U.S. Excess Deaths Project

Jason Rutberg

# Import Packages

library(ggplot2)  
library(forecast)

## Registered S3 method overwritten by 'quantmod':  
## method from  
## as.zoo.data.frame zoo

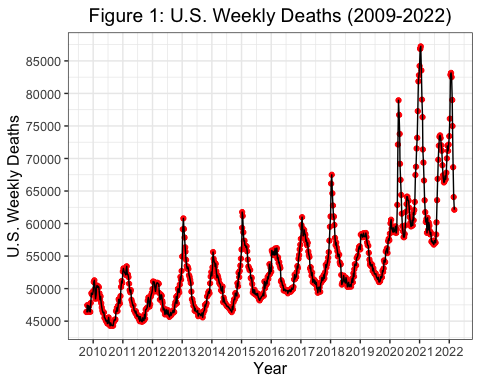
library(urca)

# Import Data and Split into Training and Testing Data Sets

data <- read.csv("US\_Weekly\_Deaths.csv", as.is=T)  
training\_data <- as.ts(data[1:535,]) # Before COVID deaths  
training\_data <- ts(training\_data, start = 2009.7665982, frequency = 52.17857)  
testing\_data <- as.ts(data[536:dim(data)[1],]) # Week 536 has first COVID death  
testing\_data <- ts(testing\_data, start = 2020.03832991, frequency = 52.17857)

# Figure 1

# Plot of US Weekly Deaths (2009-2022)  
ggplot(data, aes(x = Date, y = All.Deaths)) +   
 geom\_point(col = "red") +   
 geom\_line() +  
 labs(x = "Year",  
 y = "U.S. Weekly Deaths",  
 title = "Figure 1: U.S. Weekly Deaths (2009-2022)") +  
 scale\_x\_continuous(breaks = seq(2009, 2022, by = 1)) +  
 scale\_y\_continuous(breaks = seq(45000, 85000, by = 5000)) +  
 theme\_bw() +  
 theme(text = element\_text(size = 12),   
 plot.title = element\_text(hjust = 0.5))



# Time Series for All Deaths  
Deaths <- training\_data[,"All.Deaths"]  
Year <- training\_data[,"Year"]  
Week <- training\_data[,"Week"]

# Summary Statistics

summary(Deaths) # Summary of Weekly Deaths Pre-COVID (Training Dataset)

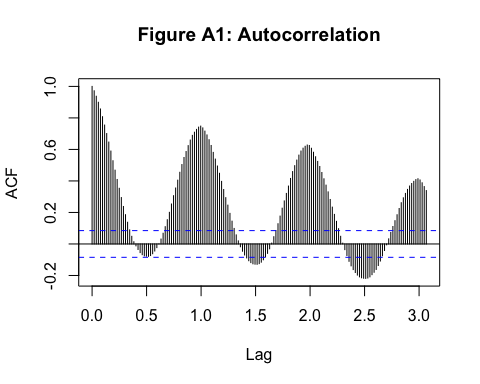
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 44281 47912 50749 51070 53468 67495

summary(testing\_data[,"All.Deaths"]) # Summary of Weekly Deaths During COVID (Testing Dataset)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 56764 59570 63576 66336 71969 87233

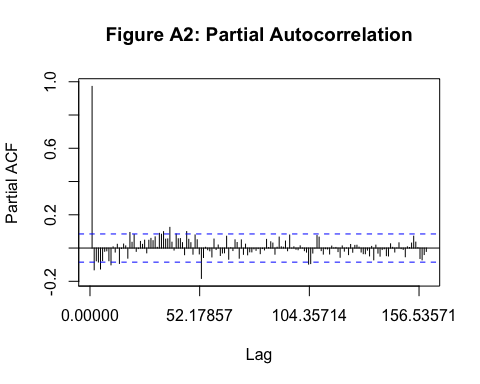
# Figure A1

# Autocorrelation  
acf(Deaths, lag = 160, main = 'Figure A1: Autocorrelation') # acf



# Figure A2

# Partial Autocorrelation  
Pacf(Deaths, lag = 160, main = 'Figure A2: Partial Autocorrelation')



# Test for Stationarity Using Original Data from Training Set

# Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test for Stationarity  
summary(ur.kpss(Deaths)) # reject null (NOT stationary)

