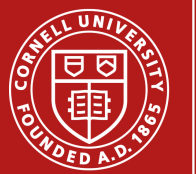


Johnson
Cornell
SC Johnson College of Business

Research Methods for the 2020s:

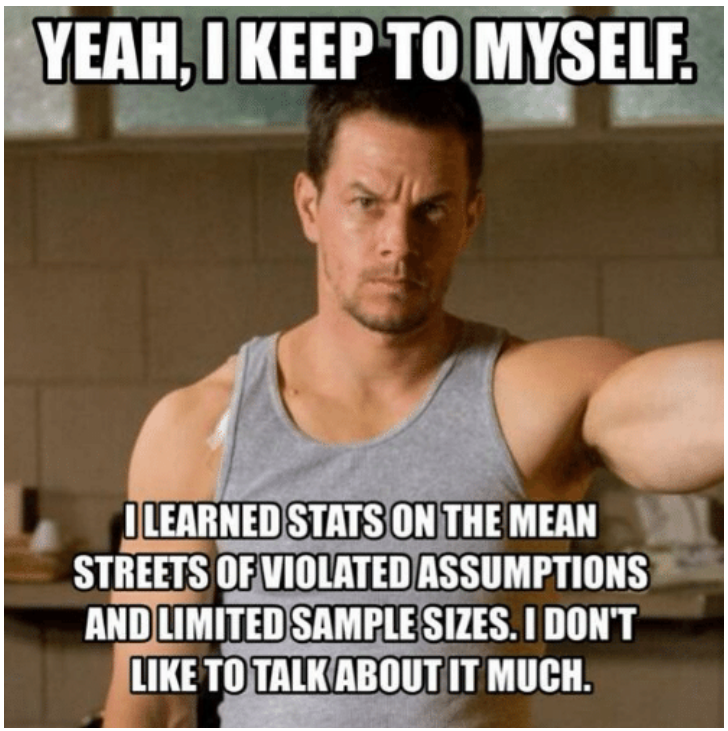
Experimental Design and Analysis for the Discerning Modern Behavioralist

Ryan Guggenmos

2020 AAA ABO Midyear Meeting PhD Consortium - My Basement

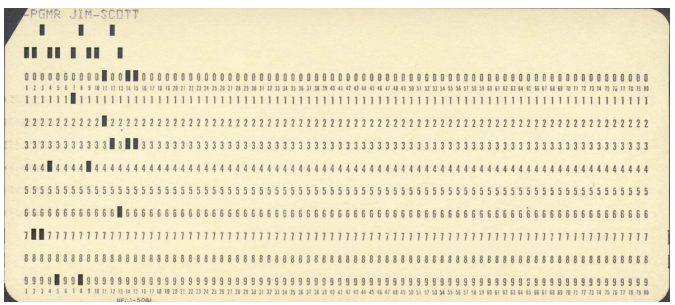
What exactly are "Research Methods for the 2020s?"

Gifs and Memes?

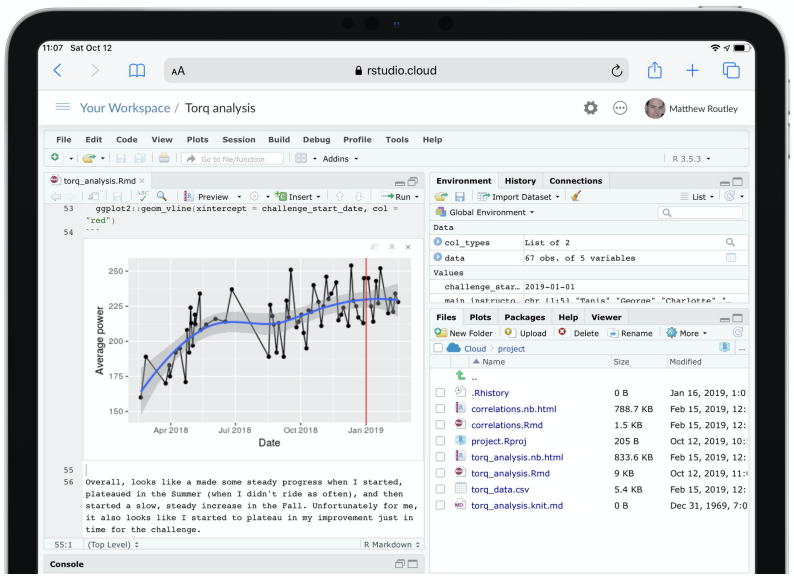


Methods for the 2020s

Advances in methodology are being driven by advances in computing and other technologies.



vs.



This gives us more options than ever, so what do you need to know?

What can methods do for us?

Strong experimental design and analyses, allow us to:

- Makes inferences about causal relationships
- Learn about how, why, and when an effect may occur
- Test theories about outcomes and processes

What can methods do for us?

Strong experimental design and analyses, allow us to:

- Makes inferences about causal relationships
- Learn about how, why, and when an effect may occur
- Test theories about outcomes and processes

Leverage the competitive advantages of your research method(s).

What can methods do for us?

Strong experimental design and analyses, allow us to:

- Makes inferences about causal relationships
- Learn about how, why, and when an effect may occur
- Test theories about outcomes and processes

Leverage the competitive advantages of your research method(s).

Design and analysis issues can keep us from being able to capitalize on our competitive advantage.

Inferences from these experiments may be incorrect or incomplete.

What can't analysis do for us?

What can't analysis do for us?

- Correct for poor experimental design

What can't analysis do for us?

- Correct for poor experimental design
- Allow you to draw strong casual inference from observational data

What can't analysis do for us?

- Correct for poor experimental design
- Allow you to draw strong casual inference from observational data
- Tell us much about magnitudes

What can't analysis do for us?

- Correct for poor experimental design
- Allow you to draw strong casual inference from observational data
- Tell us much about magnitudes
- Get your paper accepted (but it can help it not get rejected)

What can't analysis do for us?

- Correct for poor experimental design
- Allow you to draw strong casual inference from observational data
- Tell us much about magnitudes
- Get your paper accepted (but it can help it not get rejected)
- Make you more popular at parties

What can't analysis do for us?

- Correct for poor experimental design
- Allow you to draw strong casual inference from observational data
- Tell us much about magnitudes
- Get your paper accepted (but it can help it not get rejected)
- Make you more popular at parties
- Show you "truth" in the world

What can't analysis do for us?

