Understanding the basics of quantitative research methodology

Define tomorrow.

UNISA | college of education

Start here →

Defining and characterising

Quantitative research

Paradigm

Positivist

Research design

Descriptive, Correlational, Quasi-experimental, Experimental

Participants

Population, sampling techniques, and sample size

Data collection

Questionnaires, validity and reliability (psychometric properties)

Data analysis

Statistical techniques / procedures

Quantitative research emphasizes **objective measurements** and the **statistical, mathematical, or numerical analysis of data** collected through **polls, questionnaires, and surveys**.

Quantitative research focuses on **gathering numerical data and generalizing** it across groups of people or to explain a particular phenomenon (Babbie, 2010).

Your goal…

(More often than not) is to determine the relationship between one thing [an independent variable] and another [a dependent or outcome variable] within a population (this is a bivariate analysis).

Retrieved from: https://libguides.usc.edu/writingguide/quantitative

The main characteristics are…

- The data is usually gathered using **structured research instruments** (such as questionnaires).
- Data are in the form of **numbers and statistics**, often arranged in tables, charts, figures, or other non-textual forms.
- The results are based on **larger sample sizes** that are representative of the population.
- Results can be used to **generalize concepts more widely, predict future results, or investigate causal relationships.**
- The research study can usually be replicated or repeated, given its high reliability.

Retrieved from: https://libguides.usc.edu/writingguide/quantitative

Defining and characterising

Quantitative research

Next → **Paradigm**

Positivist

Research design

Descriptive, Correlational, Quasi-experimental, Experimental

Participants

Population, sampling techniques, and sample size

Data collection

Questionnaires, interviews (structured), meta-analysis

Valid and reliable (psychometric properties)

Data analysis

Statistical techniques / procedures

Ontology: How do you view knowledge? Is knowledge constructed by an individual, between individuals, or is knowledge law-like and stable regardless of who views it? Do you believe knowledge exists outside of the knower, or do you believe the knower constructs knowledge?

Epistemology: How will the relationship between the you (as the researcher) and the phenomena (you wish to study) be characterised?

Methodology: How you will go about investigating that which is to be known?

Mason (2018)

PARADIGM/DIMENSION	ONTOLOGY	EPISTEMOLOGY	METHODOLOGY
Positivist (Quantitative)	• **Stable external reality** • **Law like**	• **Objective** • **Detached observer**	• **Experimental** • **Quantitative** • **Hypothesis testing**
Interpretive	• Internal reality • Subjective experience	• Empathetic • Observer • Intersubjective	• Interactional • Interpretive • Qualitative
Constructionist	• Socially constructed reality • Discourse	• Suspicious • Political • Observer constructed versions	• Deconstruction • Textual analysis • Discourse analysis

Mason (2018)

In a nutshell…

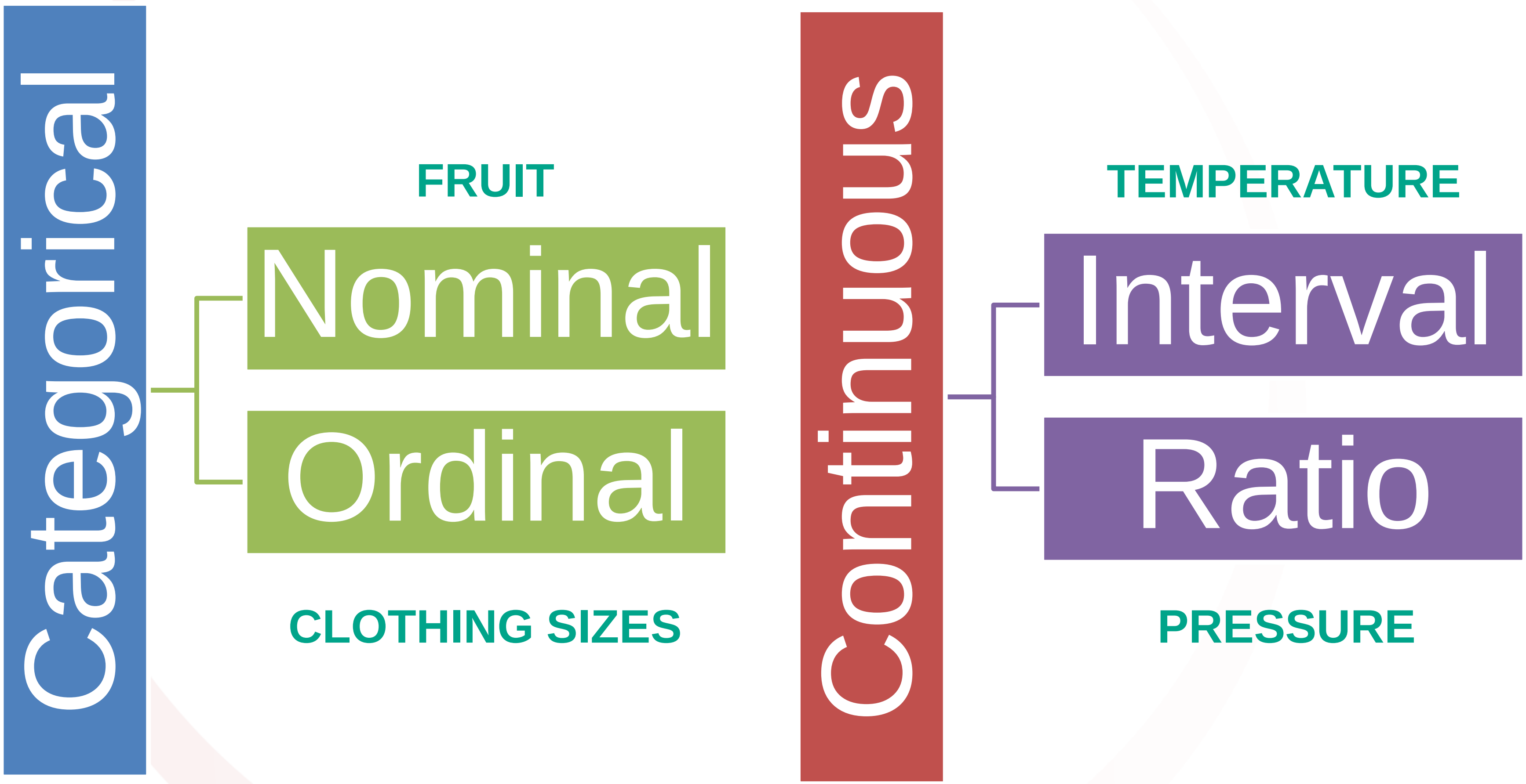
…The quantitative researcher regards knowledge as existing objectively (ontology) and therefore adopts the role of an impartial and detached observer (epistemology) who utilises valid and reliable instruments, such as questionnaires (methodology) to collect data (Clark Carter, 2010)…

…The researcher collects data objectively, assumes a detached position and adopts methods which allow for generalisations to the broader population…

