Mahoney, Mike

Take Home Exam #2

DA6823

Time Series Analysis

Due 5/7/18 midnight via email

This is the second of two take home exams. I want to give you plenty of time to work on the exam and be able to ask questions and get guidance where you might need it. This task should also help you hone your technical writing skills – an important part of the data analysis process.

The objective of this exam is for you to locate several time series data sets of interest and analyze them using the tools that you have learned in lecture. You may use whatever analytic platform(s) you wish – SAS, R, Python, GRETL, STATA, etc. Once again I want a “stream of consciousness” set of answers so that I can follow your thinking. Here are the components of the assignment:

1. Select a scientific, biomedical, business or other issue that appeals to you and go looking online for relevant time series data sets. The good news here is that there are tons of free and interesting time series data sets online. If you have problems locating them let me know and I will help. Tell me what issues you are investigating and cite the data sets so that I know where they came from. I want you to select at **least two different data sets** and perform the following analyses on each data set. For the first data set find one **without seasonality** in it. You will use this data set in steps 1 through 8. **The second of these data sets has to have seasonality in it and will only be used in step 9.** Suck the data set into your platform(s) of choice.

I found a dataset on barley production in England in Wales from 1884 to 1939.

https://datamarket.com/data/set/22rj/annual-barley-yields-per-acre-in-england-wales-1884-1939#!ds=22rj&display=line

**The following steps 2 through 8 are done only on data set #1 (the one without seasonality).**

1. Plot out your time series variable. Tell me using your Mark I eyeball whether or not you think the time series data set is stationary in terms of **constant mean** and also **constant variance**.

No, my dataset does not have a constant mean nor constant variance. My data has a low in the 1890’s, the mean is dropped drastically in the 1910’s, and the variance sways violently in the mid 1920’s to the end of my times series. It is annual, so no yearly seasonality is present. It does span 55 years with inconsistent swings, in the barley\_yields over time.

1. Plot the ACF for the time series data set. Looking at ACF, does it look like there may be a trend or non-constant mean for each time series?

No, there does not appear to be a trend, lag 0 through 3, show an exponential decay. All of our lags (except 0) are with in our confidence interval. I am confident there is not a trend in my barley data.

1. Now let’s examine the time series data set using unit root tests. First use the KPSS test for the time series data set and tell me if the test suggests if there is a constant mean or not. Then see if you can confirm your KPSS evaluation using the Augmented Dickey Fuller (ADF) or the ADF-GLS test for each time series. What is your decision concerning constant mean?

The KPSS test (using the tseries and urca libraries) both indicate I should reject the null hypothesis (that the data in question is stationary). The ADF test (from the tseries library) indicates we fail to reject the null hypothesis (that our data in question is non stationary). However the ADF (from the urca library) just barely indicates I should reject the null hypothesis. 3 out of 4 test ran, indicate the mean is non-constant. I am confident that my barley data does not have a constant mean (and therefore should be differenced).

1. Review the decisions in step #4. If the test suggests that there is a non-constant mean then use differencing to create a new differenced variable for the time series data set.
   1. Plot out the data for the new differenced data set. Tell me if it looks like the differencing got rid of the trend or non-constant mean.

Yes, the mean is now constant. All values are between 4 and -4, a drastic improvement over our original dataset

* 1. Plot the ACF for the differenced time series. Tell me if this new ACF plot looks like there now is no trend.

There is no trend in our barley\_diff\_df, as indicated by a “flip-flop” pattern in the ACF.

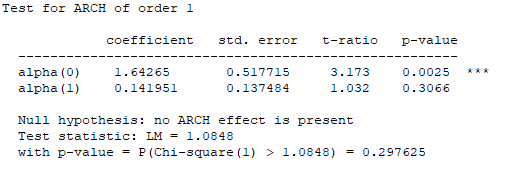
* 1. Apply the KPSS test and the ADF or ADF-GLS test to the differenced data – does the trend disappear?

Yes, the trend is gone in the differenced dataset. In both KPSS for the differenced dataset, I failed to reject the Null Hypothesis meaning my differenced data is stationary. In the ADF Test, I rejected the Null Hypothesis both times for the differenced dataset. I am confident that my differenced dataset should be utilized for the rest of part 1.

**Note: From this point onward through step 8, if the time series was differenced, use the differenced time series data set for all the rest of the questions. Otherwise you can use the undifferenced data set.**

1. Test each of the time series data sets for constant variance using the ARCH test (GRETL does this nicely). Tell me which ones might have issues with constant variance and so not be so nicely stationary. Note that we will not do anything about this issue for the moment, but it’s good to know.

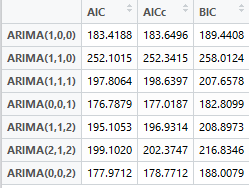
I used GRETL >>



1. Plot the PACF for the time series data sets. Using the combined information from the ACF you plotted earlier along with the information in the PACF, tell me if you see autoregressive and/or moving average processes in the data set. To help with interpretation you may want to refer to online resources – here is a decent resource from Duke University [**https://people.duke.edu/~rnau/411arim3.htm**](https://people.duke.edu/~rnau/411arim3.htm) or Penn State https://onlinecourses.science.psu.edu/stat510/node/64

I see a Moving Average process in the differenced data. I have a negative first lag in the PACF and the ACF is showing a sharp decline. I am should use a first order MA term in my ARIMA model.

1. For your time series data set, experiment with different ARIMA models for them. As you try them, list out the results of the various models and
   1. Comment on how each one is working and compare it to the previous model using various metrics such as SBC, BIC, Box Leung, etc. Most students end up creating a small able with these statistics across the models tried so it is easy to compare them.

