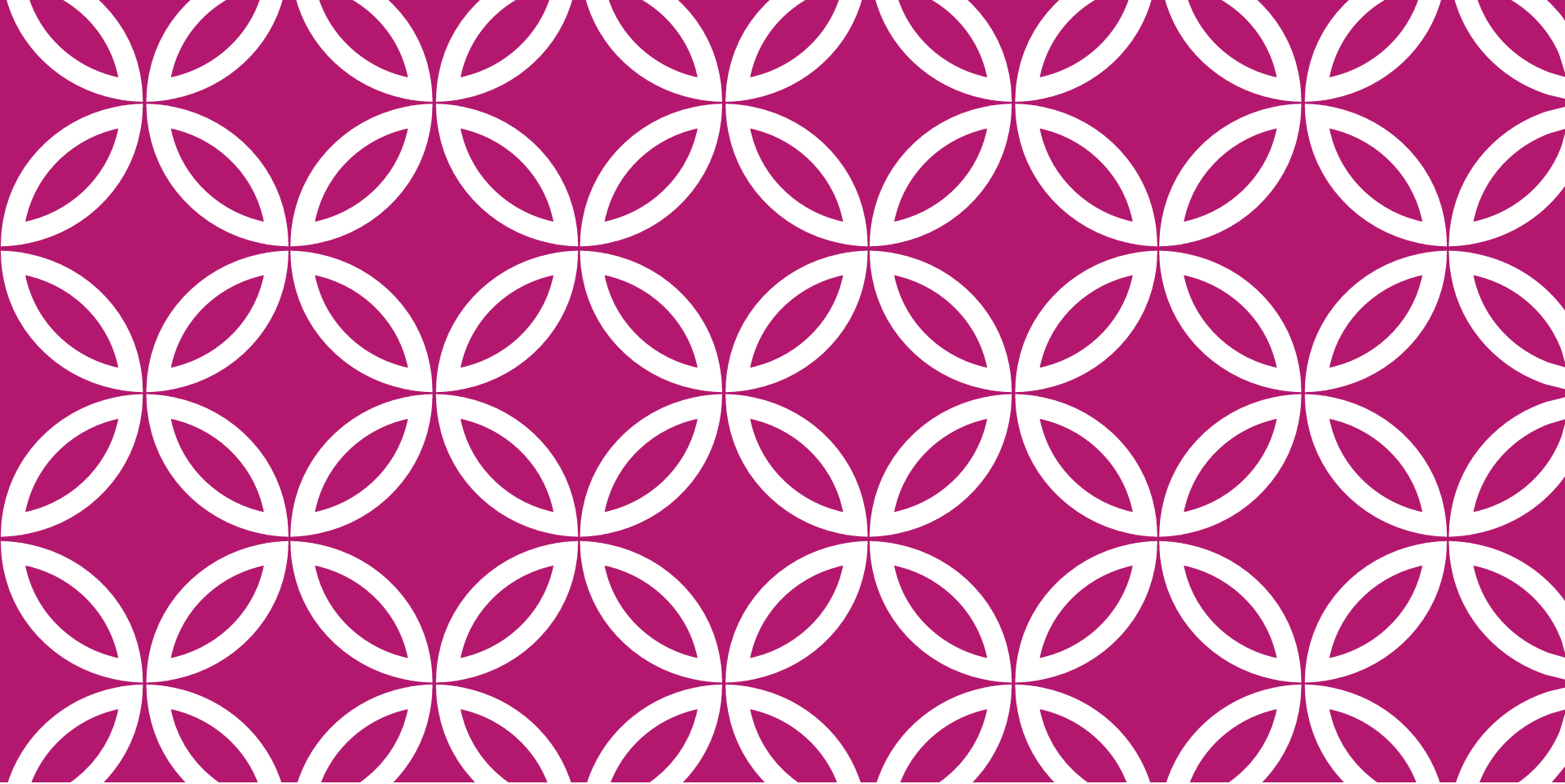




Quantitative Research

WHAT USE QUANTITATIVE RESEARCH AND WHY MIGHT I USE IT?

Quantitative research is the process of collecting and analyzing numerical data. It can be used to find patterns and averages, make predictions, test relationships, test hypothesis and potentially generalise results.

Quantitative research question examples

What is the demographic makeup of Wales in 2023?

How has school exclusion rates changed in Wales over the last 10 years?

Does gamification affect the engagement levels of pupils?

Does pupil group work increase productivity for lower ability learners?

Does regular retrieval practice impact learners test scores?

WHY USE QUANTITATIVE RESEARCH METHODS?

