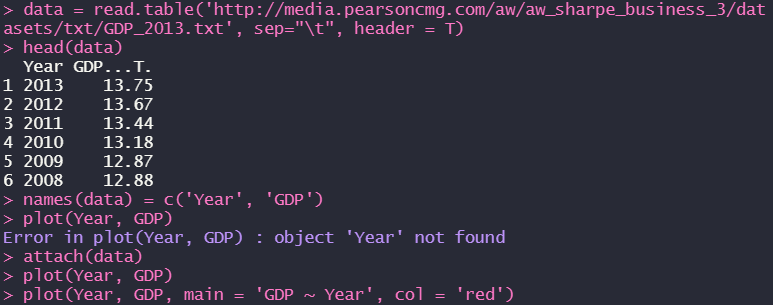
**STAT 40001/ STAT 59800 Statistical Computing Fall 2020**

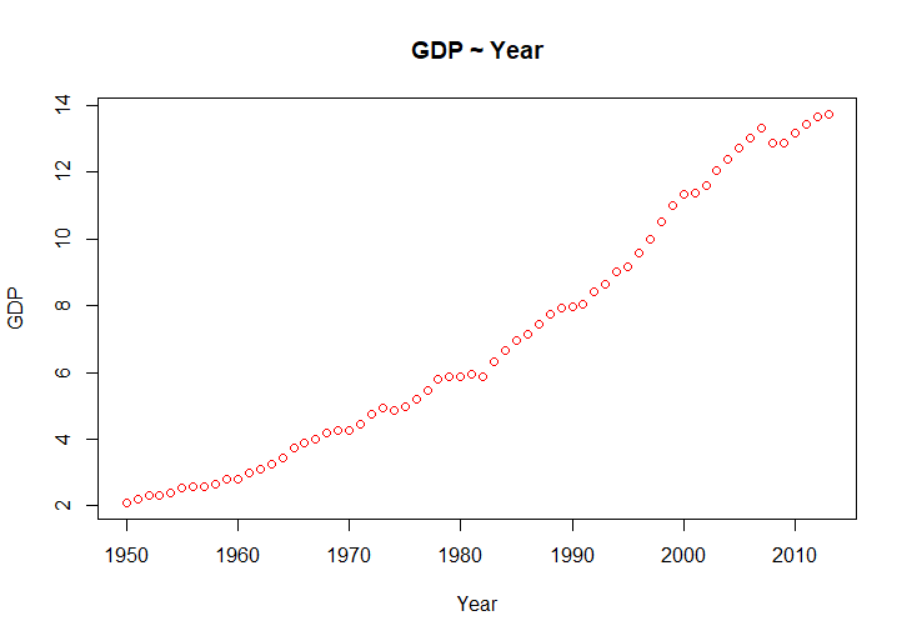
**Lab -18**

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

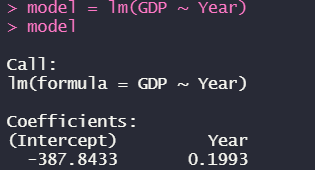
<http://media.pearsoncmg.com/aw/aw_sharpe_business_3/datasets/txt/GDP_2013.txt>

a) Display the data using a scatterplot.

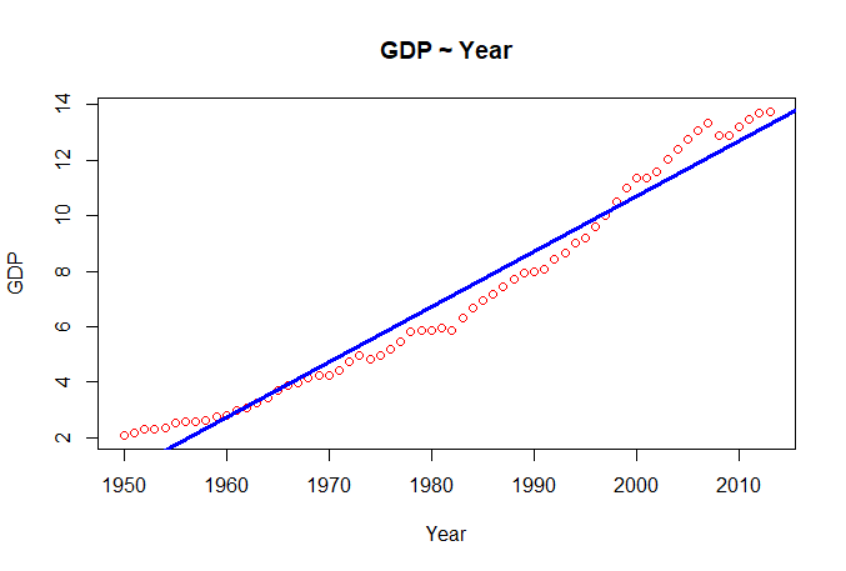




b) Fit a simple linear regression model.