- Correct for poor experimental design
- Allow you to draw strong casual inference from observational data
- Tell us much about magnitudes
- Get your paper accepted (but it can help it not get rejected)
- Make you more popular at parties
- Show you "truth" in the world
- Make your variables of interest interesting (esp. jointly interesting)

What can't analysis do for us?

- Correct for poor experimental design
- Allow you to draw strong casual inference from observational data
- Tell us much about magnitudes
- Get your paper accepted (but it can help it not get rejected)
- Make you more popular at parties
- Show you "truth" in the world
- Make your variables of interest interesting (esp. jointly interesting)
- Give you an interesting accounting research question*

* with some small exceptions (e.g., audit sampling, methods papers)


IBM
Presents
You Make The Call

You make the call!

H1: The difference in investors' willingness to invest under low versus high levels of stress will be greater when the humidity is high versus when it is low.

Is H1 supported?

You make the call!

H1: The difference in investors' willingness to invest under low versus high levels of stress will be greater when the humidity is high versus when it is low.

ANOVA	
Stress	$p = 0.56$
Humidity	$p = 0.34$
Stress x Humidity	$p = 0.21$

Is H1 supported?

You make the call!

H1: The difference in investors' willingness to invest under low versus high levels of stress will be greater when the humidity is high versus when it is low.

ANOVA	
Stress	$p = 0.56$
Humidity	$p = 0.34$
Stress x Humidity	$p = 0.21$

Tests of follow-up simple effects:

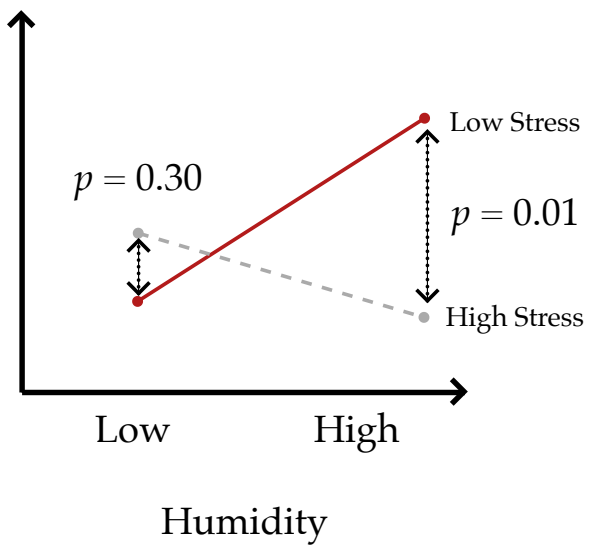
The simple effect of stress given low humidity is not significant ($p = 0.30$). The simple effect of stress given high humidity is highly significant ($p = 0.01$).

Is H1 supported?

You make the call!

H1: The difference in investors' willingness to invest under low versus high levels of stress will be greater when the humidity is high versus when it is low.

ANOVA	
Stress	$p = 0.56$
Humidity	$p = 0.34$
Stress x Humidity	$p = 0.21$



Tests of follow-up simple effects:

The simple effect of stress given low humidity is not significant ($p = 0.30$). The simple effect of stress given high humidity is highly significant ($p = 0.01$).

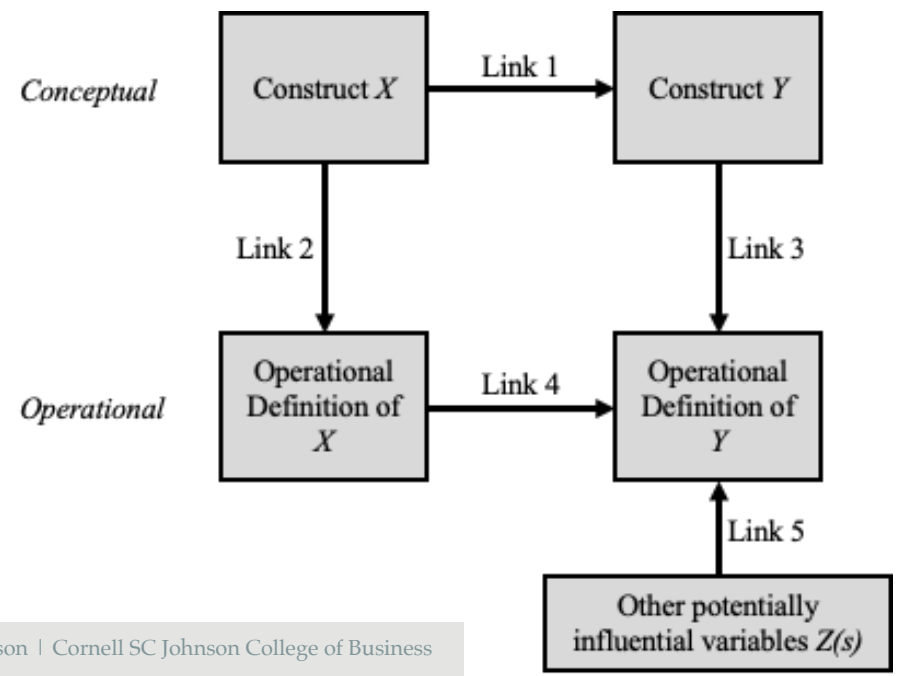
Is H1 supported?

Methods: Design and Analysis

Predictive Validity Framework (a.k.a. Libby Boxes)

"Link 4 assesses the relations between the operational independent and dependent variables."

Libby, Bloomfield, and Nelson (2002)



Using Stats!

The missing link...

Link 4 is Statistical Conclusion Validity (SCV).

Statistical Conclusion Validity (SCV) requires that:

Shadish, Cook, and Campbell (2002)

The missing link...

Link 4 is Statistical Conclusion Validity (SCV).

Statistical Conclusion Validity (SCV) requires that:

- the statistical analysis chosen matches the design employed; and

Shadish, Cook, and Campbell (2002)

The missing link...

Link 4 is Statistical Conclusion Validity (SCV).

Statistical Conclusion Validity (SCV) requires that:

- the statistical analysis chosen matches the design employed; and
- the analysis is applied in a way that does not distort the expected probability of Type I or Type II error.

Shadish, Cook, and Campbell (2002)

The missing link...

Link 4 is Statistical Conclusion Validity (SCV).

