Applied Statistical Analysis

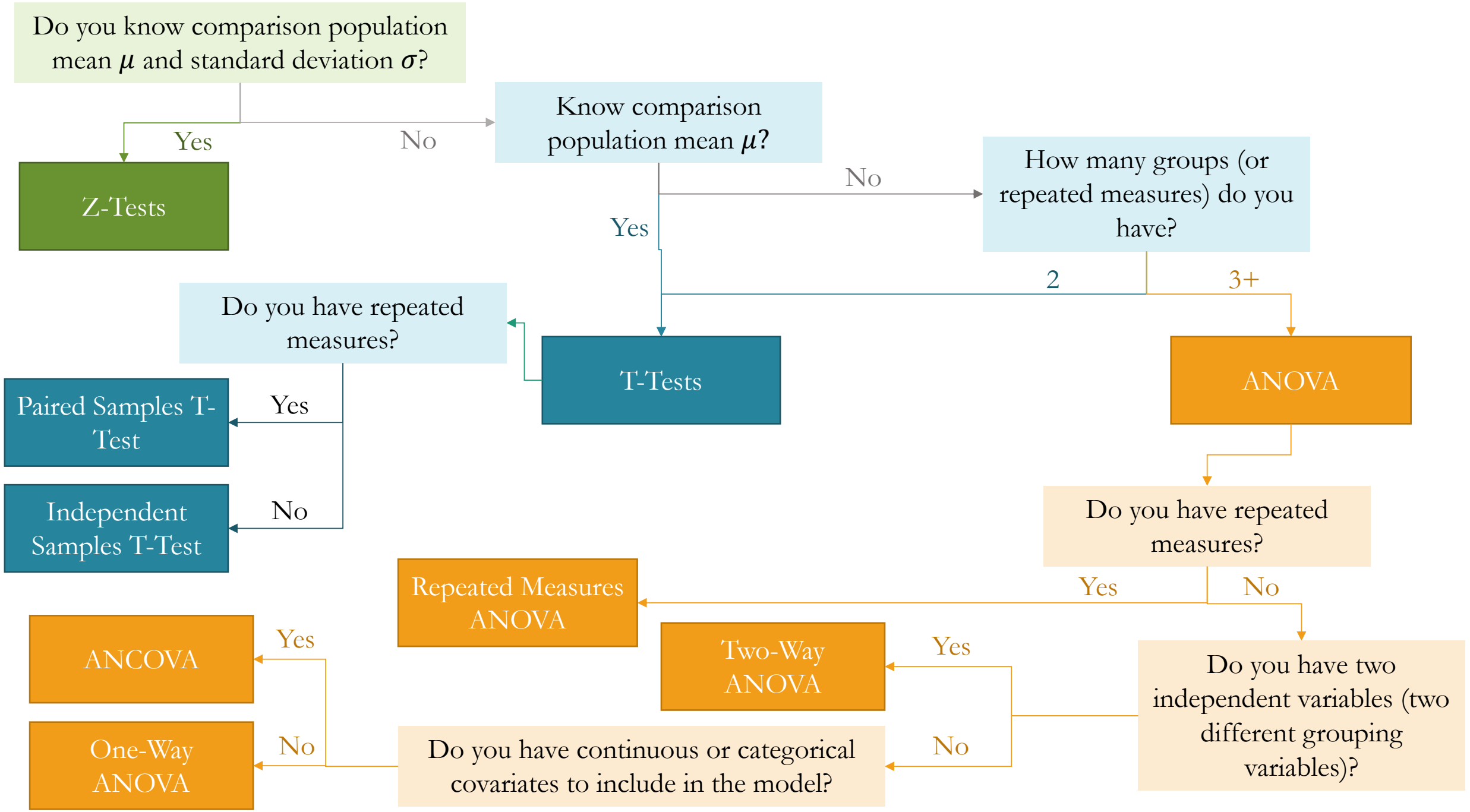
EDUC 6050

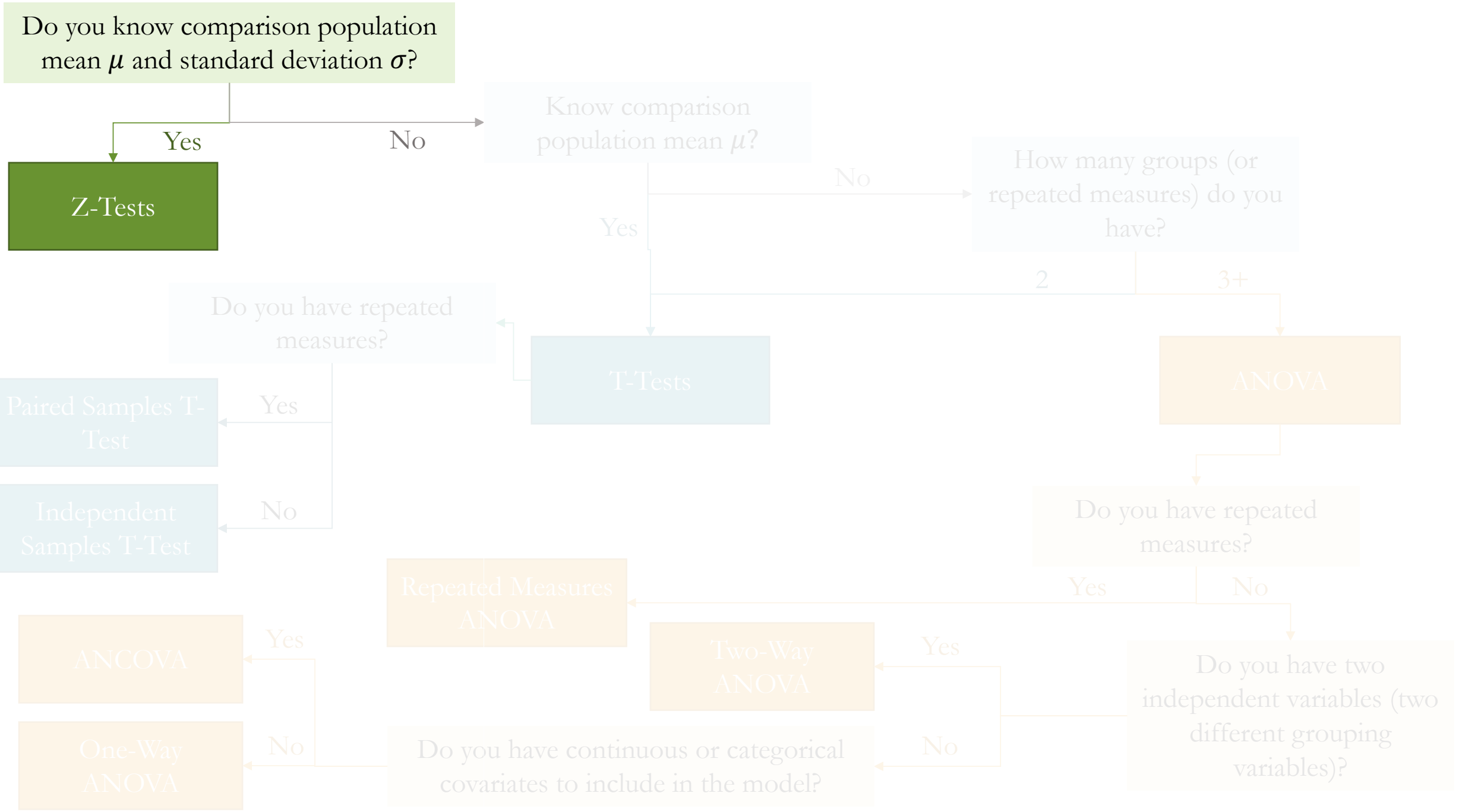
Week 6

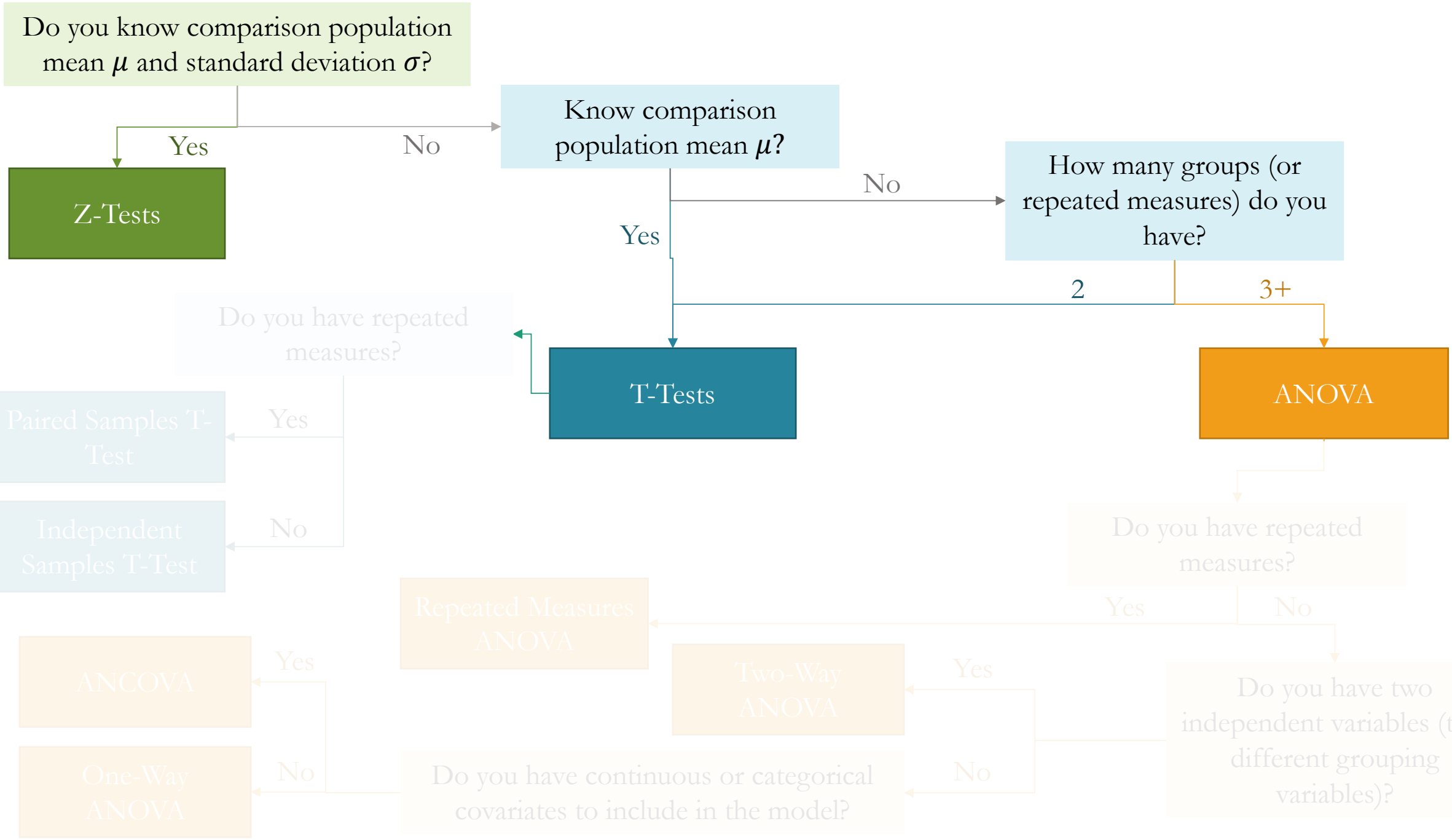
Finding clarity using data

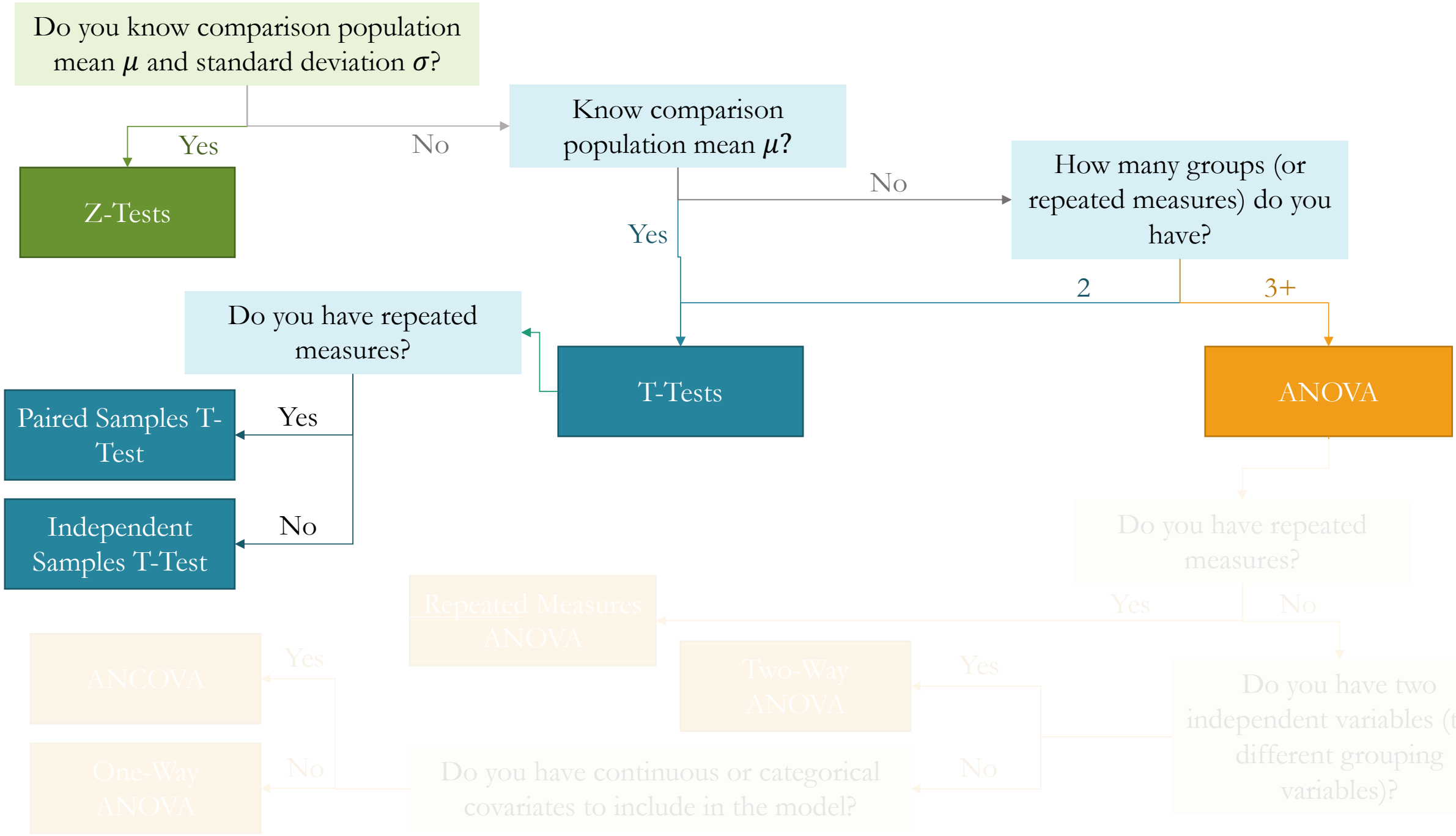
Today

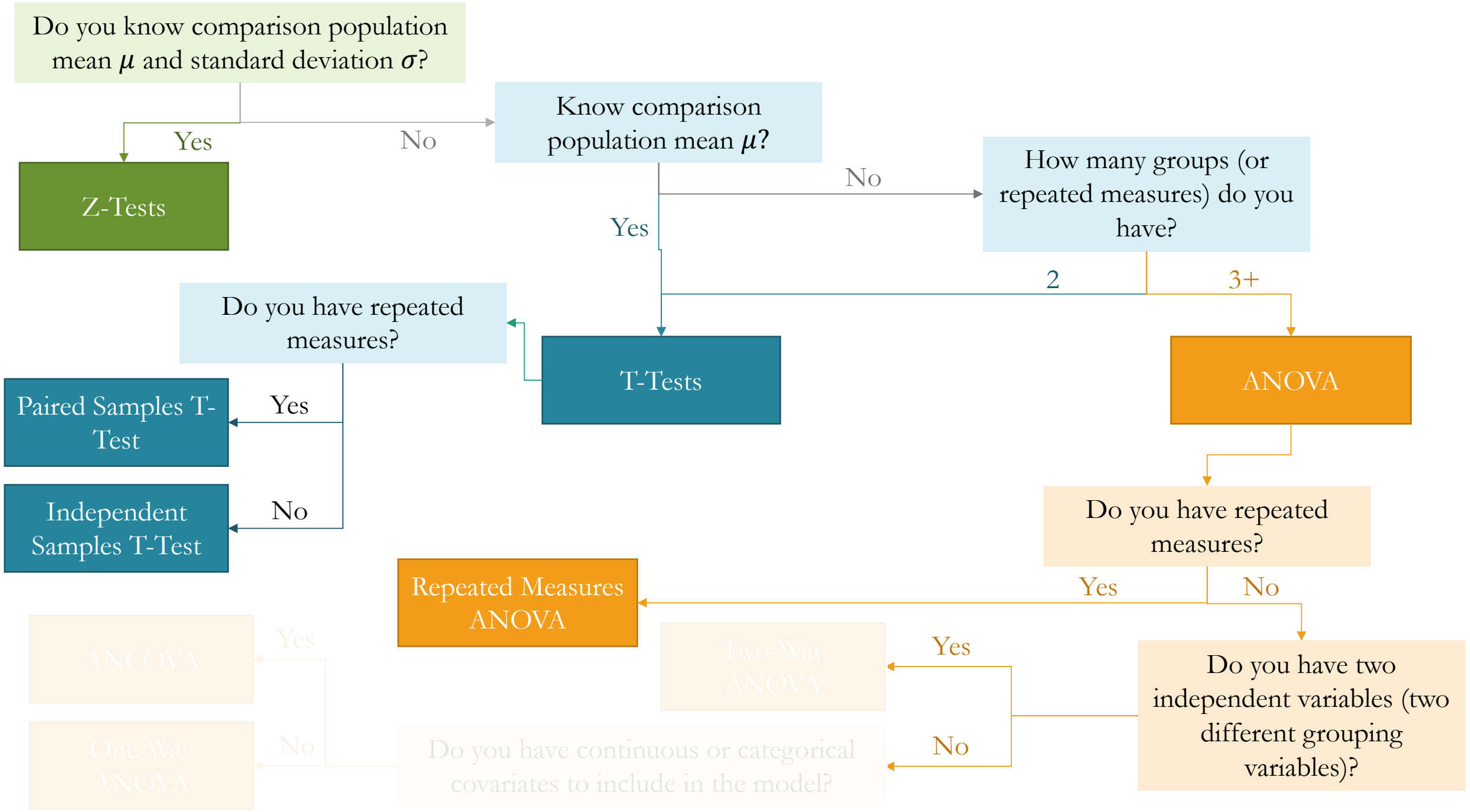
1. **Hypothesis Testing** with ANOVA
   - One-Way
   - Two-Way

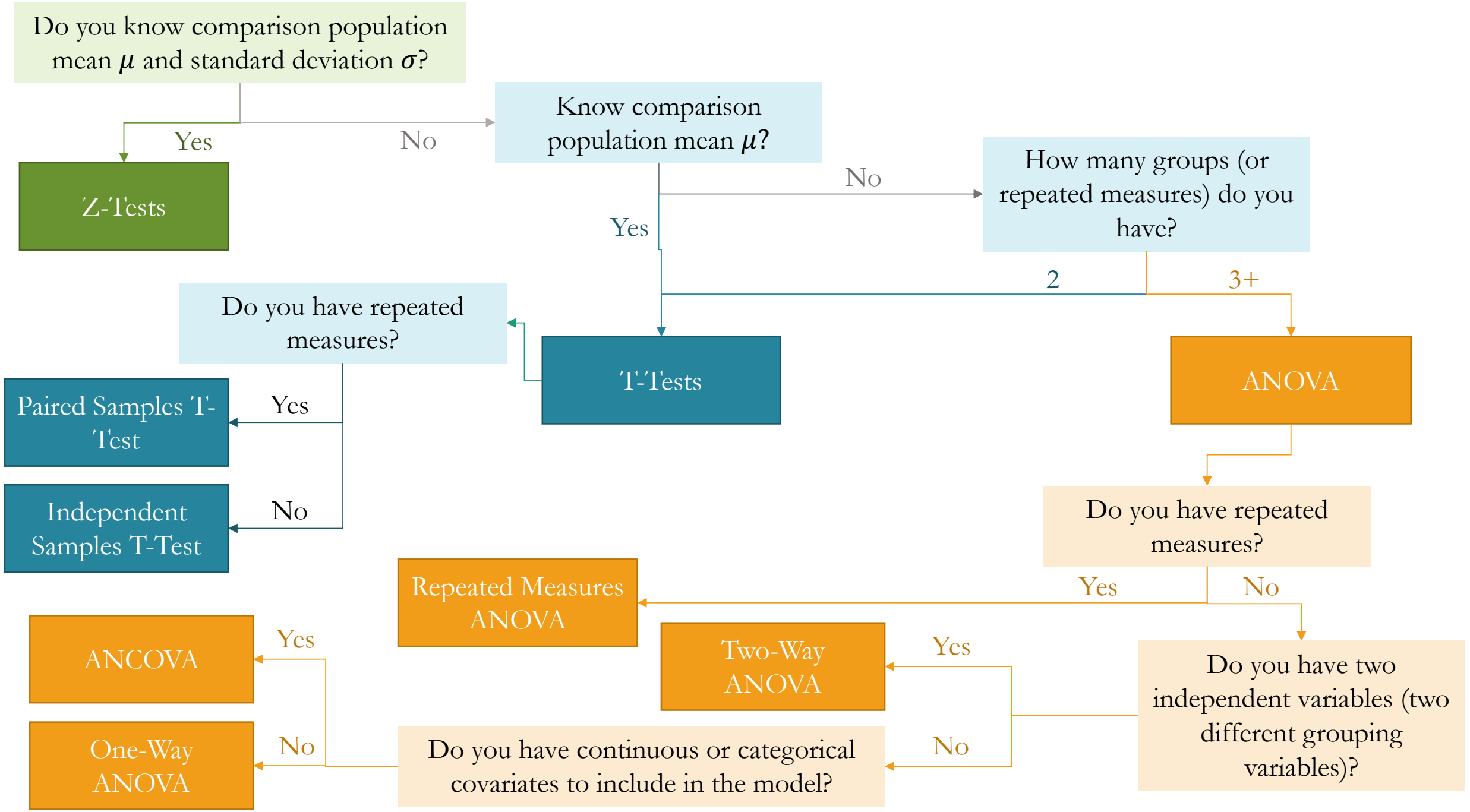

Do you know comparison population mean μ and standard deviation σ?
Yes
No
Z-Tests
Know comparison population mean μ?
No
How many groups (or repeated measures) do you have?
Yes
2
3+
Do you have repeated measures?
T-Tests
ANOVA
Paired Samples T-Test
Yes
Independent Samples T-Test
