Unit 12 For Live Session

Chad Madding

March 22, 2020

## Question Of Interest

The data set (LA\_Cmort\_Study.csv) is a portion of the data taken from a study (Shumway 1988) on the possible effects of pollution and temperature on weekly cardiac mortality (heart attacks) in Los Angeles County.

**Your goal is to utilize all given information to provide the most useful forecasts for the next 20 weeks of cardiac mortality.**

You should include plots, tables, and charts to help make your analysis and inferences clear to your peers.

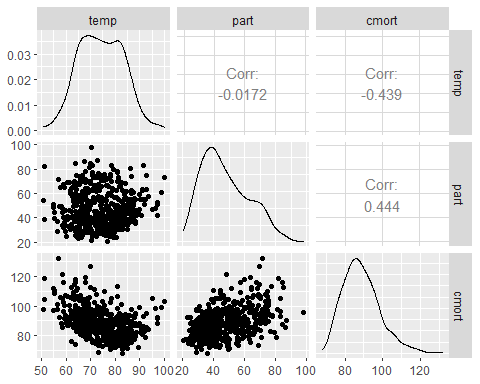
Start by reading in the data and see what it looks like.

CM = read.csv("la\_cmort\_study.csv", header = TRUE)  
#Look at the first few rows  
head(CM)

## Week temp part cmort  
## 1 1 72.38 72.72 97.85  
## 2 2 67.19 49.60 104.64  
## 3 3 62.94 55.68 94.36  
## 4 4 72.49 55.16 98.05  
## 5 5 74.25 66.02 95.85  
## 6 6 67.88 44.01 95.98

## Visualaze The Data

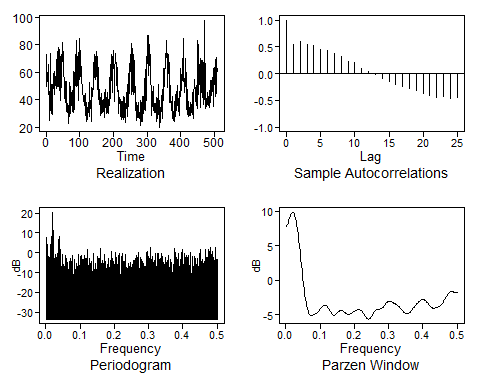
Make a matrix of plots with the LA\_Cmort data set.



#### ARIMA 1 MLR with Cor Errors (no lag, no seasonl categorical variable)

First we will forecast the Particles (part). The $freq show a peek around 0.0192 showing some sign of a 52 week annual trend.

#plot just the Particles (part)  
plotts.sample.wge(CM$part) #peek in freq near .0192/1=52.08 (annual)



Using multivariate to forecast

CMnoX=CM[2:4]  
#install.packages("vars")  
# VAR and VARselect are from CRAN package vars  
VARselect(CMnoX, lag.max = 15, type = "const",season = NULL, exogen = NULL)

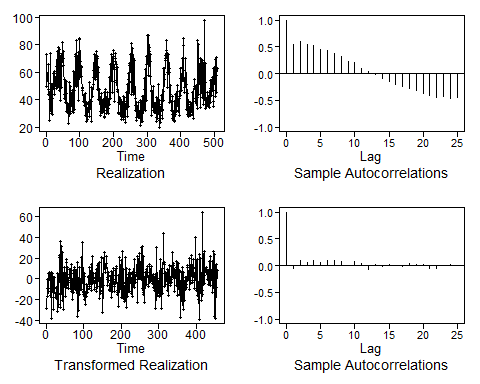
## $selection  
## AIC(n) HQ(n) SC(n) FPE(n)   
## 9 5 2 9   
##   
## $criteria  
## 1 2 3 4 5 6  
## AIC(n) 11.83703 11.35715 11.32459 11.28105 11.22936 11.20971  
## HQ(n) 11.87717 11.42740 11.42495 11.41152 11.38994 11.40039  
## SC(n) 11.93927 11.53607 11.58020 11.61334 11.63834 11.69536  
## FPE(n) 138278.93633 85575.42617 82834.75080 79307.32148 75314.81580 73852.10868  
## 7 8 9 10 11 12  
## AIC(n) 11.21438 11.19428 11.17282 11.17835 11.18313 11.17328  
## HQ(n) 11.43517 11.44518 11.45383 11.48947 11.52436 11.54461  
## SC(n) 11.77672 11.83330 11.88852 11.97074 12.05221 12.11903  
## FPE(n) 74202.71501 72731.98582 71195.36032 71599.87436 71954.70797 71262.38608  
## 13 14 15  
## AIC(n) 11.17307 11.17335 11.17635  
## HQ(n) 11.57451 11.60490 11.63801  
## SC(n) 12.19551 12.27247 12.35215  
## FPE(n) 71263.64357 71302.45977 71537.90535

#VARselect picks p=2 (using BIC)  
lsfit=VAR(CMnoX,p=2,type="const")  
preds=predict(lsfit,n.ahead=20)  
# preds$fcst$ are the VAR forecasts for x1. Similar for x2  
#this will list out just the forcast  
preds$fcst$temp

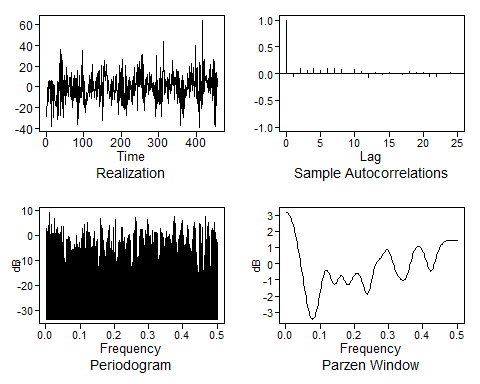
## fcst lower upper CI  
## [1,] 71.25383 59.19228 83.31538 12.06155  
## [2,] 69.90957 57.54192 82.27723 12.36765  
## [3,] 70.19691 56.98563 83.40820 13.21129  
## [4,] 69.65798 56.07989 83.23608 13.57810  
## [5,] 69.77733 55.74554 83.80912 14.03179  
## [6,] 69.64721 55.23042 84.06400 14.41679  
## [7,] 69.75578 54.95005 84.56152 14.80574  
## [8,] 69.81715 54.65589 84.97841 15.16126  
## [9,] 69.96410 54.47041 85.45778 15.49368  
## [10,] 70.11506 54.32006 85.91006 15.79500  
## [11,] 70.29838 54.23133 86.36544 16.06705  
## [12,] 70.49071 54.18112 86.80030 16.30959  
## [13,] 70.69483 54.17058 87.21908 16.52425  
## [14,] 70.90282 54.19003 87.61560 16.71279  
## [15,] 71.11275 54.23547 87.99004 16.87729  
## [16,] 71.32107 54.30113 88.34102 17.01994  
## [17,] 71.52578 54.38283 88.66872 17.14294  
## [18,] 71.72492 54.47651 88.97334 17.24842  
## [19,] 71.91718 54.57878 89.25557 17.33839  
## [20,] 72.10147 54.68670 89.51624 17.41477

Transform the Particles (part) to remove the seasonality.

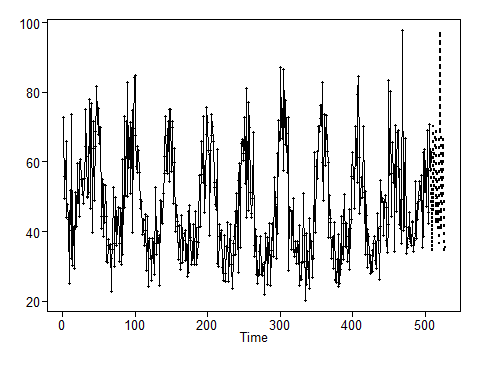
CM\_52 = artrans.wge(CM$part, c(rep(0,51),1))



plotts.sample.wge(CM\_52) #looks like some low freq?

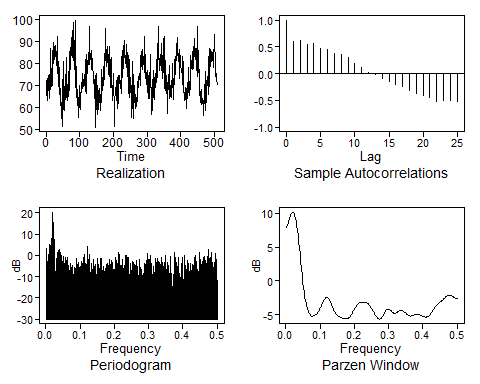


aic5.wge(CM\_52) #picks ARMA(2,1) assume stationary  
aic5.wge(CM\_52,type = "bic") #picks ARMA(2,1)   
ljung.wge(CM\_52)$pval #FTR Ho  
ljung.wge(CM\_52, K = 48)$pval #FTR Ho  
#Going with white noise despite peak at 0 in Spec D.   
#est = est.arma.wge(CM\_52, p = 3, q = 2)  
#CM\_52\_AR2\_MA1 = artrans.wge(CM\_52,est$phi)  
predsPart = fore.aruma.wge(CM$part,s = 52, n.ahead = 20, limits = F)

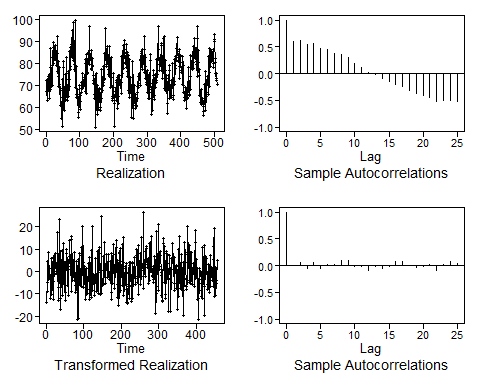


Next we will forecast temperature (Temp).

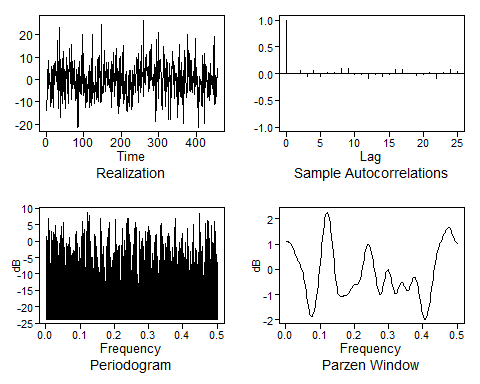
#forecast Temp  
plotts.sample.wge(CM$temp) #freq near .0192 (annual)



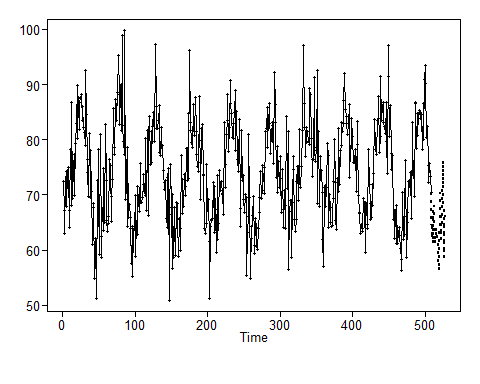
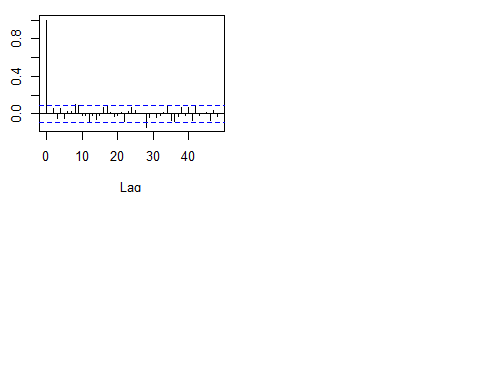
CM\_52 = artrans.wge(CM$temp, c(rep(0,51),1))



plotts.sample.wge(CM\_52) #looks like some low freq?



aic5.wge(CM\_52) #picks ARMA(0,0)  
aic5.wge(CM\_52,type = "bic") #picks ARMA(0,0)   
ljung.wge(CM\_52)$pval  
ljung.wge(CM\_52, K = 48)$pval #barely rejects  
acf(CM\_52,lag.max = 48) # acf looks consistent with white noise  
predsTemp = fore.aruma.wge(CM$temp,s = 52, n.ahead = 20, limits = F)



Looking at cardiac mortality (heart attacks) in Los Angeles County based on temperature, particles and Week.