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: mu with 6 lags.   
##   
## Value of test-statistic is: 3.3102   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.347 0.463 0.574 0.739

# Take Seasonal Difference of Weekly Deaths for Training Data Set

SD\_Deaths <- diff(Deaths, lag = 52, differences = 1)

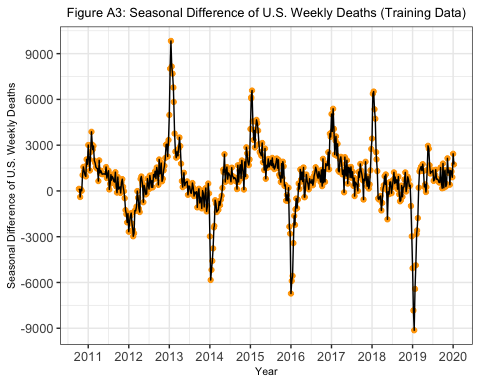
# Test for Stationarity Using Seasonally-Differenced Data from Training Set

summary(ur.kpss(SD\_Deaths)) # stationary

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: mu with 5 lags.   
##   
## Value of test-statistic is: 0.1032   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.347 0.463 0.574 0.739

# Figure A3

# Seasonal Difference of U.S. Weekly Deaths  
SD\_deaths\_data <- data.frame(Year = training\_data[53:length(Deaths),"Date"], SD\_Deaths)  
  
ggplot(SD\_deaths\_data, aes(x = Year, y = SD\_Deaths)) +   
 geom\_point(col = "orange") +   
 geom\_line() +  
 labs(x = "Year",  
 y = "Seasonal Difference of U.S. Weekly Deaths",  
 title = "Figure A3: Seasonal Difference of U.S. Weekly Deaths (Training Data)") +  
 scale\_x\_continuous(breaks = seq(2009, 2020, by = 1)) +  
 scale\_y\_continuous(breaks = seq(-15000, 15000, by = 3000)) +  
 theme\_bw() +  
 theme(text = element\_text(size = 12),   
 plot.title = element\_text(hjust = 0.5),  
 title = element\_text(size = 8))



# Model Selection

# Seasonal Component (0, 1, 0)  
M000 <- arima(Deaths, order = c(0,0,0), seasonal = list(order = c(0,1,0), period = 52))  
M001 <- arima(Deaths, order = c(1,0,0), seasonal = list(order = c(0,1,0), period = 52))  
M002 <- arima(Deaths, order = c(0,0,1), seasonal = list(order = c(0,1,0), period = 52))  
M003 <- arima(Deaths, order = c(1,0,1), seasonal = list(order = c(0,1,0), period = 52))  
M004 <- arima(Deaths, order = c(2,0,0), seasonal = list(order = c(0,1,0), period = 52))  
M005 <- arima(Deaths, order = c(0,0,2), seasonal = list(order = c(0,1,0), period = 52))  
M006 <- arima(Deaths, order = c(2,0,1), seasonal = list(order = c(0,1,0), period = 52))  
M007 <- arima(Deaths, order = c(1,0,2), seasonal = list(order = c(0,1,0), period = 52))  
M008 <- arima(Deaths, order = c(2,0,2), seasonal = list(order = c(0,1,0), period = 52))  
M009 <- arima(Deaths, order = c(3,0,0), seasonal = list(order = c(0,1,0), period = 52))  
M010 <- arima(Deaths, order = c(0,0,3), seasonal = list(order = c(0,1,0), period = 52))  
M011 <- arima(Deaths, order = c(3,0,1), seasonal = list(order = c(0,1,0), period = 52))  
M012 <- arima(Deaths, order = c(1,0,3), seasonal = list(order = c(0,1,0), period = 52))  
M013 <- arima(Deaths, order = c(3,0,2), seasonal = list(order = c(0,1,0), period = 52))  
M014 <- arima(Deaths, order = c(2,0,3), seasonal = list(order = c(0,1,0), period = 52))  
M015 <- arima(Deaths, order = c(3,0,3), seasonal = list(order = c(0,1,0), period = 52))  
  
