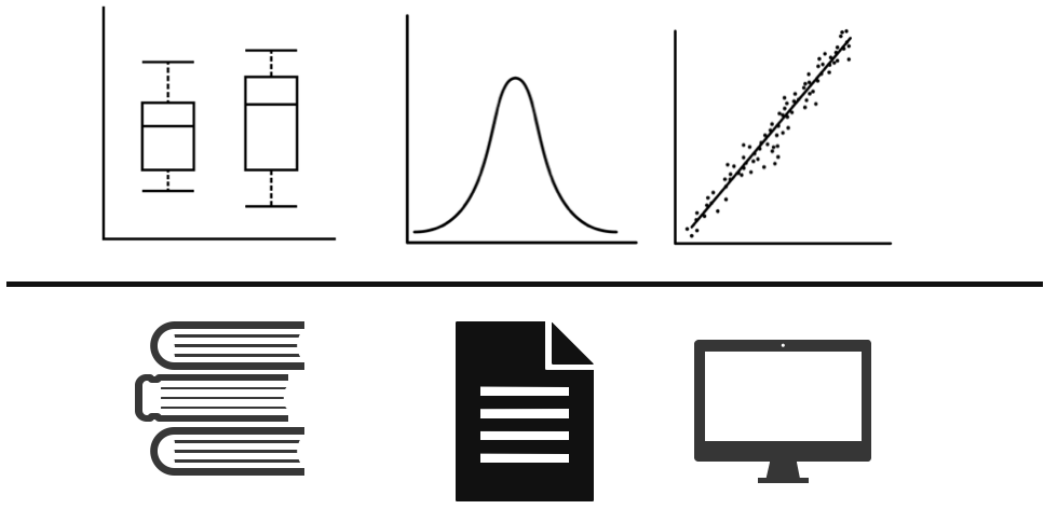
PSYC 2300

Introduction to Statistics



Lecture 13: Differences Between Many Factors

Outline for today

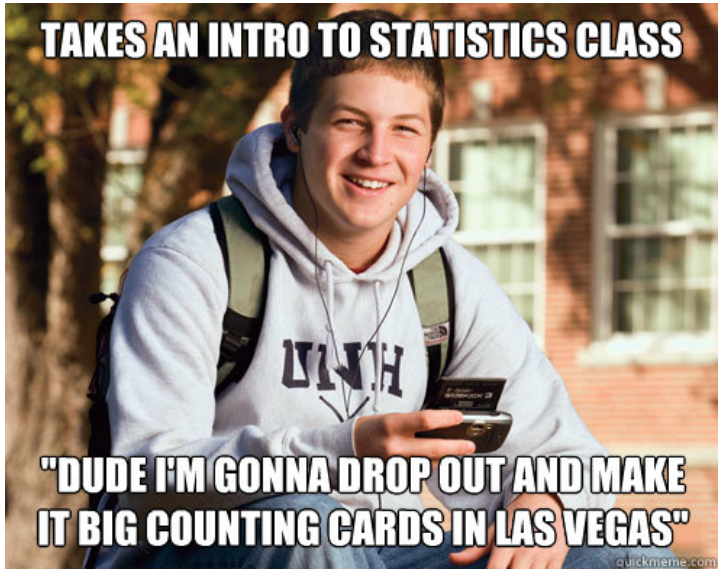
- More Extra Credit
- Application Project
- Review parts of last class
- Factorial ANOVA
- Factorial ANOVAs in JASP
    - Download `Stats Class 15 Dataset (Factorial ANOVA).jasp`
- One-Way ANOVA Practice Activity

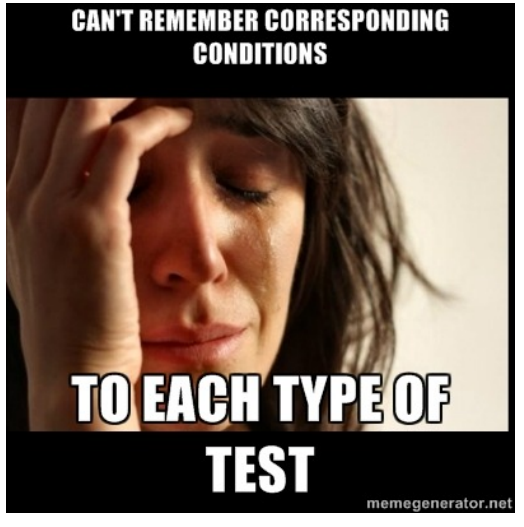
Extra Credit: Make a Meme

Merriam-Webster definition

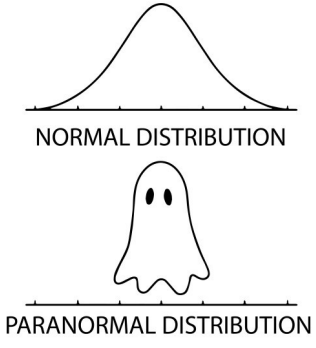
> **Meme**: An amusing or interesting item (such as a captioned picture or video) or genre of items that is spread widely online especially through social media

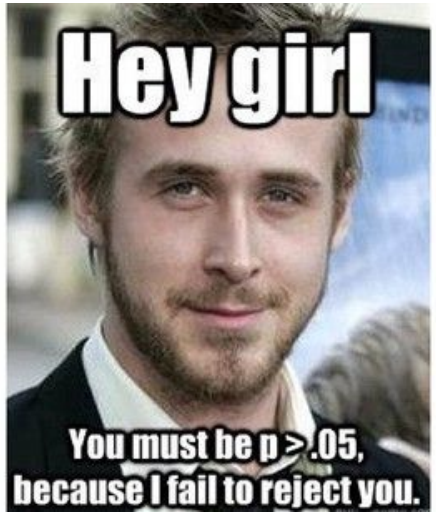
Up to 1% of extra credit for submitting a statistics-related meme on Canvas.

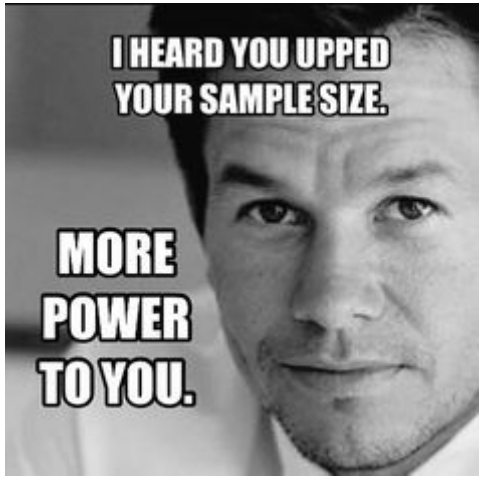


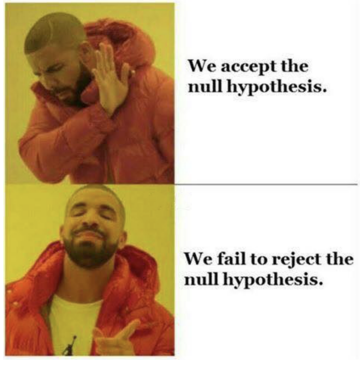


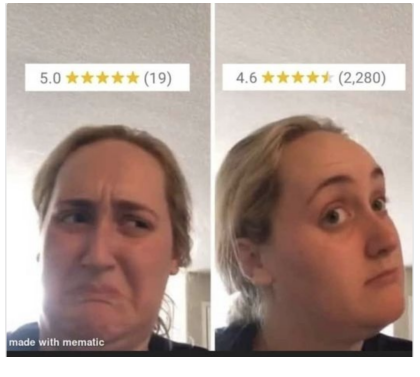
Extra Credit: Make a Meme

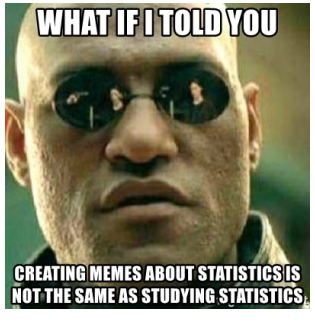












Schedule

Date	Week	Class	Topic	Reading	What's due
1/3/22	1	1	Introduction. Why study statistics in Psychology?	Syllabus, Ch.1	
1/5/22		2	Measures of central tendency and variability	Ch.2, Ch.3	
1/10/22	2	3	Visualizing data using graphs	Ch.4	
1/12/22		4	Correlations, Reliability, and Validity	Ch.5, Ch.6	
1/17/22	3	5	NO CLASS; MLK JR DAY		
1/19/22		6	Using hypotheses to test questions	Ch.7	Quiz 1 (ch.1-5)
1/24/22	4	7	Probability and normal curve	Ch.8	
1/26/22		8	Introduction to statistical significance	Ch.9	Quiz 2 (ch.6-8)
1/31/22	5	9	Midterm review		
2/2/22		10	Midterm Exam		Mini-report
2/7/22	6	11	Differences from the population	Ch.10	
2/9/22		12	Differences between two groups I	Ch.11	Quiz 3 (ch.9-10)
2/14/22	7	13	Difference between two groups II	Ch.12	
2/16/22		14	Difference between many groups	Ch.13	Quiz 4 (ch.11-12)
2/21/22	8	15	Difference between many factors	Ch.14	
2/23/22		16	Testing relationships using correlations	Ch.15	Quiz 5 (ch.13-14)
2/28/22	9	17	Making predictions using regression	Ch.16	
3/2/22		18	Analyzing data using JASP I		Quiz 6 (ch.15-16)
3/7/22	10	19	Analyzing data using JASP II		
3/9/22		20	Final Review		Application Project
3/14/22	11	21	Final Exam		

Application Project

Goal: Apply the information we've learned in class to a real dataset

You'll use JASP to:

- Compute descriptive statistics (e.g., mean and standard deviation)
- Create data visualizations
- Compute inferential statistics (e.g., *t*-tests, ANOVAs, correlation)
- Interpret the statistics that you compute
- You are free and *encouraged* to work in groups

Application Project

Step 1: Download the required dataset on Canvas

- You'll have 2-3 options to choose from. Each will focus on a different research question.

Step 2: Read about the dataset

- I'll provide any necessary information (e.g., what each variable means).

Step 3 Respond to the questions

- For each question, you will need to (1) run the requested analysis in JASP, (2) label that analysis in JASP, and (3) add notes in JASP to indicate key values and interpret those values when requested.

Step 4: Submit your Application Project

- Save and export your results to `.pdf` or `.html` file and submit on Canvas

Review

ANOVA: Terminology

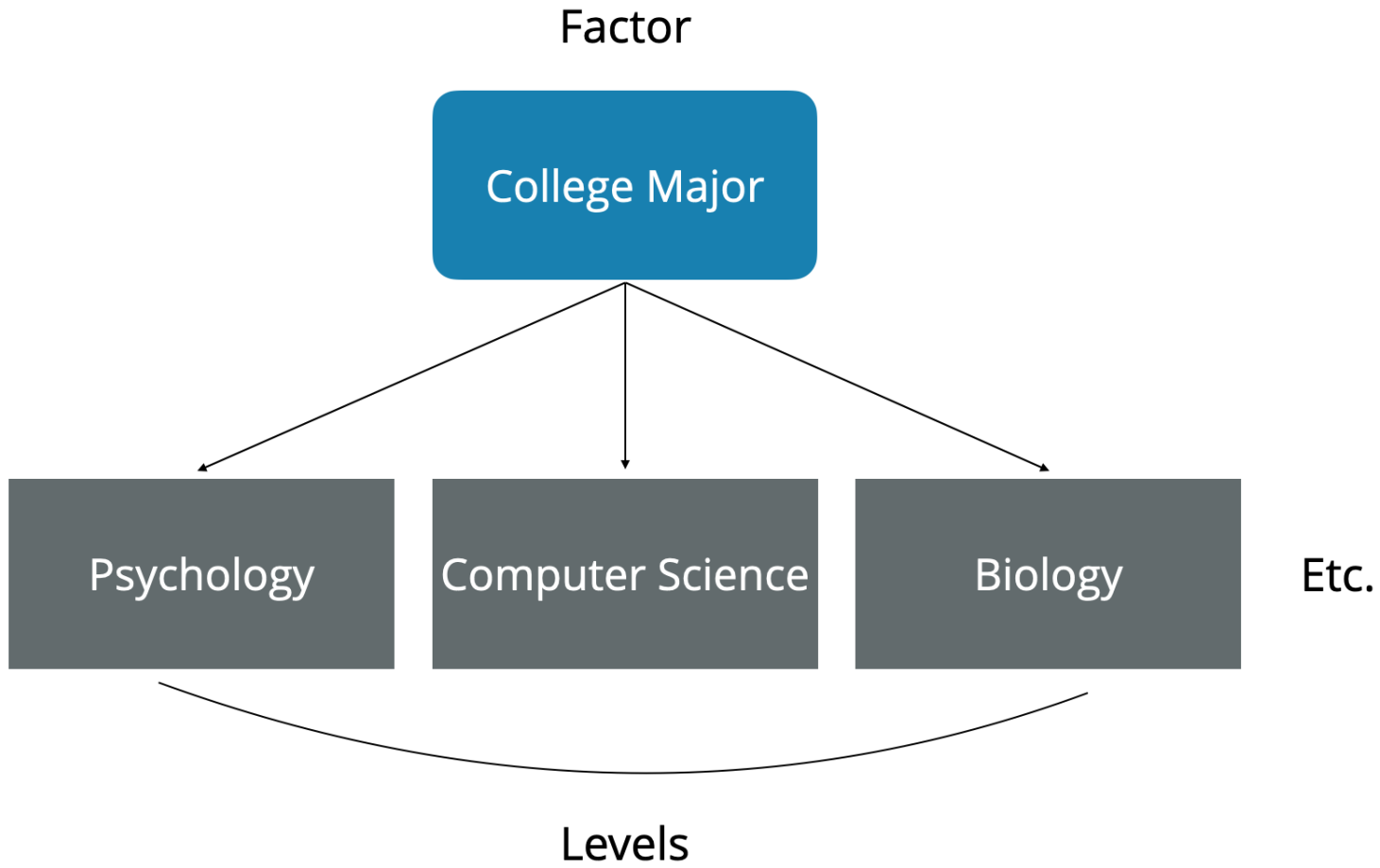
Factor: Independent variable(s) in the study

Level(s): Groups within each independent variable

Response: The dependent variable in the study

Today, we'll learn about *two-way ANOVAs*, which have two factors

Example: One factor, several levels



One-way ANOVA

What does it do?