No
Do you have repeated measures?
Yes
No
Repeated Measures ANOVA
Two-Way ANOVA
Yes
Do you have two independent variables (two different grouping variables)?
ANCOVA
Yes
One-Way ANOVA
No
Do you have continuous or categorical covariates to include in the model?
No


Do you know comparison population mean μ and standard deviation σ?
Yes
No
Z-Tests
Know comparison population mean μ?
No
Yes
How many groups (or repeated measures) do you have?
2
3+
Do you have repeated measures?
T-Tests
ANOVA
Paired Samples T-Test
Yes
Independent Samples T-Test
No
Do you have repeated measures?
Yes
No
Repeated Measures ANOVA
ANCOVA
Yes
Two-Way ANOVA
Yes
Do you have two independent variables (two different grouping variables)?
One-Way ANOVA
No
Do you have continuous or categorical covariates to include in the model?
No


Do you know comparison population mean μ and standard deviation σ?
Yes
No
Z-Tests
Know comparison population mean μ?
No
Yes
How many groups (or repeated measures) do you have?
2
3+
T-Tests
ANOVA
Do you have repeated measures?
Paired Samples T-Test
Yes
Independent Samples T-Test
No
Do you have repeated measures?
Yes
No
Repeated Measures ANOVA
ANCOVA
Yes
One-Way ANOVA
No
Two-Way ANOVA
Yes
Do you have two independent variables (two different grouping variables)?
Do you have continuous or categorical covariates to include in the model?
No


