Factorial ANOVA Demo

Professor Christopher J. Schmank

2025-04-01

Factorial ANOVA | Research Context Prompt

A local dating expert was interested in studying first-date satisfaction and whether it depended on variables like (1) date attire (i.e., Athletic Wear, Leisure Wear, or Dressed Up) and (2) dating application used (OKCupid, Tinder).

Participants consented to be in their study after setting up their first date using either OKCupid or Tinder (this variable could not be randomly assigned). Upon meeting inclusion criteria and joining the study, the dating expert randomly assigned each participant to one of the three attire conditions.

The research question was: Does first-date satisfaction differ interactively when considering dating attire and the mobile application used?

Factorial ANOVA | Data

```r
library(psych)
library(tidyverse)
library(jmv)
library(ggpubr)
library(apaTables)
library(ez)
library(rstatix)

dat_date <- read.csv("FACTORIAL_DEMO.csv")

# Set IVs to factors using as.factor() or factor()
dat_date$Application<-as.factor(dat_date$Application)
dat_date$Outfit<-as.factor(dat_date$Outfit)

# Assess number of participants and variables in data set
dim(dat_date)
```

```
## [1] 120   3
```

```r
head(dat_date,4)
```

```
##   Application       Outfit Satisfaction
## 1     OkCupid AthleticWear            6
## 2     OkCupid AthleticWear            6
## 3     OkCupid AthleticWear            7
## 4     OkCupid AthleticWear            8
```

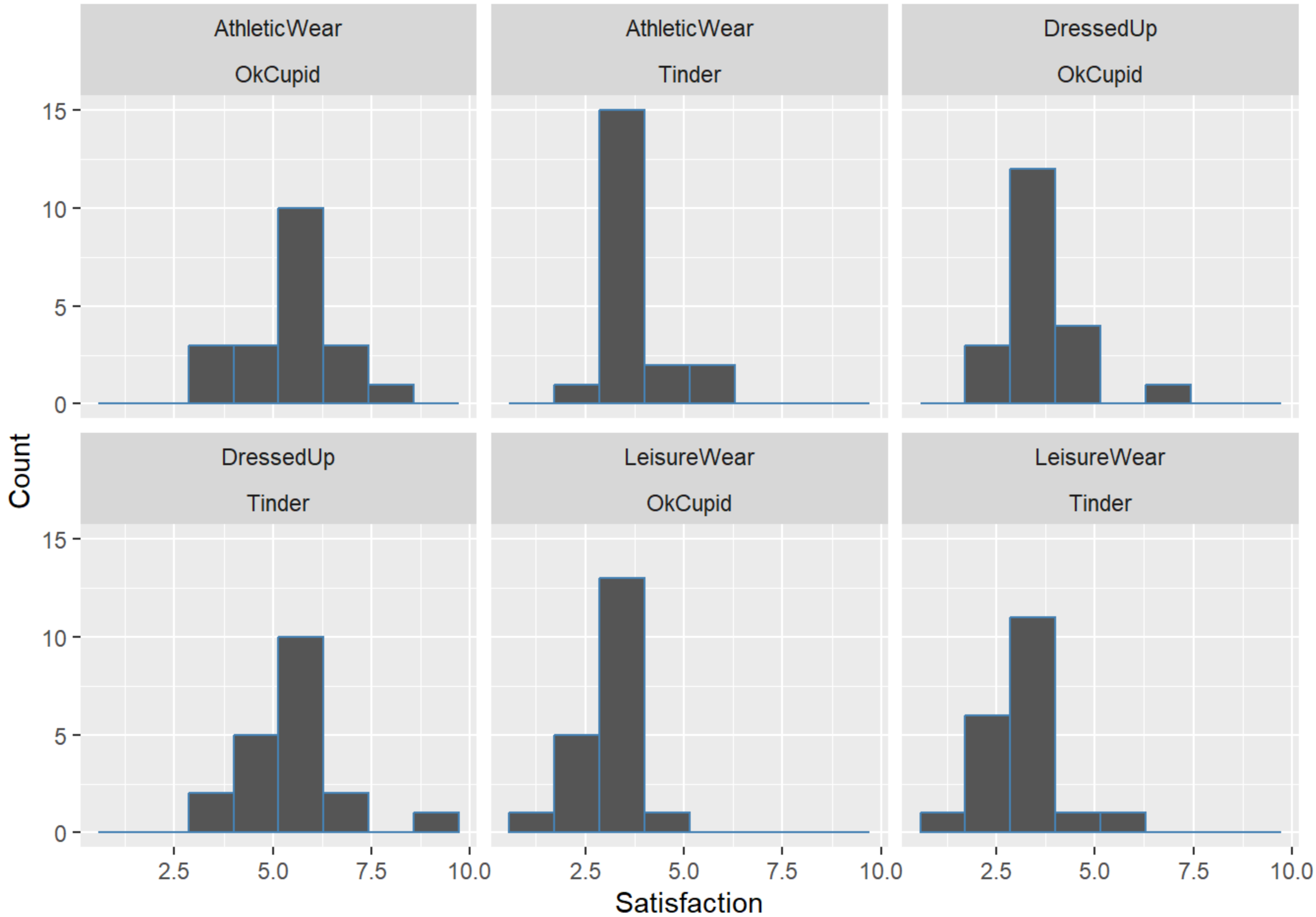
Factorial ANOVA | Descriptive Stats

```
# Descriptive Stats
describeBy(dat_date$Satisfaction,
           list(dat_date$Outfit,dat_date$Application),
           mat=TRUE)
```

```
##     item       group1  group2 vars  n mean        sd median trimmed    mad min
## X11    1 AthleticWear OkCupid    1 20 5.80 1.0563094      6  5.8125 0.7413   4
## X12    2    DressedUp OkCupid    1 20 3.75 1.2513151      4  3.6875 1.4826   2
## X13    3  LeisureWear OkCupid    1 20 2.90 0.9119095      3  2.8750 0.7413   1
## X14    4 AthleticWear  Tinder    1 20 3.90 1.0208356      4  3.8125 1.4826   2
## X15    5    DressedUp  Tinder    1 20 5.80 1.1050125      6  5.7500 0.7413   4
## X16    6  LeisureWear  Tinder    1 20 3.05 1.1909748      3  2.9375 1.4826   1
##     max range       skew    kurtosis        se
## X11   8     4 -0.1323583 -0.49459060 0.2361980
## X12   7     5  0.6028951  0.08746646 0.2798026
## X13   5     4  0.1819801 -0.09670045 0.2039092
## X14   6     4  0.4681251 -0.29305249 0.2282658
## X15   9     5  0.8182151  1.42879608 0.2470883
## X16   6     5  0.6202259 -0.05741963 0.2663101
```

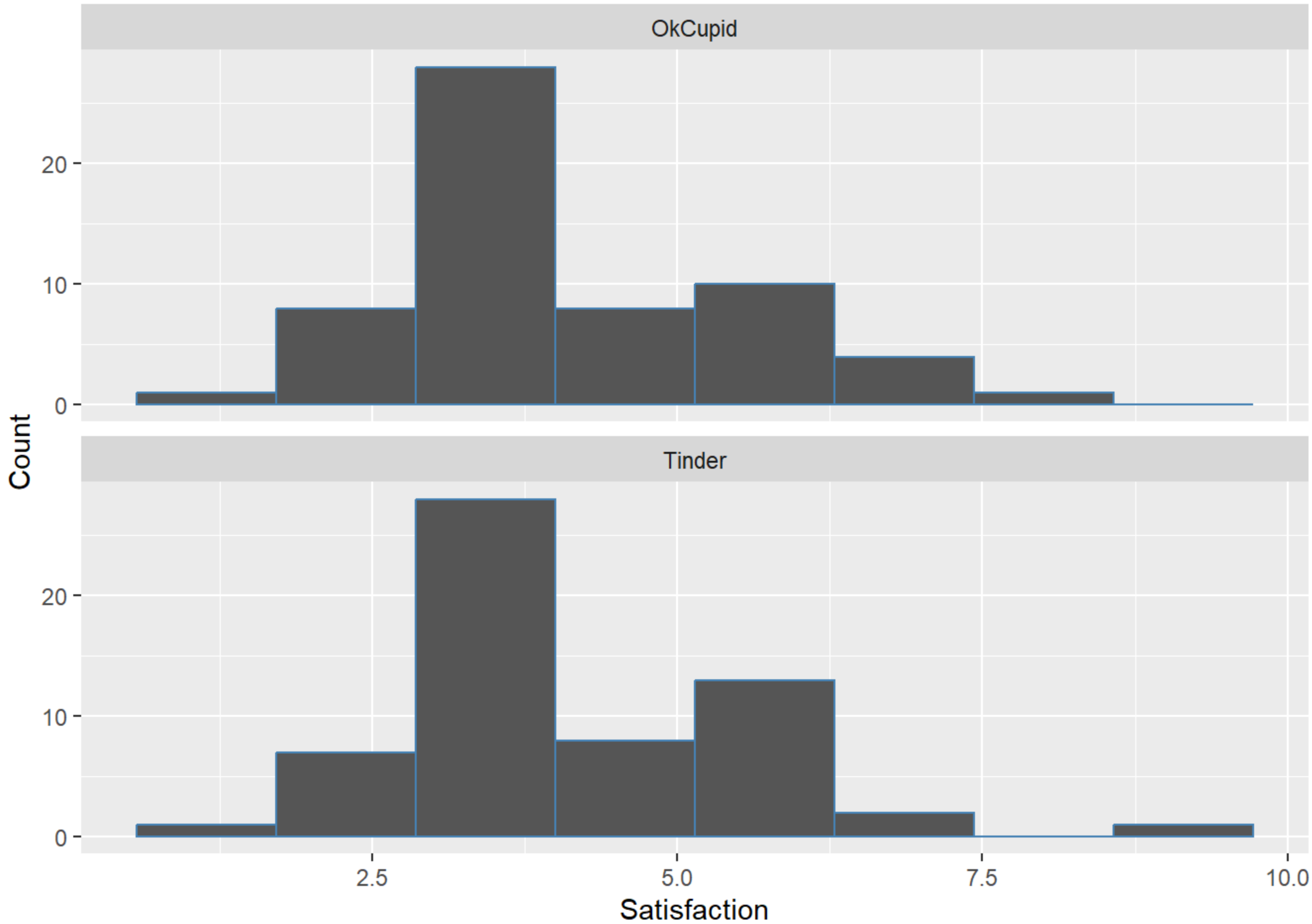
Factorial ANOVA | Histogram

```
library(ggplot2)
ggplot(data = dat_date,
       mapping = aes(x = Satisfaction)) +
  geom_histogram(bins = 8, color = "steelblue") +
  labs(y = "Count", x = "Satisfaction")+
  facet_wrap(~dat_date$Outfit+dat_date$Application)
```



