# EDP308: STATISTICAL LITERACY

The University of Texas at Austin, Fall 2020
RAZ: Rebecca A. Zárate, MA

# Overview

- From t to F
- The F Distribution
  - Degrees of Freedom
    - N-k
    - k-1
- Sum of Squares Refresher
- ANOVA Table Output
- Hypothesis Testing for ANOVA
- Pairwise Comparisons
- ANOVA in R
  - Data

# From t to F, ANOVA

- t-tests have t-statistics

I know… I am sorry.

- ANOVA has an F-statistic
  - The larger the F-statistic, the more likely you are to reject the null hypothesis
- The idea behind the hypothesis test is still the same, figure out your cut off score and determine if the statistic you compute is in the critical region
  - Critical values are now based on two types of degrees of freedom
    - Degrees of Freedom for Sample Size (Within), N-k
      - N = total sample size
    - Degrees of Freedom for Number of Groups (Between), k-1
      - k = number of groups

# Degrees of Freedom

- Degrees of Freedom for Sample Size, N-k
  - This is the degrees of freedom related to WITHIN, the denominator
    - $n_1 + n_2 + n_3 \ldots n_k = N$
    - $df_W$ = N (total sample size) – k (the number of groups)
- Degrees of Freedom for Number of Groups, k-1
  - This is the degrees of freedom related to BETWEEN, the numerator
    - $df_B$ = k (number of groups) – 1

# The F Distribution

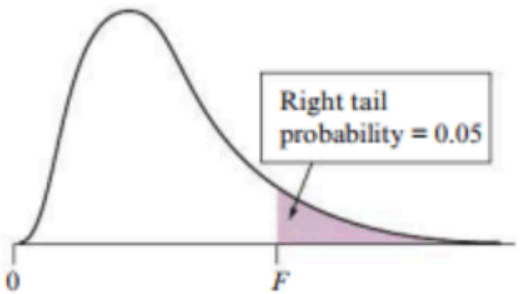



**Table D** *F* **Distribution for Values of Right-Tail Probability = 0.05**

Number of Groups (k-1)

Number of People (N-k)

| $df_2$ | $df_1$: 1 | 2 | 3 | 4 | 5 | 6 | 8 | 12 | 24 | ∞ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 161.45 | 199.50 | 215.71 | 224.58 | 230.16 | 233.99 | 238.88 | 243.91 | 249.05 | 254.31 |
| 2 | 18.51 | 19.00 | 19.16 | 19.25 | 19.30 | 19.33 | 19.37 | 19.41 | 19.45 | 19.50 |
| 3 | 10.13 | 9.55 | 9.28 | 9.12 | 9.01 | 8.94 | 8.85 | 8.74 | 8.64 | 8.53 |
| 4 | 7.71 | 6.94 | 6.59 | 6.39 | 6.26 | 6.16 | 6.04 | 5.91 | 5.77 | 5.63 |
| 5 | 6.61 | 5.79 | 5.41 | 5.19 | 5.05 | 4.95 | 4.82 | 4.68 | 4.53 | 4.37 |
| 6 | 5.99 | 5.14 | 4.76 | 4.53 | 4.39 | 4.28 | 4.15 | 4.00 | 3.84 | 3.67 |
| 7 | 5.59 | 4.74 | 4.35 | 4.12 | 3.97 | 3.87 | 3.73 | 3.57 | 3.41 | 3.23 |
| 8 | 5.32 | 4.46 | 4.07 | 3.84 | 3.69 | 3.58 | 3.44 | 3.28 | 3.12 | 2.93 |
| 9 | 5.12 | 4.26 | 3.86 | 3.63 | 3.48 | 3.37 | 3.23 | 3.07 | 2.90 | 2.71 |
| 10 | 4.96 | 4.10 | 3.71 | 3.48 | 3.33 | 3.22 | 3.07 | 2.91 | 2.74 | 2.54 |

In this table:

$df_1$ refers to $df_B$

$df_2$ refers to $df_W$.

To find a critical value for $\alpha = .05$, simply locate the F value for the appropriate degrees of freedom.

What is the critical F value for:
N = 10
k = 3

# The F Distribution

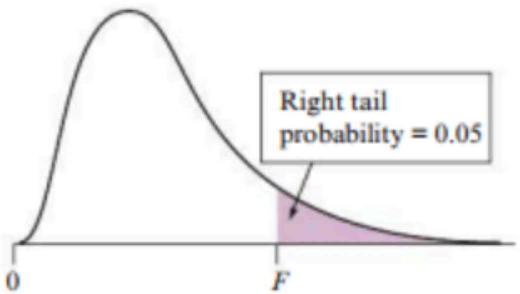



**Table D** *F* Distribution for Values of Right-Tail Probability = 0.05

Number of Groups (k-1)

Number of People (N-k)

| $df_2$ \ $df_1$ | 1 | 2 | 3 | 4 | 5 | 6 | 8 | 12 | 24 | ∞ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 161.45 | 199.50 | 215.71 | 224.58 | 230.16 | 233.99 | 238.88 | 243.91 | 249.05 | 254.31 |
| 2 | 18.51 | 19.00 | 19.16 | 19.25 | 19.30 | 19.33 | 19.37 | 19.41 | 19.45 | 19.50 |
| 3 | 10.13 | 9.55 | 9.28 | 9.12 | 9.01 | 8.94 | 8.85 | 8.74 | 8.64 | 8.53 |
| 4 | 7.71 | 6.94 | 6.59 | 6.39 | 6.26 | 6.16 | 6.04 | 5.91 | 5.77 | 5.63 |
| 5 | 6.61 | 5.79 | 5.41 | 5.19 | 5.05 | 4.95 | 4.82 | 4.68 | 4.53 | 4.37 |
| 6 | 5.99 | 5.14 | 4.76 | 4.53 | 4.39 | 4.28 | 4.15 | 4.00 | 3.84 | 3.67 |
| 7 | 5.59 | 4.74 | 4.35 | 4.12 | 3.97 | 3.87 | 3.73 | 3.57 | 3.41 | 3.23 |
| 8 | 5.32 | 4.46 | 4.07 | 3.84 | 3.69 | 3.58 | 3.44 | 3.28 | 3.12 | 2.93 |
| 9 | 5.12 | 4.26 | 3.86 | 3.63 | 3.48 | 3.37 | 3.23 | 3.07 | 2.90 | 2.71 |
| 10 | 4.96 | 4.10 | 3.71 | 3.48 | 3.33 | 3.22 | 3.07 | 2.91 | 2.74 | 2.54 |

