**STAT 40001/MA59800 Statistical Computing Fall 2017**

**Lab -18**

The gross domestic product (GDP) of the United States in trillions of dollars from 1950- 2013 are provided in the link below

http://media.pearsoncmg.com/aw/aw\_sharpe\_business\_3/datasets/txt/GDP\_2013.txt

a) Display the data using a scatterplot.

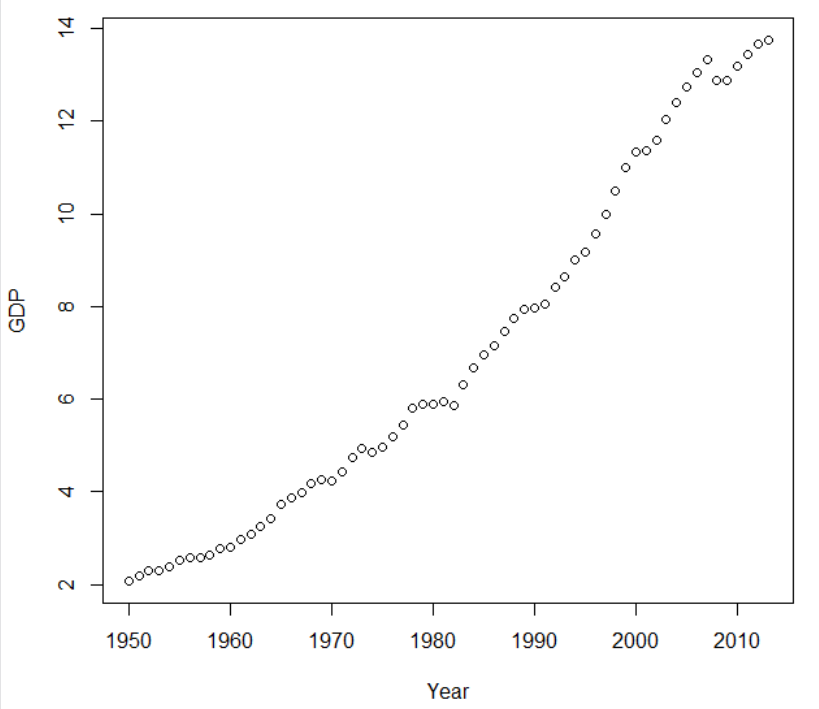
> data = read.table('C:\\Users\\Administrator\\Documents\\git\\R\\statistical computing\\lab18\\GDP.txt',header = T)

> attach(data)

> scatter(GDP~Year)

Error in scatter(GDP ~ Year) : could not find function "scatter"

> plot(GDP~Year)



b) Fit a simple linear regression model.

> model = lm(GDP~Year)

> model

Call:

lm(formula = GDP ~ Year)

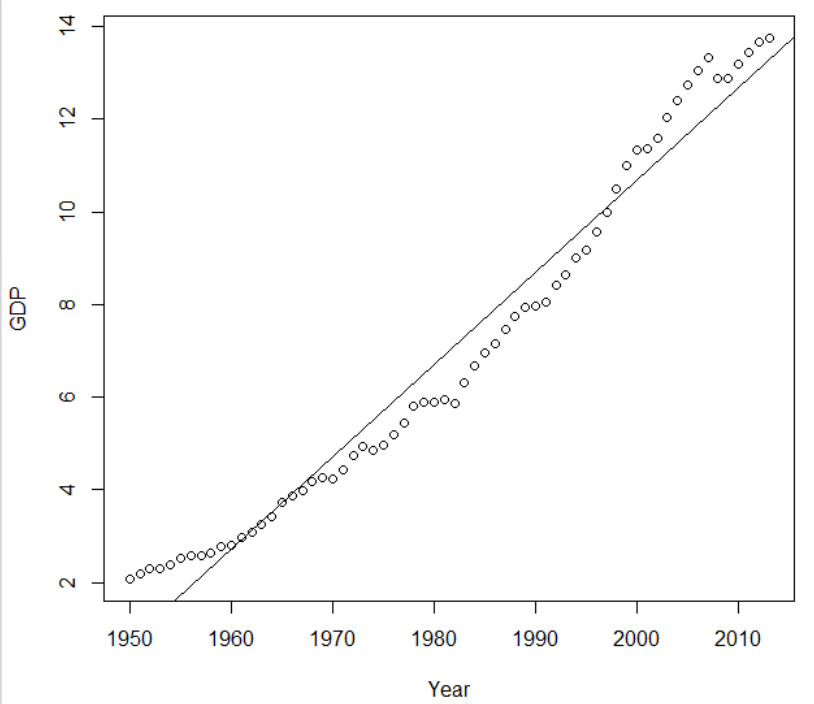
Coefficients:

(Intercept) Year

-387.8433 0.1993

c) Add the fitted line to the scatter plot.