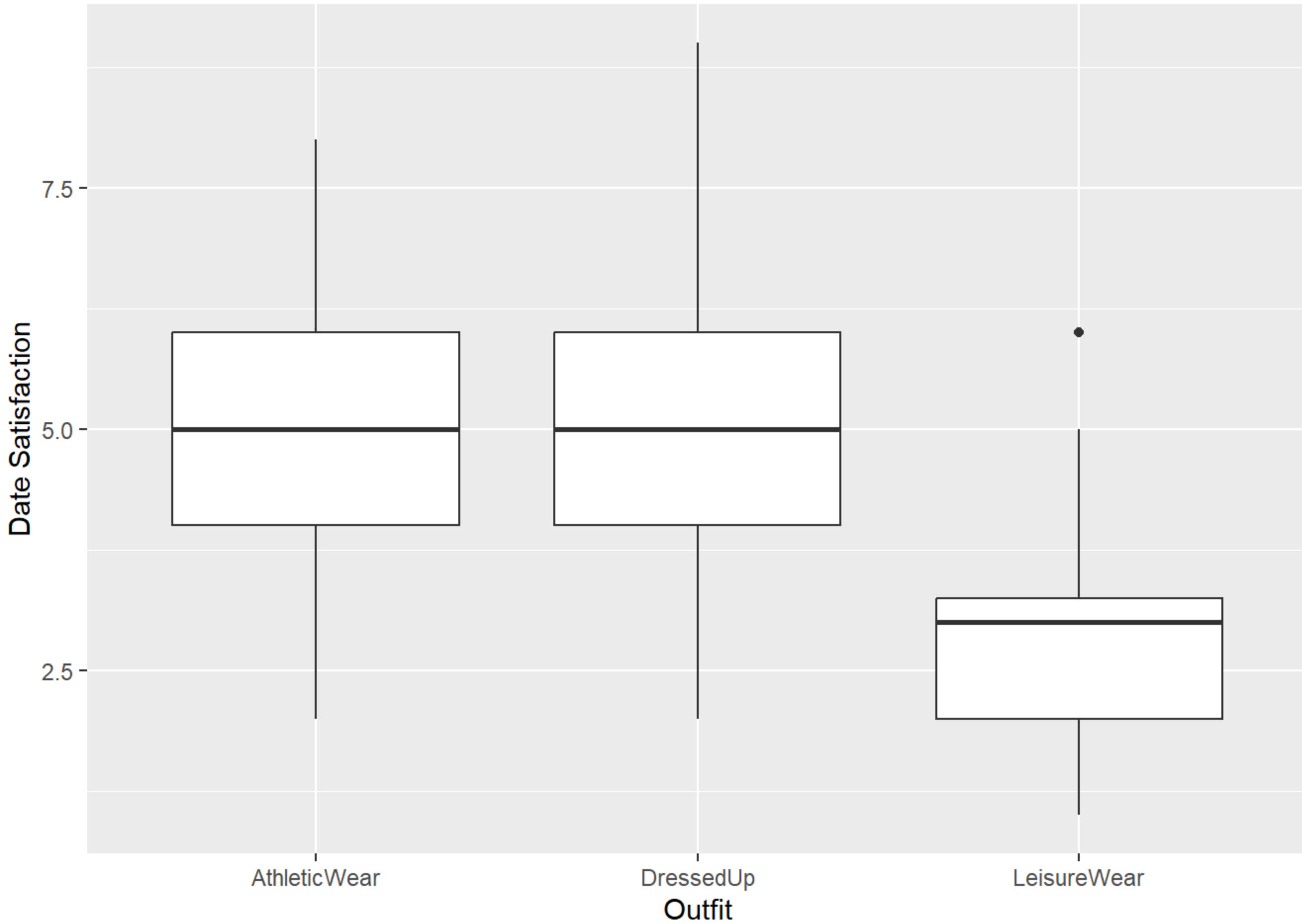
Factorial ANOVA | Histogram

```
ggplot(data = dat_date,
       mapping = aes(x = Satisfaction)) +
  geom_histogram(bins = 8, color = "steelblue") +
  labs(y = "Count", x = "Satisfaction")+
  facet_wrap(dat_date$Application,ncol=1)
```



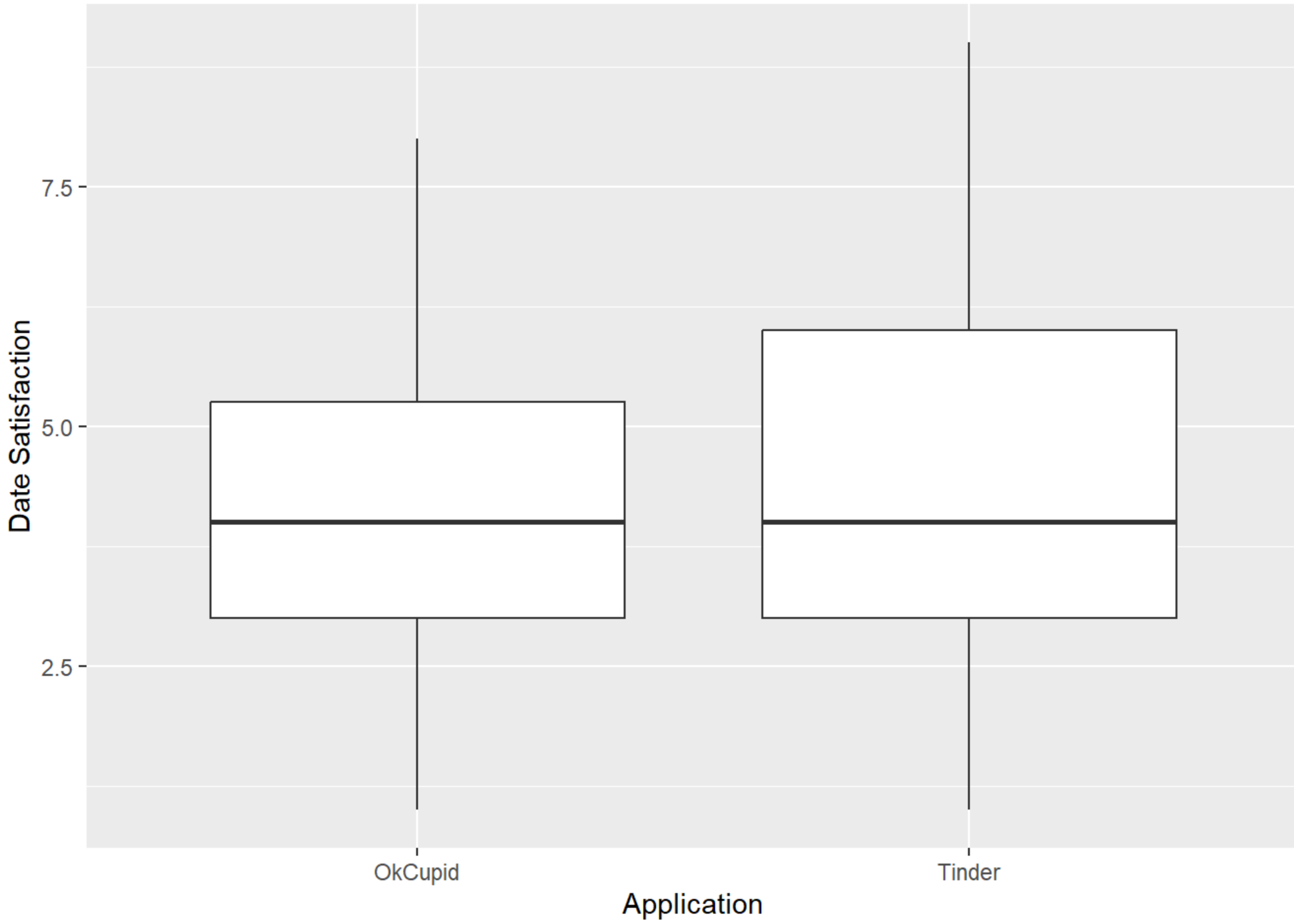
Factorial ANOVA | Boxplot

```
ggplot(data = dat_date,
       mapping = aes(y = Satisfaction, x = Outfit)) +
  geom_boxplot() +
    labs(y = "Date Satisfaction")
```



