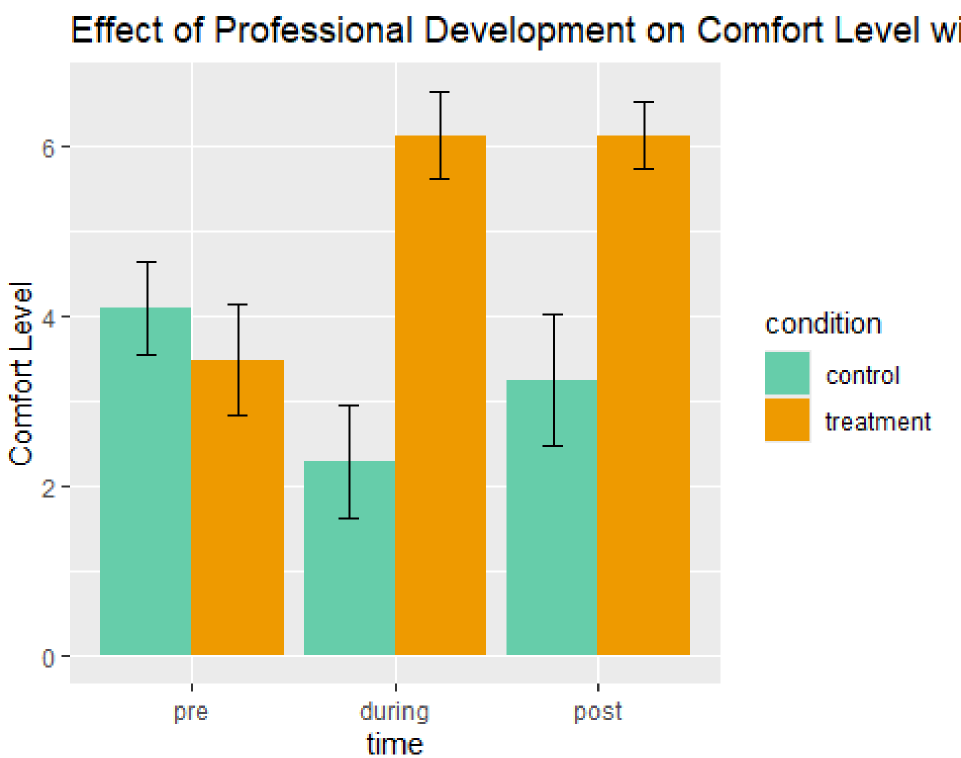
PSYCH 308B - DA6 - 2024

A local school district realized that none of their elementary teachers taught science (true story). In an attempt to remedy the situation, they developed a professional development course designed to teach their teachers science and how to teach science to their classes. This district sent their teachers to this course in cohorts. In order to determine if the PD course was working they had the first cohort of teachers complete a survey which asked them to rate, among other things, how comfortable they were teaching science to their elementary school class on a scale of 1 to 7. This survey was administered just before the PD course began, during the course, and one month after the course was completed. A CGU student begged them to get a comparison group and so, the same survey was administered to the second cohort of teachers at the same time as it was administered to the first cohort of teachers. This second cohort still has not received the PD course.

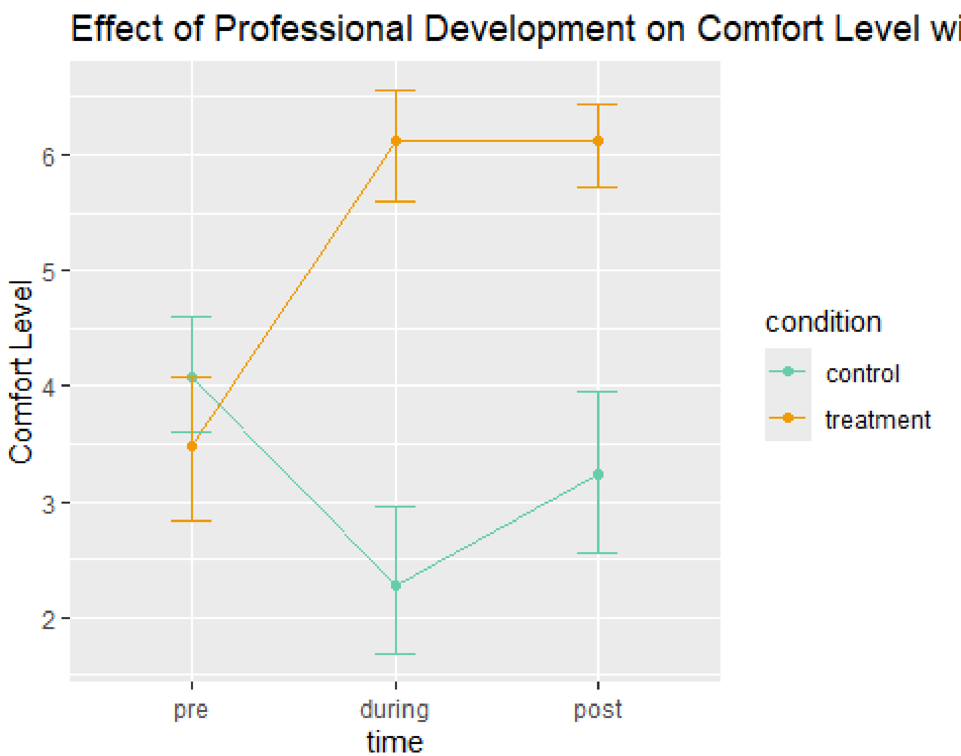
```r
#Pre-work
#Set up your libraries
library(psych)
library(ez)
library(car)
library(reshape)
library(jmv)
library(ggplot2)
library(heplots) # <- this library is needed for Box's test
library(tidyr) # <- this one is needed for pivot-wider code
library(pastecs)


#Pre-work
#Set up your data
science <- read.csv("308B.Data.DA6.csv")
science$Subject <- as.factor(science$Subject)
science$time <- factor(science$time, levels = c("pre", "during", "post"))

#convert data from long to wide so you have it ready for the ANOVA code
science.wide <- pivot_wider(science, names_from = time, values_from = value)
```

```r
#creating a bar graph
barscience <- ggplot(science, aes(time, value, fill = condition)) #order of the variables matters here!! If it's backwards, the bars are gonna go sideways

barscience + stat_summary(fun.y = mean, geom = "bar", position = "dodge") + stat_summary(fun.data = mean_cl_normal, geom = "errorbar", position = position_dodge(width = 0.90), width = 0.2) + labs(x = "time", y = "Comfort Level", fill = "condition") + ggtitle('Effect of Professional Development on Comfort Level with Teaching Over Time') + scale_fill_manual("condition", values = c("aquamarine3", "orange2"))
```




```r
#create a line graph
linescience <- ggplot(science, aes(time, value, colour = condition))
```

```
linescience + stat_summary(fun.y = mean, geom = "point") + stat_summary(fun.y = mean, geom = "line", aes(gr
oup = condition)) + stat_summary(fun.data = mean_cl_boot, geom = "errorbar", width = 0.2) + labs(x = "time"
, y = "Comfort Level", colour = "condition") + ggtitle('Effect of Professional Development on Comfort Level
with Teaching Over Time') + scale_color_manual("condition", values = c("aquamarine3", "orange2"))
```

