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

str(task)

library(caTools)

split<-sample.split(task$Scores,SplitRatio=0.7)

trainingset=subset(task,split==TRUE)

testset=subset(task,split==FALSE)

lm.r=lm(formula=Scores~Hours,data=trainingset)

coef(lm.r)

ypred<-predict(lm.r,newdata=testset)

library(ggplot2)

ggplot()+geom\_point(aes(x=trainingset$Hours,y=trainingset$Scores),colour='red')+geom\_line(aes(x=trainingset$Hours,y=predict(lm.r,newdata=trainingset)),colour='blue')+ggtitle('Linear Regression(Training set)')+xlab('Hours')+ylab('Scores')

ggplot()+geom\_point(aes(x=testset$Hours,y=testset$Scores),colour='red')+geom\_line(aes(x=testset$Hours,y=predict(lm.r,newdata=testset)),colour='blue')+ggtitle('Linear Regression(Test set)')+xlab('Hours')+ylab('Scores')