STAT 520, Fall 2015: Homework 4

1. Generate three time series data sets, each of length $n = 200$, including (i) an AR(1) with $\phi = -0.6$, (ii) an MA(1) with $\theta = 0.8$, and (iii) an ARMA(1,1) with $\phi = -0.6$ and $\theta = 0.8$. For each one,

   (a) plot the observed time series.

   (b) plot the sample ACF, the sample PACF, and the sample EACF.

   (c) use the `armasubsets` function in R to identify the best model in terms of the BIC.

   Do the plots in part (b) agree with what you know to be true? Remember, you know the correct models! That is, you are assessing here whether the sample identification functions agree with the truth. Does the BIC identify the correct model as the "best" model in each case? If not, where is the correct model ranked, if at all?
   Curiosity: What happens when we increase the sample size? Repeat everything above when $n = 1000$. Do your findings change? Comment. You don't have include a bunch of new figures in your homework; just write a detailed summary describing your findings.

2. I have put four data sets on the course web site:
   - `ibm`: daily closing IBM stock prices (dates not given)
   - `thermal`: daily temperatures data used in midterm
   - `gasprices`: average price (US dollars per gallon) for regular gasoline in the United States; there are $n = 145$ weekly observations collected from 1/5/2009 to 10/10/2011 (**Source:** Rajon Coles, Fall 2011).
   - `supremecourt`: data used in Example 1.15 on p. 16 of the notes.

   All data sets can be read in, e.g.

   ```
   d1=ts(read.table("http://people.stat.sc.edu/hansont/stat520/ibm.txt"),start=1)
   ```

   Using the methods from Chapter 6, identify a small set of candidate ARIMA(p, d, q) models for each data set. You may need to transform the data before considering differencing. There may be a single model that emerges as a "clear favorite," or there may not. For guidance, use the summary described in Section 6.7 (notes) and follow it exactly. For each data set, write up detailed notes that describe how you decided on the model(s) you did. Your summary should convince me that your model(s) is (are) worthy of further consideration. Note that we are *not* considering modeling trends here, but rather differencing the original (possibly transformed) series only, to remove nonstationarity.

Note: This problem is important because your class project will involve you finding data of your own, specifying a model (or models), fitting the model(s), and diagnosing model fit. This problem will help you in the model specification phase of your project.

Important!! Remember your selected candidate model(s) for each data set. You will consider these models further in Homework 5.

3. Provide a description of your final project data, including a time series plot.