For live session we are going to discuss the general idea of dealing with correlated residuals.  Attached is an excel file that includes Melanoma and Sunspot data over time.  A quick search on google for sunspots may be helpful for a reference. The melanoma variable is the rate of melanoma occurrences.

We are going to use PROC Autoreg here, but I will discuss some additional models that can be done using PROC ARIMA so be aware of this that other procs exist to deal with more modeling scenarios.

Melanoma

1. Plot Melanoma versus Years
2. Take a look back at what it means for a time series to be stationary. Does the melanoma time series look stationary to you?
3. The first model below runs a regular linear regression of melanoma vs time without any timeseries modeling. Its just a regular regression run.  Use the diagnostic graphs (ACF and PACF plots) to assess if there is any evidence of autocorrelation (aka correlated errors).  Be prepared to discuss your basic understanding of what the graph is telling you and what you wished it looked like to have independent data.
4. The second model runs a regression on time but now with an autoregressive process assumed.
   1. Check the residual diagnostic ACF and PACF plot. What do you make of it compared to #1?
   2. Check the regression coefficient and standard error on the “time” predictor and compare it to what is reported in #1.

Sunspot

1. The sunspot data has a cyclical behavior. What we are going to do here is explore how an Autoregressive model can actually capture the cyclical behavior without any covariates present.
   1. Plot Sunspot versus Years
   2. Using the code for melanoma as an example, fit a simple regression model to Sunspot with just an intercept (model sunspot= / nlag= in SAS) Comment on the ACF and PCF plots
   3. Fit an AR(1), AR(2), AR(3), and AR(4) model by specifying the nlag option to 1,2,3, or 4.
      1. Examine and compare the ACF and PACF plots
      2. Locate the AIC statistic for each of the models and compare them
   4. Try to forecast the next 10-20 years using the model that has the lowest AIC from part c. Once you have the predictions, try to add them to your plot from part A so we can see what is going on. If you are stumped on how to predict future values of the time series, check out the Output and Predicted statement and options within Proc Autoreg or some of the examples:

<http://support.sas.com/documentation/cdl/en/etsug/63939/HTML/default/viewer.htm#etsug_autoreg_sect042.htm>

 SAS code below to get you started.

 data Melanoma;

infile 'C:\Users\e80100\Desktop\Melanomatimeseries.csv' dlm=',' firstobs=2;

input Year Melanoma Sunspot;

run;

proc autoreg data=Melanoma all plots(unpack);

model Melanoma=year;

run;quit;

                                                                                                                                                                                                                                             proc autoreg data=Melanoma all plots(unpack);

model Melanoma=Year  / nlag=1;

run;quit;