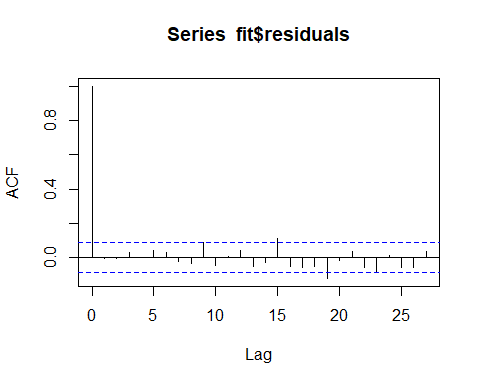
# Model cmort based on predicted part and temp using MLR with Cor Erros  
ksfit = lm(cmort~temp+part+Week, data = CM)  
phi = aic.wge(ksfit$residuals)

Fit an ARIMA with the phi from above (2). Binding the temperature, particle and week data with cbind.

fit = arima(CM$cmort,order = c(phi$p,0,0), seasonal = list(order = c(1,0,0), period = 52), xreg = cbind(CM$temp, CM$part, CM$Week))

Visually the acf is showing white noise. Now we need to check for whiteness of residuals with the Ljung-Box Test.

#First visualize the residuals  
acf(fit$residuals)



#Run the test with the default 24 and 48 maximum lag for sample autocorrelations to be used.  
ljung.wge(fit$residuals) # pval = .048

## Obs -0.008773997 -0.006778 0.02847037 0.001592821 0.04173621 0.03014614 -0.02519791 -0.03614844 0.08896799 -0.04873154 0.003349583 0.04272511 -0.05378803 -0.03121467 0.1104465 -0.0511791 -0.05781976 -0.05203572 -0.1226383 -0.01661192 0.03274971 -0.06125045 -0.08319911 0.009083812

## $test  
## [1] "Ljung-Box test"  
##   
## $K  
## [1] 24  
##   
## $chi.square  
## [1] 36.59418  
##   
## $df  
## [1] 24  
##   
## $pval  
## [1] 0.04801085

ljung.wge(fit$residuals, K = 48) # pval = .002

## Obs -0.008773997 -0.006778 0.02847037 0.001592821 0.04173621 0.03014614 -0.02519791 -0.03614844 0.08896799 -0.04873154 0.003349583 0.04272511 -0.05378803 -0.03121467 0.1104465 -0.0511791 -0.05781976 -0.05203572 -0.1226383 -0.01661192 0.03274971 -0.06125045 -0.08319911 0.009083812 -0.06018696 -0.05789479 0.03458024 -0.1371038 -0.02600092 -0.07277285 -0.03981533 -0.03157352 0.009442916 -0.06678764 -0.05777429 -0.004599118 -0.01492029 -0.03032822 -0.004028157 -0.06378089 0.01029915 0.03886128 -0.006943152 0.04753463 0.04475963 0.0230044 0.1033364 0.110844

## $test  
## [1] "Ljung-Box test"  
##   
## $K  
## [1] 48  
##   
## $chi.square  
## [1] 79.51156  
##   
## $df  
## [1] 48  
##   
## $pval  
## [1] 0.002852882

Load the forecast Part and Temp data in a data frame. This will extend the week from 509 to 528 with **seq**.

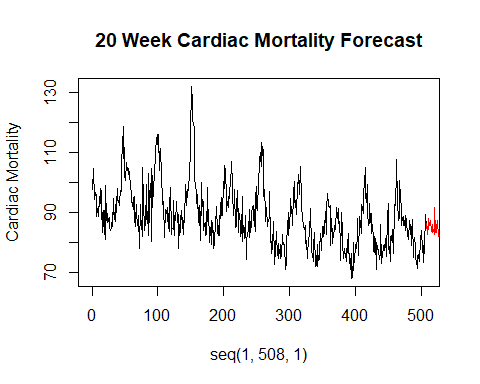
next20 = data.frame(temp = predsTemp$f, part = predsPart$f, Week = seq(509,528,1))

This will get the predictions

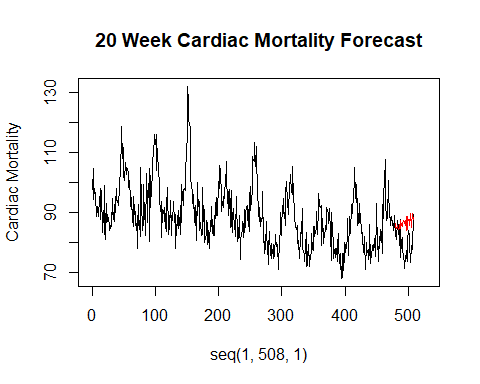
#get predictions  
predsCMort = predict(fit,newxreg = next20)

plot next 20 cmort wrt time

plot(seq(1,508,1), CM$cmort, type = "l",xlim = c(0,508), ylab = "Cardiac Mortality", main = "20 Week Cardiac Mortality Forecast")  
lines(seq(509,528,1), predsCMort$pred, type = "l", col = "red")



#Find ASE Need to forecast last 30 of known series using dplyr lag.   
CMsmall = CM[1:478,]  
CMsmall$temp\_1 = dplyr::lag(CMsmall$temp,1)  
CM$temp\_1 = dplyr::lag(CM$temp,1)  
ksfit = lm(cmort~temp\_1+part+Week, data = CMsmall)  
phi = aic.wge(ksfit$residuals)  
  
fit = arima(CMsmall$cmort,order = c(phi$p,0,0), seasonal = list(order = c(1,0,0), period = 52), xreg = cbind(CMsmall$temp, CMsmall$part, CMsmall$Week))  
  
last30 = data.frame(temp = CM$temp\_1[479:508], part = CM$part[479:508], Week = seq(479,508,1))  
#get predictions  
predsCMort = predict(fit,newxreg = last30)  
  
plot(seq(1,508,1), CM$cmort, type = "l",xlim = c(0,528), ylab = "Cardiac Mortality", main = "20 Week Cardiac Mortality Forecast")  
lines(seq(479,508,1), predsCMort$pred, type = "l", col = "red")

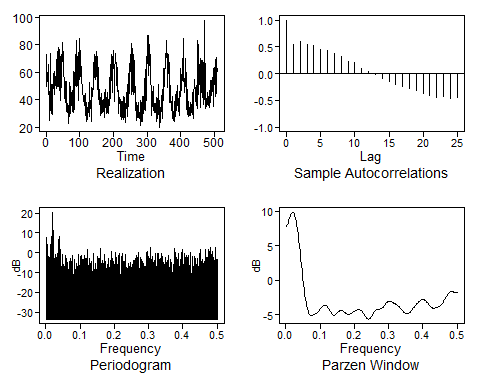


ASE\_ARMA1 = mean((CM$cmort[479:508] - predsCMort$pred)^2)  
ASE\_ARMA1

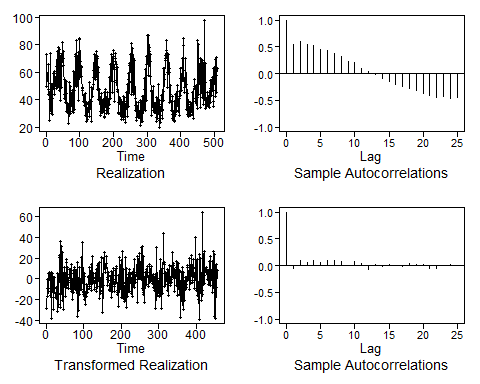
#### ARIMA2: attempt at categorical variable for week but arima takes only continuous variables

forecast Particles

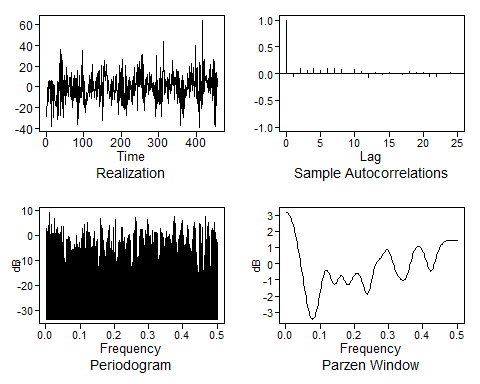
CM = read.csv("la\_cmort\_study.csv", header = TRUE)  
plotts.sample.wge(CM$part) #freq near .0192 (annual)



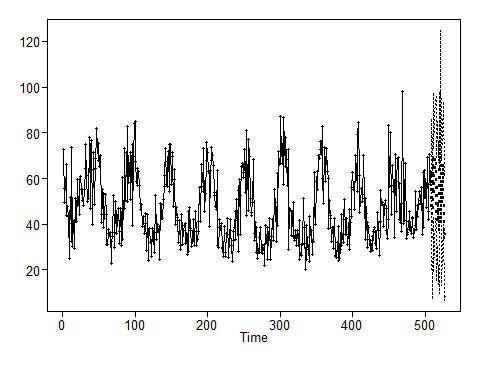
CM\_52 = artrans.wge(CM$part, c(rep(0,51),1))



plotts.sample.wge(CM\_52) #looks like some low freq?

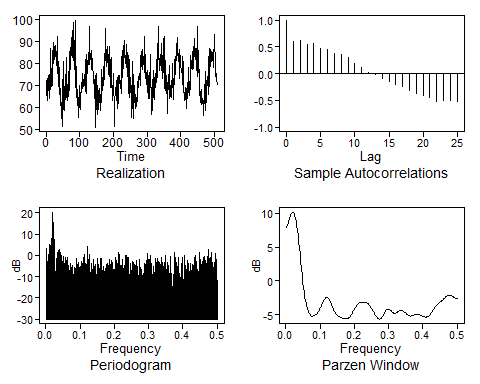


aic5.wge(CM\_52) #picks ARMA(2,1) assume stationary  
aic5.wge(CM\_52,type = "bic") #picks ARMA(0,0)   
ljung.wge(CM\_52)$pval #FTR Ho  
ljung.wge(CM\_52, K = 48)$pval #FTR Ho  
#Going with white noise despite peak at 0 in Spec D.   
#est = est.arma.wge(CM\_52, p = 3, q = 2)  
#CM\_52\_AR2\_MA1 = artrans.wge(CM\_52,est$phi)  
predsPart = fore.aruma.wge(CM$part,s = 52, n.ahead = 20)

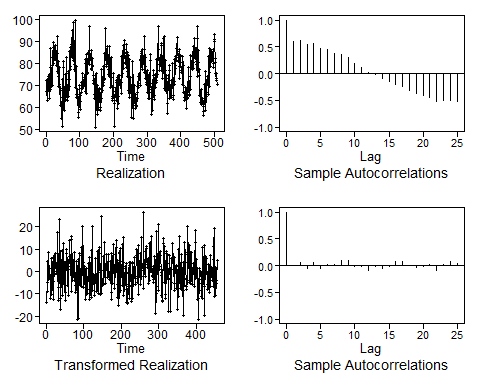


forecast Temp

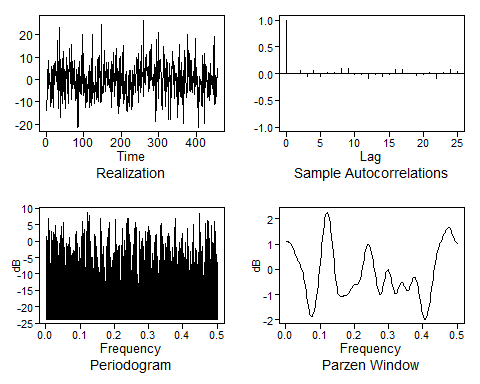
plotts.sample.wge(CM$temp) #freq near .0192 (annual)