> abline(model)



d) Determine the coefficient of determination.

> coef(model)

(Intercept) Year

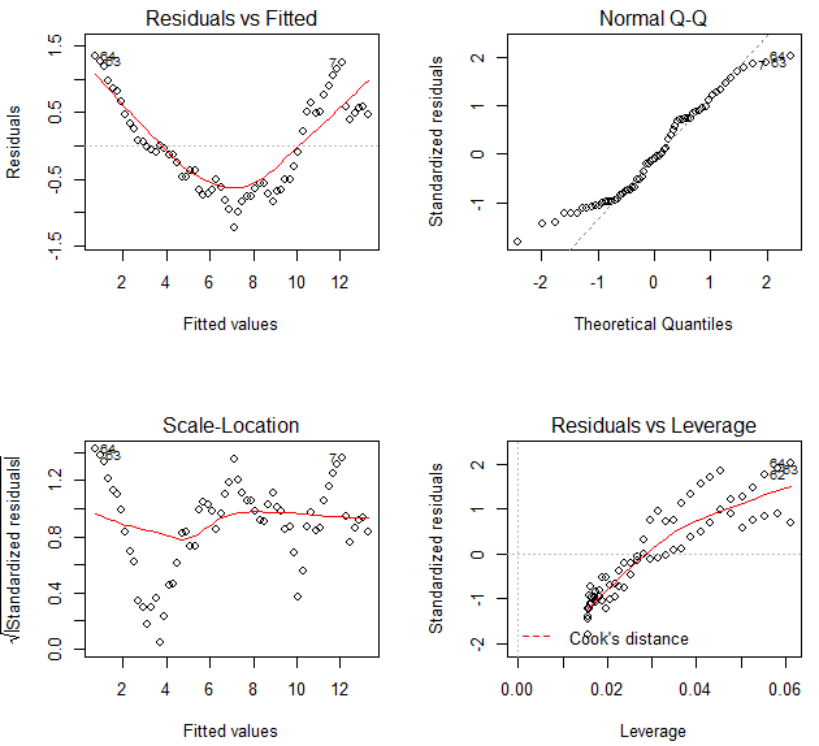
-387.8433063 0.1992681

Y\_hat = -387.84 + 0.199\*Year

d) Analyze the residual plots. Is your model questionable?

> par(mfrow=c(2,2))

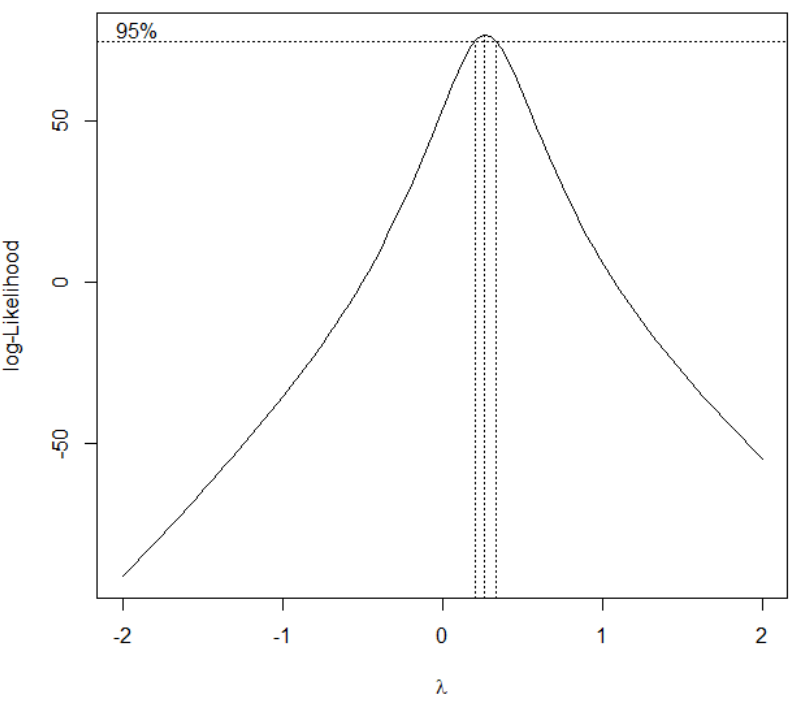
> plot(model)



Model is questionable

e) Use Box-Cox Transformation to see whether this model can be improved.

> b = boxcox(model)



λ = 0.25

> GDP1=GDP^(0.25)

> plot(Year,GDP1)