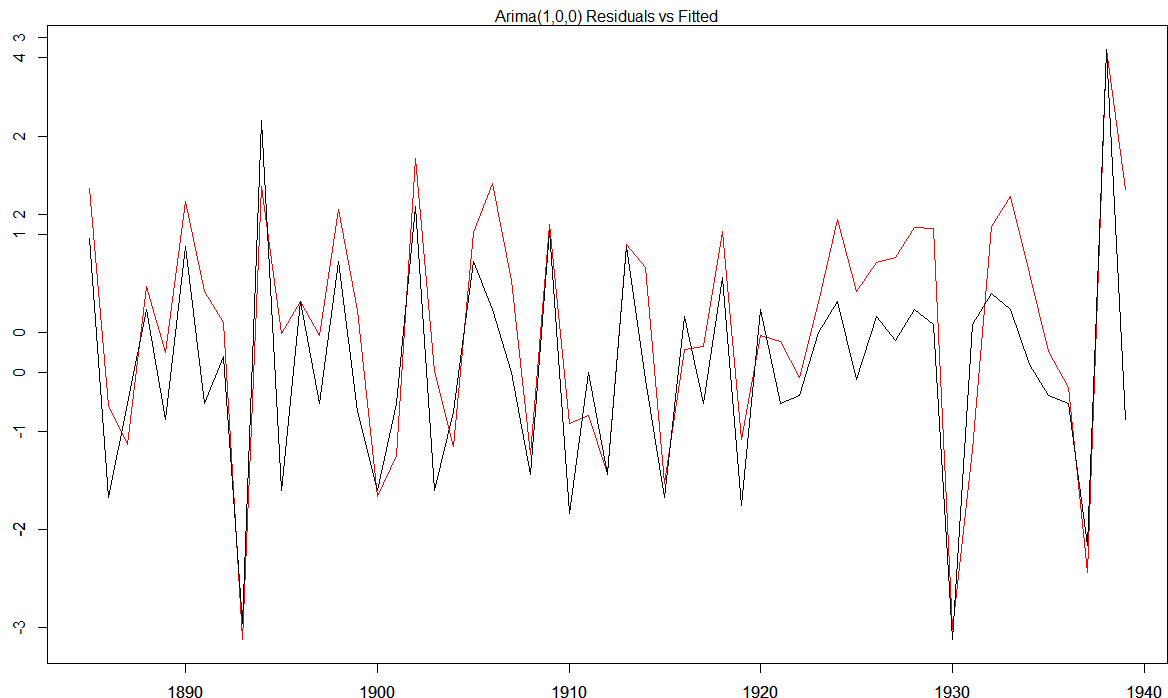
****

In practice we want to reduce the AIC and BIC (or SBC), I will go over the models from worst to best. The worst performing model is the ARIMA(1,1,0), or the first order Auto Regressive and first order differencing, which had the highest AIC and BIC denoting poor model fits. The next worst performer was the ARIMA(2,1,2) with a high BIC.

The next highest AIC and BIC was ARIMA(1,1,1) and ARIMA(1,1,2), the former had a higher AIC and the latter the higher BIC. The ARIMA(1,0,0) model is getting better but does not have the lowest AIC and BIC. The ARIMA(0,0,2) is very close to minimizing the AIC and BIC. However it is the ARIMA(0,0,1) that maintains the lowest AIC and BIC, which is exactly what our PACF told us. Although the AICc is given, the value is very close to the AIC (so the sample size must be large enough to not receive such a large penalty)

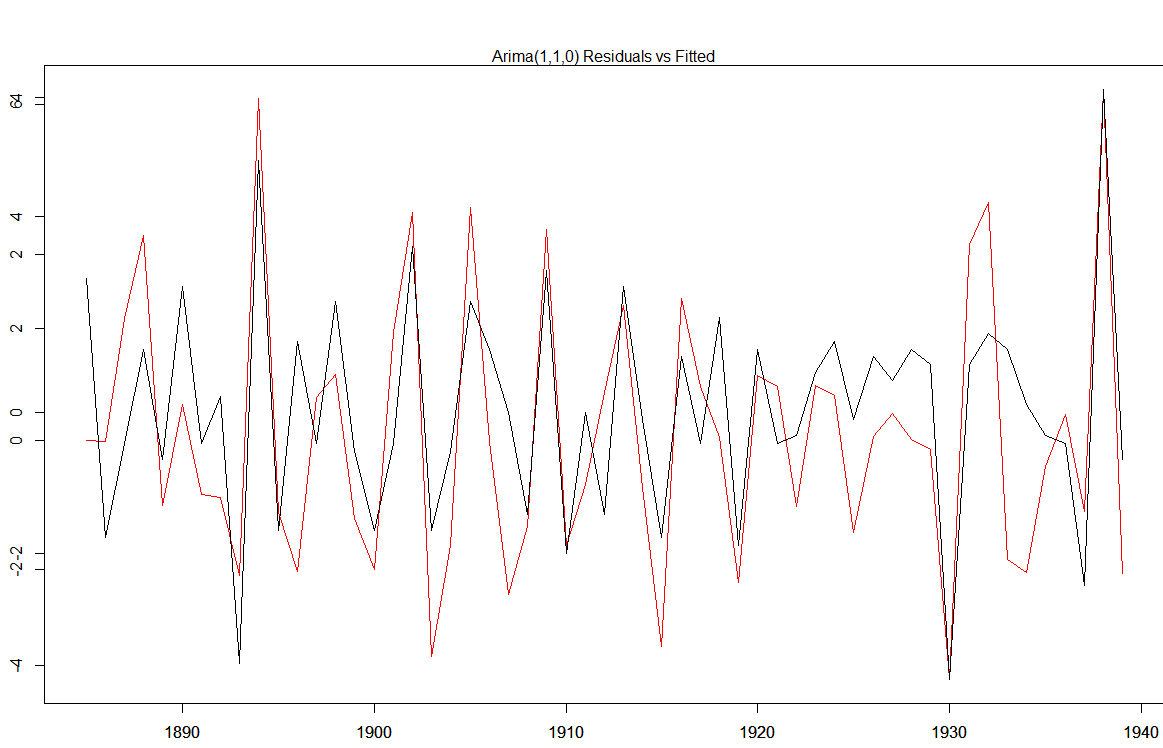
* 1. Plot the observed versus fitted data for the time series data set and comment on how well the model seems to be working >>> Note: For all plots, Black = Fit, Red = Residuals

ARIMA 1



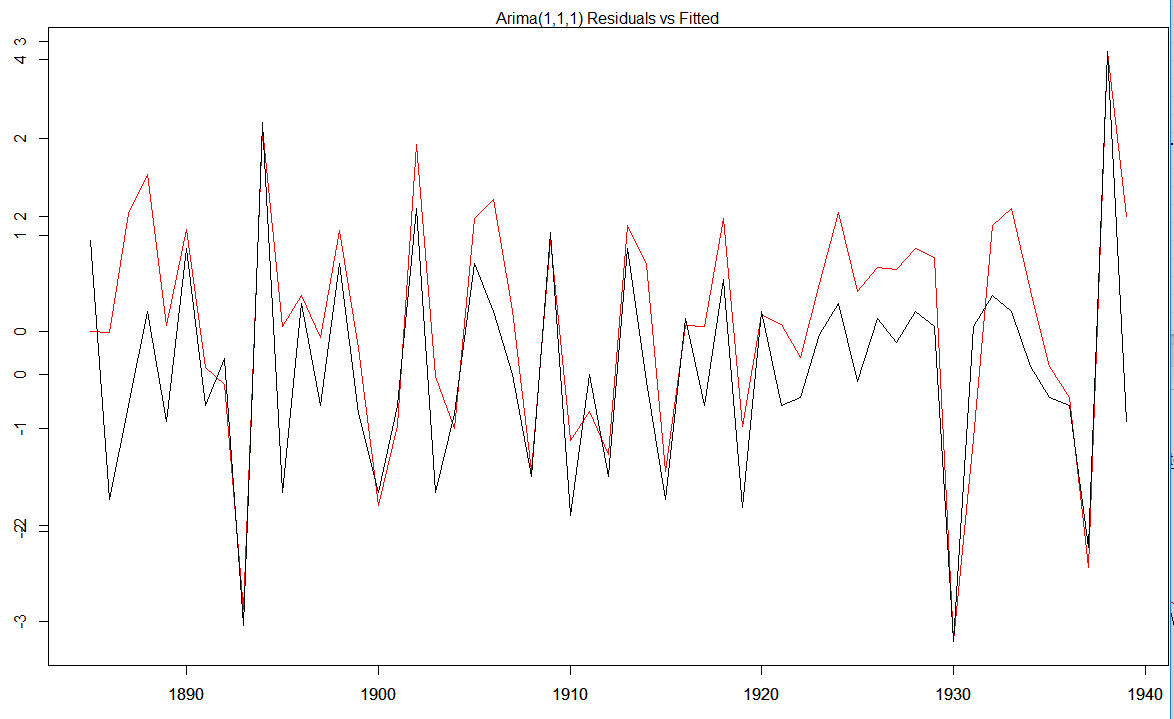
The model seems to minimize the residuals in the valleys but misses the peaks.

