

Class 7 - Elements IV: Measures and data

Agenda

- Skills corner: Descriptive statistics (15 minutes)
- Conceptual grounding (10 minutes)
- Core paper discussion (30 minutes)
- *Break*
- Final Compare-contrast presentation (40 minutes)
- Tying it together: From elements to execution (15 minutes)

Skills corner

The power of descriptives

Quick reminder:

- Descriptive statistics: Procedures for depicting the main aspects of sample data, without necessarily inferring to a larger population
- Inferential statistics: Techniques that allow inferences about characteristics of a population to be drawn from a sample of data from that population

The power of descriptives

Descriptive statistics help you understand and tell the **story** of your data and the sample it is drawn from

Inferential statistics provide you the means to generalize that story beyond your sample given certain assumptions (often, but not always, by appealing to a parametric model)

Essentials: Distributions and summary statistics

Probability distribution: A mathematical description of a random phenomenon in terms of its sample space and the probabilities of events (subsets of the sample space)

Many of the properties of a distribution can be summarized by its moments:

- Mean or expected value
- Variance (or deviations from the mean)

But other descriptives are often very informative, such as interquartile range (IQR), range (Max - Min), and others.

Example 1: A normal distribution

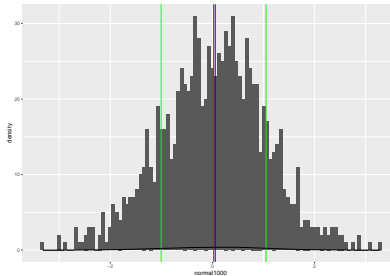


Figure 1: A normal distribution

Note the mean (red), median (blue), and standard deviations (green). A normal distribution can be completely summarized by its mean and variance.

Example 2: A log-normal distribution

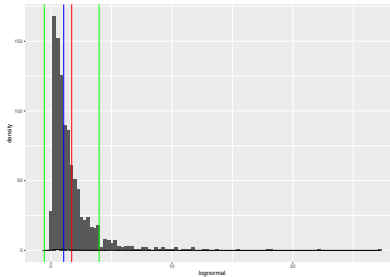


Figure 2: A log-normal distribution

By contrast, in this distribution the mean and median diverge, and the standard deviations are wonky - more information is required to characterize this distribution.

Essentials: Distributions and summary statistics

Note how the summary statistics distill key features without seeing the whole thing.

```
[1] "Normal Distribution"
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-3.32078	-0.64316	0.05468	0.02295	0.70855	3.30415

```
[1] "SD: 1.027"
```

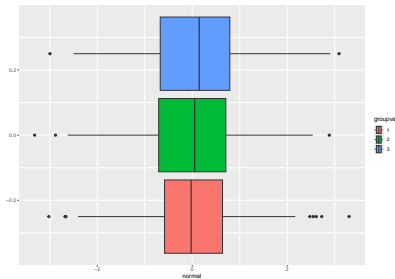
```
[1] "Log-Normal Distribution"
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.03612	0.52563	1.05620	1.71768	2.03104	27.22542

```
[1] "SD: 2.267"
```

Essentials: Univariate visualizations

Histograms (shown above) are very useful, but other tools like boxplots can be just as helpful to identify issues, such as outliers.



Essentials: Correlation matrices

When we move to multiple variables, correlation matrices take center stage, often complemented by scatterplots (Bedian 2014).

TABLE 1
Descriptive Statistics and Correlations of Study Variables

Variables	M	SD	1	2	3	4	5	6	7	8	9	10
<i>Dependent variable</i>												
1. Workplace complaining	1.86	.96	(.96) ^a									
<i>Mediating variable</i>												
2. Organization-based self-esteem	4.01	.49	-.33	(.88) ^a								
<i>Independent variables</i>												
3. Job satisfaction	3.20	.72	-.23	.46	(.80) ^a							
4. Affective organizational commitment	3.63	.73	-.27	.62	.52	(.80) ^a						
5. Procedural justice	3.26	.72	-.26	.63	.64	.56	(.88) ^a					
6. Distributive justice	3.01	.98	-.24	.45	.72	.41	.65	(.94) ^a				
7. Leader-member exchange quality	3.25	.78	-.29	.67	.66	.62	.80	.65	(.90) ^a			
<i>Control variables</i>												
8. Social desirability	.65	.21	-.01	.22	.21	.16	.16	.10	.17	(.70) ^b		
9. Negative affectivity	2.96	.71	.15	-.21	-.13	-.08	-.10	-.07	-.15	-.35	(.86) ^a	
10. Gender (Male = 0; Female = 1) ^c	.82	.38	.06	-.02	.03	.03	-.07	-.02	-.07	-.01	.08	(NA)

Note. $n = 290$. Abbreviations: Correlations $\geq |.12|$ are significant at $p < .05$ (two-tailed test).