  
*(GDP = -387.8433 + 0.1993 \* Year)*

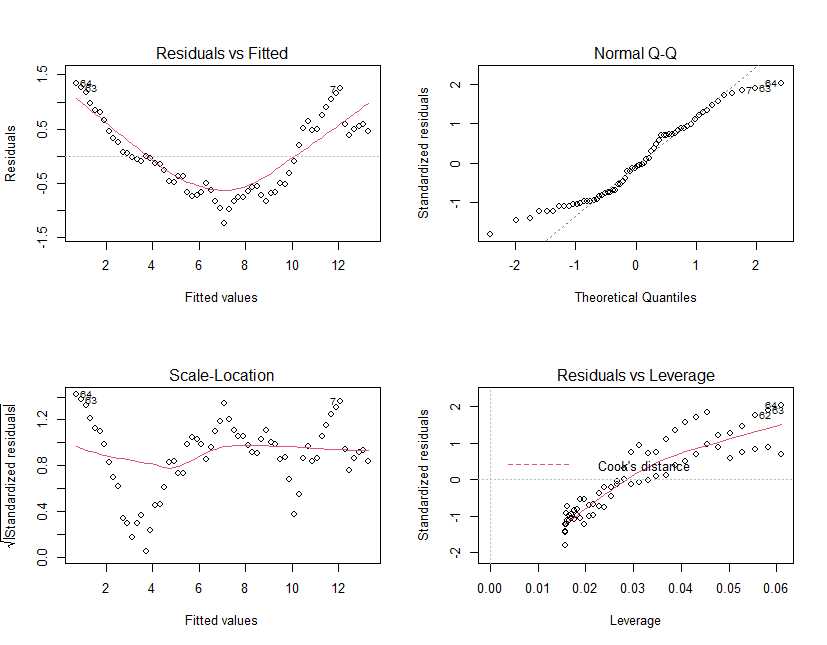
c) Add the fitted line to the scatter plot.

d) Determine the coefficient of determination.

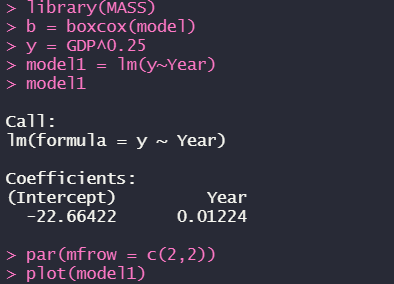


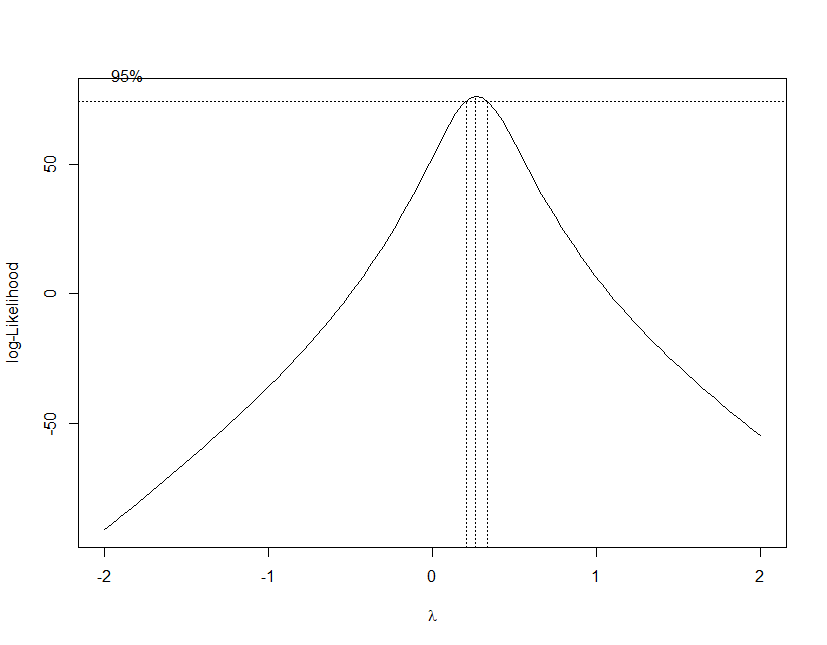
d) Analyze the residual plots. Is your model questionable?

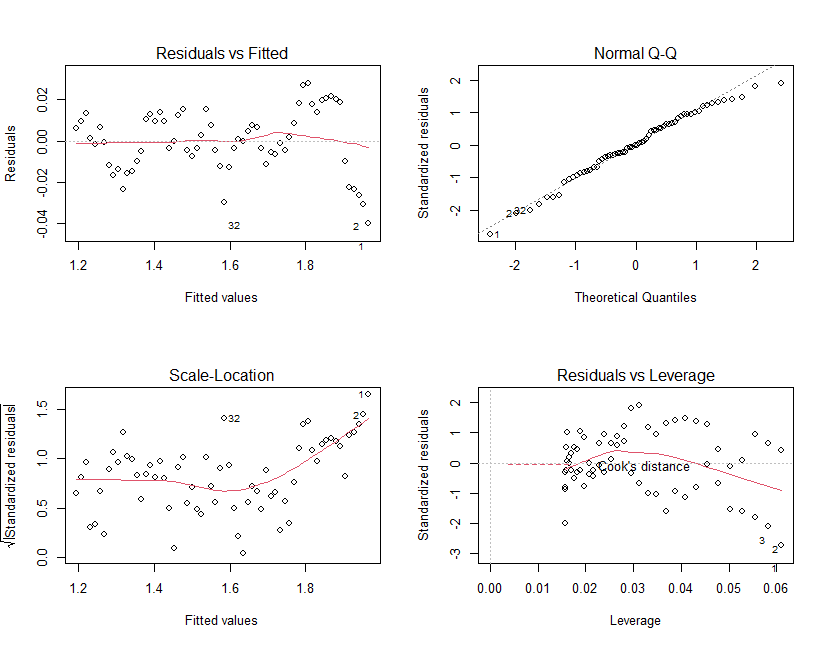
  


*(For the first plot, we can see that the pattern is U-shaped rather linear, which means that the model can be optimized)*

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







*(After changing the λ value into ¼ of the previous GDP, from the residual plot we can see that the model has been obviously improved)*