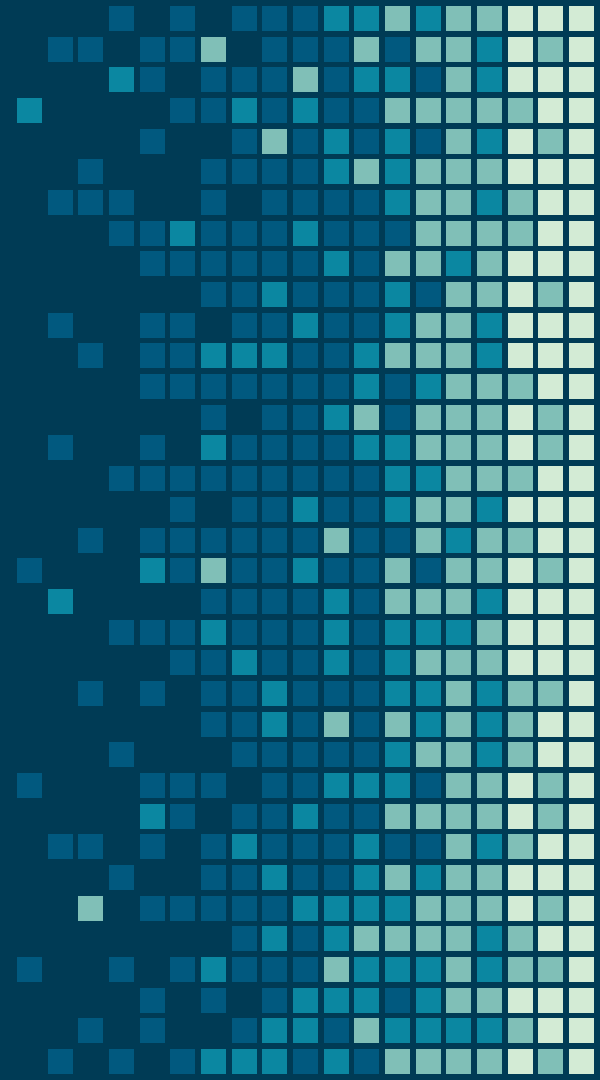


Overcoming arithmophobia:

A statistics primer

Laura Hamonic & Maria Tan

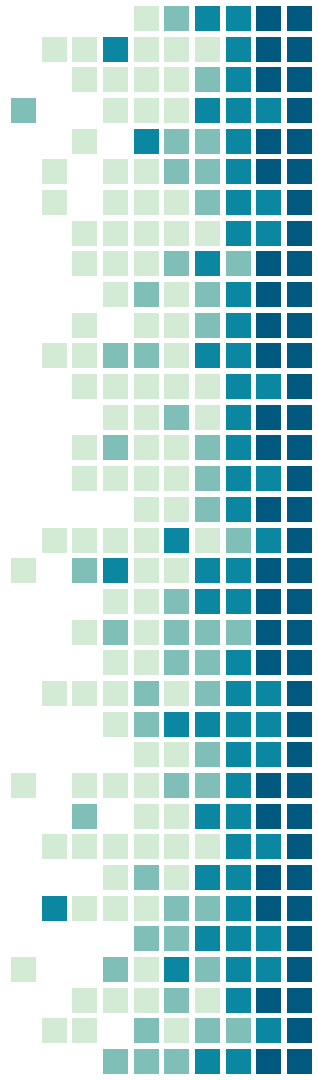
February 2018



Purpose

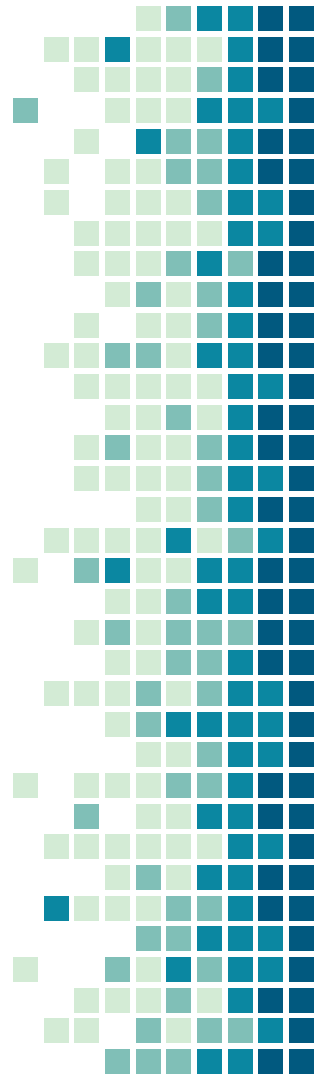
To refresh your memory about basic descriptive stats you will encounter in survey research

To help you to decide what types of stats are appropriate for types of survey data



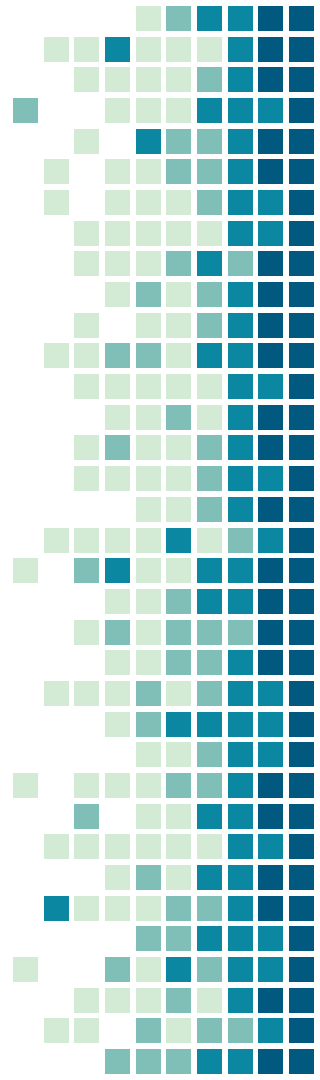
Learning Outcomes

- Recognize the appropriate level of measurement when looking at survey items
- Identify which type of data lends itself best to which descriptive statistics (e.g., central tendency)
- Interpret the meaning of measures of dispersion



