STOR 455 Homework #9

Due 11/29

library(readr)  
library(Stat2Data)  
"SleepStudy" <- read\_csv("https://raw.githubusercontent.com/JA-McLean/STOR455/master/data/SleepStudy.csv")

## Rows: 253 Columns: 27

## -- Column specification --------------------------------------------------------  
## Delimiter: ","  
## chr (5): LarkOwl, DepressionStatus, AnxietyStatus, Stress, AlcoholUse  
## dbl (22): Gender, ClassYear, NumEarlyClass, EarlyClass, GPA, ClassesMissed, ...

##   
## i Use `spec()` to retrieve the full column specification for this data.  
## i Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

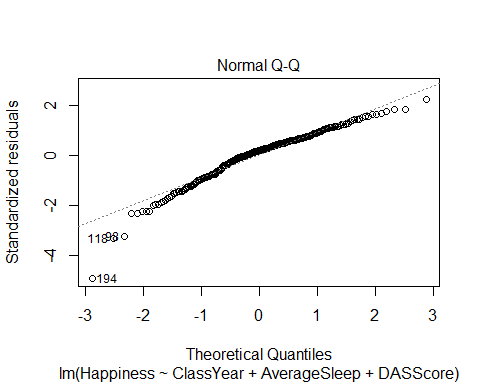
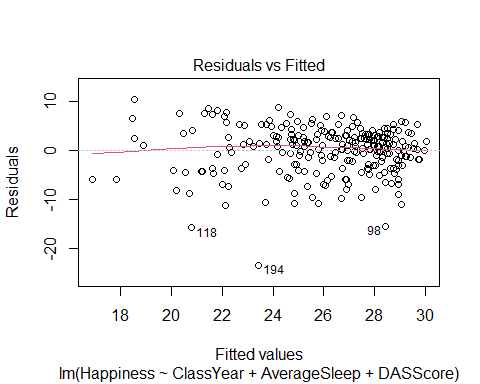
1. Multiple Regression: Building a Model & Assessing Conditions
2. Construct a linear model with *Happiness* as the response and *ClassYear* , *AverageSleep* , *DASScore* as the predictors. Include the output for the summary of the model.

mod1 = lm(Happiness~ClassYear + AverageSleep + DASScore, data=SleepStudy)  
summary(mod1)

##   
## Call:  
## lm(formula = Happiness ~ ClassYear + AverageSleep + DASScore,   
## data = SleepStudy)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -23.4348 -2.8943 0.8034 3.0193 10.4424   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 25.34113 2.68833 9.426 < 2e-16 \*\*\*  
## ClassYear 0.39920 0.29229 1.366 0.173   
## AverageSleep 0.37165 0.31511 1.179 0.239   
## DASScore -0.15870 0.01839 -8.630 7.37e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.81 on 249 degrees of freedom  
## Multiple R-squared: 0.2412, Adjusted R-squared: 0.2321   
## F-statistic: 26.39 on 3 and 249 DF, p-value: 7.432e-15

1. Asses the conditions for linearity of the model you constructed in question 1

plot(mod1, 1:2)

 Looking at our residuals vs. fitted plot, we can see that there seem to be some issues with constant variance, as the data is more compact towards the larger fitted values. Again looking at the residuals vs. fitted plot, there are not issues with linearity as there is not a clear, defined non-linear pattern. Looking at the normal quantile plot, there seem to be issues of normality with the tails of the data, especially the lower tail, as it does not seem to quite fit a normal distribution.