Assignment 6

In this assignment, you will analyze time series data. For the time being, the analysis will be in the time domain.

The data is contained in the file timeseries2015.xlsx on the website. It contains data from a quantitative diary of an individual suffering from bipolar disorder, from each day in 2015. There are 8 columns, which represent – in order: The date, the number of hours slept that day, daily mood on a scale from -2.5 to 2.5, energy levels on a scale from 0 to 2, inspiration (the individual is an artist) on a scale from 0 to 2, number of hours worked that day, minutes of REM sleep that night and minutes of deep sleep that night. All tasks below are to be executed within a single script. This is the .m file that you should upload. You can assume that the people reviewing your code (peer reviewers and the TA) will run the script from the same folder that contains the data file.

Specification of the tasks that the script needs to be able to do (in bold):

1) **Load the data into the workspace.** Note that the raw data file in an excel file.

2) **Deal with missing data**. Some of the data will be missing, in particular the data on sleep quality, as the electrodes didn't work every night. This can be done by *data pruning* (removal of missing data) or *imputation* (guessing what the missing value should have been), the latter is usually done by interpolation from neighboring points. This is your call, but for time series data, pruning is risky and you can use the information from temporal autocorrelations to inform the imputation. Basic: Prune. Ambitious: Impute.

3) Basic: Make a **figure** with 2x3 subplots of all time courses (sleep, mood, energy, inspiration, work, REM and Deep), smoothed with a 7 day kernel. Combine sleep and work in one plot, as they are on the same scale and have a similar magnitude. Ambitious: Add shaded error bands that represent the moving standard error (same 7 day kernel).

4) Basic: **Correlate** all variables with all others. What is the highest positive (and negative!) correlation between any two variables? Is REM sleep associated with good or bad mood? How about deep sleep? Ambitious: Create a multiple linear regression model – how well can you predict mood from the other variables (what are the beta weights and what is R^2?)

5) Basic: Make a single plot where you plot all 7 smoothed – again, use a 7 day kernel – but **normalized** (z-scored) time courses. Does what you see mesh with the correlations you observed in 4)? Ambitious: Are the correlations between the variables even stable over time? Make a plot where you superimpose the time course of the *correlations* between the variables. I recommend using a longer kernel – e.g. 28 days, for this.

6) Basic: Calculate average sleep, mood and work values per weekday (Monday to Sunday) and make figures that show bar graphs for these. Ambitious: Is there such a thing as a Monday mood? Friday mood? Do suitable significance tests to substantiate or refute this notion.