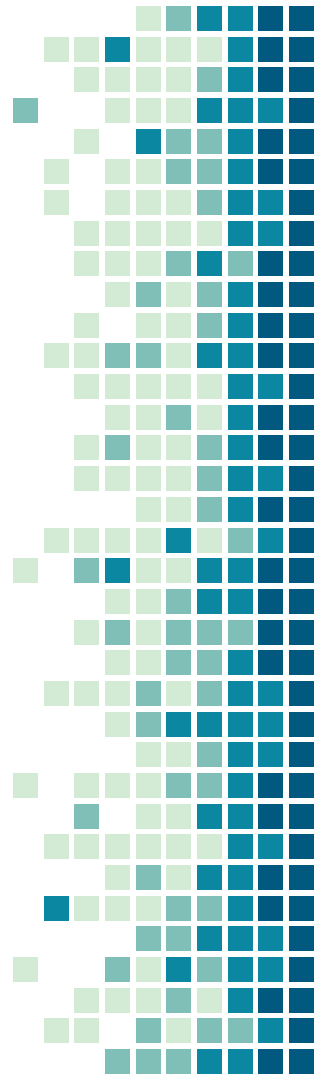
Outline

- The Role of Stats
- Variables
- Levels of Measurement (NOIR)
- Descriptive Stats
 - Frequency distributions
 - Central tendency
 - Dispersion
- Inferential Stats & Hypotheses



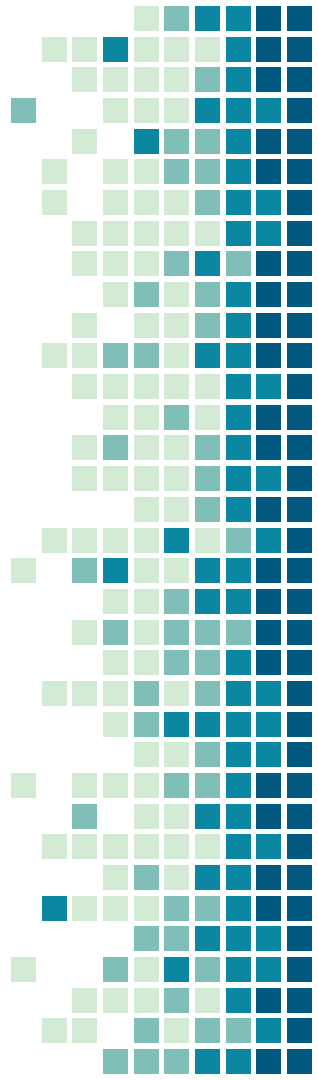
Role of statistics

- What are stats?
 - Data vs Stats
- Why are they important?
 - Getting to know your data



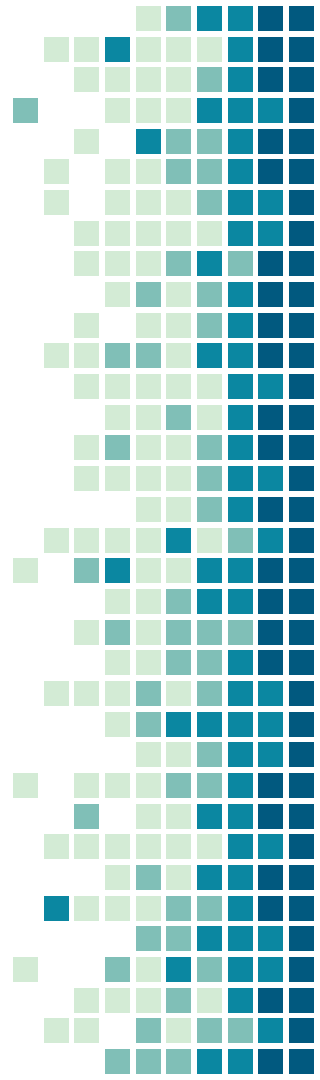
Variables

- Independent vs. dependent variables
- Discrete vs. Continuous variables
- Univariate vs. Bivariate



Levels/scales of measurement (NOIR)

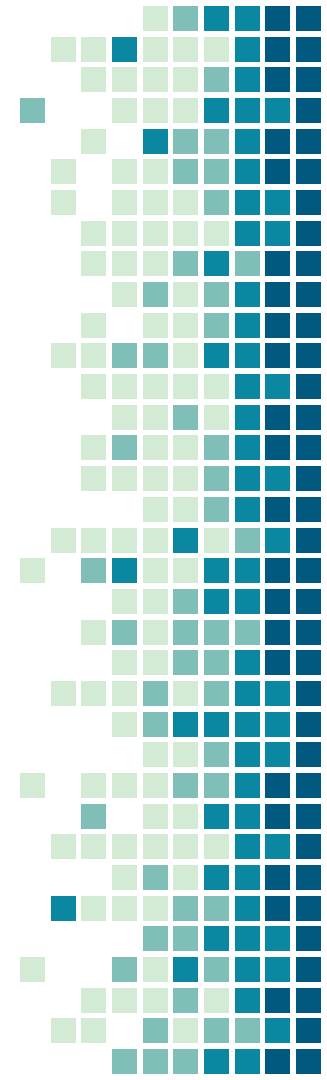
- **N**ominal
- **O**rdinal
- **I**nterval
- **R**atio



Nominal (aka categorical)

Please select the area of practice that best applies to you:

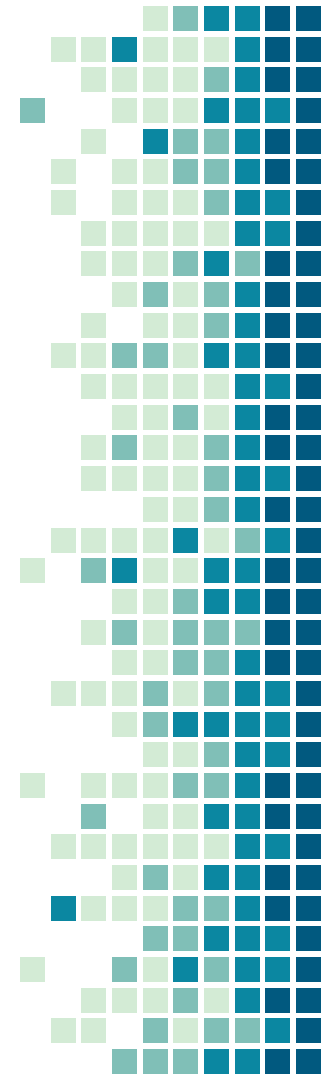
- ☐ Academic Institution
- ☐ Hospital Library
- ☐ Public Library
- ☐ Government Library
- ☐ Other