ARIMA 2



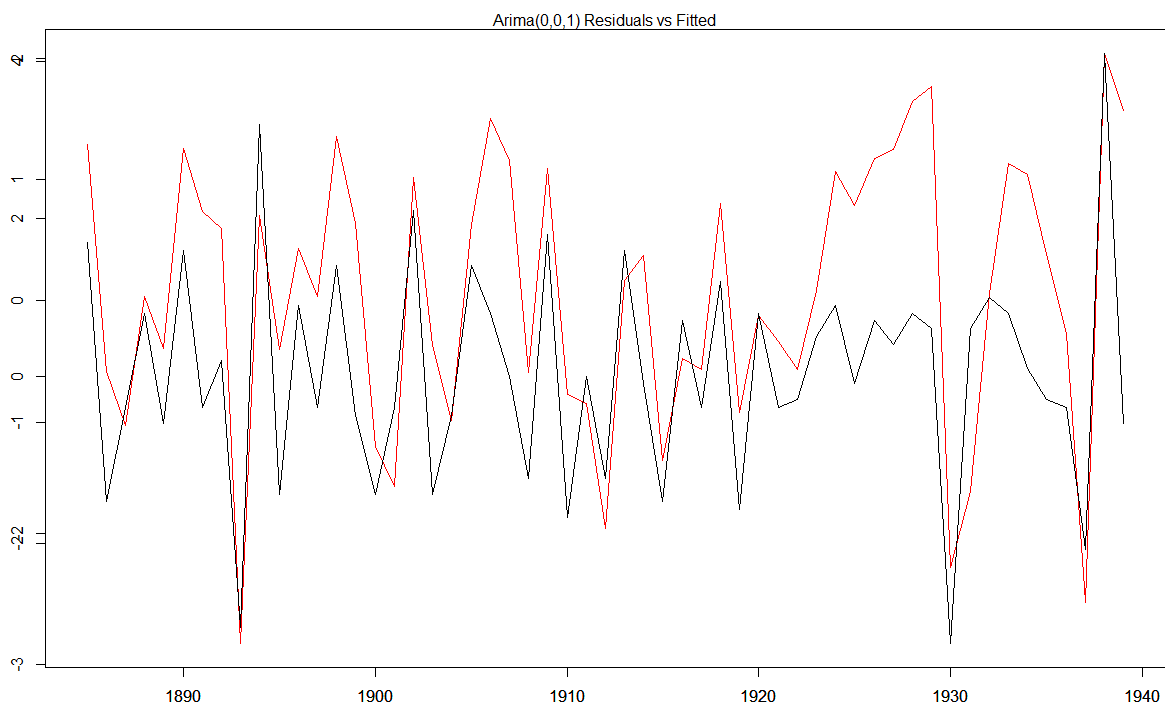
I do not think this model works well, the residuals are drastically off from the fit.

