

Analysis of Variance Tutorials

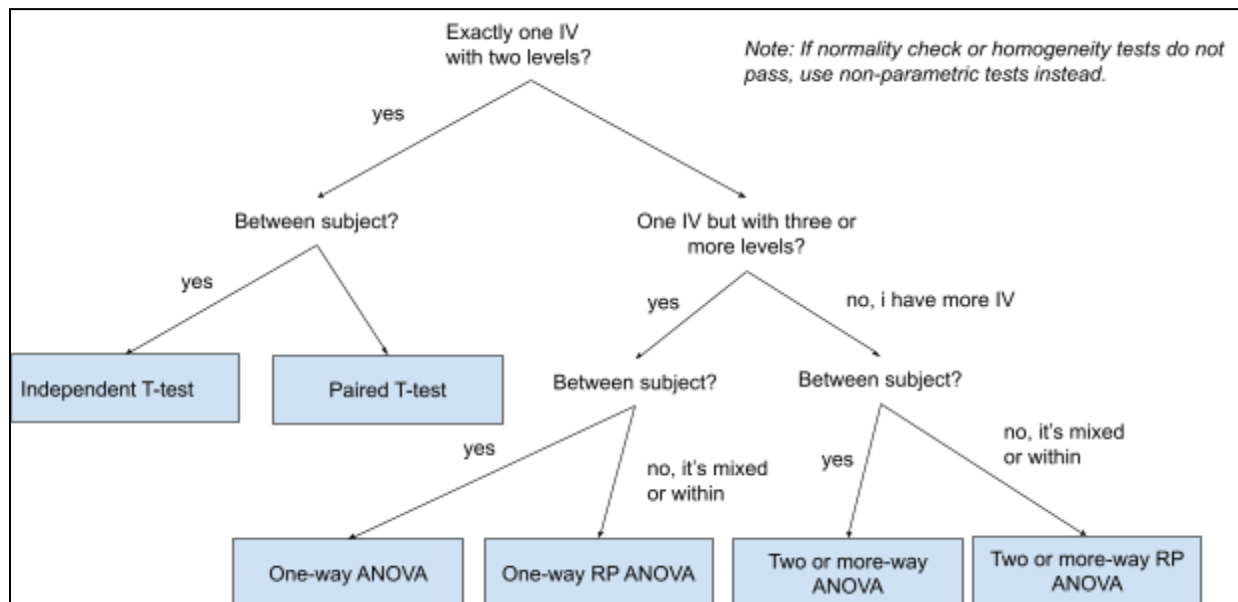
Task 1: T-Test	2
Task 2: One way ANOVA with two levels	3
Task 3: One-way ANOVA with 3 or more levels	4
Task 4: Two-way ANOVA	6
Task 5: Repeated Measures ANOVA	7
Task 6: Mixed Design ANOVA	9

This tutorial can be found in the `Tutorial` folder. It shall prepare you to be successful in performing appropriate analysis in an empirical controlled experiment.

All data files can be found in Course `GDrive > Downloads > ps4hci-Wobbrock`.

You can use R (free), or JASP (free) to complete this.

Before that, please understand this graph.



Task 1: T-Test

File used: posts.sav

The file contains a hypothetical study of 40 college students' Facebook posting behavior using one of two platforms: Apple's iOS or Google's Android OS. The data show the number of Facebook posts subjects made during a particular week using their mobile platform.

Here, this is a one-way design with two levels and thus T-test is appropriate

JASP: Go to T-Tests > Independent Samples T-Test

1. Put Posts as Variables, and Platform as Grouping Variable.
2. Tick Assumption Checks > Normality as well as Equality of Variance
3. If the data is normal ($p > 0.05$) and pass the Equality of Variance ($p > 0.05$) then
 - a. Select Tests as Welch (parametric)
 - b. Otherwise, use Wilcoxon (non-parametric)
4. Tick Effect Size > Cohen's d
5. Tick Descriptive

R: After you load the datasets, do

```
t.test(posts$Posts~posts$Platform, var.equal = TRUE)
```

For normality check:

To do a Shapiro-Wilk test, do

```
shapiro.test(posts$Posts)
```

To do a Kolmogorov-Smirnov test, do

```
ks.test(posts$Posts, "pnorm", mean=mean(posts$Posts),  
sd=sd(posts$Posts))
```