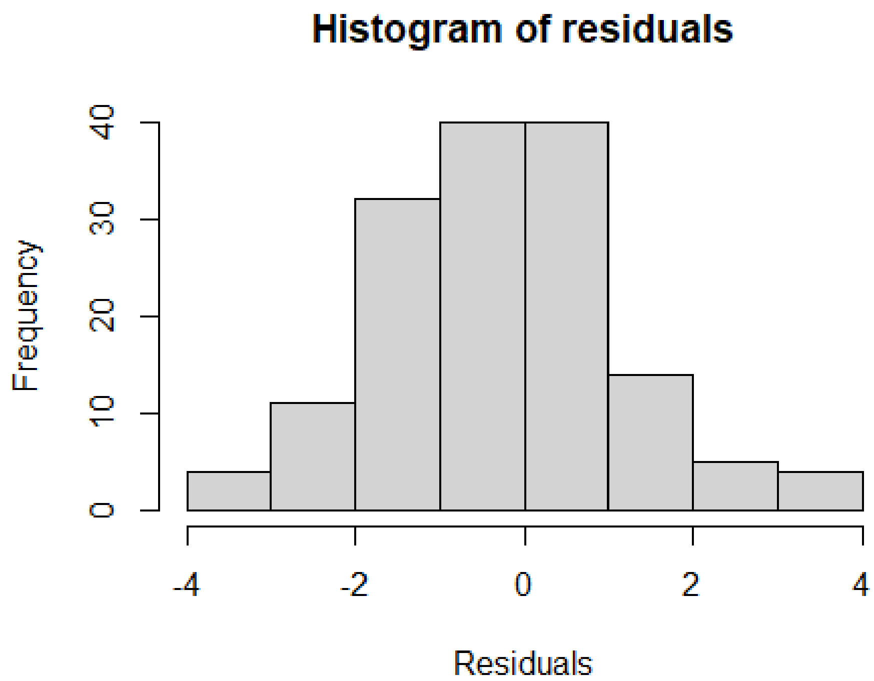



```
#This code specifies your model...
modeltc<-aov(value~as.factor(condition)*as.factor(time),data=science)

#....and this one returns all of the means (cell, marginal, and grand) for the model you specified
model.tables(modeltc, type="means")
```

```
Tables of means
Grand mean

4.22

 as.factor(condition)
as.factor(condition)
  control treatment
     3.20      5.24

 as.factor(time)
as.factor(time)
   pre during   post
  3.78   4.20   4.68

 as.factor(condition):as.factor(time)
                    as.factor(time)
as.factor(condition) pre  during post
           control   4.08 2.28   3.24
           treatment 3.48 6.12   6.12
```