ARIMA 3

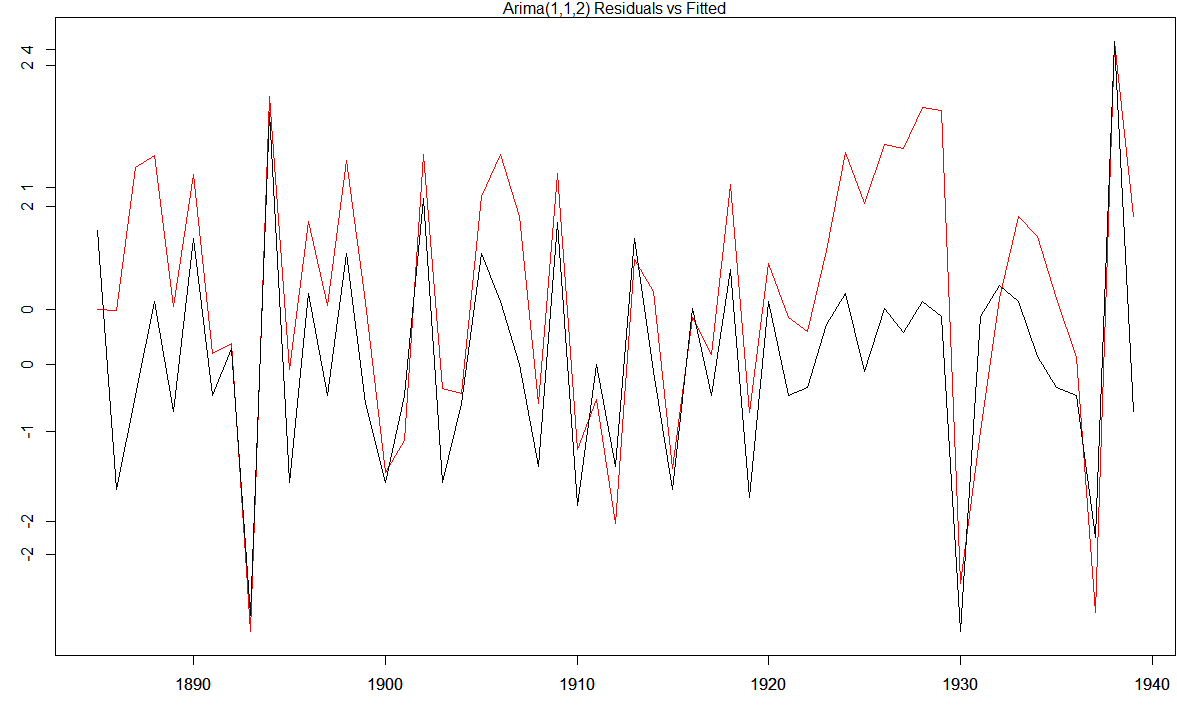


I think this is the best fitting ARIMA model

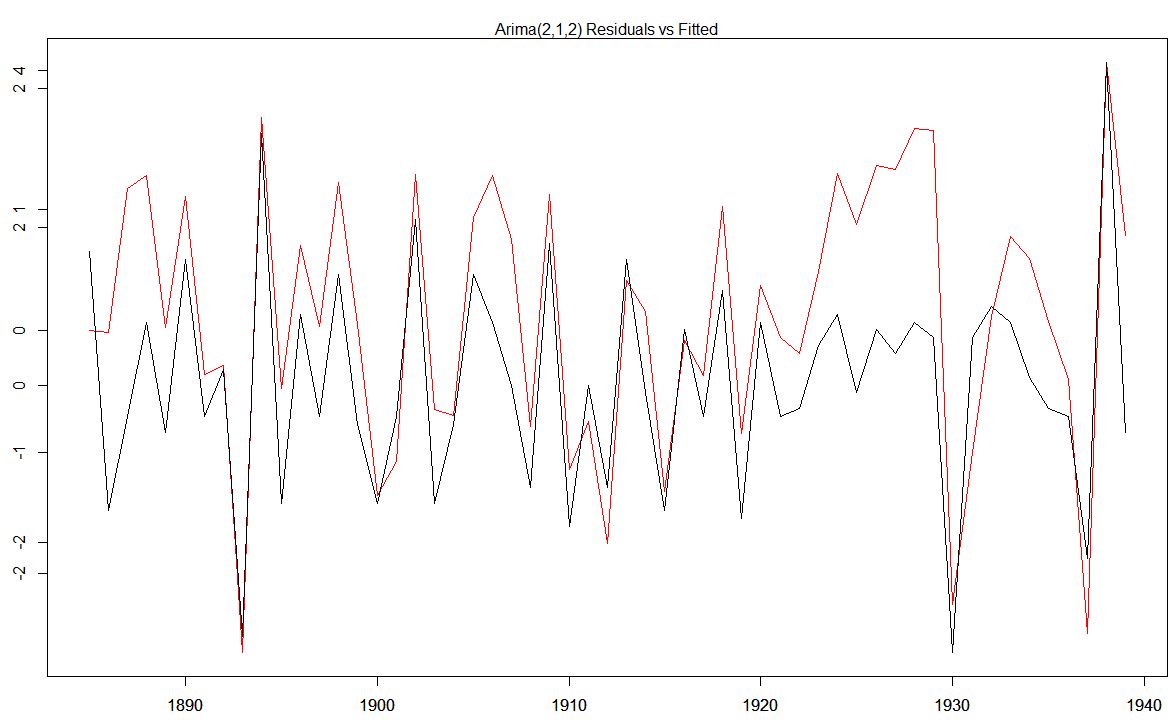
Arima 4



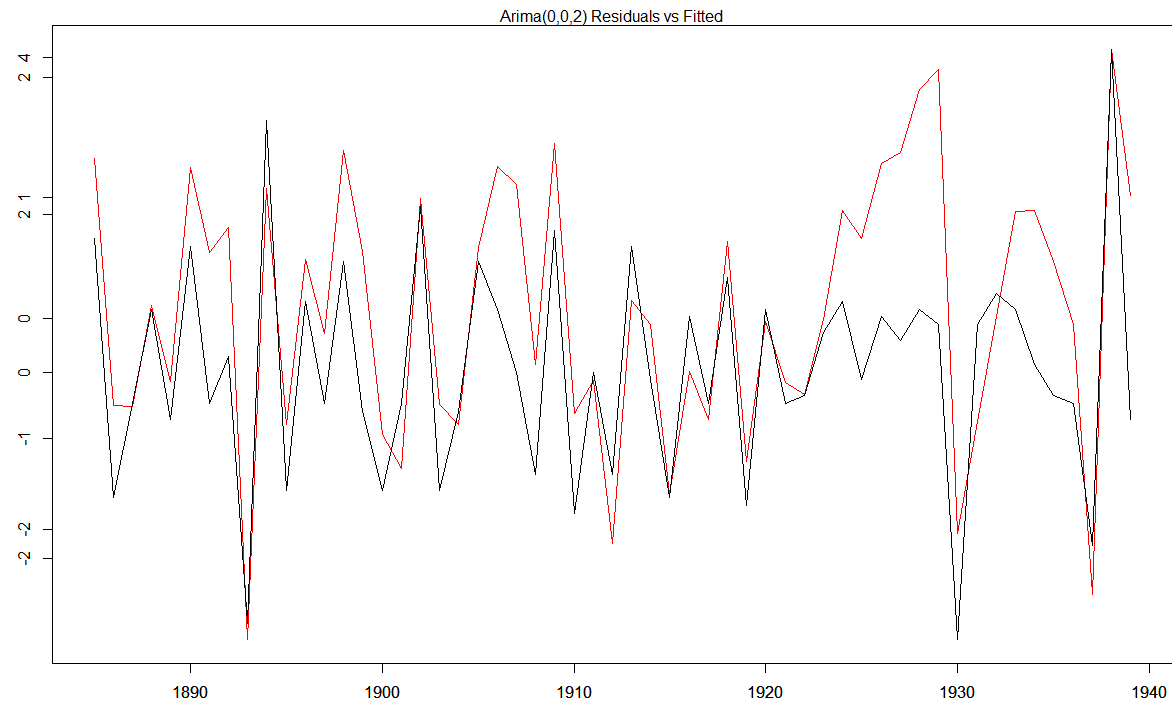
Arima 5



Arima 6



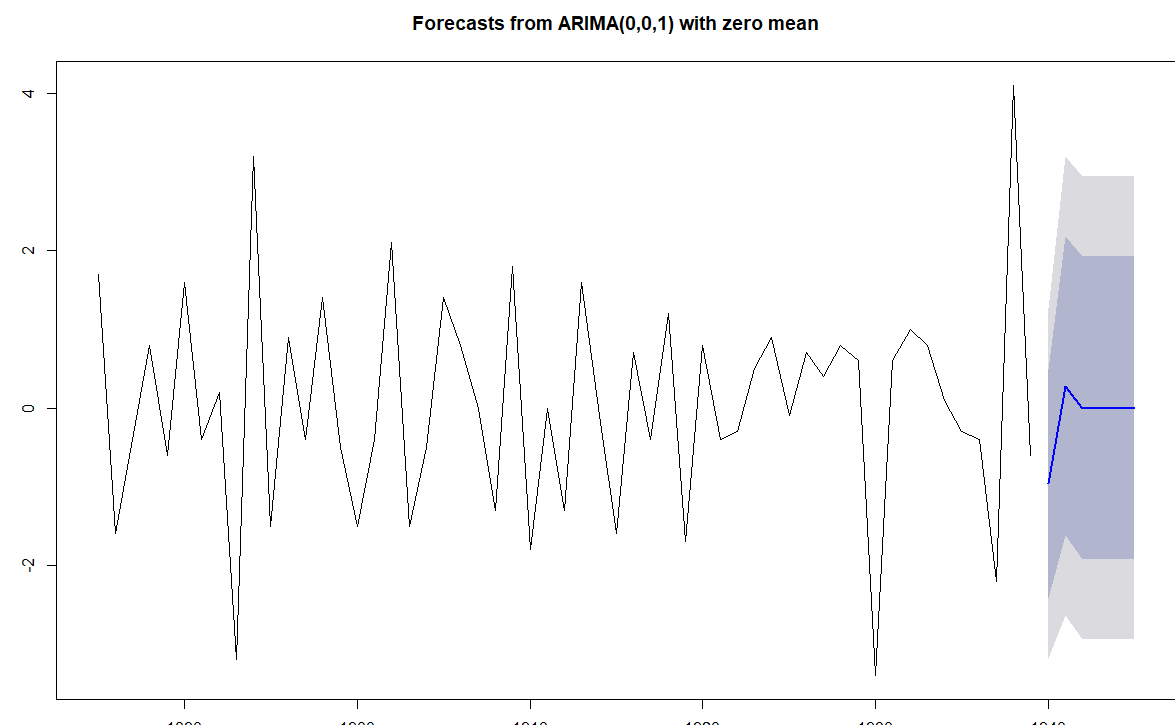
Arima 7



* 1. Pick one of the models as your favorite and tell me why you like that one the best.

I think the best fitting model is ARIMA 4, or the first order moving average. It has the lowest AIC and BIC, and is the only model not drastically off for the 1920-1930’s time period.

* 1. Forecast out your favorite model for the next 6 time periods and plot your time series plus the forecasted data. Does it look good or funky?

****

I think it looks funky, the forecast is pretty much the mean, and the confidence intervals, are almost the highest and lowest peaks and valleys. The model prediction really is not telling us much.

**Now switch to data set #2 – the one with seasonality in it. Perform the following steps.**

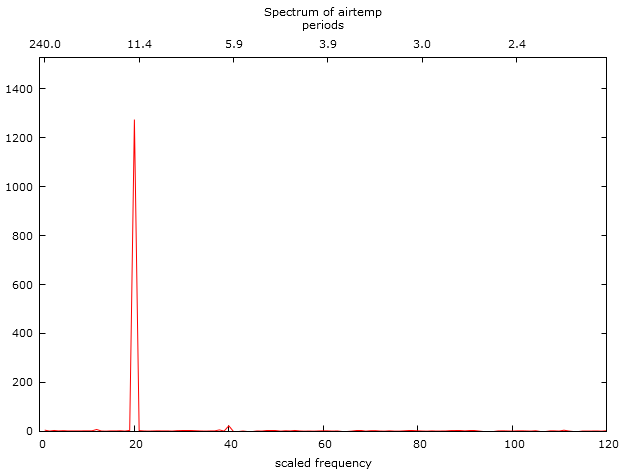
1. **For the time series data set with seasonality** **– start with the raw data before differencing here please!!!**
   1. Plot out the time series and suggest whether a type 1, type 2 or type 3 Holt Winter model should be applied and why.

I think my data is type 2 or double exponential smoothing because there is definitely trend and the data is seasonal. However there appears to be a mini spike in the valley, some years and not others, but the valleys and peaks do not appear to be directly multiples of one and other (like in type 3).