# Seasonal Component (1, 1, 0)  
M100 <- arima(Deaths, order = c(0,0,0), seasonal = list(order = c(1,1,0), period = 52))  
M101 <- arima(Deaths, order = c(1,0,0), seasonal = list(order = c(1,1,0), period = 52))  
M102 <- arima(Deaths, order = c(0,0,1), seasonal = list(order = c(1,1,0), period = 52))  
M103 <- arima(Deaths, order = c(1,0,1), seasonal = list(order = c(1,1,0), period = 52))  
M104 <- arima(Deaths, order = c(2,0,0), seasonal = list(order = c(1,1,0), period = 52))  
M105 <- arima(Deaths, order = c(0,0,2), seasonal = list(order = c(1,1,0), period = 52))  
M106 <- arima(Deaths, order = c(2,0,1), seasonal = list(order = c(1,1,0), period = 52))  
M107 <- arima(Deaths, order = c(1,0,2), seasonal = list(order = c(1,1,0), period = 52))  
M108 <- arima(Deaths, order = c(2,0,2), seasonal = list(order = c(1,1,0), period = 52))  
M109 <- arima(Deaths, order = c(3,0,0), seasonal = list(order = c(1,1,0), period = 52))  
M110 <- arima(Deaths, order = c(0,0,3), seasonal = list(order = c(1,1,0), period = 52))  
M111 <- arima(Deaths, order = c(3,0,1), seasonal = list(order = c(1,1,0), period = 52))  
M112 <- arima(Deaths, order = c(1,0,3), seasonal = list(order = c(1,1,0), period = 52))  
M113 <- arima(Deaths, order = c(3,0,2), seasonal = list(order = c(1,1,0), period = 52))  
M114 <- arima(Deaths, order = c(2,0,3), seasonal = list(order = c(1,1,0), period = 52))  
M115 <- arima(Deaths, order = c(3,0,3), seasonal = list(order = c(1,1,0), period = 52))  
  
# Seasonal Component (0, 1, 1)  
M200 <- arima(Deaths, order = c(0,0,0), seasonal = list(order = c(0,1,1), period = 52))  
M201 <- arima(Deaths, order = c(1,0,0), seasonal = list(order = c(0,1,1), period = 52))  
M202 <- arima(Deaths, order = c(0,0,1), seasonal = list(order = c(0,1,1), period = 52))  
M203 <- arima(Deaths, order = c(1,0,1), seasonal = list(order = c(0,1,1), period = 52))  
M204 <- arima(Deaths, order = c(2,0,0), seasonal = list(order = c(0,1,1), period = 52))  
M205 <- arima(Deaths, order = c(0,0,2), seasonal = list(order = c(0,1,1), period = 52))  
M206 <- arima(Deaths, order = c(2,0,1), seasonal = list(order = c(0,1,1), period = 52))  
M207 <- arima(Deaths, order = c(1,0,2), seasonal = list(order = c(0,1,1), period = 52))  
M208 <- arima(Deaths, order = c(2,0,2), seasonal = list(order = c(0,1,1), period = 52))  
M209 <- arima(Deaths, order = c(3,0,0), seasonal = list(order = c(0,1,1), period = 52))  
M210 <- arima(Deaths, order = c(0,0,3), seasonal = list(order = c(0,1,1), period = 52))  
M211 <- arima(Deaths, order = c(3,0,1), seasonal = list(order = c(0,1,1), period = 52))  
M212 <- arima(Deaths, order = c(1,0,3), seasonal = list(order = c(0,1,1), period = 52))  
M213 <- arima(Deaths, order = c(3,0,2), seasonal = list(order = c(0,1,1), period = 52))  
M214 <- arima(Deaths, order = c(2,0,3), seasonal = list(order = c(0,1,1), period = 52))  
M215 <- arima(Deaths, order = c(3,0,3), seasonal = list(order = c(0,1,1), period = 52))  
  