Do you know comparison population mean μ and standard deviation σ?
Yes
No
Z-Tests
Know comparison population mean μ?
No
How many groups (or repeated measures) do you have?
Yes
2
3+
Do you have repeated measures?
T-Tests
ANOVA
Paired Samples T-Test
Yes
Independent Samples T-Test
No
Do you have repeated measures?
Yes
No
Repeated Measures ANOVA
ANCOVA
Yes
Two-Way ANOVA
Yes
Do you have two independent variables (two different grouping variables)?
One-Way ANOVA
No
Do you have continuous or categorical covariates to include in the model?
No


Do you know comparison population mean μ and standard deviation σ?
Yes
No
Z-Tests
Know comparison population mean μ?
No
Yes
How many groups (or repeated measures) do you have?
2
3+
Do you have repeated measures?
T-Tests
Paired Samples T-Test
Yes
Independent Samples T-Test
No
ANOVA
Do you have repeated measures?
Yes
No
Repeated Measures ANOVA
Do you have two independent variables (two different grouping variables)?
Yes
Two-Way ANOVA
No
Do you have continuous or categorical covariates to include in the model?
Yes
ANCOVA
No
One-Way ANOVA

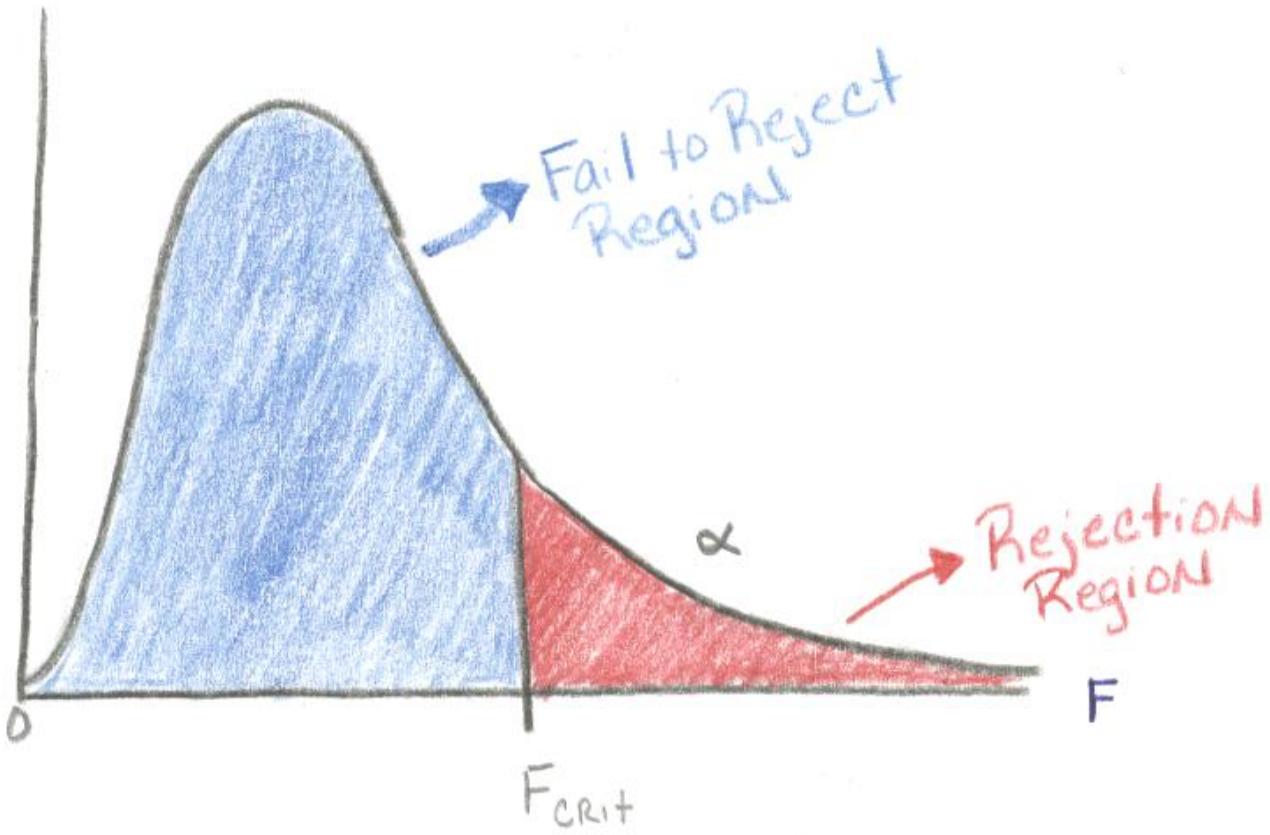

Do you know comparison population mean μ and standard deviation σ?
Yes
No
Z-Tests
Know comparison population mean μ?
No
How many groups (or repeated measures) do you have?
Yes
2
3+
Do you have repeated measures?
T-Tests
ANOVA
Paired Samples T-Test
Yes
Independent Samples T-Test
No
Do you have repeated measures?
Yes
No
Repeated Measures ANOVA
Two-Way ANOVA
Yes
Do you have two independent variables (two different grouping variables)?
ANCOVA
Yes
One-Way ANOVA
No
Do you have continuous or categorical covariates to include in the model?
No

ANalysis Of VAriance

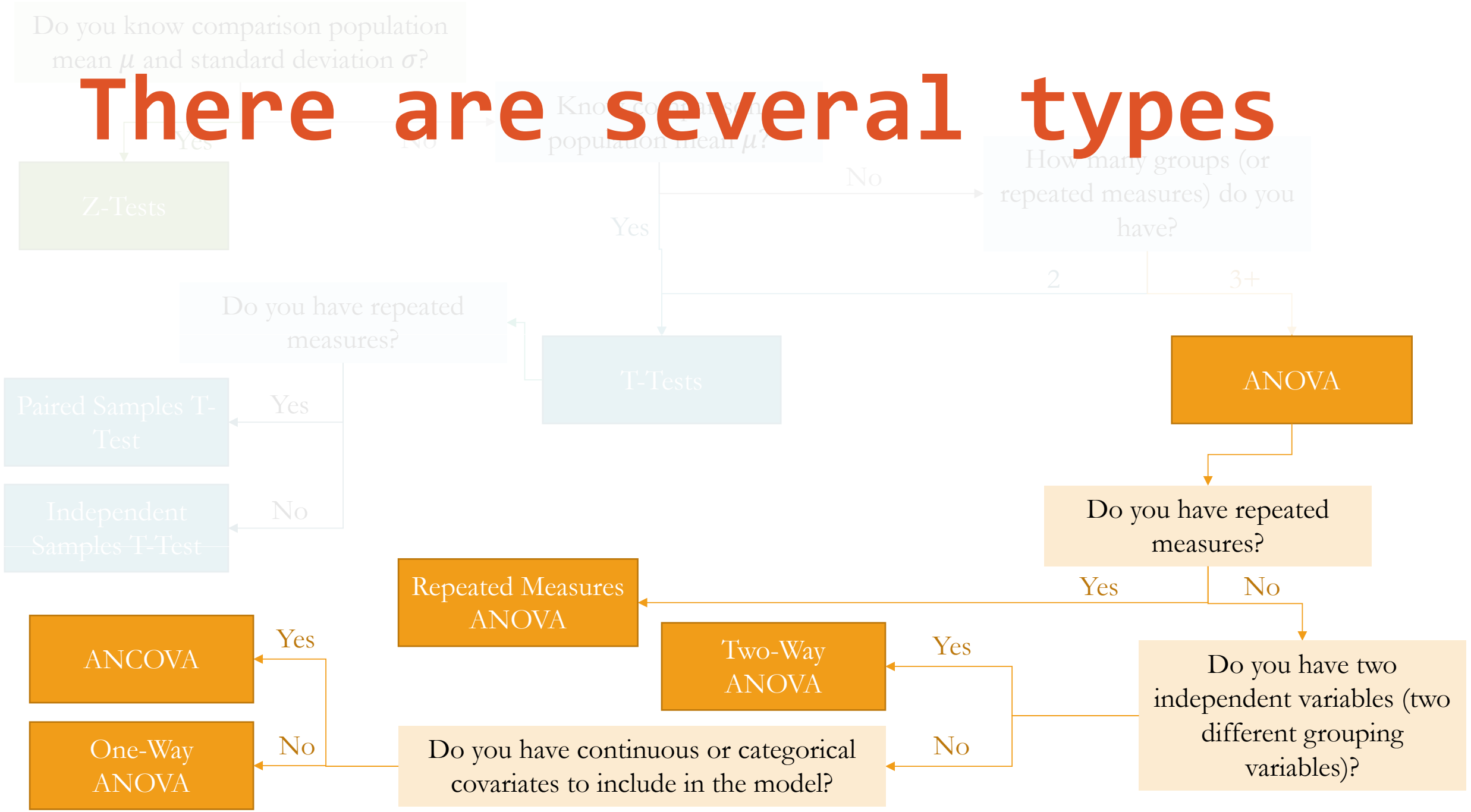
- It is a whole class of methods
- Generally used with experimental designs
- Has similar assumptions to t-test
- Gives an “omnibus” result