Check Assumptions

```
#Histograms: You already specified your model above, so in the code below we're just plotting the residuals
(how much the model over or underestimates someones feedback score.You only need to check this one histogra
m!
restc<-modeltc$residuals
hist(restc, main="Histogram of residuals",xlab="Residuals")
```




```
#confirm skew and kurtosis
by(science$value, list(science$time, science$condition), basic = FALSE, norm = TRUE, stat.desc)

: pre
: control
      median         mean      SE.mean CI.mean.0.95          var      std.dev
   4.0000000    4.0800000    0.2640707    0.5450151    1.7433333    1.3203535
    coef.var     skewness     skew.2SE     kurtosis     kurt.2SE   normtest.W
   0.3236161    0.1728305    0.1863669   -0.1412354   -0.0783144    0.9230784
  normtest.p
   0.0602361
------------------------------------------------------------
```

```
: during
: control
       median          mean       SE.mean  CI.mean.0.95           var 
 2.0000000000  2.2800000000  0.3241398875  0.6689918476  2.6266666667 
      std.dev      coef.var      skewness      skew.2SE      kurtosis 
 1.6206994375  0.7108330866  1.0288620478  1.1094443150 -0.1675416321 
     kurt.2SE    normtest.W    normtest.p 
-0.0929010838  0.7852087638  0.0001301105 
------------------------------------------------------------ 
: post
: control
      median         mean      SE.mean CI.mean.0.95          var      std.dev 
  3.00000000   3.24000000   0.37094474   0.76559231   3.44000000   1.85472370 
    coef.var     skewness     skew.2SE     kurtosis     kurt.2SE   normtest.W 
  0.57244559   0.52794926   0.56929916  -0.85511909  -0.47415970   0.90891373 
  normtest.p 
  0.02882748 
------------------------------------------------------------ 
: pre
: treatment
      median         mean      SE.mean CI.mean.0.95          var      std.dev 
  4.00000000   3.48000000   0.31685959   0.65396605   2.51000000   1.58429795 
    coef.var     skewness     skew.2SE     kurtosis     kurt.2SE   normtest.W 
  0.45525803  -0.29684142  -0.32009056  -1.08220766  -0.60007930   0.90412488 
  normtest.p 
  0.02258554 
------------------------------------------------------------ 
: during
: treatment
       median          mean       SE.mean  CI.mean.0.95           var 
 7.000000e+00  6.120000e+00  2.471167e-01  5.100238e-01  1.526667e+00 
      std.dev      coef.var      skewness      skew.2SE      kurtosis 
 1.235584e+00  2.018927e-01 -1.362090e+00 -1.468772e+00  7.870201e-01 
     kurt.2SE    normtest.W    normtest.p 
 4.363991e-01  7.269201e-01  1.715429e-05 
------------------------------------------------------------ 
```

```
: post
: treatment
       median          mean       SE.mean  CI.mean.0.95           var
 6.0000000000  6.1200000000  0.1942506971  0.4009137344  0.9433333333
      std.dev      coef.var      skewness      skew.2SE      kurtosis
 0.9712534856  0.1587015499 -1.2745294451 -1.3743528106  1.8119708424
     kurt.2SE    normtest.W    normtest.p
 1.0047297081  0.7881593646  0.0001452044
```

```
#Box's Test
#This code run Box's test to check homogeneity of the co-variance matrices. Box's M follows the chi square distribution, so we report is as x2(df) = X.XX, p = .XXX. Note that it uses the wide data
BoxTestScience<-boxM(science.wide[,4:6],science.wide$condition) #use the wide format data for Box's test
BoxTestScience$cov

$control
              pre      during        post
pre     1.7433333 -0.10666667 -1.31166667
during -0.1066667  2.62666667 -0.02833333
post   -1.3116667 -0.02833333  3.44000000

$treatment
              pre      during        post
pre     2.5100000 -0.81000000 -0.39333333
during -0.8100000  1.52666667  0.06833333
post   -0.3933333  0.06833333  0.94333333

BoxTestScience


    Box's M-test for Homogeneity of Covariance Matrices

data:  science.wide[, 4:6]
Chi-Sq (approx.) = 14.491, df = 6, p-value = 0.02461
```

```
#Mixed Factorial ANOVA
#Questions 2-3: Main Effect for Time, Main Effect for Condition, and Interaction
model.rm <- anovaRM(data = science.wide,
                 rm = list(list(label = 'Time',
                                levels = c('pre', 'during', 'post'))),
                 rmCells = list(list(measure = 'pre', cell = 'pre'),
                                list(measure = 'during', cell = 'during'),
                                list(measure = 'post', cell = 'post')),
                 rmTerms = list('Time'),
                 bs = 'condition',
                 bsTerms = list('condition'),
                 effectSize = c('partEta'),
                 leveneTest = TRUE,
                 spherTests = TRUE,
                 spherCorr = c('none','GG'),
                 postHoc = list('Time', 'condition'),
                 postHocCorr = list('holm', 'tukey'),
                 emMeans = ~ Time + condition + Time:condition,
                 emmTables = T)
model.rm
```

REPEATED MEASURES ANOVA

Within Subjects Effects

	Sphericity Correction	Sum of Squares	df	Mean Square	F	p	η^2-p
Time	None	20.28000	2	10.140000	3.957931	0.0222993	0.0761757
	Greenhouse-Geisser	20.28000	1.921398	10.554813	3.957931	0.0238539	0.0761757
Time:condition	None	136.44000	2	68.220000	26.628212	< .0000001	0.3568116
	Greenhouse-Geisser	136.44000	1.921398	71.010783	26.628212	< .0000001	0.3568116
Residual	None	245.94667	96	2.561944			
	Greenhouse-Geisser	245.94667	92.227120	2.666750			

Note. Type 3 Sums of Squares

Between Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η^2-p
condition	156.06000	1	156.060000	122.7745	< .0000001	0.7189276
Residual	61.01333	48	1.271111			

Note. Type 3 Sums of Squares

ASSUMPTIONS

Tests of Sphericity

	Mauchly's W	p	Greenhouse-Geisser ε	Huynh-Feldt ε
Time	0.9590914	0.3747229	0.9606992	0.9999058

Homogeneity of Variances Test (Levene's)

	F	df1	df2	p
pre	1.674873	1	48	0.2017993
during	2.764349	1	48	0.1029033
post	10.746707	1	48	0.0019470

POST HOC TESTS

Post Hoc Comparisons - Time

Time		Time	Mean Difference	SE	df	t	p-tukey	p-holm
pre	-	during	-0.4200000	0.3200000	48.00000	-1.312500	0.3952720	0.2106991
	-	post	-0.9000000	0.3470831	48.00000	-2.593039	0.0330659	0.0377185
during	-	post	-0.4800000	0.2908035	48.00000	-1.650599	0.2347467	0.2106991

Post Hoc Comparisons - condition

condition		condition	Mean Difference	SE	df	t	p-tukey	p-holm
control	-	treatment	-2.040000	0.1841095	48.00000	-11.08036	< .0000001	< .0000001

