Assignment 9

Submission Instructions

1. You are free to use whichever development environment you wish to create and submit the assignment answers.
2. Create a directory using your Net ID as the directory name.
3. Place your Python code in this directory. *Make sure there are no files outside this directory that will be committed to the repository.*
4. There should be at least one file called `assignment9.py` at the top level that contains the main program that will generate your answers.
5. Fork the `assignment9` repository from the ds-gs-1007 user on GitHub
6. Clone this repository onto your local system.
7. Place your directory into the working directory of this repository either using PyDev or manually.
8. Add your directory to the staging area, commit, and push to the remote repository.
9. Submit a pull request to the repository owner (ds-ga-1007).

Questions

1. Load the `countries.csv` file (located in the `assignment9` repository) into a pandas `DataFrame` and name this data set `countries`.

   Download the "Income per person (GDP/capita, PPP$ inflation-adjusted)" data set from http://www.gapminder.org/data and load this into a `DataFrame` called `income`.

   Hint: You can convert the data set to CSV, or consider using the pandas function `pandas.read_excel()` to read in the .xlsx file directly

2. Transform the data set to have years as the rows and countries as the columns. Show the head of this data set when it is loaded.

3. Graphically display the distribution of income per person across all countries in the world for any given year (e.g. 2000). Choose the best kind of plot to display this information.

4. Write a function called `merge_by_year(year)` to merge the `countries` and `income` data sets for any given year. The result should be a `DataFrame` with three columns titled `'Country'`, `'Region'`, and `'Income'`.

5. Use exploratory data analysis tools such as histograms and boxplots to explore the distribution of the income per person by region data set from question 4 for a given year. Describe how these change through the recent years.