For homogeneity of variances, do

```
library(car) //you may need to first install package  
leveneTest(posts$Posts, posts$Platform, center=mean)
```

For Mann-Whitney,

```
wilcox.test(posts$Posts ~ posts$Platform)
```

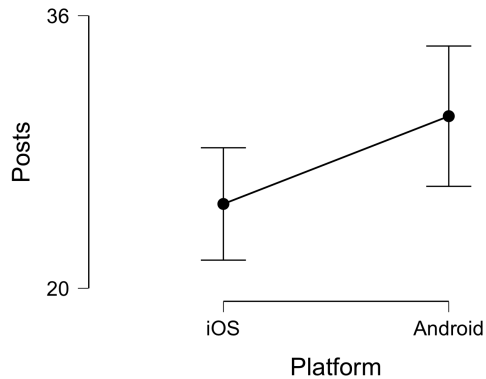
To show descriptives, do

```
summary(posts$Posts)  
mean(posts$Posts)  
sd(posts$Posts)  
sd(posts$Posts) / sqrt(length(posts$Posts))
```

Answer the followings:

1. Is this a between-subjects or within-subjects experiment? Why?
a. Between subject
2. What is the independent variable named?
a. Platform
3. How many levels does the independent variable have? What are they?
a. 2
4. How many subjects were in this experiment?
a. 40
5. How many subjects were exposed to each level of the independent variable? Is the design balanced (i.e., are the numbers equal)?
a. 20
6. What are the mean and standard deviation number of posts for each level of the independent variable?
 - a. iOS - $M = 24.950$, $SD = 7.045$
 - b. Android - $M = 30.100$, $SD = 8.795$
7. Assuming equal variances, what is the t statistic for this t-test? (Hint: this is also called the t Ratio.)
a. -2.044
8. How many degrees of freedom are there (dfs)?
a. 36.271
9. What is the two-tailed p-value resulting from this t-test? Is it significant at the $\alpha = .05$ level?
a. Yes, $p < .05$
10. The formulation for expressing a significant t-test result is: $t(df) = \text{test statistic}$, $p < p\text{-value thresholds}$, $d = \text{Cohen's } d$. For an insignificant result, it is: $t(df) = \text{t-statistic}$, n.s. Write your result just like you would do in a research paper. Read <https://shengdongzhao.com/newSite/how-to-report-statistics-in-apa-format/> for more detail how to report.

We found a significant effect of platform ($t(36.271) = -2.044, p < .05, d = -0.646$) ($t_{36.271} = -2.044, p < .05, d = -0.646$) on the number of facebook posts between iOS ($M = 24.950, SD = 7.045$) and Android ($M = 30.100, SD = 8.795$) (see Figure X)



11. The equivalent of a between-subjects (independent samples) t-test using nonparametric statistics is the Mann-Whitney U test. The Mann-Whitney U test is a test for an experiment containing one between-subjects factor with two-levels. The formulation for expressing a Mann-Whitney U-test is $U = \text{test statistic}$, $p < \text{p-value thresholds}$, $r = \text{rank biserial correlation}$. Note that when you use a non-parametric test, report the median instead of the mean. Write your result just like you would do in a research paper.

We fail to find a significant effect of platform ($U = 135.500$, *n.s.*) on the number of facebook posts between iOS ($Md = 24.500$) and Android ($Md = 27.500$) (see Figure X)

Task 2: One way ANOVA with two levels

File used: posts.sav

F-test (another name for analysis of variances or ANOVA), which can do everything a t-test can do, and more. An F-test, which is the most common analysis of variance, can handle multiple independent variables, or factors, and these factors can have more than two levels. By comparison, a t-test can only have one factor with two levels, which is not very useful for many experiment designs.

JASP: Select ANOVA.

