# multiple regression

ozone.pollution <- read.csv("c:\\temp\\ozone.data.csv")

attach(ozone.pollution)

names(ozone.pollution)

pairs(ozone.pollution,panel=panel.smooth)

library(mgcv)

par(mfrow=c(2,2))

model <- gam(ozone~s(rad)+s(temp)+s(wind))

plot(model,col= "blue")

par(mfrow=c(1,1))

library(tree)

model <- tree(ozone~.,data=ozone.pollution)

plot(model)

text(model)

model1 <- lm(ozone~temp\*wind\*rad+I(rad^2)+I(temp^2)+I(wind^2))

summary(model1)

model2 <- update(model1,~. – temp:wind:rad)

summary(model2)

model3 <- update(model2,~. - wind:rad)

summary(model3)

model4 <- update(model3,~. - temp:wind)

summary(model4)

model5 <- update(model4,~. - I(rad^2))

summary(model5)

model6 <- update(model5,~. - temp:rad)

summary(model6)

plot(model6)

# start all over again with a new transfrmation of the response

model7 <- lm(log(ozone)~temp\*wind\*rad+I(rad^2)+I(temp^2)+I(wind^2))

model8 <- step(model7)

summary(model8)

plot(model8)

# a more tricky example

pollute <- read.csv("c:\\temp\\sulphur.dioxide.csv")

attach(pollute)

names(pollute)

pairs(pollute,panel=panel.smooth)

par(mfrow=c(1,1))

library(tree)

model <- tree(Pollution~.,data=pollute)

plot(model)

text(model)

model1 <- lm(Pollution~Temp+I(Temp^2)+Industry+I(Industry^2)+Population+I(Population^2)+Wind+I(Wind^2)+Rain+I(Rain^2)+Wet.days+I(Wet.days^2))

summary(model1)

model2 <- step(model1)

summary(model2)

model3 <- update(model2, ~.- Rain-I(Wind^2))

summary(model3)

interactions <- c("ti","tp","tw","tr","td","ip","iw",

"ir","id","pw","pr","pd","wr","wd","rd")

sample(interactions)

model4 <- lm(Pollution~Temp+Industry+Population+Wind+Rain+Wet.days+Wind:Rain+Wind:Wet.days+Industry:Wet.days+Industry:Rain+Rain:Wet.days)

model5 <- lm(Pollution~Temp+Industry+Population+Wind+Rain+Wet.days+Population:Rain+Temp:Population+Population:Wind+Temp:Industry+Industry:Wind)

model6 <- lm(Pollution~Temp+Industry+Population+Wind+Rain+Wet.days+Temp:Wind+Population:Wet.days+Temp:Rain+Temp:Wet.days+Industry:Population)

model7 <- lm(Pollution~Temp+Industry+Population+Wind+Rain+Wet.days+Wind:Rain+Wind:Wet.days+Population:Wind+Temp:Rain)

summary(model7)

model8 <- update(model7,~.-Temp:Rain)

summary(model8)

model9 <- update(model8,~.-Population:Wind)

summary(model9)

plot(model9)

model10 <- update(model9,~. + Wind:Rain:Wet.days)

summary(model10)