- Investigates whether any of the population means of our groups differ
- It doesn't tell which group(s) is(are) different; you need post-hoc tests for this (we'll come back to this later today)

Null Hypothesis H_0

There is no difference

The means are equal

$$\mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k$$

Alternative Hypothesis H_a

There is a difference

The means are not all equal

Why not just use many *t*-tests?

- *Type I Error rate* increases drastically as you run additional analyses
    - Each test has a risk of a Type I error, and the more tests you do, the more risk there is (**alpha escalation**)

Experiment-wise Significance Level

$$1 - (1 - \alpha)^k$$

$$1 - (1 - 0.05)^3 = 0.14$$

*You do not need to know this for the final or any quizzes.

Building the *F*-statistic

$$F = \frac{MS_{between}}{MS_{within}}$$

$$F = \frac{MS_{between}}{MS_{within}}$$

$$MS_{between} = \frac{SS_{between}}{df_{between}}$$

$$MS_{within} = \frac{SS_{within}}{df_{within}}$$

$$SS_{between} = \sum \frac{(\sum x)^2}{n} - \frac{(\sum \sum x)^2}{nT}$$

$$SS_{within} = \sum \sum (x^2) - \sum \frac{(\sum x)^2}{n}$$

$$df_{between} = k - 1$$

$$df_{within} = nT - k$$

Sum of Squares: One-way ANOVA

For each group, first calculate

n

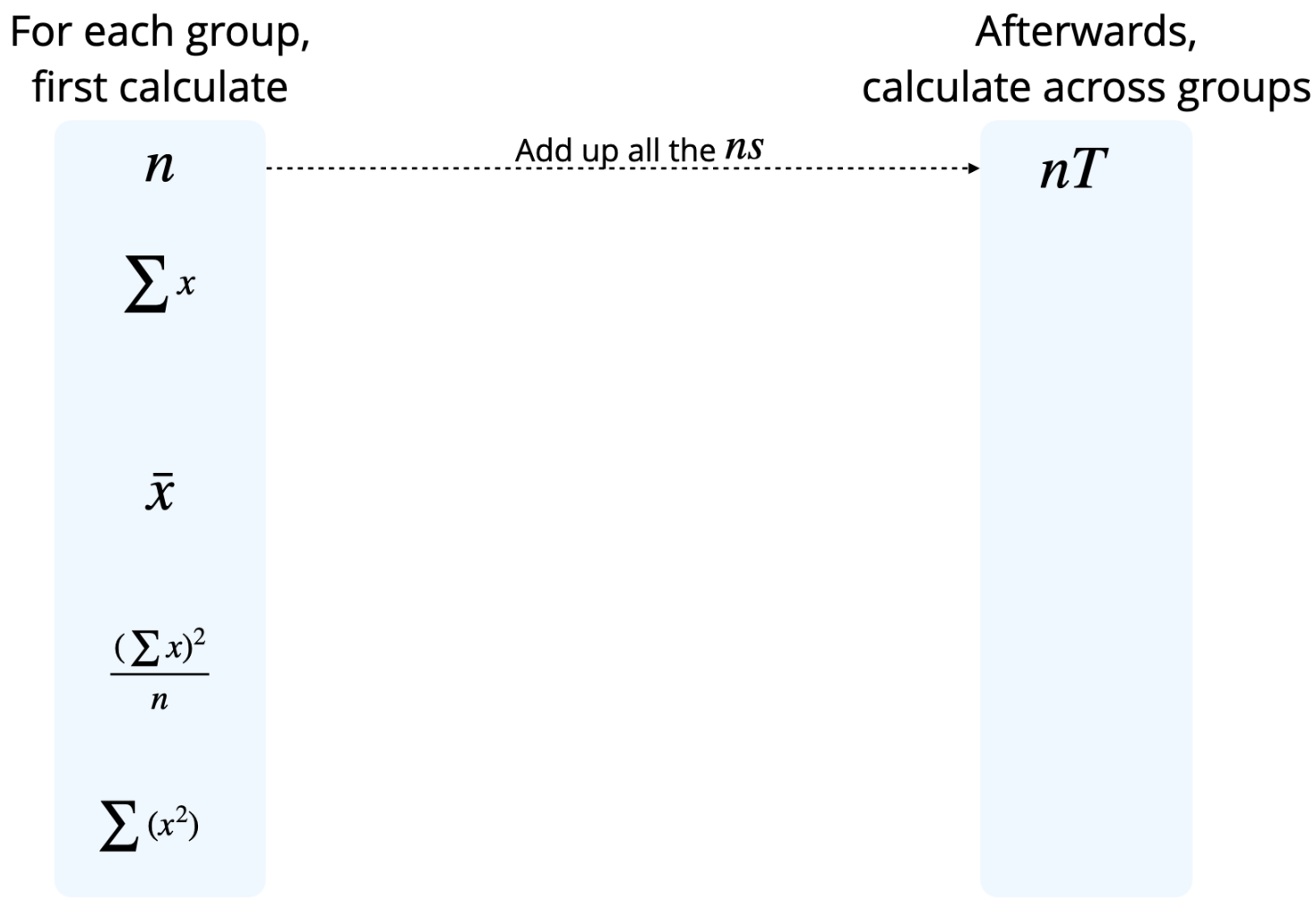
$\sum x$

$\bar{x}$

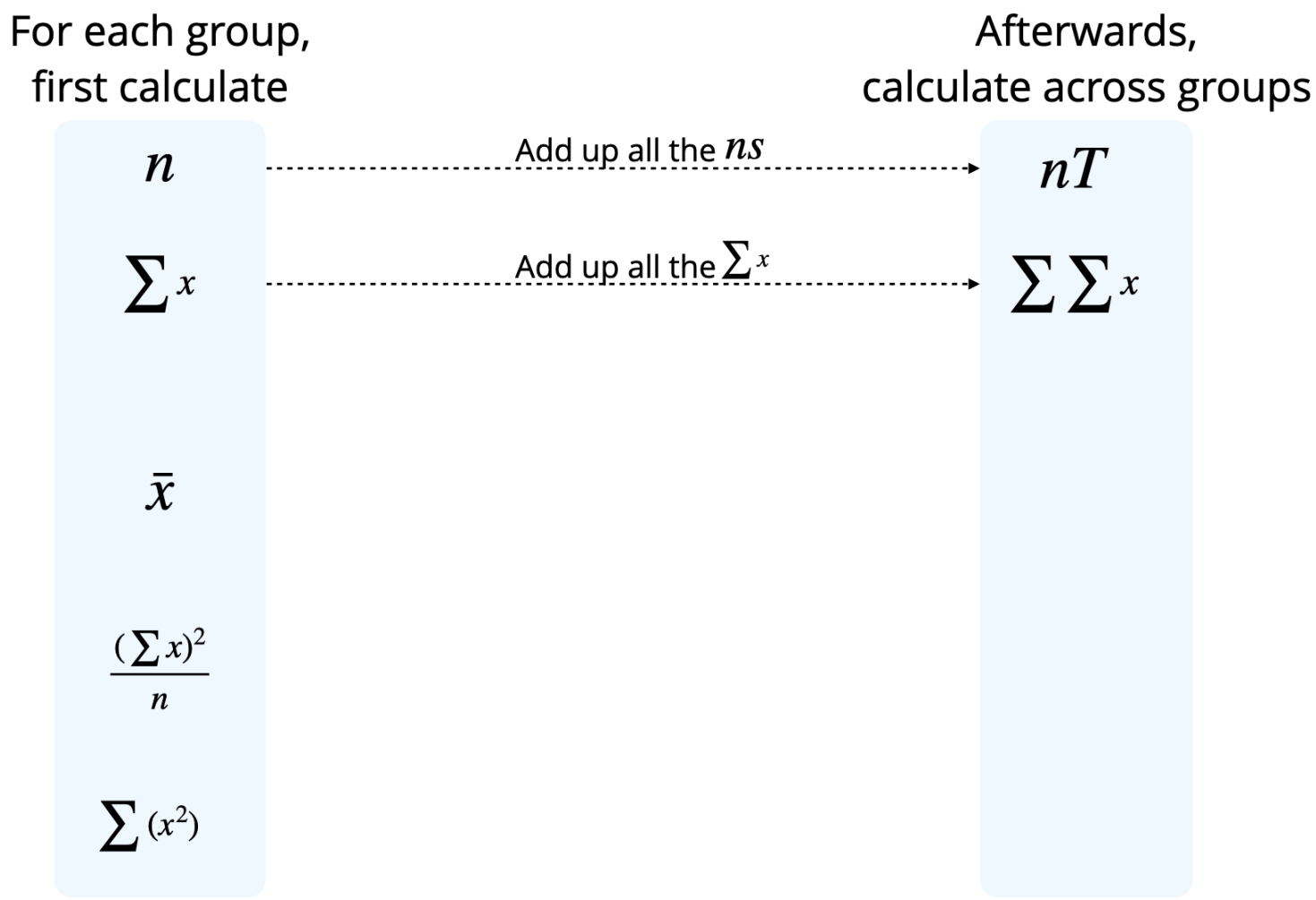