Statistical Conclusion Validity (SCV) requires that:

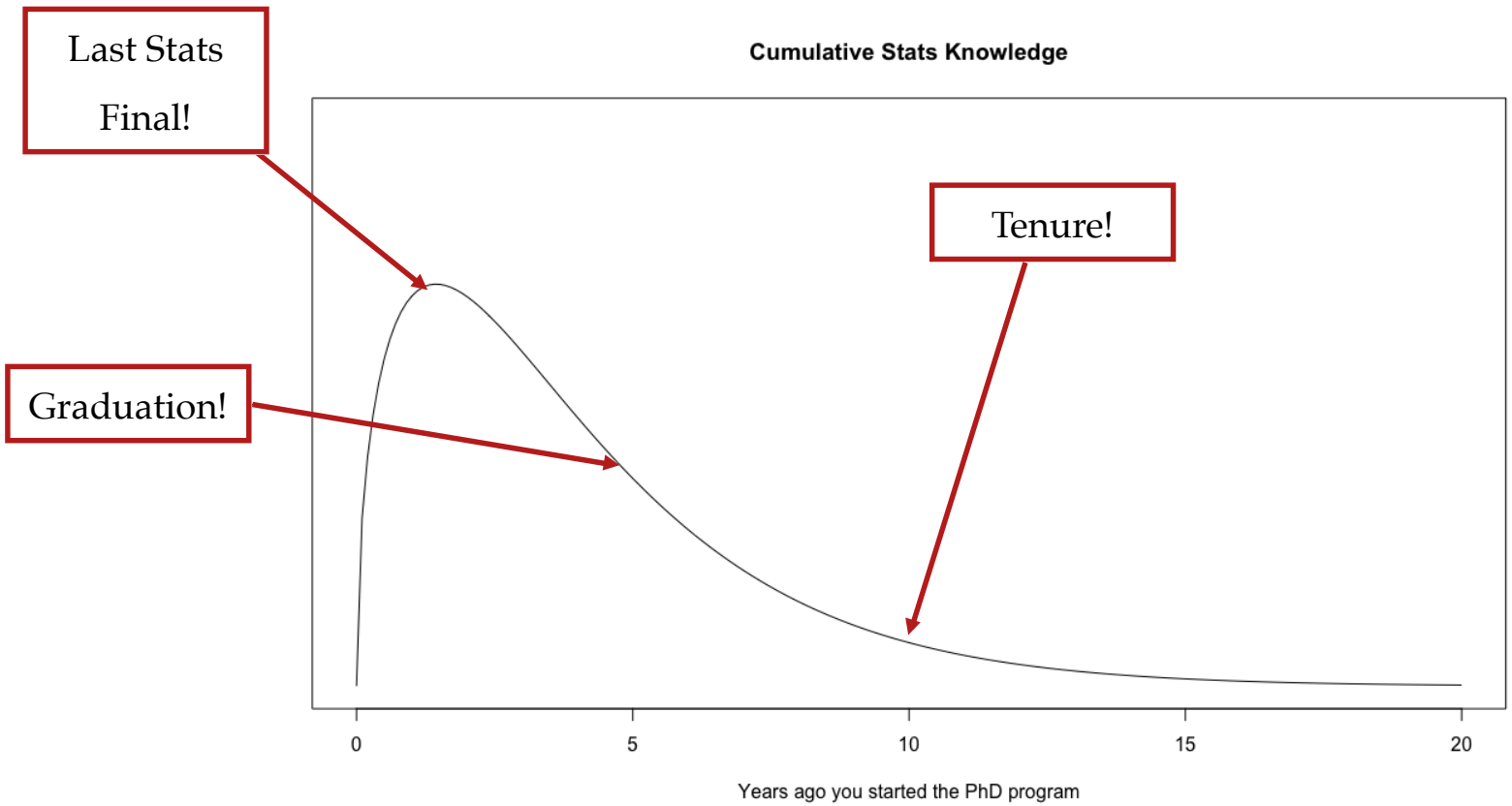
- the statistical analysis chosen matches the design employed; and
- the analysis is applied in a way that does not distort the expected probability of Type I or Type II error.

Without adequate SCV, we might not be able to trust our results.

This is a really big deal, but how do we know whether we've met these requirements?

Shadish, Cook, and Campbell (2002)

Don't be this person - #1



Read the plan and instructions!

First, know your data.

Read the plan and instructions!

First, know your data.

- What am I trying to show with these data?

Read the plan and instructions!

First, know your data.

- What am I trying to show with these data?
- How is my data distributed? Is this as expected?

Read the plan and instructions!

First, know your data.

- What am I trying to show with these data?
- How is my data distributed? Is this as expected?
- Do I have outliers, missing data, or other issues?

Read the plan and instructions!

First, know your data.

- What am I trying to show with these data?
- How is my data distributed? Is this as expected?
- Do I have outliers, missing data, or other issues?
- What kind of variables do I have? (fixed vs. random, ordinal, etc.)

Read the plan and instructions!

First, know your data.

- What am I trying to show with these data?
- How is my data distributed? Is this as expected?
- Do I have outliers, missing data, or other issues?
- What kind of variables do I have? (fixed vs. random, ordinal, etc.)

What do I need to know about the technique?

Read the plan and instructions!

First, know your data.

- What am I trying to show with these data?
- How is my data distributed? Is this as expected?
- Do I have outliers, missing data, or other issues?
- What kind of variables do I have? (fixed vs. random, ordinal, etc.)

What do I need to know about the technique?

- What are the test's assumptions?

Read the plan and instructions!

First, know your data.

- What am I trying to show with these data?
- How is my data distributed? Is this as expected?
- Do I have outliers, missing data, or other issues?
- What kind of variables do I have? (fixed vs. random, ordinal, etc.)

What do I need to know about the technique?

- What are the test's assumptions?
- What happens if those assumptions are violated?

Read the plan and instructions!

First, know your data.

- What am I trying to show with these data?
- How is my data distributed? Is this as expected?
- Do I have outliers, missing data, or other issues?
- What kind of variables do I have? (fixed vs. random, ordinal, etc.)

What do I need to know about the technique?

- What are the test's assumptions?
- What happens if those assumptions are violated?
- From a high level, how does the technique work?

Read the plan and instructions!

First, know your data.

- What am I trying to show with these data?
- How is my data distributed? Is this as expected?
- Do I have outliers, missing data, or other issues?
- What kind of variables do I have? (fixed vs. random, ordinal, etc.)

What do I need to know about the technique?

- What are the test's assumptions?
- What happens if those assumptions are violated?
- From a high level, how does the technique work?
- How is the test conducted in my software package? Defaults?

Read the plan and instructions!

First, know your data.

- What am I trying to show with these data?
- How is my data distributed? Is this as expected?
- Do I have outliers, missing data, or other issues?
- What kind of variables do I have? (fixed vs. random, ordinal, etc.)