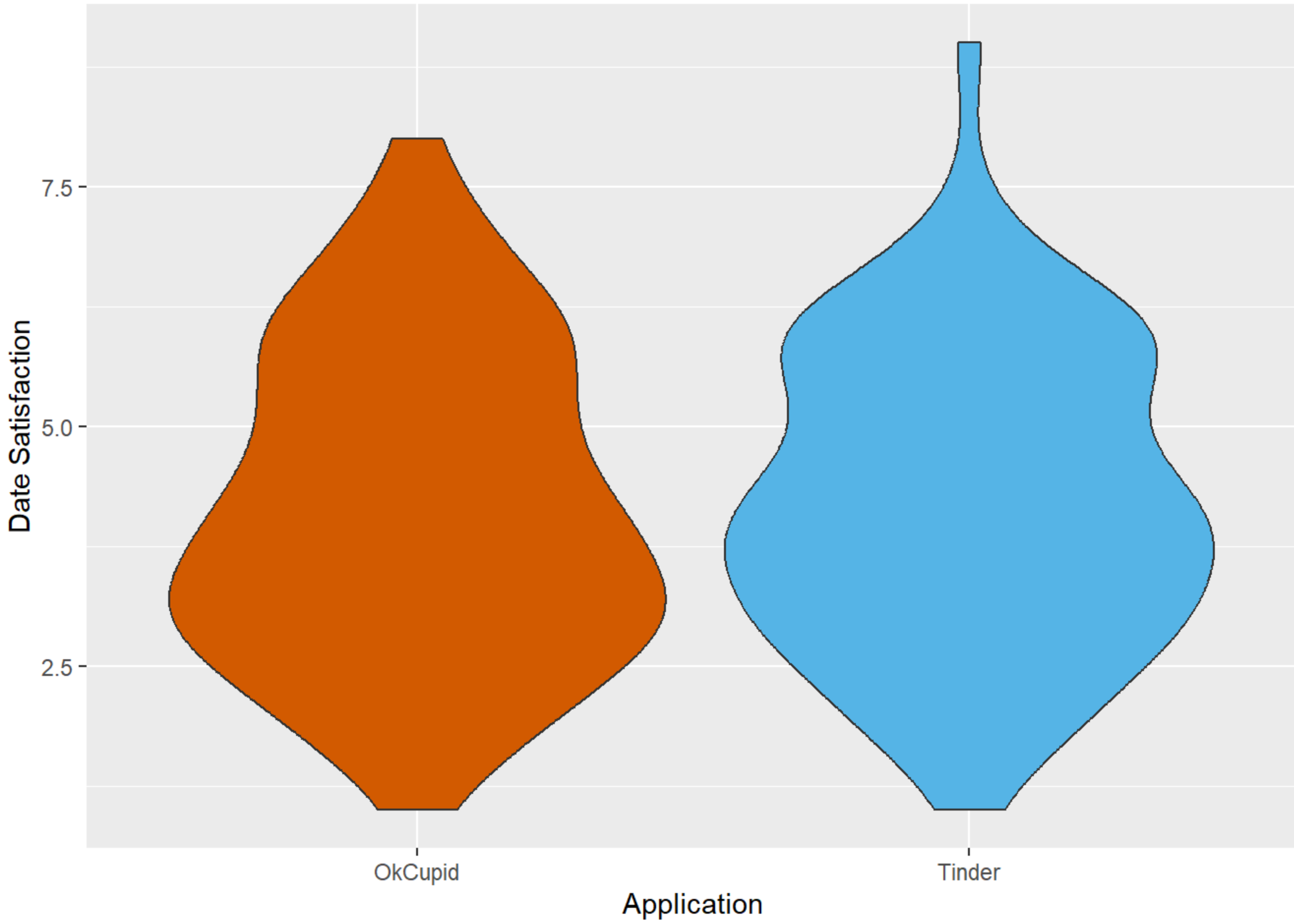
Factorial ANOVA | Boxplot

```
ggplot(data = dat_date,
       mapping = aes(y = Satisfaction, x = Application)) +
  geom_boxplot() +
    labs(y = "Date Satisfaction")
```

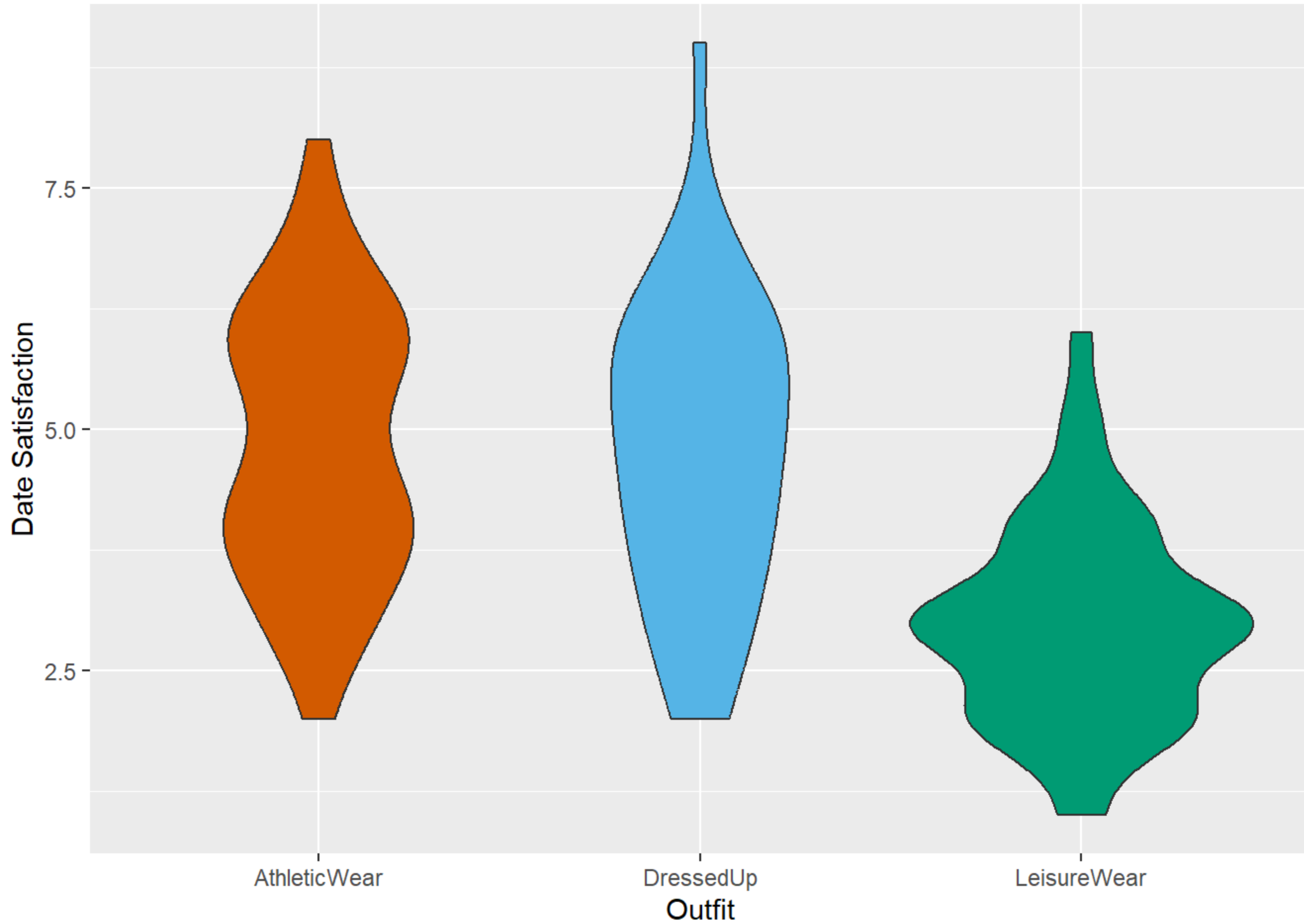


Factorial ANOVA | Violin Plot

```
ggplot(data = dat_date,
       aes(x = Application, y = Satisfaction,
           fill = Application))+
  geom_violin(show.legend=FALSE)+
  ylab("Date Satisfaction")+
  xlab("Application")+
  scale_fill_manual(values=c("#D55e00","#56B4E9","#009E73"))
```