$\frac{(\sum x)^2}{n}$

$\sum (x^2)$

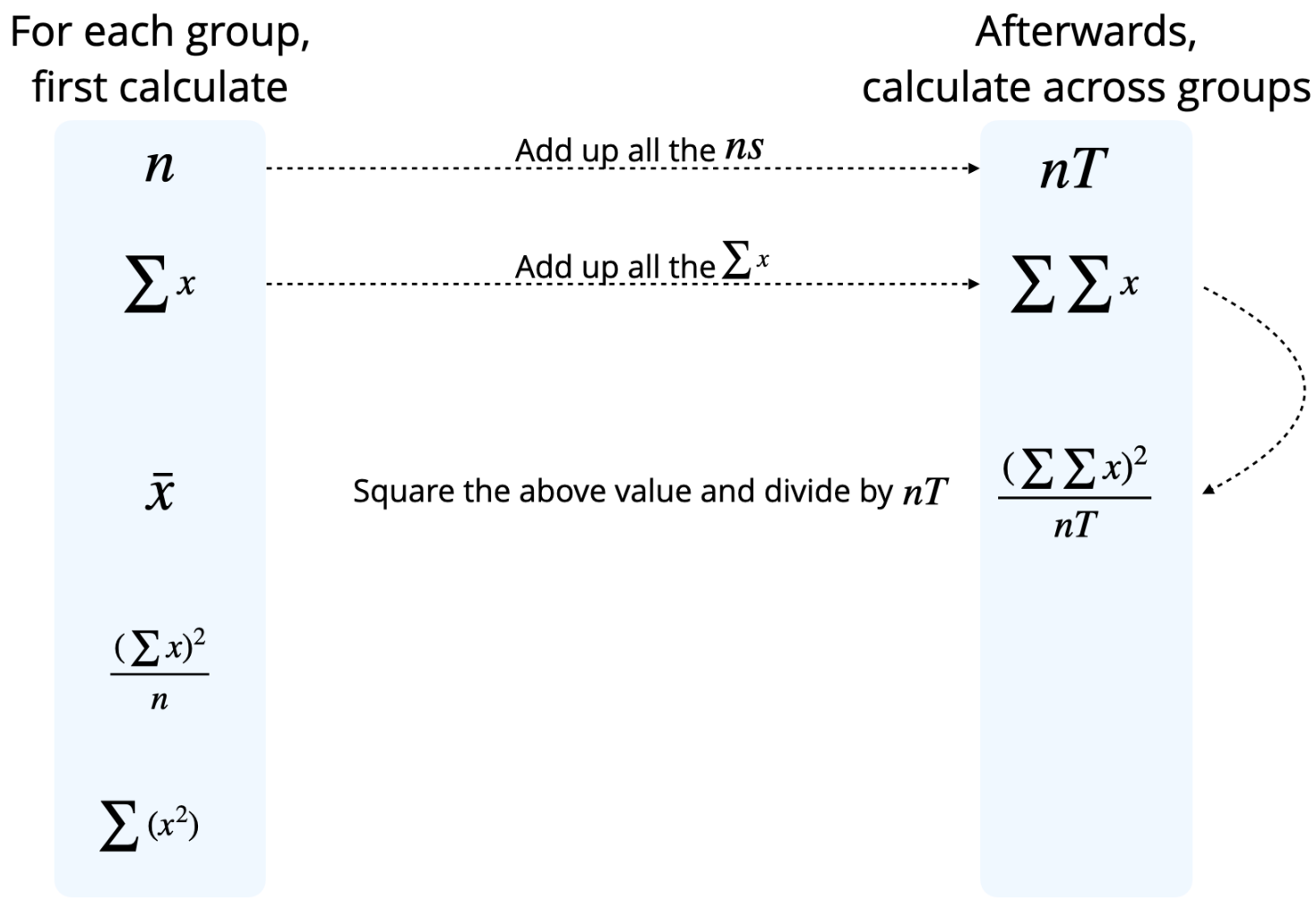
Sum of Squares: One-way ANOVA



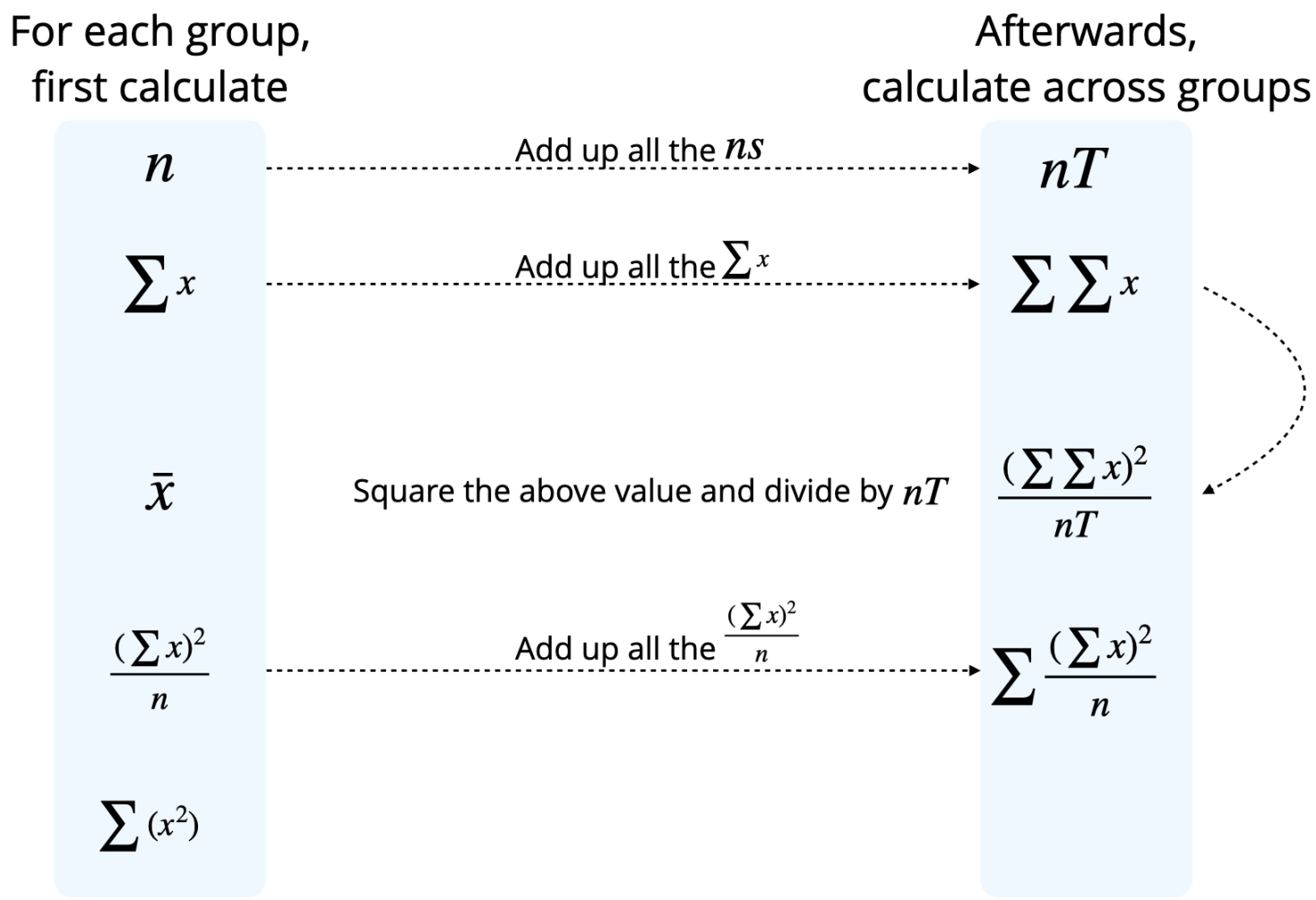
Sum of Squares: One-way ANOVA



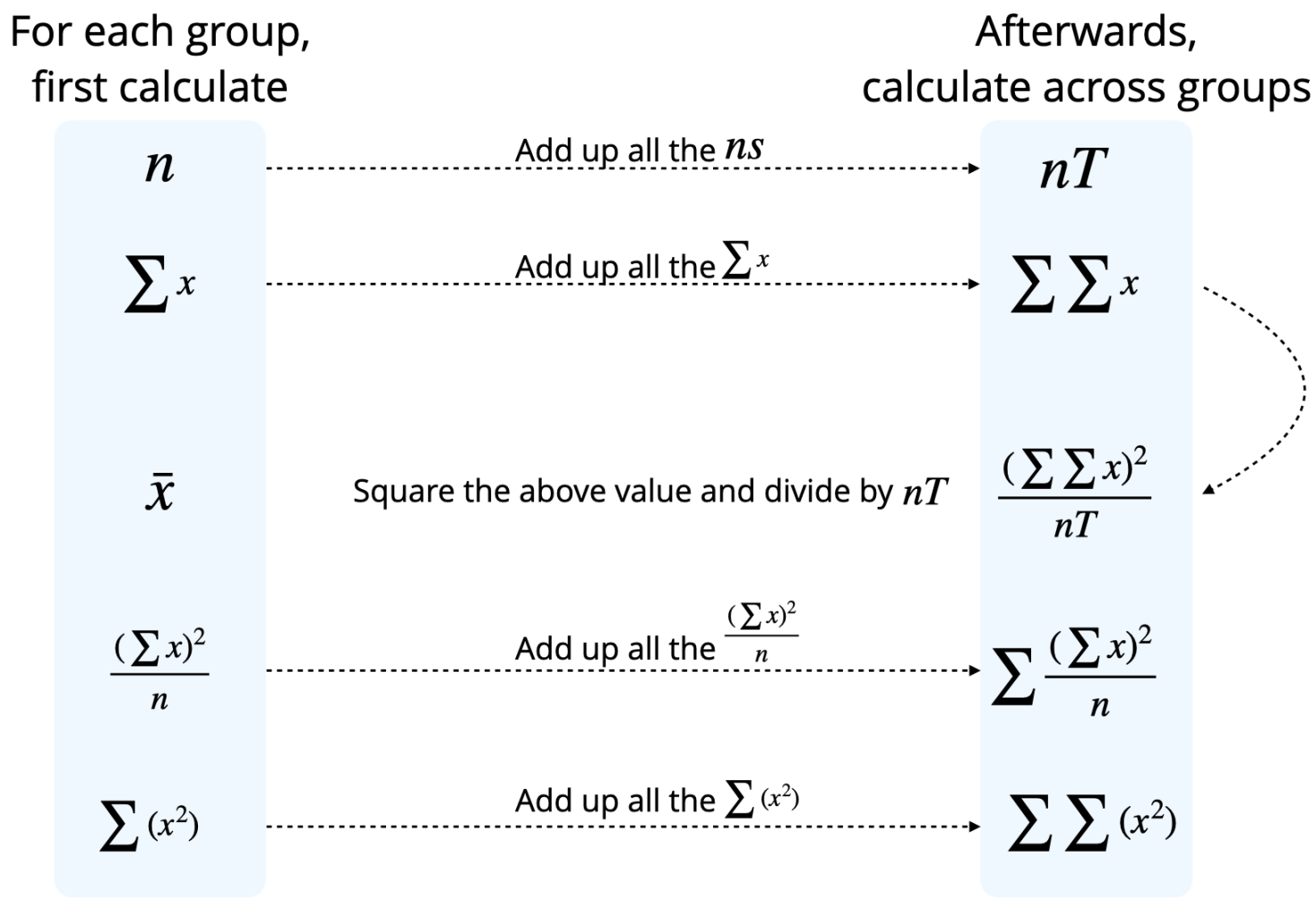
Sum of Squares: One-way ANOVA



Sum of Squares: One-way ANOVA



Sum of Squares: One-way ANOVA



$$F = \frac{MS_{between}}{MS_{within}}$$

$$MS_{between} = \frac{SS_{between}}{df_{between}}$$

$$MS_{within} = \frac{SS_{within}}{df_{within}}$$

$$SS_{between} = \sum \frac{(\sum x)^2}{n} - \frac{(\sum \sum x)^2}{nT}$$

$$SS_{within} = \sum \sum (x^2) - \sum \frac{(\sum x)^2}{n}$$

$$df_{between} = k - 1$$

$$df_{within} = nT - k$$

Effect size: One-way ANOVA

Effect size for One-way ANOVA

$$\eta^2 = \frac{SS_{between}}{SS_{Total}}$$

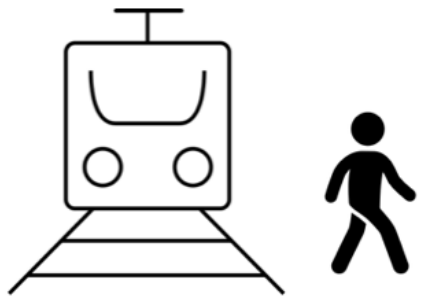
$$SS_{total} = SS_{between} + SS_{within} = \sum\sum(x^2) - \frac{(\sum\sum x)^2}{nT}$$

Two-way (Factorial) ANOVA

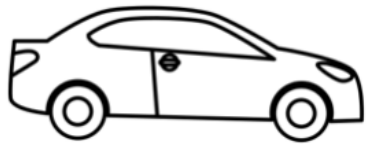
Two-way (Factorial) ANOVA

Goal: To be able to compare whether differences exist between *multiple* factors, as well as *interactions* between the levels of different factors

Research question: What is the fastest way for me to get to campus?

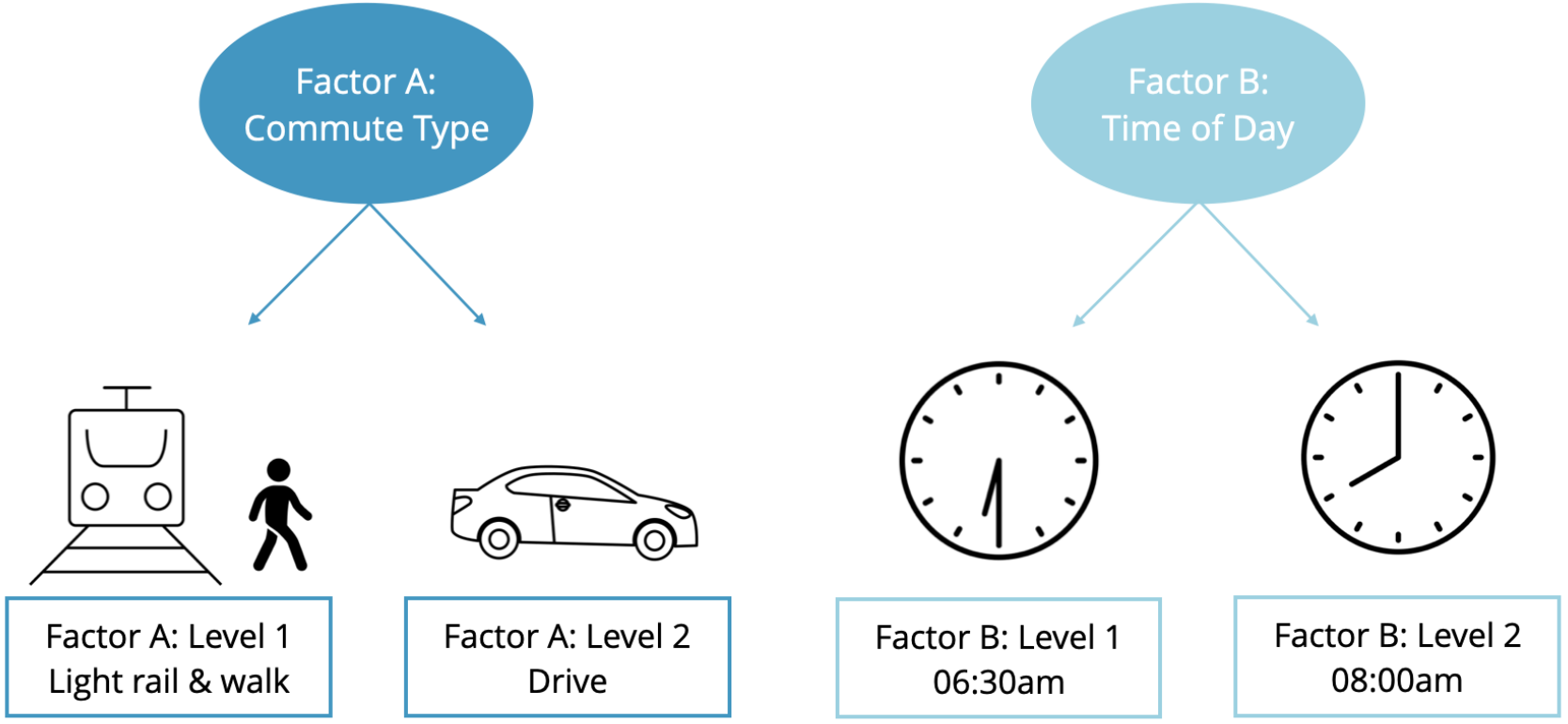


vs.

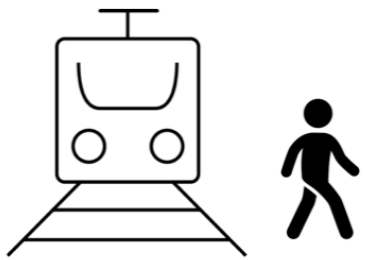


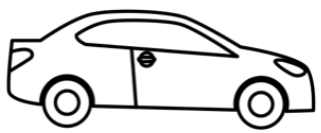
vs.