What is the critical F value for:

$N = 10$

$k = 3$

$df_B = 3\text{-}1 = 2$

$df_W = 10\text{-}3 = 7$

ANOVA always uses a one-tailed test formulation, because the F-statistic can only be positive

# Sum of Squares Refresher…

| X (sample unit) | - μ (sample mean) = | Deviation | Deviation Squared |
|---|---|---|---|
| 8 - | 8 = | $0^2$ | 0 |
| 7 - | 8 = | $-1^2$ | 1 |
| 5 - | 8 = | $-3^2$ | 9 |
| 6 - | 8 = | $-2^2$ | 4 |
| 10 - | 8 = | $2^2$ | 4 |
| 9 - | 8 = | $1^2$ | 1 |
| 7 - | 8 = | $-1^2$ | 1 |
| 9 - | 8 = | $1^2$ | 1 |
| 8 - | 8 = | $0^2$ | 0 |
| 11 - | 8 = | $3^2$ | 9 |
| Σ = 80 | $\bar{x}$ =8 | Σ = | Σ = 30 |

"Squares"

Sum the "Squares"

# Sum of Squares Refresher…

- Remember, sum of squares (SS) is just another measure of variability.
- Variance is like an average of the Sum of Squares.

SS
(Sum of Squares)

$$\frac{\text{SS (Sum of Squares)}}{\text{df}} = \text{VARIANCE} \quad s^2$$

$$\sqrt{\text{VARIANCE}} = \text{STANDARD DEVATION} \quad s$$

# ANOVA Table

- This is how an ANOVA table would look if you ran an F-test

This is just VARIANCE!

(I apologize on behalf of the statistics naming committee…)

This is the ratio of variance we are really interested in.

| Source of Variability | Sum of Squares | Degrees of Freedom | Mean Square | F Statistic |
|---|---|---|---|---|
| Group (Between) | $SSB$ Sum of Squares Between | $df_B = k - 1$ | $MSB = \frac{SSB}{df_B}$ | $F = \frac{MSB}{MSW}$ |
| Error (Within) | $SSW$ Sum of Squares Within | $df_W = N - k$ | $MSW = \frac{SSW}{df_W}$ | |
| Total | $SST = \text{SSB} + \text{SSW}$ | $df_T = df_B + df_W$ | | |

(I will not make you calculate SS by hand.)

# Back to ANOVA Studying Example

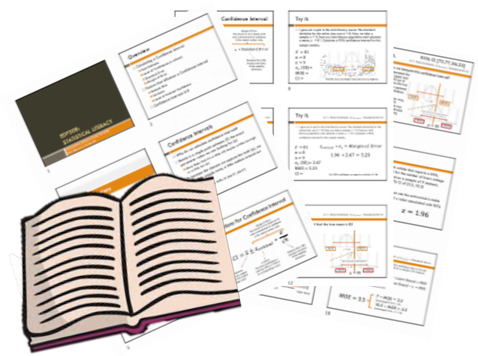
Population 1

Sample 1
(Book Only)
$n_1 = 10$
$\bar{x}_1 = 7.60$

Population 2

Sample 2
(Lecture Notes)
$n_2 = 10$
$\bar{x}_2 = 17.60$

Population 3



Sample 3
(Lecture Notes and Book)
$n_3 = 10$
$\bar{x}_3 = 19.30$

# ANOVA Studying Example

Run an ANOVA to test whether any of the true average scores differs for the three groups, using a significance level of $\alpha = .05$.

You are given the following information:

$$SSB = 799.27$$
$$SSW = 158.90$$

# Step 1: State the Hypotheses

Step 1:

$$H_0\!: \mu_{BookOnly} = \mu_{LectureNotes} = \mu_{Book+LectureNotes}$$

$$OR$$

$$H_0\!: The\ true\ mean\ score\ is\ the\ same$$
$$for\ each\ group.$$

$$H_1\!: \mu_{BookOnly} \neq \mu_{LectureNotes} \neq \mu_{Book+LectureNotes}$$

$$OR$$

$$H_1\!: At\ least\ one\ of\ the\ true\ group\ means\ is\ different.$$

# Step 2: Significant and Statistical Test

Step 2:

$$\alpha = .05$$

Step 3:

$$F = \frac{MSB}{MSW} = \frac{SSB/_{df_B}}{SSW/_{df_W}}$$

# Step 4: Find the Critical Value

Step 4:

$$\alpha = .05$$

$$k = 3 \text{ and } N = 30$$

$$df_B = 3 - 1 = 2$$

$$df_W = 30 - 3 = 27$$

$$F_{crit} = 3.35$$

ANOVA always uses a one-tailed test formulation, because the F-statistic can only be positive