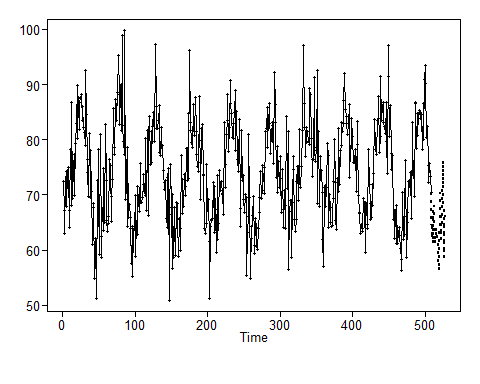
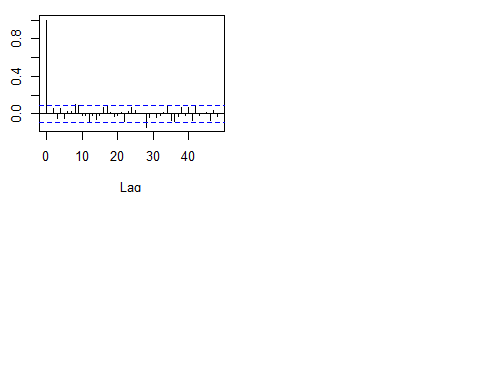
CM\_52 = artrans.wge(CM$temp, c(rep(0,51),1))



plotts.sample.wge(CM\_52) #looks like some low freq?



aic5.wge(CM\_52) #picks ARMA(0,0)  
aic5.wge(CM\_52,type = "bic") #picks ARMA(0,0)   
ljung.wge(CM\_52)$pval  
ljung.wge(CM\_52, K = 48)$pval #barely rejects  
acf(CM\_52,lag.max = 48) # acf looks consistent with white noise  
predsTemp = fore.aruma.wge(CM$temp,s = 52, n.ahead = 20, limits = F)

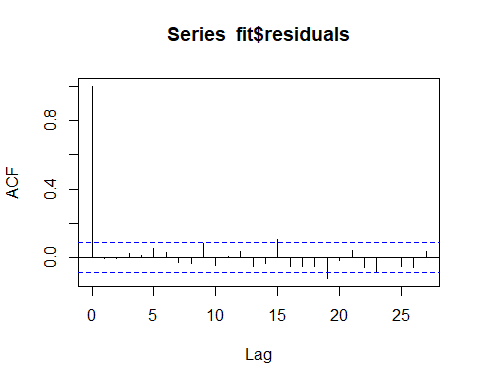


Model **cmort** based on predicted **part** and **temp** using Multiple Linear Regression (MLR) with correlation errors.

CM$FWeek = as.factor(CM$Week%%52)  
ksfit = lm(cmort~temp+part+Week+FWeek, data = CM)  
phi = aic.wge(ksfit$residuals)  
fit = arima(CM$cmort,order = c(phi$p,0,0), xreg = cbind(CM$temp, CM$part, CM$Week, CM$FWeek))

Check for whiteness of residuals with a visual test and the Ljung-Box Test

acf(fit$residuals)



ljung.wge(fit$residuals) # pval = .066

## Obs -0.006650986 -0.006552996 0.02276148 0.01126156 0.05088593 0.02941397 -0.02768429 -0.03359994 0.08138044 -0.04601947 0.003502337 0.03271662 -0.05237642 -0.03357099 0.1025354 -0.05214469 -0.05469956 -0.05537602 -0.1223643 -0.01733948 0.03854912 -0.0613006 -0.08496323 -0.00119778

## $test  
## [1] "Ljung-Box test"  
##   
## $K  
## [1] 24  
##   
## $chi.square  
## [1] 35.17006  
##   
## $df  
## [1] 24  
##   
## $pval  
## [1] 0.06591877

ljung.wge(fit$residuals, K = 48) # pval = .0058

## Obs -0.006650986 -0.006552996 0.02276148 0.01126156 0.05088593 0.02941397 -0.02768429 -0.03359994 0.08138044 -0.04601947 0.003502337 0.03271662 -0.05237642 -0.03357099 0.1025354 -0.05214469 -0.05469956 -0.05537602 -0.1223643 -0.01733948 0.03854912 -0.0613006 -0.08496323 -0.00119778 -0.05200078 -0.05592817 0.03251483 -0.1351858 -0.03402655 -0.0747738 -0.03617307 -0.03211621 -0.0003643746 -0.06623262 -0.06188258 -0.00565433 -0.005550979 -0.03445902 -0.006195876 -0.05997411 0.01078523 0.0299213 -0.00176937 0.03759249 0.03635303 0.01699562 0.1045691 0.1114073

## $test  
## [1] "Ljung-Box test"  
##   
## $K  
## [1] 48  
##   
## $chi.square  
## [1] 76.25112  
##   
## $df  
## [1] 48  
##   
## $pval  
## [1] 0.005835517

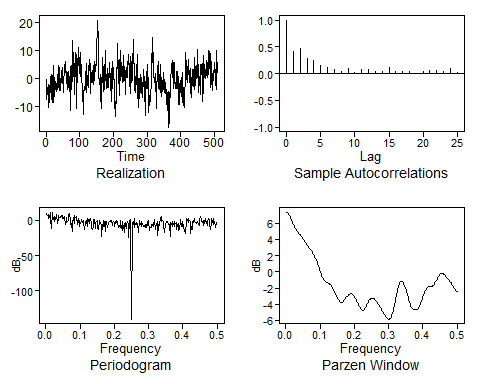
load the forecasted Part and Temp in a data frame

next20 = data.frame(temp = predsTemp$f, part = predsPart$f, Week = seq(509,528,1), FWeek = as.factor(seq(509,528,1)%%52))

get predictions - predict residuals

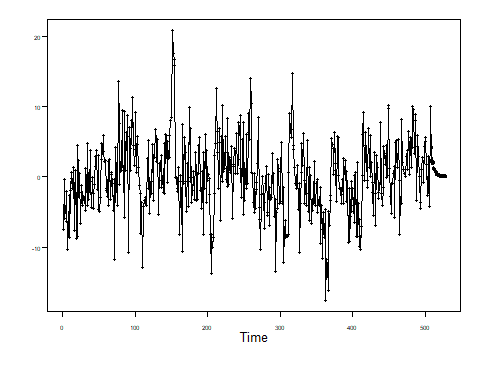
plotts.sample.wge(ksfit$residuals)