Before getting into the designs, let's have a look at some common terms…

- What is a variable?
- Categorical and continuous variables
- Levels of measurement
- Independent variables and dependent variables (Predictor and outcome variables)
- Mediating and moderating variables

A ***variable*** is an entity that ***varies*** and can take on a ***variety of values***



Understanding variables in quantitative research

Independent variable (IV)

Variable(s) that affect or cause the outcome to occur

The variable that is being manipulated

Another name in a Regression study = predictor variable

Another name in a True experimental design = cause

Dependant variable (DV)

Variable that is influenced by, or acted upon, by the independent variable

The variable that is not being manipulated

Another name in a Regression study = outcome variable

Another name in a True experimental design = effect

Mason (2018)

Mediating variables

Explains the relationship between the DV and the IV
Accounts for the relationship

Moderating variables

Influence the direction and the strength of the relationship between the IV and the DV

Check out this useful video which explains in detail

https://www.youtube.com/watch?v=WZr1jlKi_s0

Defining and characterising

Quantitative research

Paradigm

Positivist

Next →

Research design

Descriptive, Correlational, Quasi-experimental, Experimental

Participants

Population, sampling techniques, and sample size

Data collection

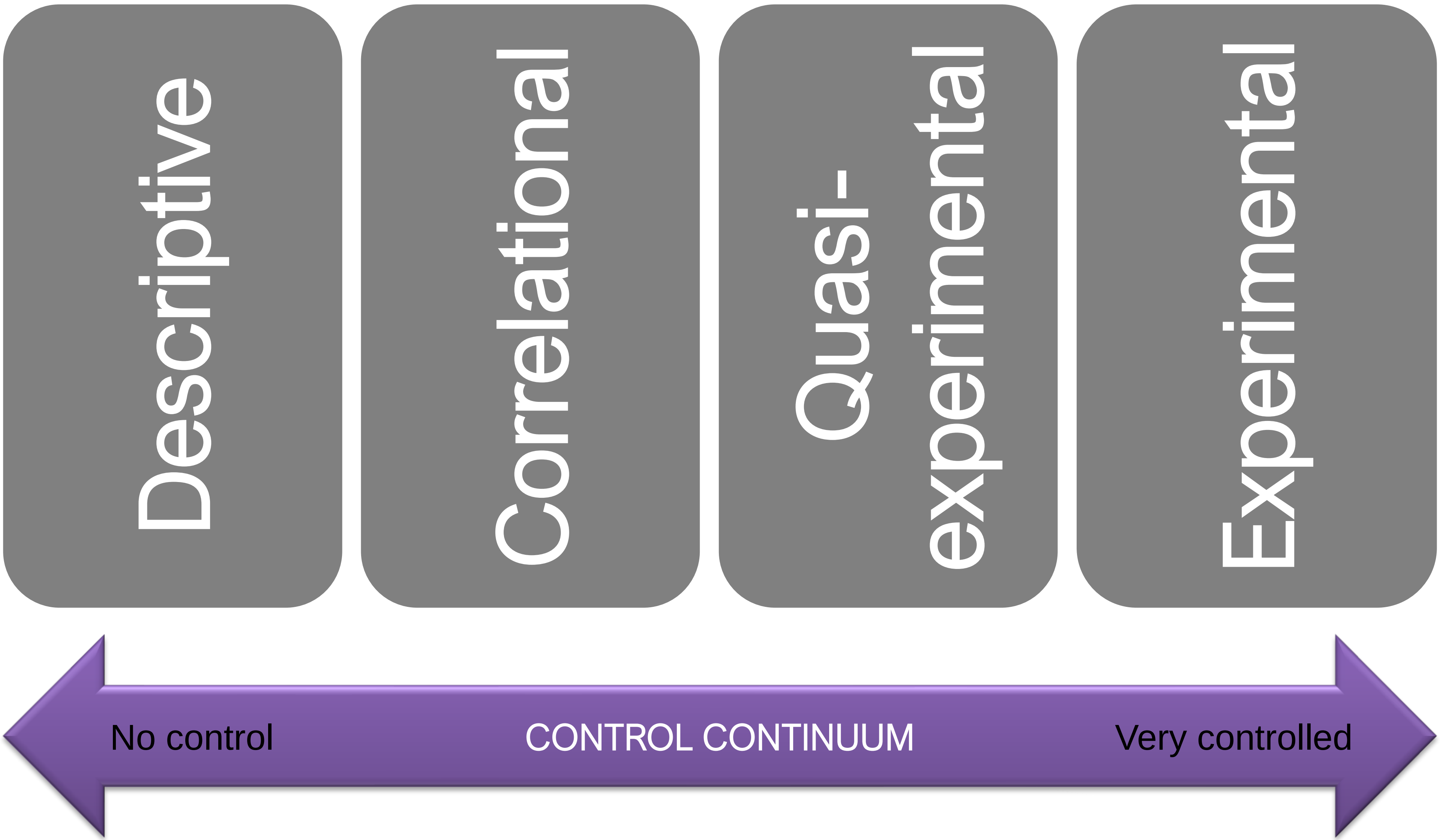
Questionnaires, interviews (structured), meta-analysis

Valid and reliable (psychometric properties)

Data analysis

Statistical techniques / procedures

Four prominent {quantitative} designs



Experimental

Problem = External validity

- Determines causality (i.e., cause and effect)
- Manipulation or intervention of the IV
- Random assignment to control and experimental groups
- *Often (not always)* conducted in laboratory settings

Correlational

Problem = Cannot determine causality

- Does not determine causality. This design measures relationships between two or more variables
- Focused on the collection of numerical data that can be analysed using statistics
- No manipulation or intervention of the IV

Quasi-experimental / causal-comparative

Problem = Internal validity

- Shares similarities with the experimental design, but it specifically lacks the element of random assignment to the treatment or control group (these groups already exist)
- Intervention
- *Often* no control group
- Utilised outside laboratory contexts; in real world settings

