

Exploring TA Grading Patterns

Andrea Malek Ash

12/6/2020

Load required packages.

```
if (!require("stats", "psych", "agricolae",  
"graphics"))install.packages("stats", "psych", "agricolae", "graphics", repos  
= 'http://cran.us.r-project.org')  
library("stats")  
library("psych")  
library("agricolae")  
library("graphics")
```

Load data. Save student scores as csv file (rows = Students, columns = Score, TA)

```
scores <- read.csv("/Users/andreaash/Student scores.csv", header=TRUE)  
scores$TA <- factor(scores$TA) #if your TA identifiers are numeric, this will  
make them factors
```

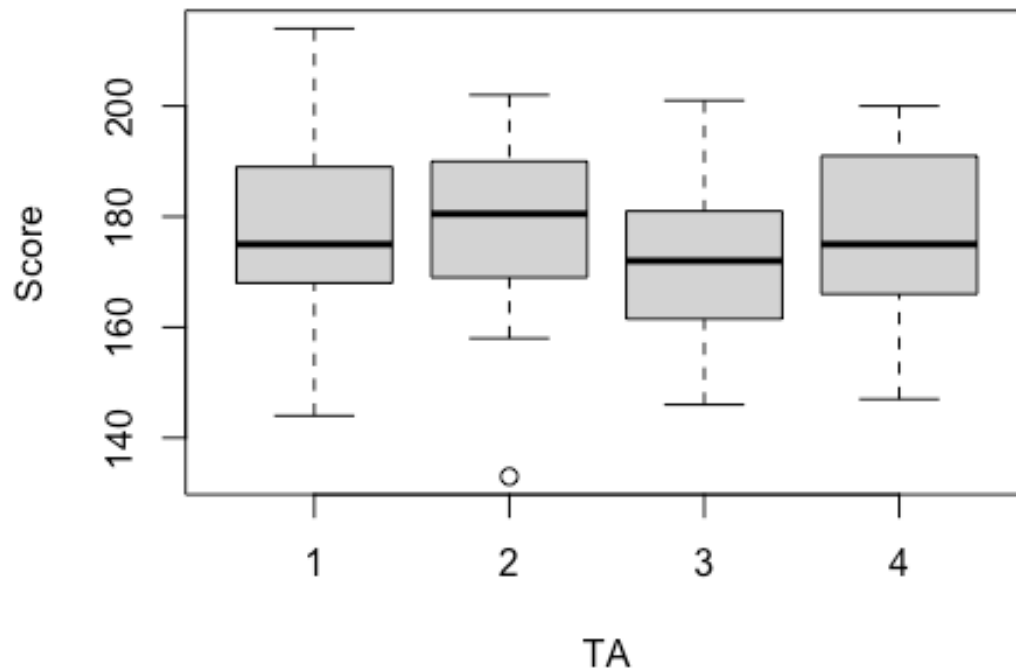
Get a summary of each groups's mean and variance

```
group.summary <- describeBy(x = scores$Score, group = scores$TA, mat=TRUE)  
#"mat=TRUE" creates a data frame  
print(group.summary)
```

##	item	group1	vars	n	mean	sd	median	trimmed	mad	min	max
##	X11	1	1	89	178.5169	14.89975	175.0	177.7534	13.3434	144	214
##	X12	2	2	22	178.9091	15.28120	180.5	180.3333	14.8260	133	202
##	X13	3	3	92	172.3913	12.57676	172.0	171.8919	14.8260	146	201
##	X14	4	4	29	175.3448	14.63673	175.0	175.7200	17.7912	147	200
##											
##			skew		kurtosis		se				
##	X11		0.4866955		-0.2791554		1.579370				
##	X12		-1.0423663		1.3695689		3.257963				
##	X13		0.2557085		-0.5222357		1.311218				
##	X14		-0.0889932		-1.1226152		2.717973				

Make a boxplot to visualize your data and get a sense of the difference between TA's

```
boxplot(Score ~ TA, data = scores)
```

