Introduction to Data Science – Fall 2016

Lab Assignment 20

Visit the Data.gov website (https://catalog.data.gov/dataset?res_format=CSV) and download the CSV data file `usersshareddsfratebrthsyaw1819raceethncty20002012.csv` from "My Brother's Keeper Key Statistical Indicators on Boys and Men of Color" containing datasets with key indicator - cross tabulated for race and gender.

1. Read the data file and explore the dataset using shape and `head()` functionality
2. Use the visualization methods provided by Pandas and matplotlib to plot the histogram for *Rate of birth to women ages 18-19.*
3. Show the boxplot depicting "Rate of birth to women ages 20-24"
4. Use Side-by-side boxplot to visualize the "`Rate of birth to women ages 18-19`" by `Year`.
5. Use barplots to illustrate the "`Rate of birth to women ages 18-19`" distribution within the different ethnic groups captured by the dataset.

For the next two problems use the `mtcars` dataset provided in blackboard/Course Documents.

6. Explore the dataset using shape and `head()` functionality
7. Depict the dependencies between `mpg` and wt by plotting the corresponding scatterplot.
8. Use the following dictionary to create a Datframe

   md = {"radius": [1.0, 2.0, 3.0, 4.0, 5.0, 6.0],
   "area" : [3.14159, 12.56636, 28.27431, 50.26544, 78.53975, 113.09724]}
   With columns Radius and Area correspondingly

   Plot the line plot displaying the relationship between the area and the radius.

What to turn in: Download your Jupyter notebook in HTML format and upload the file Lab20.html to blackboard. This HTML file should contain the following information:

- All code used and the result of the corresponding computation
- Text explanations for the code, as necessary.