Ordinal (aka ranked)

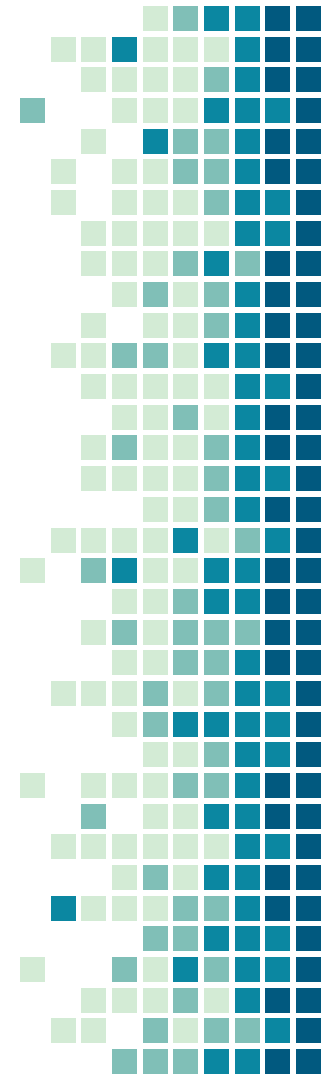
Please select the answer that best describes your level of education

- ☐ LIS Doctorate
- ☐ LIS Masters
- ☐ LIS Bachelors
- ☐ LIS Technical Diploma



Interval

Some temperatures (Celsius & Fahrenheit)



Ratio

Race Results	
Lewis	2:30
Kamran	2:32
Rainier	5:00

What about Likert Scales?

Compared to before this session, how confident are you about which library to approach for specific resources and services?

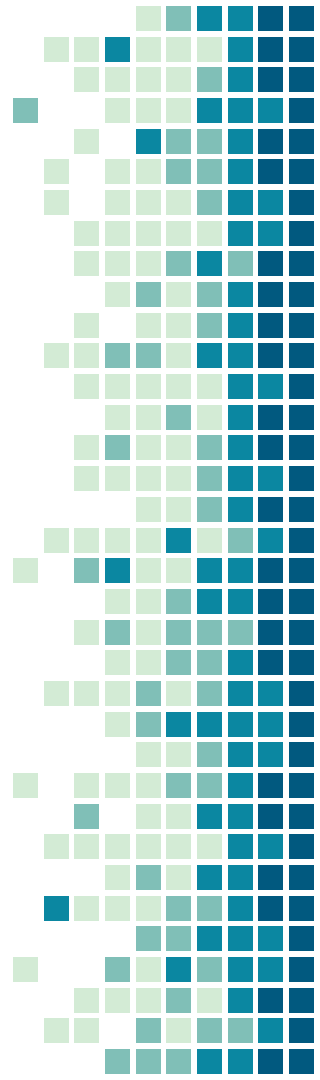
- ☐ Much more confident
- ☐ Somewhat more confident
- ☐ No change
- ☐ Somewhat less confident
- ☐ Not at all confident

How well did the session meet your expectations?

	1	2	3	4	5	
Not at all	☐	☐	☐	☐	☐	Extremely well

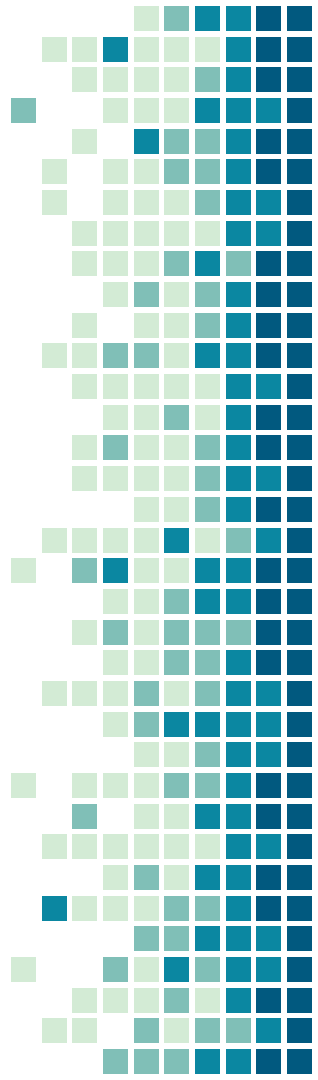
NOIR Activity, Part 1

- How many bags of m&ms do you eat in a week?
Less than 1, 1-5, 5-10, More than 10
- What is your age in numbers?
- What is your favourite colour of m&m?



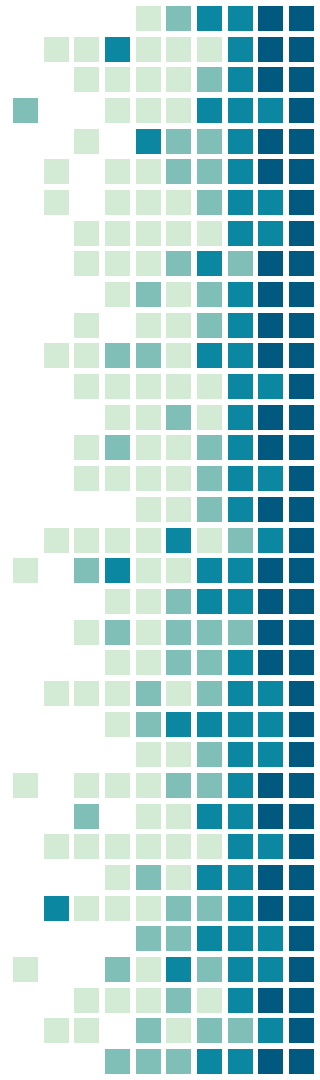
NOIR Activity, Part 1

- On a scale of 1-4, rate how much you like m&ms, where 1 is not at all, and 4 is very much
- What is the ideal temperature in celsius at which to eat m&ms?