Defining and characterising

Quantitative research

Paradigm

Positivist

Research design

Descriptive, Correlational, Quasi-experimental, Experimental

Next →

Participants

Population, sampling techniques, and sample size

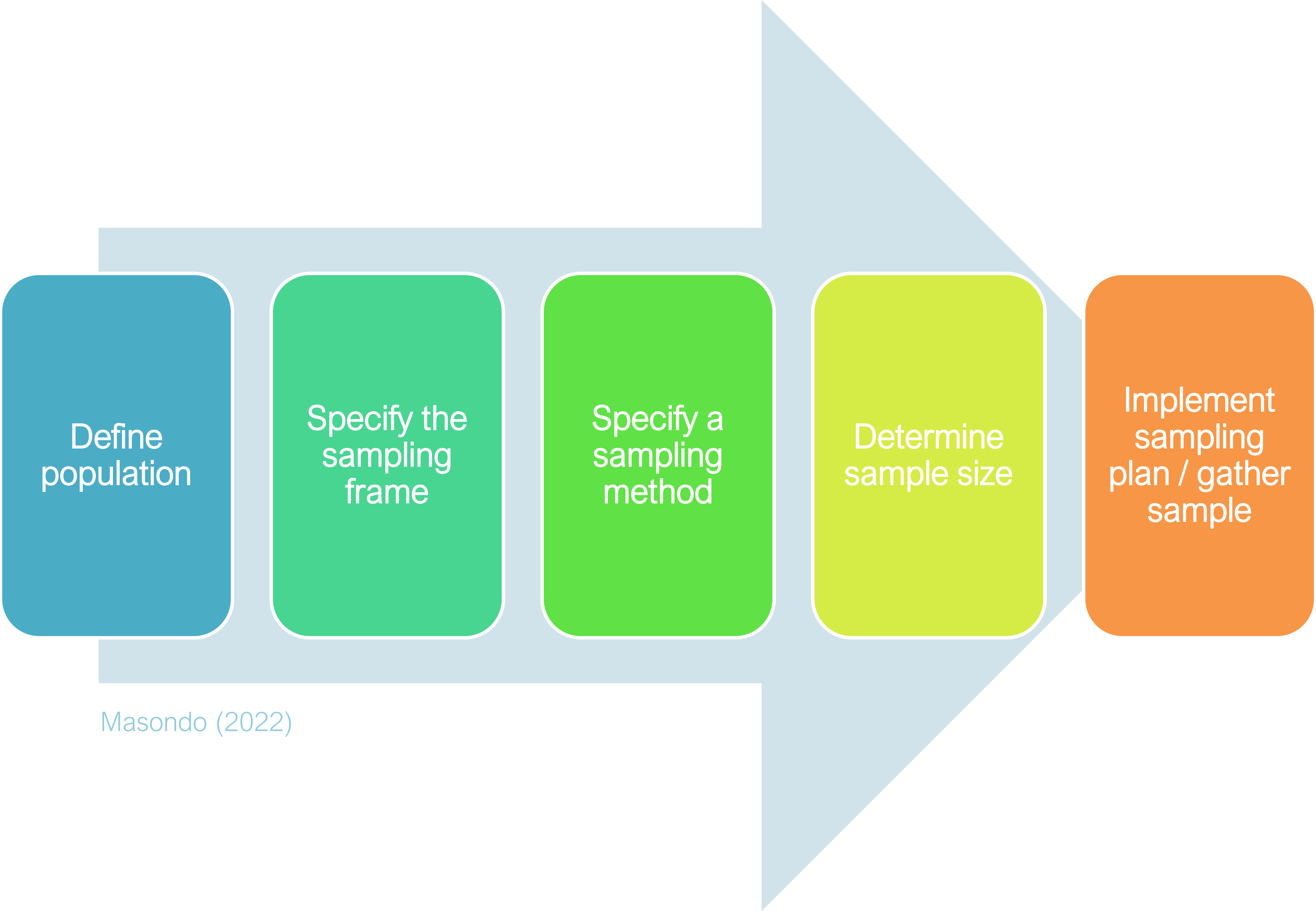
Data collection

Questionnaires, interviews (structured), meta-analysis

Valid and reliable (psychometric properties)

Data analysis

Statistical techniques / procedures



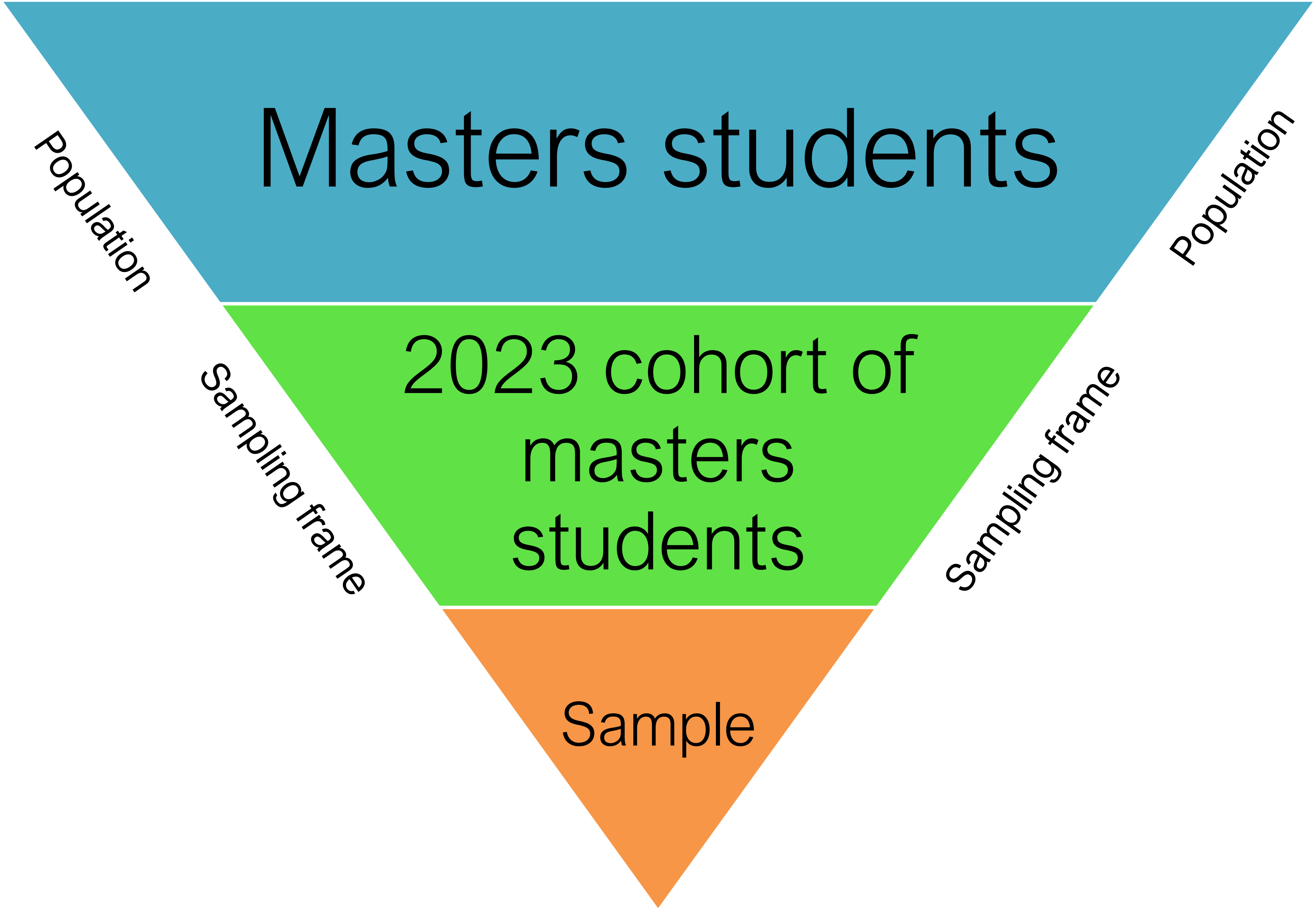
Masondo (2022)

A population is a group of individuals with some commonality (i.e., they all have something in common);

A population is a set of all possible units of interest that meet the designated criteria of investigation;

A research population is also known as a well-defined collection of individuals or objects known to have similar characteristics. All individuals or objects within a certain population usually have a common, binding characteristic or trait.

