**Assignment 1 [Submission Date: 4th April 2024 ]**

**Introduction**

This assignment uses data from the [UC Irvine Machine Learning Repository](http://archive.ics.uci.edu/ml/), a popular repository for machine learning datasets. In particular, we will be using the "Individual household electric power consumption Data Set" which I have made available on the course web site:

* **Dataset**: [Electric power consumption](https://d396qusza40orc.cloudfront.net/exdata%2Fdata%2Fhousehold_power_consumption.zip) [20Mb]
* **Description**: Measurements of electric power consumption in one household with a one-minute sampling rate over a period of almost 4 years. Different electrical quantities and some sub-metering values are available.

The following descriptions of the 9 variables in the dataset are taken from the [UCI web site](https://archive.ics.uci.edu/ml/datasets/Individual+household+electric+power+consumption):

1. **Date**: Date in format dd/mm/yyyy
2. **Time**: time in format hh:mm:ss
3. **Global\_active\_power**: household global minute-averaged active power (in kilowatt)
4. **Global\_reactive\_power**: household global minute-averaged reactive power (in kilowatt)
5. **Voltage**: minute-averaged voltage (in volt)
6. **Global\_intensity**: household global minute-averaged current intensity (in ampere)
7. **Sub\_metering\_1**: energy sub-metering No. 1 (in watt-hour of active energy). It corresponds to the kitchen, containing mainly a dishwasher, an oven and a microwave (hot plates are not electric but gas powered).
8. **Sub\_metering\_2**: energy sub-metering No. 2 (in watt-hour of active energy). It corresponds to the laundry room, containing a washing-machine, a tumble-drier, a refrigerator and a light.
9. **Sub\_metering\_3**: energy sub-metering No. 3 (in watt-hour of active energy). It corresponds to an electric water-heater and an air-conditioner.

**Loading the data**

When loading the dataset into PANDAS , please consider the following:

* The dataset has 2,075,259 rows and 9 columns. First calculate a rough estimate of how much memory the dataset will require in memory before reading into a DataFrame. Make sure your computer has enough memory (most modern computers should be fine).
* We will only be using data from the dates 2007-02-01 and 2007-02-02. One alternative is to read the data from just those dates rather than reading in the entire dataset and subsetting to those dates.
* Note that in this dataset missing values are coded as ?.

**Making Plots**

Our overall goal here is simply to examine how household energy usage varies over a 2-day period in February, 2007. Your task is to reconstruct the following plots below, all of which were constructed using the base plotting system.

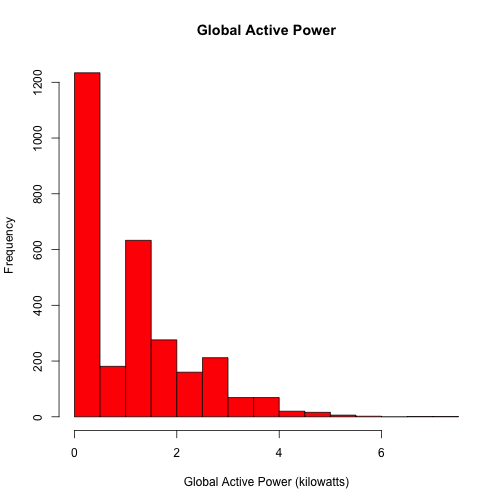
For each plot you should

* Construct the plot and save it to a PNG file with a width of 480 pixels and a height of 480 pixels.
* Name each of the plot files as plot1.png, plot2.png, etc.
* Create a separate Python code file (plot1.py, plot2.py, etc.) that constructs the corresponding plot, i.e. code in plot1.py constructs the plot1.png plot. Your code file **should include code for reading the data** so that the plot can be fully reproduced. You should also include the code that creates the PNG file.
* Add the PNG file and R code file to your git repository

When you are finished with the assignment, push your git repository to GitHub so that the GitHub version of your repository is up to date. There should be four PNG files and four Python code files. In the submission report include the URL of your git repository.

The four plots that you will need to construct are shown below.

**Plot 1**

[](https://github.com/rdpeng/ExData_Plotting1/blob/master/figure/unnamed-chunk-2.png)

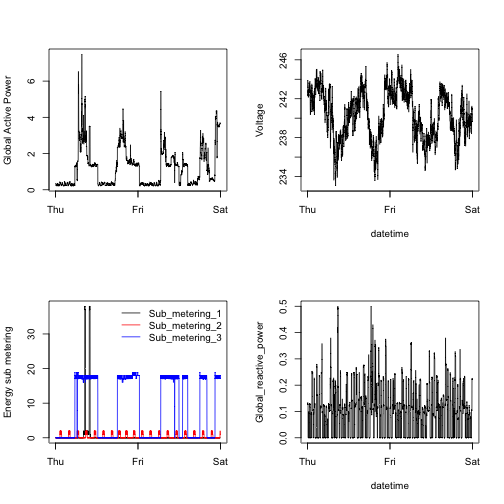
**Plot 2**

[](https://github.com/rdpeng/ExData_Plotting1/blob/master/figure/unnamed-chunk-3.png)

**Plot 3**

[](https://github.com/rdpeng/ExData_Plotting1/blob/master/figure/unnamed-chunk-4.png)

**Plot 4**

[](https://github.com/rdpeng/ExData_Plotting1/blob/master/figure/unnamed-chunk-5.png)