	Sum of Squares	df	Mean Square	F	p
Children	26.8	3	8.92	2.38	.087
Residuals	127.5	34	3.75		

The F statistic



There are several types



General Requirements

1. Need a DV on an interval/ratio scale,
2. IV defines 2+ different groups (or time points)

ID	Outcome	Group
1	8	1
2	8	1
3	9	1
4	7	1
5	7	2
6	9	2
7	5	2
8	5	2

General Requirements

1. Need a DV on an interval/ratio scale,
2. IV defines 2+ different groups (or time points)

ID	Time 1	Time 2
1	8	7
2	8	8
3	9	6
4	7	6
5	7	8
6	9	5
7	5	3
8	5	3

Hypothesis Testing with ANOVA

The same 6 step approach!

1. Examine Variables to Assess Statistical Assumptions
2. State the Null and Research Hypotheses (symbolically and verbally)
3. Define Critical Regions
4. Compute the Test Statistic
5. Compute an Effect Size and Describe it
6. Interpreting the results

1 Examine Variables to Assess Statistical Assumptions

Basic Assumptions

1. Independence of data
2. Appropriate measurement of variables for the analysis
3. Normality of distributions
4. Homogeneity of variance

1 Examine Variables to Assess Statistical Assumptions

Basic Assumptions

1. **Independence** of data
2. Appropriate for the al
3. Normality
4. Homogeneity of variance

Individuals are independent of each other (one person's scores does not affect another's)

1 Examine Variables to Assess Statistical Assumptions

Basic Assumptions

1. Independence of data
2. **Appropriate measurement** of variables for the analysis
3. Normality of distributions
4. Homogeneity

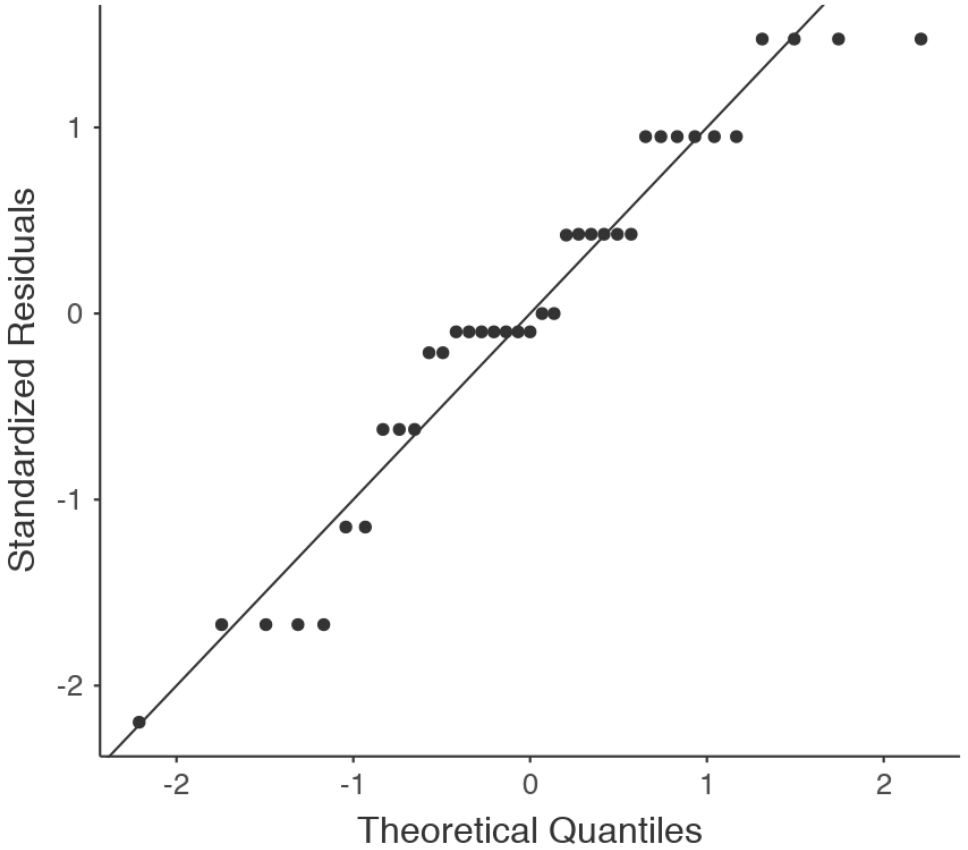
Here we need interval/ratio DV

1 Examine Variables to Assess Statistical Assumptions

Basic Assumptions

1. Inde[illegible]
2. Appr[illegible]pria[illegible]
   for [illegible]he analysis
3. Normality of distributio
4. Homogeneity of variance

Normality of the residuals



1 Examine Variables to Assess Statistical Assumptions

Basic Assumptions

1. Independence of data
2. Appropri… for the…
3. Normalit…
4. **Homogeneity** of variance

F	df1	df2	P
2.86	3	34	.051

The variances of groups should be equal (not strict if each group has similar sample sizes)

1 Examine Variables to Assess Statistical Assumptions

Examining the Basic Assumptions

1. **Independence:** random sample
2. **Appropriate measurement:** know what your variables are
3. **Normality:** Histograms, **Q-Q**, skew and kurtosis
4. **Homogeneity:** Levene's Test

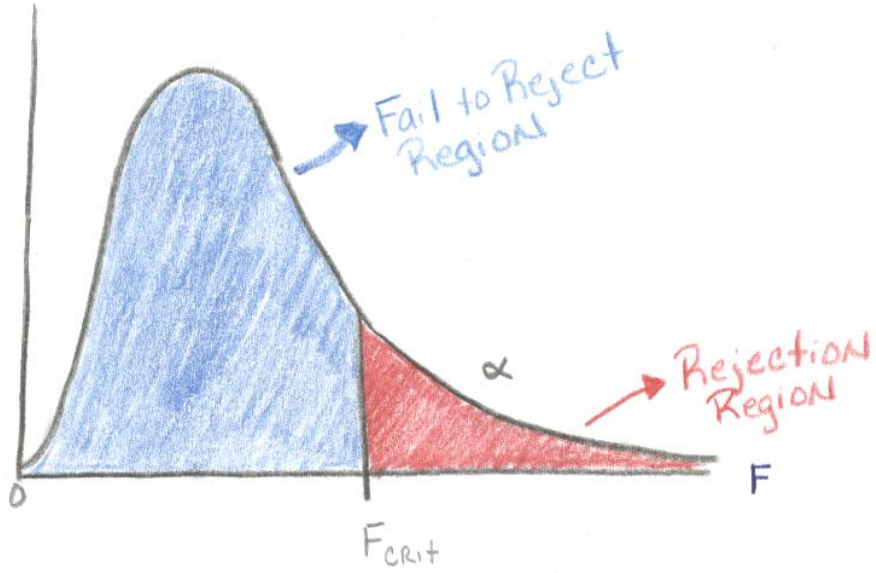
2 State the Null and Research Hypotheses (symbolically and verbally)

Hypothesis Type	Symbolic	Verbal	Difference between means created by:
Research Hypothesis	At least one μ is different than the others	One of the groups' means is different than the others	True differences
Null Hypothesis	All μ's are the same	There is no *real* difference between the groups	Random chance (sampling error)

3 Define Critical Regions

How much evidence is enough to believe the null is not true?