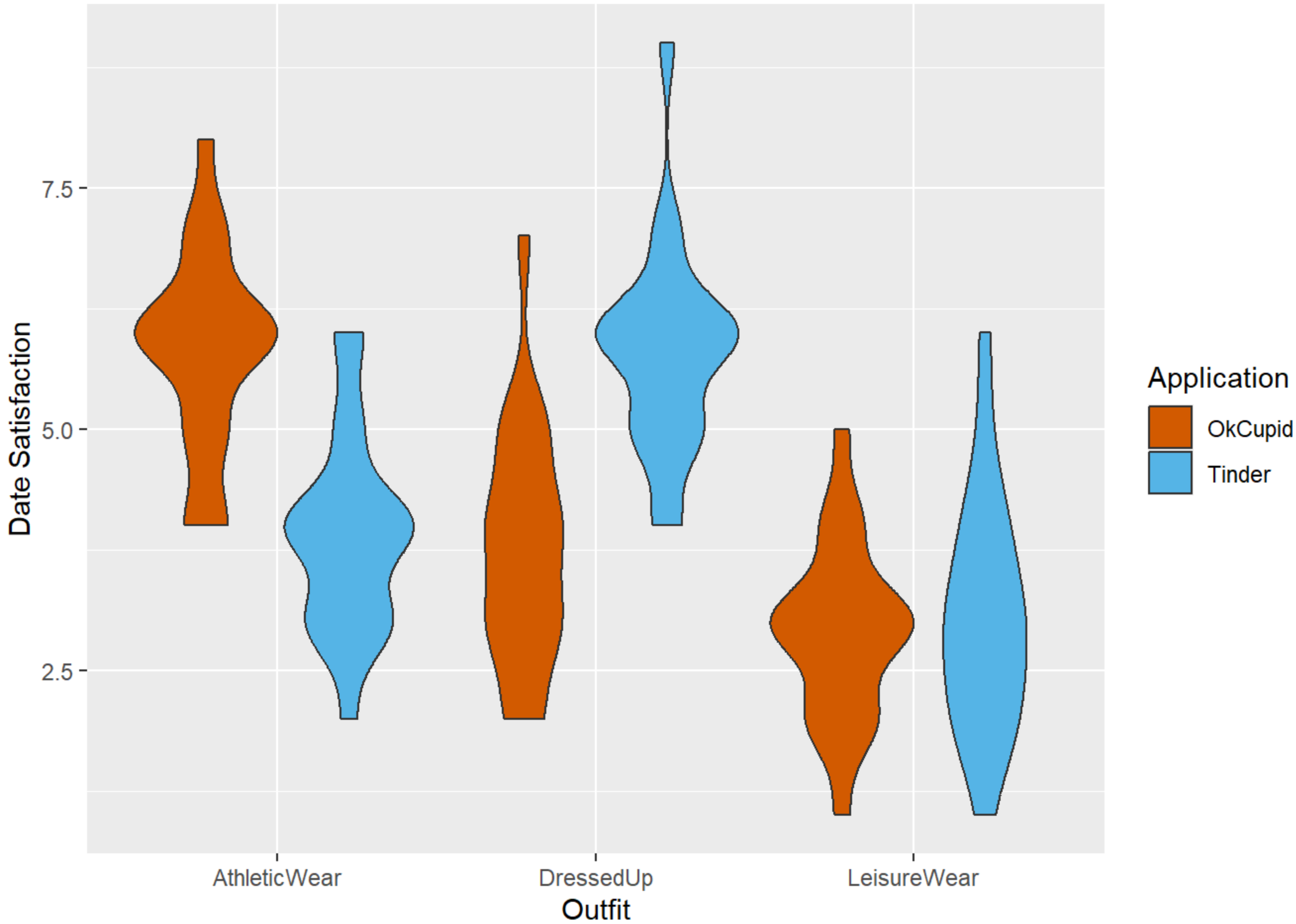



```
ggplot(data = dat_date,
       aes(x = Outfit, y = Satisfaction,
           fill = Outfit))+
  geom_violin(show.legend=FALSE)+
  ylab("Date Satisfaction")+
  xlab("Outfit")+
  scale_fill_manual(values=c("#D55e00","#56B4E9","#009E73"))
```



Factorial ANOVA | Violin Plot

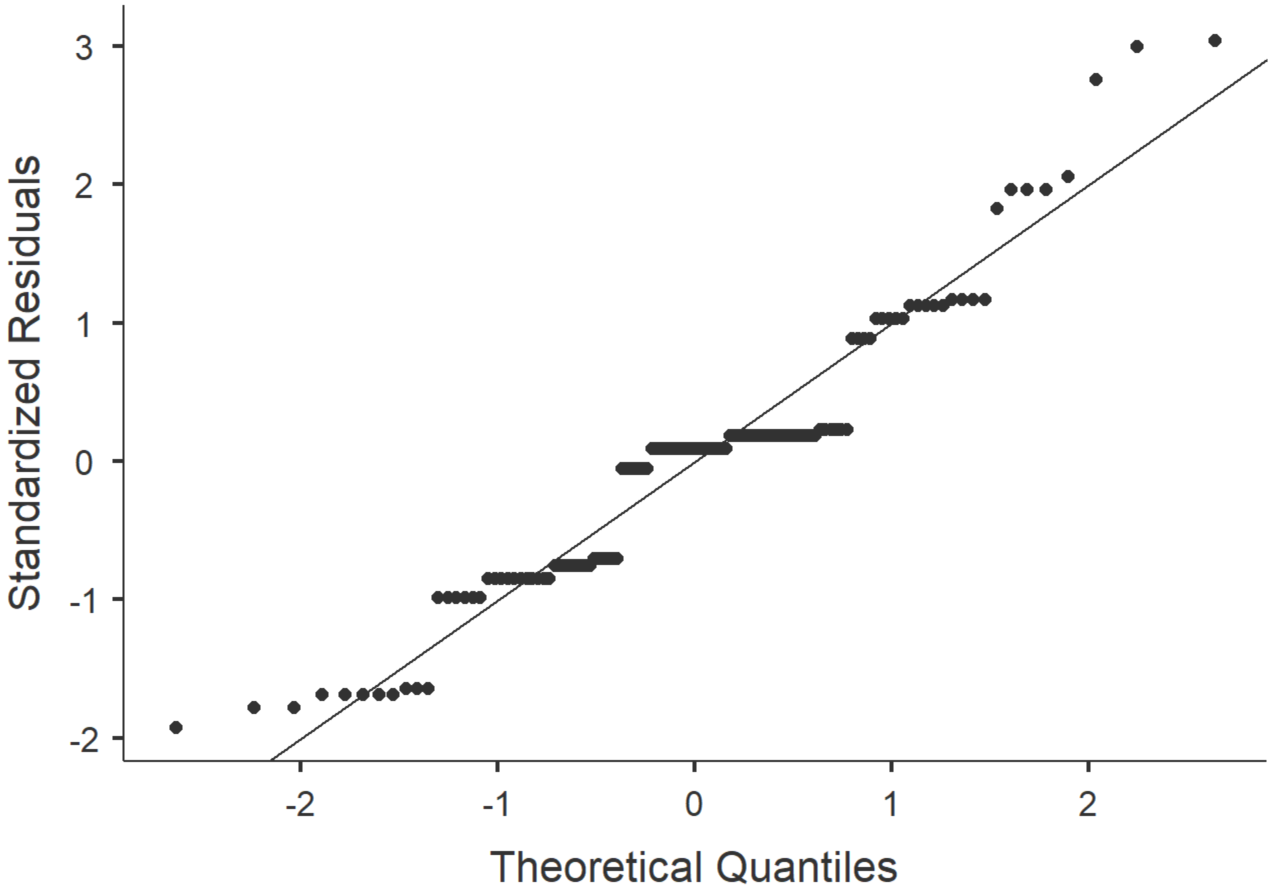
```
ggplot(data = dat_date,
       aes(x = Outfit, y = Satisfaction,
           fill = Application))+
  geom_violin(show.legend=TRUE)+
  ylab("Date Satisfaction")+
  xlab("Outfit")+
  scale_fill_manual(values=c("#D55e00","#56B4E9","#009E73"))
```



Factorial ANOVA | ANOVA() Omnibus Assumption Checks

```
# Omnibus Assumption Checking
ANOVA(data = dat_date,
      dep = 'Satisfaction',
      factors = c('Application','Outfit'),
      homo = TRUE,
      norm = TRUE,
      qq = TRUE)
```

```
##
##  ANOVA
##
##  ANOVA - Satisfaction
##  ──────────────────────────────────────────────────────────────────────────────────────────
##                               Sum of Squares    df     Mean Square    F             p
##  ──────────────────────────────────────────────────────────────────────────────────────────
##    Application                     0.3000000      1      0.3000000     0.2501829     0.6179105
##    Outfit                         90.1500000      2     45.0750000    37.5899781    < .0000001
##    Application:Outfit             78.0500000      2     39.0250000    32.5446233    < .0000001
##    Residuals                     136.7000000    114      1.1991228
##  ──────────────────────────────────────────────────────────────────────────────────────────
##
##
##  ASSUMPTION CHECKS
##
##  Homogeneity of Variances Test (Levene's)
##  ────────────────────────────────────────
##    F            df1    df2    p
##  ────────────────────────────────────────
##    0.5146008      5    114    0.7647589
##  ────────────────────────────────────────
##
##
##  Normality Test (Shapiro-Wilk)
##  ─────────────────────────────
##    Statistic    p
##  ─────────────────────────────
##    0.9379552    0.0000312
##  ─────────────────────────────
```



