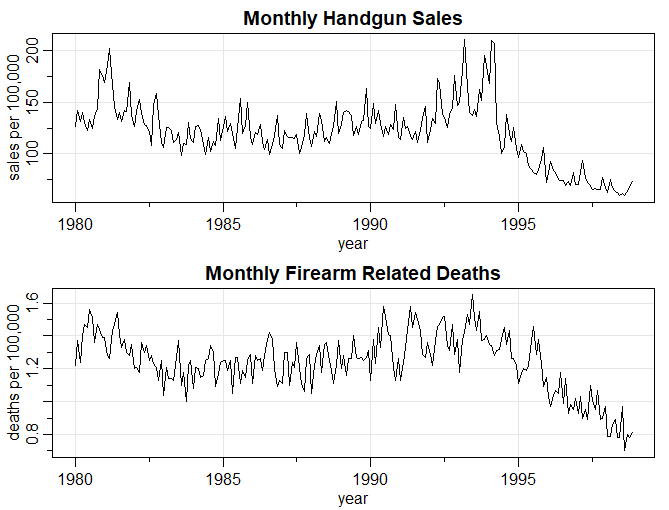
STA 137 Time Series Analysis

Final Project

**Introduction**

In recent American history, the conversation around firearms quickly leads to that of gun control reform as gun violence in our country seems all too familiar. In order to take a better look at this issue, monthly handgun sales and firearms related deaths recorded from 1980 to 1998 in the state of California. We will be interested in better understanding the trend of these two measurements as a function of time, in order to make inference about how the two vary in relation to each other. Due to the dependence that these observations have on time have we make use of time series methods to analyze this data in order to make inference on the trend and variance of the data. Again, because of the dependence that the data have on time, the usual statistical methods that require independence would not be appropriate.

**Material and Methods**

 The data consisted of monthly handgun sales and firearm related deaths in California were recorded from 1980 to 1998. Stationarity processes have constant mean and variance overtime, by looking at the data we see a trend, so we know that the data is non-stationary. Due to the trend and non-stationarity, transformations will have to be made in order to analyze the monthly

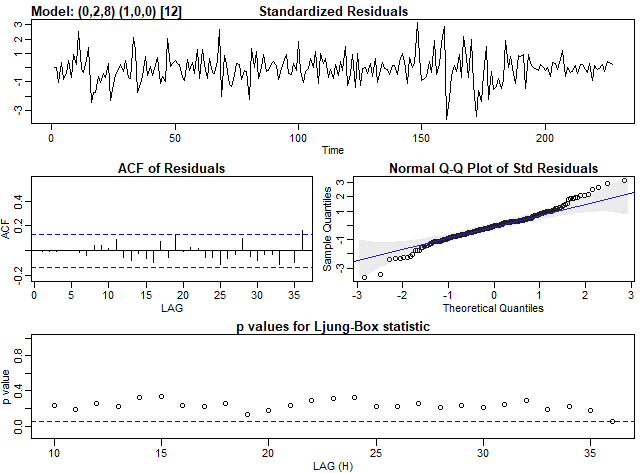
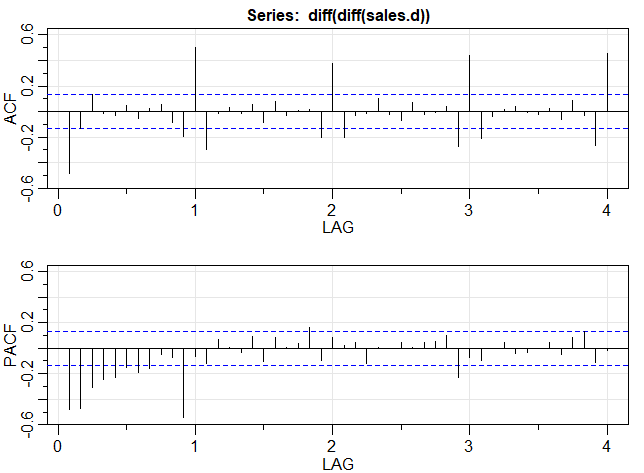
*Figure 1: time series plot.*

handgun sales and firearm related deaths. For both time series, the second order difference of each series will be taken in order to remove the trend and ensure stationarity of the data. We can also observe that the two series have similar patterns of sharp increases and decreases at the same points in time.