* 1. Eyeball the size of the period. What do you think it might be? Why is that?

The period is definitely 12, or monthly data represented yearly. The valleys or lows appear to be coinciding with the winters, and the peaks with the summers. It is air temperature data, so its most likely in the northern hemisphere; and gives validity to a 12 frequency period.

* 1. Use GRETL to do a periodogram for the data. What does the periodogram suggest might be the period length for the data?

****

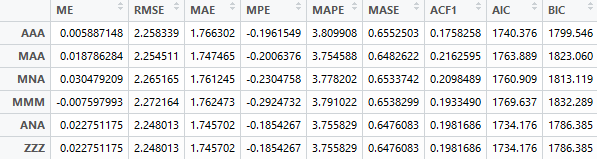
The periodogram suggest a frequency of slightly above 11.4 as indicated by the large peak, slightly to the left of 11.4; The scaled frequency on the bottom of the x-axis, is 20. This makes sense as I know my data is seasonal and I have yearly swings, over a 20 year period. There is a mini-peak around 5.9 and 40 on the bottom of the x-axis, indicating 2 main seasonal points, a high and a low.

* 1. If a type 2 or type 3 model, then apply a KPSS or ADF test to test for trend.

I ran two KPSS tests and two ADF test with different libraries. All four test unanimously show the temp data is stationary and there is a trend.

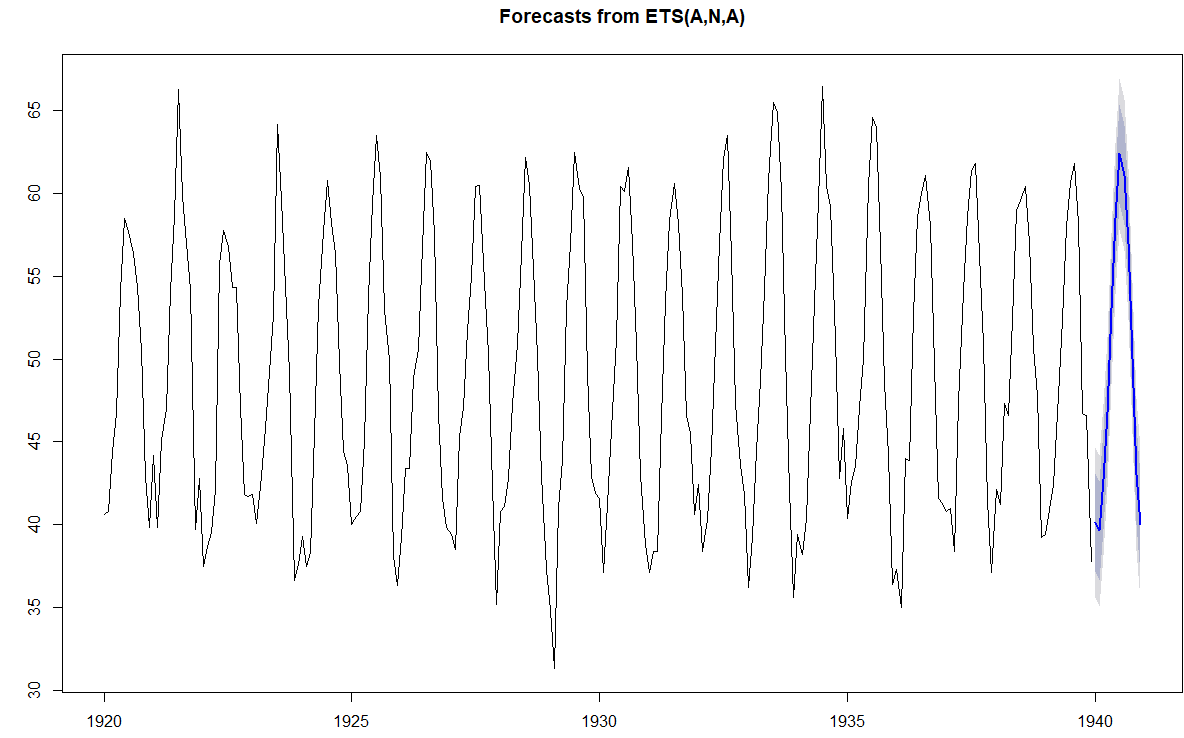
* 1. Decide the weights you will use for the three components of Winter Holt smoothing – constant, trend and seasonality – why these values? If you are using SAS then read about these weights in the proc docs, otherwise fish around in the R or Python docs.