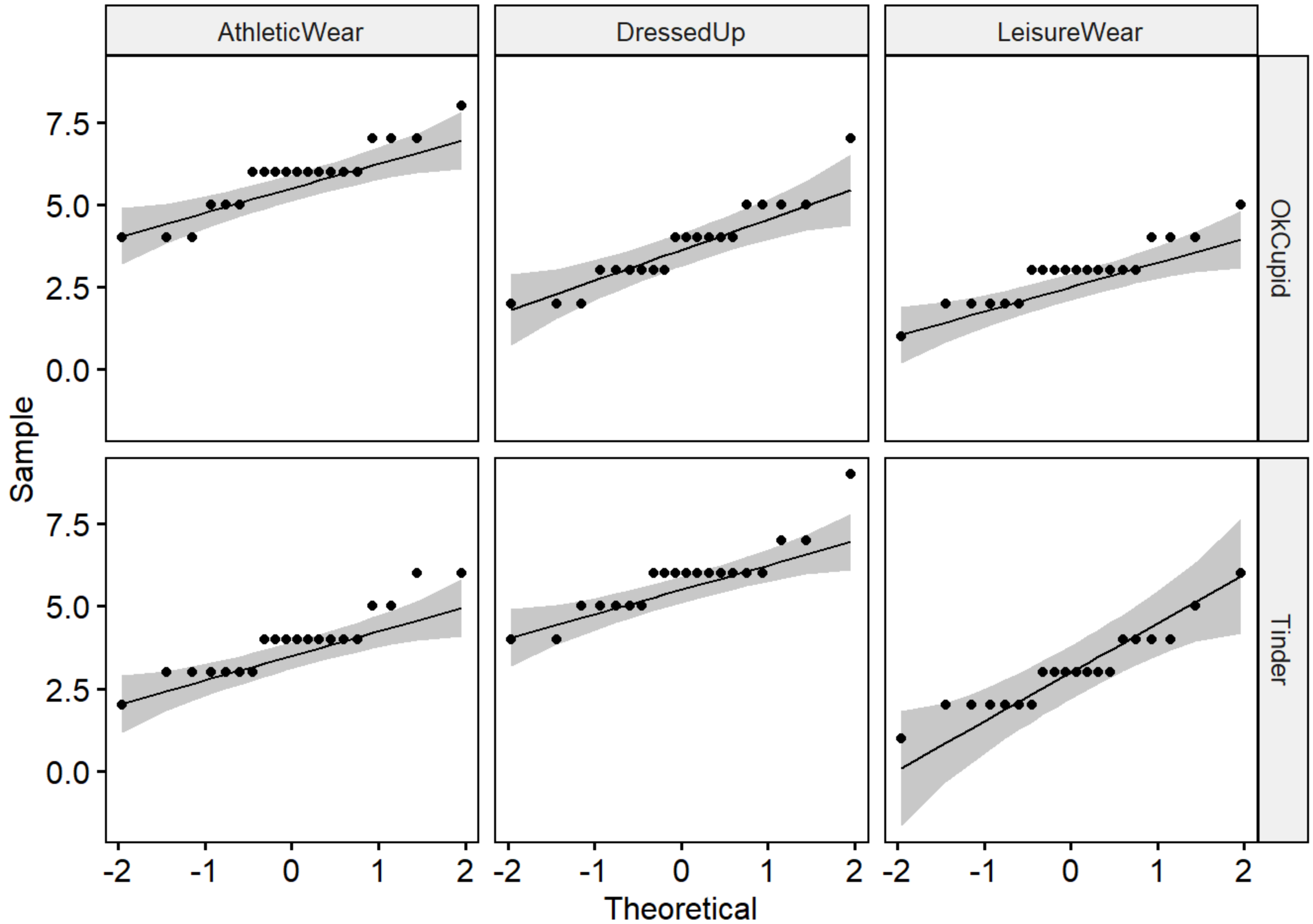
Factorial ANOVA | Group Level Assumption Checks

```
dat_date %>%
  group_by(Application,Outfit) %>%
  shapiro_test(Satisfaction)
```

```
## # A tibble: 6 × 5
##   Application Outfit       variable     statistic       p
##   <fct>       <fct>        <chr>            <dbl>   <dbl>
## 1 OkCupid     AthleticWear Satisfaction     0.890 0.0273
## 2 OkCupid     DressedUp    Satisfaction     0.909 0.0623
## 3 OkCupid     LeisureWear  Satisfaction     0.898 0.0381
## 4 Tinder      AthleticWear Satisfaction     0.887 0.0235
## 5 Tinder      DressedUp    Satisfaction     0.856 0.00663
## 6 Tinder      LeisureWear  Satisfaction     0.915 0.0796
```

Factorial ANOVA | Group Level Assumption Checks

```
ggqqplot(dat_date,
         'Satisfaction',
         facet.by = c("Application","Outfit"))
```



Factorial ANOVA | Group Level Assumption Checks

```
# If you have not already turned your variables into factors you need to do so before this step!

dat_date %>%
  levene_test(Satisfaction ~ Application*Outfit,
              center = mean)
```

```
## # A tibble: 1 × 4
##     df1   df2 statistic     p
##   <int> <int>     <dbl> <dbl>
## 1     5   114     0.515 0.765
```

Factorial ANOVA | Conducting the ANOVA

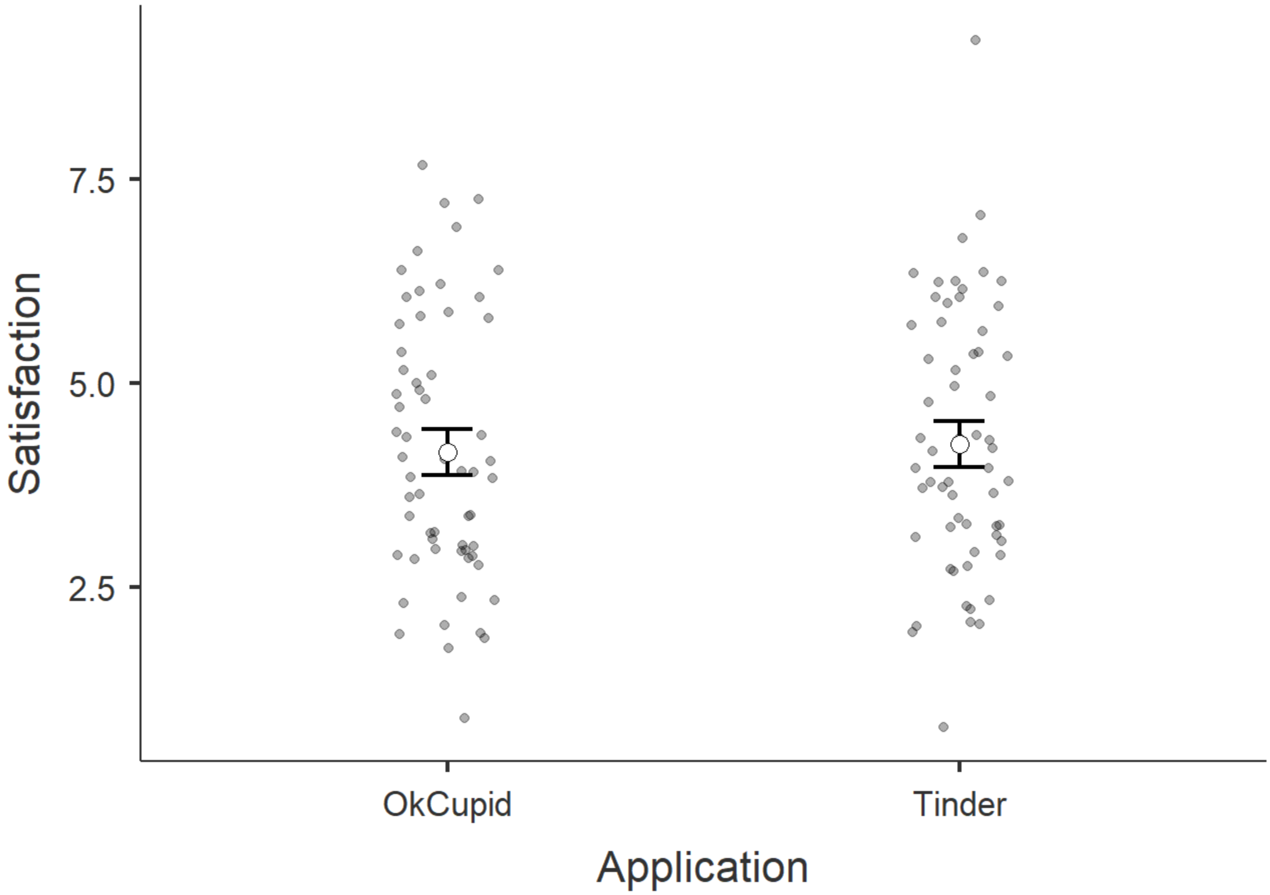
```
options(digits = 3)
ANOVA(data = dat_date,
      dep = 'Satisfaction',
      factors = list('Application','Outfit'),
      effectSize = 'partEta',
      postHoc = ~Outfit + Application:Outfit,
      postHocCorr = 'bonf',
      postHocES = 'd',
      postHocEsCi = TRUE,
      emMeans = ~Application + Outfit + Application:Outfit,
      emmPlots = TRUE,
      emmPlotData = TRUE,
      emmTables = TRUE)
```