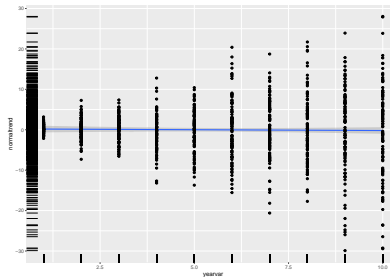
^a Cronbach's alpha (α) reliability coefficient.

^b K-R 20 reliability coefficient.

^c Point-biserial correlation.

Essentials: Scatterplots and trendlines

Here is an example **multi-graph**, with “marginal” rug plot, scatterplot, and linear trend.



Catching issues

Doing a thorough review of your data through a descriptive lens can help identify issues that will turn up in your analysis, such as:

- Deceptive descriptives - a.k.a. the datasaurus dozen
- Range restriction or selection issues - no data for certain conditions
- Non-linearity - potential cut-off or non-linear effects
- Heteroskedasticity - uneven variances across the distribution or group-level effects
- Outliers - influential observations that throw off a trendline (n.b.: when are outliers noise and when are they the signal?)

Catching issues: Anscombe's quartet

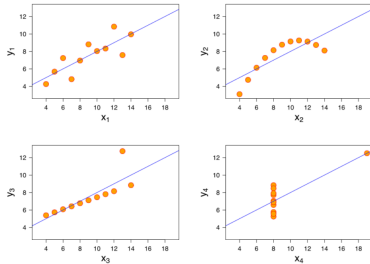


Figure 3: Anscombes' Quartet

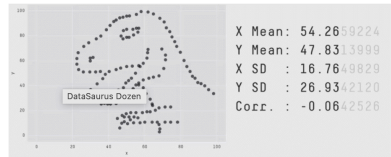
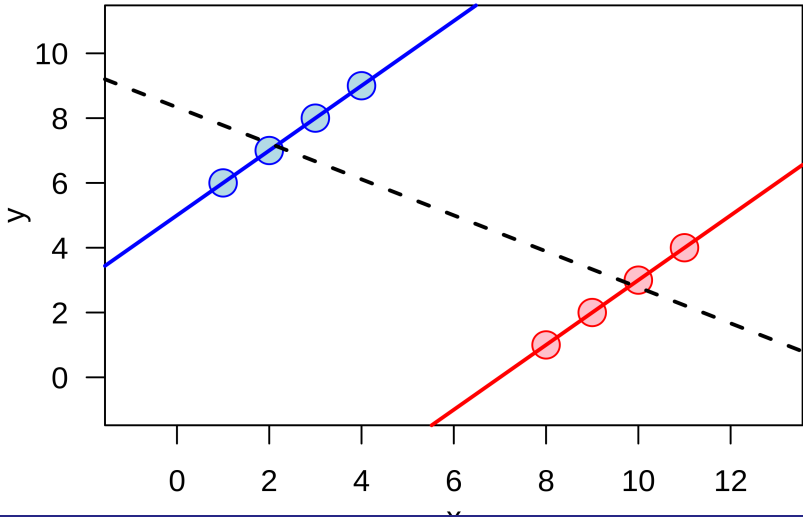


Figure 4: The Datasaurus Dozen

Catching issues: Simpson's paradox



Telling stories with data

Beyond being a prelude to inferential analyses, descriptives can also help to tell your story directly

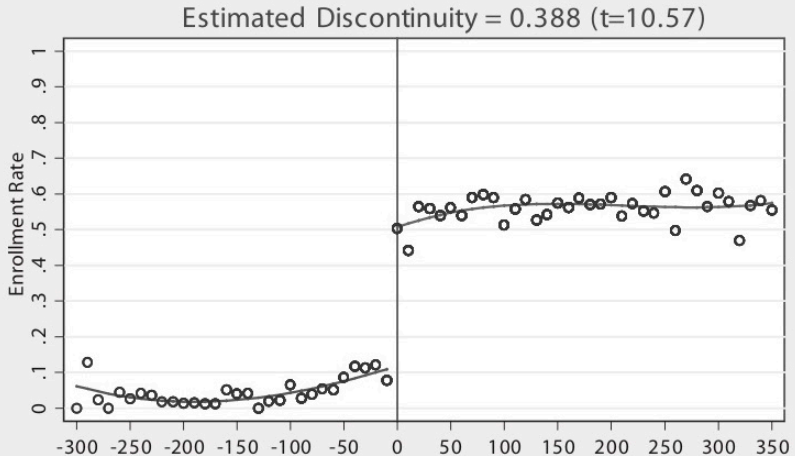
- Cross-tabulations (i.e., 2x2s)
- Group comparisons (e.g., pre- / post- intervention)
- “Existence proofs” (presence and/or variation of a construct)

