R Notebook

library(readr)  
  
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.

Sleepstudy

## # A tibble: 253 x 27  
## Gender ClassYear LarkOwl NumEarlyClass EarlyClass GPA ClassesMissed  
## <dbl> <dbl> <chr> <dbl> <dbl> <dbl> <dbl>  
## 1 0 4 Neither 0 0 3.6 0  
## 2 0 4 Neither 2 1 3.24 0  
## 3 0 4 Owl 0 0 2.97 12  
## 4 0 1 Lark 5 1 3.76 0  
## 5 0 4 Owl 0 0 3.2 4  
## 6 1 4 Neither 0 0 3.5 0  
## 7 1 2 Lark 2 1 3.35 2  
## 8 0 2 Lark 0 0 3 0  
## 9 0 1 Neither 2 1 4 0  
## 10 0 4 Neither 2 1 2.9 0  
## # ... with 243 more rows, and 20 more variables: CognitionZscore <dbl>,  
## # PoorSleepQuality <dbl>, DepressionScore <dbl>, AnxietyScore <dbl>,  
## # StressScore <dbl>, DepressionStatus <chr>, AnxietyStatus <chr>,  
## # Stress <chr>, DASScore <dbl>, Happiness <dbl>, AlcoholUse <chr>,  
## # Drinks <dbl>, WeekdayBed <dbl>, WeekdayRise <dbl>, WeekdaySleep <dbl>,  
## # WeekendBed <dbl>, WeekendRise <dbl>, WeekendSleep <dbl>,  
## # AverageSleep <dbl>, AllNighter <dbl>

summary(Sleepstudy)

## Gender ClassYear LarkOwl NumEarlyClass   
## Min. :0.0000 Min. :1.000 Length:253 Min. :0.000   
## 1st Qu.:0.0000 1st Qu.:2.000 Class :character 1st Qu.:0.000   
## Median :0.0000 Median :2.000 Mode :character Median :2.000   
## Mean :0.4032 Mean :2.478 Mean :1.735   
## 3rd Qu.:1.0000 3rd Qu.:3.000 3rd Qu.:3.000   
## Max. :1.0000 Max. :4.000 Max. :5.000   
## EarlyClass GPA ClassesMissed CognitionZscore   
## Min. :0.000 Min. :2.000 Min. : 0.000 Min. :-1.62e+00   
## 1st Qu.:0.000 1st Qu.:3.000 1st Qu.: 0.000 1st Qu.:-4.80e-01   
## Median :1.000 Median :3.300 Median : 1.000 Median :-1.00e-02   
## Mean :0.664 Mean :3.244 Mean : 2.209 Mean :-3.95e-05   
## 3rd Qu.:1.000 3rd Qu.:3.500 3rd Qu.: 3.000 3rd Qu.: 4.40e-01   
## Max. :1.000 Max. :4.000 Max. :20.000 Max. : 1.96e+00   
## PoorSleepQuality DepressionScore AnxietyScore StressScore   
## Min. : 1.000 Min. : 0.000 Min. : 0.000 Min. : 0.000   
## 1st Qu.: 4.000 1st Qu.: 1.000 1st Qu.: 1.000 1st Qu.: 3.000   
## Median : 6.000 Median : 3.000 Median : 4.000 Median : 8.000   
## Mean : 6.257 Mean : 5.202 Mean : 5.372 Mean : 9.466   
## 3rd Qu.: 8.000 3rd Qu.: 7.000 3rd Qu.: 8.000 3rd Qu.:14.000   
## Max. :18.000 Max. :35.000 Max. :26.000 Max. :37.000   
## DepressionStatus AnxietyStatus Stress DASScore   
## Length:253 Length:253 Length:253 Min. : 0.00   
## Class :character Class :character Class :character 1st Qu.: 7.00   
## Mode :character Mode :character Mode :character Median :16.00   
## Mean :20.04   
## 3rd Qu.:28.00   
## Max. :82.00   
## Happiness AlcoholUse Drinks WeekdayBed   
## Min. : 0.00 Length:253 Min. : 0.000 Min. :21.80   
## 1st Qu.:24.00 Class :character 1st Qu.: 3.000 1st Qu.:24.20   
## Median :28.00 Mode :character Median : 5.000 Median :24.80   
## Mean :26.11 Mean : 5.569 Mean :24.85   
## 3rd Qu.:30.00 3rd Qu.: 8.000 3rd Qu.:25.50   
## Max. :35.00 Max. :24.000 Max. :29.10   
## WeekdayRise WeekdaySleep WeekendBed WeekendRise   
## Min. : 5.500 Min. : 3.000 Min. :21.50 Min. : 5.25   
## 1st Qu.: 8.000 1st Qu.: 7.200 1st Qu.:24.88 1st Qu.: 9.25   
## Median : 8.500 Median : 7.950 Median :25.50 Median :10.25   
## Mean : 8.586 Mean : 7.866 Mean :25.58 Mean :10.20   
## 3rd Qu.: 9.150 3rd Qu.: 8.600 3rd Qu.:26.25 3rd Qu.:11.00   
## Max. :12.020 Max. :10.970 Max. :30.25 Max. :15.00   
## WeekendSleep AverageSleep AllNighter   
## Min. : 4.000 Min. : 4.950 Min. :0.0000   
## 1st Qu.: 7.250 1st Qu.: 7.430 1st Qu.:0.0000   
## Median : 8.250 Median : 8.000 Median :0.0000   
## Mean : 8.217 Mean : 7.966 Mean :0.1344   
## 3rd Qu.: 9.250 3rd Qu.: 8.590 3rd Qu.:0.0000   
## Max. :12.750 Max. :10.620 Max. :1.0000