ESTIMATED MARGINAL MEANS
TIME

Estimated Marginal Means - Time

Time	Mean	SE	Lower	Upper
pre	3.780000	0.2062361	3.365335	4.194665
during	4.200000	0.2037973	3.790238	4.609762
post	4.680000	0.2093641	4.259045	5.100955

CONDITION

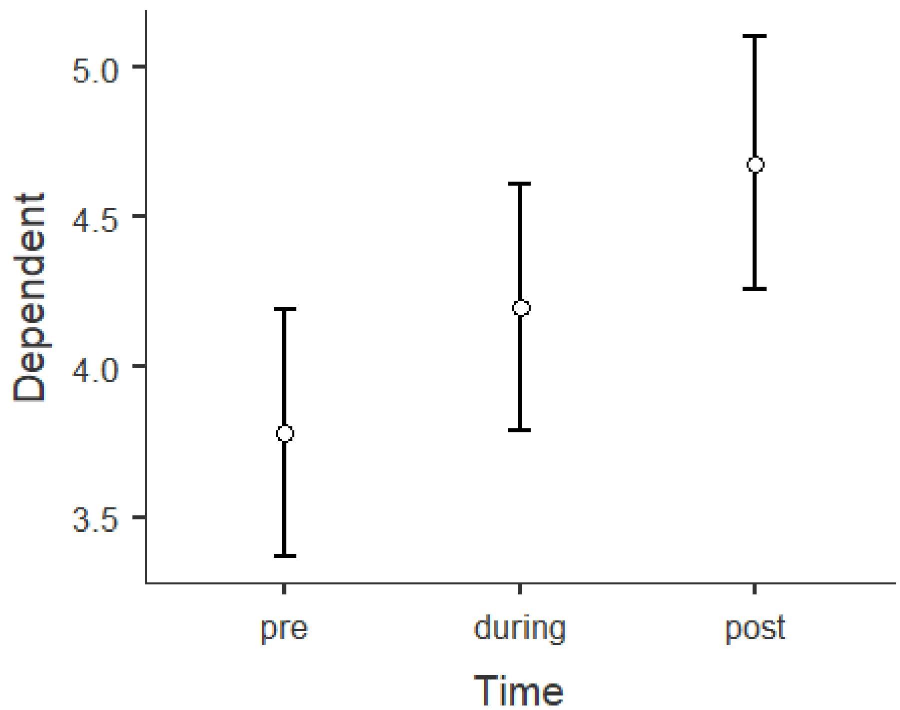
Estimated Marginal Means - condition

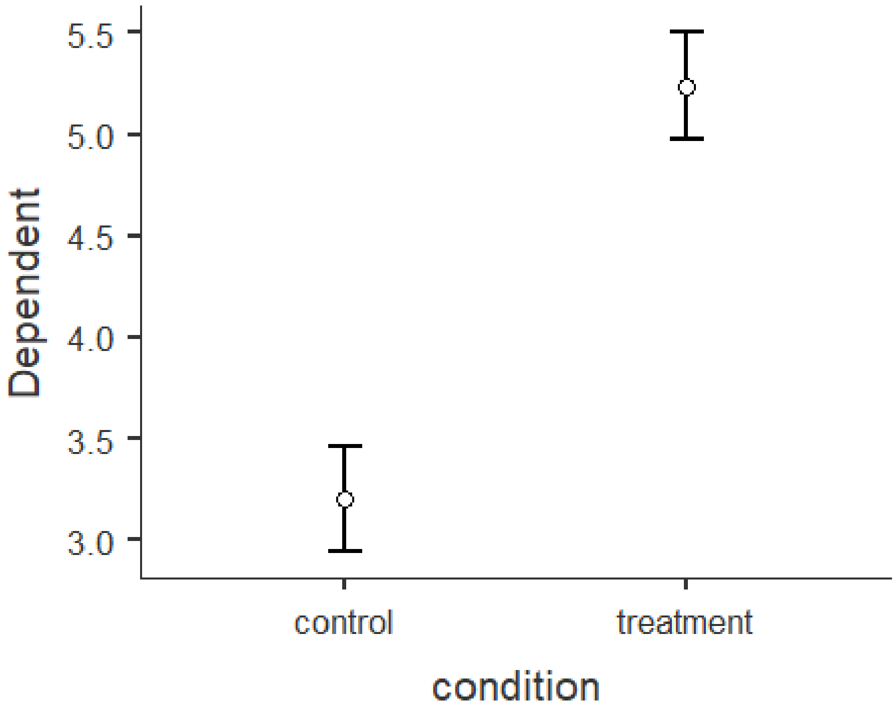
condition	Mean	SE	Lower	Upper
control	3.200000	0.1301851	2.938245	3.461755
treatment	5.240000	0.1301851	4.978245	5.501755

TIME:CONDITION

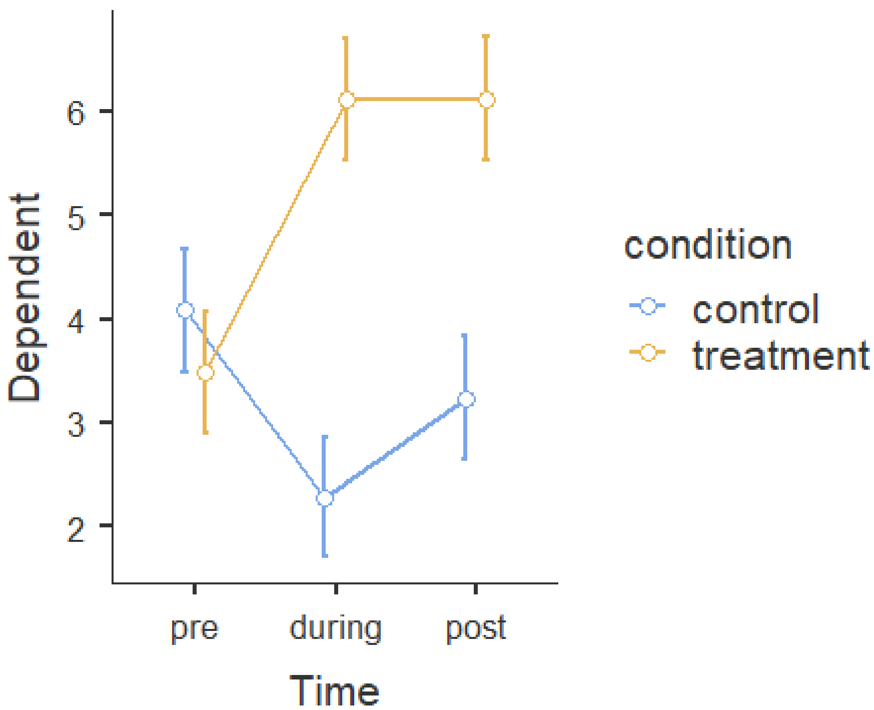
Estimated Marginal Means - Time:condition

