**Assignment #2 Using Pandas to processing big data (10 points)**

Big data set: choose a data set in .csv format (pickle or json file also okay). You are encouraged to use the dataset that you discovered in Assignment #1. Data set should be large (say at least 1MB?)

Analytics steps:

1. Read in the data set.
2. Display the first 50 data entries/rows as well as last 50 entries/rows.
3. Display a quick statistical information on all numerical columns such as count, mean, std, min, max, etc.
4. Select a subset of rows (you decide which subset to select or which criteria to use for selection.) Display the first 10 data entries selected.
5. Similar to 4, but select a subset of columns (from original data). Display the first 10 data entries with selected columns.
6. From original data, filter out some data, for example, filter out those salary lower than certain amount. After filtering out the data, display the first 50 data entries.
7. From original data, find out all entries with missing values. Display the first 10 entries.
8. Manipulate the original data by changing numerical values of specific column(s) (for example, give everyone 10% raise!) Display the first 10 entries before update and after update.
9. Sort the data set resulted from step 8 in certain way (e.g. descending order of salaries)
10. Group the data set from step 9 based on certain category (e.g. group based on rank of Professor, Assoc Professor, etc.)
11. (optional) Plot data (or subset of data) in at least three different ways such as vertical bar graph, horizontal bar graph, curve, …

Submission:

1. Save all program source code in a .py file
2. Include data set or a link to dataset (use a link if the dataset is 1MB or larger.)
3. A powerpoint file that briefly describes the data set followed by step-by-step analysis results (for each step, code followed by output in screen shots or other image format)
4. You may upload the above three items (source code, data set or link to data set, and powerpoint file) to blackboard in a zipped file or a link to your github project or host it on your webpage. (Note: if you host your project on github or via your webpage you only need to submit a link to your project on blackboard under Assignment #2.)