Mason (2018)


Masters students
2023 cohort of masters students
Sample
Population
Sampling frame
Population
Sampling frame

There are two main types of sampling techniques:

1. probability sampling
2. non-probability sampling

What's the difference between the two approaches?

Answer: Randomization

Randomization occurs when all members of the sampling frame (or population) have an *equal opportunity (or probability)* of being selected to participate in the study. This is a feature characteristic of **probability sampling**.

Probability sampling

SIMPLE RANDOM SAMPLING

1. Get a list of all possible participants (i.e. your sampling frame)
2. Generate random numbers for each
3. Sort the random numbers by size and then select the required amount of participants from the 1st row

STRATIFIED RANDOM SAMPLING

1. Separate your sampling frame into groups (also called strata)
2. Do either a simple random sample or a system random sample from there

SYSTEMATIC RANDOM SAMPLING

1. Get your sampling frame
2. Select a random start and then select every *kth* person
3. *k* is determined by dividing the population (or sampling frame) by the desired sample size

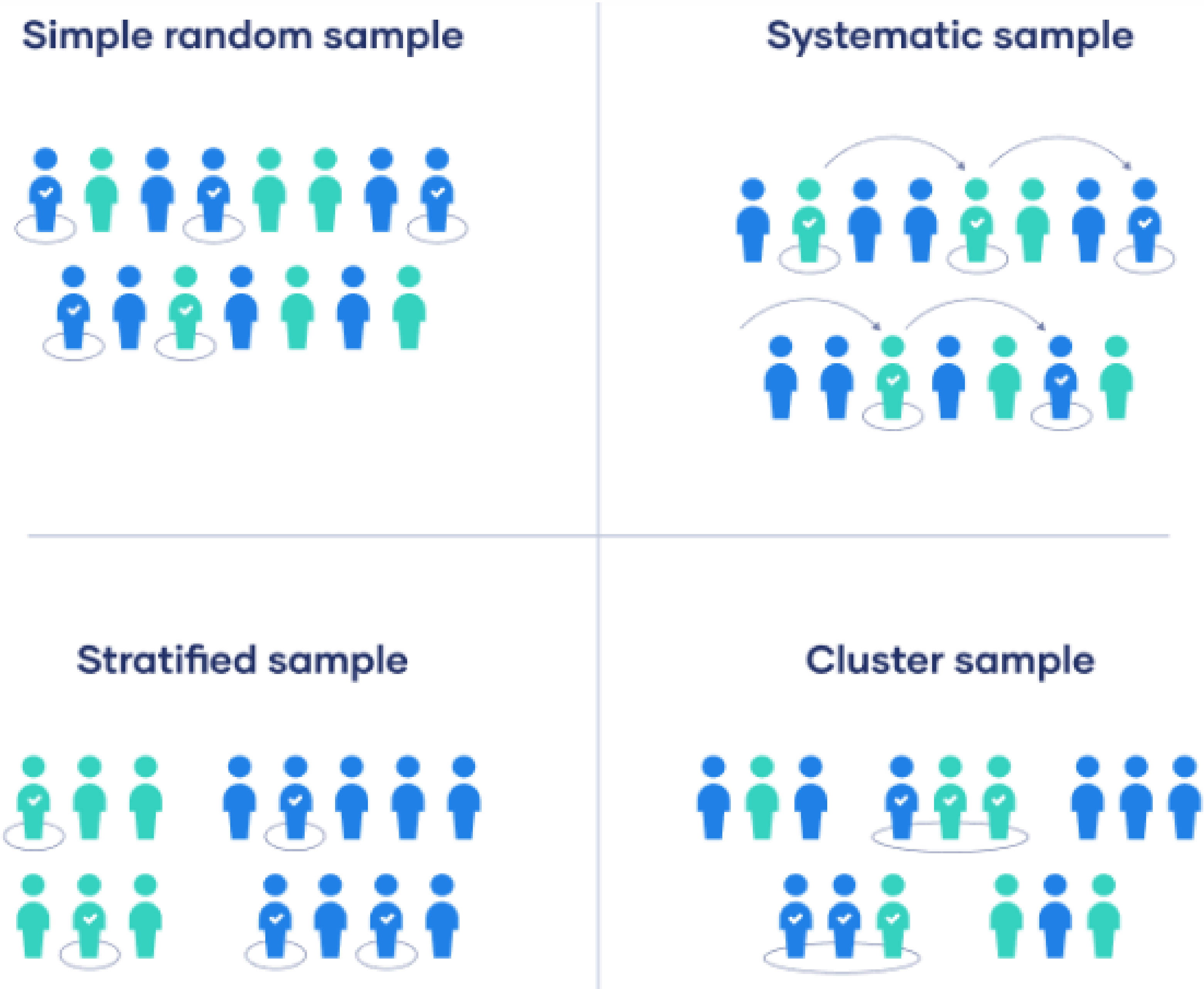


Image retrieved from https://www.scribbr.com/methodology/sampling-methods/

Non-probability Sampling

CONVENIENCE SAMPLING

1. Find some people that are easy to find and sample them

SNOWBALL SAMPLING

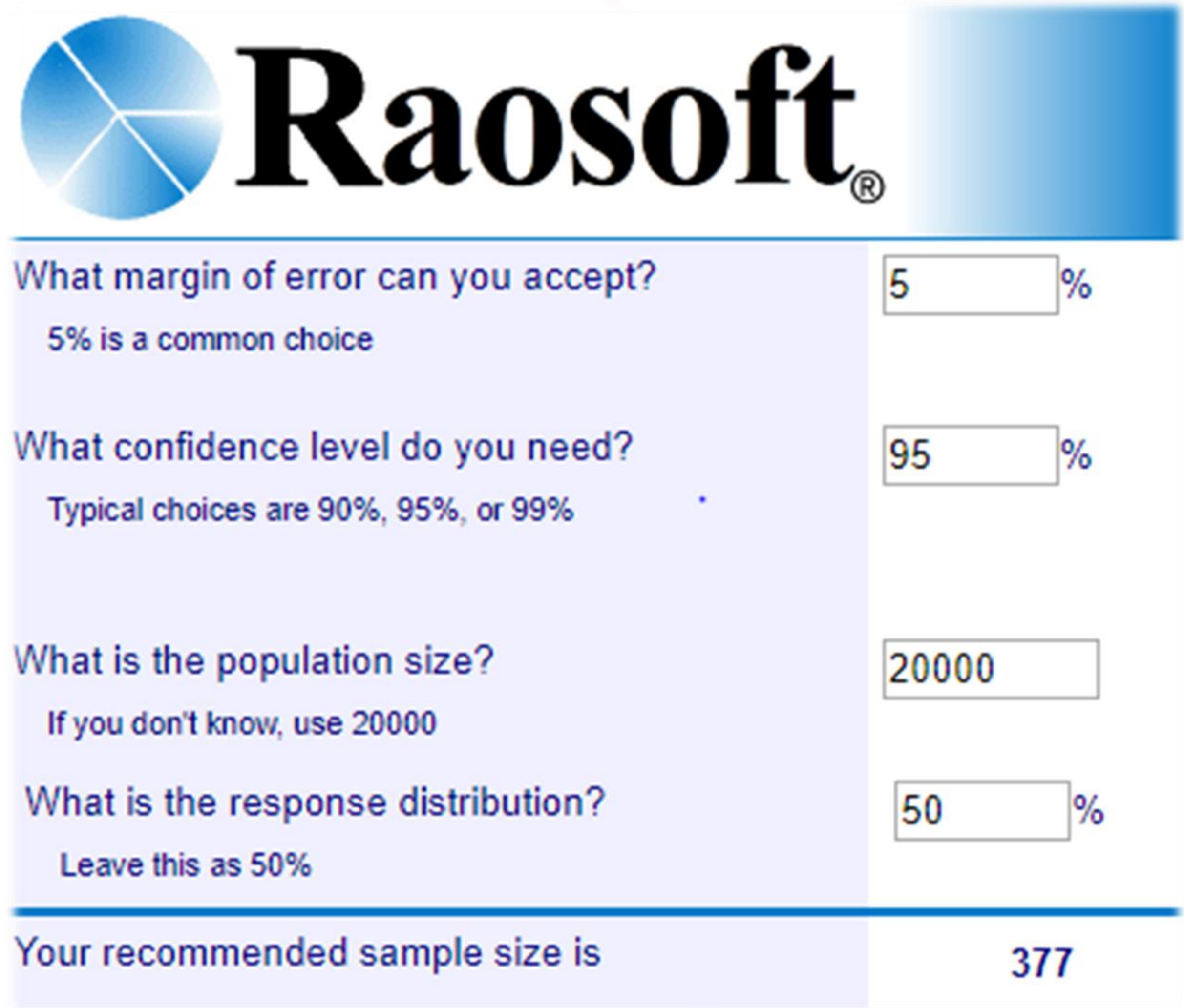
1. Find a few people that are relevant to your topic
2. Ask them to refer you to more of them

QUOTA SAMPLING

1. Determine what the population looks like in terms of specific qualities
2. Create quotas based on those qualities
3. Select people from each quota

Now that you have a better idea of sampling techniques, you can use this handy website to determine the sample size needed for your study.

http://www.raosoft.com/samplesize.html



Defining and characterising