## Warning in plotts.sample.wge(ksfit$residuals): NaNs produced



## $autplt  
## [1] 1.000000000 0.421428548 0.467044098 0.291570176 0.244137951 0.157916276  
## [7] 0.120183538 0.078910436 0.037244352 0.104623249 0.014204075 0.078592862  
## [13] 0.081097096 0.033909834 0.038497668 0.111155979 0.034992164 0.048976564  
## [19] 0.033233122 0.004707779 0.049969797 0.060358759 0.050439274 0.044482940  
## [25] 0.092000915 0.030255425  
##   
## $freq  
## [1] 0.001968504 0.003937008 0.005905512 0.007874016 0.009842520 0.011811024  
## [7] 0.013779528 0.015748031 0.017716535 0.019685039 0.021653543 0.023622047  
## [13] 0.025590551 0.027559055 0.029527559 0.031496063 0.033464567 0.035433071  
## [19] 0.037401575 0.039370079 0.041338583 0.043307087 0.045275591 0.047244094  
## [25] 0.049212598 0.051181102 0.053149606 0.055118110 0.057086614 0.059055118  
## [31] 0.061023622 0.062992126 0.064960630 0.066929134 0.068897638 0.070866142  
## [37] 0.072834646 0.074803150 0.076771654 0.078740157 0.080708661 0.082677165  
## [43] 0.084645669 0.086614173 0.088582677 0.090551181 0.092519685 0.094488189  
## [49] 0.096456693 0.098425197 0.100393701 0.102362205 0.104330709 0.106299213  
## [55] 0.108267717 0.110236220 0.112204724 0.114173228 0.116141732 0.118110236  
## [61] 0.120078740 0.122047244 0.124015748 0.125984252 0.127952756 0.129921260  
## [67] 0.131889764 0.133858268 0.135826772 0.137795276 0.139763780 0.141732283  
## [73] 0.143700787 0.145669291 0.147637795 0.149606299 0.151574803 0.153543307  
## [79] 0.155511811 0.157480315 0.159448819 0.161417323 0.163385827 0.165354331  
## [85] 0.167322835 0.169291339 0.171259843 0.173228346 0.175196850 0.177165354  
## [91] 0.179133858 0.181102362 0.183070866 0.185039370 0.187007874 0.188976378  
## [97] 0.190944882 0.192913386 0.194881890 0.196850394 0.198818898 0.200787402  
## [103] 0.202755906 0.204724409 0.206692913 0.208661417 0.210629921 0.212598425  
## [109] 0.214566929 0.216535433 0.218503937 0.220472441 0.222440945 0.224409449  
## [115] 0.226377953 0.228346457 0.230314961 0.232283465 0.234251969 0.236220472  
## [121] 0.238188976 0.240157480 0.242125984 0.244094488 0.246062992 0.248031496  
## [127] 0.250000000 0.251968504 0.253937008 0.255905512 0.257874016 0.259842520  
## [133] 0.261811024 0.263779528 0.265748031 0.267716535 0.269685039 0.271653543  
## [139] 0.273622047 0.275590551 0.277559055 0.279527559 0.281496063 0.283464567  
## [145] 0.285433071 0.287401575 0.289370079 0.291338583 0.293307087 0.295275591  
## [151] 0.297244094 0.299212598 0.301181102 0.303149606 0.305118110 0.307086614  
## [157] 0.309055118 0.311023622 0.312992126 0.314960630 0.316929134 0.318897638  
## [163] 0.320866142 0.322834646 0.324803150 0.326771654 0.328740157 0.330708661  
## [169] 0.332677165 0.334645669 0.336614173 0.338582677 0.340551181 0.342519685  
## [175] 0.344488189 0.346456693 0.348425197 0.350393701 0.352362205 0.354330709  
## [181] 0.356299213 0.358267717 0.360236220 0.362204724 0.364173228 0.366141732  
## [187] 0.368110236 0.370078740 0.372047244 0.374015748 0.375984252 0.377952756  
## [193] 0.379921260 0.381889764 0.383858268 0.385826772 0.387795276 0.389763780  
## [199] 0.391732283 0.393700787 0.395669291 0.397637795 0.399606299 0.401574803  
## [205] 0.403543307 0.405511811 0.407480315 0.409448819 0.411417323 0.413385827  
## [211] 0.415354331 0.417322835 0.419291339 0.421259843 0.423228346 0.425196850  
## [217] 0.427165354 0.429133858 0.431102362 0.433070866 0.435039370 0.437007874  
## [223] 0.438976378 0.440944882 0.442913386 0.444881890 0.446850394 0.448818898  
## [229] 0.450787402 0.452755906 0.454724409 0.456692913 0.458661417 0.460629921  
## [235] 0.462598425 0.464566929 0.466535433 0.468503937 0.470472441 0.472440945  
## [241] 0.474409449 0.476377953 0.478346457 0.480314961 0.482283465 0.484251969  
## [247] 0.486220472 0.488188976 0.490157480 0.492125984 0.494094488 0.496062992  
## [253] 0.498031496 0.500000000  
##   
## $db  
## [1] 10.20659694 7.46052360 7.83617148 2.26983509 4.44627998  
## [6] 11.46962664 2.36784400 -11.37202935 10.72714304 2.08820874  
## [11] 3.78108283 -1.44741479 4.75331571 0.83762607 6.88393814  
## [16] 8.73191021 -4.26011886 1.02276832 6.12952463 3.53589180  
## [21] 3.73025354 4.92914047 6.51631490 2.79825296 7.45447536  
## [26] 3.11490623 2.30503781 5.46313559 -5.26474583 3.55816134  
## [31] -3.23239939 -2.19731159 -0.32373133 6.50276793 8.97827428  
## [36] -0.28567397 -6.01190313 -5.23762451 -15.95336325 -2.81332405  
## [41] 3.47206276 -4.13943022 6.16668197 4.55186848 5.27195503  
## [46] -5.91833480 -1.59616360 2.25092112 -9.28627727 -0.30989689  
## [51] -4.79077157 -1.18519546 1.11087039 -3.03687059 -0.71495525  
## [56] -1.11541204 -12.49955913 -3.45912127 -3.20971892 -0.37005896  
## [61] -8.93561629 2.54513927 0.47407617 -4.51618775 -0.13763438  
## [66] 2.44456068 -4.84372851 -10.89434714 -12.54362763 -8.37741931  
## [71] 1.39042500 0.92776183 -11.10569486 -4.83496543 -1.69224285  
## [76] -0.37449350 -11.54025678 -13.65469954 -0.54546176 -5.87458241  
## [81] -2.82267518 -6.14913529 -23.09653359 -5.93372166 -8.60991890  
## [86] -0.23188249 -1.15556879 -16.22773311 -1.21949538 -1.99442300  
## [91] -10.90874240 -4.03358161 -2.98969416 -1.48583713 -1.00747767  
## [96] -0.34365593 -0.15368161 -7.81317859 -3.65016041 -1.51956372  
## [101] -3.69983746 -8.59665741 -2.14747098 -3.09713287 -0.86392543  
## [106] -1.62486028 -10.26669354 -7.20303674 -6.91080007 -10.25819378  
## [111] -0.78136386 -11.51392665 -1.47437960 -14.20737139 -10.82923416  
## [116] -3.94934045 -17.60613611 -3.36646328 -11.10105729 -5.76562696  
## [121] -2.31428092 -4.03447303 -24.13475808 0.03708916 1.47945306  
## [126] -3.75471799 -140.20091319 -2.83198327 -6.30847005 -6.44637850  
## [131] 2.41966178 -10.49245350 -5.01419281 -3.73108937 -10.57921790  
## [136] -4.26781295 -17.69315101 0.20356243 -5.05857600 -5.21426176  
## [141] -8.62264416 -1.05434283 -4.03097645 -8.82165422 -7.28270450  
## [146] -7.87328420 -9.66956539 -6.35862090 -9.98803700 -1.24463906  
## [151] -7.63591130 -0.07848241 -14.34523530 -12.94646001 -10.76747421  
## [156] -9.77630147 -6.34970155 -5.31426321 -9.24445388 -7.80198660  
## [161] 0.69316609 -12.77196995 -7.21771095 -7.77583391 -1.77549866  
## [166] -21.04583219 -8.55441330 -1.59009377 -7.02137149 4.16377511  
## [171] 6.05116660 -1.77495753 -0.39626699 -1.36931158 -5.26225613  
## [176] -16.93557986 -4.61422146 -3.72258512 -2.18825512 -1.87184650  
## [181] -6.13790510 -6.35351696 -6.12954409 -22.47980931 -3.38325599  
## [186] -4.24899873 -8.55677317 -2.15133387 -4.20976861 -9.29680097  
## [191] -2.18758335 -4.26801144 -2.78059522 -13.50819830 -6.25589978  
## [196] -4.23886796 -14.58477824 -12.74060102 -1.41188920 1.50317828  
## [201] -8.10169253 -9.08860401 -1.26573950 -7.05472208 -15.58988015  
## [206] -14.13663894 -18.60948663 -2.18073510 3.02562870 0.48161663  
## [211] -3.78103592 1.95250131 0.31316799 -1.60673267 -18.72638536  
## [216] -9.08819867 -13.14283965 -1.71348146 -2.97010978 2.49393069  
## [221] -2.18305438 -2.60702564 -5.20739470 -2.40482551 -2.29215813  
## [226] -4.19595656 -6.97837410 -0.50620624 5.90723862 2.10256527  
## [231] 0.35581504 -3.60400046 -0.42513995 -7.45415928 -3.56403293  
## [236] 2.16550610 -2.17964299 -1.80745945 -7.33168125 -1.06280850  
## [241] 2.34656138 0.92644865 -5.88929376 -12.60332989 -5.20096926  
## [246] 3.19754970 -2.83047001 2.35993403 -6.34913961 -2.21667044  
## [251] -8.85221051 -9.77770283 -2.67152109 NaN  
##   
## $dbz  
## [1] 7.3556594 7.3057608 7.2240695 7.1127972 6.9750336 6.8147136  
## [7] 6.6365410 6.4458459 6.2483577 6.0498871 5.8559304 5.6712356  
## [13] 5.4993961 5.3425474 5.2012356 5.0744899 4.9600860 4.8549464  
## [19] 4.7555987 4.6586149 4.5609739 4.4603160 4.3550812 4.2445416  
## [25] 4.1287441 4.0083847 3.8846377 3.7589566 3.6328710 3.5077945  
## [31] 3.3848613 3.2648026 3.1478670 3.0337854 2.9217745 2.8105700  
## [37] 2.6984816 2.5834648 2.4632066 2.3352263 2.1969963 2.0460841  
## [43] 1.8803214 1.6979990 1.4980856 1.2804631 1.0461648 0.7975904  
## [49] 0.5386596 0.2748445 0.0130121 -0.2389946 -0.4730732 -0.6818338  
## [55] -0.8596793 -1.0037438 -1.1144244 -1.1953346 -1.2526892 -1.2943100  
## [61] -1.3285222 -1.3631829 -1.4049809 -1.4590404 -1.5287908 -1.6160330  
## [67] -1.7211346 -1.8432936 -1.9808272 -2.1314480 -2.2924969 -2.4611064  
## [73] -2.6342783 -2.8088696 -2.9815044 -3.1484466 -3.3054912 -3.4479411  
## [79] -3.5707374 -3.6687855 -3.7374729 -3.7733093 -3.7745557 -3.7416800  
## [85] -3.6775043 -3.5869907 -3.4767226 -3.3542107 -3.2271813 -3.1029692  
## [91] -2.9880795 -2.8879306 -2.8067504 -2.7475866 -2.7123902 -2.7021411  
## [97] -2.7169929 -2.7564222 -2.8193723 -2.9043822 -3.0096924 -3.1333143  
## [103] -3.2730533 -3.4264722 -3.5907909 -3.7627235 -3.9382749 -4.1125352  
## [109] -4.2795404 -4.4322863 -4.5629947 -4.6637026 -4.7271737 -4.7480064  
## [115] -4.7236860 -4.6552699 -4.5474683 -4.4080774 -4.2469385 -4.0747209  
## [121] -3.9018110 -3.7374834 -3.5893995 -3.4633932 -3.3634634 -3.2918962  
## [127] -3.2494573 -3.2356188 -3.2488038 -3.2866397 -3.3462198 -3.4243666  
## [133] -3.5178864 -3.6237978 -3.7395120 -3.8629428 -3.9925300 -4.1271753  
## [139] -4.2661023 -4.4086677 -4.5541594 -4.7016207 -4.8497315 -4.9967652  
## [145] -5.1406236 -5.2789258 -5.4091091 -5.5284858 -5.6341985 -5.7230409  
## [151] -5.7911559 -5.8336881 -5.8445370 -5.8164125 -5.7413867 -5.6120358  
## [157] -5.4230452 -5.1728767 -4.8649268 -4.5077073 -4.1139554 -3.6990035  
## [163] -3.2789613 -2.8691873 -2.4832824 -2.1326094 -1.8262048 -1.5709240  
## [169] -1.3716840 -1.2317147 -1.1527657 -1.1352446 -1.1782776 -1.2796936  
## [175] -1.4359418 -1.6419612 -1.8910360 -2.1746907 -2.4827049 -2.8033488  
## [181] -3.1239358 -3.4317390 -3.7152050 -3.9652409 -4.1762272 -4.3464202  
## [187] -4.4775787 -4.5739247 -4.6407588 -4.6831088 -4.7046985 -4.7073739  
## [193] -4.6910000 -4.6537673 -4.5928157 -4.5050627 -4.3880956 -4.2409731  
## [199] -4.0647827 -3.8628494 -3.6405732 -3.4049613 -3.1639815 -2.9258761  
## [205] -2.6985395 -2.4890167 -2.3031303 -2.1452105 -2.0178960 -1.9219725  
## [211] -1.8562350 -1.8173815 -1.7999763 -1.7965559 -1.7979734 -1.7940758  
## [217] -1.7747538 -1.7312812 -1.6577007 -1.5518950 -1.4160026 -1.2560369  
## [223] -1.0808361 -0.9006793 -0.7259316 -0.5659706 -0.4284852 -0.3191155  
## [229] -0.2413443 -0.1965498 -0.1841552 -0.2018443 -0.2458399 -0.3112615  
## [235] -0.3925765 -0.4841442 -0.5808191 -0.6785405 -0.7748038 -0.8689087  
## [241] -0.9619137 -1.0562869 -1.1553112 -1.2623431 -1.3800326 -1.5095975  
## [247] -1.6502219 -1.7986430 -1.9490006 -2.0930506 -2.2208453 -2.3219308  
## [253] -2.3869609 -2.4094186

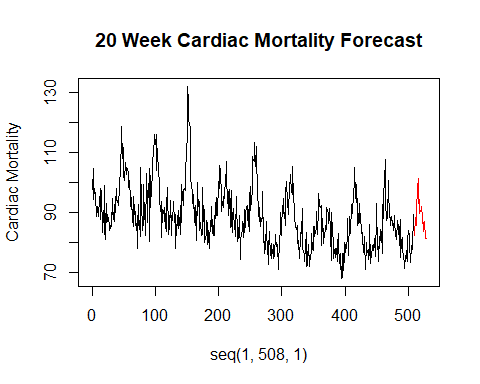
phi = aic.wge(ksfit$residuals)  
resids = fore.arma.wge(ksfit$residuals,phi = phi$phi,n.ahead = 20, limits = F)



#predict trend manually  
preds = predict(ksfit, newdata = next20)  
  
predsFinal = preds + resids$f

Plot the next 20 cmort wrt time

plot(seq(1,508,1), CM$cmort, type = "l",xlim = c(0,528), ylab = "Cardiac Mortality", main = "20 Week Cardiac Mortality Forecast")  
lines(seq(509,528,1), predsFinal, type = "l", col = "red")



length(predsFinal)

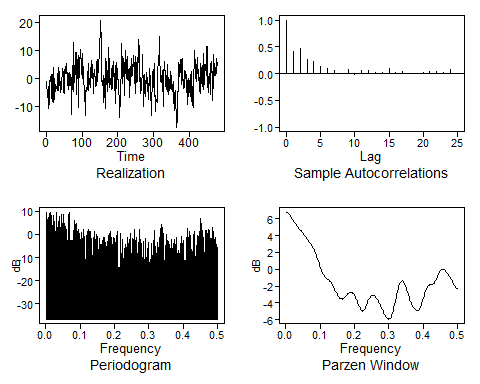
## [1] 20

To Find the ASE we need to forecast last 30 of the known series.