condition	Time	Mean	SE	Lower	Upper
control	pre	4.080000	0.2916619	3.493574	4.666426
	during	2.280000	0.2882129	1.700509	2.859491
	post	3.240000	0.2960856	2.644680	3.835320
treatment	pre	3.480000	0.2916619	2.893574	4.066426
	during	6.120000	0.2882129	5.540509	6.699491
	post	6.120000	0.2960856	5.524680	6.715320


5.0
4.5
4.0
3.5
Dependent
pre
during
post
Time



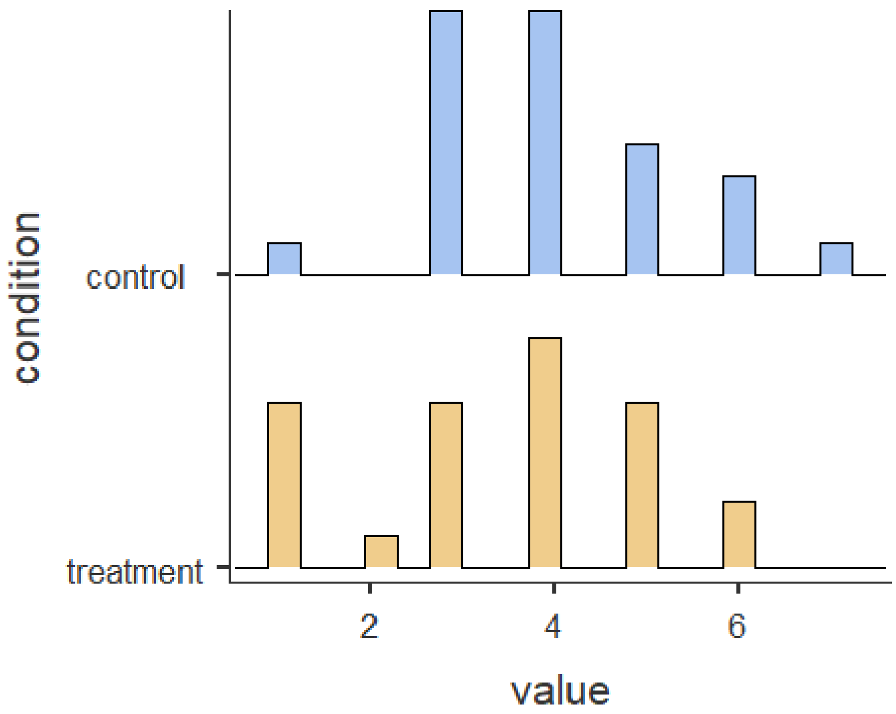
5.5
5.0
4.5
4.0
3.5
3.0
Dependent
control
treatment
condition




```
#Simple Effects Option 1
#Question 4: Assumptions for simple effect analyses
#Normal distribution, but to do this, you can first create a subset for each time point
Pre <- subset(science, science$time == "pre")
During <- subset(science, science$time == "during")
Post <- subset(science, science$time == "post")

# Then run descriptives of each time point by condition
descriptives(Pre, vars = c('value'), splitBy = c('condition'), skew = TRUE, kurt = TRUE, hist = TRUE)


 DESCRIPTIVES
```

Descriptives

	condition	value
N	control	25
	treatment	25
Missing	control	0
	treatment	0
Mean	control	4.080000
	treatment	3.480000
Median	control	4
	treatment	4
Standard deviation	control	1.320353
	treatment	1.584298
Minimum	control	1
	treatment	1
Maximum	control	7
	treatment	6
Skewness	control	0.1956867
	treatment	-0.3360976
Std. error skewness	control	0.4636835
	treatment	0.4636835
Kurtosis	control	0.4103199
	treatment	-0.8488039
Std. error kurtosis	control	0.9017205
	treatment	0.9017205