**TABLE E** F critical values (continued)

| Degrees of freedom in the denominator | p | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Degrees of freedom in the numerator | | | | | | | | |
| 18 | .100 | 3.01 | 2.62 | 2.42 | 2.29 | 2.20 | 2.13 | 2.08 | 2.04 | 2.00 |
| | .050 | 4.41 | 3.55 | 3.16 | 2.93 | 2.77 | 2.66 | 2.58 | 2.51 | 2.46 |
| | .025 | 5.98 | 4.56 | 3.95 | 3.61 | 3.38 | 3.22 | 3.10 | 3.01 | 2.93 |
| | .010 | 8.29 | 6.01 | 5.09 | 4.58 | 4.25 | 4.01 | 3.84 | 3.71 | 3.60 |
| | .001 | 15.38 | 10.39 | 8.49 | 7.46 | 6.81 | 6.35 | 6.02 | 5.76 | 5.56 |
| 19 | .100 | 2.99 | 2.61 | 2.40 | 2.27 | 2.18 | 2.11 | 2.06 | 2.02 | 1.98 |
| | .050 | 4.38 | 3.52 | 3.13 | 2.90 | 2.74 | 2.63 | 2.54 | 2.48 | 2.42 |
| | .025 | 5.92 | 4.51 | 3.90 | 3.56 | 3.33 | 3.17 | 3.05 | 2.96 | 2.88 |
| | .010 | 8.18 | 5.93 | 5.01 | 4.50 | 4.17 | 3.94 | 3.77 | 3.63 | 3.52 |
| | .001 | 15.08 | 10.16 | 8.28 | 7.27 | 6.62 | 6.18 | 5.85 | 5.59 | 5.39 |
| 20 | .100 | 2.97 | 2.59 | 2.38 | 2.25 | 2.16 | 2.09 | 2.04 | 2.00 | 1.96 |
| | .050 | 4.35 | 3.49 | 3.10 | 2.87 | 2.71 | 2.60 | 2.51 | 2.45 | 2.39 |
| | .025 | 5.87 | 4.46 | 3.86 | 3.51 | 3.29 | 3.13 | 3.01 | 2.91 | 2.84 |
| | .010 | 8.10 | 5.85 | 4.94 | 4.43 | 4.10 | 3.87 | 3.70 | 3.56 | 3.46 |
| | .001 | 14.82 | 9.95 | 8.10 | 7.10 | 6.46 | 6.02 | 5.69 | 5.44 | 5.24 |
| 21 | .100 | 2.96 | 2.57 | 2.36 | 2.23 | 2.14 | 2.08 | 2.02 | 1.98 | 1.95 |
| | .050 | 4.32 | 3.47 | 3.07 | 2.84 | 2.68 | 2.57 | 2.49 | 2.42 | 2.37 |
| | .025 | 5.83 | 4.42 | 3.82 | 3.48 | 3.25 | 3.09 | 2.97 | 2.87 | 2.80 |
| | .010 | 8.02 | 5.78 | 4.87 | 4.37 | 4.04 | 3.81 | 3.64 | 3.51 | 3.40 |
| | .001 | 14.59 | 9.77 | 7.94 | 6.95 | 6.32 | 5.88 | 5.56 | 5.31 | 5.11 |
| 22 | .100 | 2.95 | 2.56 | 2.35 | 2.22 | 2.13 | 2.06 | 2.01 | 1.97 | 1.93 |
| | .050 | 4.30 | 3.44 | 3.05 | 2.82 | 2.66 | 2.55 | 2.46 | 2.40 | 2.34 |
| | .025 | 5.79 | 4.38 | 3.78 | 3.44 | 3.22 | 3.05 | 2.93 | 2.84 | 2.76 |
| | .010 | 7.95 | 5.72 | 4.82 | 4.31 | 3.99 | 3.76 | 3.59 | 3.45 | 3.35 |
| | .001 | 14.38 | 9.61 | 7.80 | 6.81 | 6.19 | 5.76 | 5.44 | 5.19 | 4.99 |
| 23 | .100 | 2.94 | 2.55 | 2.34 | 2.21 | 2.11 | 2.05 | 1.99 | 1.95 | 1.92 |
| | .050 | 4.28 | 3.42 | 3.03 | 2.80 | 2.64 | 2.53 | 2.44 | 2.37 | 2.32 |
| | .025 | 5.75 | 4.35 | 3.75 | 3.41 | 3.18 | 3.02 | 2.90 | 2.81 | 2.73 |
| | .010 | 7.88 | 5.66 | 4.76 | 4.26 | 3.94 | 3.71 | 3.54 | 3.41 | 3.30 |
| | .001 | 14.20 | 9.47 | 7.67 | 6.70 | 6.08 | 5.65 | 5.33 | 5.09 | 4.89 |
| 24 | .100 | 2.93 | 2.54 | 2.33 | 2.19 | 2.10 | 2.04 | 1.98 | 1.94 | 1.91 |
| | .050 | 4.26 | 3.40 | 3.01 | 2.78 | 2.62 | 2.51 | 2.42 | 2.36 | 2.30 |
| | .025 | 5.72 | 4.32 | 3.72 | 3.38 | 3.15 | 2.99 | 2.87 | 2.78 | 2.70 |
| | .010 | 7.82 | 5.61 | 4.72 | 4.22 | 3.90 | 3.67 | 3.50 | 3.36 | 3.26 |
| | .001 | 14.03 | 9.34 | 7.55 | 6.59 | 5.98 | 5.55 | 5.23 | 4.99 | 4.80 |
| 25 | .100 | 2.92 | 2.53 | 2.32 | 2.18 | 2.09 | 2.02 | 1.97 | 1.93 | 1.89 |
| | .050 | 4.24 | 3.39 | 2.99 | 2.76 | 2.60 | 2.49 | 2.40 | 2.34 | 2.28 |
| | .025 | 5.69 | 4.29 | 3.69 | 3.35 | 3.13 | 2.97 | 2.85 | 2.75 | 2.68 |
| | .010 | 7.77 | 5.57 | 4.68 | 4.18 | 3.85 | 3.63 | 3.46 | 3.32 | 3.22 |
| | .001 | 13.88 | 9.22 | 7.45 | 6.49 | 5.89 | 5.46 | 5.15 | 4.91 | 4.71 |
| 26 | .100 | 2.91 | 2.52 | 2.31 | 2.17 | 2.08 | 2.01 | 1.96 | 1.92 | 1.88 |
| | .050 | 4.23 | 3.37 | 2.98 | 2.74 | 2.59 | 2.47 | 2.39 | 2.32 | 2.27 |
| | .025 | 5.66 | 4.27 | 3.67 | 3.33 | 3.10 | 2.94 | 2.82 | 2.73 | 2.65 |
| | .010 | 7.72 | 5.53 | 4.64 | 4.14 | 3.82 | 3.59 | 3.42 | 3.29 | 3.18 |
| | .001 | 13.74 | 9.12 | 7.36 | 6.41 | 5.80 | 5.38 | 5.07 | 4.83 | 4.64 |
| 27 | .100 | 2.90 | [illegible] | 2.30 | 2.17 | 2.07 | 2.00 | 1.95 | 1.91 | 1.87 |
| | .050 | 4.21 | 3.35 | 2.96 | 2.73 | 2.57 | 2.46 | 2.37 | 2.31 | 2.25 |
| | .025 | 5.63 | [illegible] | 3.65 | 3.31 | 3.08 | 2.92 | 2.80 | 2.71 | 2.63 |
| | .010 | 7.68 | 5.49 | 4.60 | 4.11 | 3.78 | 3.56 | 3.39 | 3.26 | 3.15 |
| | .001 | 13.61 | 9.02 | 7.27 | 6.33 | 5.73 | 5.31 | 5.00 | 4.76 | 4.57 |