Before analyzing the data, we define the critical regions (generally based on an alpha = .05)



3 Define Critical Regions

We decide on an alpha level first

↳ Then calculate the critical value (based on sample size)

3 Define Critical

We decide on an alpha le

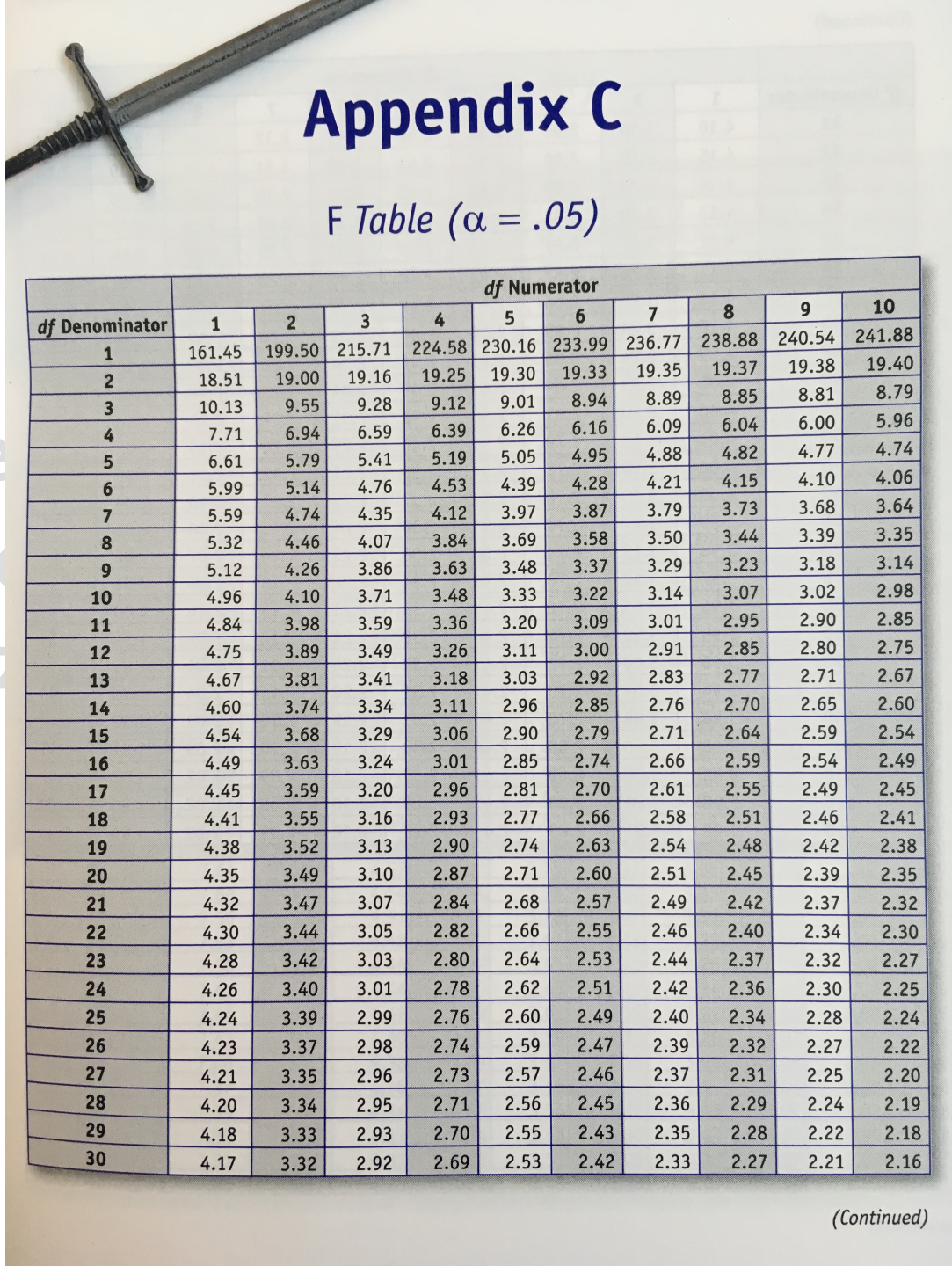
↳ Then calculate the

Use the table in the book

- Base on alpha and 2 specific df's

$df_{num} = g - 1$ where g is number of groups

$df_{den} = N - g$



Appendix C

F *Table* ($\alpha = .05$)

df Denominator	*df* Numerator									
	1	2	3	4	5	6	7	8	9	10
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16

(Continued)

3 Define Critical Regions

We decide on an alpha level first

↳ Then calculate the critical value (based on sample size)

$$F_{critical}(2, 29) = 3.33$$

So our critical region is defined as:

$$\alpha = .05$$

$$F_{critical}(2, 29) = 3.33$$

4 Compute the Test Statistic

Source of Variance	SS	df	MS	F	η_p^2
Between Groups	$n(\sum M^2 - \frac{(\sum M)^2}{g}$	$g - 1$	$\frac{SS_{between}}{df_{between}}$	$\frac{MS_{between}}{MS_{error}}$	$\frac{SS_{between}}{SS_{between} + SS_{residual}}$
Within Groups (Residual)	$\sum SS_{each\ treatment}$	$N - g$	$\frac{SS_{residual}}{df_{residual}}$		
Total	$SS_{each\ treatment} + SS_{residual}$				

4 Compute the Test Statistic

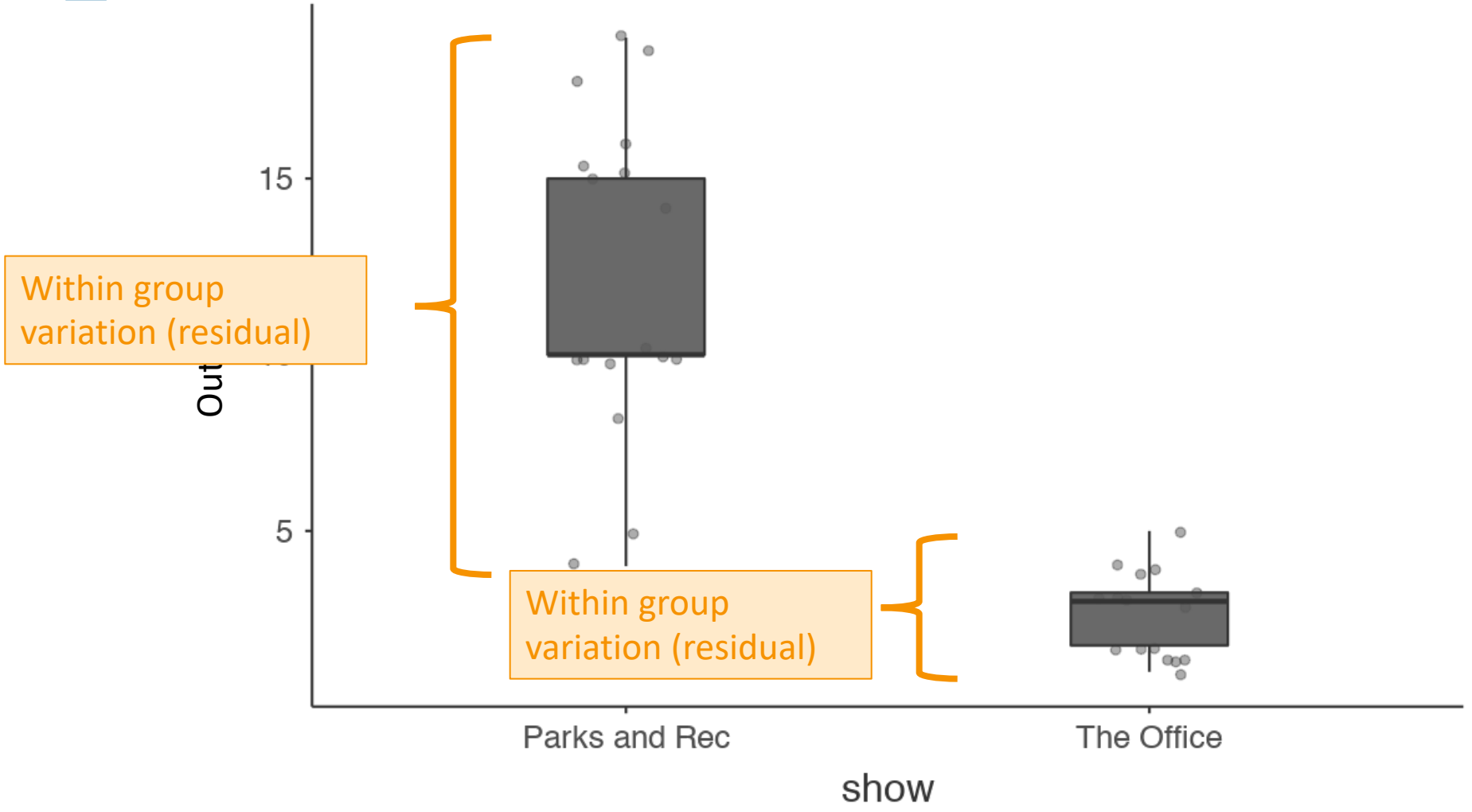
"Sum of Squares" – adding up all the squared deviations (essentially like SD)