NOIR Activity, Part 2

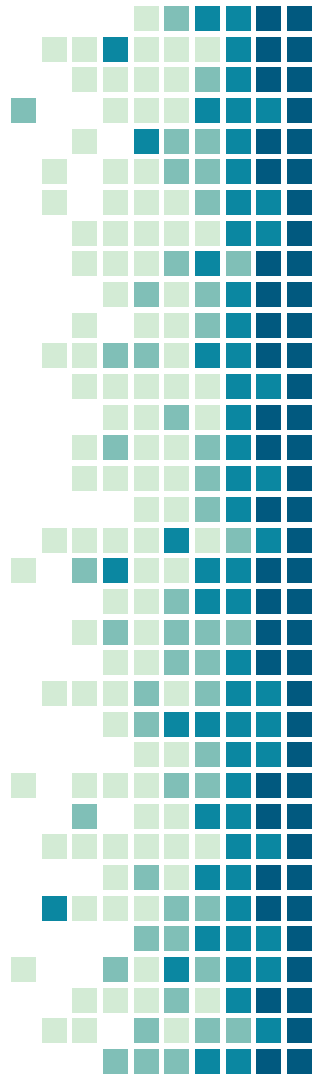
- What is your independent & dependent variable
- Identify the types of data (in terms of levels of measurement) you would want to answer your study's question



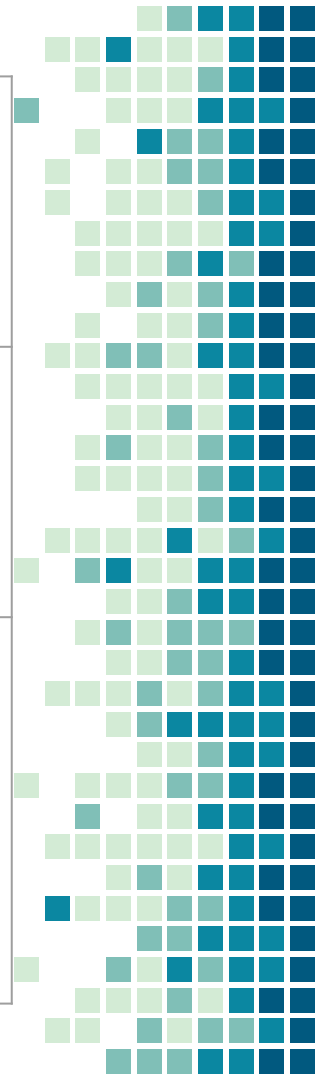
Central tendency

"Average" = 1 value that best represents the entire group of scores

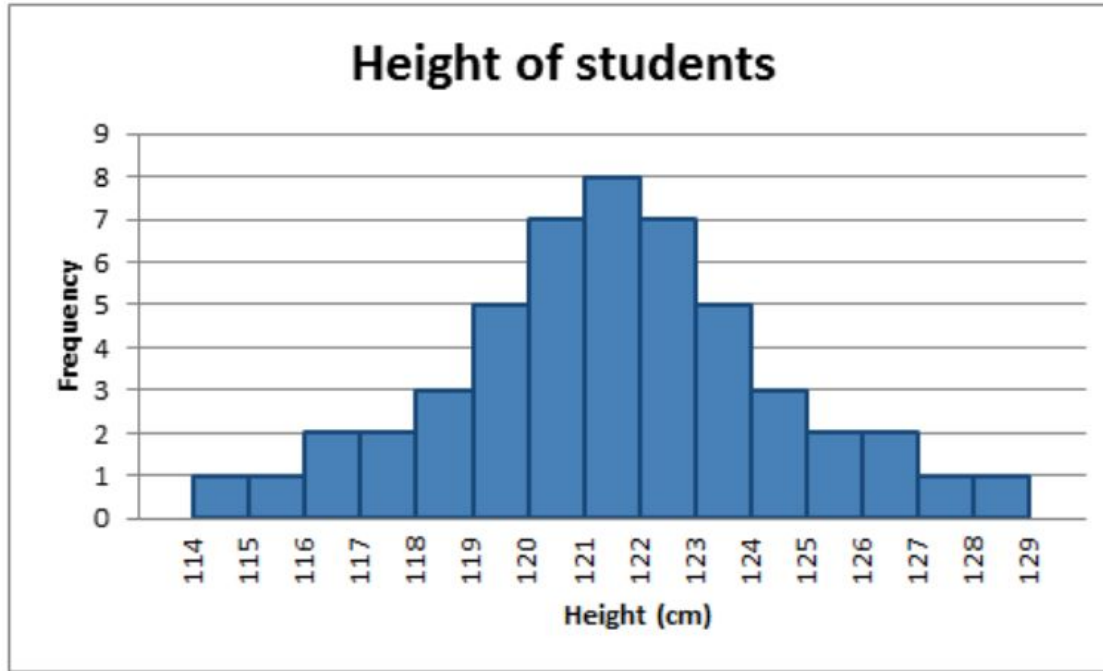
- **Mode:** Value that occurs the most often
- **Median:** Value of the middle case
- **Mean:** Sum of the individual values divided by the number of cases



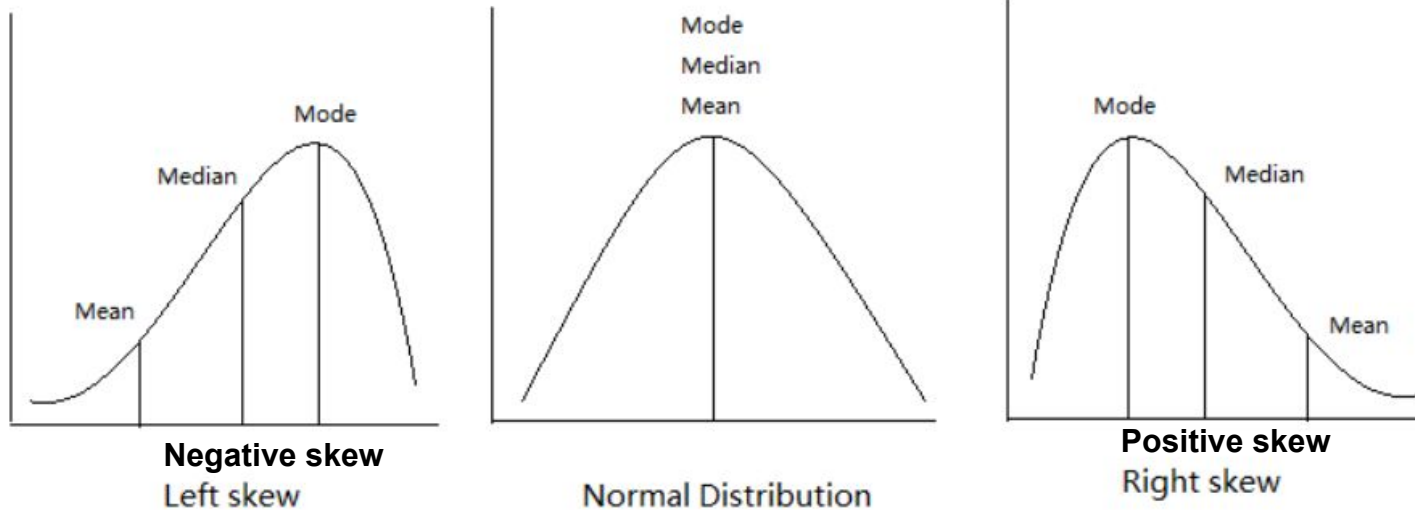
Mode	Which value occurs the most often Can have >1 mode	Most general, least precise	Nominal & ordinal data - mutually exclusive categories
Median	Middle/Centremost case Focus: # of cases, not their value	Not sensitive to extreme scores	When have extreme scores
Mean	Point at which all the values are equal in weight on both sides of the distribution Focus: value of the cases	Very sensitive to extreme scores	Interval & Ratio data No extreme scores