I used the ets function, testing different combinations of “M”, “A”, “N”. “M” == Multiplicative, “A” == Additive, “N” == None, and “Z” == Unknown (the model picks)



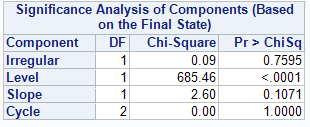
The two best models are the AAA and the ANA, which implies our data is additive and not multiplicative as the residuals stay constant and do not increase with time. The ANA gives the best RMSE, AIC and BIC implying it is the best model. The ANA model gives an alpha of 0.04214 and gamma of 0.0001 with no weight for trend.

* 1. Run the Holt Winter model and then using sgplot or your other favorite plotting poison plot the actual data and the fitted/forecast data on the same graph. How did the Holt-Winter model do in terms of forecasting?

****

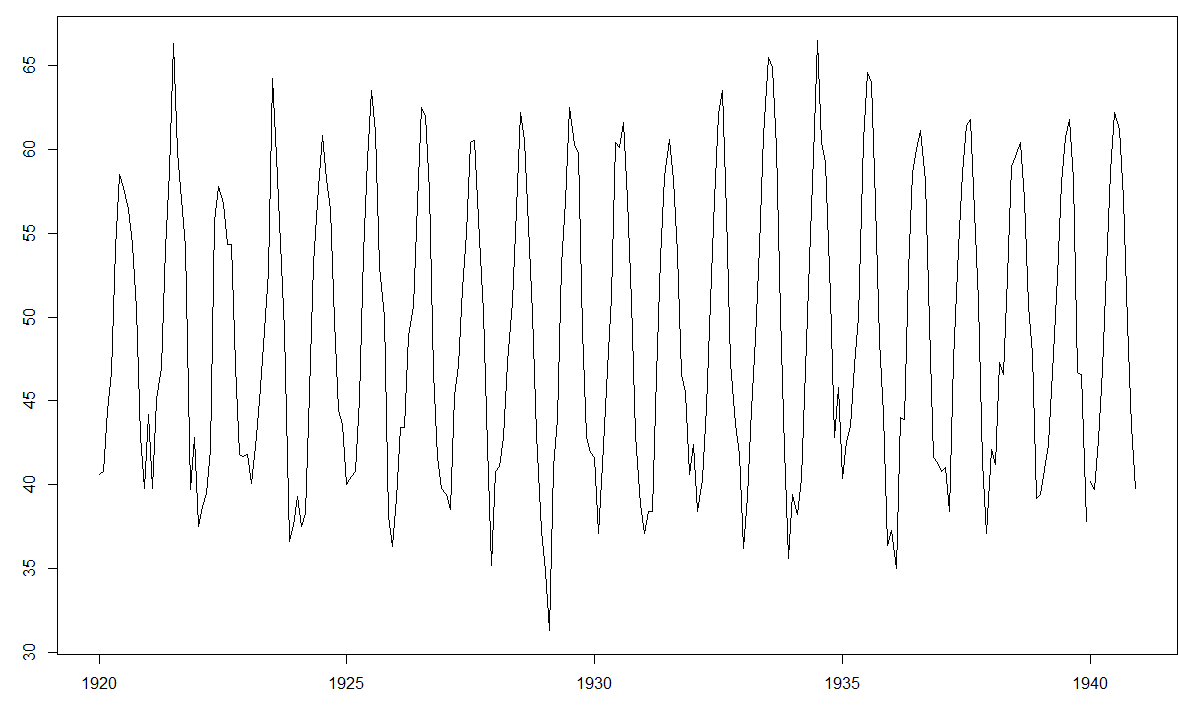
I think my model did reasonably well, my forecast seems to be roughly an average of the last 5 years, which makes sense because more weight is placed on the years closer to the forecasts. I am confident that my model did considerably well forecasting one year into the future.

* 1. Next run the same data set using the Unobserved Components Model time series analytic technique. Interpret the significance analysis of components table (based on final components in terms of trend, irregular (ARMA) and seasonality components in the data set – that is, which components are statistically significant?

****

The only component that appears to be statistically significant is the error component (level). The trend (Slope), irregular and seasonality (cycle) are not significant with p-values above 0.05

* 1. Have the UCM model produce fitted values for the existing data and forward 12 periods into the future and plot the original time series as well as the fitted/forecast data as well.

****

**Appendix >>> R:**

## Reading in my dataset >> Barley production from 1884 to 1939 >> Dtatset:

#https://datamarket.com/data/set/22rj/annual-barley-yields-per-acre-in-england-wales-1884-1939#!ds=22rj&display=line

barley\_test <- read.table("C:/Users/Mike/Downloads/annual-barley-yields-per-acre-in.csv",

nrows =56, header = TRUE, sep = ",", row.names = 1)

##### 2 >> Converting to Time Series

barley\_ts <- ts(barley\_test, frequency = 1, start=c(1884))

plot.ts(barley\_ts, type="l")

##### 3 >> Plotting the acf for barley\_ts

par(mfrow=c(1,1)) # So I only get one plot at a time

#barley\_acf <- acf(barley\_test$barley\_yields)

acf(barley\_test$barley\_yields)

pacf(barley\_test$barley\_yields)

##### 4 >> KPSS and ADF test

library(tseries)

library(urca)

#KPSS Test

barley\_kpss <- kpss.test(barley\_test$barley\_yields, null="Trend");barley\_kpss

barley\_kpss$statistic; ## Reject Null >> barley\_kpss$statistic is larger than p-value

barley\_kpss\_urca <- ur.kpss(barley\_test$barley\_yields, type="tau")#;barley\_kpss\_urca

summary(barley\_kpss\_urca) ## Reject Null at .1 and .05 sig levels

#ADF test

barley\_adf <- adf.test(barley\_test$barley\_yields, alternative = "stationary");barley\_adf # Fail to reject H0

barley\_adf\_urca <- ur.df(barley\_test$barley\_yields)

summary(barley\_adf\_urca) # Reject the Null Hypo

##### 5 >> Differencing the Data

barley\_diff <- diff(barley\_test$barley\_yields);barley\_diff

class(barley\_diff)

year <- 1885:1939

# Creating a new data.frame with our differenced data

barley\_diff\_df <- data.frame(barley\_diff, row.names = year)