Source of Variance	SS	df	MS	F	η_p^2
Between Groups	$n(\sum M^2 - \frac{(\sum M)^2}{g}$	$g - 1$	$\frac{SS_{between}}{df_{between}}$	$\frac{MS_{between}}{MS_{error}}$	$\frac{SS_{between}}{SS_{between} + SS_{residual}}$
Within Groups (Residual)	$\sum SS_{each\ treatment}$	$N - g$	$\frac{SS_{residual}}{df_{residual}}$		
Total	$SS_{each\ treatment} + SS_{residual}$				

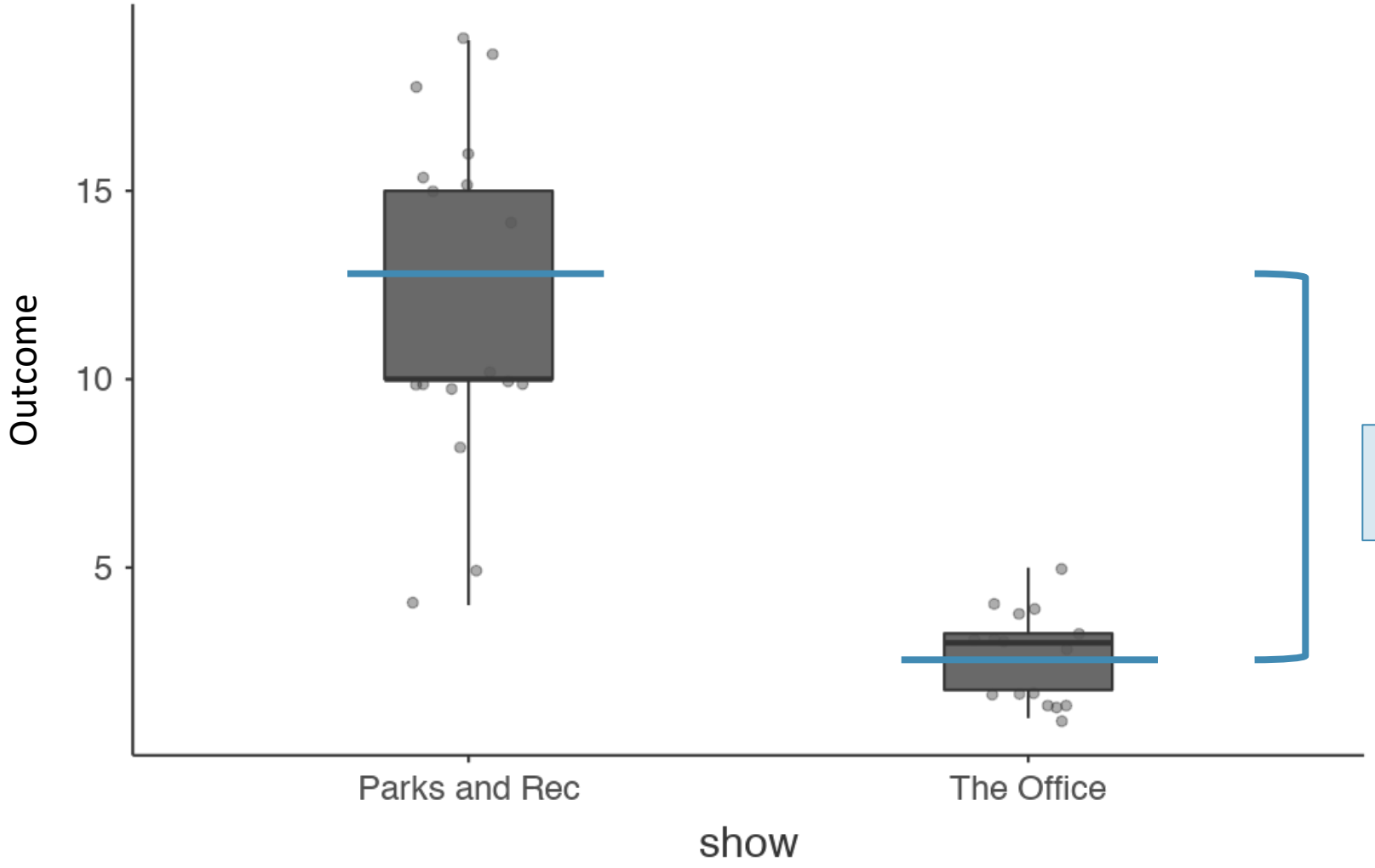
Gets split by where the variation is coming from

- Between the groups (the differences in the groups)
- Within the groups

4 Compute the Test Statistic



4 Compute the Test Statistic



4 Compute the Test Statistic

Source of Variance	SS	df	MS	F	η_p^2
Between Groups	$n(\sum M^2 - \frac{(\sum M)^2}{g}$	$g - 1$	$\frac{SS_{between}}{df_{between}}$	$\frac{MS_{between}}{MS_{error}}$	$\frac{SS_{between}}{SS_{between} + SS_{residual}}$
Within Groups (Residual)	$\sum SS_{each\ treatment}$	$N - g$	$\frac{SS_{residual}}{df_{residual}}$		
Total	$SS_{each\ treatment} + SS_{residual}$				

Is the amount of difference in the means big compared to how much variability there is in the groups?

Answers

These are all ratios

4 Compute the Test Statistic

F-statistic and p-value tell you if one of the groups is different than the others

But it doesn't tell you which ones are different...

Post Hoc Tests

4 Compute the Test Statistic

Post Hoc Tests (or Contrasts)

Post hoc usually refers to comparing all groups with each other (and making an adjustment for the multiple comparisons)

Contrasts usually refers to comparing some of the groups with each other (or a combination of groups with each other)

4 Compute the Test Statistic

Post Hoc Tests (or Contrasts)

Contrasts	Estimate	SE	t	p
1 – 0	-1.19	1.41	-0.841	.406
2 – 0	2.15	1.17	1.835	.075

Post Hoc Comparisons	Estimate	SE	t	p_{tukey}
0 – 1	1.19	1.41	-0.841	.680
0 – 2	-2.15	1.17	-1.835	.174
1 – 2	-3.33	1.77	-1.885	.158

4 Compute the Test Statistic

Post Hoc Tests (or Contrasts)

Contrasts	Estimate	SE	t	p
1 – 0	-1.19	1.41	-0.841	.406
1 – 0	2.15	1.17	1.835	.075

Post Hoc Comparisons	Estimate	SE	t	p_{tukey}
0 – 1	1.19	1.41	-0.841	.680
0 – 2	-2.15	1.17	-1.835	.174
1 – 2	-3.33	1.77	-1.885	.158

4 Compute the Test Statistic

Source of Variance	SS	df	MS	F	η_p^2
Between Groups	$n(\sum M^2 - \frac{(\sum M)^2}{g}$	$g - 1$	$\frac{SS_{between}}{df_{between}}$	$\frac{MS_{between}}{MS_{error}}$	$\frac{SS_{between}}{SS_{between} + SS_{residual}}$
Within Groups (Residual)	$\sum SS_{each\ treatment}$	$N - g$	$\frac{SS_{residual}}{df_{residual}}$		
Total	$SS_{each\ treatment} + SS_{residual}$				Answers

How much of the variability is accounted for by the groups vs everything else?