ANOVA: factors and levels



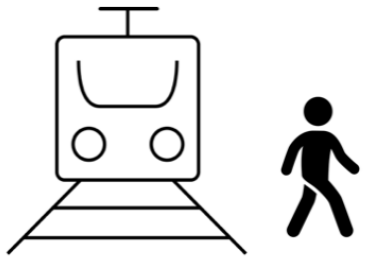
Example Experiment

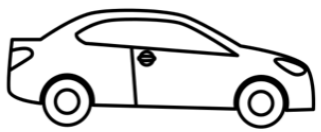




	6:30am	08:00am
Light rail & walk		
Drive		

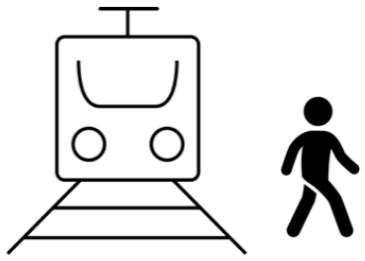
Example Experiment

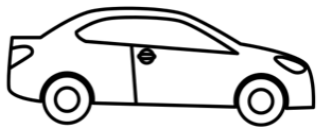




	6:30am	08:00am
Light rail & walk	Average time for light rail and walking at 6:30am	
Drive		

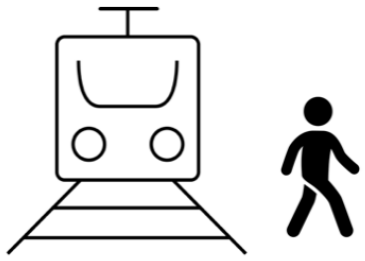
Example Experiment

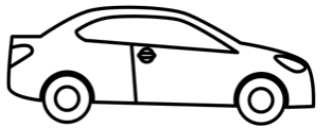




	6:30am	08:00am
Light rail & walk	Average time for light rail and walking at 6:30am	Average time for light rail and walking at 8:00am
Drive		

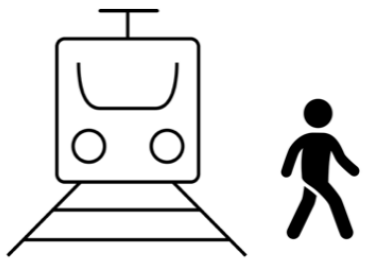
Example Experiment

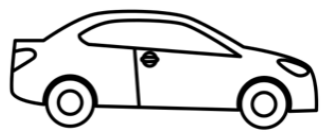




	6:30am	08:00am
Light rail & walk	Average time for light rail and walking at 6:30am	Average time for light rail and walking at 8:00am
Drive	Average time for driving at 6:30am	

Example Experiment

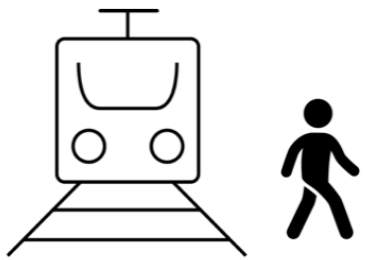


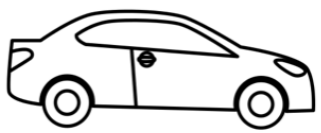


	6:30am	**08:00am**
Light rail & walk	Average time for light rail and walking at 6:30am	Average time for light rail and walking at 8:00am
Drive	Average time for driving at 6:30am	Average time for driving at 8:00am

Effect of Commute type

Possible test #1: Effect of **Commute Type**





Light rail & walk	Average time for light rail and walking
Drive	Average time for driving

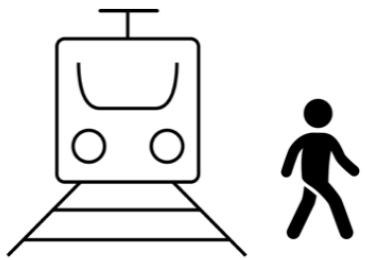
Effect of Time of day

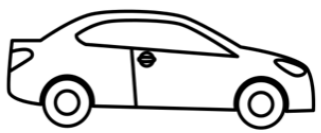
Possible test #2: Effect of **Time of Day**

6:30am	**08:00am**
Average commute at 6:30am	Average commute at 8:00am

Commute type × Time of day interaction

Possible test #3: Effect of specific **Commute Time** and **Time of Day** combinations

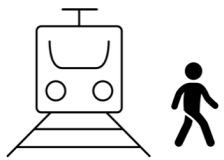




	6:30am	**08:00am**
Light rail & walk	Average time for light rail and walking at 6:30am	Average time for light rail and walking at 8:00am
Drive	Average time for driving at 6:30am	Average time for driving at 8:00am

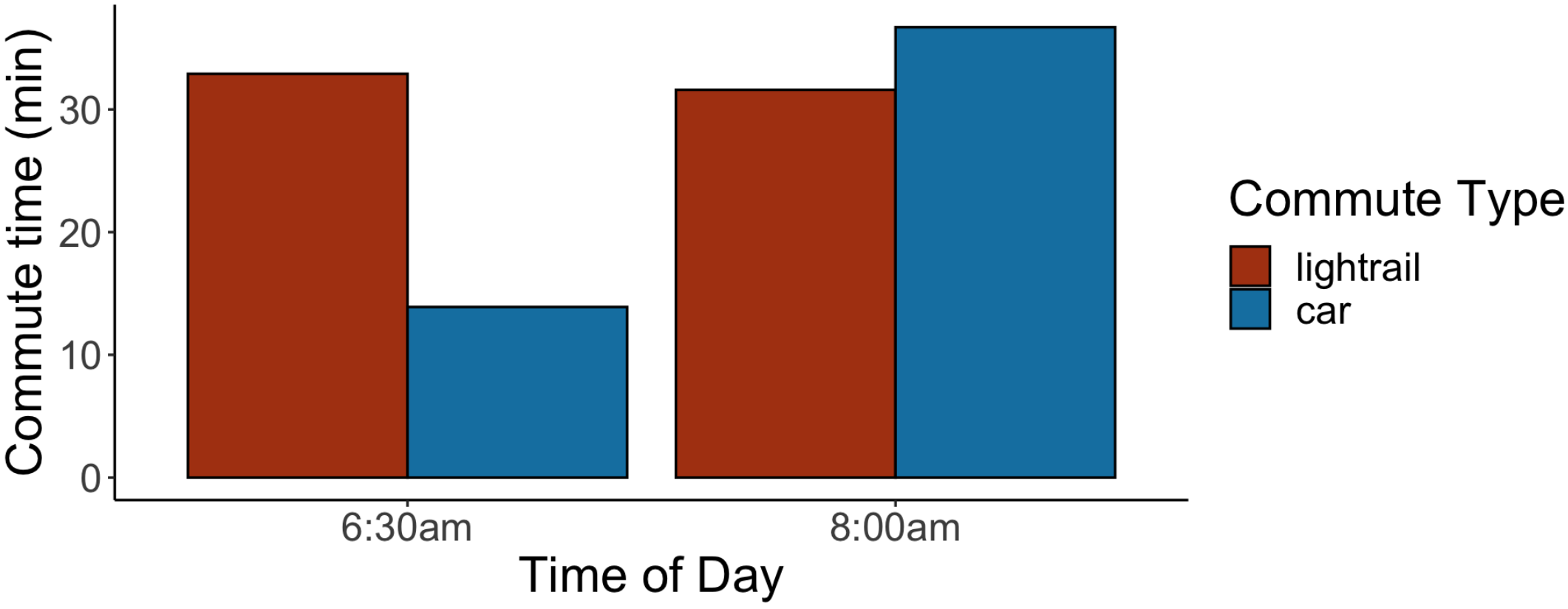
Factorial Designs

Factorial Designs: Designs in which two or more factors are completely crossed (i.e., measurements are taken for every combination of factor levels)

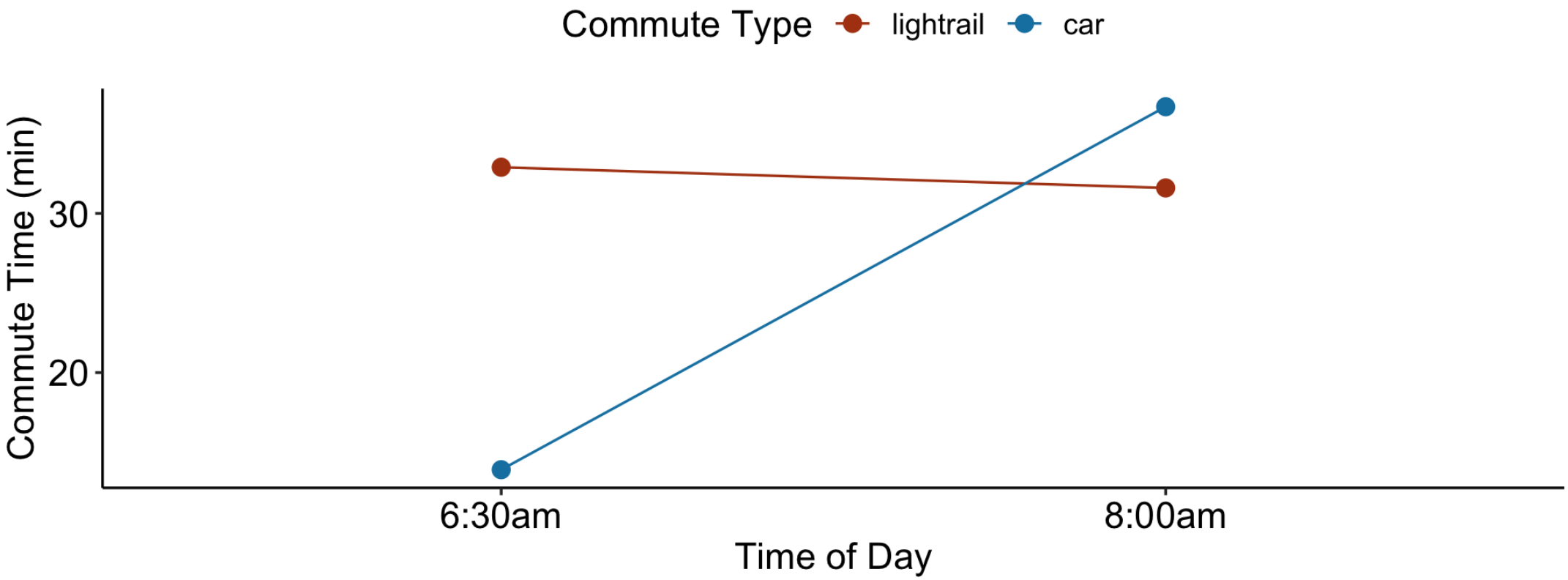


	6:30am	08:00am
Light rail & walk	Average time for light rail and walking at 6:30am	Average time for light rail and walking at 8:00am
Drive	Average time for driving at 6:30am	Average time for driving at 8:00am

Results: Bar Graph

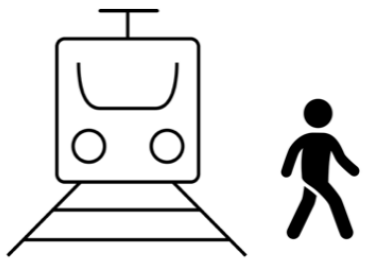


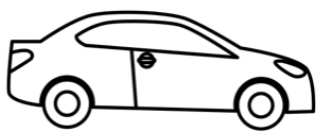
Results: Line Plot



Effect of Commute type

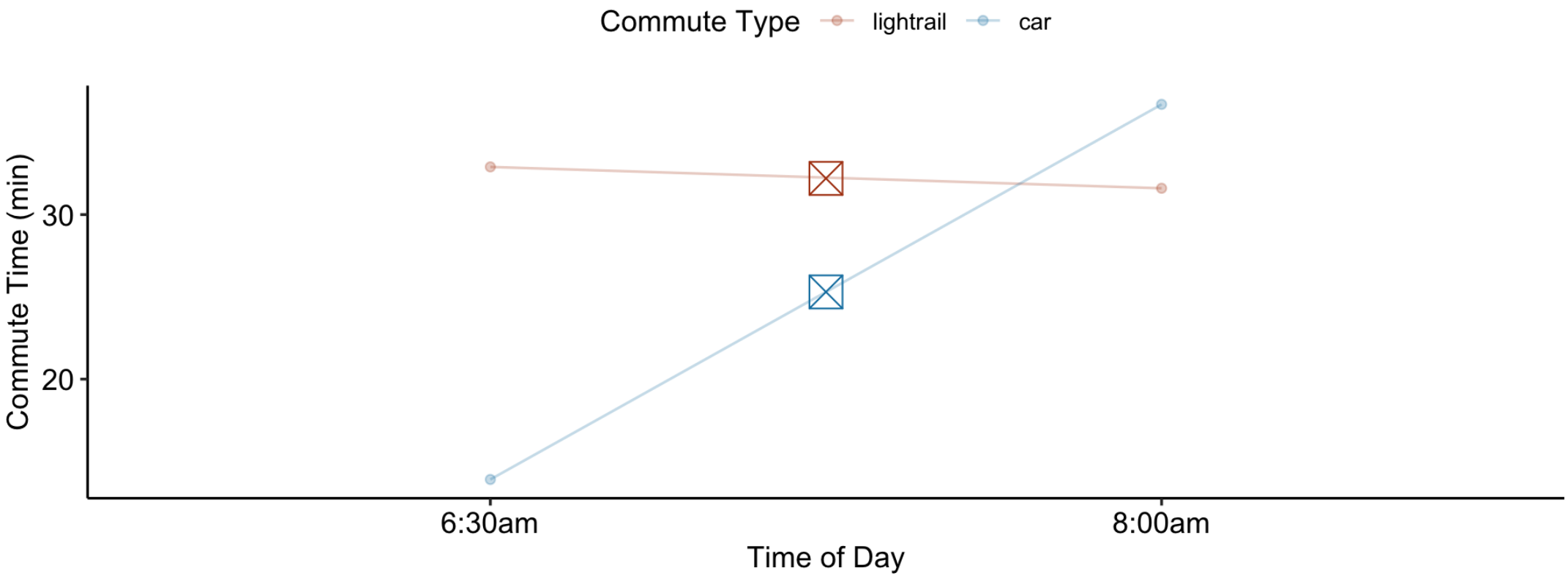
Possible test #1: Effect of **Commute Type**





Light rail & walk	Average time for light rail and walking
Drive	Average time for driving