# Step 5: Calculate Test Statistic

- Let's fill in what we know…

| Source of Variability | Sum of Squares | Degrees of Freedom | Mean Square | F Statistic |
|---|---|---|---|---|
| Group (Between) | $SSB$ | $df_B = k - 1$ | $MSB = \frac{SSB}{df_B}$ | $F = \frac{MSB}{MSW}$ |
| Error (Within) | $SSW$ | $df_W = N - k$ | $MSW = \frac{SSW}{df_W}$ | |
| Total | $SST = \text{SSB} + \text{SSW}$ | $df_T = df_B + df_W$ | | |

| Source of Variability | Sum of Squares | Degrees of Freedom | Mean Square | F Statistic |
|---|---|---|---|---|
| Group (Between) | $SSB = 799.27$ | $df_B = 2$ | $MSB = \frac{799.27}{2}$ | $F = \frac{399.64}{5.89}$ |
| Error (Within) | $SSW = 158.9$ | $df_W = 27$ | $MSW = \frac{158.9}{27}$ | |
| Total | $SST = 958.17$ | $df_T = 29$ | | |

# Step 5: Calculate Test Statistic

| Source of Variability | Sum of Squares | Degrees of Freedom | Mean Square | F Statistic |
|---|---|---|---|---|
| Group (Between) | $SSB$ | $df_B = k - 1$ | $MSB = \frac{799.27}{2}$ | $F = \frac{399.64}{5.89}$ |
| Error (Within) | $SSW$ | $df_W = N - k$ | $MSW = \frac{158.9}{27}$ | |
| Total | $SST = \text{SSB} + \text{SSW}$ | $df_T = df_B + df_W$ | | |

| Source of Variability | Sum of Squares | Degrees of Freedom | Mean Square | F Statistic |
|---|---|---|---|---|
| Group (Between) | $SSB = 799.27$ | $df_B = 2$ | $MSB = 399.64$ | $F = 67.85$ |
| Error (Within) | $SSW = 158.9$ | $df_W = 27$ | $MSW = 5.89$ | |
| Total | $SST = 958.17$ | $df_T = 29$ | | |

# Source of Variability?

- What do you think we can be determine from this F-statistic, before we even compare to the critical value, knowing what you know about ratios?
  - Where is more of the variability coming from?

$$F = \frac{Between\ \ Group\ \ Variability}{Within\ \ Group\ \ Variability} = 67.85$$

| Source of Variability | Sum of Squares | Degrees of Freedom | Mean Square | F Statistic |
|---|---|---|---|---|
| Group (Between) | $SSB = 799.27$ | $df_B = 2$ | $MSB = 399.64$ | $F = 67.85$ |
| Error (Within) | $SSW = 158.9$ | $df_W = 27$ | $MSW = 5.89$ | |
| Total | $SST = 958.17$ | $df_T = 29$ | | |

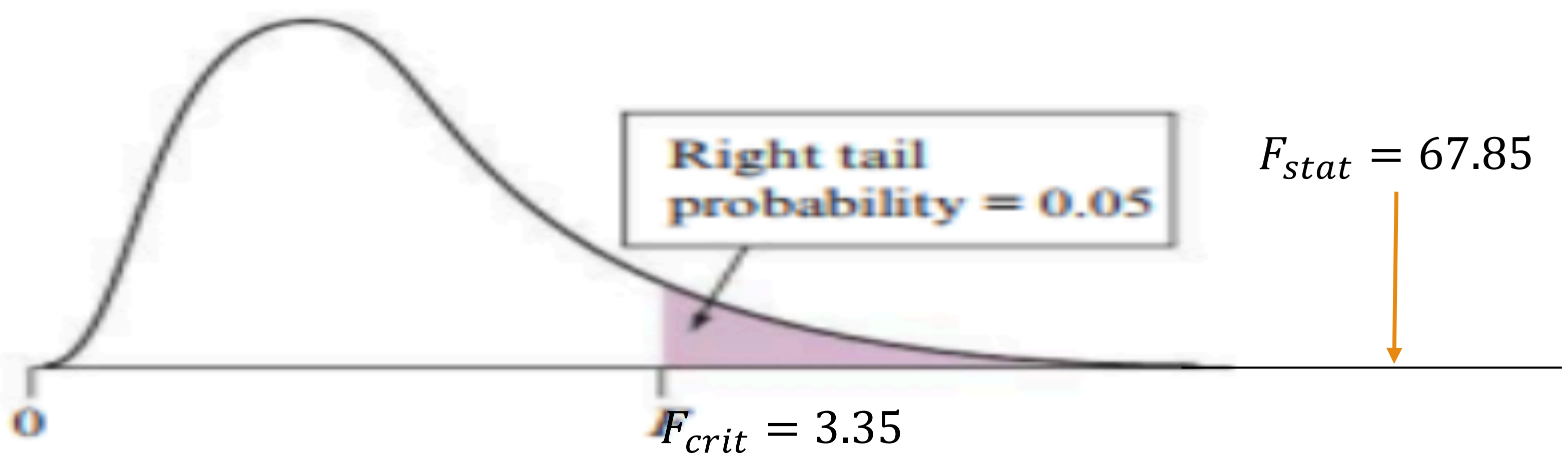
# Step 6: Draw Conclusions

Step 6:

Our $F_{stat} = 67.85$, and our $F_{crit} = 3.35$

Our $F_{stat}$ is past our $F_{crit}$, so we reject $H_0$

We reject $H_0$. There is enough evidence to reject the null hypothesis that the true mean score for each of the three groups is the same. Instead, we conclude that the true mean score is different for at least one of the groups.



# Significant F-Statistic, Then What?

- "We conclude that the true mean score <u>is different for at least one of the groups</u>"

Do we know which group(s)?

Book vs. Lecture Notes?