CMsmall = CM[2:478,]  
ksfit = lm(cmort~temp+part+Week+FWeek, data = CMsmall)  
phi = aic.wge(ksfit$residuals)  
fit = arima(CMsmall$cmort,order = c(phi$p,0,0), seasonal = list(order = c(1,0,0), period = 52), xreg = cbind(CMsmall$temp1, CMsmall$part, CMsmall$Week, CMsmall$FWeek))  
  
last30 = data.frame(temp = CM$temp[479:508], part = CM$part[479:508], Week = seq(479,508,1), FWeek = as.factor(seq(479,508,1)%%52))

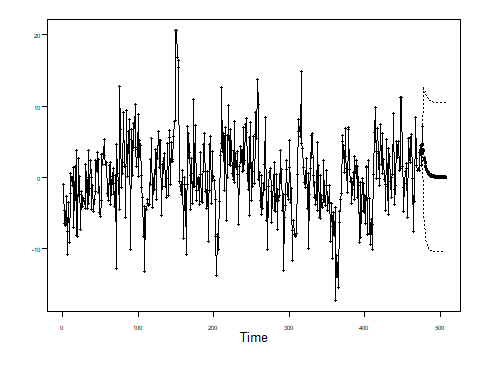
Predict residuals manually

plotts.sample.wge(ksfit$residuals)



## $autplt  
## [1] 1.000000000 0.418101024 0.462636935 0.271796012 0.228536503  
## [6] 0.128871501 0.095743981 0.053004143 0.012448232 0.080022179  
## [11] -0.015257437 0.065053444 0.058531747 0.015627633 0.018652761  
## [16] 0.094852410 0.018937017 0.036661164 0.006313430 -0.004489214  
## [21] 0.030478221 0.034794407 0.032589896 0.015039533 0.077488751  
## [26] 0.010734371  
##   
## $freq  
## [1] 0.002096436 0.004192872 0.006289308 0.008385744 0.010482180 0.012578616  
## [7] 0.014675052 0.016771488 0.018867925 0.020964361 0.023060797 0.025157233  
## [13] 0.027253669 0.029350105 0.031446541 0.033542977 0.035639413 0.037735849  
## [19] 0.039832285 0.041928721 0.044025157 0.046121593 0.048218029 0.050314465  
## [25] 0.052410901 0.054507338 0.056603774 0.058700210 0.060796646 0.062893082  
## [31] 0.064989518 0.067085954 0.069182390 0.071278826 0.073375262 0.075471698  
## [37] 0.077568134 0.079664570 0.081761006 0.083857442 0.085953878 0.088050314  
## [43] 0.090146751 0.092243187 0.094339623 0.096436059 0.098532495 0.100628931  
## [49] 0.102725367 0.104821803 0.106918239 0.109014675 0.111111111 0.113207547  
## [55] 0.115303983 0.117400419 0.119496855 0.121593291 0.123689727 0.125786164  
## [61] 0.127882600 0.129979036 0.132075472 0.134171908 0.136268344 0.138364780  
## [67] 0.140461216 0.142557652 0.144654088 0.146750524 0.148846960 0.150943396  
## [73] 0.153039832 0.155136268 0.157232704 0.159329140 0.161425577 0.163522013  
## [79] 0.165618449 0.167714885 0.169811321 0.171907757 0.174004193 0.176100629  
## [85] 0.178197065 0.180293501 0.182389937 0.184486373 0.186582809 0.188679245  
## [91] 0.190775681 0.192872117 0.194968553 0.197064990 0.199161426 0.201257862  
## [97] 0.203354298 0.205450734 0.207547170 0.209643606 0.211740042 0.213836478  
## [103] 0.215932914 0.218029350 0.220125786 0.222222222 0.224318658 0.226415094  
## [109] 0.228511530 0.230607966 0.232704403 0.234800839 0.236897275 0.238993711  
## [115] 0.241090147 0.243186583 0.245283019 0.247379455 0.249475891 0.251572327  
## [121] 0.253668763 0.255765199 0.257861635 0.259958071 0.262054507 0.264150943  
## [127] 0.266247379 0.268343816 0.270440252 0.272536688 0.274633124 0.276729560  
## [133] 0.278825996 0.280922432 0.283018868 0.285115304 0.287211740 0.289308176  
## [139] 0.291404612 0.293501048 0.295597484 0.297693920 0.299790356 0.301886792  
## [145] 0.303983229 0.306079665 0.308176101 0.310272537 0.312368973 0.314465409  
## [151] 0.316561845 0.318658281 0.320754717 0.322851153 0.324947589 0.327044025  
## [157] 0.329140461 0.331236897 0.333333333 0.335429769 0.337526205 0.339622642  
## [163] 0.341719078 0.343815514 0.345911950 0.348008386 0.350104822 0.352201258  
## [169] 0.354297694 0.356394130 0.358490566 0.360587002 0.362683438 0.364779874  
## [175] 0.366876310 0.368972746 0.371069182 0.373165618 0.375262055 0.377358491  
## [181] 0.379454927 0.381551363 0.383647799 0.385744235 0.387840671 0.389937107  
## [187] 0.392033543 0.394129979 0.396226415 0.398322851 0.400419287 0.402515723  
## [193] 0.404612159 0.406708595 0.408805031 0.410901468 0.412997904 0.415094340  
## [199] 0.417190776 0.419287212 0.421383648 0.423480084 0.425576520 0.427672956  
## [205] 0.429769392 0.431865828 0.433962264 0.436058700 0.438155136 0.440251572  
## [211] 0.442348008 0.444444444 0.446540881 0.448637317 0.450733753 0.452830189  
## [217] 0.454926625 0.457023061 0.459119497 0.461215933 0.463312369 0.465408805  
## [223] 0.467505241 0.469601677 0.471698113 0.473794549 0.475890985 0.477987421  
## [229] 0.480083857 0.482180294 0.484276730 0.486373166 0.488469602 0.490566038  
## [235] 0.492662474 0.494758910 0.496855346 0.498951782  
##   
## $db  
## [1] 9.29045594 3.92249535 7.18988838 0.27659482 9.22776358  
## [6] 9.50455827 0.33482492 8.19888119 -0.11139890 8.23453993  
## [11] -5.16038198 5.52143622 -2.80843374 4.08433802 9.61689976  
## [16] -2.60736815 2.03188738 3.34561004 7.65939016 -1.96098919  
## [21] 8.74170683 -1.72036522 0.69193171 8.52060643 -6.05693542  
## [26] 4.83079703 3.12016450 1.73198146 -0.06679052 -5.98770627  
## [31] 1.35163489 7.81930377 9.33353758 -13.67129861 -15.18155904  
## [36] -2.67254129 -6.88144294 1.92018347 -2.05278415 5.98446180  
## [41] 3.43186374 4.60379745 0.79598881 0.46009215 2.18560419  
## [46] -10.98330716 0.36374752 -4.29951977 -1.23416879 0.73859204  
## [51] -5.16001531 2.00146441 -2.90759934 -2.40857285 -20.59740795  
## [56] -0.42602409 -3.26588074 0.43743054 2.29335503 -6.65848563  
## [61] 1.47295720 1.32817455 -4.16277225 -19.90136950 -14.31519809  
## [66] -3.86203745 1.13425408 -2.10277356 -3.72934718 0.20296984  
## [71] -0.81675737 -6.13629820 -11.35624538 -1.54800770 -7.87149936  
## [76] -4.35576336 -4.35681827 -17.44762131 -3.95135335 -1.45835084  
## [81] -6.81786408 -1.42635331 -5.00661576 -6.84728989 -3.14132507  
## [86] -8.85541995 -0.53879136 -1.15830908 -0.36598783 1.14972431  
## [91] -3.43074778 -11.93582812 -7.20774470 -1.27503253 -1.79277795  
## [96] -4.22162252 -11.52490807 -1.82644126 1.29208219 -5.72719815  
## [101] -21.98343139 -15.66905932 -12.47638411 0.36019079 -16.11834000  
## [106] -2.18969875 -17.01380069 -14.52887375 -5.88002330 -36.56873831  
## [111] -4.30318905 -6.32611076 -6.25941746 -1.21897959 -5.44368473  
## [116] -2.58698126 -1.19663377 0.58471503 -16.19887448 -1.79648714  
## [121] -8.30781728 -6.48009489 2.15074981 -6.56754285 -6.58989715  
## [126] -1.20968429 -14.41302121 -9.40769640 -9.08017945 1.99843616  
## [131] -5.64444384 -8.32581091 -3.99045792 -16.37263077 -2.88312708  
## [136] -3.67077409 -17.36130661 -18.86258924 -1.69026837 -17.18041234  
## [141] -6.00026290 -6.45831490 -3.48811158 -19.88459344 -17.09914647  
## [146] -3.90085949 -12.35820874 -9.27798776 -13.67584705 -6.88230471  
## [151] 1.64671383 -11.28901744 -8.99299641 -13.07170077 -3.61572345  
## [156] -26.01320659 -6.60593984 -4.77147777 2.78047827 -7.36869906  
## [161] 4.12409968 2.58884848 -1.30567105 0.07564399 -21.88255263  
## [166] -9.29992032 -0.14679193 -0.85061295 -3.95932561 -3.10857812  
## [171] -11.15398105 -4.36264157 -7.17459426 -8.36014073 -2.07652436  
## [176] -9.17299166 -4.80137139 -3.05201941 -2.15835746 -7.03938104  
## [181] -1.15098520 -11.64790057 -8.93789219 -6.75881440 -13.30941231  
## [186] -9.45622560 -6.31978696 -0.77954425 -7.17528310 -5.23803740  
## [191] -2.67332007 -19.59028386 -14.26452664 -17.14176035 -5.19232170  
## [196] 2.45276309 2.73599758 -4.74445356 1.21336745 1.31976160  
## [201] -1.77565540 -10.87024320 -16.57834957 -12.85130171 -1.50718314  
## [206] 1.29422471 -6.43437097 -11.75141399 -9.01464691 -0.47338108  
## [211] -28.99485271 -0.88810771 -6.62691781 -3.71932560 7.06767718  
## [216] 3.09348726 1.43903841 -4.83106149 -1.40565148 -17.41406187  
## [221] 1.62906925 -6.46003393 1.59910590 -13.46678822 -0.98956171  
## [226] 2.18167503 1.70591013 -4.14176633 -9.83352423 -6.47160030  
## [231] 3.04431098 -1.21935201 1.38486984 -4.57844203 -1.44840043  
## [236] -24.01341877 -4.32633128 -6.69287101  
##   
## $dbz  
## [1] 6.797041165 6.760714725 6.701185055 6.619977154 6.519223807  
## [6] 6.401643030 6.270482334 6.129414733 5.982375202 5.833335210  
## [11] 5.686027205 5.543647956 5.408584765 5.282215052 5.164822875  
## [16] 5.055654995 4.953109954 4.855025865 4.759015046 4.662790288  
## [21] 4.564436579 4.462597984 4.356566441 4.246273776 4.132199084  
## [26] 4.015211232 3.896371246 3.776722022 3.657092559 3.537940176  
## [31] 3.419246937 3.300476855 3.180590701 3.058107689 2.931199910  
## [36] 2.797806228 2.655756609 2.502903423 2.337261259 2.157159758  
## [41] 1.961414043 1.749514304 1.521829800 1.279812845 1.026175055  
## [46] 0.764992179 0.501678244 0.242762158 -0.004587723 -0.233300273  
## [51] -0.437327521 -0.612590999 -0.757714031 -0.874322807 -0.966813891  
## [56] -1.041651684 -1.106393402 -1.168681821 -1.235397982 -1.312073782  
## [61] -1.402578612 -1.509041347 -1.631949540 -1.770368772 -1.922233218  
## [66] -2.084665193 -2.254284608 -2.427471390 -2.600549697 -2.769876224  
## [71] -2.931837074 -3.082784899 -3.218972960 -3.336556047 -3.431721401  
## [76] -3.500981517 -3.541608077 -3.552126127 -3.532743462 -3.485586279  
## [81] -3.414658175 -3.325519768 -3.224765450 -3.119418739 -3.016366025  
## [86] -2.921912066 -2.841492548 -2.779538589 -2.739463824 -2.723735449  
## [91] -2.733991263 -2.771169604 -2.835624695 -2.927204605 -3.045272708  
## [96] -3.188656967 -3.355515901 -3.543117664 -3.747541768 -3.963335429  
## [101] -4.183191475 -4.397762696 -4.595777906 -4.764647707 -4.891690960  
## [106] -4.965928036 -4.980093083 -4.932258975 -4.826463141 -4.672066498  
## [111] -4.482093295 -4.271154880 -4.053567963 -3.842026504 -3.646892989  
## [116] -3.475987990 -3.334698959 -3.226252431 -3.152045403 -3.111979863  
## [121] -3.104778188 -3.128275898 -3.179695484 -3.255904386 -3.353654833  
## [126] -3.469796170 -3.601444227 -3.746089888 -3.901631998 -4.066328206  
## [131] -4.238669649 -4.417198233 -4.600295392 -4.785975791 -4.971717382  
## [136] -5.154350544 -5.330015215 -5.494178747 -5.641693435 -5.766868196  
## [141] -5.863540841 -5.925168835 -5.945000178 -5.916420437 -5.833563787  
## [146] -5.692197697 -5.490748588 -5.231191699 -4.919481723 -4.565316213  
## [151] -4.181259864 -3.781480803 -3.380438891 -2.991801689 -2.627719535  
## [156] -2.298456242 -2.012293233 -1.775602757 -1.592998460 -1.467496868  
## [161] -1.400648144 -1.392613760 -1.442182699 -1.546728949 -1.702124475  
## [166] -1.902636456 -2.140857020 -2.407736376 -2.692808634 -2.984698035  
## [171] -3.271950156 -3.544133706 -3.793019294 -4.013523435 -4.204094286  
## [176] -4.366354367 -4.504052816 -4.621591901 -4.722481141 -4.808034927  
## [181] -4.876538881 -4.923037244 -4.939850664 -4.917868955 -4.848505648  
## [186] -4.725939060 -4.549016829 -4.322179166 -4.055078460 -3.761090632  
## [191] -3.455299645 -3.152579687 -2.866164047 -2.606790422 -2.382310515  
## [196] -2.197582399 -2.054485385 -1.951957646 -1.886026370 -1.849868348  
## [201] -1.834002550 -1.826764708 -1.815218930 -1.786580608 -1.730032934  
## [206] -1.638563432 -1.510273534 -1.348690638 -1.161961073 -0.961228958  
## [211] -0.758747052 -0.566217815 -0.393632589 -0.248639721 -0.136332768  
## [216] -0.059314417 -0.017919122 -0.010525106 -0.033928577 -0.083778370  
## [221] -0.155074451 -0.242719599 -0.342085279 -0.449521034 -0.562716857  
## [226] -0.680834034 -0.804355671 -0.934662916 -1.073396510 -1.221698195  
## [231] -1.379438837 -1.544540163 -1.712500589 -1.876250099 -2.026472236  
## [236] -2.152505697 -2.243825834 -2.291890846