What do I need to know about the technique?

- What are the test's assumptions?
- What happens if those assumptions are violated?
- From a high level, how does the technique work?
- How is the test conducted in my software package? Defaults?
- What does the test tell me?

Read the plan and instructions!

First, know your data.

- What am I trying to show with these data?
- How is my data distributed? Is this as expected?
- Do I have outliers, missing data, or other issues?
- What kind of variables do I have? (fixed vs. random, ordinal, etc.)

What do I need to know about the technique?

- What are the test's assumptions?
- What happens if those assumptions are violated?
- From a high level, how does the technique work?
- How is the test conducted in my software package? Defaults?
- What does the test tell me?
- Are there other tests that I should consider?

Must know: Core Analyses

Must know: Core Analyses

- ANOVA
- Regression

Must know: Core Analyses

- ANOVA
- Regression

- Simple Effects Tests
- Simple comparisons between cells
- Basic Contrast Testing

Must know: Core Analyses

- ANOVA
- Regression

- Simple Effects Tests
- Simple comparisons between cells
- Basic Contrast Testing

- Descriptive Statistics / Data visualization
- Common non-parametric tests (e.g., Mann-Whitney U, χ^2)

Must know: Core Analyses

- ANOVA
- Regression

- Simple Effects Tests
- Simple comparisons between cells
- Basic Contrast Testing

- Descriptive Statistics / Data visualization
- Common non-parametric tests (e.g., Mann-Whitney U, χ^2)

Plus: Any technique used in a paper authored or co-authored by you. This goes double for your JMP.

Don't be this person - #2

Cite a (methods) paper, without reading (and understanding) the cited paper.

Don't be this person - #2

Cite a (methods) paper, without reading (and understanding) the cited paper.

Example:

Don't be this person - #2

Cite a (methods) paper, without reading (and understanding) the cited paper.

Example:

1. You read "Paper 1."

Don't be this person - #2

Cite a (methods) paper, without reading (and understanding) the cited paper.

Example:

1. You read "Paper 1."
2. Paper 1 cites "Methods Paper A."

Don't be this person - #2

Cite a (methods) paper, without reading (and understanding) the cited paper.

Example:

1. You read "Paper 1."
2. Paper 1 cites "Methods Paper A."
3. You use the same technique as "Paper 1", so you cite "Methods Paper A" without reading it.

Don't be this person - #2

Cite a (methods) paper, without reading (and understanding) the cited paper.

Example:

1. You read "Paper 1."
2. Paper 1 cites "Methods Paper A."
3. You use the same technique as "Paper 1", so you cite "Methods Paper A" without reading it.
4. 

Don't be this person - #2

Cite a (methods) paper, without reading (and understanding) the cited paper.

Example:

1. You read "Paper 1."
2. Paper 1 cites "Methods Paper A."
3. You use the same technique as "Paper 1", so you cite "Methods Paper A" without reading it.
4. 

This isn't a good look for you and could be interpreted as a "signal" by the reviewer.

Should know: Common Analyses

Should know: Common Analyses

- ANOVA with random effects (e.g., within-SS)
- ANCOVA
- MANOVA
- Non-linear regression techniques

Should know: Common Analyses

- ANOVA with random effects (e.g., within-SS)
- ANCOVA
- MANOVA
- Non-linear regression techniques

- Custom Contrast Testing

Should know: Common Analyses

- ANOVA with random effects (e.g., within-SS)
- ANCOVA
- MANOVA
- Non-linear regression techniques

- Custom Contrast Testing

- Factor Analysis (EFA, CFA)
- Principal Components Analysis (PCA)
- Rotations (e.g. varimax vs. oblimax)

Should know: Common Analyses

- ANOVA with random effects (e.g., within-SS)
- ANCOVA
- MANOVA
- Non-linear regression techniques

- Custom Contrast Testing

- Factor Analysis (EFA, CFA)
- Principal Components Analysis (PCA)
- Rotations (e.g. varimax vs. oblimax)

These are different!

Should know: Common Analyses

- ANOVA with random effects (e.g., within-SS)
- ANCOVA
- MANOVA
- Non-linear regression techniques

- Custom Contrast Testing

- Factor Analysis (EFA, CFA)
- Principal Components Analysis (PCA)
- Rotations (e.g. varimax vs. oblimax)

These are different!

- Basic Structural Equation Modeling (SEM)
- Basic Simultaneous Regression (PROCESS)
- Basic Bayesian null effect testing

Good to know: Other Topics

These are good to know if they are used in your topical area (or for generally being a good consumer of research).

Good to know: Other Topics

These are good to know if they are used in your topical area (or for generally being a good consumer of research).

- Conditional Indirect Effects (i.e., Advanced PROCESS)

Good to know: Other Topics

These are good to know if they are used in your topical area (or for generally being a good consumer of research).

- Conditional Indirect Effects (i.e., Advanced PROCESS)
- Analysis of incomplete and other complex designs

Good to know: Other Topics

These are good to know if they are used in your topical area (or for generally being a good consumer of research).

- Conditional Indirect Effects (i.e., Advanced PROCESS)
- Analysis of incomplete and other complex designs
- Mixed design analysis (both random and fixed effects)

Good to know: Other Topics

These are good to know if they are used in your topical area (or for generally being a good consumer of research).

- Conditional Indirect Effects (i.e., Advanced PROCESS)
- Analysis of incomplete and other complex designs
- Mixed design analysis (both random and fixed effects)
- Hierarchical design analysis (HLM and nested analysis)

Good to know: Other Topics

These are good to know if they are used in your topical area (or for generally being a good consumer of research).

- Conditional Indirect Effects (i.e., Advanced PROCESS)
- Analysis of incomplete and other complex designs
- Mixed design analysis (both random and fixed effects)
- Hierarchical design analysis (HLM and nested analysis)
- Survey Analysis Techniques (e.g., stratification)

Good to know: Other Topics

These are good to know if they are used in your topical area (or for generally being a good consumer of research).

- Conditional Indirect Effects (i.e., Advanced PROCESS)
- Analysis of incomplete and other complex designs
- Mixed design analysis (both random and fixed effects)
- Hierarchical design analysis (HLM and nested analysis)
- Survey Analysis Techniques (e.g., stratification)
- Archival techniques:
  - Fixed Effects and Standard Error Structures
  - Propensity Score Matching
  - Diff-and-Diff designs

Good to know: Other Topics

These are good to know if they are used in your topical area (or for generally being a good consumer of research).

