Between-subjects Designs2

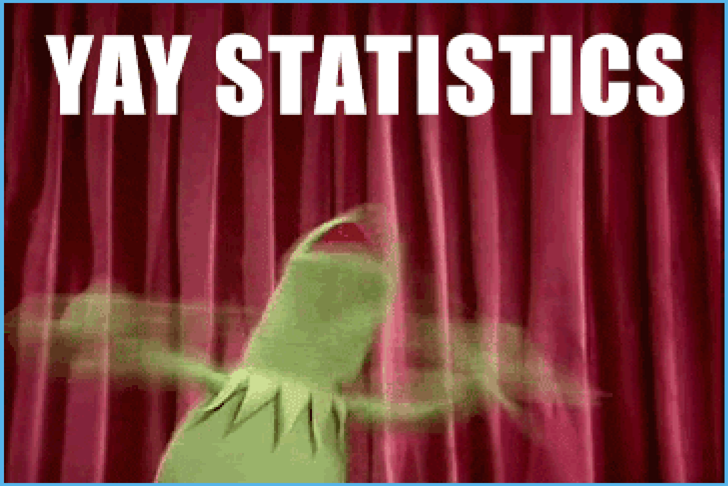
PSY 4433

Jonathan A. Pedroza, PhD

Cal Poly Pomona

2022-02-23

Analyses - Between-Subjects Designs



Group Project Analyses

- For your project, you'll have
    - two IVs
    - both IVs should have at least two conditions
    - one IV should be manipulated, other can another variable that literature states is important
    - continuous DV
    - 1 to 3 hypotheses
    - IV1 --> DV main effect
    - IV2 --> DV main effect
    - IV1*IV2 --> DV Interaction

Example

```
##       groups    male_groups     female_groups
## 1 treatment male_treatment female_treatment
## 2   control   male_control   female_control
```

