## Introduction

Fine particulate matter (PM2.5) is an ambient air pollutant for which there is strong evidence that it is harmful to human health. In the United States, the Environmental Protection Agency (EPA) is tasked with setting national ambient air quality standards for fine PM and for tracking the emissions of this pollutant into the atmosphere. Approximatly every 3 years, the EPA releases its database on emissions of PM2.5. This database is known as the National Emissions Inventory (NEI). You can read more information about the NEI at the EPA National Emissions Inventory web site.

For each year and for each type of PM source, the NEI records how many tons of PM2.5 were emitted from that source over the course of the entire year. The data that you will use for this assignment are for 1999, 2002, 2005, and 2008.

## Data

The data for this assignment are available from the course web site as a single zip file: \* Dataset: Data for Peer Assessment [29Mb]

The zip file contains two files:

PM2.5 Emissions Data (summarySCC\_PM25.rds): This file contains a data frame with all of the PM2.5 emissions data for 1999, 2002, 2005, and 2008. For each year, the table contains number of tons of PM2.5 emitted from a specific type of source for the entire year. Here are the first few rows.

The following descriptions of the 9 variables in the dataset are taken from the UCI web site:

Date: Date in format dd/mm/yyyy

Time: time in format hh:mm:ss

Global\_active\_power: household global minute-averaged active power (in kilowatt)

Global\_reactive\_power: household global minute-averaged reactive power (in kilowatt)

Voltage: minute-averaged voltage (in volt)

Global\_intensity: household global minute-averaged current intensity (in ampere)

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).

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.

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)

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

plot of chunk unnamed-chunk-2

### Plot 2

plot of chunk unnamed-chunk-3

### Plot 3

plot of chunk unnamed-chunk-4

### Plot 4

plot of chunk unnamed-chunk-5