- Conditional Indirect Effects (i.e., Advanced PROCESS)
- Analysis of incomplete and other complex designs
- Mixed design analysis (both random and fixed effects)
- Hierarchical design analysis (HLM and nested analysis)
- Survey Analysis Techniques (e.g., stratification)
- Archival techniques:
  - Fixed Effects and Standard Error Structures
  - Propensity Score Matching
  - Diff-and-Diff designs
- Bayesian data analysis

Remember!

As amazing as learning stats is, you are (probably) not a statistician.

Remember!

As amazing as learning stats is, you are (probably) not a statistician.

You're institution will likely give you credit for:

- Publishing ***Accounting*** research
- Teaching well
- Service to your department, school, and field
- Being a good colleague / not being a jerk

Remember!

As amazing as learning stats is, you are (probably) not a statistician.

You're institution will likely give you credit for:

- Publishing ***Accounting*** research
- Teaching well
- Service to your department, school, and field
- Being a good colleague / not being a jerk

You are less likely to get credit for a cutting-edge technique or a really cool analysis - so try not to fall down a rabbit hole.

(I'm still working on taking this advice...)


IBM
Presents
You Make The Call

You make the call!

You make the call!

You have conducted a 2 x 2 between-subjects experiment with the following IVs and DV.

You make the call!

You have conducted a 2 x 2 between-subjects experiment with the following IVs and DV.

	Description	**Levels**
Factor 1	Inspection Likelihood?	High, Low
Factor 2	Partner based out of your office?	Yes, No
Dependent Variable	Sought Additional Evidence?	Yes, No

You make the call!

You have conducted a 2 x 2 between-subjects experiment with the following IVs and DV.

	Description	Levels
Factor 1	Inspection Likelihood?	High, Low
Factor 2	Partner based out of your office?	Yes, No
Dependent Variable	Sought Additional Evidence?	Yes, No

What analysis should you use to test for main and interaction effects?

You make the call!

You have conducted a 2 x 2 between-subjects experiment with the following IVs and DV.

	Description	Levels
Factor 1	Inspection Likelihood?	High, Low
Factor 2	Partner based out of your office?	Yes, No
Dependent Variable	Sought Additional Evidence?	Yes, No

What analysis should you use to test for main and interaction effects?

a. ANOVA

b. Linear Regression

c. Logistic Regression

d. Structural Equation Modeling

e. Other

You make the call!

You have conducted a 2 x 2 between-subjects experiment with the following IVs and DV.

	Description	Levels
Factor 1	Inspection Likelihood?	High, Low
Factor 2	Partner based out of your office?	Yes, No
Dependent Variable	Sought Additional Evidence?	Yes, No

What analysis should you use to test for main and interaction effects?

a. ANOVA

b. Linear Regression

c. Logistic Regression

d. Structural Equation Modeling

e. Other

Gomila (2020 Journal of Exp. Psych: General)

Testing Process Theories

(Asay, Guggenmos, Kadous, Koonce, and Libby 2019 WP)

Testing Process Theories

(Asay, Guggenmos, Kadous, Koonce, and Libby 2019 WP)

- Process testing has become nearly ubiquitous in experimental accounting research.

Testing Process Theories

(Asay, Guggenmos, Kadous, Koonce, and Libby 2019 WP)

- Process testing has become nearly ubiquitous in experimental accounting research.
- We argue that process evidence is not always necessary to make a significant contribution to the accounting literature, but can add value to a paper.

Testing Process Theories

(Asay, Guggenmos, Kadous, Koonce, and Libby 2019 WP)

- Process testing has become nearly ubiquitous in experimental accounting research.
- We argue that process evidence is not always necessary to make a significant contribution to the accounting literature, but can add value to a paper.
- The ability of a process analysis to provide useful evidence rests on the researchers' ability to conduct the analysis in a way that minimizes the threats to validity that are common to the method used.

What is Process Analysis?

What is Process Analysis?

Process analyses shed light on the mechanisms linking cause and effect and the circumstances that alter the direction and magnitude of those links.

What is Process Analysis?

Process analyses shed light on the mechanisms linking cause and effect and the circumstances that alter the direction and magnitude of those links.

- When does this effect occur or not occur?

What is Process Analysis?

Process analyses shed light on the mechanisms linking cause and effect and the circumstances that alter the direction and magnitude of those links.

- When does this effect occur or not occur?
- How does this effect "work"?

What is Process Analysis?

Process analyses shed light on the mechanisms linking cause and effect and the circumstances that alter the direction and magnitude of those links.

- When does this effect occur or not occur?
- How does this effect "work"?
- What conditions exacerbate or mitigate the effect?

What is Process Analysis?

Process analyses shed light on the mechanisms linking cause and effect and the circumstances that alter the direction and magnitude of those links.

- When does this effect occur or not occur?
- How does this effect "work"?
- What conditions exacerbate or mitigate the effect?

This information is often interesting or helpful,
but may or may not be a study's main research question.

Providing Process Evidence

There are many ways to provide process evidence.

(if it's needed!)

Providing Process Evidence

There are many ways to provide process evidence.

(if it's needed!)

- Moderation-of-Process designs
  - Manipulated Moderators
  - Measured Moderators

Providing Process Evidence

There are many ways to provide process evidence.

(if it's needed!)

- Moderation-of-Process designs
  - Manipulated Moderators
  - Measured Moderators
- Mediation-by-Measurement designs
  - Less Obtrusive Measurement
  - More Obtrusive Measurement
  - Coded Responses

Providing Process Evidence

There are many ways to provide process evidence.

(if it's needed!)

- Moderation-of-Process designs
  - Manipulated Moderators
  - Measured Moderators
- Mediation-by-Measurement designs
  - Less Obtrusive Measurement
  - More Obtrusive Measurement
  - Coded Responses
- Triangulation
  - Multiple experiments (within or across papers)
  - Multiple methods (within or across papers)

Providing Process Evidence

Our paper does not say that you:

Providing Process Evidence

Our paper does not say that you:

- need process evidence for your paper.

Providing Process Evidence

Our paper does not say that you:

- need process evidence for your paper.
- need to use multiple experiments.

Providing Process Evidence

Our paper does not say that you:

- need process evidence for your paper.
- need to use multiple experiments.
- need to use multiple methods.

Providing Process Evidence

Our paper does not say that you:

- need process evidence for your paper.
- need to use multiple experiments.
- need to use multiple methods.
- should't use PEQs for mediation.