#barley\_diff <- diff(barley\_test);barley\_diff

### a >> Plotting the differenced data

barley\_diff\_ts <- ts(barley\_diff, frequency = 1, start = c(1885))

plot.ts(barley\_diff\_ts, type="l")

### b >> Plotting ACF for diffeenced time series

acf(barley\_diff\_ts)

acf(barley\_diff\_df$barley\_diff) ## This is identical to the previous >> I just did double work

### c >> Applying KPSS and ADF to the barley\_diff\_df (or barley\_diff\_ts >> its the same thing)

barley\_diff\_kpss <- kpss.test(barley\_diff\_df$barley\_diff, null="Trend");

barley\_diff\_kpss$statistic ## Fail to reject the H0

barley\_diff\_kpss\_urca <- ur.kpss(barley\_diff\_df$barley\_diff, type="tau")#;barley\_kpss\_urca

summary(barley\_diff\_kpss\_urca) # Fail to reject the Null Hypo

barley\_diff\_adf <- adf.test(barley\_diff\_df$barley\_diff, alternative = "stationary");barley\_diff\_adf # reject the H0

barley\_diff\_adf\_urca <- ur.df(barley\_diff\_df$barley\_diff)

summary(barley\_diff\_adf\_urca) # reject the H0

##### 6 >> ARCH test >> Acutally performed in GRETL

#write.csv(barley\_diff\_ts, "C:/Users/Mike/Documents/Practicum 1/barley\_diff\_ts.csv")#, sep=",")

#write.csv(barley\_diff\_df, "C:/Users/Mike/Documents/Practicum 1/barley\_diff\_df.csv")

##### 7 >> PACF (diff data)

par(mfrow=c(2,1))

pacf(barley\_diff\_ts)

acf(barley\_diff\_ts)

##### 8 >> ARIMA models

library(lmtest)

## Very handy function, I found to compute the aicc >> I did not create this I found it

aicc = function(model){

n = model$nobs

p = length(model$coef)

aicc = model$aic + 2\*p\*(p+1)/(n-p-1)

return(aicc)

}

barley\_arima\_1 <- arima(barley\_diff\_ts, order = c(1,0,0), seasonal = list(order=c(1,0,0), period=1), include.mean=FALSE)

barley\_arima\_1\_info <- summary(barley\_arima\_1)

attributes(barley\_arima\_1)

barley\_arima\_1$aic

barley\_arima\_1$AICc

AIC(barley\_arima\_1)

BIC(barley\_arima\_1)

AIC(barley\_arima\_1, k =log(length(barley\_diff\_ts)))

coeftest(barley\_arima\_1)

confint(barley\_arima\_1)

fitarima(barley\_arima\_1)

barley\_arima\_1$

barley\_arima\_1$aic

BIC(barley\_arima\_1)

aicc(barley\_arima\_1)

barley\_arima\_2 <- arima(barley\_diff\_ts, order = c(1,1,0), seasonal = list(order=c(1,1,0), period=1), include.mean=FALSE)

barley\_arima\_2\_info <- summary(barley\_arima\_2)

barley\_arima\_3 <- arima(barley\_diff\_ts, order = c(1,1,1), seasonal = list(order=c(1,1,1), period=1), include.mean=FALSE)

barley\_arima\_3\_info <- summary(barley\_arima\_3)

barley\_arima\_4 <- arima(barley\_diff\_ts, order = c(0,0,1), seasonal = list(order=c(0,0,1), period=1), include.mean=FALSE)

barley\_arima\_4\_info <- summary(barley\_arima\_4)

barley\_arima\_5 <- arima(barley\_diff\_ts, order = c(1,1,2), seasonal = list(order=c(1,1,2), period=1), include.mean=FALSE)

barley\_arima\_5\_info <- summary(barley\_arima\_5)

barley\_arima\_6 <- arima(barley\_diff\_ts, order = c(2,1,2), seasonal = list(order=c(2,1,2), period=1), include.mean=FALSE)

barley\_arima\_6\_info <- summary(barley\_arima\_6)

barley\_arima\_7 <- arima(barley\_diff\_ts, order = c(0,0,2), seasonal = list(order=c(0,0,2), period=1), include.mean=FALSE)

barley\_arima\_7\_info <- summary(barley\_arima\_7)

barley\_arima\_7$aic

BIC(barley\_arima\_7)

aicc(barley\_arima\_7)

arima\_aic <- c(AIC(barley\_arima\_1), AIC(barley\_arima\_2), AIC(barley\_arima\_3), AIC(barley\_arima\_4), AIC(barley\_arima\_5), AIC(barley\_arima\_6), barley\_arima\_7$aic)

arima\_aicc <- c(aicc(barley\_arima\_1), aicc(barley\_arima\_2), aicc(barley\_arima\_3), aicc(barley\_arima\_4), aicc(barley\_arima\_5), aicc(barley\_arima\_6), aicc(barley\_arima\_7))

arima\_bic <- c(BIC(barley\_arima\_1), BIC(barley\_arima\_2), BIC(barley\_arima\_3), BIC(barley\_arima\_4), BIC(barley\_arima\_5), BIC(barley\_arima\_6), BIC(barley\_arima\_7))

arima\_metrics <- data.frame(arima\_aic, arima\_aicc, arima\_bic)

colnames(arima\_metrics) <- c("AIC", "AICc", "BIC")

rownames(arima\_metrics) <- c("ARIMA(1,0,0)","ARIMA(1,1,0)", "ARIMA(1,1,1)", "ARIMA(0,0,1)", "ARIMA(1,1,2)", "ARIMA(2,1,2)", "ARIMA(0,0,2)")

### B >> Plotting observed vs fitted data

par(mfrow=c(1,1))

# Arima 1

plot(barley\_arima\_1$residuals, type="l", col="red")

par(new=TRUE)

plot(barley\_diff\_ts, type="l", col = "black")

mtext("Arima(1,0,0) Residuals vs Fitted")

# Arima 2

plot(barley\_arima\_2$residuals, type="l", col="red")

par(new=TRUE)

plot(barley\_diff\_ts, type="l", col = "black")

mtext("Arima(1,1,0) Residuals vs Fitted")

# Arima 3

plot(barley\_arima\_3$residuals, type="l", col="red")

par(new=TRUE)

plot(barley\_diff\_ts, type="l", col = "black")

mtext("Arima(1,1,1) Residuals vs Fitted")

# Arima 4

plot(barley\_arima\_4$residuals, type="l", col="red")

par(new=TRUE)

plot(barley\_diff\_ts, type="l", col = "black")

mtext("Arima(0,0,1) Residuals vs Fitted")