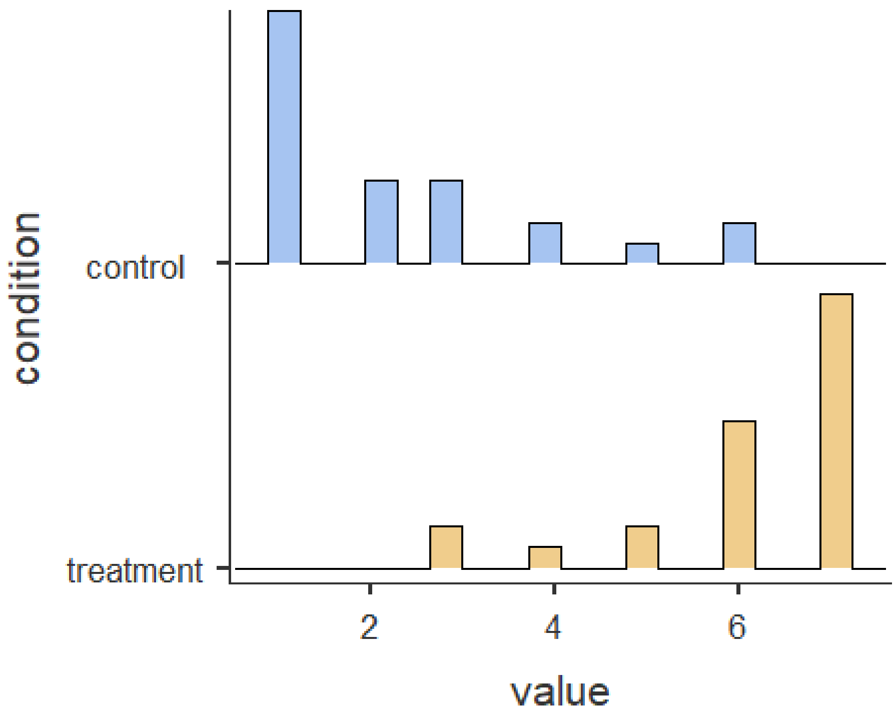



```
descriptives(During, vars = c('value'), splitBy = c('condition'), skew = TRUE, kurt = TRUE, hist = TRUE)
```

DESCRIPTIVES

Descriptives

	condition	value
N	control	25
	treatment	25
Missing	control	0
	treatment	0
Mean	control	2.280000
	treatment	6.120000
Median	control	2
	treatment	7
Standard deviation	control	1.620699
	treatment	1.235584
Minimum	control	1
	treatment	3
Maximum	control	6
	treatment	7
Skewness	control	1.164925
	treatment	-1.542222
Std. error skewness	control	0.4636835
	treatment	0.4636835
Kurtosis	control	0.3751193
	treatment	1.652427
Std. error kurtosis	control	0.9017205
	treatment	0.9017205

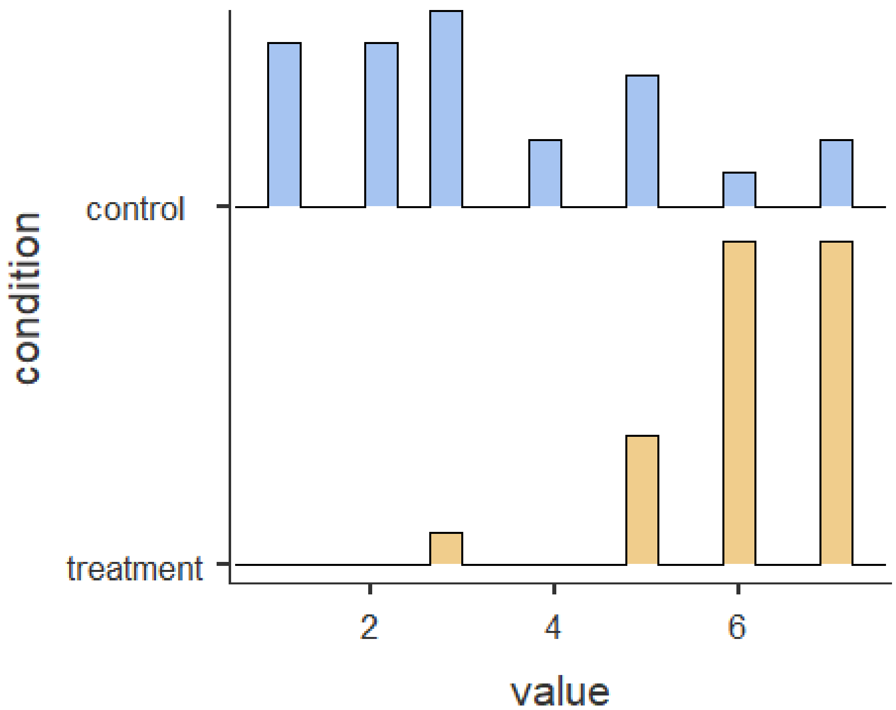



```
descriptives(Post, vars = c('value'), splitBy = c('condition'), skew = TRUE, kurt = TRUE, hist = TRUE)
```

DESCRIPTIVES

Descriptives

	condition	value
N	control	25
	treatment	25
Missing	control	0
	treatment	0
Mean	control	3.240000
	treatment	6.120000
Median	control	3
	treatment	6
Standard deviation	control	1.854724
	treatment	0.9712535
Minimum	control	1
	treatment	3
Maximum	control	7
	treatment	7
Skewness	control	0.5977686
	treatment	-1.443081
Std. error skewness	control	0.4636835
	treatment	0.4636835
Kurtosis	control	-0.5449346
	treatment	3.023923
Std. error kurtosis	control	0.9017205
	treatment	0.9017205




```
#Simple Effects Option 1
# Question 5: Simple Effects Analyses
# Note: notice that a Welch's correction was added to the code for the third t-test since levene's was sign
ificant for those two groups (i.e., treatment and control for the "Post" time point)
ttestIS(data=Pre, vars = 'value', group = 'condition', eqv = T, effectSize = T, desc = TRUE)
```

INDEPENDENT SAMPLES T-TEST

Independent Samples T-Test

		Statistic	df	p		Effect Size
value	Student's t	1.454643	48.00000	0.1522766	Cohen's d	0.4114353

Note. H_a μ _{control} ≠ μ _{treatment}

ASSUMPTIONS

Homogeneity of Variances Test (Levene's)