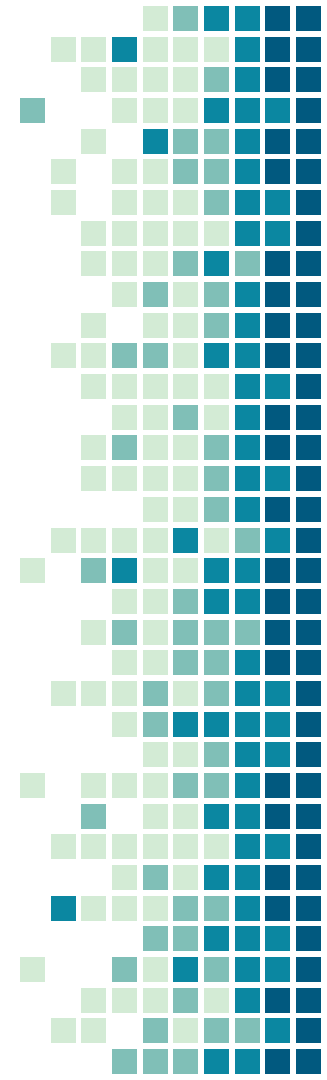


Normal Distribution



Central tendency & Skew





Central tendency – What are all these weird symbols

Population Mean

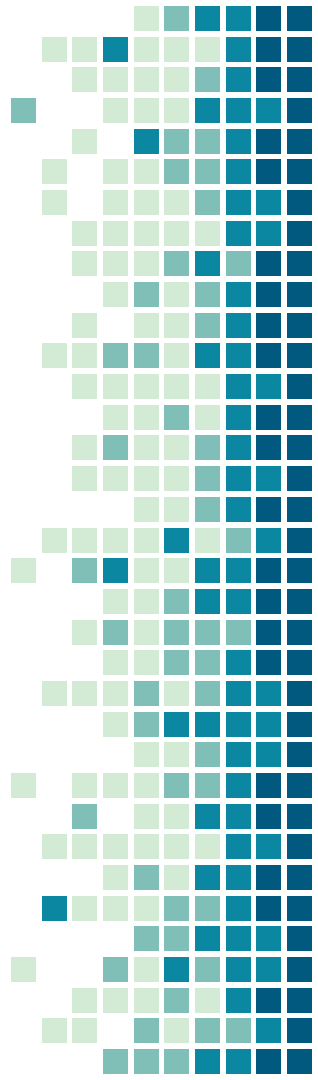
$$\mu = \frac{\sum x}{N}$$

Sample Mean

$$\bar{X} = \frac{\sum x}{n}$$

Central tendency – M&M activity

- Pour out your m&ms and organize your data points by colour
- Mode?
- Frequency?
- Median?



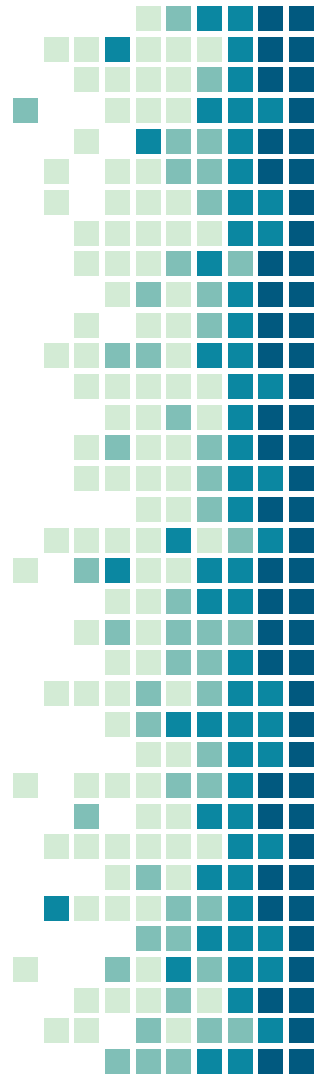
Central tendency – M&M activity

- Does it work for mean?

Mode can be used for nominal, ordinal, interval, ratio.

Median can be used for ordinal, interval, and ratio data

Mean can only be used for interval and ratio data

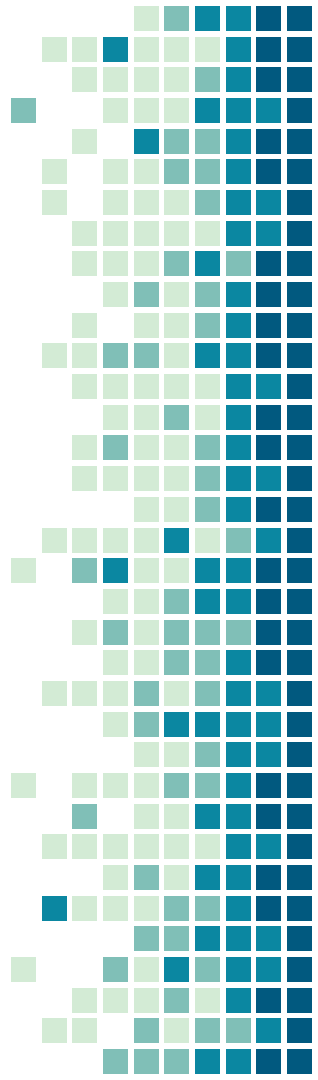


Dispersion

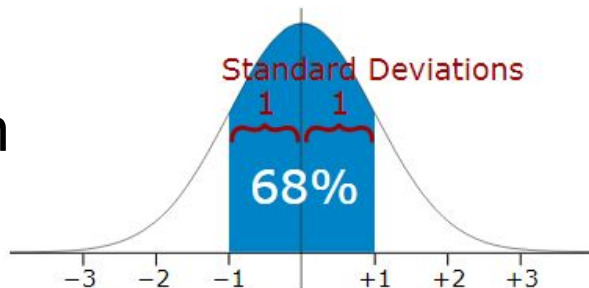
- Range = difference between highest and lowest values
- Standard Deviation = distance from the mean

