Battery Example: Two-way Factorial Analysis

In this example, we consider 3 levels of temperature (temp) and 3 levels of material (mat) for a total of 9 treatment combinations. There are n = 4 reps per treatment combination for a total of 36 observations. The response variable is battery life (in hours).

```
library(dplyr)
library(ggplot2)
library(car)
library(emmeans)
Battery <- read.csv("C:/hess/STAT512/RNotes/ExpDesign2/ED2_Battery.csv")
str(Battery)
```

```
## 'data.frame':    36 obs. of  4 variables:
##  $ temp: int  1 1 1 1 1 1 1 1 1 1 ...
##  $ mat : int  1 1 1 1 2 2 2 2 3 3 ...
##  $ k   : int  1 2 3 4 1 2 3 4 1 2 ...
##  $ life: int  130 155 74 180 150 188 159 126 138 110 ...
```

```
#Important: Need to define Temp and Mat as.factors!!!!
Battery$temp<-as.factor(Battery$temp)
Battery$mat<-as.factor(Battery$mat)
str(Battery)
```

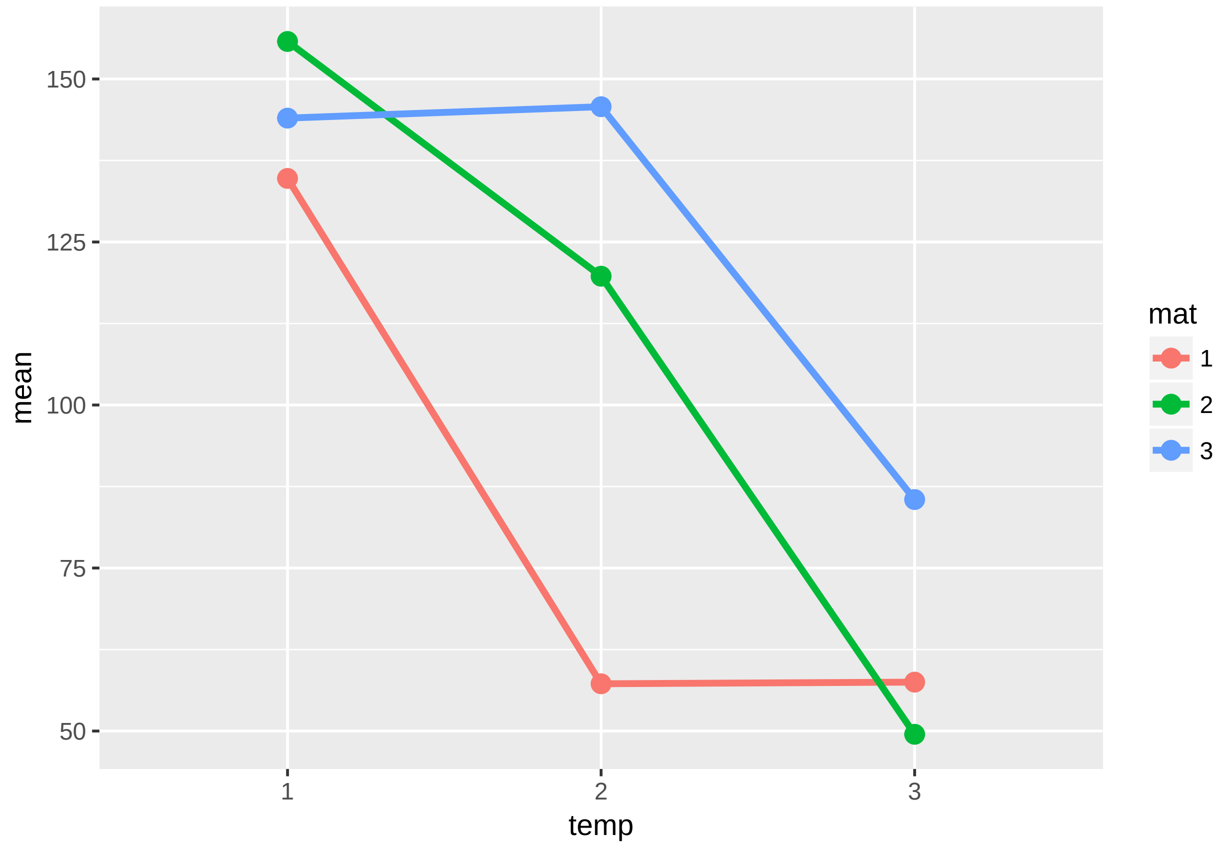
```
## 'data.frame':    36 obs. of  4 variables:
##  $ temp: Factor w/ 3 levels "1","2","3": 1 1 1 1 1 1 1 1 1 1 ...
##  $ mat : Factor w/ 3 levels "1","2","3": 1 1 1 1 2 2 2 2 3 3 ...
##  $ k   : int  1 2 3 4 1 2 3 4 1 2 ...
##  $ life: int  130 155 74 180 150 188 159 126 138 110 ...
```