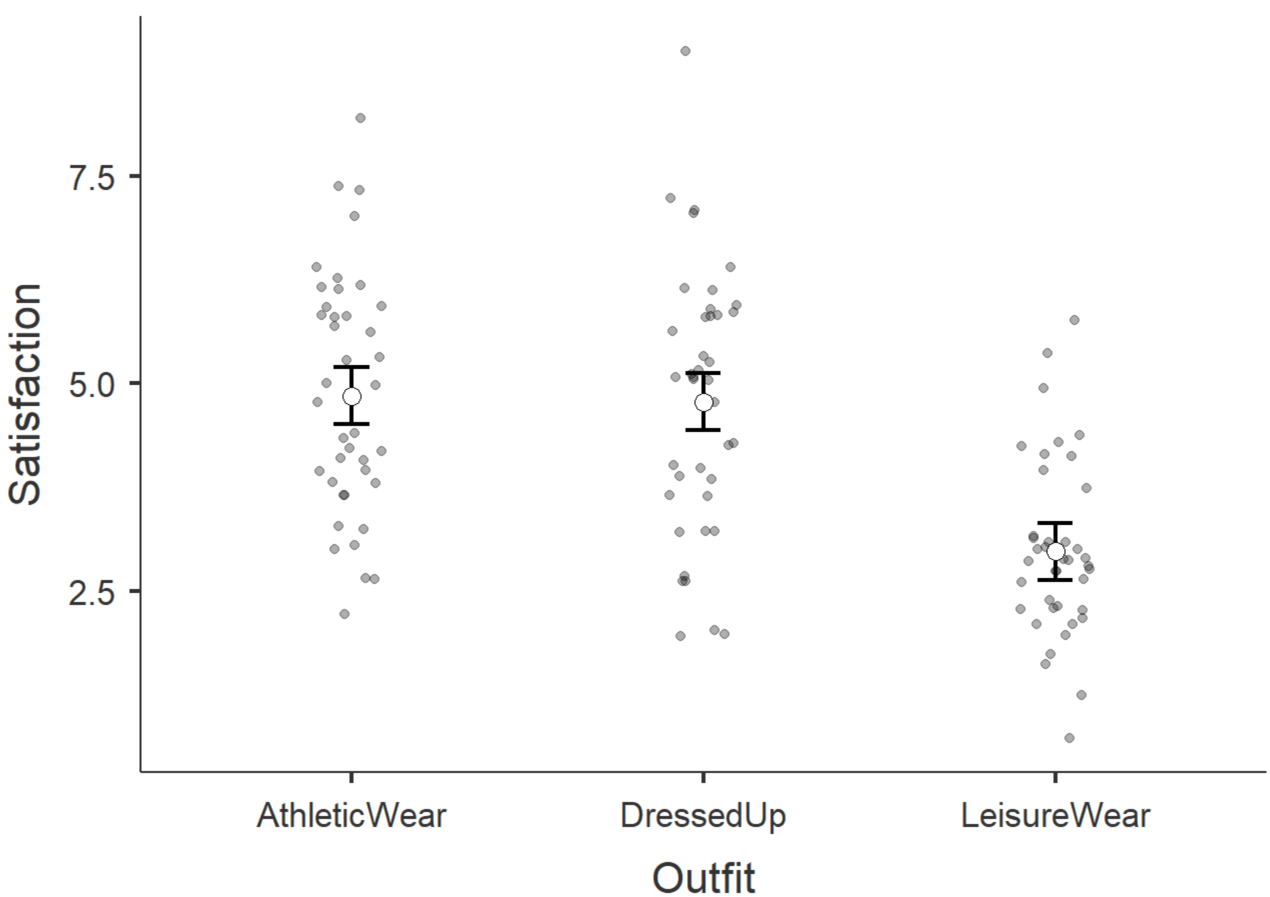
```
## NOTE: Results may be misleading due to involvement in interactions
```

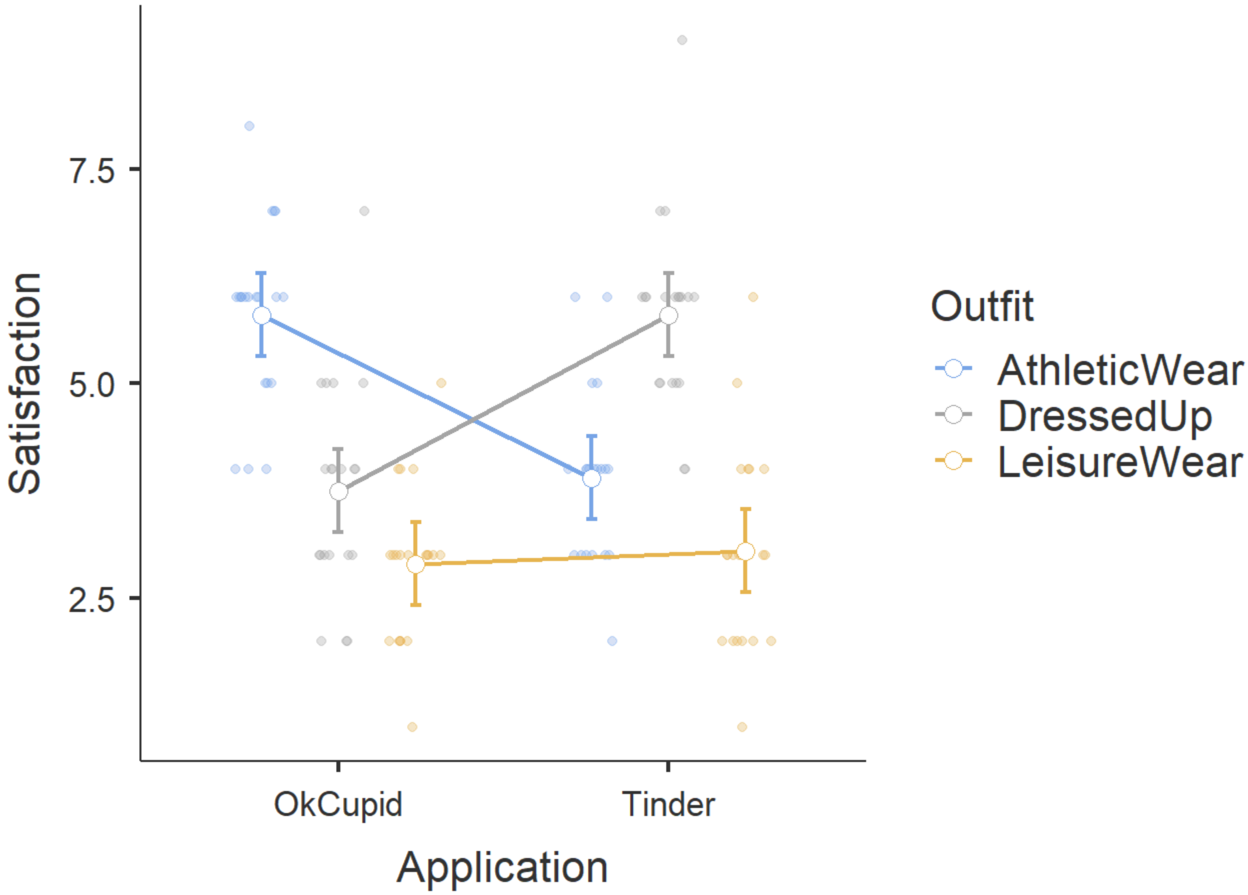
```
##
##  ANOVA
##
##  ANOVA - Satisfaction
##  ─────────────────────────────────────────────────────────────────────────────────────────────
##                               Sum of Squares    df     Mean Square    F         p         η²p
##  ─────────────────────────────────────────────────────────────────────────────────────────────
##    Application                         0.300      1          0.300     0.250     0.618    0.002
##    Outfit                             90.150      2         45.075    37.590    < .001    0.397
##    Application:Outfit                 78.050      2         39.025    32.545    < .001    0.363
##    Residuals                         136.700    114          1.199
##  ─────────────────────────────────────────────────────────────────────────────────────────────
##
##
##  POST HOC TESTS
##
##  Post Hoc Comparisons - Outfit
##  ──────────────────────────────────────────────────────────────────────────────────────────────
────────────────────────────────────
##    Outfit               Outfit         Mean Difference    SE        df     t         p-bonferro
ni    Cohen's d    Lower     Upper
##  ──────────────────────────────────────────────────────────────────────────────────────────────
────────────────────────────────────
##    AthleticWear    -    DressedUp               0.0750    0.245    114    0.306           1.0
00       0.0685    -0.375    0.512
##                    -    LeisureWear             1.8750    0.245    114    7.657    < .001
1.7123     1.216    2.209
##    DressedUp       -    LeisureWear             1.8000    0.245    114    7.351    < .001
1.6438     1.151    2.136
##  ──────────────────────────────────────────────────────────────────────────────────────────────
────────────────────────────────────
##    Note. Comparisons are based on estimated marginal means
##
##
##  Post Hoc Comparisons - Application:Outfit
##  ──────────────────────────────────────────────────────────────────────────────────────────────
──────────────────────────────────────────────────────────────────────────
##    Application    Outfit              Application    Outfit          Mean Difference    SE
df     t            p-bonferroni    Cohen's d    Lower       Upper
##  ──────────────────────────────────────────────────────────────────────────────────────────────
──────────────────────────────────────────────────────────────────────────
##    OkCupid        AthleticWear    -    OkCupid        DressedUp                  2.050    0.34
6    114        5.920    < .001        1.872     1.19920     2.545
##                                   -    OkCupid        LeisureWear                2.900    0.34
6    114        8.375    < .001        2.648     1.93195     3.365
##                                   -    Tinder         AthleticWear               1.900    0.34
6    114        5.487    < .001        1.735     1.06857     2.402
##                                   -    Tinder         DressedUp              -1.94e-15    0.34
6    114    -5.61e-15           1.000    -8.88e-16    -0.62644     0.626
##                                   -    Tinder         LeisureWear                2.750    0.34
6    114        7.941    < .001        2.511     1.80351     3.219
##                   DressedUp       -    OkCupid        LeisureWear                0.850    0.34
```

```
6   114       2.455          0.234         0.776     0.14156     1.411
##                             -    Tinder        AthleticWear              -0.150    0.34
6   114      -0.433          1.000         0.137    -0.48972     0.764
##                             -    Tinder        DressedUp                 -2.050    0.34
6   114      -5.920    < .001        -1.872    -2.54494    -1.199
##                             -    Tinder        LeisureWear                0.700    0.34
6   114       2.021          0.684         0.639     0.00721     1.271
##                LeisureWear     -    Tinder        AthleticWear              -1.000    0.34
6   114      -2.888          0.070         0.913     0.27541     1.551
##                             -    Tinder        DressedUp                 -2.900    0.34
6   114      -8.375    < .001         2.648     1.93195     3.365
##                             -    Tinder        LeisureWear               -0.150    0.34
6   114      -0.433          1.000        -0.137    -0.76368     0.490
##   Tinder         AthleticWear    -    Tinder        DressedUp                 -1.900    0.34
6   114      -5.487    < .001        -1.735    -2.40161    -1.069
##                             -    Tinder        LeisureWear                0.850    0.34
6   114       2.455          0.234         0.776     0.14156     1.411
##                DressedUp       -    Tinder        LeisureWear                2.750    0.34
6   114       7.941    < .001         2.511     1.80351     3.219
##  ─────────────────────────────────────────────────────────────────────────────────────────
──────────────────────────────────────────────────────────────────────
##   Note. Comparisons are based on estimated marginal means
##
##
##  ESTIMATED MARGINAL MEANS
##
##  APPLICATION
##
##  Estimated Marginal Means - Application
##  ──────────────────────────────────────────────────
##   Application    Mean    SE       Lower    Upper
##  ──────────────────────────────────────────────────
##   OkCupid        4.15    0.141     3.87     4.43
##   Tinder         4.25    0.141     3.97     4.53
##  ──────────────────────────────────────────────────
##
##
##  OUTFIT
##
##  Estimated Marginal Means - Outfit
##  ───────────────────────────────────────────────────
##   Outfit          Mean    SE       Lower    Upper
##  ───────────────────────────────────────────────────
##   AthleticWear    4.85    0.173     4.51     5.19
##   DressedUp       4.78    0.173     4.43     5.12
##   LeisureWear     2.98    0.173     2.63     3.32
##  ───────────────────────────────────────────────────
##
##
##  APPLICATION:OUTFIT
##
##  Estimated Marginal Means - Application:Outfit
```

```
## ───────────────────────────────────────────────────────────────
##   Outfit          Application    Mean    SE       Lower    Upper
## ───────────────────────────────────────────────────────────────
##   AthleticWear    OkCupid        5.80    0.245     5.31     6.29
##                   Tinder         3.90    0.245     3.41     4.39
##   DressedUp       OkCupid        3.75    0.245     3.26     4.24
##                   Tinder         5.80    0.245     5.31     6.29
##   LeisureWear     OkCupid        2.90    0.245     2.41     3.39
##                   Tinder         3.05    0.245     2.56     3.54
## ───────────────────────────────────────────────────────────────
```