# Seasonal Component (1, 1, 1)  
M300 <- arima(Deaths, order = c(0,0,0), seasonal = list(order = c(1,1,1), period = 52))  
M301 <- arima(Deaths, order = c(1,0,0), seasonal = list(order = c(1,1,1), period = 52))  
M302 <- arima(Deaths, order = c(0,0,1), seasonal = list(order = c(1,1,1), period = 52))  
M303 <- arima(Deaths, order = c(1,0,1), seasonal = list(order = c(1,1,1), period = 52))  
M304 <- arima(Deaths, order = c(2,0,0), seasonal = list(order = c(1,1,1), period = 52))  
M305 <- arima(Deaths, order = c(0,0,2), seasonal = list(order = c(1,1,1), period = 52))  
M306 <- arima(Deaths, order = c(2,0,1), seasonal = list(order = c(1,1,1), period = 52))  
M307 <- arima(Deaths, order = c(1,0,2), seasonal = list(order = c(1,1,1), period = 52))  
M308 <- arima(Deaths, order = c(2,0,2), seasonal = list(order = c(1,1,1), period = 52))  
M309 <- arima(Deaths, order = c(3,0,0), seasonal = list(order = c(1,1,1), period = 52))  
M310 <- arima(Deaths, order = c(0,0,3), seasonal = list(order = c(1,1,1), period = 52))  
M311 <- arima(Deaths, order = c(3,0,1), seasonal = list(order = c(1,1,1), period = 52))  
M312 <- arima(Deaths, order = c(1,0,3), seasonal = list(order = c(1,1,1), period = 52))  
M313 <- arima(Deaths, order = c(3,0,2), seasonal = list(order = c(1,1,1), period = 52))  
M314 <- arima(Deaths, order = c(2,0,3), seasonal = list(order = c(1,1,1), period = 52))  
M315 <- arima(Deaths, order = c(3,0,3), seasonal = list(order = c(1,1,1), period = 52))  
  
# Seasonal Component (2, 1, 0)  
M400 <- arima(Deaths, order = c(0,0,0), seasonal = list(order = c(2,1,0), period = 52))  
M401 <- arima(Deaths, order = c(1,0,0), seasonal = list(order = c(2,1,0), period = 52))  
M402 <- arima(Deaths, order = c(0,0,1), seasonal = list(order = c(2,1,0), period = 52))  
M403 <- arima(Deaths, order = c(1,0,1), seasonal = list(order = c(2,1,0), period = 52))  
M404 <- arima(Deaths, order = c(2,0,0), seasonal = list(order = c(2,1,0), period = 52))  
M405 <- arima(Deaths, order = c(0,0,2), seasonal = list(order = c(2,1,0), period = 52))  
M406 <- arima(Deaths, order = c(2,0,1), seasonal = list(order = c(2,1,0), period = 52))  
M407 <- arima(Deaths, order = c(1,0,2), seasonal = list(order = c(2,1,0), period = 52))  
M408 <- arima(Deaths, order = c(2,0,2), seasonal = list(order = c(2,1,0), period = 52))  
M409 <- arima(Deaths, order = c(3,0,0), seasonal = list(order = c(2,1,0), period = 52))  
M410 <- arima(Deaths, order = c(0,0,3), seasonal = list(order = c(2,1,0), period = 52))  
M411 <- arima(Deaths, order = c(3,0,1), seasonal = list(order = c(2,1,0), period = 52))  
M412 <- arima(Deaths, order = c(1,0,3), seasonal = list(order = c(2,1,0), period = 52))  
M413 <- arima(Deaths, order = c(3,0,2), seasonal = list(order = c(2,1,0), period = 52))  
M414 <- arima(Deaths, order = c(2,0,3), seasonal = list(order = c(2,1,0), period = 52))  
M415 <- arima(Deaths, order = c(3,0,3), seasonal = list(order = c(2,1,0), period = 52))  
  