1. Posts as DV, and Platform as IV.
2. Model -> Components -> Hold **Shift** and select everything and click the right arrow
3. Assumption Checks >
 - a. check Homogeneity tests and Q-Q plot of residuals.
 - b. Select Descriptives on the top bar (*not inside ANOVA, the top bar*), Put Posts on Variable, under Statistics > check Shapiro Wilk test

4. If data is normal (Shapiro Wilk p-value > 0.05) and is homogenous (Levene's test > 0.05)
 - a. Use the ANOVA results
 - b. Otherwise, inside the ANOVA, go under Nonparametrics, select Platform and click the right arrow button.
5. Descriptive Plots > check Display error bars.
6. Tick Descriptive Statistics and Estimates of effect size (check partial eta squared)

R: Do:

For ANOVA:

```
aov <- aov(posts$Posts ~ posts$Platform)
summary(aov)
```

For effect size:

```
install.packages("lsr")
library(lsr)
etaSquared(aov)
```

For Q-Q plot:

```
qqnorm(posts$Posts)
qqline(posts$Posts)
```

For other R scripts, check out previous homework.

Answer the followings:

1. What is the output of the Q-Q plot test? Is the data normal? How about homogeneity of variance tests?
 - a. Points align with the Q-Q plot, so it looks normal. A Shapiro-Wilk test confirms the normality ($p > .05$). For homogeneity of variances, a Levene test confirms the assumption check ($p > .05$)
2. Do the number of observations (N's) and means agree with those produced by the t-test? What are they? (If they do not agree, there is an error somewhere!)
 - a. Yes
3. In the ANOVA table, what is the F-statistic? What is the p-value? Is it significant at the $\alpha = .05$ level?
 - a. $F = 4.177$, $p = 0.048$, yes
4. How does this p-value compare to that produced by the t-test?
 - a. Same
5. What is the effect size in terms of partial eta squared? What is the interpretation?
 - a. $\eta_p^2 = 0.099$, since small = 0.01, medium = 0.06, large = 0.14, the effect is toward large.

6. The general formulation for expressing an F-test result is: $F(df\text{-between}, df\text{-within}) = F\text{-ratio}, p < .05$ (or n.s.), $\eta_p^2 =$ partial eta squared value. Report just like you would in a research paper. Read this for more detail - <https://shengdongzhao.com/newSite/how-to-report-statistics-in-apa-format/>

We found a main effect of Platform ($F(1, 38) = 4.177, p < .05, \eta_p^2 = 0.099$) between iOS ($M = 24.950, SD = 7.045$) and Android ($M = 30.100, SD = 8.795$) (see Figure X).

7. In case your data cannot use ANOVA, you can use Kruskal-Wallis test, where the formulation is $H(df) = \text{statistics}, p < \text{threshold}$. Report just like you would in a research paper.

We fail to find a significant effect of Platform ($H(1) = 3.053, n.s.$) on Posts between iOS ($Md = 24.500$) and Android ($Md = 27.500$) (see Figure X).

Task 3: One-way ANOVA with 3 or more levels

Review about post hoc analysis:

- Step 1: Find whether a significant difference on some factor (IV) between some levels using ANOVA
 - Step 1.1a: If no, do nothing
 - Step 1.1b: If yes, and if you have more than 2 levels, you have to do post hoc, why? → because you need to know exactly which level is really different between one another. → Tukey or **Bonferroni (highly recommended)**

File used: postctrl.sav

As noted, the F-test can handle more than one factor, and also, more than two levels per factor. A one-way ANOVA refers to a single factor design. Similarly, a two-way ANOVA refers to a two-factor design, i.e., two independent variables. In this part, we will still conduct a one-way ANOVA, but this time, our factor will have three levels. Thus, it cannot be analyzed with a t-test, which can only handle two levels of a single factor.

Open postctrl.sav. This data set is the same for the iOS and Android levels, but now has added 20 new college students as a control group who did not use a mobile device for posting on Facebook but were told to use their desktop computer instead. Thus, the Platform factor now has three levels: iOS, Android, and desktop.

JASP: Do everything as Task 2. In addition, since Platform has three levels, if the ANOVA has significance in Platform (i.e., $p < 0.05$), then inside ANOVA, under Post Hoc tests, choose Platform, and tick Tukey and Bonferroni.