1. Construct a two-way ANOVA model for the mean GPA in the dataset, *GPA*, using the *LarkOwl* and *Stress* as the predictors. Include the output showing the ANOVA table. Comment on what this output tells you about the GPA across college stduents. Provide formal hypotheses, p-values, and conclusions.

amodA = aov(GPA~LarkOwl+Stress, data=Sleepstudy)  
summary(amodA)

## Df Sum Sq Mean Sq F value Pr(>F)   
## LarkOwl 2 0.43 0.2150 1.347 0.262   
## Stress 1 1.02 1.0215 6.401 0.012 \*  
## Residuals 249 39.74 0.1596   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Hypotheses:

H0: μ1 = μ2 = μ3 (1= Lark, 2 = Owl, 3= Neither) (Mean GPA across LarkOwl status are all same) Ha: Some μi != μk (There is at least one mean GPA of one LarkOwl status that is different from the mean GPA of another LarkOwl status)

P-value: 0.262

Conclusion: Because p-value is not small enough (0.262), we fail to reject the null hypothesis. Mean GPA across LarkOwl statuses are all same.

Hypotheses:

H0: α1 = α2 = 0 (1 = Normal, 2 = High) (The affect for both stress scores are zeros) Ha: One of αi != 0 (There is at least one stress score that does not have zero for the affect)

p-value: 0.012

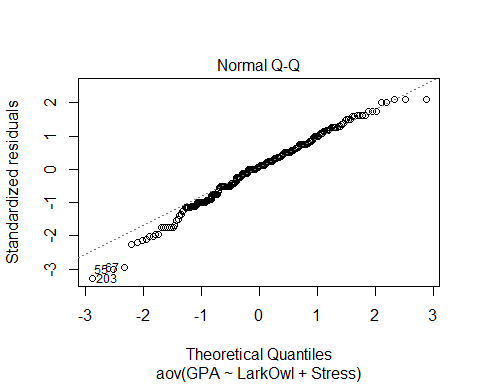
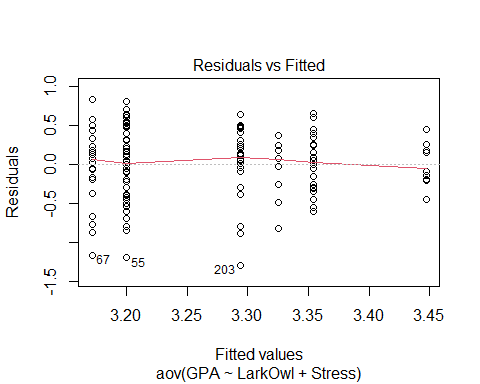
Conclusion: Because our p-value is small enough (0.012), we reject the null hypothesis. There is at least one stress score that does not have zero for the affect.

1. Construct residual plots and comment on the conditions of equality of variances and normality of residuals for the model that you created.

amodA

## Call:  
## aov(formula = GPA ~ LarkOwl + Stress, data = Sleepstudy)  
##   
## Terms:  
## LarkOwl Stress Residuals  
## Sum of Squares 0.43001 1.02147 39.73688  
## Deg. of Freedom 2 1 249  
##   
## Residual standard error: 0.399482  
## Estimated effects may be unbalanced

plot(amodA, 1:2)



tapply(Sleepstudy$GPA, Sleepstudy$LarkOwl, sd)

## Lark Neither Owl   
## 0.3972581 0.3935868 0.4413586

tapply(Sleepstudy$GPA, Sleepstudy$Stress, sd)

## high normal   
## 0.3513277 0.4123180

Variances: In general, we have constant variances, but as fitted values increase, the variances (vertical distances) are decreasing. In addition, the standard deviations for different combinations of LarkOwl and Stress predictors are roughly similar.

Normality: The normal Q-Q line shows some skewness on the left tail.