# Seasonal Component (0, 1, 2)  
M500 <- arima(Deaths, order = c(0,0,0), seasonal = list(order = c(0,1,2), period = 52))  
M501 <- arima(Deaths, order = c(1,0,0), seasonal = list(order = c(0,1,2), period = 52))  
M502 <- arima(Deaths, order = c(0,0,1), seasonal = list(order = c(0,1,2), period = 52))  
M503 <- arima(Deaths, order = c(1,0,1), seasonal = list(order = c(0,1,2), period = 52))  
M504 <- arima(Deaths, order = c(2,0,0), seasonal = list(order = c(0,1,2), period = 52))  
M505 <- arima(Deaths, order = c(0,0,2), seasonal = list(order = c(0,1,2), period = 52))  
M506 <- arima(Deaths, order = c(2,0,1), seasonal = list(order = c(0,1,2), period = 52))  
M507 <- arima(Deaths, order = c(1,0,2), seasonal = list(order = c(0,1,2), period = 52))  
M508 <- arima(Deaths, order = c(2,0,2), seasonal = list(order = c(0,1,2), period = 52))  
M509 <- arima(Deaths, order = c(3,0,0), seasonal = list(order = c(0,1,2), period = 52))  
M510 <- arima(Deaths, order = c(0,0,3), seasonal = list(order = c(0,1,2), period = 52))  
M511 <- arima(Deaths, order = c(3,0,1), seasonal = list(order = c(0,1,2), period = 52))  
M512 <- arima(Deaths, order = c(1,0,3), seasonal = list(order = c(0,1,2), period = 52))  
M513 <- arima(Deaths, order = c(3,0,2), seasonal = list(order = c(0,1,2), period = 52))  
M514 <- arima(Deaths, order = c(2,0,3), seasonal = list(order = c(0,1,2), period = 52))  
M515 <- arima(Deaths, order = c(3,0,3), seasonal = list(order = c(0,1,2), period = 52))  
  
# Seasonal Component (2, 1, 1)  
M600 <- arima(Deaths, order = c(0,0,0), seasonal = list(order = c(2,1,1), period = 52))  
M601 <- arima(Deaths, order = c(1,0,0), seasonal = list(order = c(2,1,1), period = 52))  
M602 <- arima(Deaths, order = c(0,0,1), seasonal = list(order = c(2,1,1), period = 52))  
M603 <- arima(Deaths, order = c(1,0,1), seasonal = list(order = c(2,1,1), period = 52))  
M604 <- arima(Deaths, order = c(2,0,0), seasonal = list(order = c(2,1,1), period = 52))  
M605 <- arima(Deaths, order = c(0,0,2), seasonal = list(order = c(2,1,1), period = 52))  
M606 <- arima(Deaths, order = c(2,0,1), seasonal = list(order = c(2,1,1), period = 52))  
M607 <- arima(Deaths, order = c(1,0,2), seasonal = list(order = c(2,1,1), period = 52))  
M608 <- arima(Deaths, order = c(2,0,2), seasonal = list(order = c(2,1,1), period = 52))  
M609 <- arima(Deaths, order = c(3,0,0), seasonal = list(order = c(2,1,1), period = 52))  
M610 <- arima(Deaths, order = c(0,0,3), seasonal = list(order = c(2,1,1), period = 52))  
M611 <- arima(Deaths, order = c(3,0,1), seasonal = list(order = c(2,1,1), period = 52))  
M612 <- arima(Deaths, order = c(1,0,3), seasonal = list(order = c(2,1,1), period = 52))  
M613 <- arima(Deaths, order = c(3,0,2), seasonal = list(order = c(2,1,1), period = 52))  
M614 <- arima(Deaths, order = c(2,0,3), seasonal = list(order = c(2,1,1), period = 52))  
M615 <- arima(Deaths, order = c(3,0,3), seasonal = list(order = c(2,1,1), period = 52))  
  