Telling stories: Cross-tabulations

	Category	Number of studies in category combination																																Count	Unique Articles
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32		
Context	1 Form of interdependence	22																															22	108	
	2 Sequential	56																															56		
	3 Reciprocal	39																															39		
	4 Job demands	1	3	3	6																												7		
	5 Strategic challenges	3	4	3																													11		
	6 Dynamism	2	5	5	1	1	10																										12		
	7 Hierarchical	6	21	3	4	3	6	26																									31		
	8 Relational	8	13	9	1	3	2	2	26																								31		
	9 Organizational	3	7	5	1	1	1	1	2	14																							18		
	10 Cultural	5	9	9				4	2	1	1	19																					23		
Context	11 Intra	6	41	17	3	1	7	21	8	9	11	77																					85		
	12 Parties present	9	12	4	1	4	4	3	6	1	6	4	15																				25		
	13 SL-Internal	7	3	9	2	5	3	2	12	4	2	4	6	11																		21			
	14 Downwards	7	35	12	2	2	6	19	12	6	13	41	14	4	39																		59		
	15 Upwards	4	10		1	2	2	4	6	3	2	10	4	14		15																	15		
	16 Influence direction	6	6	12	3	7	1	4	9	6	3	19	4	3		26																	26		
	17 Multidirectional	5	5	6	1	3		4	4	3	2	14	3									17										17			
	18 Constraining	2	9	4	3	6	1	5	3	1	2	12	2	4	9	2	5	2	18														18		
	19 Enabling	13	28	17	2	4	7	16	18	7	17	31	16	13	37	9	11	3				48											60		
	20 Influence pattern	5	13	7	1	1	3	6	7	4	3	19	5	2	9	3	6	5				27											27		
Context	21 Unidirectional	2	6	2	1	1		3	2	6	1	8	2	2	4	1	4	3				12											12		
	22 Socio-cognitive	1	13	5	1	1	1	3	3	2	4	8	1	1	17	1	2	4	12	4			9										15		
	23 Socio-regulatory	7	18	11	1	5	2	6	10	8	6	20	2	8	18	6	11	4	13	13	8	5	3	26									39		
	24 Influence mechanisms	3	6	5	1	2	1	1	3	1	1	8	6	1	2	10	1	3	1	3	7	2	2	3	3	6							10		
	25 Socio-cultural																															1			
	26 Socio-political	8	20	19	3	3	9	11	7	4	4	18	6	4	22	7	4	6	11	15	10	3	5	13	3		18						39		
	27 Strategic change	1	2	3		1	1	1	3		3	2	2	3		2	2	3	2	1						2							5		
	28 R&D investment	2	1					1				3		2	1	1	1	1							1								3		
	29 Proximal outcomes				1			1	2	1	1	4	1		3										1	2	1						5		
	30 Decision quality																																13		
Consequences	31 Other	3	8	1				4	3	1	3	9	5		10	2		2	9	2	1	3	4	2	1	2							13		
	32 Performance	3	14	8	2		7	8	7	4	7	31	4	3	22	1	3	2	6	14	8		9	7	3		14	1		1		28			
	33 Distal outcomes																															3			
Number of Blank Cells:		4	2	3	8	9	6	2	1	4	4	0	3	5	0	6	5	2	3	9	1	8	4	1	5	9	3	5	10	7	8	4	13		

Figure 6: Cross-tab example

Telling stories: Group comparisons



Grounding

The rubber meets the road

- Measures: Translating our variables into concrete measures (these do not need to be 1 to 1)
- (Sourcing) Data: Where are we going to find these measures?