phi = aic.wge(ksfit$residuals)  
resids = fore.arma.wge(ksfit$residuals,phi = phi$phi,n.ahead = 30)

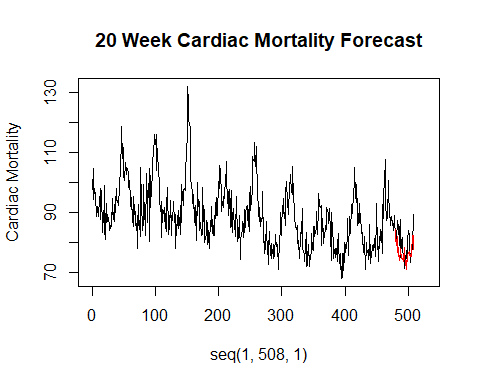


Predict trend manually

preds = predict(ksfit, newdata = last30)  
#Final using the trend and residuals  
predsFinal = preds + resids$f

Plot and print out the ASE

plot(seq(1,508,1), CM$cmort, type = "l",xlim = c(0,528), ylab = "Cardiac Mortality", main = "20 Week Cardiac Mortality Forecast")  
lines(seq(479,508,1), predsFinal, type = "l", col = "red")



ASE\_ARMA2 = mean((CM$cmort[479:508] - predsFinal)^2,na.rm = TRUE)  
ASE\_ARMA2

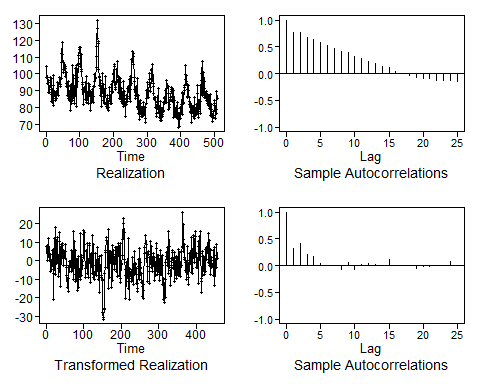
## [1] 32.45511

#### VAR Model

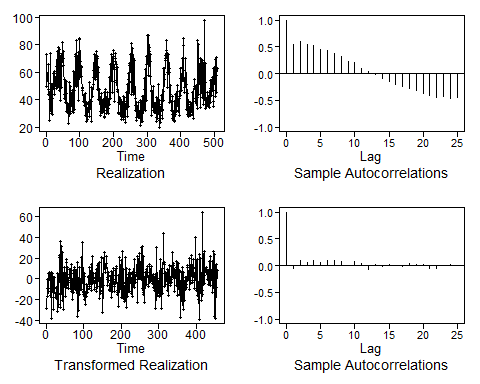
**VAR Model Forecasts Seasonally Differenced Data**

Difference all series to make them stationary which is an assumption of VAR.  
They do not have to be white, just stationary.

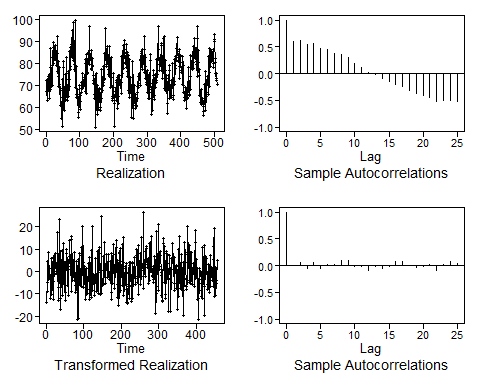
CM = read.csv("la\_cmort\_study.csv", header = TRUE)  
CM\_52 = artrans.wge(CM$cmort,c(rep(0,51),1))



Part\_52 = artrans.wge(CM$part,c(rep(0,51),1))



Temp\_52 = artrans.wge(CM$temp,c(rep(0,51),1))



#VARSelect on Differenced Data chooses 2  
VARselect(cbind(CM\_52, Part\_52, Temp\_52),lag.max = 10, type = "both")

## $selection  
## AIC(n) HQ(n) SC(n) FPE(n)   
## 4 2 2 4   
##   
## $criteria  
## 1 2 3 4 5  
## AIC(n) 12.87263 12.74039 12.74236 12.71258 12.71419  
## HQ(n) 12.92701 12.82739 12.86198 12.86483 12.89906  
## SC(n) 13.01054 12.96103 13.04575 13.09871 13.18307  
## FPE(n) 389506.60462 341260.26330 341940.07226 331918.68995 332470.16877  
## 6 7 8 9 10  
## AIC(n) 12.74034 12.75365 12.75100 12.73260 12.75196  
## HQ(n) 12.95784 13.00376 13.03374 13.04796 13.09995  
## SC(n) 13.29196 13.38800 13.46810 13.53244 13.63454  
## FPE(n) 341302.58712 345905.71709 345034.37801 338794.95293 345484.81652

#VAR with p = 2  
CMortDiffVAR = VAR(cbind(CM\_52, Part\_52, Temp\_52),type = "both",p = 2)  
preds=predict(CMortDiffVAR,n.ahead=20)  
  
#We have predicted differences .... calculate actual cardiac mortalities   
startingPoints = CM$cmort[428:457]  
CMortForcasts = preds$fcst$CM\_52[,1:3] + startingPoints  
  
#Plot  
while (!is.null(dev.list())) dev.off()  
plot(seq(1,508,1), CM$cmort, type = "l",xlim = c(0,528), ylab = "Cardiac Mortality", main = "20 Week Cardiac Mortality Forecast")  
lines(seq(509,528,1), as.data.frame(CMortForcasts)$fcst, type = "l", col = "red")  
  
ASE\_VAR1 = mean((CM$cmort[489:508] - CMortForcasts[,1])^2)  
ASE\_VAR1

## [1] 25.47226

### Modeling with the sales data set

find the ASE for forecasts of the last five observations of the sales data set example for both the MLE model (using ARIMA) and the VAR model we fit.

We fit several models using the arima() function, you should pick the one you feel is the best model (maybe with the lowest AIC?).

Your slides should include each of the following:

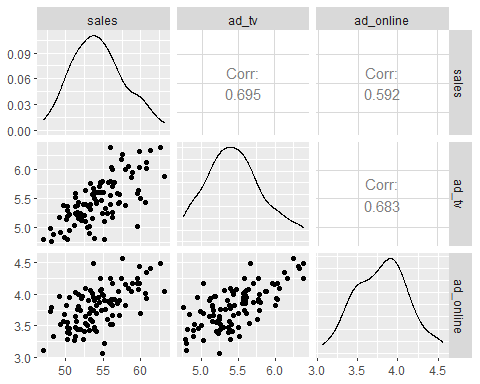
* Your code
* Enough visual aids to describe how you calculated the ASE for each model
* The ASE for each model
* Which model you feel is better in that respect, and why.

Read in and visualize the sales data

BSales = read.csv("businesssales.csv", header = TRUE)  
  
## Visualaze The Data  
head(BSales)

## X sales ad\_tv ad\_online discount  
## 1 1 52.61137 5.105617 3.646744 0.0000000  
## 2 2 50.46478 5.233105 3.381140 0.0000000  
## 3 3 50.15374 5.202212 3.550546 0.7571115  
## 4 4 49.68254 5.386699 3.570285 3.5303790  
## 5 5 50.44613 4.961060 3.460531 5.4983130  
## 6 6 51.72145 5.153207 3.801946 7.2211060

#Make a matrix of plots with the BSales data set.  
ggpairs(BSales[2:4]) #matrix of scatter plots



#### MLR Modeling

**Model 1**

All data with no lag and no trend

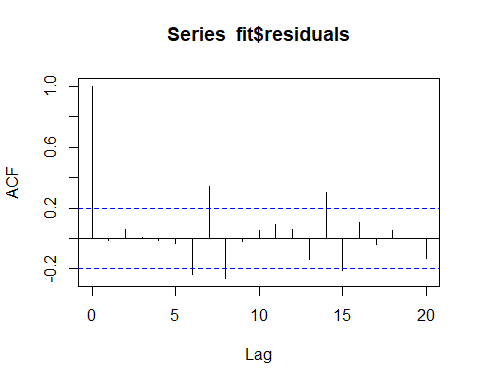
ksfit=lm(sales~ad\_tv+ad\_online+discount, data = BSales)  
aic.wge(ksfit$residuals,p=0:8, q=0) # AIC picks p=7

