The "Data Science" Specialization   Learn More (https://www.coursera.org/specialization/jhudatascience/1?utm_source=spark&utm_medium=banner)   ✕
(http://help.coursera.org/customer/widget/emails/new)

Peer Assessments (https://class.coursera.org/exdata-003/human_grading/) / Course Project 1
Help (https://class.coursera.org/exdata-003/help/peergrading?url=https%3A%2F%2Fclass.coursera.org%2Fexdata-003%2Fhuman_grading%2Fview%2Fcourses%2F972140%2Fassessments%2F3%2Fsubmissions)

---

due in 5day 14h

Submission Phase

1. Do assignment ☐ (/exdata-003/human_grading/view/courses/972140/assessments/3/submissions)

Evaluation Phase

2. Evaluate peers 🔒 (/exdata-003/human_grading/view/courses/972140/assessments/3/peerGradingSets)

Results Phase

3. See results 🔒 (/exdata-003/human_grading/view/courses/972140/assessments/3/results/mine)

☐ In accordance with the Honor Code, I certify that my answers here are my own work, and that I have appropriately acknowledged all external sources (if any) that were used in this work.

Save draft   Submit for grading

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 R, 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 R. 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.

- You may find it useful to convert the Date and Time variables to Date/Time classes in R using the `strptime()` and `as.Date()` functions.

- 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.

First you will need to fork and clone the following GitHub repository: [https://github.com/rdpeng/ExData_Plotting1 (https://github.com/rdpeng/ExData_Plotting1)](https://github.com/rdpeng/ExData_Plotting1)
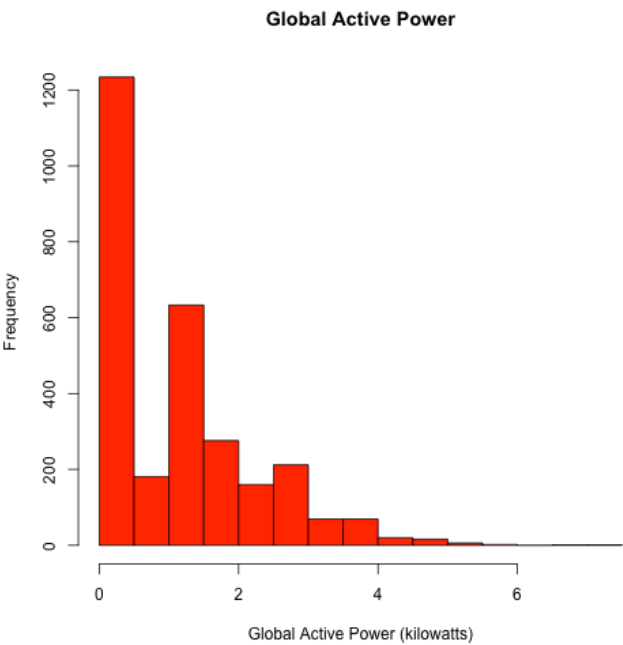
For each plot you should

- 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 R code file (`plot1.R` , `plot2.R` , etc.) that constructs the corresponding plot, i.e. code in `plot1.R` 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.

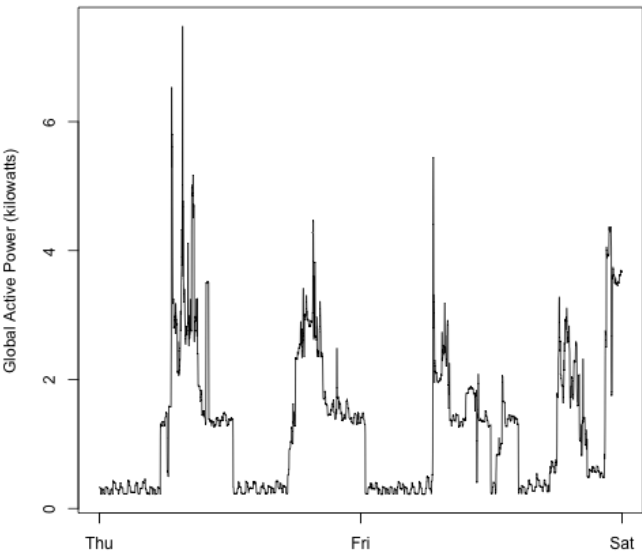- Add the PNG file and R code file to your git repository

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 R code files.

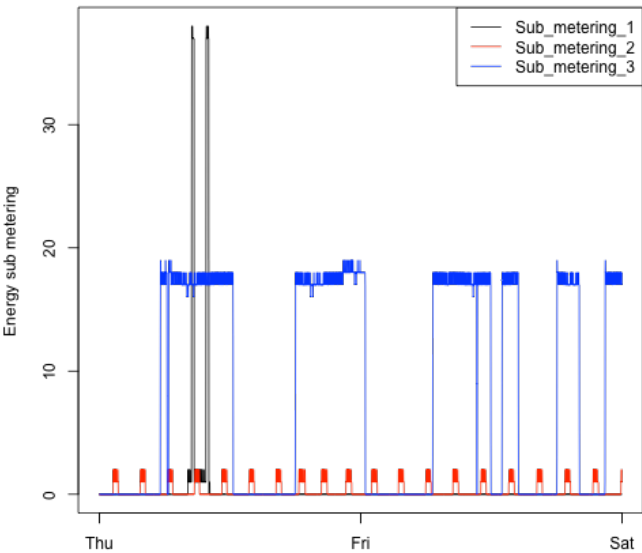The four plots that you will need to construct are shown below.

Plot 1

Global Active Power

Global Active Power (kilowatts)

Plot 2

Plot 3

Plot 4