**CODES :**

Ins<-read.csv(file.choose(),header=T)

**Q1**

head(Ins)

#to know each field of data collected through descriptive analysis

#to gain basics insight into data set

summary(Ins)

**Q2**

MDLlm<-lm(Ins$Payment~Ins$Claims+Ins$Insured)

MDLlm

summary(MDLlm)

#as per the pvalue is less,payment is related to no.of claim & no of insured policy per year

#visualisation

plot(Ins$Payment,Ins$Claims,col="red")

plot(Ins$Payment,Ins$Insured,col="blue")

**Q3**

MDL1lm2<-lm(Ins$Payment~.,data=Ins)

MDL1lm2

summary(MDL1lm2)

##results

##increase or decrease of payment related to

#distance,insured,claims,less related to location

**Q4**

grpzone<-apply(Ins[,c(5,6,7)],2,function(x)tapply(x,Ins$Zone,mean))

grpzone

#results

#zone 4 has highest nos of claims, payments & insured policy

grpkm<-apply(Ins[,c(5,6,7)],2,function(x)tapply(x,Ins$Kilometres,mean)

grpkm

#result

#claims and payments have highest value at kilometer group 2

#the insured no of years has highest value at group 1

grpbns<-apply(Ins[,c(5,6,7)],2,function(x)tapply(x,Ins$Bonus,mean))

grpbns

#result

#bonus group 7 has maximum insured policies, claims & payments.

**Q5**

MDL5lm5<-lm(Claims~Kilometres+Zone+Bonus+Make+Insured,data=Ins)

summary(MDL5lm5)

#result

#zone,bonus,make,insured policies affected highly to claims, distance affected less.