Readings for Today

Common Readings

- 1 Stevens, S. S. 1946. On the Theory of Scales of Measurement. Science, New Series, 103, No. 2684, 677-680.
- 2 Bedian, A. G. 2014. "More Than Meets the Eye": A Guide to Interpreting the Descriptive Statistics and Correlation Matrices Reported in Management Research. Academy of Management Learning & Education, 13, No. 1, 121-135.
- 3 Heggstad, E. D., Scheaf, D. J., Banks, G. C., Monroe Hausfeld, M., Tonidandel, S., & Williams, E. B. (2019). Scale Adaptation in Organizational Science Research: A Review and Best-Practice Recommendations. Journal of Management, 45(6), 2596-2627.

Stevens (1946)

Paraphrasing N. R. Campbell (Final Report, p. 340), we may say that measurement, in the broadest sense, is defined as the assignment of numerals to objects or events according to rules. The fact that numerals can be assigned under different rules leads to different kinds of scales and different kinds of measurement. (Stevens 1946, 677)

SCIENCE

Vol. 103, No. 2684

Friday, June 7, 1946

On the Theory of Scales of Measurement

S. S. Stevens

Director, Psycho-Genetic Laboratory, Harvard University

FOR SEVEN YEARS A COMMITTEE of the British Association for the Advancement of Science debated the problem of measurement. Appointed in 1932 to represent Section A (Mathematical and Physical Sciences) and Section J (Psychology), the committee was instructed to consider and report upon the possibility of "quantitative estimates of sensory events"—meaning simply: Is it possible to measure human sensations? Deliberation led only to disagreement, mainly about what is meant by the term measurement. An interim report in 1938 found one member insisting that his colleague "came out by that same door as they went in" and in order to have another try at agreement, the committee hoped to be continued for another year.

For its final report (1940) the committee chose a common basis for its consideration, directing its arguments at a concrete example of a sensory scale. This was the *Bone scale* of Lorden (S. S. Stevens and H. Davis, *Hearings*, New York: Wiley, 1938), which purports to measure the subjective magnitude of an auditory sensation against a scale having the formal properties of other basic scales, such as those used to measure length and weight. Again the 19 members of the committee came out by the same door they entered, and their views ranged widely between two extremes. One member submitted "that any law purporting to express a quantitative relation between sensation intensity and stimulus intensity is not merely false but is in fact meaningless unless and until a meaning can be given to the concept of addition as applied to sensation" (Final Report, p. 145).

It is plain from this and from other statements by the committee that the real issue in the measuring of measurement. This, to be sure, is a semantic issue, but one impossible of orderly discussion. Perhaps agreement can better be achieved if we recognize that measurement enters in a variety of forms and that scales of measurement fall into certain definite classes. These classes are determined both by the empirical operations involved in the process of "measuring" and

by the formal (mathematical) properties of the scales. Furthermore—and this is of great concern to several of the members—the statistical manipulations that are legitimately applied to empirical data depend upon the type of scale against which the data are ordered.

A CLASSIFICATION OF SCALES OF MEASUREMENT

Paraphrasing N. R. Campbell (Final Report, p. 340), we may say that measurement, in the broadest sense, is defined as the assignment of numerals to objects or events according to rules. The fact that numerals can be assigned under different rules leads to different kinds of scales and different kinds of measurement. The problem then becomes that of making explicit: (a) the various rules for the assignment of numerals, (b) the mathematical properties (or group structure) of the resulting scales, and (c) the statistical operations applicable to measurements made with each type of scale.

Scales are possible in the first place only because there is a certain isomorphism between what we see as with the aspects of objects and the properties of the numeral series. In dealing with the aspects of objects we invoke empirical operations for determining equality (classifying), for rank-ordering, and for determining when differences and when ratios between the aspects of objects are equal. The conventional series of numerals yields to analogous operations: We can identify the members of a numeral series and classify them. We know their order as given by succession. We can determine equal differences, as 4-3=4-2, and equal ratios, as 8/4=2/1. The isomorphism between these properties of the numeral series and certain empirical operations which we perform with objects permits the use of the series as a model to represent aspects of the empirical world.

The type of scale selected depends upon the character of the basic empirical operations performed. These operations are limited originally by the nature of the thing being asked and by our choice of procedures, but, once selected, the operations determine

Stevens (1946)

Discussion Questions

- Are there types of measurements that are not discussed in this article?
- What would happen if you make inferences about a measurement using the wrong type of scale?