Summary Statistics and Graphs

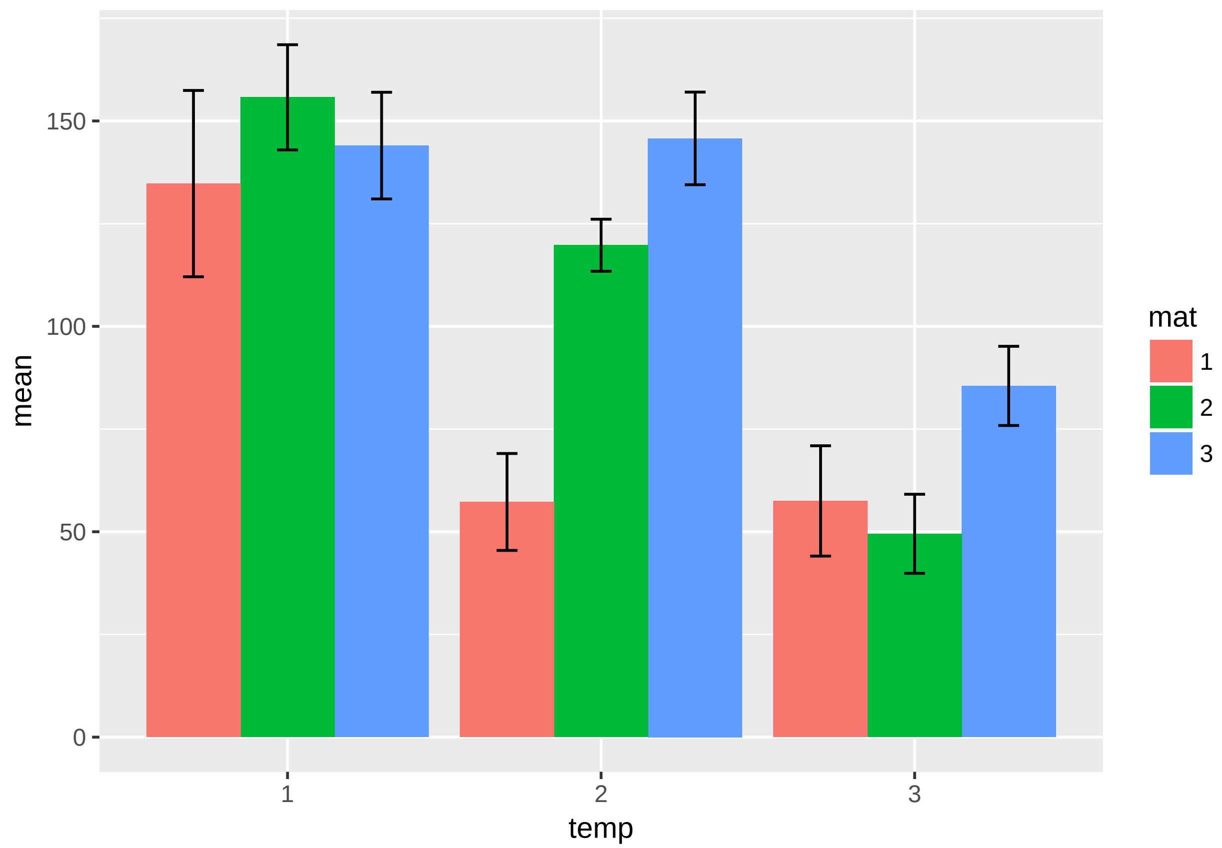
```
SumStats <- summarize(group_by(Battery, temp, mat),
                n = n(),
                mean = mean(life),
                sd   = sd(life),
                SE   = sd/sqrt(n))
SumStats
```

```
## # A tibble: 9 x 6
## # Groups: temp [?]
##   temp   mat        n  mean    sd    SE
##   <fctr> <fctr> <int> <dbl> <dbl> <dbl>
## 1 1      1          4 135    45.4 22.7
## 2 1      2          4 156    25.6 12.8
## 3 1      3          4 144    26.0 13.0
## 4 2      1          4  57.2  23.6 11.8
## 5 2      2          4 120    12.7  6.33
## 6 2      3          4 146    22.5 11.3
## 7 3      1          4  57.5  26.9 13.4
## 8 3      2          4  49.5  19.3  9.63
## 9 3      3          4  85.5  19.3  9.64
```

```r
#Line Plot
p <- qplot(x = temp, y = mean, colour = mat, group = mat, data = SumStats)
p + geom_line(size=1.2) + geom_point(size=3)
```




```r
#Grouped Bar Plot with SE bars
ggplot(SumStats, aes(x=temp, y=mean, fill=mat)) +
    geom_bar(position=position_dodge(), stat="identity") +
    geom_errorbar(aes(ymin=mean-SE, ymax=mean+SE),
                  width=.2,  # Width of the error bars
                  position=position_dodge(.9))
```