Commute Type: Plot



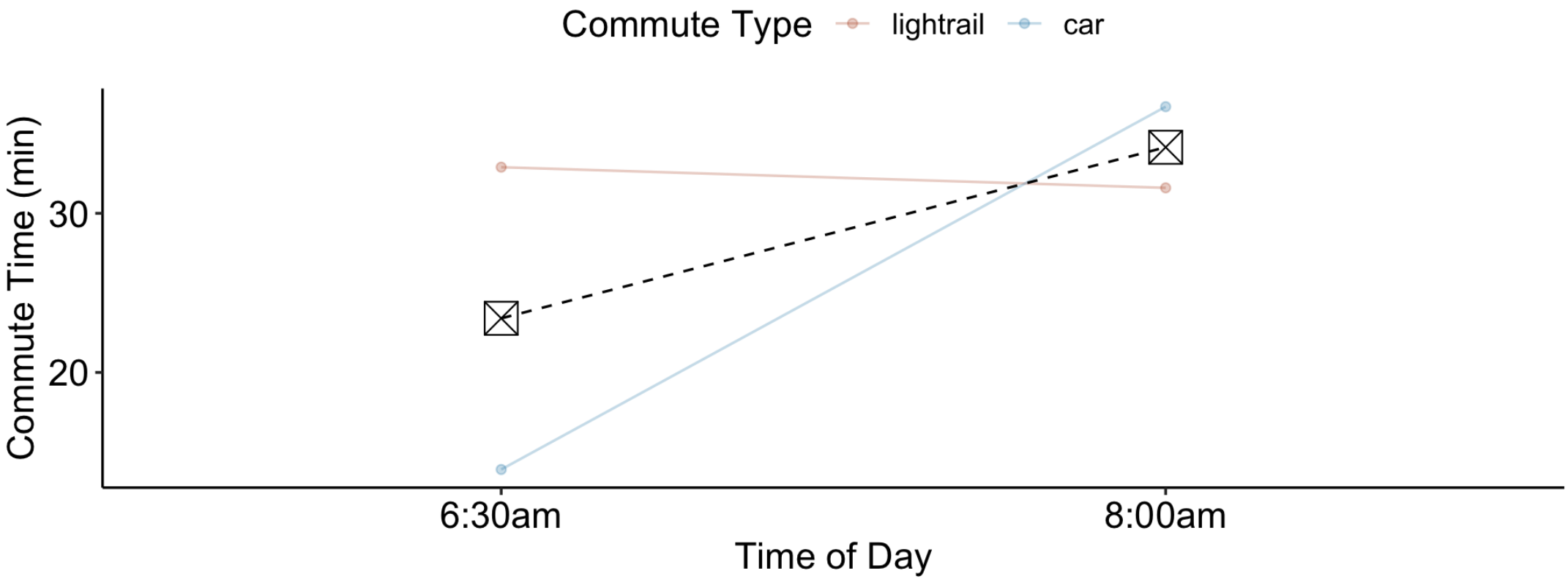
Result: Taking the lightrail and walking takes longer than driving

Effect of Time of day

Possible test #2: Effect of **Time of Day**

6:30am	**08:00am**
Average commute at 6:30am	Average commute at 8:00am

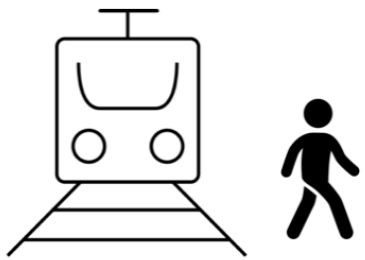
Time of Day: Plot

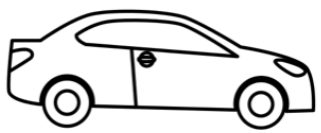


Result: It takes longer to get to campus when leaving at 8:00am than at 6:30am

Commute type × Time of day interaction

Possible test #3: Effect of specific **Commute Time** and **Time of Day** combinations



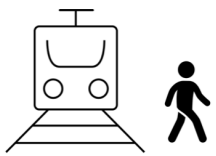


	6:30am	**08:00am**
Light rail & walk	Average time for light rail and walking at 6:30am	Average time for light rail and walking at 8:00am
Drive	Average time for driving at 6:30am	Average time for driving at 8:00am

Interaction

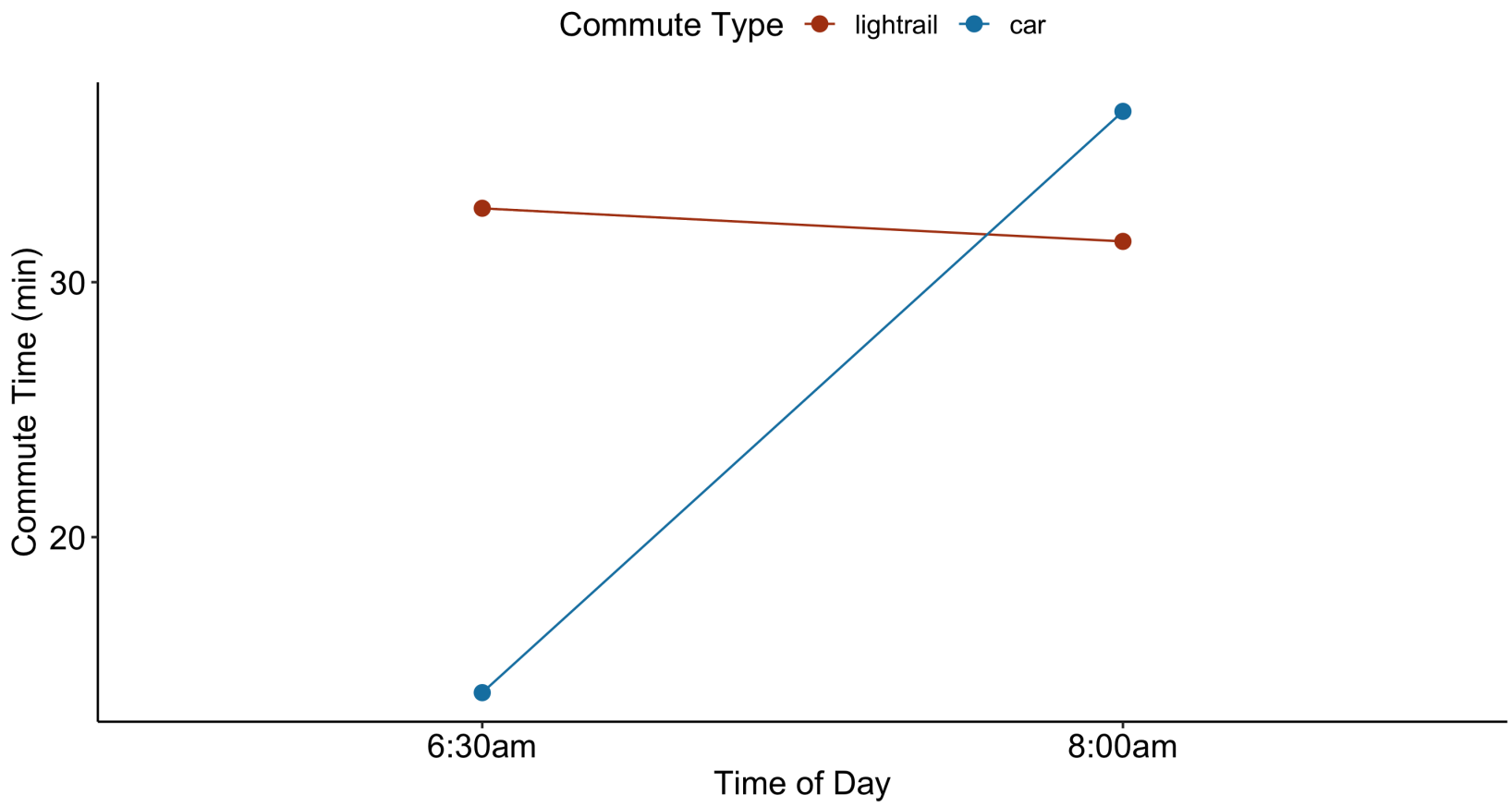
Interaction: describes the degree to which the effect of *one factor* depends on the *level* of the other factor

Possible test #3: Effect of specific **Commute Time** and **Time of Day** combinations

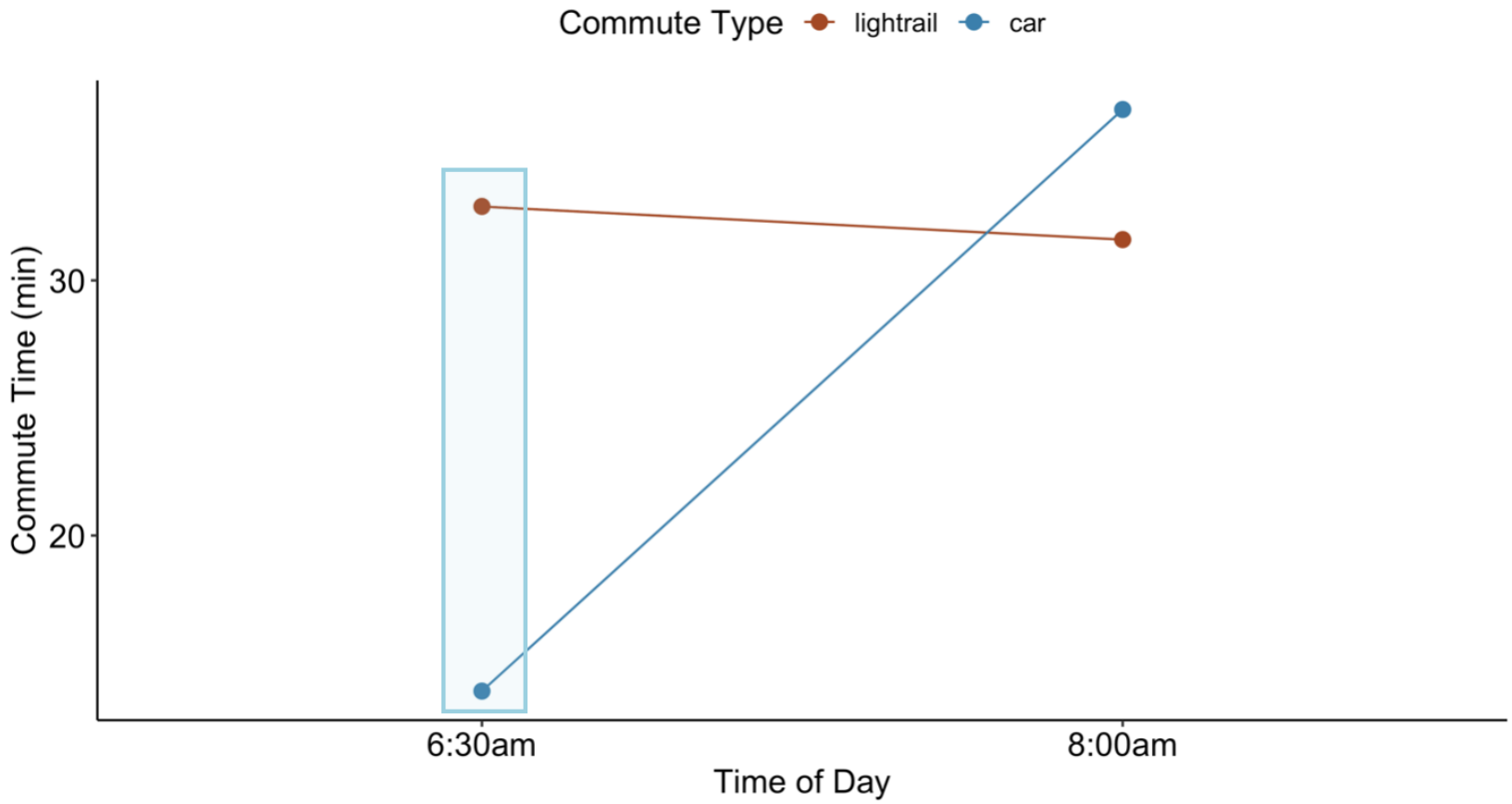


	6:30am	08:00am
Light rail & walk	Average time for light rail and walking at 6:30am	Average time for light rail and walking at 8:00am
Drive	Average time for driving at 6:30am	Average time for driving at 8:00am

Testing the Interaction

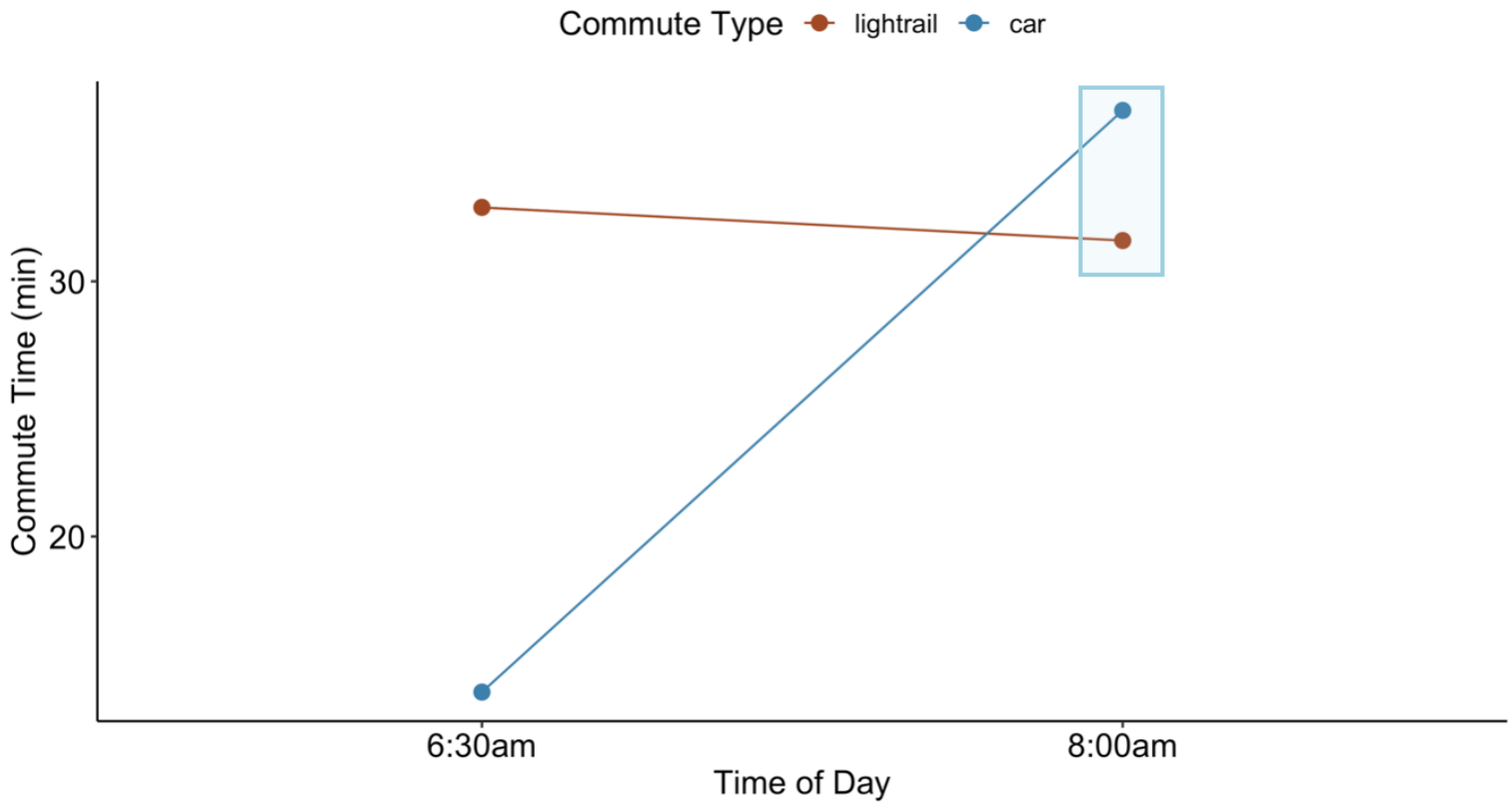


Testing the Interaction



At **6:30am**, **driving** is much *faster* than taking the **light rail and walking**