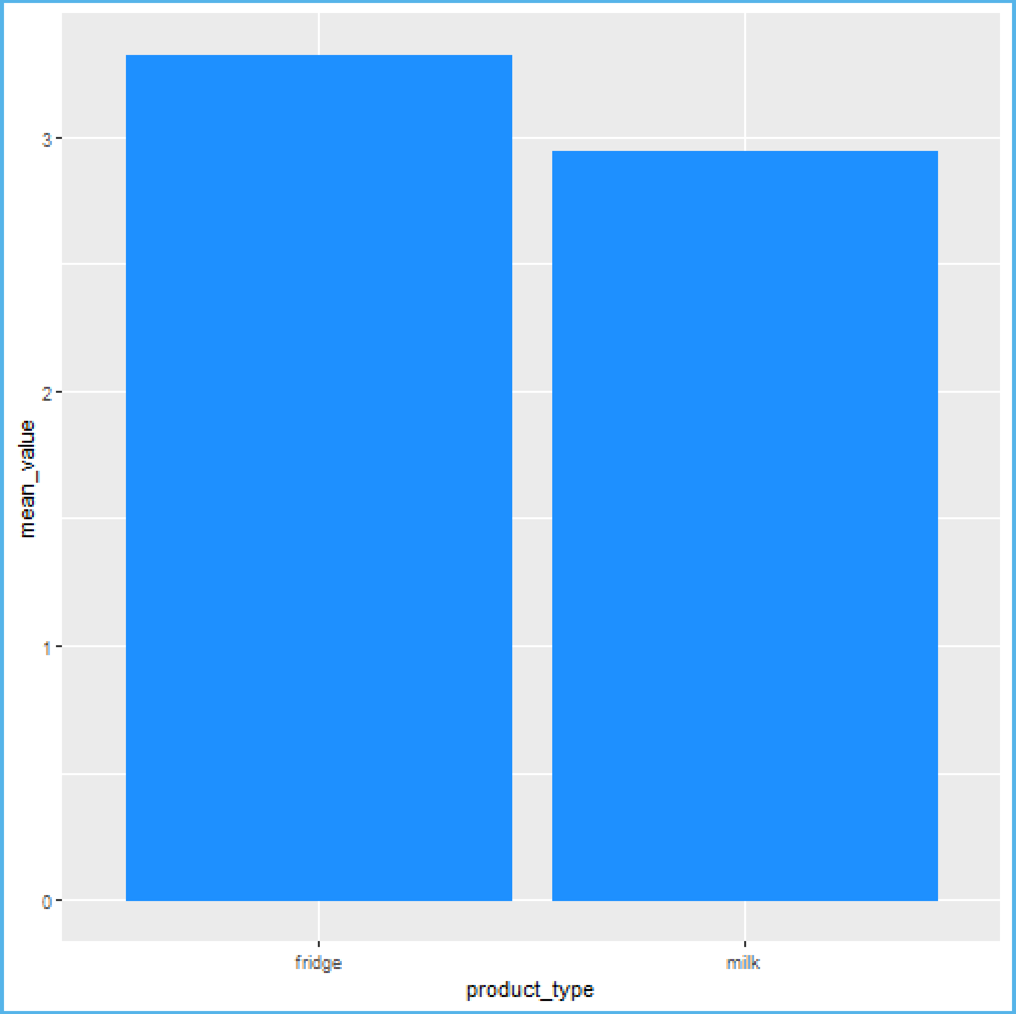
Two-Group Mean Difference

- the simplest experiment is when you have two groups that you are comparing
  - we will take this one step further in your project
- the design is often called a **single-factor two-group design** or a **two-group design**
- advantage is its simplicity
  - comparing the means of two different groups (Control and Treatment Groups)
- disadvantage is the limited amount of information you have and get from the comparison
  - your treatment group is better at DV than control group
  - if you have more than 2 groups, then this is not very helpful because you would have to compare each group
  - problematic because of false positives

Two-Group Mean Difference

- can use either a independent-samples t-test or One-way ANOVA

```
##
##      Two Sample t-test
##
## data:  condition_values by product_type
## t = 2.7166, df = 198, p-value = 0.00718
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.1026262 0.6462427
## sample estimates:
## mean in group fridge   mean in group milk
##             3.318034             2.943600
```


mean_value
3
2
1
0
fridge
milk
product_type

Comparing Means for More Than Two Groups

- comparing more than two groups
  - **single-factor multiple-group design**
- uses only a one-way ANOVA