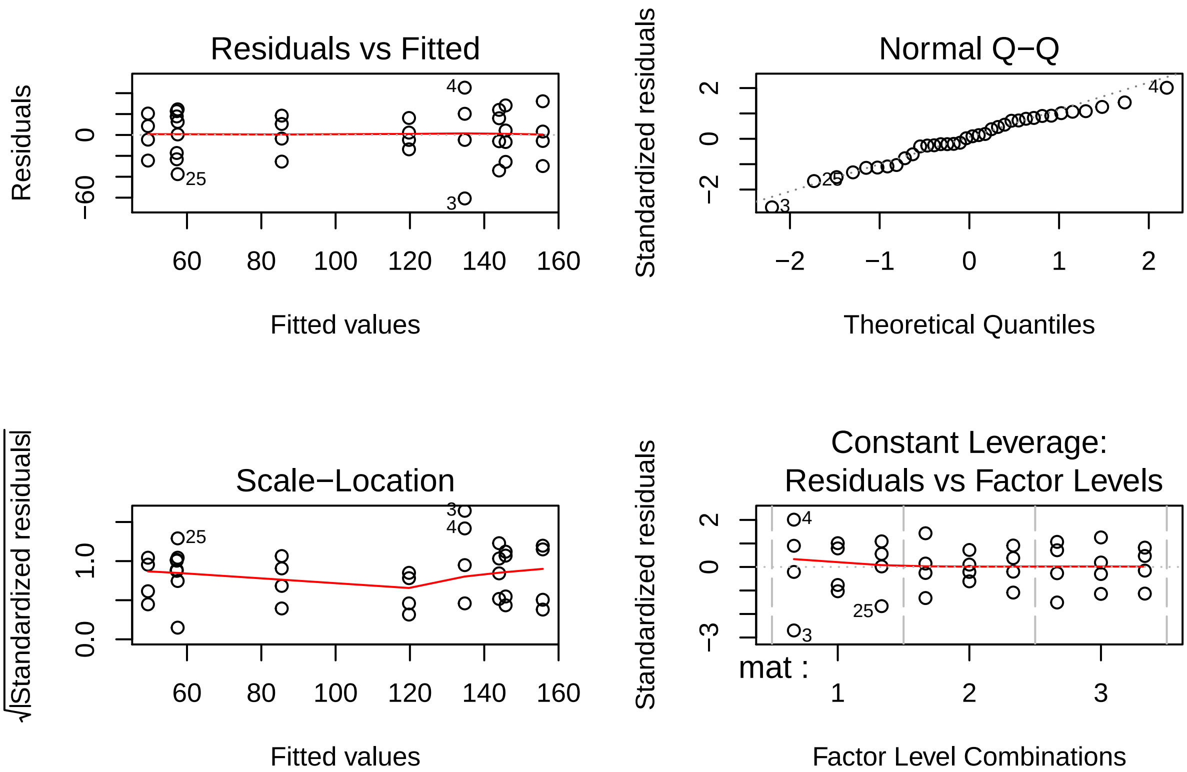
Two-way model

Typical research questions are addressed using Type 3 tests (using Anova() from the car package) and pairwise comparisons (using lsmeans() from the lsmeans package). Important: Change contrasts options to get meaningful Type 3 tests!

```
options(contrasts=c("contr.sum", "contr.poly"))
Model1 <- lm(life ~ mat*temp, data = Battery)
#Equivalent to lm(life ~ mat + temp + mat:temp, data = Battery)
Anova(Model1, type = 3)
```

```
## Anova Table (Type III tests)
##
## Response: life
##             Sum Sq Df  F value    Pr(>F)
## (Intercept) 400900  1 593.7386 < 2.2e-16 ***
## mat          10684  2   7.9114  0.001976 **
## temp         39119  2  28.9677 1.909e-07 ***
## mat:temp      9614  4   3.5595  0.018611 *
## Residuals    18231 27
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
par(mfrow=c(2, 2))
plot(Model1)
```



Pairwise Comparisons

Pairwise comparisons on materials at EACH level of Temp. This type of comparison is reasonable given the significant interaction.

```r
emmeans(Model1, pairwise ~ mat|temp)
```

```