# Arima 5

plot(barley\_arima\_5$residuals, type="l", col="red")

par(new=TRUE)

plot(barley\_diff\_ts, type="l", col = "black")

mtext("Arima(1,1,2) Residuals vs Fitted")

# Arima 6

plot(barley\_arima\_6$residuals, type="l", col="red")

par(new=TRUE)

plot(barley\_diff\_ts, type="l", col = "black")

mtext("Arima(2,1,2) Residuals vs Fitted")

# Arima 7

plot(barley\_arima\_7$residuals, type="l", col="red")

par(new=TRUE)

plot(barley\_diff\_ts, type="l", col = "black")

mtext("Arima(0,0,2) Residuals vs Fitted")

### D >> Forcasting with ARIMA model 4

library(forcats)

forecast\_arima\_4 <- forecast(barley\_arima\_4, h = 6)

plot(forecast\_arima\_4)

################### Dataset 2

## Air temperatures >> https://datamarket.com/data/set/22li/mean-monthly-air-temperature-deg-f-nottingham-castle-1920-1939#!ds=22li&display=line

temp <- read.csv("C:/Users/Mike/Downloads/mean-monthly-air-temperature-deg.csv", header=TRUE, row.names = 1)

tem <- read.csv("C:/Users/Mike/Downloads/mean-monthly-air-temperature-deg.csv", header=TRUE)

##### 9 - A >> Plot time series

temp\_ts <- ts(temp, frequency = 12, start=c(1920))

ts.plot(temp\_ts, type="l")

##### D >> Applying the KPSS and ADF test

temp\_kpss <- kpss.test(temp\_ts, null="Trend");temp\_kpss # Fail to reject the H0, stat less than p-value

temp\_kpss\_urca <- ur.kpss(temp\_ts, type="tau")

summary(temp\_kpss\_urca) # Fail to reject the H0, test-stat less than any critical value

temp\_adf <- adf.test(temp\_ts, alternative = "stationary");temp\_adf # Reject the H0 >>

temp\_adf\_urca <- ur.df(temp$air.temp)

summary(temp\_adf\_urca) # Reject H0 >> Astronomically low P-value

##### E >> Weights for Holt Winters Smoothing

temp\_hw\_smooth\_1 <- ets(temp\_ts, model = "AAA")

temp\_hw\_smooth\_1\_error <- summary(temp\_hw\_smooth\_1)

checkresiduals(temp\_hw\_smooth\_1)

temp\_hw\_smooth\_2 <- ets(temp\_ts, model = "MAA")

temp\_hw\_smooth\_2\_error <- summary(temp\_hw\_smooth\_2)

checkresiduals(temp\_hw\_smooth\_2)

temp\_hw\_smooth\_3 <- ets(temp\_ts, model = "MNA")

temp\_hw\_smooth\_3\_error <- summary(temp\_hw\_smooth\_3)

checkresiduals(temp\_hw\_smooth\_3)

temp\_hw\_smooth\_4 <- ets(temp\_ts, model = "MMM")

temp\_hw\_smooth\_4\_error <- summary(temp\_hw\_smooth\_4)

checkresiduals(temp\_hw\_smooth\_4)

temp\_hw\_smooth\_5 <- ets(temp\_ts, model = "ANA")

temp\_hw\_smooth\_5\_error <- summary(temp\_hw\_smooth\_5)

checkresiduals(temp\_hw\_smooth\_5)

temp\_hw\_smooth\_6 <- ets(temp\_ts, model = "ZZZ")

temp\_hw\_smooth\_6\_error <- summary(temp\_hw\_smooth\_6)

checkresiduals(temp\_hw\_smooth\_6)

error\_row <- NULL

temp\_hw\_error\_df <- rbind(temp\_hw\_smooth\_1\_error, temp\_hw\_smooth\_2\_error, temp\_hw\_smooth\_3\_error, temp\_hw\_smooth\_4\_error, temp\_hw\_smooth\_5\_error, temp\_hw\_smooth\_6\_error)

temp\_hw\_aic <- rbind(temp\_hw\_smooth\_1$aic, temp\_hw\_smooth\_2$aic, temp\_hw\_smooth\_3$aic, temp\_hw\_smooth\_4$aic, temp\_hw\_smooth\_5$aic, temp\_hw\_smooth\_6$aic)

temp\_hw\_bic <- rbind(temp\_hw\_smooth\_1$bic, temp\_hw\_smooth\_2$bic, temp\_hw\_smooth\_3$bic, temp\_hw\_smooth\_4$bic, temp\_hw\_smooth\_5$bic, temp\_hw\_smooth\_6$bic)

colnames(temp\_hw\_aic) <- "AIC"; colnames(temp\_hw\_bic) <- "BIC"

temp\_hw\_error\_df <- cbind(temp\_hw\_error\_df, temp\_hw\_aic, temp\_hw\_bic)

rownames(temp\_hw\_error\_df) <- c("AAA", "MAA", "MNA", "MMM", "ANA", "ZZZ")

temp\_hw\_smooth\_1$par

temp\_hw\_smooth\_5$par

##### F >> Forecasting with the ANA Holt winters model

temp\_hw\_smooth\_5\_forecast <- forecast(temp\_hw\_smooth\_5, h = 12)

plot(temp\_hw\_smooth\_5\_forecast)

##### H >> UCM

library(rucm)

temp\_ucm <- ucm(temp\_ts~0, temp\_ts, level=TRUE, season=TRUE, season.length=12);temp\_ucm

temp\_ucm <- ucm(temp\_ts~0, temp\_ts, season.length = 12, season=TRUE, slope = TRUE)

temp\_ucm$est.var.season

temp\_ucm$est.var.slope

temp\_ucm$est.var.level

temp\_ucm$irr.var[1]

temp\_ucm\_stats <- data.frame(temp\_ucm$est.var.season, temp\_ucm$est.var.slope, temp\_ucm$irr.var[1])

rownames(temp\_ucm\_stats) <- "temp\_ucm"

colnames(temp\_ucm\_stats) <- c("seasonality", "trend", "irregular")

temp\_ucm\_stats

temp\_ucm\_predict <- predict(temp\_ucm$model, n.ahead = 12)

ts.plot(temp\_ts, temp\_ucm\_predict)

time\_col <- 1:240

temp\_df <- data.frame(time\_col, temp)

colnames(temp\_df) <- c("time", "temp")

tem$test <- tem$Month

tem$test <-paste(tem$test,"-01", sep="")

colnames(tem) <- c("time", "temp", "date")

#write.csv(tem, "C:/Users/Mike/Documents/Practicum 1/temp\_df\_sas.csv")

Appendix >> SAS!!

**proc** **ucm** data=WORK.TEMP\_DF\_SAS\_0000;

id date interval=month;

model temp;

irregular;

level;

slope;

cycle;

forecast lead=**12** plot=decomp;

**run**;