STA371G Homework Assignment 10

(40 Points. Due in class. Group homework.) Please write down the NAME and EID of each group member. Each group consists of up to three members. Each member needs to submit his/her own report

Problem 1: Wine Sales (I) (20 points)

Read the case "Northern Napa Valley Winery, Inc." in the course packet. The data file is available in the course website. The file contains the monthly wine sales for the Northern Napa Valley Winery for the period January, 1988 through August, 1996.

The goal of this case is to provide monthly forecasts for wine sales in the next twelve months, i.e. September, 1996 through August, 1997, and to combine the monthly forecasts to provide a forecast of annual sales for these twelve months.

(a) Using the appropriate time series plots, decide whether sales or the percentage change in sales, denoted PctChange, should be analyzed. Construct a regression model on PctChange that accounts for the trend and seasonality in the data.

(b) Are the residuals from this model independent? To test if the residuals are independent, you may run the following regression:

$$Residual_t = \beta_0 + \beta_1 Residual_{t-1} + \epsilon_t.$$

If you reject the null hypothesis that $\beta_1 = 0$, then consider modifying the model. (Hint: construct a trend+seasonality+AR(1) model)

(c) Using the final model obtained above, provide numerical forecasts of wine sales in September, 1996 and October, 1996.

(d) Explain briefly how you would obtain forecasts of wine sales for the months November, 1996 through August, 1997. You do not have to do the actual calculations for these forecasts because they are a bit tedious to do by hand. You just need to explain briefly how you would do it.

(e) Given the forecasts for the months September, 1996 through August, 1997, explain briefly how you could get a forecast of annual sales for the year encompassing these twelve months. You do not have to do the actual calculations, just provide a brief explanation of what you would do if the twelve monthly forecasts were available.

(f) Provide a 68% plug-in prediction interval for the forecast of sales in September, 1996.

(g) Explain briefly why it is difficult to get prediction intervals for the monthly forecasts of wine sales in October, 1996 through August, 1997.

Sales

Percentage Change in Sales
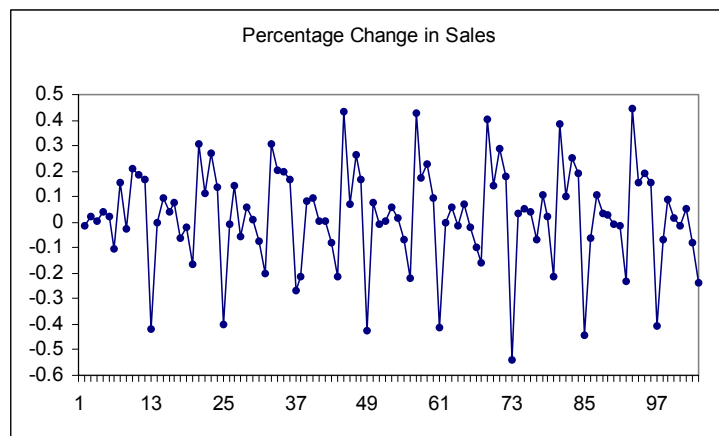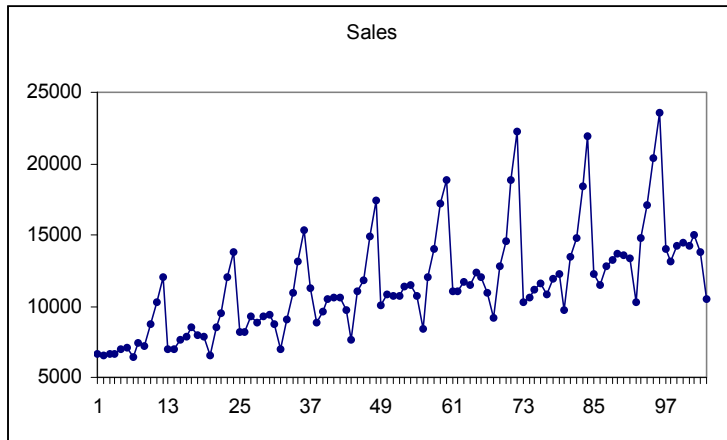
Problem 2: Wine Sales (II) (20 points)

In this problem, we will try log transformation to stabilize the variance and then build a regression model to forecast wine sales.

Run the regression

$$Log(Sales_t) = \beta_0 + \beta_1 t + \beta_2 Jan + \cdots + \beta_{12} Nov + \epsilon_t$$

where t is the trend variable and $Jan, Feb, \cdots, Nov$ are dummies variables for each month of the year.

(a) What would be the predicted increase of Wine Sales (in percentage) from December 1996 to January 1997?

(b) What would be the predicted increase of Wine Sales (in percentage) from July 1997 to August 1997?

(c) Plot the residuals against time. Do you see any clear patterns in the plot? Are there any model assumptions violated?

Add the term t^2 and run the regression

$$Log(Sales_t) = \beta_0 + \beta_1 t + \beta_2 Jan + \cdots + \beta_{12} Nov + \beta_{13} t^2 + \epsilon_t$$

(d) Plot the residuals against time. Do you see any clear patterns in the plot?

(e) Using the final model obtained above, provide numerical forecasts of wine sales in September, 1996 and October, 1996.

(f) Provide a 95% plug-in prediction interval for the forecast of sales in September, 1996.

(g) (Optional) What would be the predicted increase of Wine Sales (in percentage) from March 1997 to April 1997?

(h) (Optional) According to the model, without considering the seasonal effects, do the sales increase at a faster rate as time increases?

(i) (Optional) Build an Excel spreadsheet model or write R code to forecast the sales of the next 12 months.