Standard deviation explained and visualized

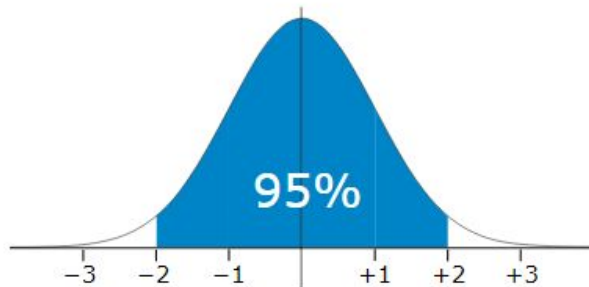
<https://www.youtube.com/watch?v=MRqtXL2WX2M>



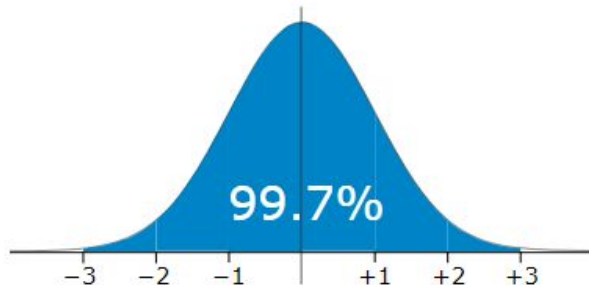
Standard Deviation



68% of values are within
1 standard deviation of the mean



95% of values are within
2 standard deviations of the mean



99.7% of values are within
3 standard deviations of the mean

Dispersion – More weird symbols

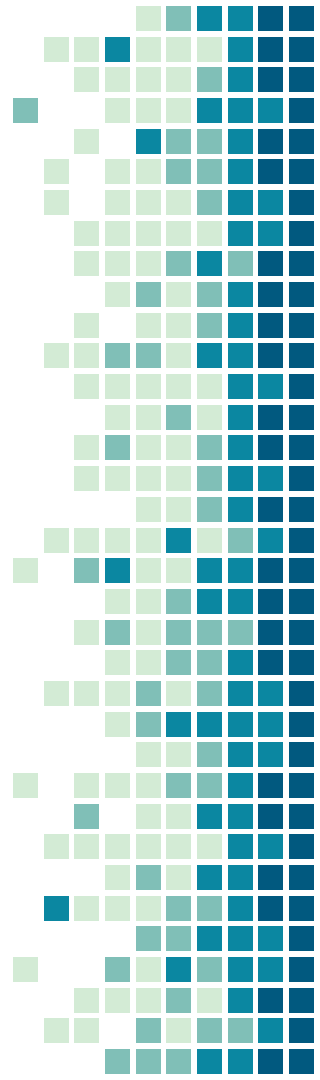
Variance (population): $\sigma^2 = \frac{\sum(x-\mu)^2}{n}$

Standard Deviation (population): $\sigma = \sqrt{\frac{\sum(x-\mu)^2}{n}}$

Variance (sample): $s^2 = \frac{\sum(x-\bar{x})^2}{n-1}$

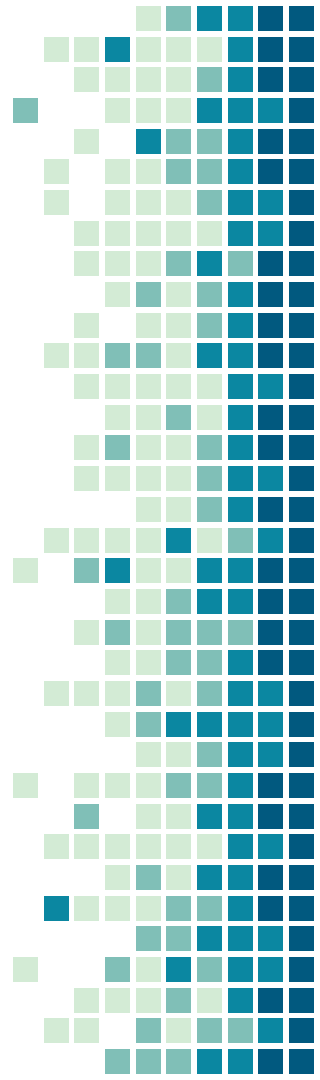
Standard Deviation (sample): $s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$

Range = maximum-minimum



Central tendency & Dispersion

- What measures can be done in your articles?
(Activity)

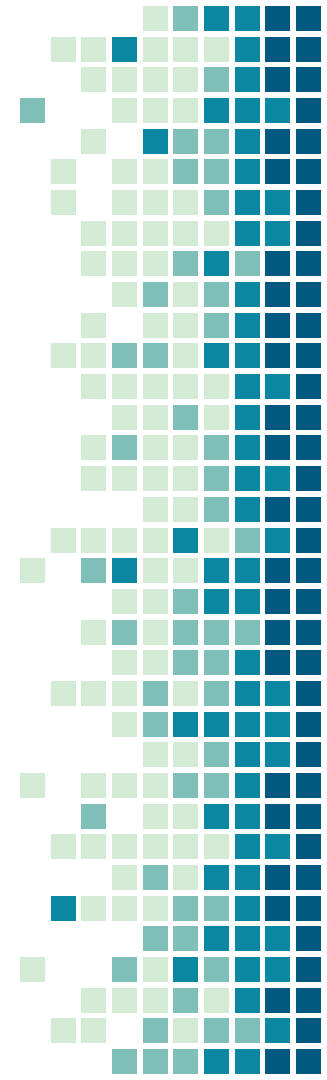


Describing relationships between variables: Correlations

How does the value of 1 variable changes when another variable changes?