  - determine differences in means of the outcome for each group/condition
- problem is whether you have a hypothesized group that will do differently than a group specifically (control)
  - or you just test everything (this is bad)

Comparing Means for More Than Two Groups

```
##                 Df Sum Sq Mean Sq F value   Pr(>F)    
## as.factor(cyl)  2  824.8   412.4    39.7 4.98e-09 ***
## Residuals      29  301.3    10.4                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

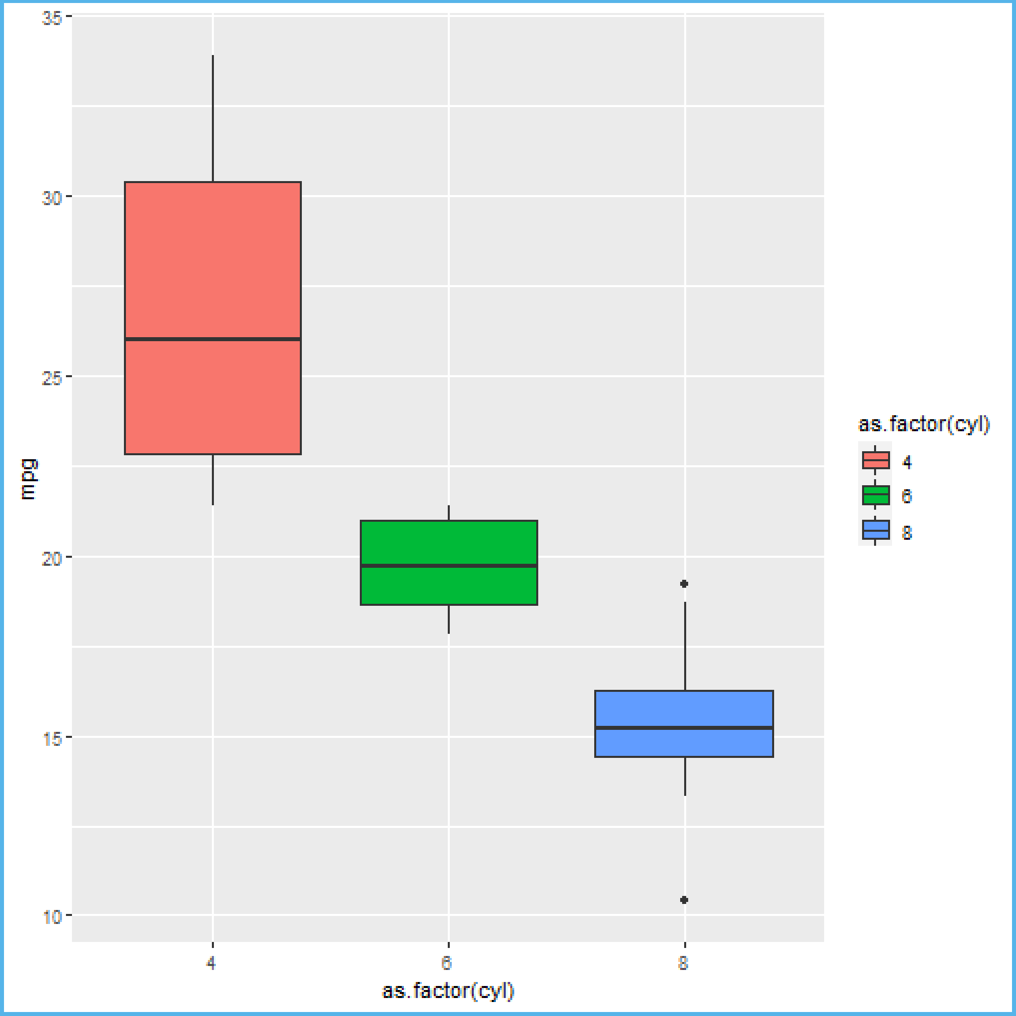
```
## Anova Table (Type II tests)
##
## Response: mpg
##                 Sum Sq Df F value    Pr(>F)
## as.factor(cyl) 824.78  2  39.697 4.979e-09 ***
## Residuals      301.26 29
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Anova Table (Type III tests)
##
## Response: mpg
##                 Sum Sq Df F value    Pr(>F)
## (Intercept)     7820.4  1 752.808 < 2.2e-16 ***
## as.factor(cyl)   824.8  2  39.697 4.979e-09 ***
## Residuals        301.3 29
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = mpg ~ as.factor(cyl), data = mtcars)
##
## $`as.factor(cyl)`
##           diff        lwr        upr     p adj
## 6-4  -6.920779 -10.769350 -3.0722086 0.0003424
## 8-4 -11.563636 -14.770779 -8.3564942 0.0000000
## 8-6  -4.642857  -8.327583 -0.9581313 0.0112287
```