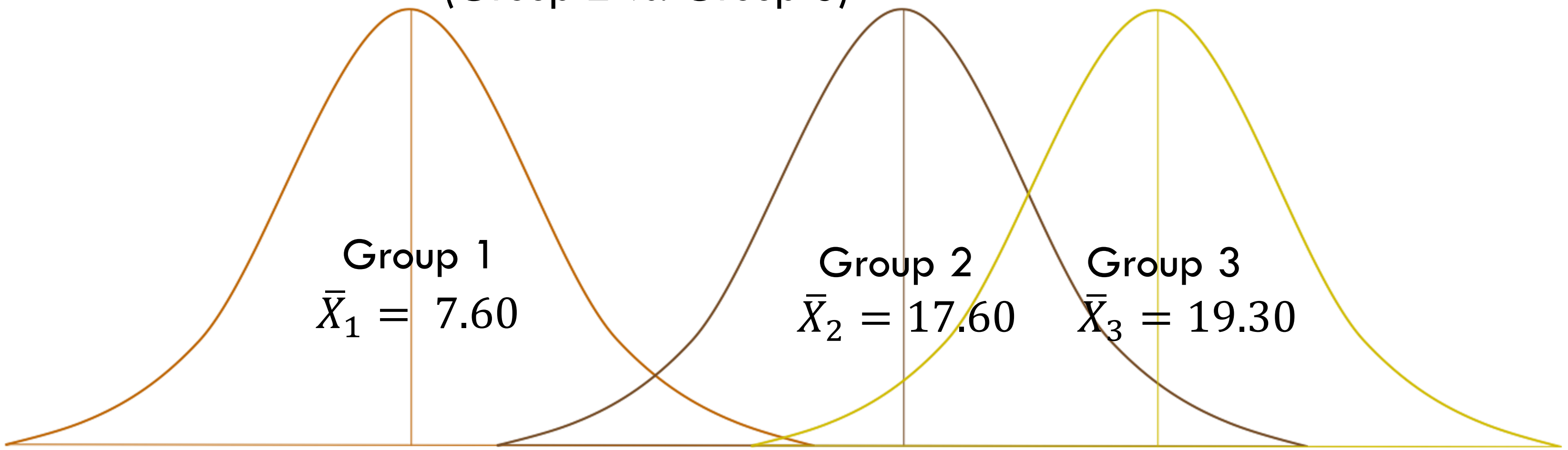
(Group 1 vs. Group 2)

Book vs. Book + Lecture Notes?

(Group 1 vs. Group 3)

Lecture Notes vs. Book + Lecture Notes?

(Group 2 vs. Group 3)



# Pairwise Comparisons

# Pairwise Comparisons

- Now we will consider each pair and see if there are significantly different from each other
- Various methods to do this, they differ in how conservative (harder to reach significant) or liberal (easier to reach significance) each test is
  - Tukey's HSD
  - Bonferroni

# Pairwise Comparisons

- What can we determine from this output?
  - * = significant

**Multiple Comparisons**

Dependent Variable: Test Score

| | (I) Group | (J) Group | Mean Difference (I–J) | Std. Error | Sig. | 95% Confidence Interval Lower Bound | 95% Confidence Interval Upper Bound |
|---|---|---|---|---|---|---|---|
| Tukey HSD | Book | Lecture | −10.00* | 1.085 | .000 | −12.69 | −7.31 |
| | | Lecture +Book | −11.70* | 1.085 | .000 | −14.39 | −9.01 |
| | Lecture | Book | 10.00* | 1.085 | .000 | 7.31 | 12.69 |
| | | Lecture +Book | −1.70 | 1.085 | .277 | −4.39 | .99 |
| | Lecture +Book | Book | 11.70* | 1.085 | .000 | 9.01 | 14.39 |
| | | Lecture | 1.70 | 1.085 | .277 | −.99 | 4.39 |

# Pairwise Comparisons

- There is a significant different between:
  - Book vs. Lecture Notes
  - Book vs. Book + Lecture Notes
- The difference between Lecture Notes vs. Lecture Notes + Book was NOT significant.

Notice how some 95% CI contains zero… These are not significant.

**Multiple Comparisons**

Dependent Variable: Test Score

| | (I) Group | (J) Group | Mean Difference (I–J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| | | | | | | Lower Bound | Upper Bound |
| Tukey HSD | Book | Lecture | −10.00* | 1.085 | .000 | −12.69 | −7.31 |
| | | Lecture +Book | −11.70* | 1.085 | .000 | −14.39 | −9.01 |
| | Lecture | Book | 10.00* | 1.085 | .000 | 7.31 | 12.69 |
| | | Lecture +Book | −1.70 | 1.085 | .277 | −4.39 | 0 .99 |
| | Lecture +Book | Book | 11.70* | 1.085 | .000 | 9.01 | 14.39 |
| | | Lecture | 1.70 | 1.085 | .277 | −.99 | 0 4.39 |

# Step 6: Draw Conclusions (Again)

- The overall F-test was significant, indicating that at least two of the different types of study material groups were significantly different with the scores they got on an exam ($F = 67.8, p < .05$).
  - APA: $F(2, 27) = 67.8, p < .05$
- The Book Only group had significantly lower mean score than the Lecture Notes and Lecture Notes + Book. There was no difference in the exam scores between the Lecture Notes and Lecture Notes + Book.

# APA Notation for ANOVA

$$F(2, 27) = 67.8, p < .05)$$

Degree of Freedom Between

Degree of Freedom Within

F-statistic (the ratio you calculated)

APA Notation: F(df1, df2) = f-statistic, p < .05)

# Practice

# Television Viewing

New research suggests that watching television, especially medical shows, can result in more concern about personal health. Surveys were administered to 18 college students measuring their personal health concerns on a scale of 0-10. For the following data, students were grouped based on their television viewing habits.

Test whether the true personal health concerns are the same for the three groups, using $\alpha = .05$. You are given: $SSB = 36.99$ and $SSW = 49$.

| **Television Viewing Habits** | | |
|---|---|---|
| **Little to none** | **Moderate** | **Substantial** |
| 3 | 5 | 6 |
| 2 | 7 | 7 |
| 1 | 3 | 6 |
| 5 | 4 | 6 |
| 3 | 8 | 8 |
| 7 | 6 | 9 |
| $\bar{X}_1 = 3.5$ | $\bar{X}_2 = 5.5$ | $\bar{X}_3 = 7.0$ |

# Step 1: State the Hypotheses

Step 1:

$$H_0\text{: } \mu_{Little} = \mu_{Moderate} = \mu_{Substantial}$$

$$OR$$

$$H_0\text{: } The\ true\ mean\ level\ of\ personal\ health\ concerns\ is\ the\ same\ for\ each\ group.$$

$$H_0\text{: } \mu_{Little} \neq \mu_{Moderate} \neq \mu_{Substantial}$$

$$OR$$

$$H_1\text{: } At\ least\ one\ of\ the\ true\ group\ mean\ level\ of\ health\ concerns\ is\ different.$$