Do the variables move in the same direction, or in opposite directions?

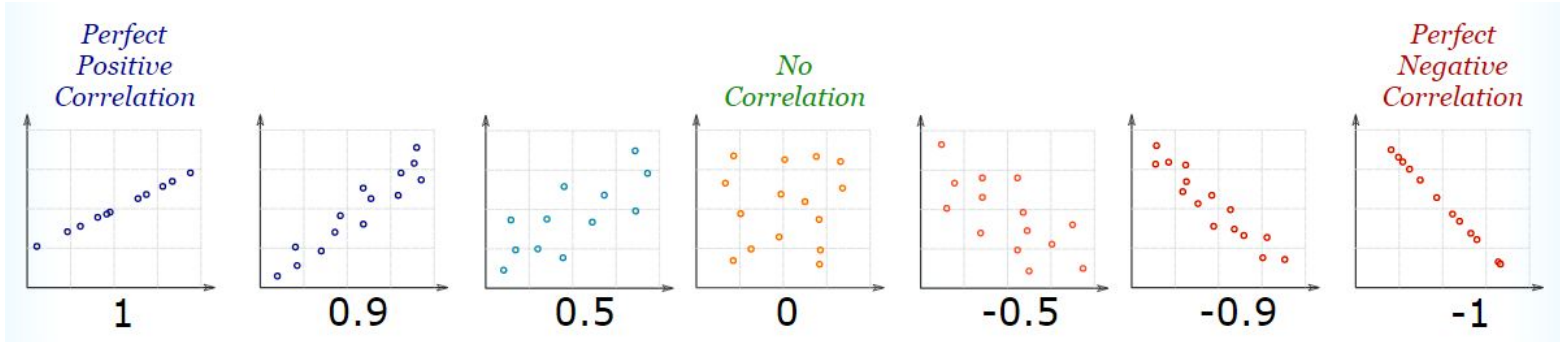
How strong is that relationship between the variables?



Correlations

Positive correlation: values move **together**

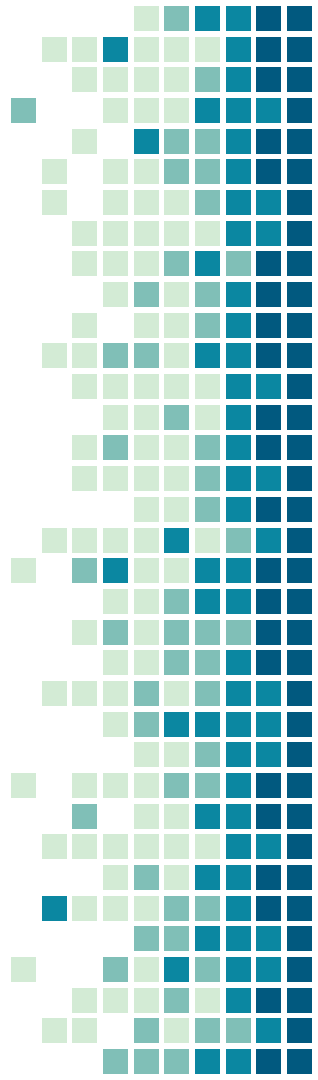
Negative correlation: values move in **opposite directions**



Adapted from <https://www.mathsisfun.com/data/correlation.html>

Correlation & Variability

Sample diversity can affect correlation strength



Determining Correlation Strength

- Eyeballing method →
- Coefficient of determination (calculation)

Correlation Coefficient Scale

+ r values	Positive	- r values	Negative
 1.0	Perfect +	 -1.0	Perfect -
 .8 to .99	Very strong+	 -.8 to -.99	Very strong-
 .6 to .8	Strong +	 -.6 to -.8	Strong -
 .4 to .6	Moderate +	 -.4 to -.6	Moderate -
 .2 to .4	Weak +	 -.2 to -.4	Weak -
 0 to .2	Very weak +	 0 to -.2	Very weak -

Source: <https://spencermath.weebly.com/home/interpreting-the-correlation-coefficient>

Calculating Correlations

TABLE 5.3 Correlation Coefficient Shopping, Anyone?

<i>Level of Measurement and Examples</i>		<i>Type of Correlation</i>	<i>Correlation Being Computed</i>
<i>Variable X</i>	<i>Variable Y</i>		
Nominal (voting preference, such as Republican or Democrat)	Nominal (sex, such as male or female)	Phi coefficient	The correlation between voting preference and sex
Nominal (social class, such as high, medium, or low)	Ordinal (rank in high school graduating class)	Rank biserial coefficient	The correlation between social class and rank in high school
Nominal (family configuration, such as intact or single parent)	Interval (grade point average)	Point biserial	The correlation between family configuration and grade point average
Ordinal (height converted to rank)	Ordinal (weight converted to rank)	Spearman rank coefficient	The correlation between height and weight
Interval (number of problems solved)	Interval (age in years)	Pearson correlation coefficient	The correlation between of number problems solved and age in years

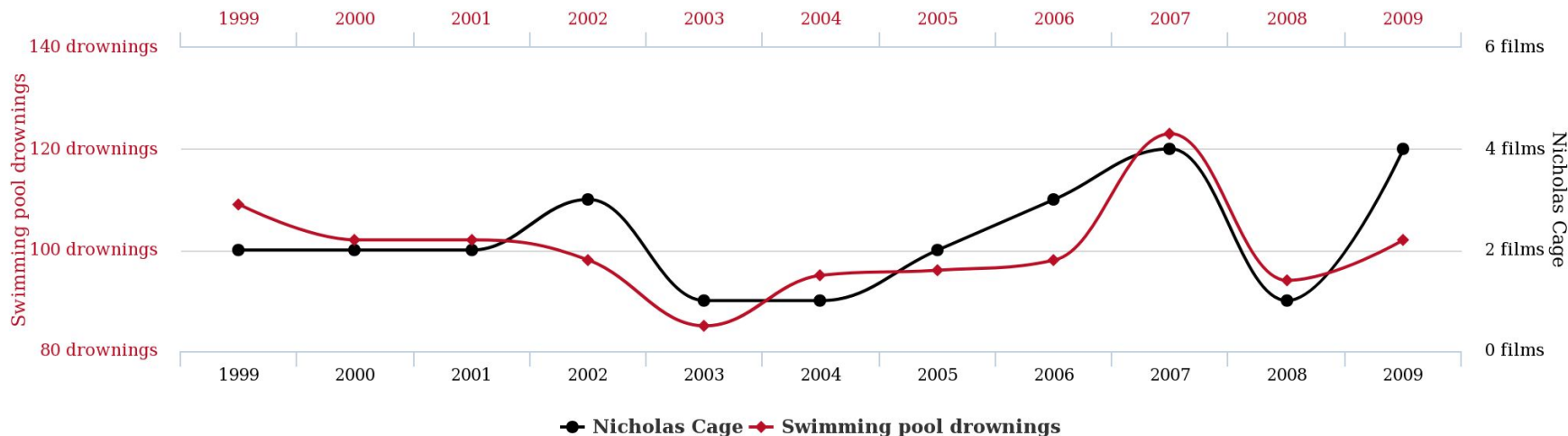
Source: Salkind, N. J. (2008). *Statistics for people who (think they) hate statistics*. 3rd ed. Thousand Oaks: SAGE Publications. (p.89)