R: Do:

Set the levels (unfortunately, R does not recognize the levels of Platforms):

```
postsctrl$Platform <- ordered(postsctrl$Platform, levels = c("1", "2", "3"))
```

To check the levels are set correctly:

```
levels(postsctrl$Platform)
```

Perform anova:

```
aov <- aov(Posts ~ Platform, data = postsctrl)
summary(aov)
```

Perform Post Hoc:

```
TukeyHSD(aov) //Tukey test
pairwise.t.test(postsctrl$Posts, postsctrl$Platform, p.adj="bonferroni")
//pairwise t test with Bonferroni correction
```

Answer the followings:

1. What is the output of the normality test? Is the data normal?
 - a. Q-Q plot looks aligned. Shapiro wilk test confirms the normality ($p > .05$).
2. Was this a one-way, two-way, or three-way analysis of variance? What is/are the factor(s)? What are each factor's levels?
 - a. One way because of only one IV
 - b. Platform - 3
3. How many data points are there for each level of Platform? What are the means and standard deviations for each level?
 - a. 20
 - b. iOS ($M = 24.950$, $SD = 7.045$)
 - c. Android ($M = 30.100$, $SD = 8.795$)
 - d. Desktop ($M = 31.700$, $SD = 7.623$)
4. In this case, the overall, or omnibus, F-test is testing for whether any differences exist among the levels of the independent variable. Is the F-test significant at the $\alpha = .05$ level? What is the F-ratio? What is the p-value?
 - a. $F = 4.033$, $p = 0.023$
5. The omnibus F-test does not tell us whether all three levels of Platform are different from one another, or whether just two levels (and which two?) are different. For this, we need post hoc comparisons, which are justified only when the omnibus F-test is significant. Examine the Post Hoc Tests output. What is the interpretation? Does Tukey and Bonferroni converge to the same interpretation?
 - a. iOS - Android (*n.s.*)

- b. iOS - Desktop ($p < .05$)
 - c. Android - Desktop (*n.s.*)
6. What is the effect size in terms of partial eta squared? What is the interpretation?
- a. $\eta_p^2 = 0.124$, since small = 0.01, medium = 0.06, large = 0.14, the effect is toward large.
7. Write a sentence summarizing the findings from this analysis.
- A Shapiro-Wilk test and a Levene test confirm the normality and homogeneity of variances test ($p > .05$). A one-way ANOVA test found a main effect of Platform ($F(2, 57) = 4.033$, $p < .05$, $\eta_p^2 = 0.124$) on Posts. A post hoc analysis with bonferroni corrections confirms the difference between iOS and Desktop ($p < .05$), but not between iOS-Android and Android-Desktop.

In case of significance of more than one pair:

A post hoc analysis with bonferroni corrections confirms the difference between iOS and Desktop ($p < .05$), between iOS and Android ($p < .05$), but not Android-Desktop.

Task 4: Two-way ANOVA

File used: postsbtwn.sav

It is often the case that we wish to examine the effects of more than one factor, and we also care about the interaction among factors. Because multiple factors are involved, this is called a factorial design, expressed as $N_1 \times N_2 \times \dots \times N_n$ for an arbitrary number n of factors, and where each N_i is an integer indicating the number of levels of that factor. In practice, it is difficult to interpret experiments with more than three factors, especially if those factors each have more than two levels. For this part, we will examine an augmented version of our current study that adds another factor. Open `postsbtwn.sav`. You will see another column labeled `Day` with values "weekday" and "weekend." These values correspond to the days of the week the subject was allowed to post to Facebook.

JASP: Very similar to previous homework on ANOVA. You have to simply add one more factor "Day" under Fixed Factors.

R: Do:

```
Anova(lm(Posts ~ Platform * Day, data = postsbtwn, type="III"))
```


FYI: For the car library, it explicitly requires us to build a linear model in order to use type 3 model. Type 3 is used here as it does not depend on the order of the factors (Type 1 does!). By default, aov is of type 1 while lm can specify the type. In the previous homework, we can use aov since we only have one factor.