# Step 2: Significant and Statistical Test

Step 2:

$$\alpha = .05$$

Step 3: ANOVA F-test

$$F = \frac{MSB}{MSW} = \frac{SSB/_{df_B}}{SSW/_{df_W}}$$

# Step 4: Find the Critical Value

Step 4:

$$\alpha = .05$$

$$k = 3 \text{ and } N = 18$$

$$df_B = 3 - 1 = 2$$

$$df_W = 18 - 3 = 15$$

$$F_{crit} = 3.68$$

**TABLE E** F critical values (continued)

| Degrees of freedom in the denominator | p | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Degrees of freedom in the numerator | | | | | | | | |
| 8 | .100 | 3.46 | 3.11 | 2.92 | 2.81 | 2.73 | 2.67 | 2.62 | 2.59 | 2.56 |
| | .050 | 5.32 | 4.46 | 4.07 | 3.84 | 3.69 | 3.58 | 3.50 | 3.44 | 3.39 |
| | .025 | 7.57 | 6.06 | 5.42 | 5.05 | 4.82 | 4.65 | 4.53 | 4.43 | 4.36 |
| | .010 | 11.26 | 8.65 | 7.59 | 7.01 | 6.63 | 6.37 | 6.18 | 6.03 | 5.91 |
| | .001 | 25.41 | 18.49 | 15.83 | 14.39 | 13.48 | 12.86 | 12.40 | 12.05 | 11.77 |
| 9 | .100 | 3.36 | 3.01 | 2.81 | 2.69 | 2.61 | 2.55 | 2.51 | 2.47 | 2.44 |
| | .050 | 5.12 | 4.26 | 3.86 | 3.63 | 3.48 | 3.37 | 3.29 | 3.23 | 3.18 |
| | .025 | 7.21 | 5.71 | 5.08 | 4.72 | 4.48 | 4.32 | 4.20 | 4.10 | 4.03 |
| | .010 | 10.56 | 8.02 | 6.99 | 6.42 | 6.06 | 5.80 | 5.61 | 5.47 | 5.35 |
| | .001 | 22.86 | 16.39 | 13.90 | 12.56 | 11.71 | 11.13 | 10.70 | 10.37 | 10.11 |
| 10 | .100 | 3.29 | 2.92 | 2.73 | 2.61 | 2.52 | 2.46 | 2.41 | 2.38 | 2.35 |
| | .050 | 4.96 | 4.10 | 3.71 | 3.48 | 3.33 | 3.22 | 3.14 | 3.07 | 3.02 |
| | .025 | 6.94 | 5.46 | 4.83 | 4.47 | 4.24 | 4.07 | 3.95 | 3.85 | 3.78 |
| | .010 | 10.04 | 7.56 | 6.55 | 5.99 | 5.64 | 5.39 | 5.20 | 5.06 | 4.94 |
| | .001 | 21.04 | 14.91 | 12.55 | 11.28 | 10.48 | 9.93 | 9.52 | 9.20 | 8.96 |
| 11 | .100 | 3.23 | 2.86 | 2.66 | 2.54 | 2.45 | 2.39 | 2.34 | 2.30 | 2.27 |
| | .050 | 4.84 | 3.98 | 3.59 | 3.36 | 3.20 | 3.09 | 3.01 | 2.95 | 2.90 |
| | .025 | 6.72 | 5.26 | 4.63 | 4.28 | 4.04 | 3.88 | 3.76 | 3.66 | 3.59 |
| | .010 | 9.65 | 7.21 | 6.22 | 5.67 | 5.32 | 5.07 | 4.89 | 4.74 | 4.63 |
| | .001 | 19.69 | 13.81 | 11.56 | 10.35 | 9.58 | 9.05 | 8.66 | 8.35 | 8.12 |
| 12 | .100 | 3.18 | 2.81 | 2.61 | 2.48 | 2.39 | 2.33 | 2.28 | 2.24 | 2.21 |
| | .050 | 4.75 | 3.89 | 3.49 | 3.26 | 3.11 | 3.00 | 2.91 | 2.85 | 2.80 |
| | .025 | 6.55 | 5.10 | 4.47 | 4.12 | 3.89 | 3.73 | 3.61 | 3.51 | 3.44 |
| | .010 | 9.33 | 6.93 | 5.95 | 5.41 | 5.06 | 4.82 | 4.64 | 4.50 | 4.39 |
| | .001 | 18.64 | 12.97 | 10.80 | 9.63 | 8.89 | 8.38 | 8.00 | 7.71 | 7.48 |
| 13 | .100 | 3.14 | 2.76 | 2.56 | 2.43 | 2.35 | 2.28 | 2.23 | 2.20 | 2.16 |
| | .050 | 4.67 | 3.81 | 3.41 | 3.18 | 3.03 | 2.92 | 2.83 | 2.77 | 2.71 |
| | .025 | 6.41 | 4.97 | 4.35 | 4.00 | 3.77 | 3.60 | 3.48 | 3.39 | 3.31 |
| | .010 | 9.07 | 6.70 | 5.74 | 5.21 | 4.86 | 4.62 | 4.44 | 4.30 | 4.19 |
| | .001 | 17.82 | 12.31 | 10.21 | 9.07 | 8.35 | 7.86 | 7.49 | 7.21 | 6.98 |
| 14 | .100 | 3.10 | 2.73 | 2.52 | 2.39 | 2.31 | 2.24 | 2.19 | 2.15 | 2.12 |
| | .050 | 4.60 | 3.74 | 3.34 | 3.11 | 2.96 | 2.85 | 2.76 | 2.70 | 2.65 |
| | .025 | 6.30 | 4.86 | 4.24 | 3.89 | 3.66 | 3.50 | 3.38 | 3.29 | 3.21 |
| | .010 | 8.86 | 6.51 | 5.56 | 5.04 | 4.69 | 4.46 | 4.28 | 4.14 | 4.03 |
| | .001 | 17.14 | 11.78 | 9.73 | 8.62 | 7.92 | 7.44 | 7.08 | 6.80 | 6.58 |
| 15 | .100 | 3.07 | 2.70 | 2.49 | 2.36 | 2.27 | 2.21 | 2.16 | 2.12 | 2.09 |
| | .050 | 4.54 | 3.68 | 3.29 | 3.06 | 2.90 | 2.79 | 2.71 | 2.64 | 2.59 |
| | .025 | 6.20 | [illegible] | 4.15 | 3.80 | 3.58 | 3.41 | 3.29 | 3.20 | 3.12 |
| | .010 | 8.68 | 6.36 | 5.42 | 4.89 | 4.56 | 4.32 | 4.14 | 4.00 | 3.89 |
| | .001 | 16.59 | 11.34 | 9.34 | 8.25 | 7.57 | 7.09 | 6.74 | 6.47 | 6.26 |
| 16 | .100 | 3.05 | 2.67 | 2.46 | 2.33 | 2.24 | 2.18 | 2.13 | 2.09 | 2.06 |
| | .050 | 4.49 | 3.63 | 3.24 | 3.01 | 2.85 | 2.74 | 2.66 | 2.59 | 2.54 |
| | .025 | 6.12 | 4.69 | 4.08 | 3.73 | 3.50 | 3.34 | 3.22 | 3.12 | 3.05 |
| | .010 | 8.53 | 6.23 | 5.29 | 4.77 | 4.44 | 4.20 | 4.03 | 3.89 | 3.78 |
| | .001 | 16.12 | 10.97 | 9.01 | 7.94 | 7.27 | 6.80 | 6.46 | 6.19 | 5.98 |
| 17 | .100 | 3.03 | 2.64 | 2.44 | 2.31 | 2.22 | 2.15 | 2.10 | 2.06 | 2.03 |
| | .050 | 4.45 | 3.59 | 3.20 | 2.96 | 2.81 | 2.70 | 2.61 | 2.55 | 2.49 |
| | .025 | 6.04 | 4.62 | 4.01 | 3.66 | 3.44 | 3.28 | 3.16 | 3.06 | 2.98 |
| | .010 | 8.40 | 6.11 | 5.19 | 4.67 | 4.34 | 4.10 | 3.93 | 3.79 | 3.68 |
| | .001 | 15.72 | 10.66 | 8.73 | 7.68 | 7.02 | 6.56 | 6.22 | 5.96 | 5.75 |