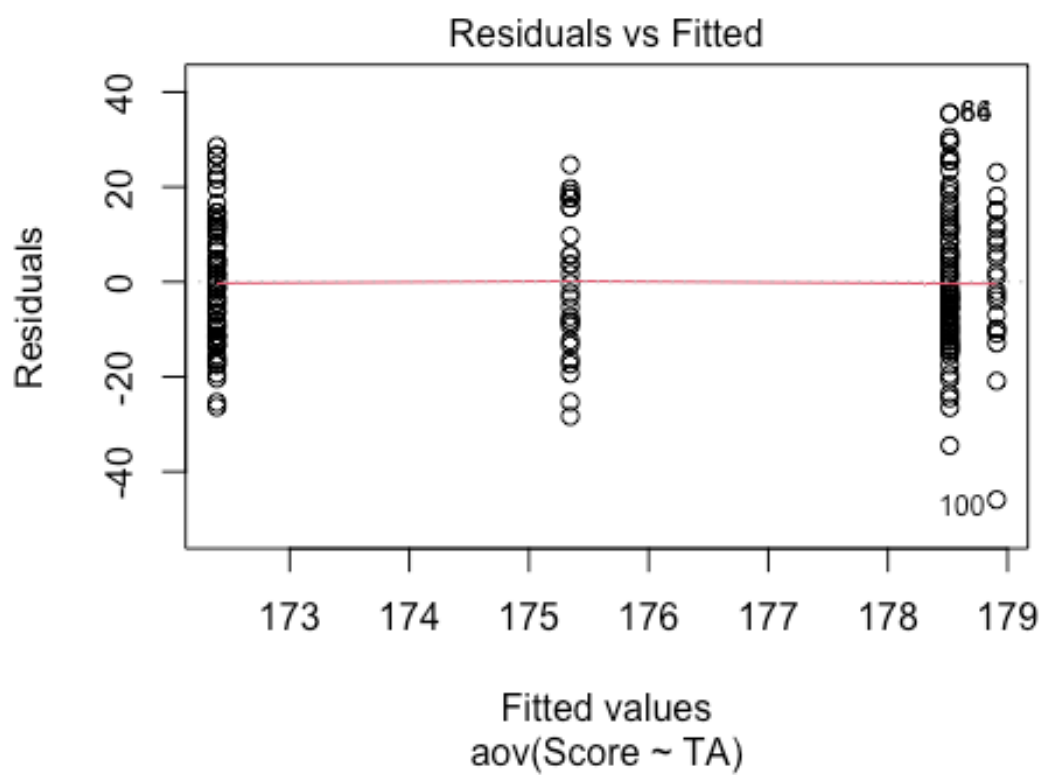


Use a simple ANOVA to compare groups of students graded by the same TA

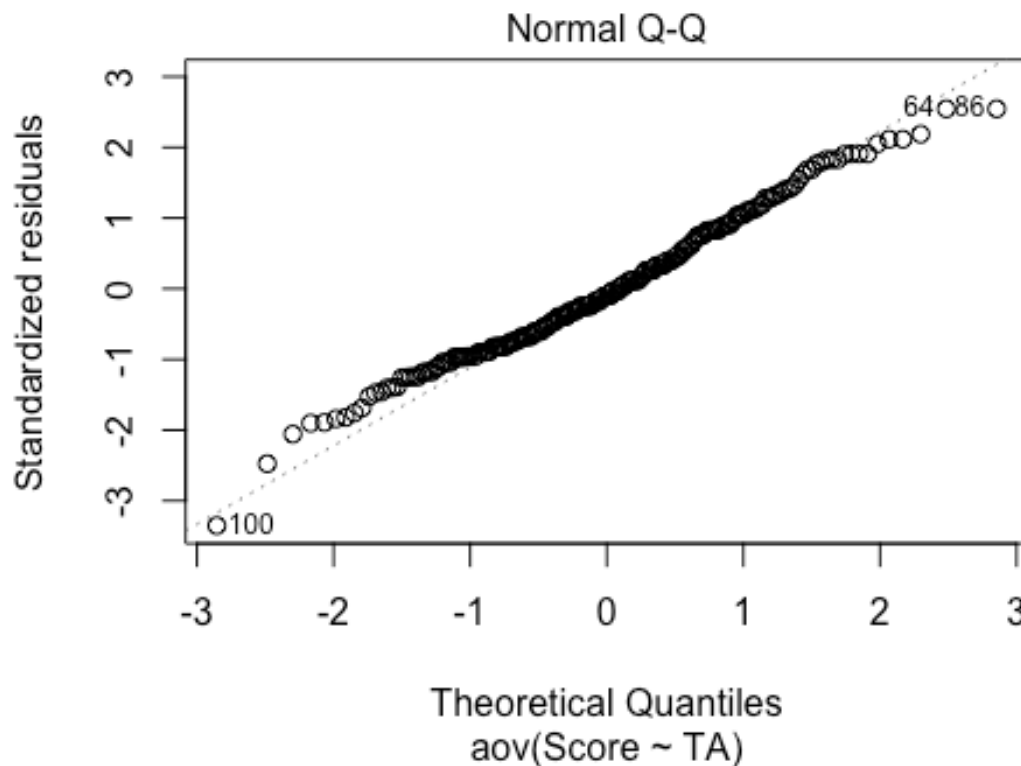
```
model.TA <- aov(Score ~ TA, data= scores)
```

Checking ANOVA assumptions and possible problem spots

```
plot(model.TA,1) #Homogeneity of variance - do some score groups vary more widely than others?
```



```
plot(model.TA,2) #Normality of score distribution - check whether the data are normally distributed.
```



ANOVA

is fairly robust to violations of homogeneity of variance. Most likely you can use Plot 1 for informational purposes only, to help you understand differences in variance between groups.

If scores are not normal (likely in course grading!)

```
skewness(scores$Score) #Check existing skewness of score distribution
## [1] 0.2403181

scores$Score.log10 <- log10(scores$Score) #Use a transformation of your
choosing (log10 is used here)
skewness(scores$Score.log10) #Check the skewness again - if improved and the
transformation makes sense, use this transformed value as your dependent
value in the ANOVA
## [1] 0.004507981
```

Repeat the ANOVA with new outcome variable, if you'd like.

```
model.TA.log10 <- aov(Score.log10 ~ TA, data= scores)
```

Remember that TA's are still confounded with student groups - you can't interpret a TA effect on its own at this point.

If significant p-value, you can continue with post-hoc analysis to get a more detailed picture of mean differences.

```
print(TukeyHSD(x = model.TA))

##    Tukey multiple comparisons of means
##      95% family-wise confidence level
##
## Fit: aov(formula = Score ~ TA, data = scores)
##
## $TA
##           diff           lwr           upr           p adj
## 2-1  0.392237  -8.248716   9.033190  0.9994188
## 3-1 -6.125550 -11.521362  -0.729737  0.0189571
## 4-1 -3.172026 -10.931879   4.587827  0.7154134
## 3-2 -6.517787 -15.130771   2.095198  0.2069429
## 4-2 -3.564263 -13.825064   6.696537  0.8053485
## 4-3  2.953523  -4.775173  10.682220  0.7558402
```

Remember, group differences could be due to TA's OR other factors like lab section, lab time, student schedules, etc. We don't know and can't interpret this - it's only for informational purposes to help understand patterns in the data.