Quantitative research

Paradigm

Positivist

Research design

Descriptive, Correlational, Quasi-experimental, Experimental

Participants

Population, sampling techniques, and sample size

Next →

Data collection

Questionnaires, interviews (structured), meta-analysis

Valid and reliable (psychometric properties)

Data analysis

Statistical techniques / procedures

Typical quantitative data gathering strategies include…

- Experiments or clinical trials;
- Observing and recording well-defined events;
- Obtaining relevant data from management information systems (such as learner management systems – LMS)
- But, the most common data gathering technique in the social sciences is the administration of surveys/questionnaires with closed-ended questions

When administering surveys…

Ensure your instrument is valid and reliable (How do you do that? You establish the psychometric properties of the instrument among your sample to determine how valid and reliable the instrument is). So, ask yourself:

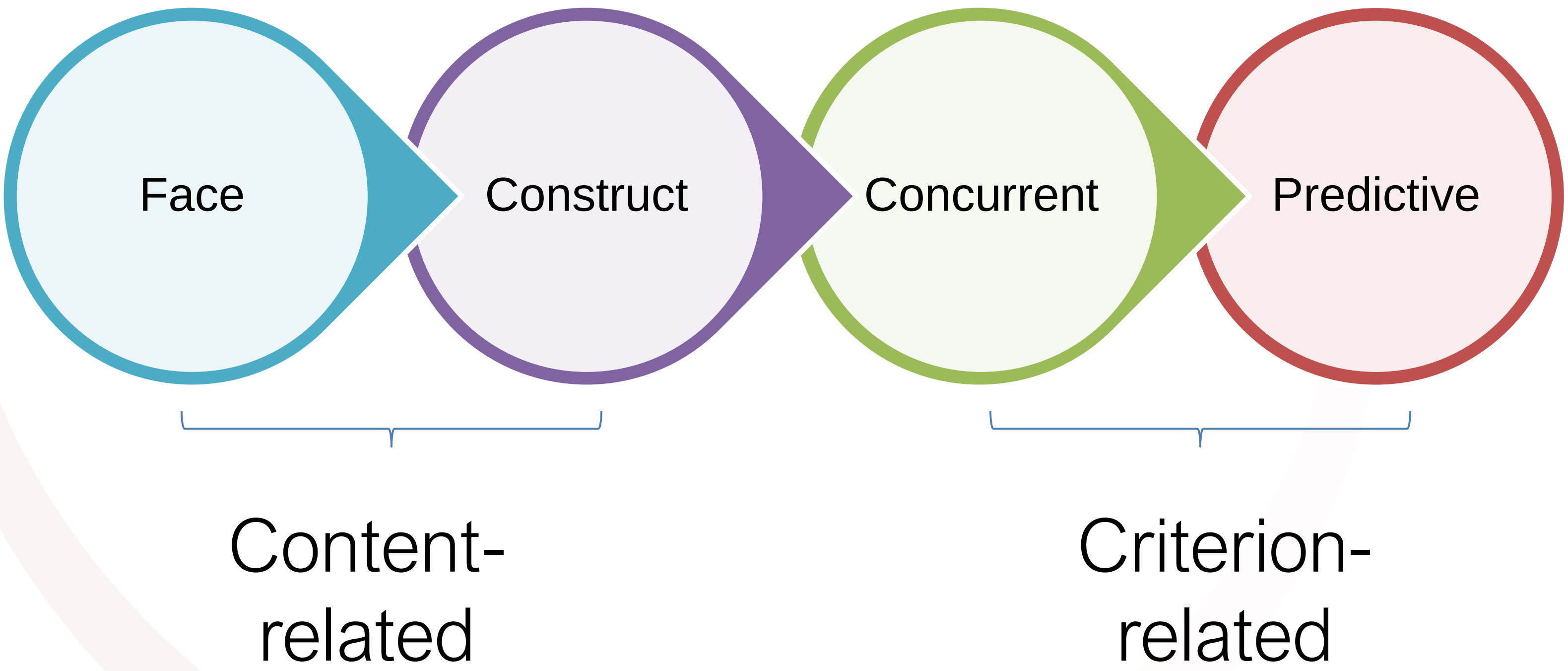
Is the instrument measuring what it intends to measure?
To determine this, you could conduct an exploratory factor analyses to explore the construct validity

Is the instrument reliable?
To determine this, you could conduct item analyses to explore the scales' reliability (also called internal consistency)

https://www.youtube.com/watch?v=UWP9OEoaNnE&list=PL7Tw2kQ2edvpCOKL40H_kHolyaiHRsW4Q

Understanding the basics of validity

Validity simply means that a test or instrument is accurately measuring what it's supposed to. Two broad categories (content- and criterion-related) include the following (McLeod, 2013):



https://www.simplypsychology.org/validity.html

See the link for explanations of each

Understanding the basics of reliability

Reliability refers to the degree to which an instrument or survey produces consistent (or similar) results.