Answer the following:

1. What is the output of the normality test? Is the data normal? How about the homogeneity tests?
 - a. A Shapiro-Wilk test and a Levene test confirm the normality and homogeneity of variances assumption check (both $p > .05$).
2. Was this a one-way, two-way, or three-way analysis of variance? What is/are the factor(s)? What are each factor's levels? Express the design using $N1 \times N2 \times \dots \times Nn$ notation.
 - a. Two-way because of two IV
 - i. Platform (2 levels)
 1. iOS
 2. Android
 - ii. Day (2 levels)
 1. Weekday
 2. Weekend
 - b. We conducted a 2 (Platform) x 2 (Day) between-subject experimental design.
3. For each identified factor, was it between-subjects or within-subjects? How do you know?
 - a. Both between subject. If the subject is exposed to all levels, there should be four columns, not two columns of IV. It uses a long format.
4. What were the means and standard deviations for the number of posts on weekdays? Weekends?
 - a. Weekday ($M = 26.300$, $SD = 8.430$)
 - b. Weekend ($M = 28.750$, $SD = 8.168$)
5. Write the F-test result for the Platform factor. Is it significant at the $\alpha = .05$ level? What is its p-value? Does this differ from the finding prior to the inclusion of the Day factor (above)? If so, how?
 - a. Platform $\rightarrow F(1, 36) = 4.058$, $p = .051$, $\eta_p^2 = 0.101$
 - b. Yes, it is STILL not significant but we explicitly write out the p-value and partial eta-square because they are very close (It's a culture in the research community)
6. Write the F-test result for the Day factor. Is it significant at the $\alpha = .05$ level? What is its p-value?
 - a. Day $\rightarrow F(1, 36) = 0.918$, *n.s.*
 - b. It is not significant
7. Write the F-test result for the Platform*Day interaction. Is it significant at the $\alpha = .05$ level? What is its p-value?
 - a. Day * Platform $\rightarrow F(1, 36) = 0.0004$, *n.s.*
 - b. It is not significant

8. Within each factor, why don't we need to perform any post hoc pairwise comparison tests?
 - a. Because each IV both has only two levels. We do post hoc only when the levels of a factor exceed 2.
9. What is the effect size? What is the interpretation?
 - a. For Platform which is near significance, the partial eta-squared is 0.1 which is toward large effect.
10. Notice in JASP, there is a non-parametric section under ANOVA. Try to put both factors under Kruskal-Wallis Test. What is the p-value? Does it differ from the parametric test? (For R people, try to find how to perform Kruskal-Wallis)

We fail to find a significant effect of both platform ($H(1) = 3.053$, *n.s.*) and day ($H(1) = 0.535$, *n.s.*) on the number of facebook posts between iOS ($Md = 24.500$) and Android ($Md = 27.500$) and between weekday ($Md = 27.000$) and weekend ($Md = 28.000$) (see Figure X).

11. Interpret these results and craft three sentences describing the results of this experiment, one for each factor and one for the interaction. What can we say about the findings from this study? (Hint: p-values between .05 and .10 are often called "trends" or "marginal results," and are often reported, although they cannot be considered strong evidence. Be wary of ever calling such results "marginally significant." A result is either significant or it is not; there is no "marginal significance.")

We found a marginally significant effect of Platform ($F_{1,36} = 4.058$, $p = .051$, $\eta_p^2 = 0.101$) between iOS ($M = 24.950$, $SD = 7.045$) and Android ($M = 30.100$, $SD = 8.795$). We failed to find a significant effect of Day ($F_{1,36} = 0.918$, *n.s.*) between weekday ($M = 26.300$, $SD = 8.430$) and weekend ($M = 28.750$, $SD = 8.168$) nor the interaction effect of Day * Platform ($F_{1,36} = 0.0004$, *n.s.*) on number of facebook posts.

Task 5: Repeated Measures ANOVA

File used: postwthn.sav

Thus far, we have only considered experiments where one subject was measured once on only one level of each factor. But often we wish to measure a subject more than once, perhaps for different levels of our factor(s), or over time, in which case time itself becomes a factor. Such designs are called "repeated measures" designs, and the factors on which we obtain repeated measures are called within-subjects factors (as opposed to between-subjects factors). For repeated measures studies, we can still use an ANOVA, but now we use a "repeated measures

ANOVA,” and our data table inevitably looks different: for a wide-format table, there are now multiple measures per row (each row still corresponds to just one subject, as it has thus far).

Our current hypothetical study on Facebook posts has been modified to be a purely within-subjects study. Imagine that each college student was issued either an iOS or Android device for one week, and then the other device for the next week. Also, each college student’s posts were counted separately on weekdays and weekends. Instead of needing 40 college students as before, we now only need 10 students for the same data, which is shown in postswthn.sav. Open those files and see the wide-format data tables.