# Seasonal Component (1, 1, 2)  
M700 <- arima(Deaths, order = c(0,0,0), seasonal = list(order = c(1,1,2), period = 52))  
M701 <- arima(Deaths, order = c(1,0,0), seasonal = list(order = c(1,1,2), period = 52))  
M702 <- arima(Deaths, order = c(0,0,1), seasonal = list(order = c(1,1,2), period = 52))  
M703 <- arima(Deaths, order = c(1,0,1), seasonal = list(order = c(1,1,2), period = 52))  
M704 <- arima(Deaths, order = c(2,0,0), seasonal = list(order = c(1,1,2), period = 52))  
M705 <- arima(Deaths, order = c(0,0,2), seasonal = list(order = c(1,1,2), period = 52))  
M706 <- arima(Deaths, order = c(2,0,1), seasonal = list(order = c(1,1,2), period = 52))  
M707 <- arima(Deaths, order = c(1,0,2), seasonal = list(order = c(1,1,2), period = 52))  
M708 <- arima(Deaths, order = c(2,0,2), seasonal = list(order = c(1,1,2), period = 52))  
M709 <- arima(Deaths, order = c(3,0,0), seasonal = list(order = c(1,1,2), period = 52))  
M710 <- arima(Deaths, order = c(0,0,3), seasonal = list(order = c(1,1,2), period = 52))  
M711 <- arima(Deaths, order = c(3,0,1), seasonal = list(order = c(1,1,2), period = 52))  
M712 <- arima(Deaths, order = c(1,0,3), seasonal = list(order = c(1,1,2), period = 52))  
M713 <- arima(Deaths, order = c(3,0,2), seasonal = list(order = c(1,1,2), period = 52))  
M714 <- arima(Deaths, order = c(2,0,3), seasonal = list(order = c(1,1,2), period = 52))  
M715 <- arima(Deaths, order = c(3,0,3), seasonal = list(order = c(1,1,2), period = 52))  
  
# Seasonal Component (2, 1, 2)  
M800 <- arima(Deaths, order = c(0,0,0), seasonal = list(order = c(2,1,2), period = 52))  
M801 <- arima(Deaths, order = c(1,0,0), seasonal = list(order = c(2,1,2), period = 52))  
M802 <- arima(Deaths, order = c(0,0,1), seasonal = list(order = c(2,1,2), period = 52))  
M803 <- arima(Deaths, order = c(1,0,1), seasonal = list(order = c(2,1,2), period = 52))  
M804 <- arima(Deaths, order = c(2,0,0), seasonal = list(order = c(2,1,2), period = 52))  
M805 <- arima(Deaths, order = c(0,0,2), seasonal = list(order = c(2,1,2), period = 52))  
M806 <- arima(Deaths, order = c(2,0,1), seasonal = list(order = c(2,1,2), period = 52))  
M807 <- arima(Deaths, order = c(1,0,2), seasonal = list(order = c(2,1,2), period = 52))  
M808 <- arima(Deaths, order = c(2,0,2), seasonal = list(order = c(2,1,2), period = 52))  
M809 <- arima(Deaths, order = c(3,0,0), seasonal = list(order = c(2,1,2), period = 52))  
M810 <- arima(Deaths, order = c(0,0,3), seasonal = list(order = c(2,1,2), period = 52))  
M811 <- arima(Deaths, order = c(3,0,1), seasonal = list(order = c(2,1,2), period = 52))  
M812 <- arima(Deaths, order = c(1,0,3), seasonal = list(order = c(2,1,2), period = 52))  
M813 <- arima(Deaths, order = c(3,0,2), seasonal = list(order = c(2,1,2), period = 52))  
M814 <- arima(Deaths, order = c(2,0,3), seasonal = list(order = c(2,1,2), period = 52))  
M815 <- arima(Deaths, order = c(3,0,3), seasonal = list(order = c(2,1,2), period = 52))