	F	df	df2	p
value	1.674873	1	48	0.2017993

Note. A low p-value suggests a violation of the assumption of equal variances

Group Descriptives

	Group	N	Mean	Median	SD	SE
value	control	25	4.080000	4.000000	1.320353	0.2640707
	treatment	25	3.480000	4.000000	1.584298	0.3168596

```
ttestIS(data=During, vars = 'value', group = 'condition', eqv = T, effectSize = T, desc = TRUE)
```

INDEPENDENT SAMPLES T-TEST

Independent Samples T-Test

		Statistic	df	p		Effect Size
value	Student's t	-9.421126	48.00000	< .0000001	Cohen's d	-2.664697

Note. H_a μ _{control} ≠ μ _{treatment}

ASSUMPTIONS

Homogeneity of Variances Test (Levene's)

	F	df	df2	p
value	2.764349	1	48	0.1029033

Note. A low p-value suggests a violation of the assumption of equal variances

Group Descriptives

	Group	N	Mean	Median	SD	SE
value	control	25	2.280000	2.000000	1.620699	0.3241399
	treatment	25	6.120000	7.000000	1.235584	0.2471167

```
ttestIS(data=Post, vars = 'value', group = 'condition', welchs = T, eqv = T, effectSize = T, desc = TRUE)
```

```
INDEPENDENT SAMPLES T-TEST

 Independent Samples T-Test
 ───────────────────────────────────────────────────────────────────────────────────────────
                            Statistic    df          p                            Effect Size
 ───────────────────────────────────────────────────────────────────────────────────────────
   value    Student's t    -6.877969    48.00000    < .0000001    Cohen's d        -1.945384
            Welch's t      -6.877969    36.24219    < .0000001    Cohen's d        -1.945384
 ───────────────────────────────────────────────────────────────────────────────────────────
   Note. Hₐ μ <sub>control</sub> ≠ μ <sub>treatment</sub>


 ASSUMPTIONS

 Homogeneity of Variances Test (Levene's)
 ───────────────────────────────────────────────
            F           df    df2    p
 ───────────────────────────────────────────────
   value    10.74671     1     48    0.0019470
 ───────────────────────────────────────────────
   Note. A low p-value suggests a violation
   of the assumption of equal variances


 Group Descriptives
 ──────────────────────────────────────────────────────────────────────────────
            Group        N     Mean        Median      SD           SE
 ──────────────────────────────────────────────────────────────────────────────
   value    control      25    3.240000    3.000000     1.854724    0.3709447
            treatment    25    6.120000    6.000000    0.9712535    0.1942507
 ──────────────────────────────────────────────────────────────────────────────


#Simple Effects Option 2
#Question 4: Assumptions for simple effect analyses (subset by condition)
```

```
#Normal distribution, but to do this, you can first create a subset for each time point
Treatment <- subset(science.wide, science.wide$condition == "treatment")
Control <- subset(science.wide, science.wide$condition == "control")

# Then run descriptives of each time point by condition
describe.by(Treatment)

Warning: describe.by is deprecated.  Please use the describeBy function

Warning in describeBy(x = x, group = group, mat = mat, type = type, ...): no
grouping variable requested

           vars  n  mean   sd median trimmed  mad min max range  skew kurtosis
Subject*      2 25 13.00 7.36     13   13.00 8.90   1  25    24  0.00    -1.34
condition*    3 25  1.00 0.00      1    1.00 0.00   1   1     0   NaN      NaN
pre           4 25  3.48 1.58      4    3.48 1.48   1   6     5 -0.30    -1.08
during        5 25  6.12 1.24      7    6.33 0.00   3   7     4 -1.36     0.79
post          6 25  6.12 0.97      6    6.24 1.48   3   7     4 -1.27     1.81
             se
Subject*   1.47
condition* 0.00
pre        0.32
during     0.25
post       0.19

describe.by(Control)

Warning: describe.by is deprecated.  Please use the describeBy function

Warning: no grouping variable requested

           vars  n  mean   sd median trimmed  mad min max range skew kurtosis
Subject*      2 25 38.00 7.36     38   38.00 8.90  26  50    24 0.00    -1.34
condition*    3 25  1.00 0.00      1    1.00 0.00   1   1     0  NaN      NaN
pre           4 25  4.08 1.32      4    4.05 1.48   1   7     6 0.17    -0.14
during        5 25  2.28 1.62      2    2.05 1.48   1   6     5 1.03    -0.17
post          6 25  3.24 1.85      3    3.10 1.48   1   7     6 0.53    -0.86
```

```
              se
Subject*   1.47
condition* 0.00
pre        0.26
during     0.32
post       0.37

#Simple Effects Option 2
# Question 5: Simple Effects Analyses - Subset by condition, Treatment

modeltreat <- anovaRM(data = Treatment,
                 rm = list(list(label = 'Time',
                                levels = c('pre', 'during', 'post'))),
                 rmCells = list(list(measure = 'pre', cell = 'pre'),
                                list(measure = 'during', cell = 'during'),
                                list(measure = 'post', cell = 'post')),
                 rmTerms = list('Time'),
                 effectSize = c('partEta', 'eta'),
                 spherTests = TRUE,
                 spherCorr = c('none','GG'),
                 postHoc = list('Time'),
                 postHocCorr = 'holm',
                 emMeans = ~ Time,
                 emmTables = T)
modeltreat
```

REPEATED MEASURES ANOVA

Within Subjects Effects

	Sphericity Correction	Sum of Squares	df	Mean Square	F	p	η²	η²-p
Time	None	116.16000	2	58.080000	28.49387	< .0000001	0.4928717	0.5428037
	Greenhouse-Geisser	116.16000	1.635153	71.039209	28.49387	0.0000001	0.4928717	0.5428037
Residual	None	97.84000	48	2.038333				
	Greenhouse-Geisser	97.84000	39.243680	2.493140				

Note. Type 3 Sums of Squares

Between Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η^2	η^2-p
Residual	21.68000	24	0.9033333				