Testing the Interaction



But at **8:00am**, **light rail and walking** is much *faster* than **driving**

Effect size: Factorial ANOVA

Omega Squared

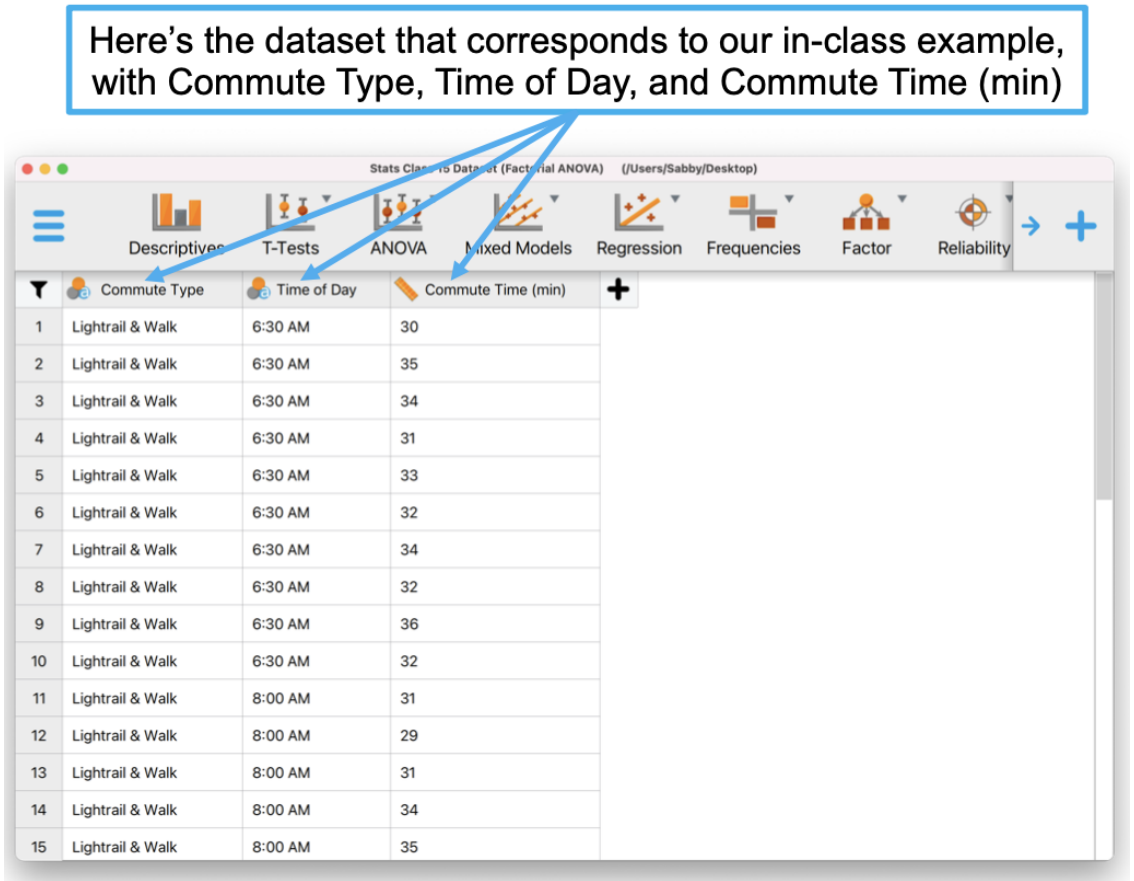
$$w^2 = \frac{SS_{between} - (df_{between})(MS_{within})}{MS_{within} + SS_{total}}$$

You can interpret this the same way as eta-squared, η^2

Note: I will **not** ask you to compute this; I only display this formula to illustrate that you can compute the effect size for a factorial ANOVA

Two-way (Factorial) ANOVA in JASP

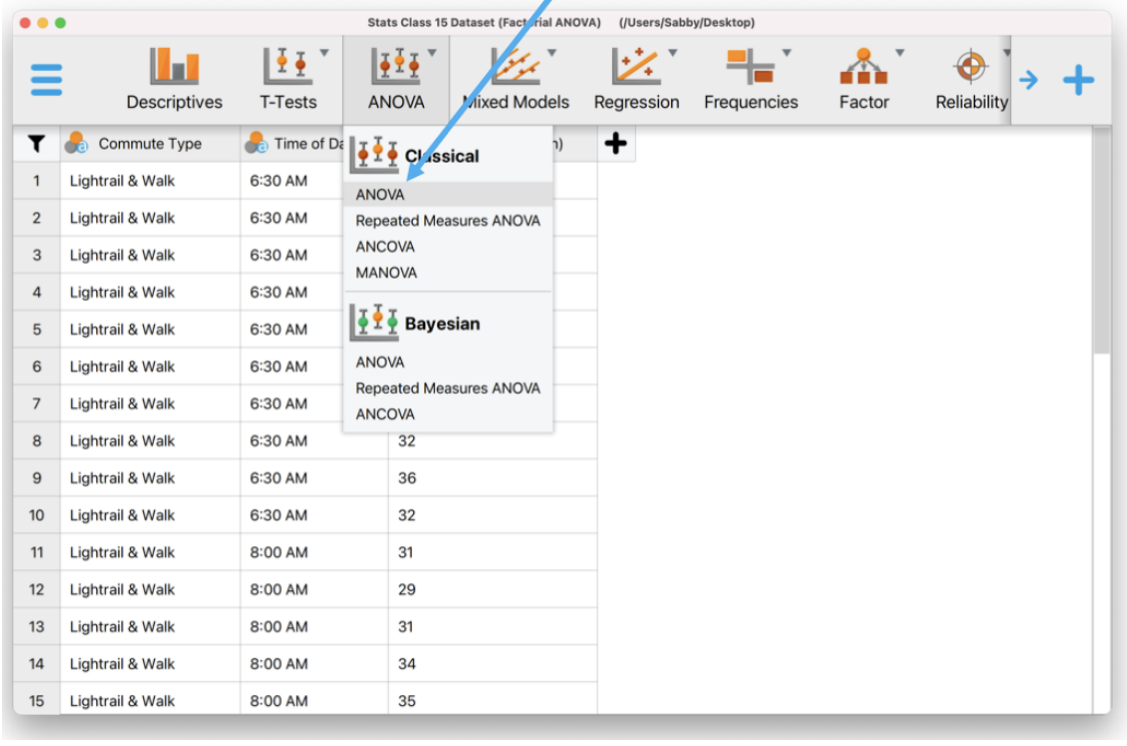
JASP: Two-way ANOVA



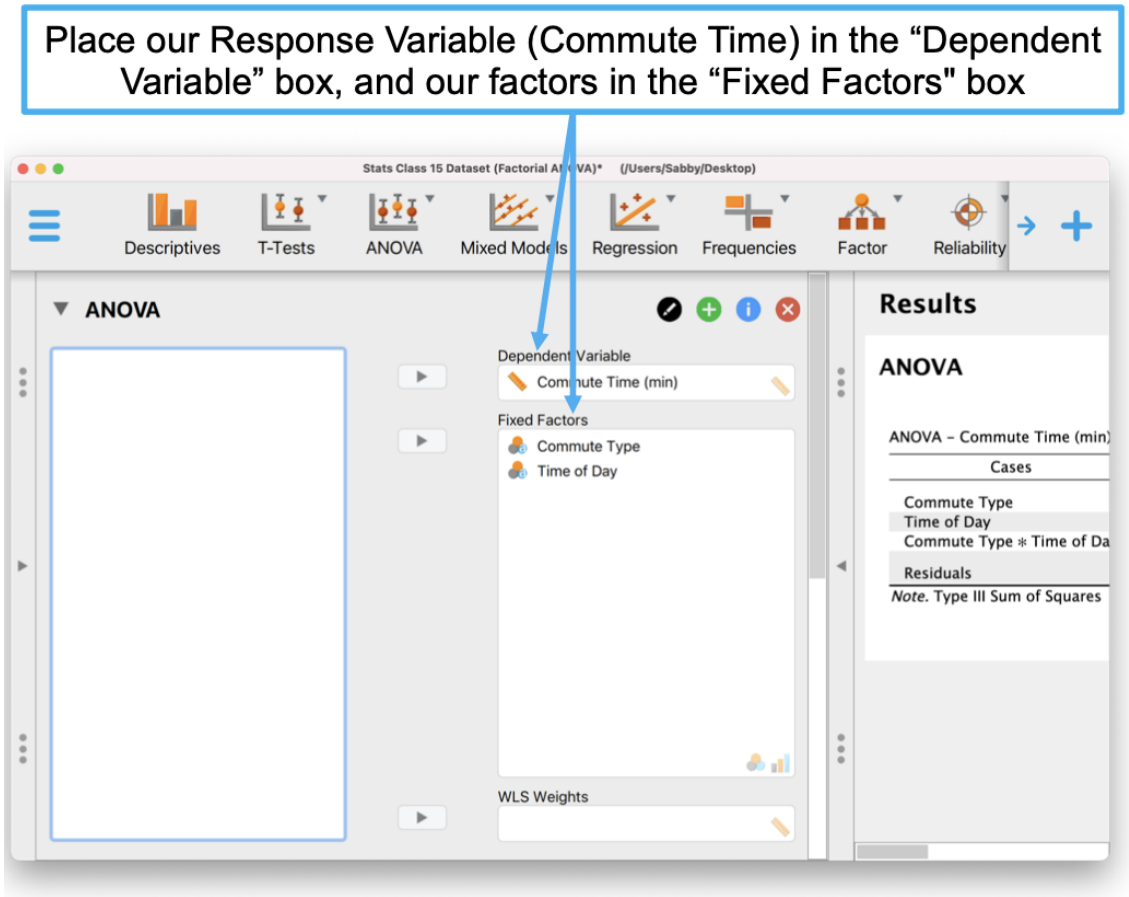
	Commute Type	Time of Day	Commute Time (min)
1	Lightrail & Walk	6:30 AM	30
2	Lightrail & Walk	6:30 AM	35
3	Lightrail & Walk	6:30 AM	34
4	Lightrail & Walk	6:30 AM	31
5	Lightrail & Walk	6:30 AM	33
6	Lightrail & Walk	6:30 AM	32
7	Lightrail & Walk	6:30 AM	34
8	Lightrail & Walk	6:30 AM	32
9	Lightrail & Walk	6:30 AM	36
10	Lightrail & Walk	6:30 AM	32
11	Lightrail & Walk	8:00 AM	31
12	Lightrail & Walk	8:00 AM	29
13	Lightrail & Walk	8:00 AM	31
14	Lightrail & Walk	8:00 AM	34
15	Lightrail & Walk	8:00 AM	35

JASP: Two-way ANOVA

To run a factorial ANOVA in JASP, begin the same way as a one-way ANOVA (i.e., ANOVA → ANOVA, as shown below)

