assign2 <- function(data, Xindex, Yindex){

#set the figure to a 2 by 2 plot

par(mfrow=c(2,2))

#X is the column in xindex, Y is the column in yindex

X <- data[[Xindex]]

Y <- sqrt(data[[Yindex]])

#plot based on X and Y

plot(X,Y,main='Raw data and smooths')

#add lines on the plot, one is the smoothing spline with default cross validation, the other is with 2 df

lines(smooth.spline(X,Y))

lines(smooth.spline(X,Y,df=2),col=2)

#store the smoothing spline to smsp.strcv

smsp.strcv<-smooth.spline(X,Y)

#approx y based on the result of smoothing spline and assign the difference to resid

smspcv.resid<-Y-approx(smsp.strcv$x,smsp.strcv$y,X)$y

#store the smoothing spline with 2 df to str2

smsp.str2<-smooth.spline(X,Y,df=2)

#approx y based on the smoothing spline with 2 df

smsp2.resid<-Y-approx(smsp.str2$x,smsp.str2$y,X)$y

#get the sum of the squared residual from full model

SSF<-sum(smspcv.resid^2)

#get the sum of the sqaured residual from nested model

SSN<-sum(smsp2.resid^2)

#get the F test result based on df=2

Fstat <- ((SSN-SSF)/(smsp.strcv$df-2))/(SSF/(nrow(NOAA)-2))

#get the p value

pfstat <- 1- pf(Fstat,floor((smsp.strcv$df-2)),(nrow(data)-2))

#qqnorm plot on resid of full model

qqnorm(smspcv.resid,main = 'residuals smooth')

#qqnorm plot on resid of nested model

qqnorm(smsp2.resid,main = 'residuals linear')

#list the f test value and p value, get the df num for numerator and get the df for denominator

list(Fstatistic=Fstat,p = pfstat, dfnum=smsp.strcv$df-2, dfden = length(data[[1]])-2)

}

> assign2(NOAA,3,2)

$Fstatistic

[1] 3.050626

$p

[1] 0.08947794

$dfnum

[1] 1.132256

$dfden

[1] 35 