Stevens (1946)

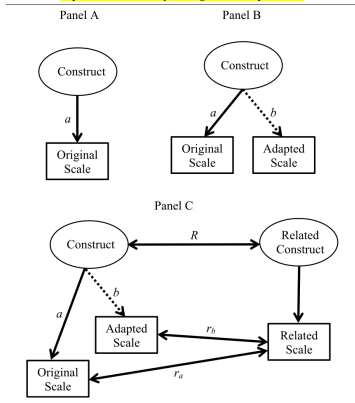
Basically, I grabbed this seminal paper to show where the different scales of measurement originated from: these are commonly referenced categories.

TABLE 1

Scale	Basic Empirical Operations	Mathematical Group Structure	Permissible Statistics (invariantive)
NOMINAL	Determination of equality	<i>Permutation group</i> $x' = f(x)$ $f(x)$ means any one-to-one substitution	Number of cases Mode Contingency correlation
ORDINAL	Determination of greater or less	<i>Isotonic group</i> $x' = f(x)$ $f(x)$ means any monotonic increasing function	Median Percentiles
INTERVAL	Determination of equality of intervals or differences	<i>General linear group</i> $x' = ax + b$	Mean Standard deviation Rank-order correlation Product-moment correlation
RATIO	Determination of equality of ratios	<i>Similarity group</i> $x' = ax$	Coefficient of variation

Heggstad et al (2019)

Figure 1
Representation of Validity for Original and Adapted Scales



Heggstad et al (2019)

Discussion questions

- What is a scale, again?
- Why would you want to change one that is already in the literature?

Heggstad et al (2019)

Key points

- Scales are very common, particularly in micro research
- They are a common basis to establish construct validity
- They are an example of where data and variables are not 1 to 1, and thus adaptations are possible
- While I would argue some of this concern is a bit overblown, the point is conceded that material changes to a scale may unmoor it from its validated basis

Bedian (2014)

Discussion Questions

- Did you know you could extract so much information from a single table?
- While this is a nice list of ‘sanity checks’, is this really all that descriptive statistics can tell us?

Bedian (2014)

The 12-point checklist

- 1 Disclosed Mean, SD, Correlations
- 2 Sensible frequency distributions
- 3 Feasible standard deviations
- 4 Reported reliabilities (multi-item scales)
- 5 Feasible correlations
- 6 Wonky looking scatterplots

Bedian (2014)

The 12-point checklist

- 7 Accounted for common data collection methods
- 8 “Correct” signs (see also Kennedy (2008))
- 9 Assessment of collinearity (e.g., VIFs)
- 10 Sensible “point-biserial” (e.g., binary) correlations?
- 11 Disclosed data missingness
- 12 Disclosed sampling procedure

Break



COFFEE BREAK

Readings for Today II

Compare / Contrast Presentations

- Combs, J. G. 2010. Big samples and small effects: Let's not trade relevance and rigor for power. *Academy of Management Journal*, 53(1): 9-13.
- Simsek, Z., Vaara, E., Parachuri, S., Nadkarni, S., & Shaw, J. D. 2019. New ways of seeing big data. *Academy of Management Journal*, 62: 971-978.

Sourcing data

Given the variation in your interests and topics, it is not productive to talk at length about data sourcing since it is unique to your circumstances

But here are some ideas on the following two slides

Sourcing data - Primary sources

Collecting specifically what you need for a study

- Field surveys
- Experiments (lab or field) with purpose-built collection instruments
- Interviews / focus groups
- Direct observation / ethnographic methods

Sourcing data - Secondary sources

Relying upon others to collect data or using 'unobtrusive' measures

- Archival datasets (e.g., Factiva, COMPUSTAT, Biocentury, SDC)
- Publicly available survey data (e.g., Kauffman Firm Survey, Census ACS)
- Industry reports (e.g., Wohler's Reports)
- Video, audio, or written artifacts (CEO speeches, earnings calls, 10-Ks, website scraping)
- Company records
- Cliometric methods (historical archives)