Providing Process Evidence

Our paper does not say that you:

- need process evidence for your paper.
- need to use multiple experiments.
- need to use multiple methods.
- should't use PEQs for mediation.
- should always choose moderation over mediation.

Providing Process Evidence

Our paper does not say that you:

- need process evidence for your paper.
- need to use multiple experiments.
- need to use multiple methods.
- should't use PEQs for mediation.
- should always choose moderation over mediation.
- should use SEM instead of PROCESS.

Providing Process Evidence

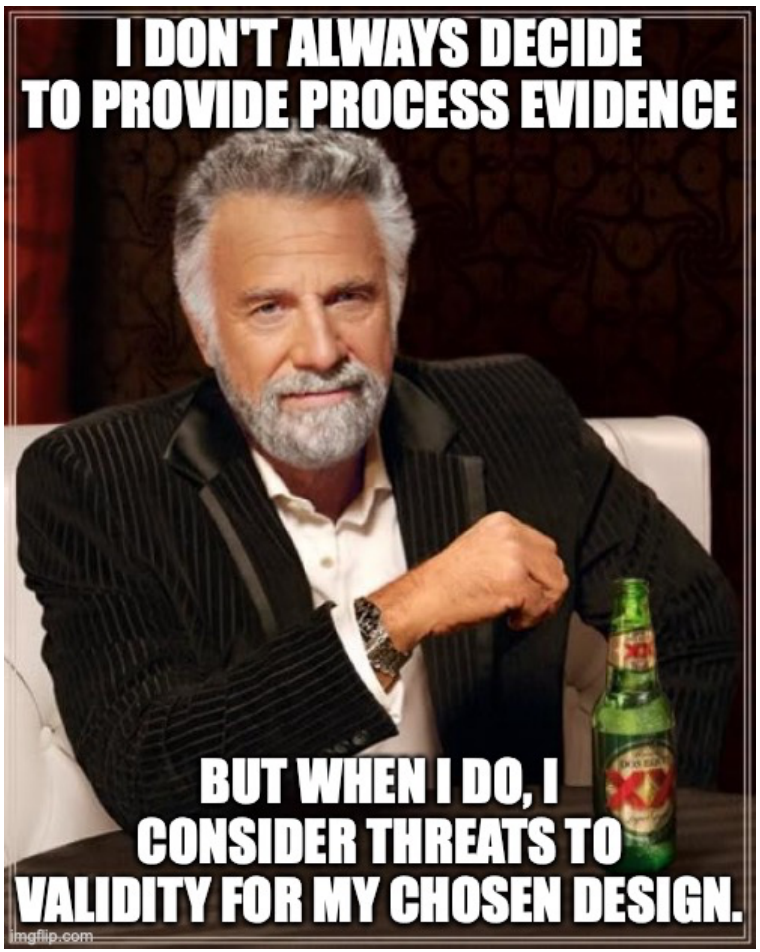
Our paper does not say that you:

- need process evidence for your paper.
- need to use multiple experiments.
- need to use multiple methods.
- should't use PEQs for mediation.
- should always choose moderation over mediation.
- should use SEM instead of PROCESS.
- should use PROCESS instead of SEM.

Providing Process Evidence

Our paper does not say that you:

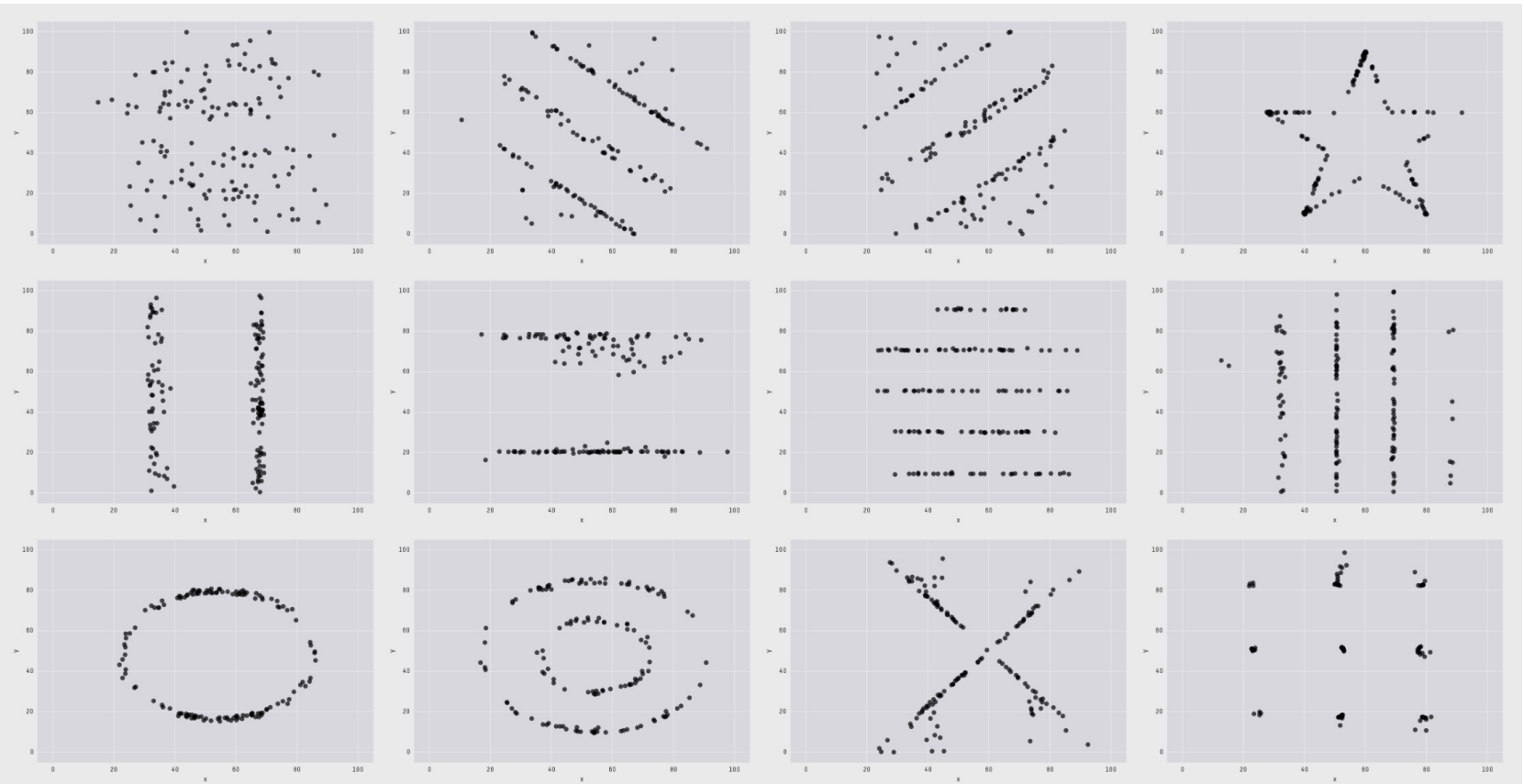
- need process evidence for your paper.
- need to use multiple experiments.
- need to use multiple methods.
- should't use PEQs for mediation.
- should always choose moderation over mediation.
- should use SEM instead of PROCESS.
- should use PROCESS instead of SEM.
- don't need process evidence for your paper.