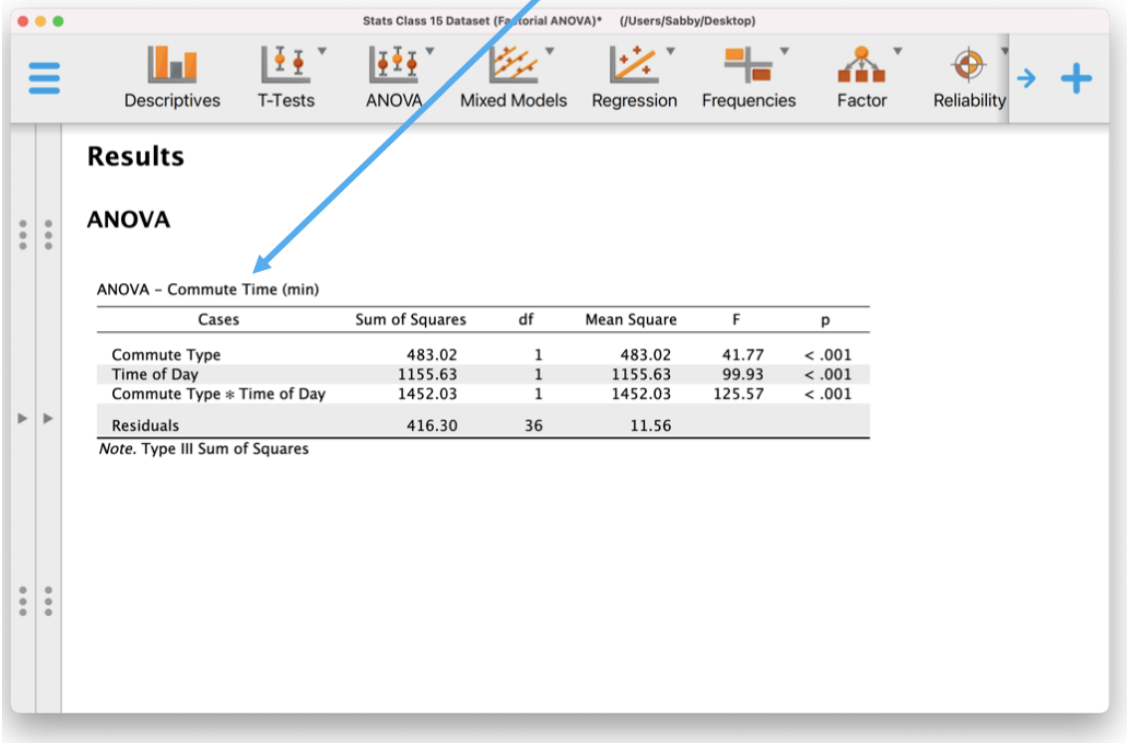


JASP: Two-way ANOVA

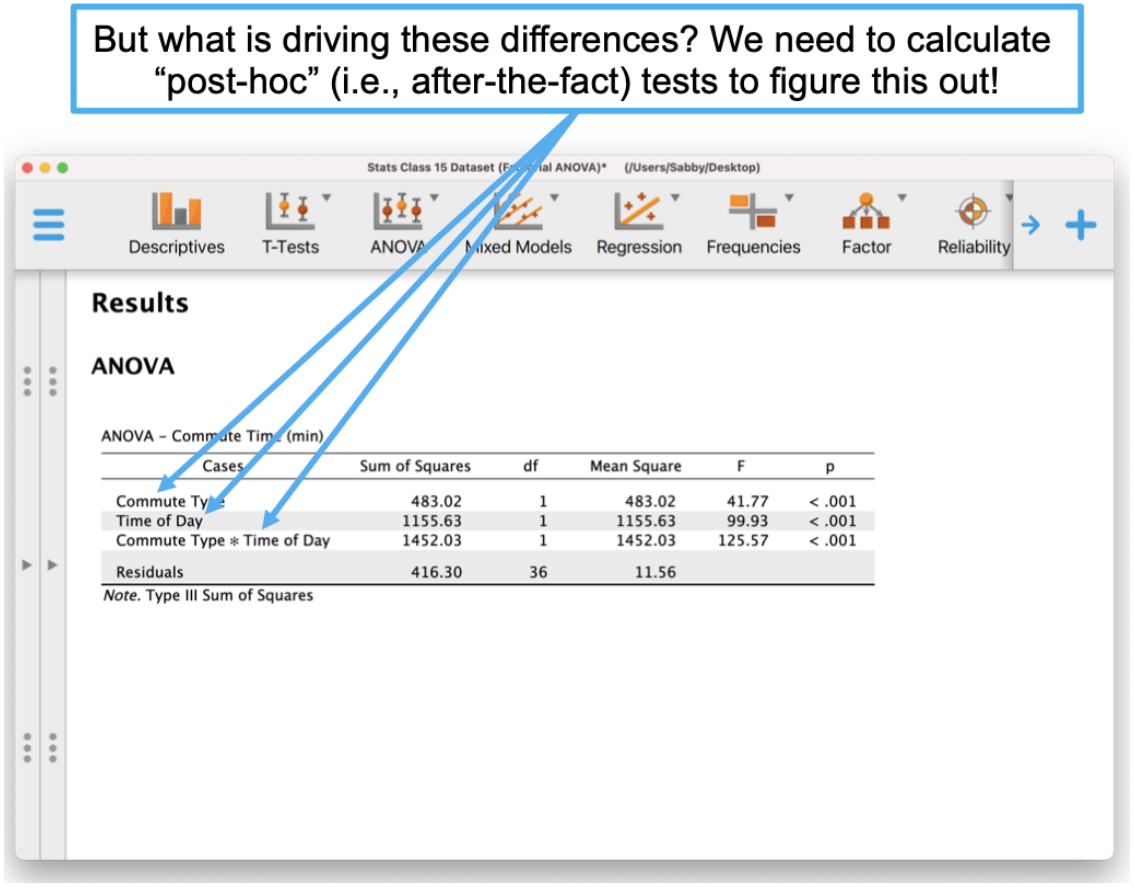


JASP: Two-way ANOVA

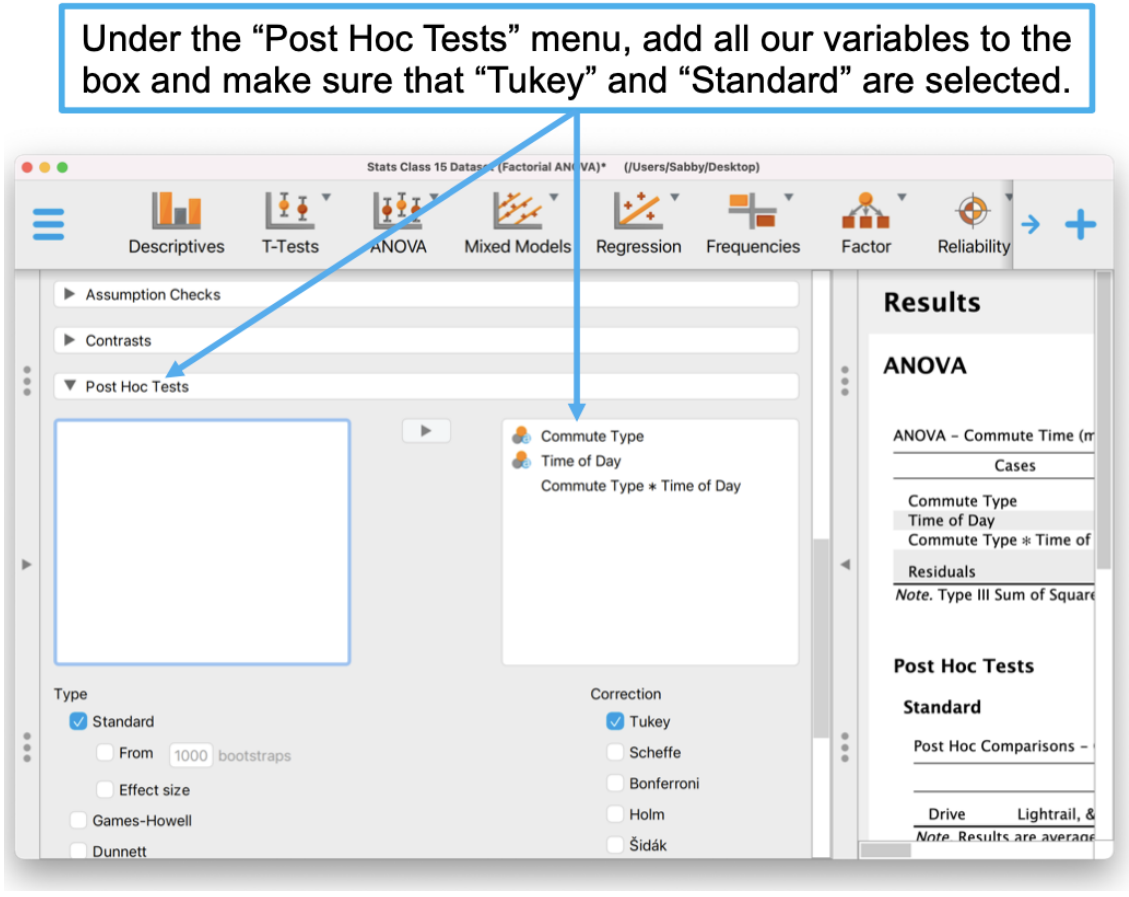
JASP calculates the factorial ANOVA for us, just like that! We have two significant main effects and a significant interaction.



JASP: Two-way ANOVA

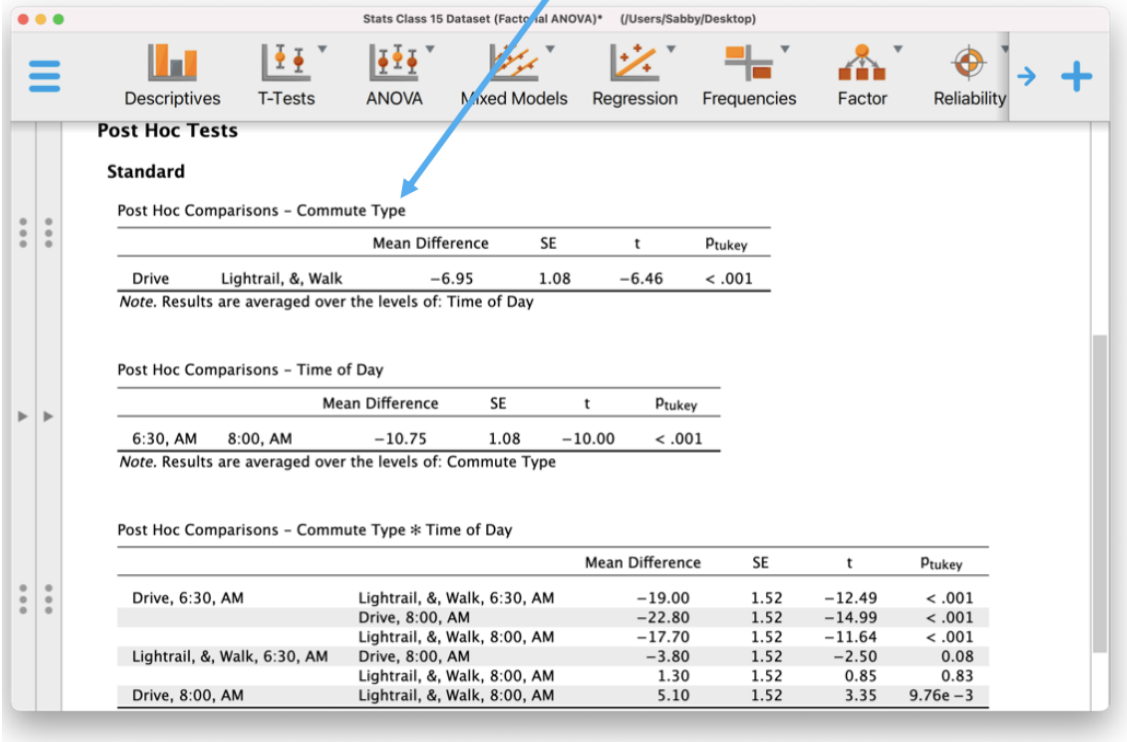


JASP: Two-way ANOVA



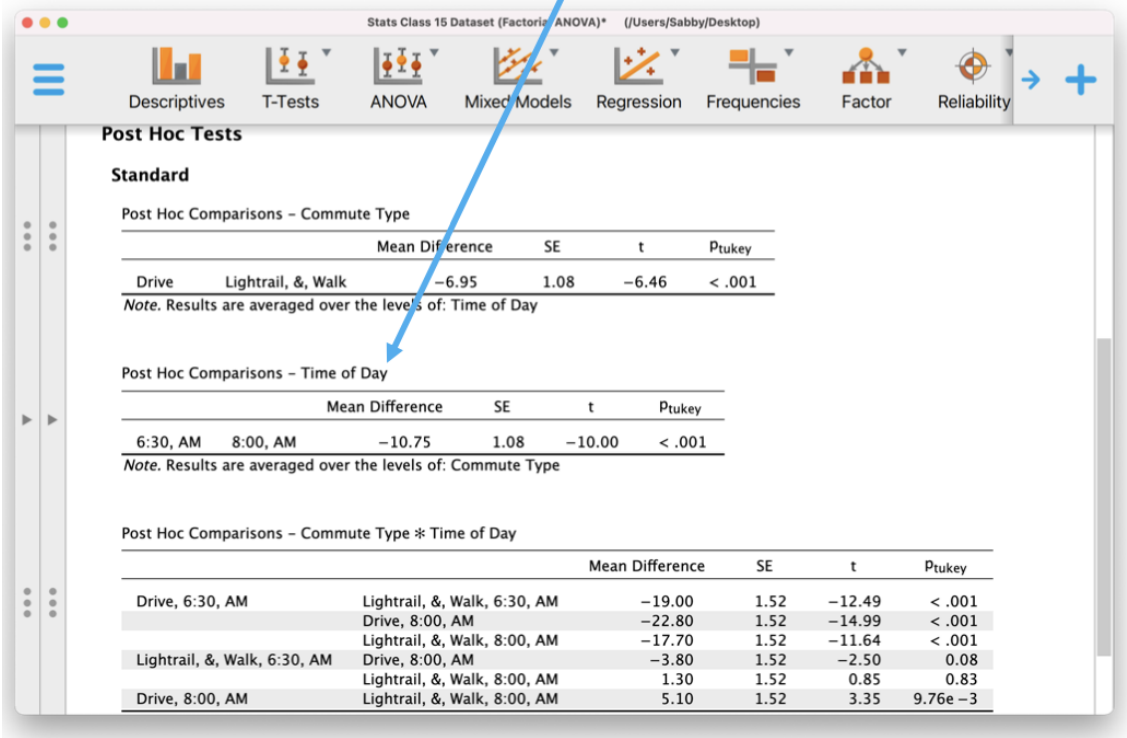
JASP: Two-way ANOVA

The first post-hoc test shows describes the significant difference between driving vs. taking the light rail (light rail takes longer).