Satisfaction
7.5
5.0
2.5
OkCupid
Tinder
Application


Satisfaction
7.5
5.0
2.5
AthleticWear
DressedUp
LeisureWear
Outfit



Factorial ANOVA | ANOVA Output: Interaction Simple Effects Code

```
model <- lm(Satisfaction~Application*Outfit, dat_date)
# Effect of Outfit across each Application
# When statistically significant interaction occurs we assess simple effects
one.way <- dat_date %>%
  group_by(Outfit) %>%
  anova_test(dv = Satisfaction,
             between = Application,
             error = model)%>%
  get_anova_table() %>%
  adjust_pvalue(method = "bonferroni")
one.way
```

```
## # A tibble: 3 × 9
##   Outfit       Effect       DFn   DFd     F            p `p<.05`   ges    p.adj
##   <fct>        <chr>       <dbl> <dbl>  <dbl>        <dbl> <chr>   <dbl>    <dbl>
## 1 AthleticWear Application     1   114 30.1   0.00000025   "*"     0.209 7.5 e-7
## 2 DressedUp    Application     1   114 35.0   0.0000000346 "*"     0.235 1.04e-7
## 3 LeisureWear  Application     1   114  0.188 0.666        ""      0.002 1    e+0
```

```
one.way2 <- dat_date %>%
  group_by(Application) %>%
  anova_test(dv = Satisfaction,
             between = Outfit,
             error = model)%>%
  get_anova_table() %>%
  adjust_pvalue(method = "bonferroni")
one.way2
```

```
## # A tibble: 2 × 9
##   Application Effect   DFn   DFd     F        p `p<.05`   ges    p.adj
##   <fct>       <chr>  <dbl> <dbl> <dbl>    <dbl> <chr>   <dbl>    <dbl>
## 1 OkCupid     Outfit     2   114  37.1 3.97e-13 *       0.394 7.94e-13
## 2 Tinder      Outfit     2   114  33.1 4.73e-12 *       0.367 9.46e-12
```

The simple main effect of Outfit on date satisfaction was statistically significant for both OKCupid and Tinder users ($p < .001$). In other words, there was a statistically significant difference in mean date satisfaction scores between OKCupid users based on date attire [$F(2,114) = 37.07$, $p < .001$]. And the same conclusion holds true for Tinder users [$F(2,114) = 33.07$, $p < .001$].

Factorial ANOVA | Professional ANOVA Visualization Code

```
# Generates Marginal Means for Visualization
pwc <- dat_date %>%
  group_by(Application) %>%
  emmeans_test(Satisfaction~Outfit,
               p.adjust.method = "bonferroni") %>%
  add_xy_position(x = "Application")

pwc
```

```
## # A tibble: 6 × 15
##   Application term   .y.          group1 group2    df statistic        p    p.adj
##   <fct>       <chr>  <chr>        <chr>  <chr>  <dbl>     <dbl>    <dbl>    <dbl>
## 1 OkCupid     Outfit Satisfacti… Athle… Dress…   114      5.92 3.46e- 8 1.04e- 7
## 2 OkCupid     Outfit Satisfacti… Athle… Leisu…   114      8.37 1.61e-13 4.83e-13
## 3 OkCupid     Outfit Satisfacti… Dress… Leisu…   114      2.45 1.56e- 2 4.68e- 2
## 4 Tinder      Outfit Satisfacti… Athle… Dress…   114     -5.49 2.50e- 7 7.50e- 7
## 5 Tinder      Outfit Satisfacti… Athle… Leisu…   114      2.45 1.56e- 2 4.68e- 2
## 6 Tinder      Outfit Satisfacti… Dress… Leisu…   114      7.94 1.55e-12 4.65e-12
## # ℹ 6 more variables: p.adj.signif <chr>, y.position <dbl>,
## #   groups <named list>, x <dbl>, xmin <dbl>, xmax <dbl>
```

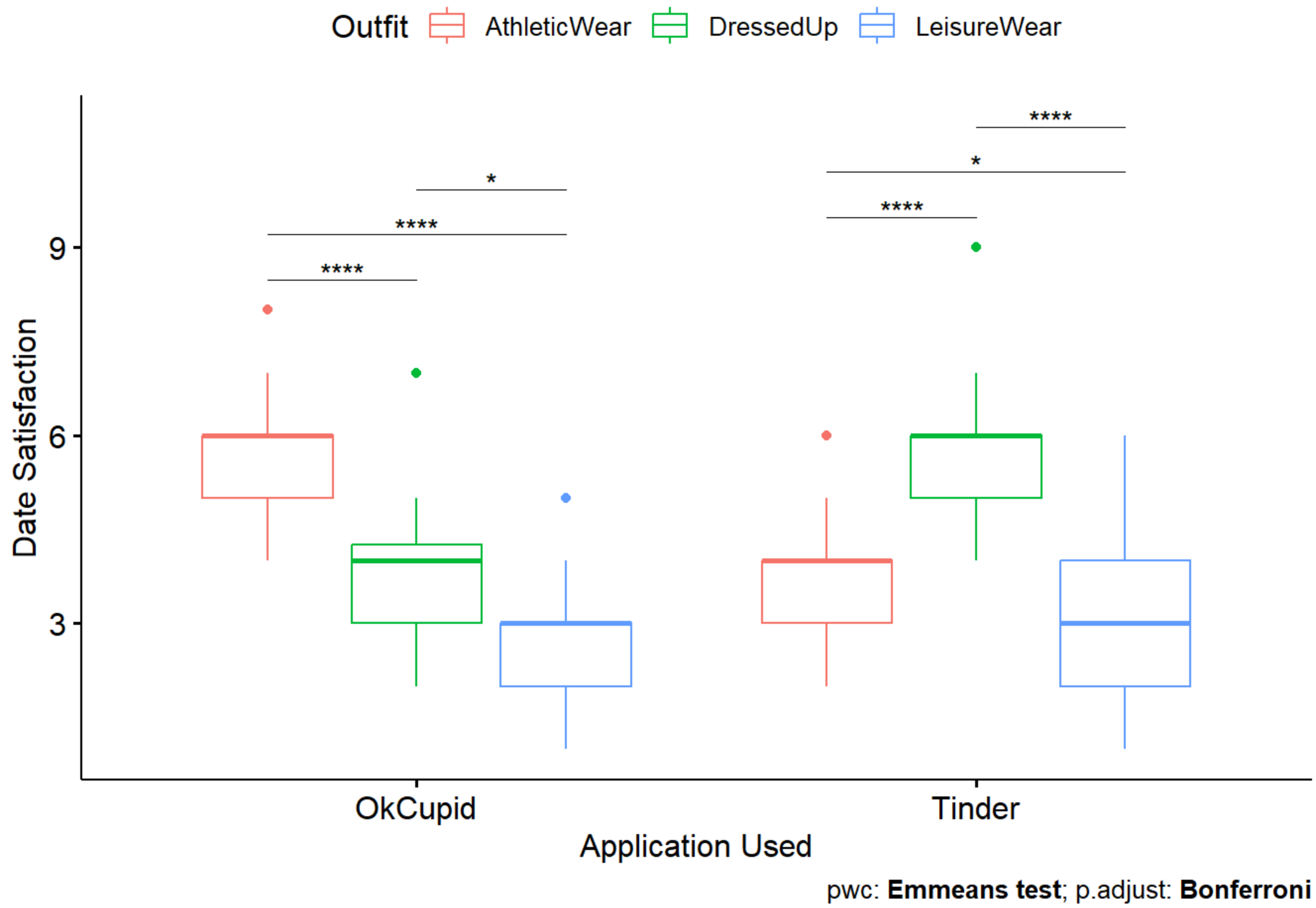
```
pwc2 <- dat_date %>%
  group_by(Outfit) %>%
  emmeans_test(Satisfaction~Application,
               p.adjust.method = "bonferroni") %>%
  add_xy_position(x = "Outfit")

pwc2
```

```
## # A tibble: 3 × 15
##   Outfit  term  .y.   group1 group2    df statistic       p   p.adj p.adj.signif
##   <fct>   <chr> <chr> <chr>  <chr>  <dbl>     <dbl>   <dbl>   <dbl> <chr>
## 1 Athlet… Appl… Sati… OkCup… Tinder   114     5.49  2.50e-7 2.50e-7 ****
## 2 Dresse… Appl… Sati… OkCup… Tinder   114    -5.92  3.46e-8 3.46e-8 ****
## 3 Leisur… Appl… Sati… OkCup… Tinder   114    -0.433 6.66e-1 6.66e-1 ns
## # ℹ 5 more variables: y.position <dbl>, groups <named list>, x <dbl>,
## #   xmin <dbl>, xmax <dbl>
```