# Step 5: Calculate Test Statistic

Step 5:

| Source of Variability | Sum of Squares | Degrees of Freedom | Mean Square | F Statistic |
|---|---|---|---|---|
| Group (Between) | $SSB = 36.99$ | $df_B = 2$ | $MSB = ?$ | $F = ?$ |
| Error (Within) | $SSW = 49$ | $df_W = 15$ | $MSW = ?$ | |
| Total | $SST = ?$ | $df_T = ?$ | | |

# Step 5: Calculate Test Statistic

Step 5:

| Source of Variability | Sum of Squares | Degrees of Freedom | Mean Square | F Statistic |
|---|---|---|---|---|
| Group (Between) | $SSB = 36.99$ | $df_B = 2$ | $MSB = 18.495$ | $F = 5.66$ |
| Error (Within) | $SSW = 49$ | $df_W = 15$ | $MSW = 3.267$ | |
| Total | $SST = 85.99$ | $df_T = 17$ | | |

# Step 6: Draw Conclusions

Step 6:

$$F_{stat} = 5.66$$
$$F_{crit} = 3.68$$

Our $F_{stat}$ is past our $F_{crit}$, so we reject $H_0$

APA Notation: $\mathrm{F}(2, 15) = 5.66, \mathrm{p} < .05)$

We reject $H_0$. There is enough evidence to reject the null hypothesis that the true mean level of personal health concerns for each of the three groups is the same. Instead, we conclude that the true mean level of personal health concerns is different for at least one of the groups.

We can proceed to do pairwise comparisons to check which of the groups actually differ from the others.

# Up Next

- We've spend a lot of time looking at the differences between things… We're going to switch gears and next look at the *associations* (relationships) between some variables.

Correlation

# ANOVA in R

# Bug Bites…

- Mosquitos are an annoying fact of life we must all live with. There are various bug sprays and products out there that claim to protect against bug bites, but which ones work best?
- Let's compare the effectiveness (as quantified by the number of bug bites) of six different bug bite prevention products.
- Natural Oils, Citronelle Candles, DEET Spray, Citronelle Spray, Picaridin, and Nothing (Control)

These are real data but made up the names.

# Bug Spray Descriptive Statistics

- Just by looking at the descriptive statistics, which products look like they might be significantly different from the others?
- Which looks like it might be the best?
  - Is there a best?

```
 Descriptive statistics by group
group: Natural Oils
       vars  n mean   sd median trimmed  mad min max range skew kurtosis   se
count     1 12 14.5 4.72     14    14.4 5.19   7  23    16 0.27    -1.13 1.36
spray*    2 12  1.0 0.00      1     1.0 0.00   1   1     0  NaN      NaN 0.00
------------------------------------------------------------------------------------
group: Citronelle Candle
       vars  n  mean   sd median trimmed  mad min max range  skew kurtosis   se
count     1 12 15.33 4.27   16.5    15.6 4.45   7  21    14 -0.35    -1.04 1.23
spray*    2 12  2.00 0.00    2.0     2.0 0.00   2   2     0   NaN      NaN 0.00
------------------------------------------------------------------------------------
group: DEET
       vars  n mean   sd median trimmed  mad min max range skew kurtosis   se
count     1 12 2.08 1.98    1.5     1.8 1.48   0   7     7 1.13     0.52 0.57
spray*    2 12 3.00 0.00    3.0     3.0 0.00   3   3     0  NaN      NaN 0.00
------------------------------------------------------------------------------------
group: Citronelle Spray
       vars  n mean  sd median trimmed  mad min max range skew kurtosis   se
count     1 12 4.92 2.5      5     4.5 1.48   2  12    10 1.68     2.56 0.72
spray*    2 12 4.00 0.0      4     4.0 0.00   4   4     0  NaN      NaN 0.00
------------------------------------------------------------------------------------
group: Picaridin
       vars  n mean   sd median trimmed  mad min max range skew kurtosis  se
count     1 12  3.5 1.73      3     3.5 2.22   1   6     5 0.05    -1.41 0.5
spray*    2 12  5.0 0.00      5     5.0 0.00   5   5     0  NaN      NaN 0.0
------------------------------------------------------------------------------------
group: Nothing
       vars  n  mean   sd median trimmed  mad min max range skew kurtosis   se
count     1 12 16.67 6.21     15    16.5 6.67   9  26    17 0.39    -1.56 1.79
spray*    2 12  6.00 0.00      6     6.0 0.00   6   6     0  NaN      NaN 0.00
```

# R ANOVA Output

- We can see our between group df = 5, so we have 6 groups we are comparing (k-1)
- We have within df = 66, so we have 72 observations, (N-k)
- Our F-value is 34.7 which comes from dividing the Mean Squares Between (533.8) by the Mean Squares Within (15.4).
- Our p-value is very significant.