5 Compute an Effect Size and Describe it

One of the main effect sizes for ANOVA is "Eta Squared"

$$\eta_p^2 = \frac{SS_{between}}{SS_{between} + SS_{residual}}$$

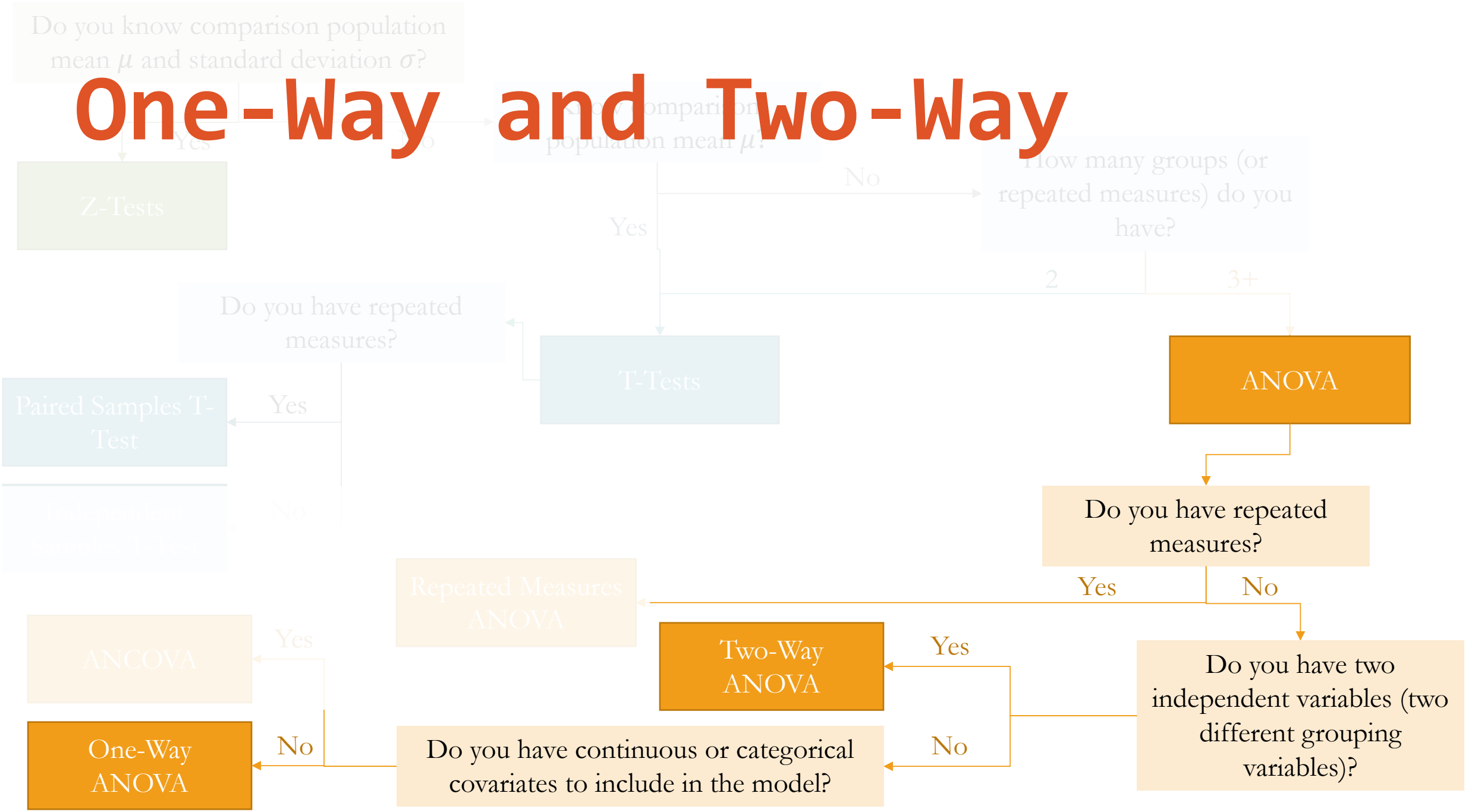
η_p^2	Estimated Size of the Effect
Close to .01	Small
Close to .06	Moderate
Close to .14	Large

6 Interpreting the results

Put your results into words

Use the example around page 382 as a template

One-Way and Two-Way



One-Way vs. Two-Way

One-way ANOVA has one predictor

Tests for any differences across the groups on one predictor

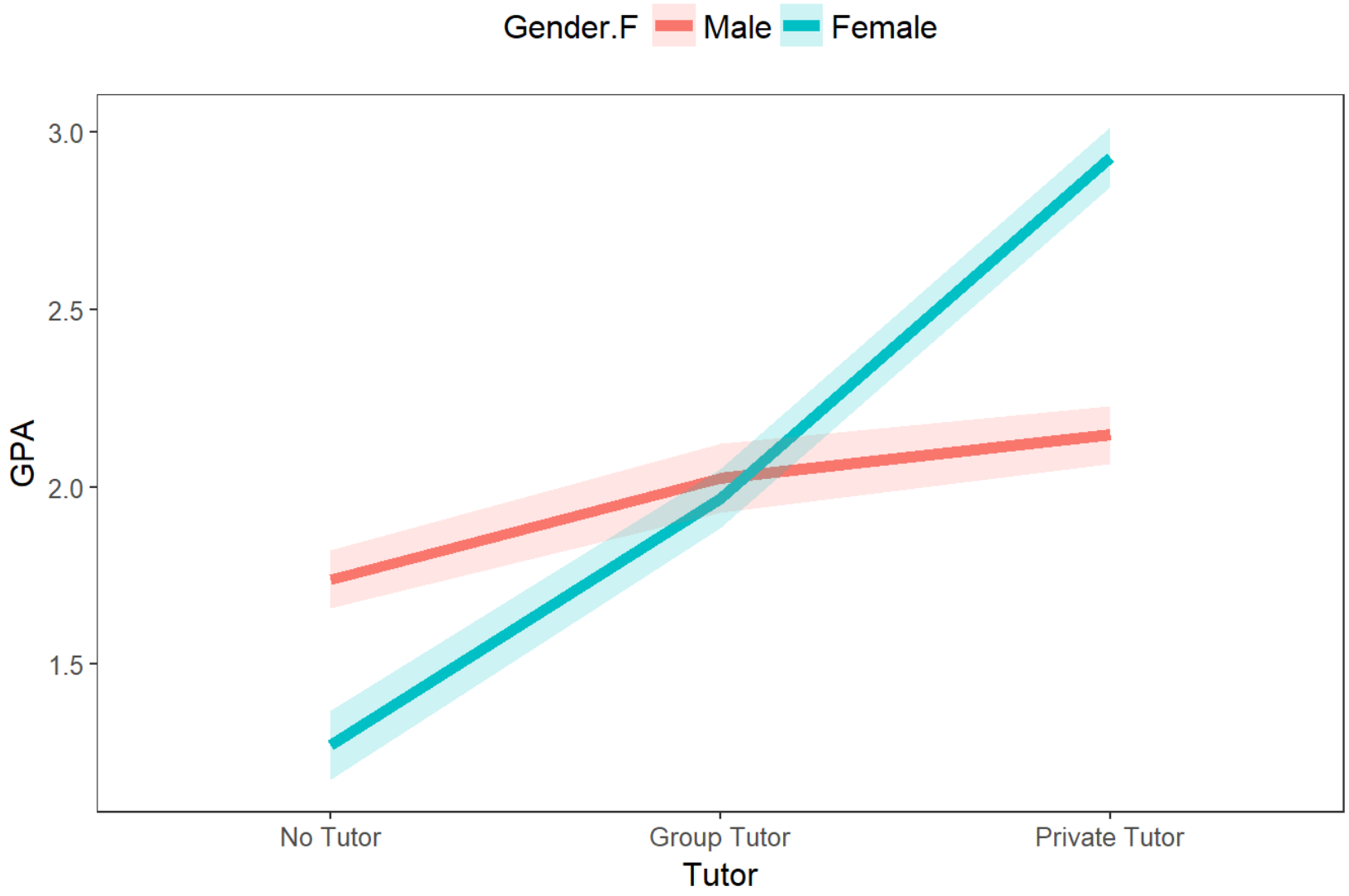
Two-way ANOVA has two predictors

Tests for any differences across the groups for both predictors (and their combinations)

"Interaction"

Interactions

When the effect of a predictor depends on another



Questions?

Please post them to the discussion board before class starts

End of Pre-Recorded Lecture Slides

In-class discussion slides

https://youtu.be/h4MhbkWJzKk

Application

Example Using
The Office/Parks and Rec Data Set

Hypothesis Test with ANOVA (One-Way and Two-Way)