I DON'T ALWAYS DECIDE
TO PROVIDE PROCESS EVIDENCE
BUT WHEN I DO, I
CONSIDER THREATS TO
VALIDITY FOR MY CHOSEN DESIGN.
imgflip.com


IBM
Presents
You Make The Call

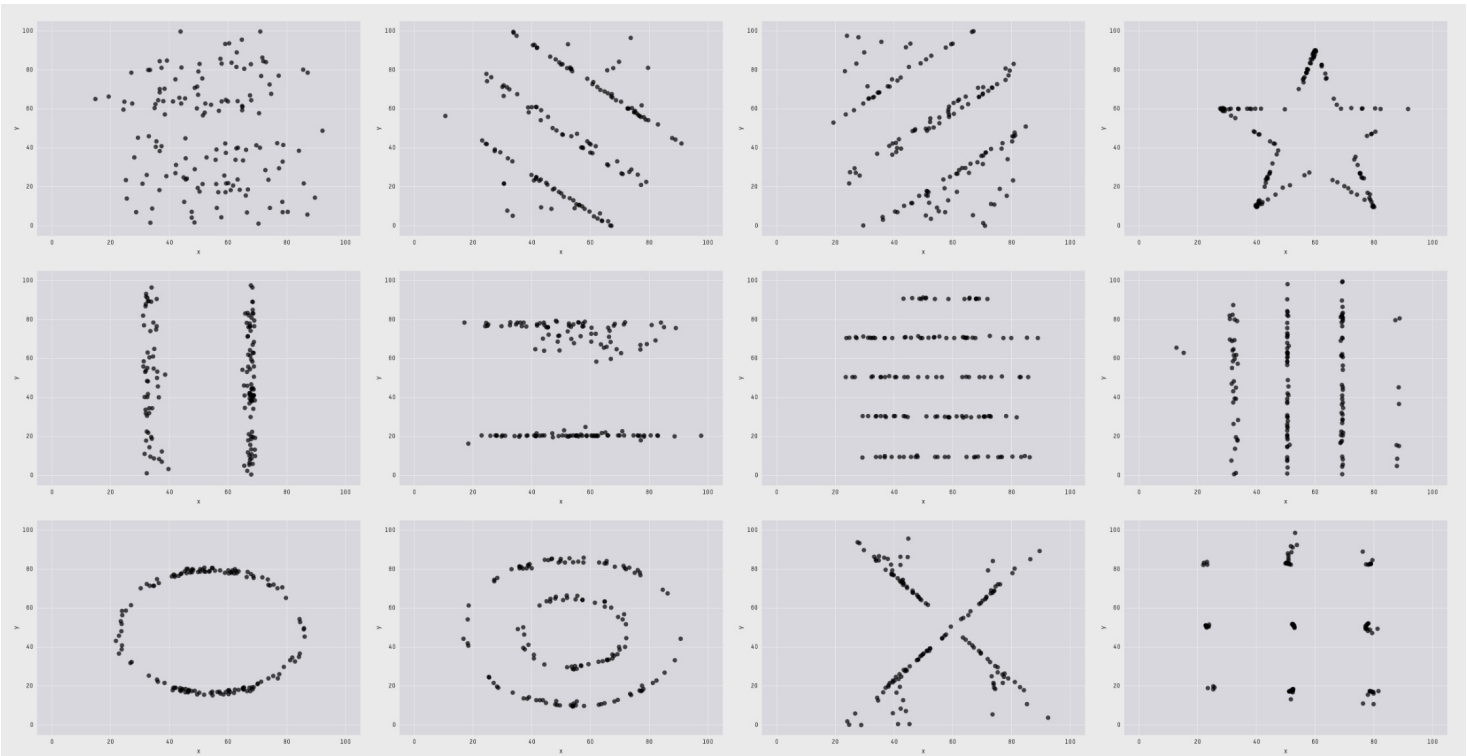
You make the call!

r = -0.06	X	Y
mean	54.26	47.83
Std. dev.	16.76	26.93



Visualize

All of them.