## $emmeans
## temp = 1:
##  mat emmean       SE df  lower.CL  upper.CL
##  1   134.75 12.99243 27 108.09174 161.40826
##  2   155.75 12.99243 27 129.09174 182.40826
##  3   144.00 12.99243 27 117.34174 170.65826
##
## temp = 2:
##  mat emmean       SE df  lower.CL  upper.CL
##  1    57.25 12.99243 27  30.59174  83.90826
##  2   119.75 12.99243 27  93.09174 146.40826
##  3   145.75 12.99243 27 119.09174 172.40826
##
## temp = 3:
##  mat emmean       SE df  lower.CL  upper.CL
##  1    57.50 12.99243 27  30.84174  84.15826
##  2    49.50 12.99243 27  22.84174  76.15826
##  3    85.50 12.99243 27  58.84174 112.15826
```

```
##
## Confidence level used: 0.95
##
## $contrasts
## temp = 1:
##  contrast estimate       SE df t.ratio p.value
##  1 - 2      -21.00 18.37407 27  -1.143  0.4967
##  1 - 3       -9.25 18.37407 27  -0.503  0.8703
##  2 - 3       11.75 18.37407 27   0.639  0.7998
##
## temp = 2:
##  contrast estimate       SE df t.ratio p.value
##  1 - 2      -62.50 18.37407 27  -3.402  0.0058
##  1 - 3      -88.50 18.37407 27  -4.817  0.0001
##  2 - 3      -26.00 18.37407 27  -1.415  0.3475
##
## temp = 3:
##  contrast estimate       SE df t.ratio p.value
##  1 - 2        8.00 18.37407 27   0.435  0.9012
##  1 - 3      -28.00 18.37407 27  -1.524  0.2959
##  2 - 3      -36.00 18.37407 27  -1.959  0.1419
##
## P value adjustment: tukey method for comparing a family of 3 estimates
```