**Results**

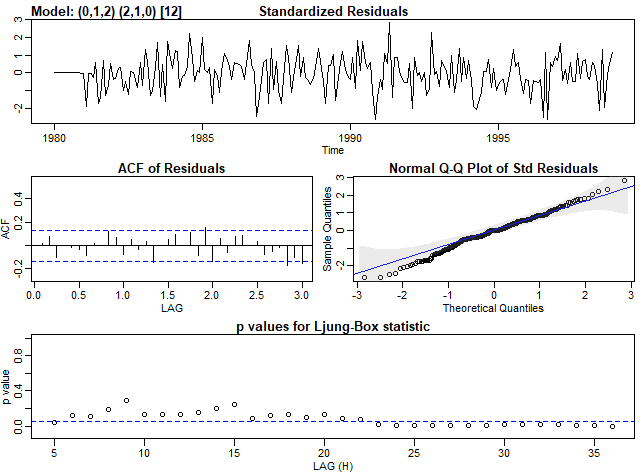
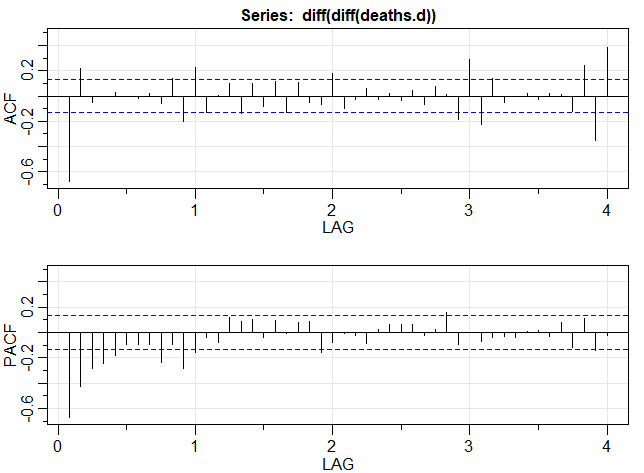
After taking the second difference of each time series to ensure the stationarity of the series, we can now plot the sample autocorrelation and partial autocorrelation functions. Once those plots are attained, we can use them to suggest a fit for an autoregressive integrated moving average (ARIMA) model and select the best fit model for our data using a certain selection criterion.

The plot of the autocorrelation function (acf) measures of the correlation between observations of a time series and observations separated by a certain lag in time. For our handgun sales series, when adjusting for the seasonal aspect of 12, we can see significant seasonal correlations in lags of 12 that tail off in the acf and cut off after lag 12 in pacf this suggests and moving average portion of this model. When looking at figure 1 only within the first lag, we observe that it cuts of after one significant point in the acf and tails off to near zero in the pacf. This observation suggests an autoregressive portion of the model. When putting all of these observations together we proposed and ARIMA model (0,2,8)(1,0,0)[12]. When looking at the standardized residuals to check if they are white noise it appears that they are. In addition, we have the acf of the residuals plot which have all but one nonsignificant lag. Noting that for the nature of this analysis we will conclude that the one slightly significant lag is not enough to invalid our model. In addition to the acf plot the p-values for the Ljung-Box statistic all appear to be significant. From the regression of this data we can conclude that this is a sufficient SARIMA model to fit this data.

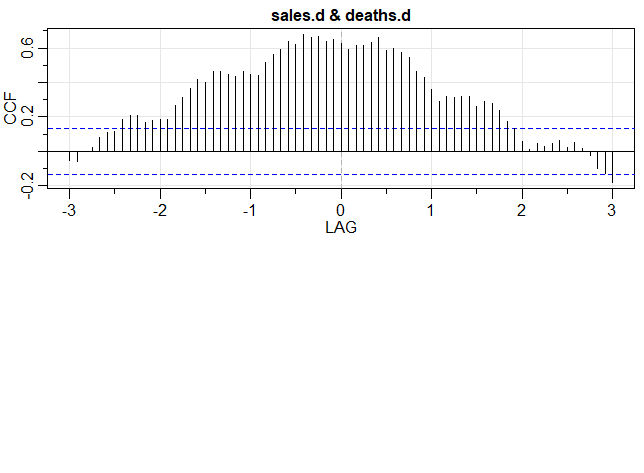


*Figure 1: sample acf and pacf of handgun sales data. Figure 2: standardized residuals for diagnostics of handgun sales data*

For the firearm related deaths series, when adjusting for a seasonal aspect of 12 for our data, we can see generally the acf tails off and the pacf cuts off after lag 1, suggesting a moving average portion our model. When putting all of these observations together we proposed and ARIMA model (0,1,2)(2,1,0)[12]. When looking at the standardized residuals to check if they are white noise it appears that they are. In addition, we have the acf of the residuals plot which have all but one nonsignificant lag. Noting that for the nature of this analysis we will conclude that the one slightly significant lag is not enough to invalid our model. In addition to the acf plot the p-values for the Ljung-Box statistic, for the purposes of this analysis enough appear to be significant. From the regression of this data we can conclude that this is a sufficient SARIMA model to fit this data. It would also be worth it to note that the models for each series that was chosen were done so by minimizing the selection criterion AIC.



*Figure 3: sample acf and pacf of firearm related deaths. Figure 4: standardized residuals for diagnostics of deaths.*

 The cross-correlation function is the correlation of two time series, in this case our handgun sales and firearm related deaths, separated by a lag in one of the series. From Figures 5 and our original time series plots of the data in Figure 1, it looks like handgun sales and firearm related deaths are positively correlated. We can tell this from the significant lag from -2.5 to 2. This means that the handgun sales and firearm related deaths from two and a half years in the past up to the two years in the future are correlated. Notice that by making inference here about the years rather than months because our data is adjusted for the 12 months in a year, so a lag of 1 displayed on the plot refers to 12 months of observations.

*Figure 5: ccf of handgun sales and fire arm related deaths*

**Conclusion and Discussion**

Based on the results of the analysis discussed above we can see that there is a link between handgun sales and firearm related deaths. Therefore, it would be worth it to note and understand that this positive correlation is present, and that it is significant enough to present to lawmakers in the event of a proposing reform.