## $type  
## [1] "aic"  
##   
## $value  
## [1] 1.101655  
##   
## $p  
## [1] 7  
##   
## $q  
## [1] 0  
##   
## $phi  
## [1] 0.3914830668 0.1263257053 -0.3764190275 -0.0005409533 0.1006855558  
## [6] 0.2148095958 0.1476114131  
##   
## $theta  
## [1] 0  
##   
## $vara  
## [1] 2.564221

fit=arima(BSales$sales,order=c(7,0,0),xreg=BSales[,3:5])  
fit

##   
## Call:  
## arima(x = BSales$sales, order = c(7, 0, 0), xreg = BSales[, 3:5])  
##   
## Coefficients:  
## ar1 ar2 ar3 ar4 ar5 ar6 ar7 intercept  
## 1.4734 -0.8921 0.0749 0.0919 0.0438 0.1865 -0.1287 54.5513  
## s.e. 0.1107 0.1963 0.2167 0.2057 0.2049 0.1852 0.1080 2.2040  
## ad\_tv ad\_online discount  
## 0.0703 -0.0934 -0.1514  
## s.e. 0.3434 0.2075 0.1315  
##   
## sigma^2 estimated as 1.411: log likelihood = -161.1, aic = 346.21

acf(fit$residuals)



ltest = ljung.wge(fit$resid)

## Obs -0.01040185 0.05898301 0.008929627 -0.01224377 -0.0296057 -0.2398486 0.3402942 -0.2623472 -0.01792799 0.04983832 0.09440642 0.05862074 -0.1409925 0.299757 -0.208223 0.1039675 -0.04145994 0.05483977 0.002771559 -0.1299524 0.228699 -0.1645907 0.09479325 -0.170837

ltest$pval

## [1] 7.445469e-06

ASE for model with no lag and no trend (last 5)

#Cut off the last 5  
BSales2 = BSales[1:95,]  
ksfit=lm(sales~ad\_tv+ad\_online+discount, data = BSales2)  
aic.wge(ksfit$residuals,p=0:8, q=0) # AIC picks p=7

## $type  
## [1] "aic"  
##   
## $value  
## [1] 1.119875  
##   
## $p  
## [1] 7  
##   
## $q  
## [1] 0  
##   
## $phi  
## [1] 0.39110624 0.13440258 -0.39244949 0.03048536 0.07647737 0.19178372  
## [7] 0.16884555  
##   
## $theta  
## [1] 0  
##   
## $vara  
## [1] 2.589471

fit=arima(BSales2$sales,order=c(7,0,0),xreg=cbind(BSales2$ad\_tv,BSales2$ad\_online,BSales2$discount))  
fit

##   
## Call:  
## arima(x = BSales2$sales, order = c(7, 0, 0), xreg = cbind(BSales2$ad\_tv, BSales2$ad\_online,   
## BSales2$discount))  
##   
## Coefficients:  
## ar1 ar2 ar3 ar4 ar5 ar6 ar7 intercept  
## 1.4819 -0.9000 0.1156 0.073 0.0122 0.2405 -0.1444 55.5453  
## s.e. 0.1129 0.2002 0.2214 0.213 0.2141 0.1924 0.1106 2.3458  
## cbind(BSales2$ad\_tv, BSales2$ad\_online, BSales2$discount)1  
## -0.0418  
## s.e. 0.3508  
## cbind(BSales2$ad\_tv, BSales2$ad\_online, BSales2$discount)2  
## -0.1579  
## s.e. 0.2143  
## cbind(BSales2$ad\_tv, BSales2$ad\_online, BSales2$discount)3  
## -0.1140  
## s.e. 0.1489  
##   
## sigma^2 estimated as 1.426: log likelihood = -153.65, aic = 331.29

preds = predict(fit, newxreg = cbind(BSales$ad\_tv[96:100],BSales$ad\_online[96:100],BSales$discount[96:100]))  
ASE1 = mean((BSales$sales[96:100] - preds$pred)^2)  
ASE1

## [1] 57.1069

while (!is.null(dev.list())) dev.off()  
  
plot(seq(1,100,1), BSales$sales[1:100], type = "l",xlim = c(0,100), ylab = "Business Sales", main = "5 Week Sales Forecast")  
lines(seq(96,100,1), preds$pred, type = "l", col = "red")

**Model 2**

ASE for model with no lag and trend (last 5)

BSales = read.csv("businesssales.csv", header = TRUE)  
#Create trend (length of the dataset)  
t=1:100  
#add trend to the BSales dataset  
BSales$t = t  
#Take off the last 5 for testing the forcast  
BSales2 = BSales[1:95,]  
#Trend (t) has been added to the model  
ksfit=lm(sales~t+ad\_tv+ad\_online+discount, data = BSales2)  
aic.wge(ksfit$residuals,p=0:8, q=0) # AIC picks p=6

## $type  
## [1] "aic"  
##   
## $value  
## [1] 0.9778159  
##   
## $p  
## [1] 6  
##   
## $q  
## [1] 0  
##   
## $phi  
## [1] 0.46249692 0.09924096 -0.45100298 -0.02869670 0.08821844 0.19665961  
##   
## $theta  
## [1] 0  
##   
## $vara  
## [1] 2.294345

fit=arima(BSales2$sales,order=c(6,0,0),xreg=cbind(BSales2$ad\_tv,BSales2$ad\_online,BSales2$t,BSales2$discount))  
fit

##   
## Call:  
## arima(x = BSales2$sales, order = c(6, 0, 0), xreg = cbind(BSales2$ad\_tv, BSales2$ad\_online,   
## BSales2$t, BSales2$discount))  
##   
## Coefficients:  
## ar1 ar2 ar3 ar4 ar5 ar6 intercept  
## 1.3938 -0.8703 0.0509 0.0300 0.1505 -0.0108 51.9501  
## s.e. 0.1159 0.2013 0.2226 0.2103 0.1828 0.1073 2.3181  
## cbind(BSales2$ad\_tv, BSales2$ad\_online, BSales2$t, BSales2$discount)1  
## 0.0051  
## s.e. 0.3654  
## cbind(BSales2$ad\_tv, BSales2$ad\_online, BSales2$t, BSales2$discount)2  
## -0.0897  
## s.e. 0.1966  
## cbind(BSales2$ad\_tv, BSales2$ad\_online, BSales2$t, BSales2$discount)3  
## 0.0615  
## s.e. 0.0168  
## cbind(BSales2$ad\_tv, BSales2$ad\_online, BSales2$t, BSales2$discount)4  
## -0.0920  
## s.e. 0.1179  
##   
## sigma^2 estimated as 1.353: log likelihood = -150.77, aic = 325.53

preds = predict(fit, newxreg = cbind(BSales$ad\_tv[96:100],BSales$ad\_online[96:100],BSales$t[96:100],BSales$discount[96:100]))  
ASE2 = mean((BSales$sales[96:100] - preds$pred)^2)  
ASE2

## [1] 62.88811

while (!is.null(dev.list())) dev.off()  
plot(seq(1,100,1), BSales$sales[1:100], type = "l",xlim = c(0,100), ylab = "Business Sales", main = "5 Week Sales Forecast")  
lines(seq(96,100,1), preds$pred, type = "l", col = "red")

#Lagging Variables  
#Example:  
#With dplyr lag function  
library(dplyr)  
df = data.frame(Y = c(1,1,2,3,4,4,5,8),X1 = c(5,6,6,7,7,8,8,9))  
df$X1\_L1 = dplyr::lag(df$X1,1)  
df$X1\_L2 = dplyr::lag(df$X1,2)  
df

## Y X1 X1\_L1 X1\_L2  
## 1 1 5 NA NA  
## 2 1 6 5 NA  
## 3 2 6 6 5  
## 4 3 7 6 6  
## 5 4 7 7 6  
## 6 4 8 7 7  
## 7 5 8 8 7  
## 8 8 9 8 8

# Model 3  
  
#Lagging BSales Ad Variables  
ad\_tv1 = dplyr::lag(BSales$ad\_tv,1)  
ad\_online1 = dplyr::lag(BSales$ad\_online,1)  
BSales$ad\_tv1= ad\_tv1  
BSales$ad\_online1 = ad\_online1  
  
  
#with trend and lagging  
  
# ASE for model with no lag and trend (last 5)  
t=1:100  
BSales$t = t  
BSales2 = BSales[2:95,]  
ksfit=lm(sales~t+ad\_tv1+ad\_online1+discount, data = BSales2)  
aic.wge(ksfit$residuals,p=0:8,q=0:0) # AIC picks p=7

## $type  
## [1] "aic"  
##   
## $value  
## [1] 0.6537305  
##   
## $p  
## [1] 7  
##   
## $q  
## [1] 0  
##   
## $phi  
## [1] -0.03007855 0.17051043 -0.12919619 -0.03898462 -0.16150496 0.21423956  
## [7] 0.19801677  
##   
## $theta  
## [1] 0  
##   
## $vara  
## [1] 1.621769

fit = arima(BSales2$sales,order = c(7,0,0), xreg = cbind(BSales2$ad\_tv1,BSales2$ad\_online1,BSales2$t,BSales2$discount))  
fit

##   
## Call:  
## arima(x = BSales2$sales, order = c(7, 0, 0), xreg = cbind(BSales2$ad\_tv1, BSales2$ad\_online1,   
## BSales2$t, BSales2$discount))  
##   
## Coefficients:  
## ar1 ar2 ar3 ar4 ar5 ar6 ar7 intercept  
## 0.0407 0.2918 -0.2106 -0.0784 -0.1343 0.2739 0.2206 17.4379  
## s.e. 0.3025 0.1176 0.1321 0.1160 0.1145 0.1212 0.1183 6.7994  
## cbind(BSales2$ad\_tv1, BSales2$ad\_online1, BSales2$t, BSales2$discount)1  
## 3.5229  
## s.e. 0.5830  
## cbind(BSales2$ad\_tv1, BSales2$ad\_online1, BSales2$t, BSales2$discount)2  
## 4.4104  
## s.e. 1.6532  
## cbind(BSales2$ad\_tv1, BSales2$ad\_online1, BSales2$t, BSales2$discount)3  
## 0.0211  
## s.e. 0.0113  
## cbind(BSales2$ad\_tv1, BSales2$ad\_online1, BSales2$t, BSales2$discount)4  
## -0.0747  
## s.e. 0.0511  
##   
## sigma^2 estimated as 1.502: log likelihood = -153.23, aic = 332.46

preds = predict(fit, newxreg = cbind(BSales$ad\_tv1[96:100],BSales$ad\_online1[96:100],BSales$t[96:100],BSales$discount[96:100]))  
ASE3 = mean((BSales$sales[96:100] - preds$pred)^2)  
ASE3

## [1] 4.933711

while (!is.null(dev.list())) dev.off()  
plot(seq(1,100,1), BSales$sales[1:100], type = "l",xlim = c(0,100), ylab = "Business Sales", main = "5 Week Sales Forecast")  
lines(seq(96,100,1), preds$pred, type = "l", col = "red")  
  
  
  
  
  
####### Forecast Features   
  
plotts.sample.wge(BSales$ad\_tv)