Additional Info: In general, we should expect within-subjects studies to be more statistically powerful than between-subjects studies because subjects are, in effect, compared to themselves, and any given subject is more like himself than he is like any other subject. Within-subjects designs therefore reduce variance compared to between-subjects designs, and thus result in more power for detecting differences. The downside of using within-subjects designs is that they are subject to carryover effects and require careful counterbalancing. This makes them impractical in many circumstances.

JASP: Go to Repeated Measures ANOVA.

1. Under Repeated Measures Factor,
 - a. for RM Factor 1, rename to Platform - for level 1, rename to iOS, for level 2, rename to Android.
 - b. For RM Factor 2, double click and name it Day, for level 1, rename to Weekday, for level 2, rename to Weekend.
2. Then under Repeated Measures Cells, transfer the variable accordingly, e.g., iOS_weekday to iOS, Weekday.
3. Under Model > Components, make sure to shift and select all factors and click the right arrow
4. Assumption Checks >
 - a. check Homogeneity tests and Sphericity tests.
 - b. Select Descriptives on the top bar (*not inside ANOVA, the top bar*) and use the Shapiro Wilk test
5. If data is normal (Shapiro Wilk p-value > 0.05) and homogenous (Levene’s test > 0.05)
 - a. If the data does not pass the sphericity tests,
 - i. Use the p-value from the correction (e.g., Greenhouse Geiser)
 - b. If the data is either not normal or not homogenous
 - i. inside the ANOVA, go under Nonparametrics, and select the the factor accordingly. Note that Friedman tests (used by JASP) only support one factor at a time (with no between subject factor), so you have to do one by one..

R:

First, let R understand your table structure:

```
Platform <- factor(c("iOS", "iOS", "Android", "Android"))
Day <- factor(c("weekday", "weekend", "weekday", "weekend"))
factor <- data.frame(Platform, Day)
factor
```

The first row will be iOS weekday, second row will be iOS weekend, third row will be Android weekday and so on. The first row must match with the first column of the table, for example, the first column is iOS_weekday, which match with the factor first row.

Second, create a multivariate model with only the intercept as predictor

```
library(car)
model = lm(cbind(postswthn$iOS_weekday, postswthn$iOS_weekend,
postswthn$Android_weekday, postswthn$Android_weekend)~1)
```

Third, run the anova (idata is the structure, while idesign is the DV)

```
aov <- Anova(model, idata=factor, idesign=~Platform*Day,
type=3)
summary(aov, multivariate = FALSE)
```

Fourth, run the Mauchly test to confirm the anova did not violate the sphericity assumptions. We have to do for Platform, Day, and Platform:Day separately:

```
mauchly.test(model, M = ~Day, X = ~1, idata=factor)
mauchly.test(model, M = ~Platform, X = ~1, idata=factor)
mauchly.test(model, M = ~Platform:Day, X = ~Platform+Day, idata=factor)
```

1. What is the output of the normality test? Is the data normal? How about the test of sphericity?
 - a. Normality: All four combinations are normal, as shown by the Shapiro-Wilk test ($p > .05$).
 - b. Sphericity: Well, since we have only two levels, sphericity will not be a concern. Sphericity is a test of difference among variances which cannot be done when there are only two levels. Similarly, homogeneity test cannot be done since there are no between-subject factors.
2. Was this a one-way, two-way, or three-way analysis of variance? What is/are the factor(s)? What are each factor's levels? Express the design using $N_1 \times N_2 \times \dots \times N_n$ notation.
 - a. A Two-way Repeated Measures ANOVA because there are two IV
 - b. A 2 (Day) \times 2 (Platform) within-subject design is conducted.
3. For each identified factor, was it between-subjects or within-subjects? How do you know?
 - a. Both Day and Platform are within-subject factors. We can easily seen from the format of the table (short table instead of long table)
4. Write the statistical result for the Platform factor