```
##
##      Pairwise comparisons using t tests with pooled SD
##
## data:  mtcars$mpg and as.factor(mtcars$cyl)
##
##   4       6
## 6 0.00036 -
## 8 2.6e-09 0.01246
##
## P value adjustment method: bonferroni
```

```
##
## Call:
## lm(formula = mpg ~ as.factor(cyl), data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.2636 -1.8357  0.0286  1.3893  7.2364
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      26.6636     0.9718  27.437  < 2e-16 ***
## as.factor(cyl)6  -6.9208     1.5583  -4.441 0.000119 ***
## as.factor(cyl)8 -11.5636     1.2986  -8.905 8.57e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.223 on 29 degrees of freedom
## Multiple R-squared:  0.7325,    Adjusted R-squared:  0.714
## F-statistic:  39.7 on 2 and 29 DF,  p-value: 4.979e-09
```


35
30
25
20
15
10
mpg
4
6
8
as.factor(cyl)
as.factor(cyl)
4
6
8

Comparing Proportions for Two+ Groups

- if your DV is nominal or ordinal (which is won't be), then you would look at the proportions within each cell
    - cannot use t-test or ANOVA for this
- use chi-square tests of independence
    - non-parametric test
- JP Note: I rarely use this, but it can be helpful for certain scenarios (check the book)

Comparing the Means for Groups Based on Another Variable

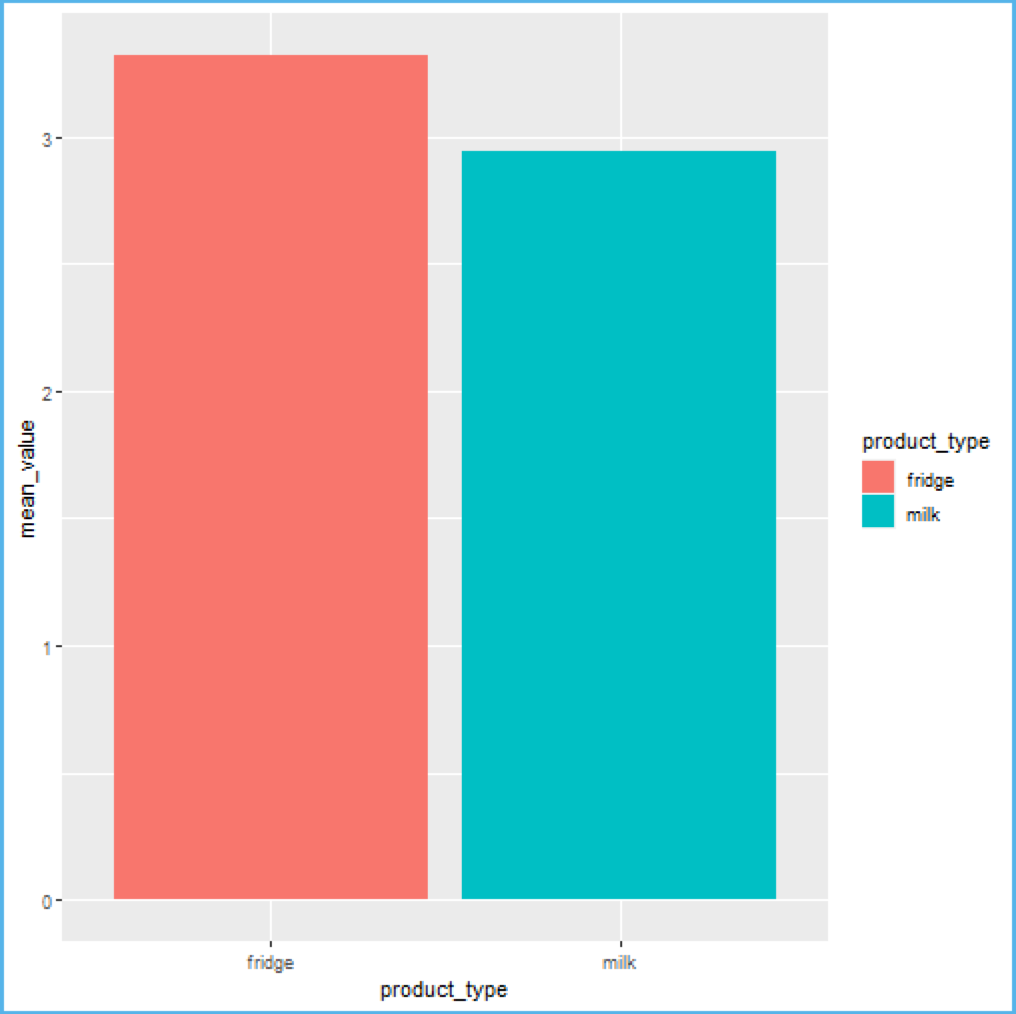
```
##
## Call:
## lm(formula = condition_values ~ product_type + celebrity, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.39917 -0.75900 -0.03952  0.73386  2.29412
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)        3.4334     0.1188  28.896  < 2e-16 ***
## product_typemilk  -0.3744     0.1372  -2.729  0.00693 **
## celebrityCeleb 2  -0.2307     0.1372  -1.682  0.09424 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9702 on 197 degrees of freedom
## Multiple R-squared:  0.04957,    Adjusted R-squared:  0.03993
## F-statistic: 5.138 on 2 and 197 DF,  p-value: 0.006682
```

```
##
## Call:
## lm(formula = condition_values ~ product_type * celebrity, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.2019 -0.6555  0.0011  0.6218  2.3751
##
## Coefficients:
##                                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)                        3.23614    0.13463  24.037  < 2e-16 ***
## product_typemilk                   0.02006    0.19040   0.105  0.91620
## celebrityCeleb 2                   0.16378    0.19040   0.860  0.39073
## product_typemilk:celebrityCeleb 2 -0.78899    0.26927  -2.930  0.00379 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.952 on 196 degrees of freedom
## Multiple R-squared:  0.08946,    Adjusted R-squared:  0.07552
## F-statistic: 6.419 on 3 and 196 DF,  p-value: 0.0003611
```

```
##                          Df Sum Sq Mean Sq F value  Pr(>F)
## product_type              1   7.01   7.010   7.735 0.00595 **
## celebrity                 1   2.66   2.661   2.937 0.08817 .
## product_type:celebrity    1   7.78   7.781   8.586 0.00379 **
## Residuals               196 177.64   0.906
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Anova Table (Type II tests)
##
## Response: condition_values
##                            Sum Sq  Df F value   Pr(>F)
## product_type                7.010   1  7.7348 0.005945 **
## celebrity                   2.661   1  2.9366 0.088175 .
## product_type:celebrity      7.781   1  8.5858 0.003790 **
## Residuals                 177.636 196
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Anova Table (Type III tests)
##
## Response: condition_values
##                           Sum Sq  Df  F value  Pr(>F)
## (Intercept)               523.63   1 577.7654 < 2e-16 ***
## product_type                0.01   1   0.0111 0.91620
## celebrity                   0.67   1   0.7399 0.39073
## product_type:celebrity      7.78   1   8.5858 0.00379 **
## Residuals                 177.64 196
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


mean_value
3
2
1
0
fridge
milk
product_type
product_type
fridge
milk

```
## `summarise()` has grouped output by 'product_type'. You can override using t
```