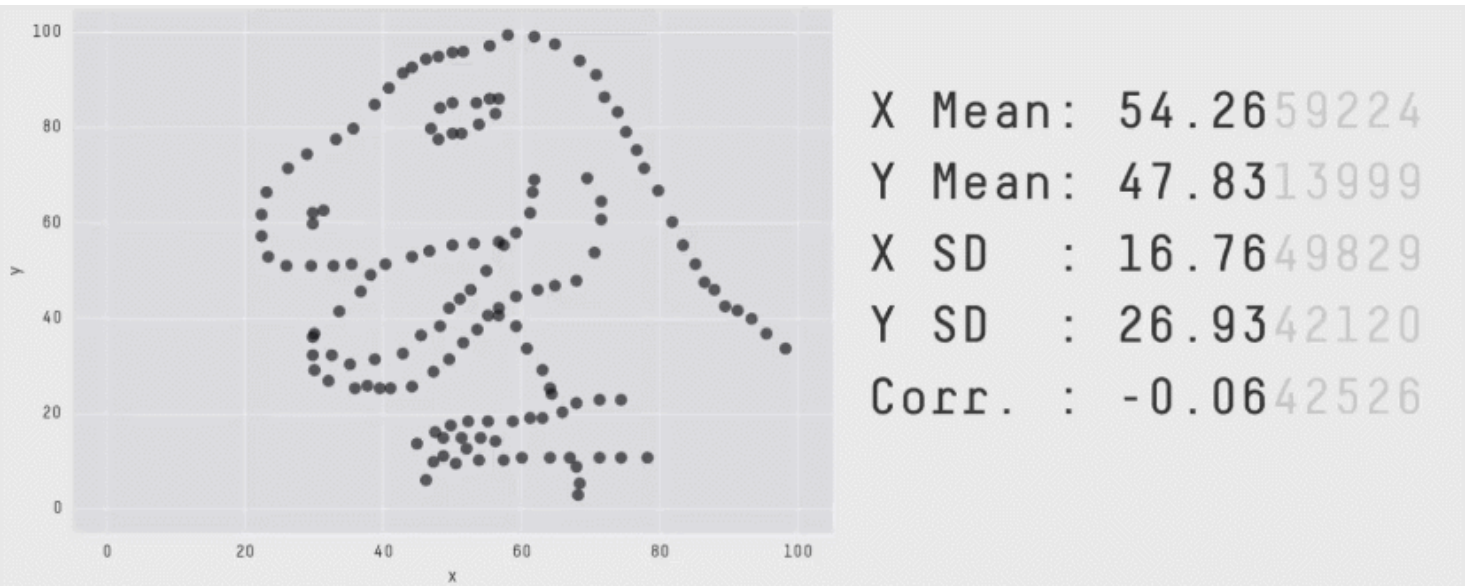


Matejka and Fitzmaurice (2017)

Data visualization isn't optional

"*...make both calculations and graphs. Both sorts of output should be studied; each will contribute to understanding."*

Anscombe (1973)



Matejka and Fitzmaurice (2017)

Be thankful...

If we need a short suggestion of what exploratory data analysis is, I would suggest that

1. It is an attitude, AND
2. A flexibility, AND
3. Some graph paper (or transparencies, or both).

No catalog of techniques can convey a willingness to look for what can be seen, whether or not anticipated. Yet this is at the heart of exploratory data analysis. The graph paper—and transparencies—are there, not as a technique, but rather as a recognition that the picture-examining eye is the best finder we have of the wholly unanticipated.

Tukey (1980 American Statistician)

Closing thoughts.

Data analysis is hard.

- Best practices change frequently.
- Innovations in experimental analysis are not frequently published in accounting journals.
- You often get little credit for it being right, but lose a lot of credit if it is wrong.

Closing thoughts.

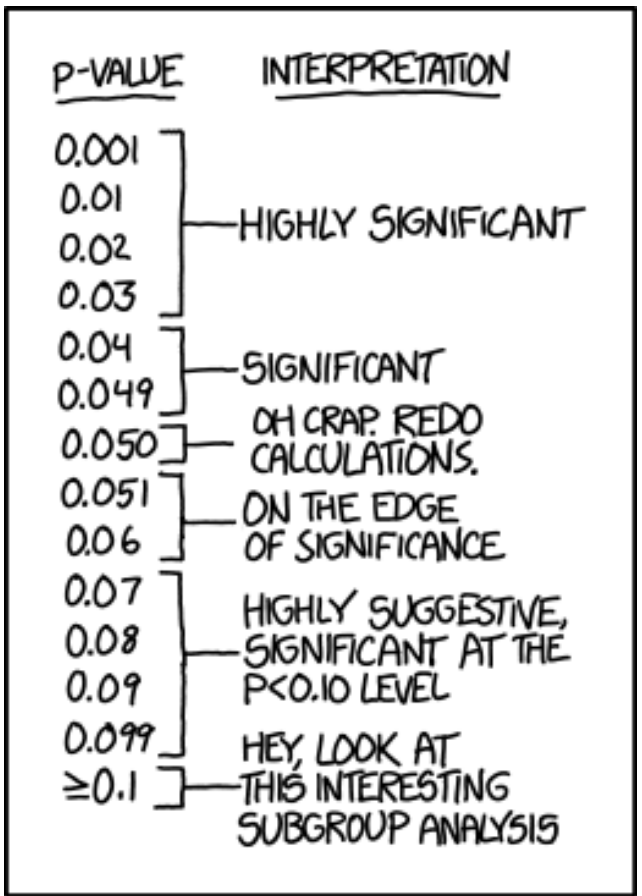
Data analysis is hard.

- Best practices change frequently.
- Innovations in experimental analysis are not frequently published in accounting journals.
- You often get little credit for it being right, but lose a lot of credit if it is wrong.

Data analysis has to be right.

- Good analysis always starts with good design.
- If analysis isn't sound, we can't trust the results.
- When experimental methods papers are published in accounting, read them closely and make sure you understand.
- Ask for help!

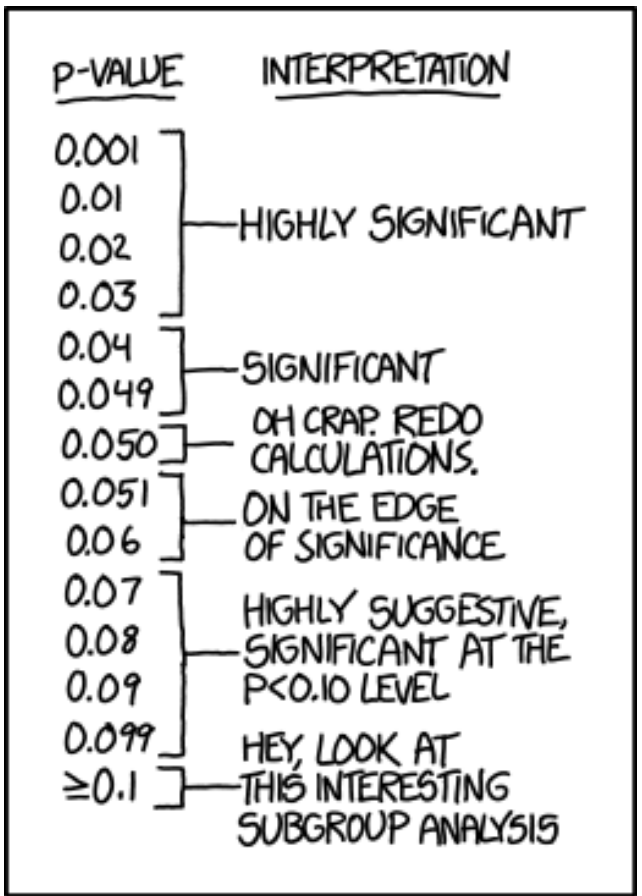
But most importantly,



www.xkcd.com

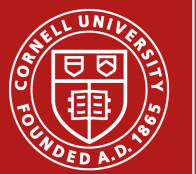
Don't be sketchy.

But most importantly,



www.xkcd.com

Don't be sketchy.



Johnson
Cornell
SC Johnson College of Business

Thank you.

And good luck with your papers!

Questions?