# Compare Model AIC’s

# Compare Models (choose one with lowest AIC)  
AIC(M000, M001, M002, M003, M004, M005, M006, M007,  
 M008, M009, M010, M011, M012, M013, M014, M015,  
 M100, M101, M102, M103, M104, M105, M106, M107,  
 M108, M109, M110, M111, M112, M113, M114, M115,  
 M200, M201, M202, M203, M204, M205, M206, M207,  
 M208, M209, M210, M211, M212, M213, M214, M215,  
 M300, M301, M302, M303, M304, M305, M306, M307,  
 M308, M309, M310, M311, M312, M313, M314, M315,  
 M400, M401, M402, M403, M404, M405, M406, M407,  
 M408, M409, M410, M411, M412, M413, M414, M415,  
 M500, M501, M502, M503, M504, M505, M506, M507,  
 M508, M509, M510, M511, M512, M513, M514, M515,  
 M600, M601, M602, M603, M604, M605, M606, M607,  
 M608, M609, M610, M611, M612, M613, M614, M615,  
 M700, M701, M702, M703, M704, M705, M706, M707,  
 M708, M709, M710, M711, M712, M713, M714, M715,  
 M800, M801, M802, M803, M804, M805, M806, M807,  
 M808, M809, M810, M811, M812, M813, M814, M815)

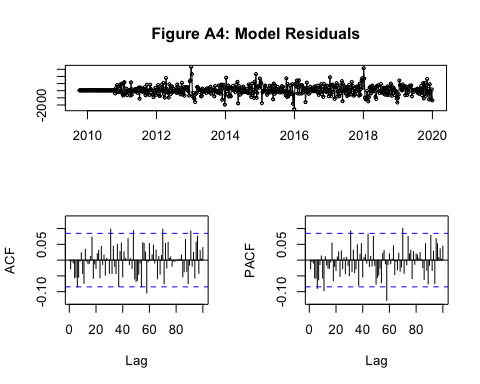
# Final Model: M513: ARIMA(3, 0, 2)(0, 1, 2) [52]

M513 <- arima(Deaths, order = c(3,0,2), seasonal = list(order = c(0,1,2), period = 52))  
summary(M513)

##   
## Call:  
## arima(x = Deaths, order = c(3, 0, 2), seasonal = list(order = c(0, 1, 2), period = 52))  
##   
## Coefficients:  
## ar1 ar2 ar3 ma1 ma2 sma1 sma2  
## -0.8976 0.8908 0.8405 1.8074 0.8188 -0.7180 0.1034  
## s.e. 0.0762 0.0263 0.0653 0.0980 0.0963 0.0499 0.0533  
##   
## sigma^2 estimated as 495982: log likelihood = -3868.77, aic = 7753.54  
##   
## Training set error measures:  
## ME RMSE MAE MPE MAPE MASE  
## Training set 80.57158 669.3204 477.5416 0.1457515 0.9107069 0.776708  
## ACF1  
## Training set -0.02794821

# Figure A4

tsdisplay(residuals(M513), lag.max = 100, main='Figure A4: Model Residuals')