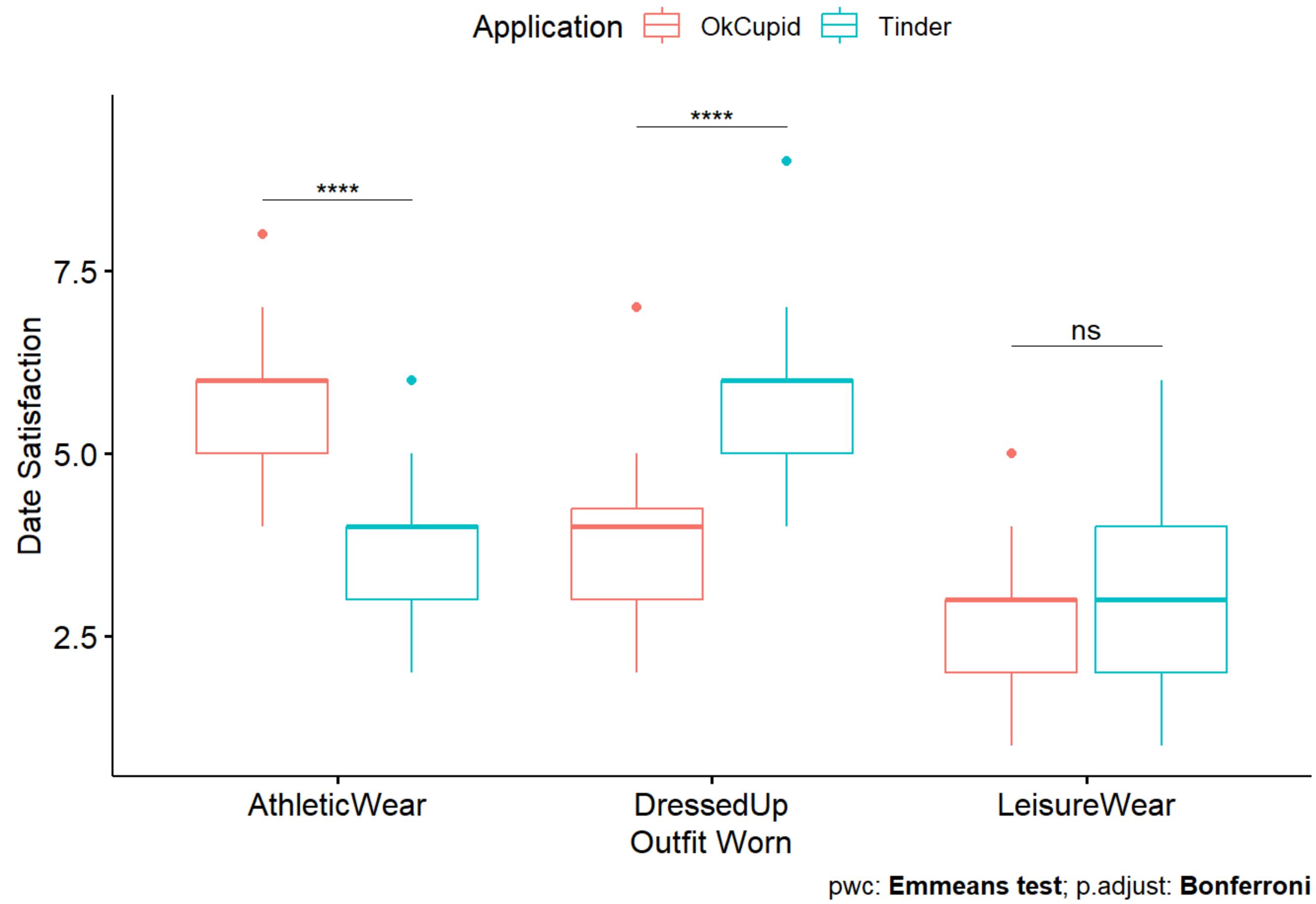
There were statistically significant differences in date satisfaction between all groups for both OKCupid and Tinder users (*p*s < .05 when adjusted using Bonferroni correction).

Factorial ANOVA | Professional ANOVA Visualization Code

```
ggboxplot(dat_date,
          x = "Application",
          y = "Satisfaction",
          color = "Outfit") +
  stat_pvalue_manual(pwc,
                     hide.ns = FALSE,
                     tip.length = 0) +
  labs(
    caption = get_pwc_label(pwc),
    y = "Date Satisfaction",
    x = "Application Used")
```




```
ggboxplot(dat_date,
          x = "Outfit",
          y = "Satisfaction",
          color = "Application") +
  stat_pvalue_manual(pwc2,
                     hide.ns = FALSE,
                     tip.length = 0) +
  labs(
    caption = get_pwc_label(pwc2),
    y = "Date Satisfaction",
    x = "Outfit Worn")
```



Factorial ANOVA | Saving the ANOVA Model Object

```
aovm3<-ANOVA(data = dat_date,
      dep = 'Satisfaction',
      factors = list('Application','Outfit'),
      effectSize = 'partEta',
      postHoc = ~Outfit,
      postHocCorr = 'bonf',
      postHocES = 'd',
      postHocEsCi = TRUE,
      emMeans = ~Application + Outfit + Application:Outfit,
      emmPlots = TRUE,
      emmPlotData = TRUE,
      emmTables = TRUE)$model
```

```
## NOTE: Results may be misleading due to involvement in interactions
```

Factorial ANOVA | APA Style ANOVA Tables

```
apa.aov.table(aovm3,
              table.number=5,
              conf.level=.95,
              filename = "FactorialANOVA.doc")
```

```
##
##
## Table 5
##
## ANOVA results using Satisfaction as the dependent variable
##
##
##             Predictor      SS  df      MS       F    p partial_eta2
##           (Intercept) 2116.80   1 2116.80 1765.29 .000
##           Application    0.30   1    0.30    0.25 .618          .00
##                Outfit   90.15   2   45.08   37.59 .000          .40
##  Application x Outfit   78.05   2   39.02   32.54 .000          .36
##                 Error  136.70 114    1.20
##  CI_95_partial_eta2
##
##          [.00, .05]
##          [.25, .50]
##          [.22, .47]
##
##
## Note: Values in square brackets indicate the bounds of the 95% confidence interval for partia
l eta-squared
```

```
#Generate Marginal Means Table
apa.2way.table(Outfit,
               Application,
               Satisfaction,
               dat_date,
               table.number=6,
               show.conf.interval = TRUE)
```

```
##
##
## Table 6
##
## Means and standard deviations for Satisfaction as a function of a 3(Outfit) X 2(Application)
design
##
##                          M     M_95%_CI   SD
##  Application:OkCupid
##               Outfit
##         AthleticWear 5.80 [5.31, 6.29] 1.06
##            DressedUp 3.75 [3.16, 4.34] 1.25
##          LeisureWear 2.90 [2.47, 3.33] 0.91
##
##   Application:Tinder
##               Outfit
##         AthleticWear 3.90 [3.42, 4.38] 1.02
##            DressedUp 5.80 [5.28, 6.32] 1.11
##          LeisureWear 3.05 [2.49, 3.61] 1.19
##
## Note. M and SD represent mean and standard deviation, respectively.
## LL and UL indicate the lower and upper limits of the
## 95% confidence interval for the mean, respectively.
## The confidence interval is a plausible range of population means
## that could have created a sample mean (Cumming, 2014).
```