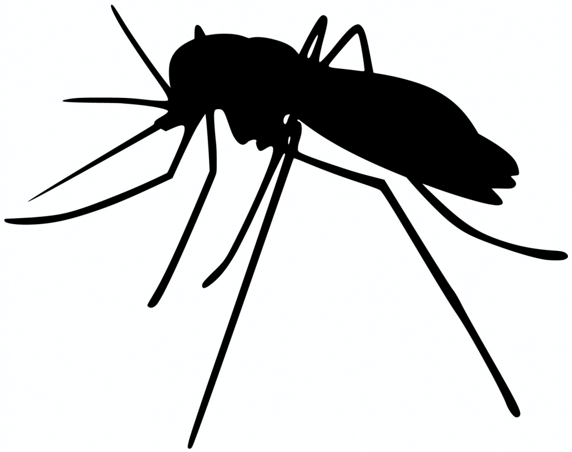
What is our conclusion?

What should we do next?

```
            Df Sum Sq Mean Sq F value Pr(>F)    
spray        5   2669   533.8    34.7 <2e-16 ***
Residuals   66   1015    15.4                   
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

# Significant F-value

- When we have a significant F-value, we know that *at least one* of the groups differs from the others, but our ANOVA does not tell us which groups…
- We need to conduct a pair-wise comparison like Tukey's HSD.



# Which groups differ significantly?

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = count ~ spray, data = bug_spray)

$spray
                                        diff        lwr       upr     p adj
Citronelle Candle-Natural Oils     0.8333333  -3.866075  5.532742 0.9951810
DEET-Natural Oils                -12.4166667 -17.116075 -7.717258 0.0000000
Citronelle Spray-Natural Oils     -9.5833333 -14.282742 -4.883925 0.0000014
Picaridin-Natural Oils           -11.0000000 -15.699409 -6.300591 0.0000000
Nothing-Natural Oils               2.1666667  -2.532742  6.866075 0.7542147
DEET-Citronelle Candle           -13.2500000 -17.949409 -8.550591 0.0000000
Citronelle Spray-Citronelle Candle -10.4166667 -15.116075 -5.717258 0.0000002
Picaridin-Citronelle Candle      -11.8333333 -16.532742 -7.133925 0.0000000
Nothing-Citronelle Candle          1.3333333  -3.366075  6.032742 0.9603075
Citronelle Spray-DEET              2.8333333  -1.866075  7.532742 0.4920707
Picaridin-DEET                     1.4166667  -3.282742  6.116075 0.9488669
Nothing-DEET                      14.5833333   9.883925 19.282742 0.0000000
Picaridin-Citronelle Spray        -1.4166667  -6.116075  3.282742 0.9488669
Nothing-Citronelle Spray          11.7500000   7.050591 16.449409 0.0000000
Nothing-Picaridin                 13.1666667   8.467258 17.866075 0.0000000
```

# Which groups differ significantly?

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = count ~ spray, data = bug_spray)

$spray
                                        diff        lwr       upr     p adj
Citronelle Candle-Natural Oils     0.8333333  -3.866075  5.532742 0.9951810
DEET-Natural Oils                -12.4166667 -17.116075 -7.717258 0.0000000
Citronelle Spray-Natural Oils     -9.5833333 -14.282742 -4.883925 0.0000014
Picaridin-Natural Oils           -11.0000000 -15.699409 -6.300591 0.0000000
Nothing-Natural Oils               2.1666667  -2.532742  6.866075 0.7542147
DEET-Citronelle Candle           -13.2500000 -17.949409 -8.550591 0.0000000
Citronelle Spray-Citronelle Candle -10.4166667 -15.116075 -5.717258 0.0000002
Picaridin-Citronelle Candle      -11.8333333 -16.532742 -7.133925 0.0000000
Nothing-Citronelle Candle          1.3333333  -3.366075  6.032742 0.9603075
Citronelle Spray-DEET              2.8333333  -1.866075  7.532742 0.4920707
Picaridin-DEET                     1.4166667  -3.282742  6.116075 0.9488669
Nothing-DEET                      14.5833333   9.883925 19.282742 0.0000000
Picaridin-Citronelle Spray        -1.4166667  -6.116075  3.282742 0.9488669
Nothing-Citronelle Spray          11.7500000   7.050591 16.449409 0.0000000
Nothing-Picaridin                 13.1666667   8.467258 17.866075 0.0000000
```

# Which groups differ significantly?

- The Less Effective Groups:
  - The Citronelle Candle, Natural Oils, and Control group do not differ significantly from each other.
- The More Effective Groups:
  - The DEET, Picaridin, and Citronelle Spray to significantly not differ from each other.
- The More Effective Groups all differ significantly from the Less Effective Group.
  - If you don't want bug bites, go for DEET, Picaridin, or Citronelle Spray.

These are real data but made up the names.

# R Code For ANOVA With Data

```
library(plyr)

# Loading in the dataset
bug_spray <- InsectSprays

# Renaming the groups, you do not have to do this
bug_spray$spray <- revalue(bug_spray$spray, c("A"= "Natural Oils", "B" = "Citronelle Candle", "C" =
"DEET", "D"= "Citronelle Spray", "E" = "Picaridin", "F" = "Nothing"))

# Get some descriptive statistics for the different groups
describeBy(bug_spray, group = "spray")

# This is the simple ANOVA function "aov()"
# The first input is the outcome variable, here it is "count" which is the number of bug bites
# The next input is the grouping variable, here is it the different types of bug "spray"
# Lastly, you tell R the name of the dataset
bug_bites_anova <- aov(count ~ spray, data = bug_spray)

# This gives you all the important summary information,
# this is also where you will see the F-statistic and signficance
summary(bug_bites_anova)

# Since our F-statistic was significant, we can look at the pairwise comparisons using "TuskeyHSD()"
function
TukeyHSD(bug_bites_anova)
```

Data Source: R Datasets "InsectSpray"

Beall, G., (1942) The Transformation of data from entomological field experiments,*Biometrika*, **29**, 243–262.

ANOVA Tutorial [https://bioinformatics-core-shared-training.github.io/linear-models-r/anova.html#section_2:_anova](https://bioinformatics-core-shared-training.github.io/linear-models-r/anova.html#section_2:_anova)