Note. Type 3 Sums of Squares

ASSUMPTIONS

Tests of Sphericity

	Mauchly's W	p	Greenhouse-Geisser ε	Huynh-Feldt ε
Time	0.7768731	0.0548310	0.8175767	0.8691952

POST HOC TESTS

Post Hoc Comparisons - Time

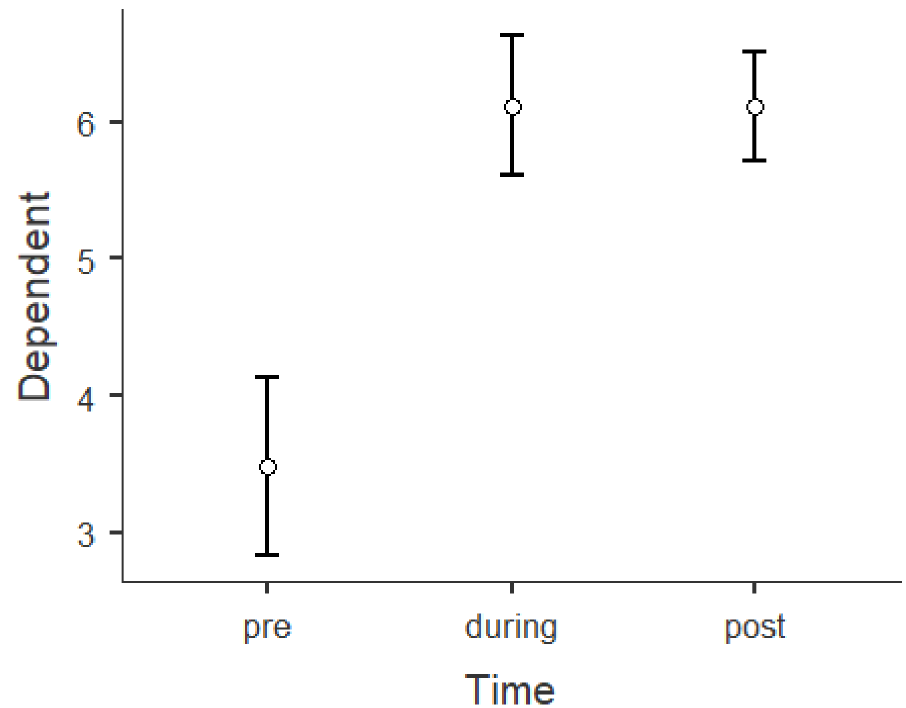
Time		Time	Mean Difference	SE	df	t	p-holm
pre	-	during	-2.640000	0.4756750	24.00000	-5.550008	0.0000208
	-	post	-2.640000	0.4118252	24.00000	-6.410487	0.0000038
during	-	post	0.000000	0.3055050	24.00000	0.000000	1.0000000

ESTIMATED MARGINAL MEANS

TIME

Estimated Marginal Means - Time

Time	Mean	SE	Lower	Upper
pre	3.480000	0.3168596	2.826034	4.133966
during	6.120000	0.2471167	5.609976	6.630024
post	6.120000	0.1942507	5.719086	6.520914



```
#Simple Effects Option 2
# Question 5: Simple Effects Analyses - Subset by condition, Control

modelcont <- anovaRM(data = Control,
                 rm = list(list(label = 'Time',
                                levels = c('pre', 'during', 'post'))),
                 rmCells = list(list(measure = 'pre', cell = 'pre'),
                                list(measure = 'during', cell = 'during'),
                                list(measure = 'post', cell = 'post')),
                 rmTerms = list('Time'),
                 effectSize = c('partEta', 'eta'),
                 spherTests = TRUE,
                 spherCorr = c('none','GG'),
                 postHoc = list('Time'),
                 postHocCorr = 'holm',
                 emMeans = ~ Time,
                 emmTables = T)
modelcont
```

```
REPEATED MEASURES ANOVA

Within Subjects Effects
 ─────────────────────────────────────────────────────────────────────────────────────────────────────────────────────────
                 Sphericity Correction    Sum of Squares    df           Mean Square    F           p           η²          η²-p
 ─────────────────────────────────────────────────────────────────────────────────────────────────────────────────────────
   Time          None                           40.56000            2      20.280000    6.572560    0.0029999   0.1778947   0.2149823
                 Greenhouse-Geisser             40.56000     1.833178      22.125511    6.572560    0.0039992   0.1778947   0.2149823

   Residual      None                          148.10667           48       3.085556
                 Greenhouse-Geisser            148.10667    43.996273       3.366346
 ─────────────────────────────────────────────────────────────────────────────────────────────────────────────────────────
   Note. Type 3 Sums of Squares
```

Between Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η²	η²-p
Residual	39.33333	24	1.638889				

Note. Type 3 Sums of Squares

ASSUMPTIONS

Tests of Sphericity

	Mauchly's W	p	Greenhouse-Geisser ε	Huynh-Feldt ε
Time	0.9089985	0.3337923	0.9165890	0.9886273

POST HOC TESTS

Post Hoc Comparisons - Time

Time		Time	Mean Difference	SE	df	t	p-holm
pre	-	during	1.8000000	0.4281744	24.00000	4.203894	0.0009431
	-	post	0.8400000	0.5588083	24.00000	1.503199	0.1458335
during	-	post	-0.9600000	0.4949074	24.00000	-1.939757	0.1284987

ESTIMATED MARGINAL MEANS

TIME

Estimated Marginal Means - Time

Time	Mean	SE	Lower	Upper
pre	4.080000	0.2640707	3.534985	4.625015
during	2.280000	0.3241399	1.611008	2.948992
post	3.240000	0.3709447	2.474408	4.005592