Some of the more common types (Middleton, 2023):

- Test-retest reliability = This is when the same test is given to the same group of people at two different times. Results from the first iteration are then compared to the results from the second iteration to determine consistency
- Parallel-forms reliability = This is when the same group of people complete two similar versions of a questionnaire. Results from the two versions are then compared in order to determine the consistency of the results.
- Internal consistency reliability = Internal consistency assesses the correlation between multiple items in a test that are intended to measure the same construct. Cronbach's Alpha (α) coefficient is the most widely used internal-consistency coefficient. See the next slide for α cut-offs…

Interpreting Cronbach's Alpha (α)

Rule of thumb:
For low stakes tests: > 0.70. Some authors suggest this figure should be above 0.80

Cronbach's alpha	Internal consistency
$\alpha \geq 0.9$	Excellent
$0.9 > \alpha \geq 0.8$	Good
$0.8 > \alpha \geq 0.7$	Acceptable
$0.7 > \alpha \geq 0.6$	Questionable
$0.6 > \alpha \geq 0.5$	Poor
$0.5 > \alpha$	Unacceptable

Rule of thumb:
For high stakes tests, this figure should be above .90.

Defining and characterising

Quantitative research

Paradigm

Positivist

Research design

Descriptive, Correlational, Quasi-experimental, Experimental

Participants

Population, sampling techniques, and sample size

Data collection

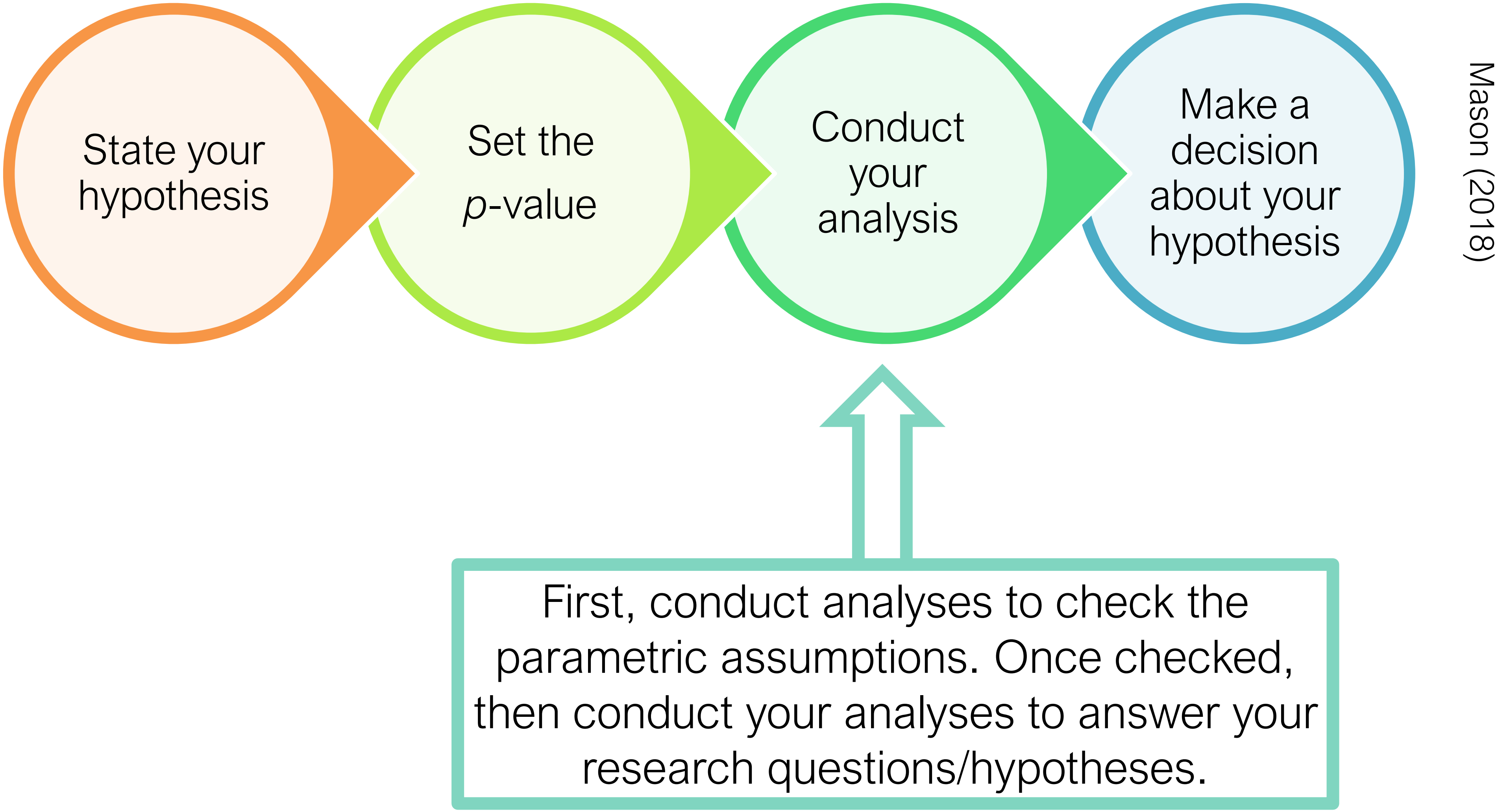
Questionnaires, interviews (structured), meta-analysis

Valid and reliable (psychometric properties)

Lastly →

Data analysis

Statistical techniques / procedures


State your hypothesis
Set the *p*-value
Conduct your analysis
Make a decision about your hypothesis
Mason (2018)
First, conduct analyses to check the parametric assumptions. Once checked, then conduct your analyses to answer your research questions/hypotheses.

What’s a hypothesis?

A **hypothesis** is a tentative, testable statement

What’s the difference between the null and the alternative hypothesis?

