Due Date: 10/17 by 10:20AM

Notes:
1. Make sure you label all plots in the title and then use them in the context.
2. Make sure you do not put R code in the middle of each problem, instead you present the entire solution as a report.

We will work with the data on median housing prices in neighborhoods in the suburbs of Boston.

install.packages("MASS")

library(MASS)

data(Boston)

Our goal is to analyze the relationship between neighborhood poverty and housing prices.

Using dataset Boston in the R package MASS answer the following:
a.) [5 points] Make a scatterplot with x as % in Poverty using variable lstat and y as the Median home prices using variable medv. Comment on the relationship.
b.) [5 points] Regress medv (Y) on lstat (X) and report the fitted model.
c.) [7.5 points] Share goodness of this fit.
d.) [15 points] Report if the LINE conditions are met or not for model in part (c).
e.) [2.5 points] Using residual plot, can you tell which assumption(s) are failing?
f.) [40 points] Log transformation: To see if taking the log of poverty percentage (lstat) and median home values (medv) adjusts the relationship below.
   - [5 points] Regress log(medv) on log(lstat) and report the fitted model.
   - [5 points] Regress log(medv) on lstat and report the fitted model.
   - [5 points] Regress medv on log(lstat) and report the fitted model.
   - [5 points] Compare models (a), (b), and (c) using goodness of fit measures.
   - [15 points] Compare models (a), (b), and (c) using LINE conditions
   - [5 points] Compare what changed in residual plot reported in part (e)
   - [5 points] Did the transformation help with your analysis, why or why not?
g.) [20 points] Polynomial model: Regress medv on lstat and lstat squared, that is a polynomial model. Report the fitted model. Share goodness of this fit. Check LINE conditions for this model. What changed in the residual plot reported in (e)? Did the transformation help with your analysis, why or why not?