of all wells in the state of Pennsylvania that are drilled in the Marcellus formation. Those wells have the word "marcellus" somewhere in the "formation" column.~~
* ~~In the background, plot the boundaries of Pennsylvania and it's neighboring states. Also, plot all the counties in Pennsylvania.~~
  + ~~since we need this background very often, it might be a good idea to create a function that attaches such a background with all the boundaries, labels, etc. to whatever plot we want.~~
* ~~Create a function that takes a well's id as input and plots its gas production as a function of time (see if you can make it work with multiple entries; if multiple id's are given to the function, it should plot several curves on the top of one another with proper legend)~~
  + Also, add an option to the function which allows the user to choose months\_from\_first instead of date for the X axis values.
* Create a function that takes a list of counties and generates a bar plot for the number of new wells that where drilled in those counties (combined) with time (this is like a histogram). You can use one of existing functions in python that takes a list of values and a number of bins and outputs the number of observations in each bins. The equivalent I used in MATLAB was [counts,centers] = hist(x, NumberOfBins). Then I use bar(centers, counts) to generate the bar plot. There should be options that automatically handle date types. Let me know if you need more clarification on this item.
  + if an empty list of counties are given to the function, it should generate the histogram for the entire state of Pennsylvania.
* ~~Write a function that takes a well's id and fits the equation that I gave you earlier and returns the fitting parameters. The equation is called Exponential Power Law Decline for future references.~~