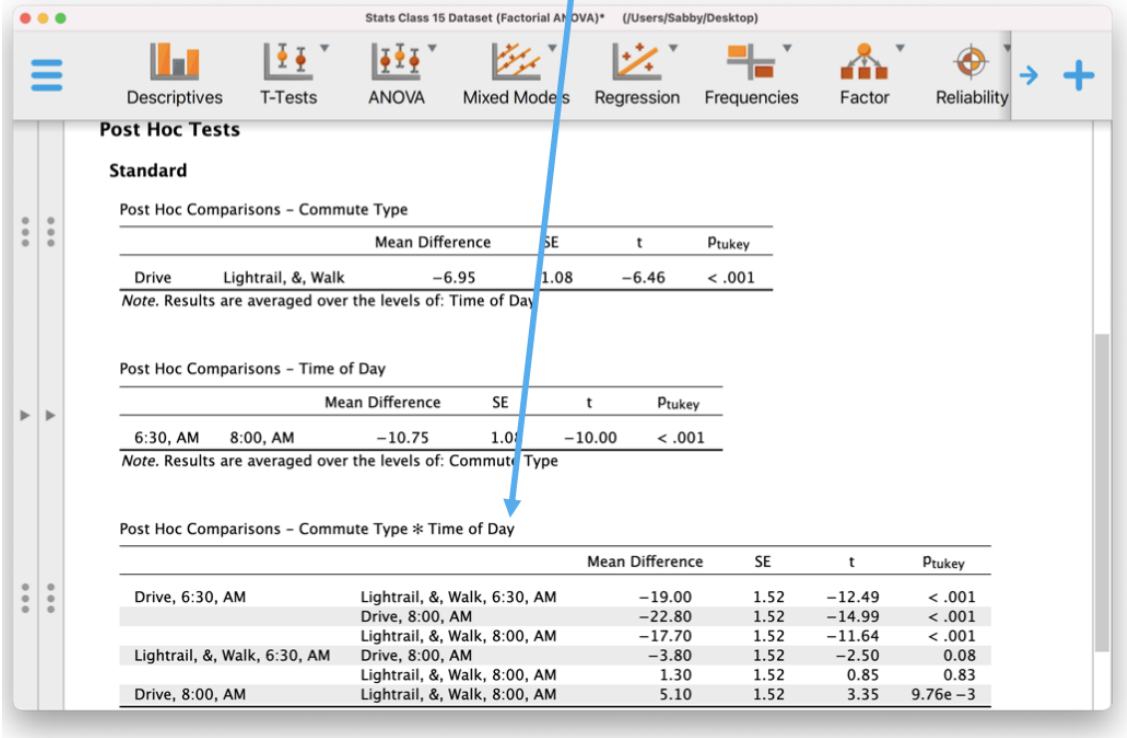
JASP: Two-way ANOVA

The next post-hoc test shows describes the significant difference between 6:30 vs. 8:00 AM departure (8:00 AM takes longer).



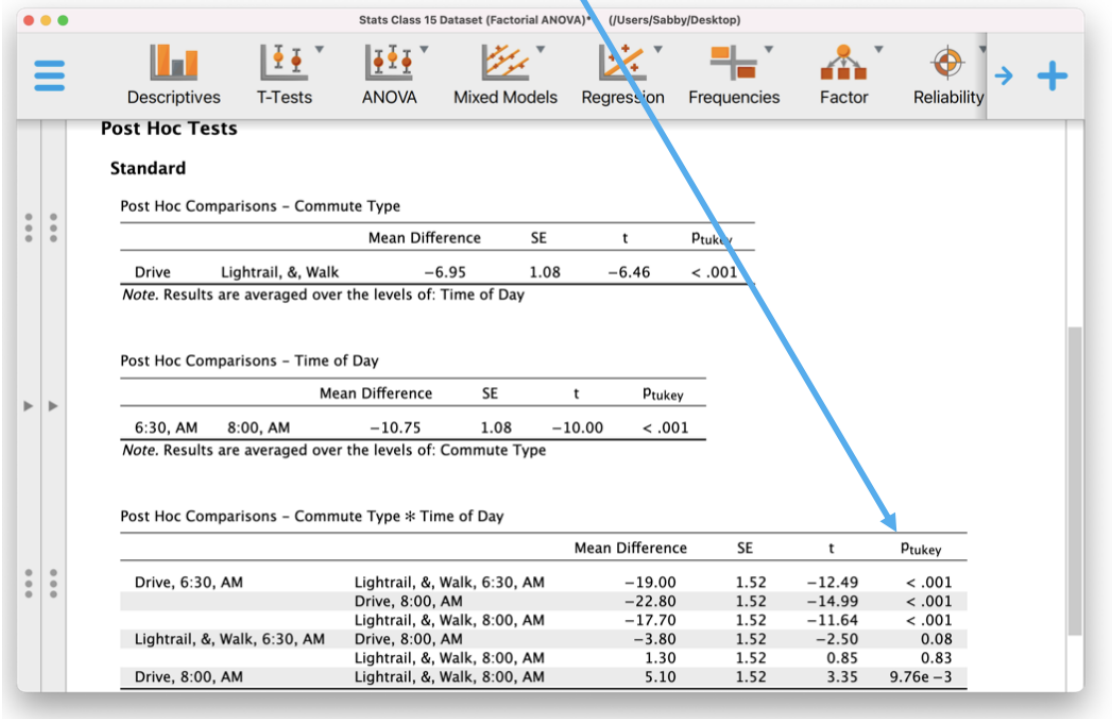
JASP: Two-way ANOVA

Next, we have the interaction, in which the Tukey test shows us every possible comparison with p-values adjusted for many tests.



JASP: Two-way ANOVA

Notice which differences are significant (e.g., the p-value for driving at 6:30 AM vs. taking the light rail at 6:30 AM is < .001).



Post Hoc Tests

Standard

Post Hoc Comparisons - Commute Type

		Mean Difference	SE	t	P_{tukey}
Drive	Lightrail, &, Walk	−6.95	1.08	−6.46	< .001

Note. Results are averaged over the levels of: Time of Day

Post Hoc Comparisons - Time of Day

		Mean Difference	SE	t	P_{tukey}
6:30, AM	8:00, AM	−10.75	1.08	−10.00	< .001

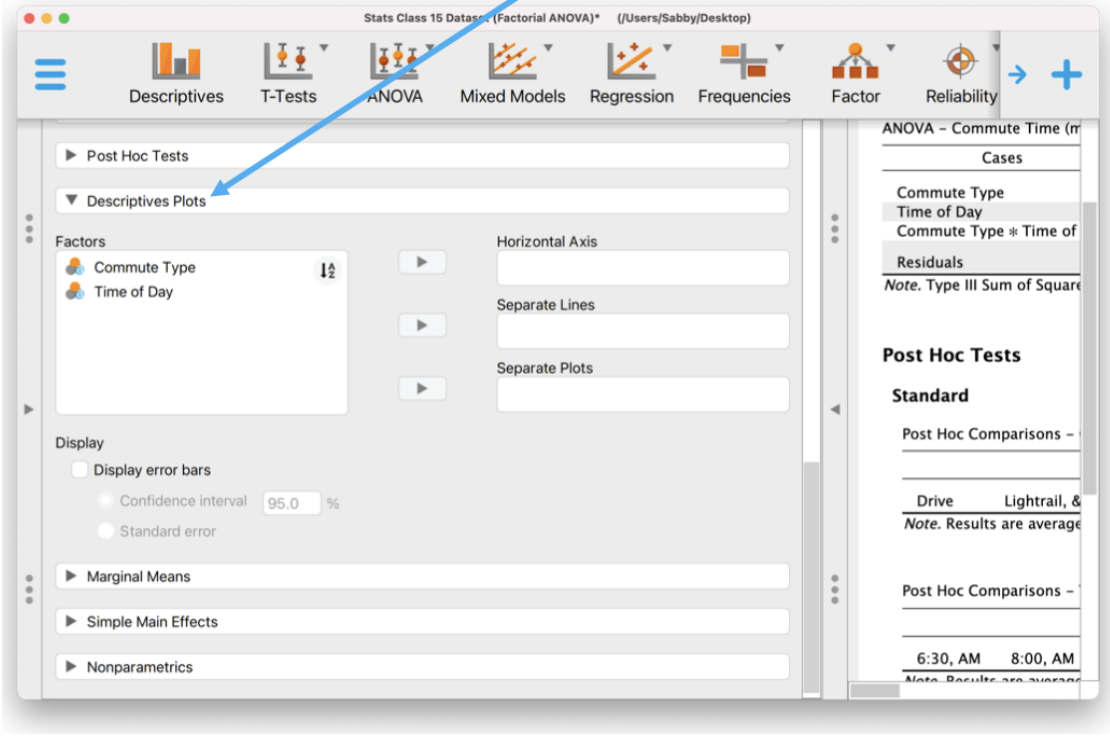
Note. Results are averaged over the levels of: Commute Type

Post Hoc Comparisons - Commute Type ✻ Time of Day

		Mean Difference	SE	t	P_{tukey}
Drive, 6:30, AM	Lightrail, &, Walk, 6:30, AM	−19.00	1.52	−12.49	< .001
	Drive, 8:00, AM	−22.80	1.52	−14.99	< .001
	Lightrail, &, Walk, 8:00, AM	−17.70	1.52	−11.64	< .001
Lightrail, &, Walk, 6:30, AM	Drive, 8:00, AM	−3.80	1.52	−2.50	0.08
	Lightrail, &, Walk, 8:00, AM	1.30	1.52	0.85	0.83
Drive, 8:00, AM	Lightrail, &, Walk, 8:00, AM	5.10	1.52	3.35	9.76e−3

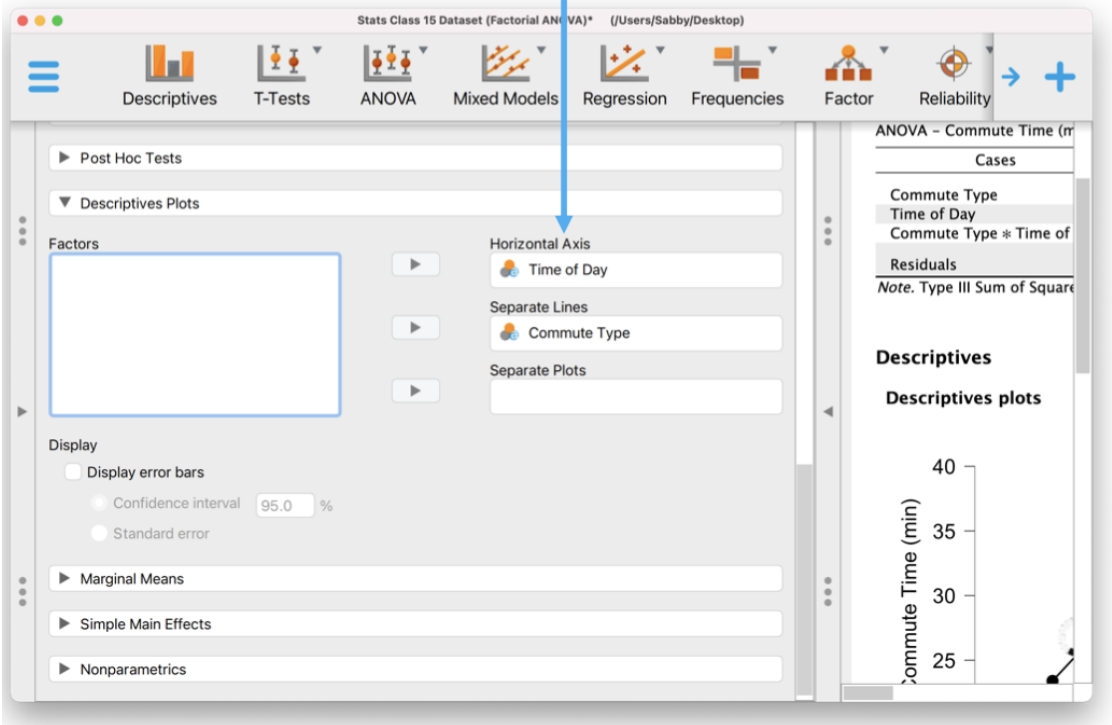
JASP: Two-way ANOVA

Let's create an interaction plot to help us and others visualize these differences by selecting the "Descriptive Plots" menu.



JASP: Two-way ANOVA

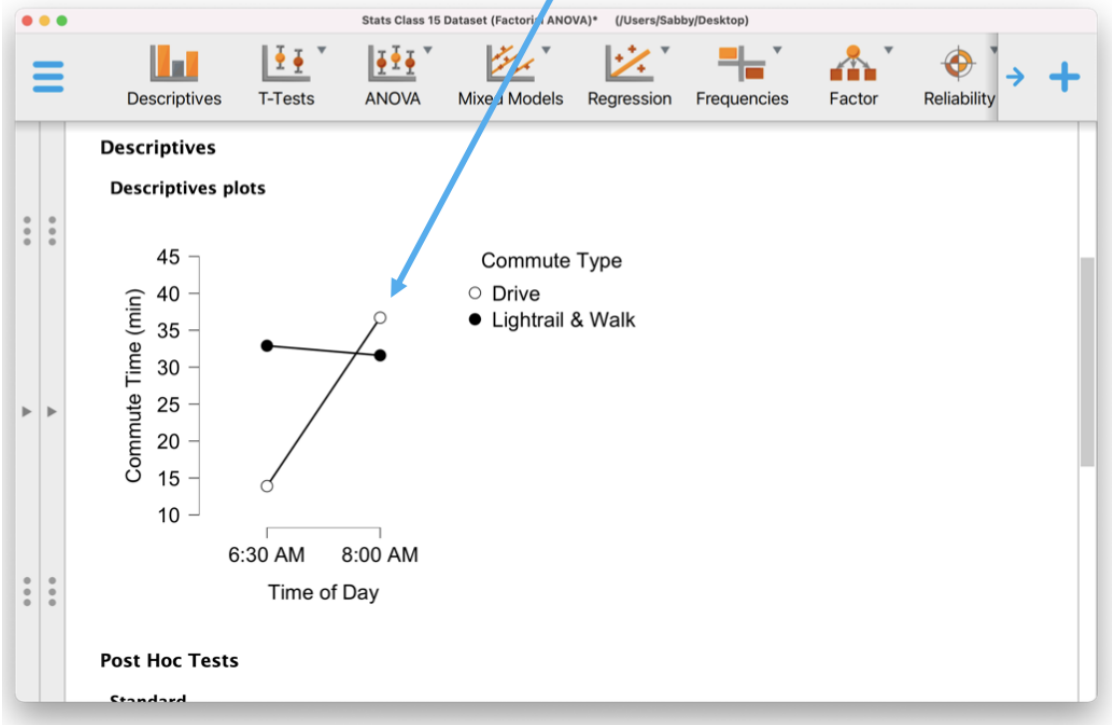
Put Time of Day in the “Horizonal Axis” box and Commute Type in the “Separate Lines” box to mimic the plot we created in class.



12

JASP: Two-way ANOVA

And here's the plot! Note that you can switch where you put the variables and it wouldn't matter, but this replicate our in-class plot.



One-way ANOVA Practice Activity

Next time

Lecture

- Testing relationships using correlations
  - *Guest lecture*: Lindsay Goolsby

Reading

- Chapter 13

Quiz 5

- Due Wednesday 2/23/2022 11:59pm MT
  - Covers Ch.13-14