## $autplt  
## [1] 1.00000000 0.58474346 0.21254108 -0.13388149 -0.21616075 -0.06681139  
## [7] 0.17784779 0.27494754 0.17066751 -0.01564928 -0.18797554 -0.14361859  
## [13] 0.16708921 0.30754572 0.28354958 0.11087656 -0.07499245 -0.17677033  
## [19] -0.06563142 0.17199060 0.28571289 0.32375925 0.08879168 -0.17600977  
## [25] -0.26663643 -0.12241449  
##   
## $freq  
## [1] 0.01 0.02 0.03 0.04 0.05 0.06 0.07 0.08 0.09 0.10 0.11 0.12 0.13 0.14 0.15  
## [16] 0.16 0.17 0.18 0.19 0.20 0.21 0.22 0.23 0.24 0.25 0.26 0.27 0.28 0.29 0.30  
## [31] 0.31 0.32 0.33 0.34 0.35 0.36 0.37 0.38 0.39 0.40 0.41 0.42 0.43 0.44 0.45  
## [46] 0.46 0.47 0.48 0.49 0.50  
##   
## $db  
## [1] 7.1430820 4.2493100 3.0895327 -3.4194121 5.0114908 2.9282282  
## [7] -6.7839088 3.5529969 4.0636201 5.1338549 -2.5945726 2.4688269  
## [13] 0.2485623 4.8389808 10.5017546 -0.8872750 -6.5241577 -2.7893890  
## [19] -33.4080685 -12.2859548 -17.4631938 -15.1244411 -0.9232950 -9.7598121  
## [25] -3.0578166 -10.7304812 -8.4430064 -11.0844285 -14.7639341 -10.5824888  
## [31] -12.9404024 -0.7225899 -2.0538719 -9.2806871 -14.0335086 -5.5639748  
## [37] -17.5222190 -7.2088663 -5.6979641 -13.9041414 -13.0927206 -1.9705313  
## [43] -3.7179589 -5.2340822 -12.4783821 -12.5221820 -6.7733554 -6.0867223  
## [49] -4.8928164 -1.9471997  
##   
## $dbz  
## [1] 3.8785041 3.6201791 3.2816363 2.9729852 2.7795374 2.7154298  
## [7] 2.7385664 2.8156304 2.9662681 3.2406447 3.6470468 4.1051464  
## [13] 4.4760555 4.6229953 4.4429210 3.8670629 2.8536352 1.3919037  
## [19] -0.4611455 -2.4809594 -4.1690942 -5.0387795 -5.2431694 -5.3214194  
## [25] -5.5902494 -6.0726580 -6.5961331 -6.8908624 -6.7797700 -6.3472261  
## [31] -5.8441869 -5.4893648 -5.3972089 -5.5902645 -6.0115228 -6.5205485  
## [37] -6.9099455 -7.0021131 -6.7837096 -6.4037040 -6.0379208 -5.7941317  
## [43] -5.6990472 -5.7149544 -5.7628492 -5.7594718 -5.6656978 -5.5117628  
## [49] -5.3724938 -5.3170922

aic5.wge(BSales$ad\_tv)

## ---------WORKING... PLEASE WAIT...   
##   
##   
## Error in aic calculation at 4 2   
## Five Smallest Values of aic

## p q aic  
## 9 2 2 -2.389467  
## 14 4 1 -2.383890  
## 17 5 1 -2.381380  
## 16 5 0 -2.379176  
## 10 3 0 -2.378777

est\_ad\_tv = est.arma.wge(BSales$ad\_tv,p = 2, q = 2)

##   
## Coefficients of Original polynomial:   
## 1.1044 -0.8356   
##   
## Factor Roots Abs Recip System Freq   
## 1-1.1044B+0.8356B^2 0.6608+-0.8718i 0.9141 0.1468  
##   
##

ad\_tvFORECAST = fore.arma.wge(BSales$ad\_tv,phi = est\_ad\_tv$phi, theta = est\_ad\_tv$theta, n.ahead = 6)  
  
plotts.sample.wge(BSales$ad\_online)

## $autplt  
## [1] 1.000000000 0.162073069 0.416557696 -0.290166127 0.089087007  
## [6] -0.079288499 0.386013567 0.178574756 0.289530025 -0.127957612  
## [11] -0.009656967 -0.145773388 0.170420828 0.108360703 0.292467163  
## [16] 0.047946804 0.022060762 -0.127256593 -0.065362555 0.075978430  
## [21] 0.110344941 0.197692825 -0.062088465 -0.030316466 -0.258688586  
## [26] -0.034049262  
##   
## $freq  
## [1] 0.01 0.02 0.03 0.04 0.05 0.06 0.07 0.08 0.09 0.10 0.11 0.12 0.13 0.14 0.15  
## [16] 0.16 0.17 0.18 0.19 0.20 0.21 0.22 0.23 0.24 0.25 0.26 0.27 0.28 0.29 0.30  
## [31] 0.31 0.32 0.33 0.34 0.35 0.36 0.37 0.38 0.39 0.40 0.41 0.42 0.43 0.44 0.45  
## [46] 0.46 0.47 0.48 0.49 0.50  
##   
## $db  
## [1] 6.5807385 4.7040395 3.6773370 -1.3866483 0.3795181 1.7304282  
## [7] -11.4115005 -0.1491254 -1.2634361 0.5687462 -7.5105025 1.9186019  
## [13] -3.2885397 5.7002771 7.8576202 1.2736339 -4.1415218 -2.3950886  
## [19] -2.5896781 -3.9761856 -7.8280630 -11.6857943 -10.9778059 -12.1187977  
## [25] -8.3029988 -5.6795550 -9.1351002 -16.1805673 -6.0080978 -7.4966094  
## [31] -9.1382952 -8.6907610 -10.9341845 -11.1489592 -7.3550812 -8.7306602  
## [37] -4.9607585 -4.8189662 -2.0550239 -5.2106387 -8.5715222 -3.7150533  
## [43] 1.9343996 -2.2426154 -5.0444611 -3.5414095 1.1722727 5.7773548  
## [49] 7.9589701 -4.9199963  
##   
## $dbz  
## [1] 3.81468286 3.39461459 2.74943581 1.96579162 1.16116138 0.46121785  
## [7] -0.02354414 -0.19687915 0.01748220 0.60942440 1.43018499 2.25879462  
## [13] 2.90564849 3.24997886 3.22629801 2.80240292 1.96581800 0.72105247  
## [19] -0.90091054 -2.81581103 -4.83885265 -6.65148596 -7.90406515 -8.47965771  
## [25] -8.57000251 -8.45828792 -8.35859080 -8.38956583 -8.57627341 -8.84358583  
## [31] -9.02379369 -8.93346826 -8.51208748 -7.86270011 -7.13420642 -6.41168082  
## [37] -5.70516424 -4.99338783 -4.26661498 -3.53443549 -2.79774223 -2.01834760  
## [43] -1.12661533 -0.07913597 1.08289736 2.24284056 3.26932707 4.06131814  
## [49] 4.55698976 4.72529149

aic5.wge(BSales$ad\_online, p = 0:10)

## ---------WORKING... PLEASE WAIT...   
##   
##   
## Five Smallest Values of aic

## p q aic  
## 19 6 0 -2.860012  
## 17 5 1 -2.842250  
## 20 6 1 -2.840006  
## 22 7 0 -2.840004  
## 15 4 2 -2.838058

est\_online = est.arma.wge(BSales$ad\_online,p = 6)

##   
## Coefficients of Original polynomial:   
## 0.2212 0.5968 -0.5007 -0.1737 0.2880 0.2575   
##   
## Factor Roots Abs Recip System Freq   
## 1-1.1259B+0.8470B^2 0.6647+-0.8596i 0.9203 0.1452  
## 1-0.9008B 1.1102 0.9008 0.0000  
## 1+0.8721B -1.1466 0.8721 0.5000  
## 1+0.9333B+0.3871B^2 -1.2056+-1.0630i 0.6222 0.3850  
##   
##

while (!is.null(dev.list())) dev.off()  
plot.ts(BSales$ad\_online[1:100])  
ad\_onlineFORECAST = fore.arma.wge(BSales$ad\_online,phi = est\_online$phi, n.ahead = 6)  
  
  
  
#with trend and lagging  
  
ad\_tvFORECAST1 = lag(ad\_tvFORECAST,1)  
  
ad\_onlineFORECAST1 = lag(ad\_onlineFORECAST,1)  
  
# ASE for model with no lag and trend (last 5)  
t=1:100  
BSales$t = t  
BSales2 = BSales[2:95,]  
ksfit=lm(sales~t+ad\_tv1+ad\_online1, data = BSales2)  
aic.wge(ksfit$residuals,p=0:8,q=0:0) # AIC picks p=7

## $type  
## [1] "aic"  
##   
## $value  
## [1] 0.6685298  
##   
## $p  
## [1] 7  
##   
## $q  
## [1] 0  
##   
## $phi  
## [1] -0.02226133 0.17382426 -0.12381067 -0.03187600 -0.15737637 0.20951958  
## [7] 0.18648969  
##   
## $theta  
## [1] 0  
##   
## $vara  
## [1] 1.645949

fit = arima(BSales2$sales,order = c(7,0,0), xreg = cbind(BSales2$ad\_tv1,BSales2$ad\_online1,BSales2$t))  
fit

##   
## Call:  
## arima(x = BSales2$sales, order = c(7, 0, 0), xreg = cbind(BSales2$ad\_tv1, BSales2$ad\_online1,   
## BSales2$t))  
##   
## Coefficients:  
## ar1 ar2 ar3 ar4 ar5 ar6 ar7 intercept  
## -0.0376 0.2970 -0.1797 -0.0798 -0.1415 0.2314 0.1992 14.8719  
## s.e. 0.3058 0.1106 0.1374 0.1129 0.1144 0.1406 0.1045 6.7215  
## cbind(BSales2$ad\_tv1, BSales2$ad\_online1, BSales2$t)1  
## 3.6084  
## s.e. 0.5746  
## cbind(BSales2$ad\_tv1, BSales2$ad\_online1, BSales2$t)2  
## 4.9147  
## s.e. 1.7783  
## cbind(BSales2$ad\_tv1, BSales2$ad\_online1, BSales2$t)3  
## 0.0216  
## s.e. 0.0113  
##   
## sigma^2 estimated as 1.548: log likelihood = -154.47, aic = 332.94

preds = predict(fit, newxreg = cbind(ad\_tvFORECAST$f[2:6],ad\_onlineFORECAST$f[2:6],BSales$t[96:100]))  
ASE3.5 = mean((BSales$sales[96:100] - preds$pred[1:5])^2)  
ASE3.5

## [1] 13.06587

plot(seq(1,100,1), BSales$sales[1:100], type = "l",xlim = c(0,100), ylab = "Business Sales", main = "5 Week Sales Forecast")  
lines(seq(96,100,1), preds$pred[2:6], type = "l", col = "red")

#### VAR MODELS

**Model 4**

BSVar = VAR(cbind(BSales2$sales,BSales2$ad\_tv1,BSales2$ad\_online1), type = "both", lag.max = 10)

## Warning in VAR(cbind(BSales2$sales, BSales2$ad\_tv1, BSales2$ad\_online1), : No column names supplied in y, using: y1, y2, y3 , instead.

preds = predict(BSVar,n.ahead = 5)  
   
ASE4 = mean((BSales$sales[96:100] - preds$fcst$y1[,1])^2)  
ASE4

## [1] 24.60046

while (!is.null(dev.list())) dev.off()  
plot(seq(1,100,1), BSales$sales[1:100], type = "l",xlim = c(0,100), ylab = "Business Sales", main = "5 Week Sales Forecast")  
lines(seq(96,100,1), preds$fcst$y1[,1], type = "l", col = "red")

**Model 5**

BSVar = VAR(cbind(BSales2$sales,BSales2$ad\_tv,BSales2$ad\_online), type = "both", lag.max = 10)

## Warning in VAR(cbind(BSales2$sales, BSales2$ad\_tv, BSales2$ad\_online), type = "both", : No column names supplied in y, using: y1, y2, y3 , instead.

preds = predict(BSVar,n.ahead = 5)  
  
ASE5 = mean((BSales$sales[96:100] - preds$fcst$y1[,1])^2)  
ASE5

## [1] 21.44551

while (!is.null(dev.list())) dev.off()  
plot(seq(1,100,1), BSales$sales[1:100], type = "l",xlim = c(0,100), ylab = "Business Sales", main = "5 Week Sales Forecast")  
lines(seq(96,100,1), preds$fcst$y1[,1], type = "l", col = "red")