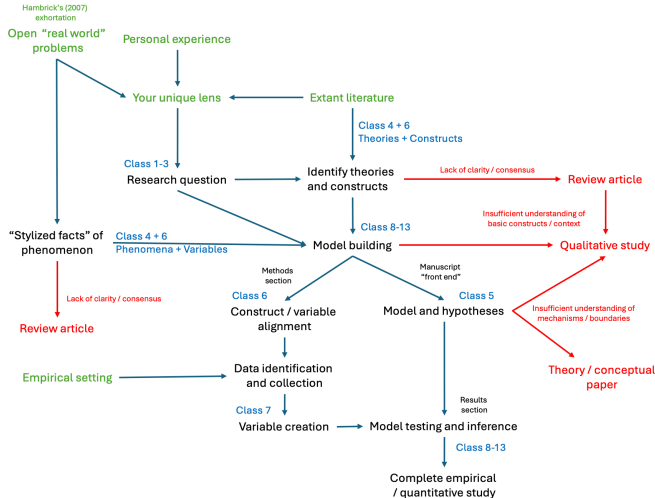
From elements to execution

A pictorial representation of the research process

On the following slide, I illustrate how the class fits together

- Note that it assumes that the intent is to complete a quantitative, empirical project, but I have indicated “offramps” to other types of contributions in red
- Class content is indicated in blue
- Information relied upon from ‘outside the system’ is shown in green

A pictorial representation of the research process



Our paradigm going forward

The balance of the class is rooted in a paradigm

- We have built a model with hypotheses to answer a research question
- That model is nested within a causal system and empirical reality
- Our hope is to recover a causal effect through an estimation process

For this to be effective, we want this estimate to be:

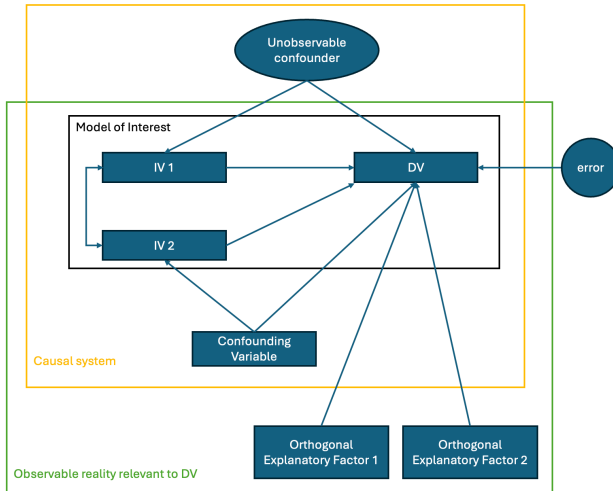
- Accurate: hopefully unbiased, but at least consistent
- Precise: an “efficient” estimator which uses sample information well

How your model fits within this paradigm

On the following slide, I illustrate how your focal model compares to the wider system of relationships

- Note that you may need to include elements outside of your model of interest to specify the causal system (as discussed in Class 5: Models + Hypotheses)
- Also note that our model may be influenced by observable and unobservable factors (as we discussed in Class 6: Constructs + Variables)
- This presumes you have a story you want to understand rather than trying to maximize the predictability of the DV (a different question to answer)

How your model fits in the big picture



Preparation for next class

Next class

Techniques I: Regression

- 1 Kennedy, P. 2008. A Guide to Econometrics (6th Edition ed.). Malden, MA: Blackwell. [Chs. 3 and 4]
- 2 Carlson, K. D., & Wu, J. 2012. The illusion of statistical control: Control variable practice in management.

Next class

Techniques I: Regression

Applications:

- 3 Replication: Simsek, Z., Fox, B., & Heavey, C. 2021. Systematicity in Organizational Research Literature Reviews: A Framework and Assessment. Organizational Research Methods, 109442812110086.
- 4 Katila, R., & Ahuja, G. 2002. Something Old, Something New: A Longitudinal Study of Search Behavior and New Product Introduction. Academy of Management Journal, 45(6), 1183-1194.

How it will work

Everyone should read the first two articles in detail, they provide grounding

Everyone should have a working familiarity with the application papers (what the study is about, how the tests were performed, key findings)

One group will actively **try to replicate** the findings with data and code that I provide and report out the process

Workshop

Preview for next class?

Given the technical nature of next week's materials, would it be helpful to have a quick preview?

Agenda

- Research Question Presentations (11 - 2p; break from 12 - 1p)
- modified to leave at 11:45, please be back no later than 1:30p
- Breakouts (2 - 3:30p)
- Reflections

References

- Bedian, Arthur G. 2014. “‘More Than Meets the Eye’: A Guide to Interpreting the Descriptive Statistics and Correlation Matrices Reported in Management Research.” *Academy of Management Learning & Education* 13, No. 1: 121–35.
- Kennedy, Peter. 2008. *A Guide to Econometrics*. Malden, MA: Blackwell.
- Stevens, S. S. 1946. “On the Theory of Scales of Measurement.” *Science New Series*, 103, No. 2684: 677–80.