Correlations $\neq$ Causation

Number of people who drowned by falling into a pool

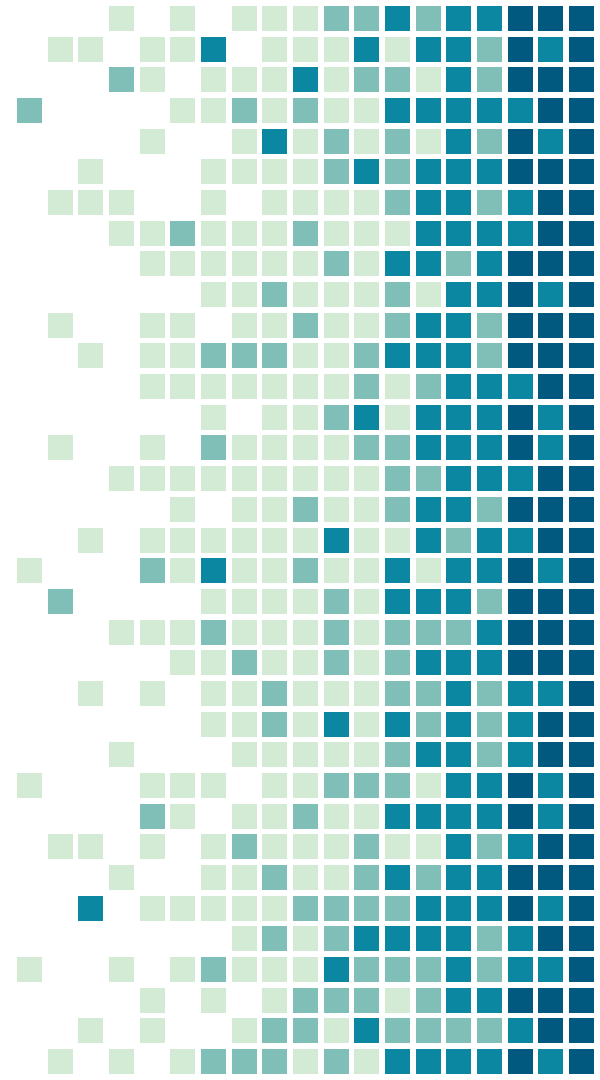
correlates with

Films Nicolas Cage appeared in



Source: <http://tylervigen.com/spurious-correlations>

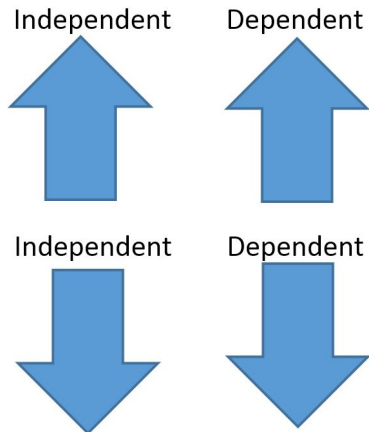
Intro to Inferential Stats



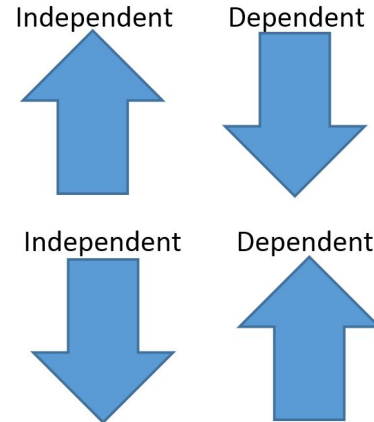
Hypotheses

- Hypothesis and null hypothesis
 - Directional and non-directional hypotheses

POSITIVE

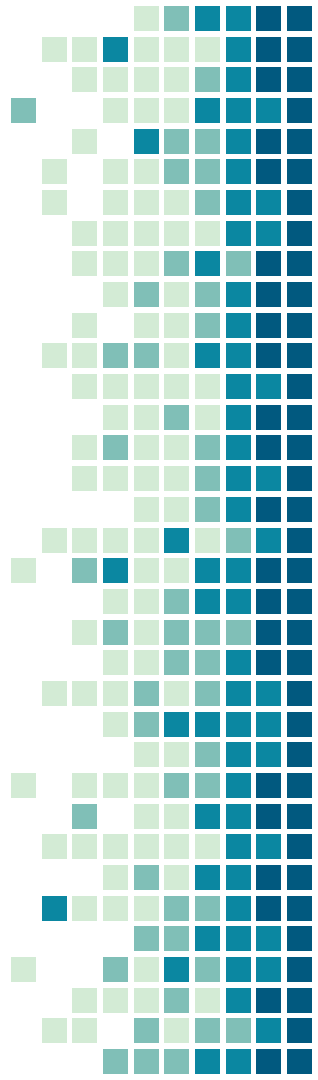


NEGATIVE



Hypotheses

- Hypothesis Testing
- Significance Testing



THANKS!

Any questions?

Resources & References

- Australian Bureau of Statistics. [Understanding Statistics](#).
- Byrne, G. (2007). [A Statistics Primer: Understanding Descriptive and Inferential Statistics](#). Evidence Based Library and Information Practice, 2(1), 32-47.
- Kranzler, J.H. (2011). Statistics for the Terrified.
- Salkind, N. J. (2011). Statistics for people who (think they) hate statistics.
- University of North Texas Libraries. [Statistics for Librarians](#). (4 videos)