Null hypothesis:

- States that no difference/relationship exists between the variables

Alternative hypothesis:

- States that a difference/relationship exists between the variables
  - Can be stated in two ways

Stating your hypothesis…

H_0 = **Null hypothesis** (states that *there is no* relationship/difference)

H_0: There is no relationship between exam performance and exam anxiety

H_1 = **Alternative hypothesis** (states that *there is* a relationship/difference).
There are two types of H_1's

Non-directional H_1 : (*doesn't* state the direction of the relationship)
There is a significant relationship between exam performance and exam anxiety

Directional H_1: (*does* state the direction of the relationship)
There is significant negative relationship between exam performance and exam anxiety

Set the *p*-value…

- Also called an alpha, level of significance, statistical significance level

- Refers to probability level (*p*-value)

- **Typically set at 0.05**

- Interpretation: 0.05 = 95% chance that the relationship/effect did not occur by chance. ***But***, there is a 5% chance that the relationship / effect / association was random

Check the assumptions associated with your planned statistical technique and how to perform them in SPSS…

- First and foremost – Normally distributed data (https://www.youtube.com/watch?v=2GRZ_d4ftoo)
- Equal variances / homogeneity (https://www.youtube.com/watch?v=O6taUIWejB0)
- Independence
- Data are measured at the interval or ratio scale

	Parametric	Non-parametric
Assumed distribution	Normal	Any
Assumed variance	Homogeneous	Any
Typical data	Ratio or interval	Ordinal or nominal
Data set relationships	Independent	Any
Statistical test		
Correlation	Pearson	Spearman
Independent measures, 2 groups	Independent t-test	Mann-Whitney test
Independent measures, >2 groups	One-way ANOVA	Kruskal-Wallis test
Repeated measures, 2 conditions	Dependent t-test	Wilcoxon test
Repeated measures, >2 conditions	One-way, repeated measures ANOVA	Friedman's test

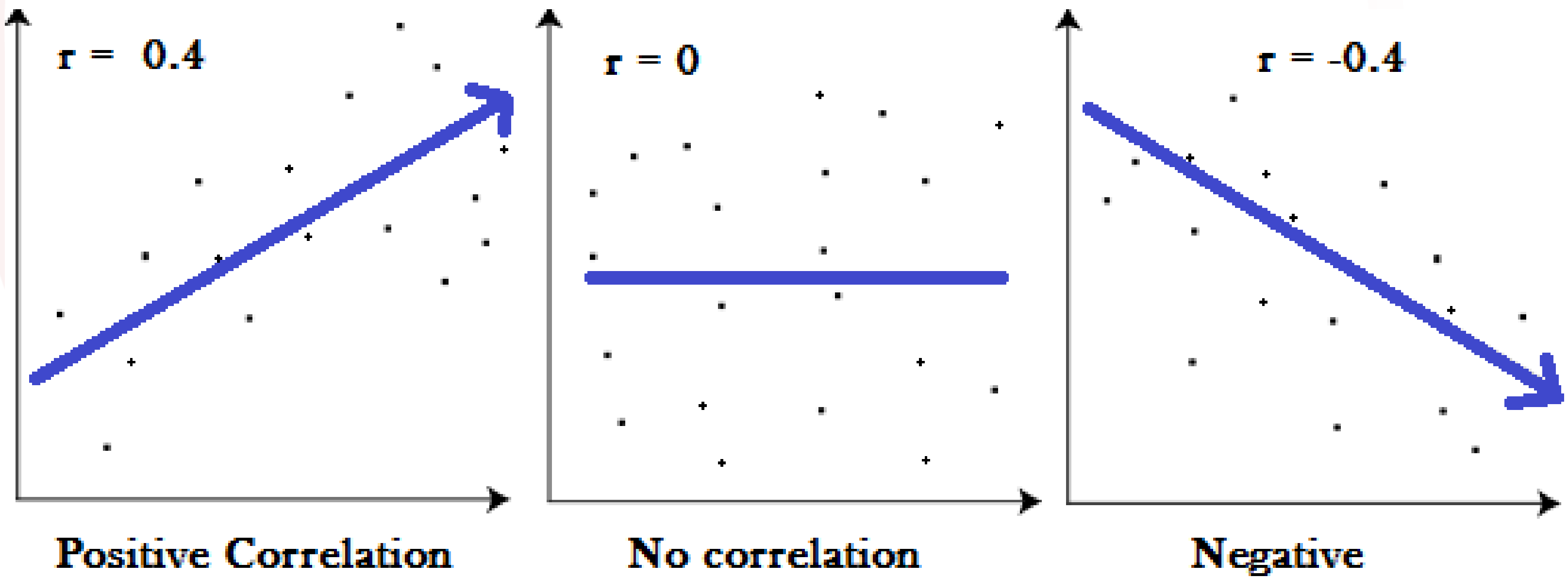
Table adapted from: http://changingminds.org/explanations/research/analysis/parametric_non-parametric.htm

Just in case you're not sure yet (on which statistical test to use for your data)…

https://www.socscistatistics.com/tests/what_stats_test_wizard.aspx

Correlation

- Statistical techniques used to describe the relationship/association between variables
  - Positive, negative or no correlation (ranging from -1 to +1)
    - Zero indicates no relationship



Regression

Taking correlations one step further … Predicting one variable from another

Simple regression: ONE predictor variable; CONTINUOUS outcome variable

Multiple regression: SEVERAL predictor variables; CONTINUOUS outcome variable

Binary logistic regression: ONE OR MORE predictor variables; CATEGORICAL outcome variable (Only two outcomes)

Multinomial logistic regression: ONE OR MORE predictor variables; CATEGORICAL outcome variable (More than two outcomes)

But, what if I'm not interested in exploring relationships or associations, but rather **assessing whether two groups are different, or have different means/averages**?

Well, check the next slide…

Independent measures (independent *t*-test, Mann-Whitney)

Used to compare the mean scores of two different groups.

Are the anxiety scores different between 12 spider-phobes who see a picture of a spider and 12 different spider-phobes who see a real-life tarantula?

IV: People who see a spider picture vs people who see a real spider (**two groups**)

DV: Scores on spider anxiety scale (**one DV**)

Repeated measures (dependent *t*-test, Wilcoxon)

Used to compare the mean scores of the same group

Is there a difference between the mean anxiety scores from spider-phobes who were exposed to a picture of a spider and a real live tarantula?

IV: Picture and real-life exposure (**two conditions**)