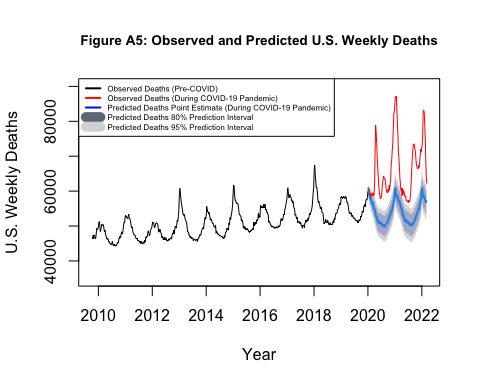


# Model Forecasting

forecast <- forecast(M513, h = 113) # 113 weeks in the testing data set

# Figure A5

plot(forecast, xlab = "Year",   
 ylab = "U.S. Weekly Deaths",   
 main = "Figure A5: Observed and Predicted U.S. Weekly Deaths",  
 ylim = c(35000, 90000), cex.main = 0.85)  
lines(testing\_data[,"All.Deaths"], col = "red")  
legend("topleft", legend = c("Observed Deaths (Pre-COVID)", "Observed Deaths (During COVID-19 Pandemic)", "Predicted Deaths Point Estimate (During COVID-19 Pandemic)", "Predicted Deaths 80% Prediction Interval", "Predicted Deaths 95% Prediction Interval"),   
 lty = 1, lwd = c(2, 2, 2, 10, 10), col = c("black", "red", "blue", "lightsteelblue4", "gray85"), cex = 0.5)



# Compute Excess Deaths

point\_estimate <- forecast$mean  
upper\_bound <- forecast$upper[,2]  
excess\_deaths <- as.numeric(testing\_data[,"All.Deaths"]) - as.numeric(upper\_bound)  
excess\_deaths\_data <- data.frame(Year = testing\_data[,"Date"], COVID\_deaths = testing\_data[,"COVID.19.Deaths"], excess\_deaths)

# Total Excess Deaths

sum(excess\_deaths) # total number of excess deaths

## [1] 883510.4

# Total COVID-19 Deaths

sum(testing\_data[,"COVID.19.Deaths"]) # total number of COVID deaths

## [1] 981196

# Correlation Between Excess Deaths and COVID-19 Deaths

round(cor(excess\_deaths, testing\_data[,"COVID.19.Deaths"]), 3)

## [1] 0.951

cor.test(excess\_deaths, testing\_data[,"COVID.19.Deaths"])

##   
## Pearson's product-moment correlation  
##   
## data: excess\_deaths and testing\_data[, "COVID.19.Deaths"]  
## t = 32.442, df = 111, p-value < 2.2e-16  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.9297232 0.9660933  
## sample estimates:  
## cor   
## 0.9511033

# Figure 2

ggplot(excess\_deaths\_data, aes(x = Year, y = excess\_deaths)) +   
 geom\_point(data = excess\_deaths\_data, aes(x = Year, y = excess\_deaths), color = "black") +  
 geom\_point(data = excess\_deaths\_data, aes(x = Year, y = COVID\_deaths), color = "black") +  
 geom\_line(data = excess\_deaths\_data, aes(x = Year, y = excess\_deaths, color = 'Excess Deaths')) +  
 geom\_line(data = excess\_deaths\_data, aes(x = Year, y = COVID\_deaths, color = 'COVID-19 Deaths')) +  
 labs(x = "Year",  
 y = "U.S. Weekly Deaths",  
 title = "Figure 2: U.S. Weekly Excess Deaths and COVID-19 Deaths (2020-2022)") +  
 scale\_x\_continuous(breaks = seq(2020, 2022.5, by = 0.5)) +  
 scale\_y\_continuous(breaks = seq(0, 30000, by = 5000)) +  
 theme\_bw() +  
 theme(text = element\_text(size = 7),  
 plot.title = element\_text(hjust = 0.5)) +  
 scale\_color\_manual(name = 'Legend',  
 breaks=c("Excess Deaths", "COVID-19 Deaths"),  
 values=c("Excess Deaths" = "blue", "COVID-19 Deaths" = "red"))

