Time Series HW 1

Due: September 2, 2016 at 9 am

## HW 1

This will not be written in report style - just provide numbered answers to each question. Include your edited R code either in-line (using something like R-markdown - the newer option to compile directly to a Word document is quite nice for controlling formatting after compiling) or as an appendix to your homework.

You will get a 5% bonus on this homework if you do this homework in a group of 2 or more, but remember that you must understand everything that is done on the assignment and are responsible for all answers. If you discuss the assignment with others but turn in separete assignments, you need to document any discussions you had and how it impacted your answers - treat this like citing your sources.

For a preliminary check and to introduce some of the statistical challenges we will confront, I want you to analyze the mean monthly maximum daily temperatures in Bozeman. You can access the original file that I downloaded from NOAA's National Centers for Environmental Information (<http://www.ncdc.noaa.gov/cdo-web/>) here that has temperatures in degrees F in the column labeled MMXT (<https://dl.dropboxusercontent.com/u/77307195/rawbozemandata.csv>) for Bozeman's MSU weather station. If the code below does not work on your computer, try removing the "s" from https or just copying the link and pasting it into a browser bar. If none of that works, email me and I'll send you the csv file directly.

Note that you are getting monthly data values and that the DATE column contains year and month information concatinated together. Along with this issue, the data may not be perfect and you should check all steps carefully, not making any assumptions about what you are getting...

1. Read in the data set and use R to make a correct date code that separates year and month. There are many ways to do this. If you can't figure out how to do this using functions in R, you can do this outside R (say in Excel) or by some sort of hand coding of the date information but will get a small deduction in points for bypassing the challenge of doing this in an efficient way in R.
2. Plot the monthly mean maximum temperatures (y-axis) vs year (x-axis), labelling the axes with the name and units of each variable.
3. Create a variable that is just the year of each observation and another for the month. Then fit a linear model with temperature as the response and year and month as explanatory variables treated correctly as either quantitative or categorical predictors. Do not consider any higher order model terms such as polynomials or interactions. For many reasons but especially for the following question, do any variable manipulations prior to fitting the model and use the general code format for your lm of:

model1<-lm(y~x1+x2,data=mydatasetname).

1. Install and load the effects package and run the following code to get effects (also better called termplots) of the model that you fit: plot(allEffects(model1)). Discuss the month effect plot in general.
2. For the "year" model component, interpret the estimated slope coefficient and report a 95% confidence interval. Also note the size of the estimated change in the mean temperature over the entire length of the data set and report and confidence interval for that result.
3. Generate a test for the month model component, write out the hypotheses, report the results (extract any pertinent numerical results from output), and write a conclusion based on these results.
4. Run the following code:

par(mfrow=c(2,2))

plot(model1)

It should produce four panels with residuals vs fitted, normall QQ, scale-location, and residuals vs leverage plots. Only discuss the normall QQ plot. What model assumptions does this help us assess and what does it suggest here?

rawbozemandata<-read.csv("https://dl.dropboxusercontent.com/u/77307195/rawbozemandata.csv",header=T)  
View(rawbozemandata)  
head(rawbozemandata)

## STATION STATION\_NAME DATE MMXT  
## 1 COOP:241044 BOZEMAN MONTANA STATE UNIVERSITY MT US 190001 37.6  
## 2 COOP:241044 BOZEMAN MONTANA STATE UNIVERSITY MT US 190002 29.9  
## 3 COOP:241044 BOZEMAN MONTANA STATE UNIVERSITY MT US 190003 47.7  
## 4 COOP:241044 BOZEMAN MONTANA STATE UNIVERSITY MT US 190004 52.7  
## 5 COOP:241044 BOZEMAN MONTANA STATE UNIVERSITY MT US 190005 66.6  
## 6 COOP:241044 BOZEMAN MONTANA STATE UNIVERSITY MT US 190006 79.1