We can look at ALL pairwise comparisons, but many of these may not be of interest.
But we pay a price for running so many tests with Tukey adjustment.

```r
emmeans(Model1, pairwise ~ mat*temp)
```

```
## $emmeans
##  mat temp emmean       SE df  lower.CL  upper.CL
##  1   1    134.75 12.99243 27 108.09174 161.40826
##  2   1    155.75 12.99243 27 129.09174 182.40826
##  3   1    144.00 12.99243 27 117.34174 170.65826
##  1   2     57.25 12.99243 27  30.59174  83.90826
##  2   2    119.75 12.99243 27  93.09174 146.40826
##  3   2    145.75 12.99243 27 119.09174 172.40826
##  1   3     57.50 12.99243 27  30.84174  84.15826
##  2   3     49.50 12.99243 27  22.84174  76.15826
##  3   3     85.50 12.99243 27  58.84174 112.15826
##
## Confidence level used: 0.95
##
## $contrasts
##  contrast  estimate       SE df t.ratio p.value
##  1,1 - 2,1   -21.00 18.37407 27  -1.143  0.9616
##  1,1 - 3,1    -9.25 18.37407 27  -0.503  0.9999
##  1,1 - 1,2    77.50 18.37407 27   4.218  0.0065
##  1,1 - 2,2    15.00 18.37407 27   0.816  0.9953
##  1,1 - 3,2   -11.00 18.37407 27  -0.599  0.9995
##  1,1 - 1,3    77.25 18.37407 27   4.204  0.0067
##  1,1 - 2,3    85.25 18.37407 27   4.640  0.0022
##  1,1 - 3,3    49.25 18.37407 27   2.680  0.2017
##  2,1 - 3,1    11.75 18.37407 27   0.639  0.9991
##  2,1 - 1,2    98.50 18.37407 27   5.361  0.0003
```

```
##  2,1 - 2,2    36.00 18.37407 27   1.959  0.5819
##  2,1 - 3,2    10.00 18.37407 27   0.544  0.9997
##  2,1 - 1,3    98.25 18.37407 27   5.347  0.0004
##  2,1 - 2,3   106.25 18.37407 27   5.783  0.0001
##  2,1 - 3,3    70.25 18.37407 27   3.823  0.0172
##  3,1 - 1,2    86.75 18.37407 27   4.721  0.0018
##  3,1 - 2,2    24.25 18.37407 27   1.320  0.9165
##  3,1 - 3,2    -1.75 18.37407 27  -0.095  1.0000
##  3,1 - 1,3    86.50 18.37407 27   4.708  0.0019
##  3,1 - 2,3    94.50 18.37407 27   5.143  0.0006
##  3,1 - 3,3    58.50 18.37407 27   3.184  0.0743
##  1,2 - 2,2   -62.50 18.37407 27  -3.402  0.0460
##  1,2 - 3,2   -88.50 18.37407 27  -4.817  0.0014
##  1,2 - 1,3    -0.25 18.37407 27  -0.014  1.0000
##  1,2 - 2,3     7.75 18.37407 27   0.422  1.0000
##  1,2 - 3,3   -28.25 18.37407 27  -1.537  0.8282
##  2,2 - 3,2   -26.00 18.37407 27  -1.415  0.8823
##  2,2 - 1,3    62.25 18.37407 27   3.388  0.0475
##  2,2 - 2,3    70.25 18.37407 27   3.823  0.0172
##  2,2 - 3,3    34.25 18.37407 27   1.864  0.6420
##  3,2 - 1,3    88.25 18.37407 27   4.803  0.0015
##  3,2 - 2,3    96.25 18.37407 27   5.238  0.0005
##  3,2 - 3,3    60.25 18.37407 27   3.279  0.0604
##  1,3 - 2,3     8.00 18.37407 27   0.435  1.0000
##  1,3 - 3,3   -28.00 18.37407 27  -1.524  0.8347
##  2,3 - 3,3   -36.00 18.37407 27  -1.959  0.5819
##
## P value adjustment: tukey method for comparing a family of 9 estimates
```

Pairwise comparisons for main effect of material primarily for illustration. Probably not of interest due to signficant interaction. Note: The warning from emmeans ("Results may be misleading due to involvement in interactions") will be displayed when considering comparisons of main effects in any model that includes an interaction.

```r
emmeans(Model1, pairwise ~ mat)
```

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

## $emmeans
##  mat    emmean       SE df  lower.CL  upper.CL
##  1    83.16667 7.501183 27  67.77551  98.55782
##  2   108.33333 7.501183 27  92.94218 123.72449
##  3   125.08333 7.501183 27 109.69218 140.47449
##
## Results are averaged over the levels of: temp
## Confidence level used: 0.95
##
## $contrasts
##  contrast  estimate       SE df t.ratio p.value
##  1 - 2    -25.16667 10.60827 27  -2.372  0.0628
##  1 - 3    -41.91667 10.60827 27  -3.951  0.0014
##  2 - 3    -16.75000 10.60827 27  -1.579  0.2718
##
## Results are averaged over the levels of: temp
## P value adjustment: tukey method for comparing a family of 3 estimates
```