**6304 Module 7 Live Lecture**

**R Script File**

**rm(list=ls())**

**library(rio)**

**library(moments)**

**library(car)**

**quebec=import("6304 Module 7 Live Lecture Data.xlsx",**

**skip=3)**

**#Fixing names**

**#Creating new variables**

**colnames(quebec)=tolower(make.names(colnames(quebec)))**

**colnames(quebec)[2]="sales"**

**quebec$year=as.numeric(format(quebec$yrmo,'%Y'))**

**quebec$month=as.numeric(format(quebec$yrmo,'%m'))**

**quebec$item=seq(1:nrow(quebec))**

**attach(quebec)**

**plot(item,sales,type="o",pch=19,**

**main="Quebec Car Sales -- Raw Data")**

**base.out=lm(sales~item,data=quebec)**

**summary(base.out)**

**points(base.out$fitted.values,type="o",pch=19,col="red")**

**plot(item,sales,type="o",pch=19,**

**main="Quebec Car Sales -- Raw Data")**

**points(base.out$fitted.values,type="l",lwd=3,col="red")**

**cor(quebec$sales,base.out$fitted.values)**

**plot(quebec$item,rstandard(base.out),pch=19,type="o")**

**abline(0,0,col="red",lwd=3)**

**#Durbin Watson Test**

**durbin.out=durbinWatsonTest(base.out)**

**durbin.out**

**#Making Seasonal Indices**

**indices=data.frame(month=1:12,average=0,index=0)**

**for(i in 1:12) {**

**count=0**

**for(j in 1:nrow(quebec)) {**

**if(i==quebec$month[j]) {**

**indices$average[i]=indices$average[i]+quebec$sales[j]**

**count=count+1**

**}**

**}**

**indices$average[i]=indices$average[i]/count**

**indices$index[i]=indices$average[i]/mean(quebec$sales)}**

**#Deseasonalizing the original data**

**for(i in 1:12){**

**for(j in 1:nrow(quebec)){**

**if(i==quebec$month[j]){**

**quebec$deseason.sales[j]=quebec$sales[j]/indices$index[i]**

**}**

**}**

**}**

**#Comparison of Seasonalized and Deseasonalized**

**plot(quebec$item,quebec$sales,type="o",pch=19,**

**main="Original and Deseasonalized Data")**

**points(quebec$item,quebec$deseason.sales,type = "o",**

**pch=19,col="red")**

**#Conducting the deseasonalized regression**

**desreg.out=lm(deseason.sales~item,data=quebec)**

**summary(desreg.out)**

**plot(quebec$item,quebec$deseason.sales,type="o",pch=19,**

**main="Deseasonalized Data and Regression Model")**

**points(quebec$item,desreg.out$fitted.values,type="l",**

**lwd=3,col="red")**

**plot(quebec$item,rstandard(desreg.out),pch=19,type="o",**

**main="Deseasonalized Forecasts -- Standardized Errors")**

**abline(0,0,col="red",lwd=3)**

**#Reseasonalizing Forecasts**

**quebec$deseason.forecast=desreg.out$fitted.values**

**for(i in 1:12){**

**for(j in 1:nrow(quebec)){**

**if(i==quebec$month[j]){**

**quebec$reseason.forecast[j]=quebec$deseason.forecast[j]\***

**indices$index[i]**

**}**

**}**

**}**

**plot(quebec$item,quebec$sales,type="o",pch=19,**

**main="Original Data and Reseasonalized Forecasts")**

**points(quebec$item,quebec$reseason.forecast,**

**type="o",pch=19,col="red")**

**#Linearity**

**plot(quebec$sales,quebec$reseason.forecast,pch=19,**

**main="Reseasonalized Data Linearity Check")**

**abline(0,1,lwd=3,col="red")**

**#Creating Residuals**

**quebec$error=quebec$sales-quebec$reseason.forecast**

**quebec$stdzd.error=scale(quebec$error)**

**#Normality**

**qqnorm(quebec$stdzd.error,pch=19)**

**qqline(quebec$stdzd.error,col="red",lwd=3)**

**hist(quebec$stdzd.error,col="red",**

**main="Reseasonalized Forecasts Standardized Errors")**

**hist(quebec$stdzd.error,col="red",**

**main="Reseasonalized Forecasts Standardized Errors"**

**,ylim=c(0,.5),probability = TRUE)**

**curve(dnorm(x,mean(quebec$stdzd.error),sd(quebec$stdzd.error)),**

**from=min(quebec$stdzd.error),**

**to=max(quebec$stdzd.error),lwd=3,add=TRUE)**

**skewness(quebec$stdzd.error)**

**kurtosis((quebec$stdzd.error))**

**#Equality of Variances**

**#Plot by Time Period (Sequence)**

**par(mfrow=c(2,2))**

**plot(quebec$item,quebec$error,pch=19,type="o",**

**xlab="Time Period",ylab="Error",**

**main="Reseasonalized Forecasts -- Errors",**

**sub="By Sequence")**

**abline(0,0,col="red",lwd=3)**

**plot(quebec$item,quebec$stdzd.error,type="o",pch=19,**

**main="Reseasonalized Forecasts -- Standardized Errors",**

**xlab="Time Period",ylab="Standardized Errors",**

**sub="By Sequence")**

**abline(0,0,col="red",lwd=3)**

**#Plot by Sales**

**plot(quebec$sales,quebec$error,pch=19,**

**main="Reseasonalized Forecasts -- Errors",**

**xlab="Sales",ylab=" Errors",**

**sub="By Sales")**

**abline(0,0,col="red",lwd=3)**

**plot(quebec$sales,quebec$stdzd.error,pch=19,**

**main="Reseasonalized Forecasts -- Standardized Errors",**

**xlab="Sales",ylab="Standardized Errors",**

**sub="By Sales")**

**abline(0,0,col="red",lwd=3)**

**par(mfrow=c(1,1))**

**# Stepwise Regression**

**rm(list=ls())**

**midwest=import("6304 Module 7 Live Lecture Data.xlsx",**

**sheet="Midwest Data")**

**reduced.midwest=midwest[,-c(1,2,3,11,13,20)]**

**attach(reduced.midwest)**

**names(reduced.midwest)**

**gilligan=lm(percchildbelowpovert~.,data=reduced.midwest)**

**summary(gilligan)**

**plot(reduced.midwest$percchildbelowpovert,gilligan$fitted.values,**

**pch=19,main=paste("Actuals v. Forecasts of Gilligan r = ",**

**round(cor(reduced.midwest$percchildbelowpovert,**

**gilligan$fitted.values),4)))**

**abline(0,1,col="red",lwd=3)**

**skipper=step(lm(percchildbelowpovert~.,**

**data=reduced.midwest),**

**direction=c("backward"))**

**summary(gilligan)**

**summary(skipper)**

**plot(reduced.midwest$percchildbelowpovert,skipper$fitted.values,**

**pch=19,main=paste("Actuals v. Forecasts of Skipper r = ",**

**round(cor(reduced.midwest$percchildbelowpovert,**

**skipper$fitted.values),4)))**

**abline(0,1,col="red",lwd=3)**

**backward.test=step(lm(percchildbelowpovert~.,**

**data=reduced.midwest),**

**direction=c("backward"))**

**forward.test=step(lm(percchildbelowpovert~.,**

**data=reduced.midwest),**

**direction=c("forward"))**

**both.test=step(lm(percchildbelowpovert~.,**

**data=reduced.midwest),**

**direction=c("both"))**

**summary(backward.test)**

**summary(forward.test)**

**summary(both.test)**