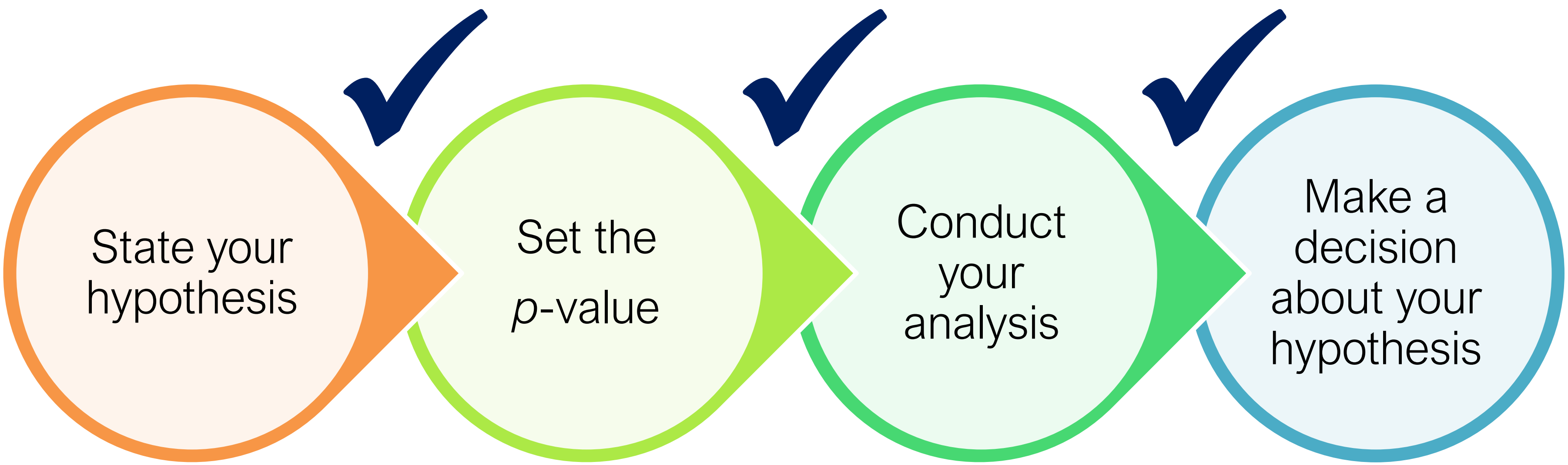
DV: Scores on spider anxiety scale (**one DV**)

One-way (independent-measures) ANOVA or Kruskal-Wallis…

Similar to *t*-test, but used with more than two independent groups.

One-way (repeated-measures) ANOVA or Friedman's ANOVA…

Similar to *t*-test, but used with more than two dependent conditions.


State your hypothesis
Set the *p*-value
Conduct your analysis
Make a decision about your hypothesis

Make a decision about the hypotheses

So, we've conducted statistical analyses on our data. Now what?

Well, each statistical procedure provides a test statistic and its associated *p*-value. You'll need to use these statistics to make a decision about your hypotheses.

The *p*-value indicates the statistical significance of the result – did it occur by chance or not?

If you set your *p*-value at 0.05 and your results reveal a *p*-value of:

$p \leq .05$ – then your results are statistically significant and you must reject the H_o

$p > 0.05$ – then your results are not statistically significant and you must accept the H_o

An example

Your research question: Is there a relationship between exam performance and exam anxiety?
Your research hypothesis: There is a significant negative relationship between exam performance and exam anxiety (H_1)

If you set your *p*-value at 0.05 and your results reveal a *p*-value of:

$p \leq .05$ – then your result is statistically significant and you should reject the H_o – *So, there* ***is*** *a real relationship between exam performance and exam anxiety.*
$p > 0.05$ – then your result is not statistically significant and you should accept the H_o - *So, there* ***is no*** *real relationship between exam performance and exam anxiety.*

		Exam performance (%)	Exam Anxiety	Time spent revising
Exam performance (%)	Pearson Correlation	1.000	-.441**	.397**
	Sig. (1-tailed)	.	.000	.000
	N	103	103	103
Exam Anxiety	Pearson Correlation	-.441**	1.000	-.709**
	Sig. (1-tailed)	.000	.	.000
	N	103	103	103

Note: the *p*-value is the 'Sig. (1-tailed)'

Researchers could make errors at this stage…

Type I error "Rejecting the H_o if it was in fact true"

Type II error "Accepting the H_o if it was false"

Type I error

- Rejecting the H_o if it was in fact true;
- Thus, there was no relationship between the two constructs, but the results and *p*-value *erroneously* suggested that there was.

Type II error

- Failure to reject the H_o if it was false;
- There was in fact a relationship between the two constructs, but the results and *p*-value indicated that there were no differences.

Some key tips to keep in mind when reporting your results:

1. **Explain the data collected as well as the results** in relation to the research problem you are investigating.
2. **Report unanticipated events** that occurred during your data collection. Explain how the actual analysis differs from the planned analysis. Explain your handling of missing data and why any missing data does not undermine the validity of your analysis.
3. **Explain the techniques** you used to clean your dataset.
4. **Describe the assumptions for each statistical test** and the steps you took to ensure that they were not violated.

Retrieved from: https://libguides.usc.edu/writingguide/quantitative

5. **Choose an appropriate statistical procedure**; provide a rationale for its use and a reference for it. Also, specify any computer programs used, such as SPSS or MS Excel.
6. **When using inferential statistics**, provide the descriptive statistics, confidence intervals, and sample sizes for each variable as well as the value of the test statistic, its direction, the degrees of freedom, and the significance level [the *p* value].
7. **Avoid inferring causality** (unless, of course, you adopted an experimental design - which proves causality)

Retrieved from: https://libguides.usc.edu/writingguide/quantitative

write

Sources

- Babbie, E. (2010) The Practice of Social Research. 12th ed. Belmont, CA: Wadsworth Cengage (as cited in https://libguides.usc.edu/writingguide/quantitative)
- Clark-Carter, D. (2010). Quantitative psychological research: The complete student's companion (3rd Ed.). New York, NY: Psychology Press.
- Field, A. (2013). Factor analysis and PCA. Retrieved from https://www.youtube.com/watch?v=UWP9OEoaNnE&list=PL7Tw2kQ2edvpCOKL40H_kHolyaiHRsW4Q
- Gaskin, J. (2018). Mediator or Moderator? Retrieved from https://www.youtube.com/watch?v=WZr1jlKi_s0
- Grande, T. (2015). Normality Tests in SPSS. Retrieved from https://www.youtube.com/watch?v=2GRZ_d4ftoo
- Löfgren, K. (2013). Test for equality of variances: Parametric and nonparametric Levene's test in SPSS. Retrieved from https://www.youtube.com/watch?v=O6taUlWejB0
- Mason, H. (2018). Research methodology. Pretoria, The South African College of Applied Psychology (SACAP)
- Masondo, P. (2022). Sample selection techniques. College of Graduate Studies, University of South Africa (UNISA)
- McLeod, S. (2013). What is validity? Simply Psychology. Retrieved from www.simplypsychology.org/validity.html
- Middleton, F. (2023). The 4 Types of Reliability in Research: Definitions & Examples. Scribbr. Retrieved from https://www.scribbr.com/methodology/types-of-reliability/
- Parametric and non-parametric tests (Table). Retrieved from http://changingminds.org/explanations/research/analysis/parametric_non-parametric.htm
- Probability sampling (Image). Retrieved from https://www.scribbr.com/methodology/sampling-methods/
- Raosoft sample size calculator (Image). Retrieved from http://www.raosoft.com/samplesize.html
- University of Southern California (2023). USC Libraries: Research Guides - Organizing Your Social Sciences Research Paper. Retrieved from https://libguides.usc.edu/writingguide/quantitative
- Which Statistics Test Should I Use? (2023). Social Science Statistics. Retrieved from https://www.socscistatistics.com/tests/what_stats_test_wizard.aspx