- a. $F(1, 9) = 22.097, p < .01, \eta_p^2 = 0.711$
5. Write the statistical result for the Day factor.
 - a. $F(1, 9) = 0.368, n.s.$
6. Write the statistical result for the Platform*Day interaction.
 - a. $F(1, 9) = 0.004, n.s.$
7. What is the effect size in terms of partial eta squared? What is the interpretation?
 - a. For platform, the partial eta squared is 0.711 which is considered large, implying a practical significance.
8. Write the result in APA format.
 - a. A Shapiro-Wilk test confirms the normality of the data ($p > .05$). A Two-Way Repeated Measures ANOVA shows the main effect of Platform ($F(1, 9) = 22.097, p < .01, \eta_p^2 = 0.711$) on Posts between iOS ($M = 49.900, SD = 8.863$) and Android ($M = 60.200, SD = 10.347$). Anyhow, the test fails to find any significance of Day ($F(1, 9) = 0.368, n.s.$) on Posts between weekday ($M = 52.600, SD = 15.918$) and weekend ($M = 57.500, SD = 15.299$) nor the interaction effect of Platform * Day ($F(1, 9) = 0.004, n.s.$).

Task 6: Mixed Design ANOVA

File used: postmix.sav

Here, the platform is a between subject factor. Use a similar approach as Task 5, but consider the platform as between subject factors.

1. What is the output of the normality test? Is the data normal? How about the test of sphericity?
 - a. **Normality:** Both combinations are normal, as shown by the Shapiro-Wilk test ($p > .05$).
2. Was this a one-way, two-way, or three-way analysis of variance? What is/are the factor(s)? What are each factor's levels? Express the design using $N1 \times N2 \times \dots \times Nn$ notation.
 - a. A Two-Way Repeated Measures ANOVA
 - b. A 2 (Platform) x 2 (Day) mixed experimental design is conducted with Platform as between-subject factor and Day as within-subject factor.
3. For each identified factor, was it between-subjects or within-subjects? How do you know?
 - a. See 2
4. Write the statistical result for the Platform factor
 - a. $F(1, 18) = 5.716, p < .05, \eta_p^2 = 0.241$
5. Write the statistical result for the Day factor.

- a. $F(1, 18) = 0.712, n.s.$
6. Write the statistical result for the Platform*Day interaction.
 - a. $F(1, 18) = 0.00003, n.s.$ (*n.s. \Rightarrow not significant*)
7. What is the effect size in terms of partial eta squared? What is the interpretation?
 - a. For platform, the partial eta squared is 0.241 which is considered large, implying a practical significance.
8. Write the result in APA format.
 - a. A Shapiro-Wilk test confirms the normality of the data. *** A Two-Way Repeated Measures ANOVA with Platform as the between-subject factor shows the main effect of Platform ($F(1, 18) = 5.716, p < .05, \eta_p^2 = 0.241$) on Posts between iOS ($M = 49.900, SD = 8.863$) and Android ($M = 60.200, SD = 10.347$). Anyhow, the test fails to find any significance of Day ($F(1, 18) = 0.712, n.s.$) on Posts between weekday ($M = 26.300, SD = 8.430$) and weekend ($M = 28.750, SD = 8.168$) nor the interaction effect of Platform * Day ($F(1, 18) = 0.00003, n.s.$).
 - b. What if sphericity test say that we violate the assumption?
 - i. A Two-Way Repeated Measures ANOVA with Greenhouse-Geisser Correction and Platform as the between-subject factors shows.....
 - c. How to write to confirm that your ANOVA passes the sphericity check/homogeneity test:
 - i. *** A sphericity test confirms the assumption of homogeneity of variances among groups ($p < .05$) -> RP ANOVA
 - ii. *** A homogeneity test confirms the assumption of homogeneity of variances among groups ($p < .05$) -> between subject ANOVA
 - d. How to report the Posthocs, if the levels are more than three
 - i. A Two-Way Repeated Measures ANOVA with Platform as the between-subject factor shows the main effect of Platform ($F(1, 18) = 5.716, p < .05, \eta_p^2 = 0.241$) on Posts between iOS ($M = 49.900, SD = 8.863$), Android ($M = 60.200, SD = 10.347$) and ChakyOS ($M = , SD =$). A posthoc test with Bonferroni correction confirms the difference between ChakyOS and Android ($p < .05$), ChakyOS and iOS ($p < .05$), but not between iOS and Android.
 - ii. A posthoc test with Bonferroni correction > A Tukey posthoc test
 - e. What if your data is not normal?
 - i. A Shapiro-Wilk found that our data is not normal ($p < .05$) thus Kruskal-Wallis test is used. The test found