**6304 Module 4**

**Live Lecture R Command File**

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

**library(rio)**

**oil=import("Module 4 Data Sets.xlsx",sheet="Oil and Gas")**

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

**attach(oil)**

**plot(crude,gasoline,pch=19,main="Oil & Gas Raw Data Plot")**

**oilout=lm(gasoline~crude,data=oil)**

**summary(oilout)**

**confint(oilout)**

**plot(crude,gasoline,pch=19,main="Oil & Gas Raw Data Plot")**

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

**plot(resid(oilout),pch=19,main="O&G Residual Plot by Order of Data")**

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

**plot(oilout$fitted.values,oilout$residuals,pch=19,main="O&G Plot by Fitted Values")**

**plot(rstandard(oilout),pch=19,main="Oil & Gas Standardized Residuals")**

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

**#Linearity**

**plot(oil$gasoline,oilout$fitted.values,pch=19,main="O&G Actual v. Fitted Values")**

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

**#Normality**

**qqnorm(oilout$residuals,pch=19,main="O&G Normality Plot")**

**qqline(oilout$residuals,col="red",lwd=3)**

**#Equality of Variances**

**plot(oilout$fitted.values,oilout$residuals,pch=19,**

**main="O&G Residuals")**

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

**#OR**

**plot(oilout$fitted.values,rstandard(oilout),pch=19,**

**main="O&G Standardized Residuals")**

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

**# New Data Set**

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

**tools=import("Module 4 Data Sets.xlsx",sheet="Cutting Tools")**

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

**attach(tools)**

**brand.a.out=lm(brand.a~speed,data=tools)**

**brand.b.out=lm(brand.b~speed,data=tools)**

**summary(brand.a.out)**

**summary(brand.b.out)**

**brand.a.out$coefficients**

**brand.b.out$coefficients**

**plot(speed,brand.a,ylim=c(0,7),xlim=c(30,80),pch=19,**

**main="Cutting Tools Plot")**

**points(speed,brand.b,col="red",pch=19)**

**abline(brand.a.out,lwd=3)**

**abline(brand.b.out,lwd=3,col="red")**

**cor(speed,brand.a);cor(speed,brand.b)**

**plot(speed,rstandard(brand.a.out),pch=19,ylim=c(-4,4),**

**main="Cutting Tools Std. Residual Plot")**

**points(speed,rstandard(brand.b.out),pch=19,col="red")**

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

**# A Data Set with Random X and Y Values**

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

**x=rnorm(1000,100,10)**

**y=rnorm(1000,200,20)**

**plot(x,y,pch=19,main="Shotgun Blast")**

**cor(x,y)**

**cor(x,y)^2**

**regout=lm(y~x)**

**summary(regout)**

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

**qqnorm(resid(regout),pch=19)**

**qqline(resid(regout),col="red",lwd=3)**

**plot(y,rstandard(regout),pch=19,main="Standardized Residuals")**

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

**# An Exponential Pattern**

**x=rnorm(1000,100,10)**

**y=x^5**

**plot(x,y,pch=19,xlim=c(0,150),**

**main="Exponential Relationship")**

**cor(x,y)**

**regout=lm(y~x)**

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

**sum(regout$residuals)**

**qqnorm(resid(regout),pch=19)**

**qqline(resid(regout),col="red",lwd=3)**

**plot(regout$fitted.values,rstandard(regout),pch=19,**

**main="Exponential Model, Standardized Residuals")**

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

**# Abuse Data**

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

**abuse=import("Module 4 Data Sets.xlsx",sheet="Child Abuse")**

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

**attach(abuse)**

**cor(under.18,victims)**

**cor(under.18,victims)^2**

**abuseout=lm(victims~under.18,data=abuse)**

**summary(abuseout)**

**plot(under.18,victims,pch=19,main="Child Abuse Data")**

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

**#Linearity**

**plot(abuse$victims,abuseout$fitted.values,pch=19,main="Abuse Actual v. Fitted Values")**

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

**#Normality**

**qqnorm(abuseout$residuals,pch=19,main="Abuse Normality Plot")**

**qqline(abuseout$residuals,col="red",lwd=3)**

**#Equality of Variances**

**plot(abuseout$fitted.values,rstandard(abuseout),pch=19,**

**main="Abuse Standardized Residuals")**

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

**#Identifying high leverage points.**

**lev=hat(model.matrix(abuseout))**

**plot(lev,pch=19,main="Leverage Plot, Abuse Data")**

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

**abuse[lev>(3\*mean(lev)),]**

**abuse[lev>(3\*mean(lev)),1]**

**newdata=data.frame(under.18=500000)**

**predict(abuseout,newdata,interval="predict")**

**predict(abuseout,newdata,interval="confidence")**