Code for Project 1:

library(MASS)

attach(Boston)

set.seed(1)

train=sample(506,253,replace=FALSE)

training=Boston[train,]

testing=Boston[-train,]

library(leaps)

regfit.full=regsubsets(medv~.,training,nvmax=13)

reg.summary=summary(regfit.full)

reg.summary

par(mfrow=c(2,2))

plot(reg.summary$rss ,xlab=" Number of Variables ", ylab=" RSS", type="l")

plot(reg.summary$adjr2 ,xlab =" Number of Variables ", ylab=" Adjusted RSq",type="l")

which.max (reg.summary$adjr2)

points (11, reg.summary$adjr2[11], col ="red",cex =2, pch =20)

lm.fit1=lm(medv~crim+zn+chas+nox+rm+dis+rad+tax+ptratio+black+lstat,data=training)

summary(lm.fit1)

plot(lm.fit1)

t.test(rm)

t.test(dis)

t.test(tax)

bestsub.pred=predict(lm.fit1,newx=x[testing,])

mean((medv-bestsub.pred)^2)

For LASSO:

library(glmnet)

x=model.matrix(medv~.,Boston)

y=Boston$medv

set.seed(1)

train=sample(1:nrow(x),nrow(x)/2)

test=(-train)

y.test=y[test]

lasso.mod=glmnet(x[train,],y[train],alpha=1)

plot(lasso.mod)

set.seed(1)

cv.out=cv.glmnet(x[train,],y[train],alpha=1)

bestlam=cv.out$lambda.min

bestlam

out=glmnet(x,y,alpha=1)

lasso.coef=predict(out,type="coefficients",s=bestlam )[1:15,]

lasso.coef

install.packages("hdi", repos="http://R-Forge.R-project.org")

library(hdi)

hdi(rm)

hdi(tax)

hdi(dis)

lasso.pred=predict(lasso.mod,s=bestlam,newx=x[test,])

mean((lasso.pred-y.test)^2)

For ridge:

ridge.mod=glmnet(x[train,],y[train],alpha=0)

set.seed(1)

cv.out1=cv.glmnet(x[train,],y[train],alpha=0)

bestlam1=cv.out1$lambda.min

bestlam1

out1=glmnet (x,y,alpha =0)

predict(out1,type="coefficients",s=bestlam1)[1:15,]

hdi(rm)

hdi(lstat)

hdi(nox)

ridge.pred=predict(ridge.mod,s=bestlam1,newx=x[test,])

mean((ridge.pred-y.test)^2)