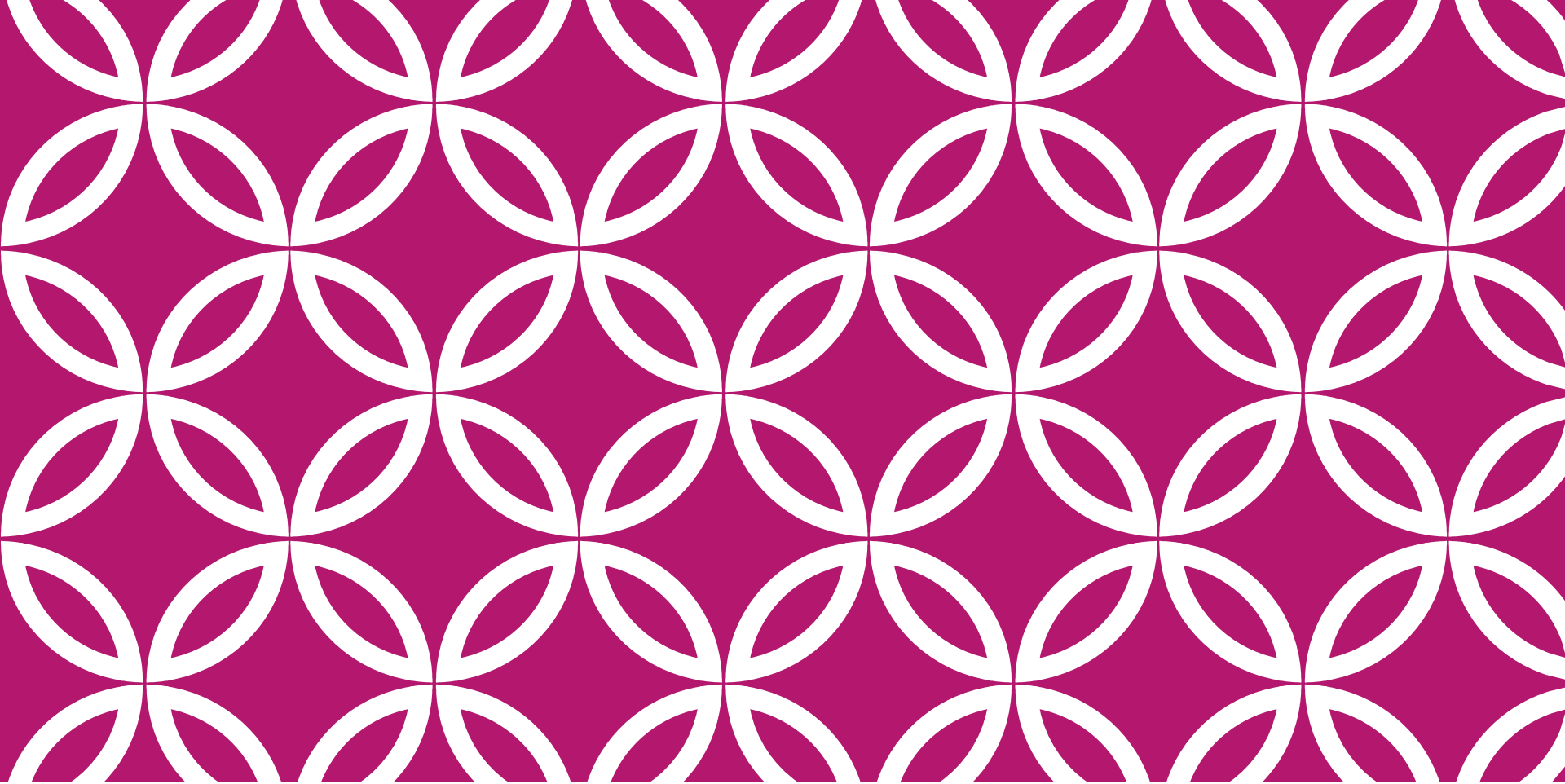
Quantitative research methods (taken from [What Is Quantitative Research? | Definition, Uses & Methods \(scribbr.com\)](#))

You can use quantitative research methods for descriptive, correlational or experimental research.

- In [descriptive research](#), you simply seek an overall summary of your study variables.
- In [correlational research](#), you investigate relationships between your study variables.
- In [experimental research](#), you systematically examine whether there is a cause-and-effect relationship between variables.

Correlational and experimental research can both be used to formally [test hypotheses](#), or predictions, using statistics. The results may be generalized to broader populations based on the [sampling method](#) used.

To collect quantitative data, you will often need to operationalise your variables, i.e. have definitions that translate abstract concepts (e.g. engagement, enjoyment, participation) into observable and quantifiable measures



Quantitative Data
Handling

Types of data

Quantitative

Qualitative

Discrete

Continuous

Levels/Scales of Measurement

Recognising the levels of measurement will help you to decide *how* to interpret the data.

Knowing the level of measurement helps you, as a researcher, decide what statistical analysis is appropriate on the values that you have collected and what graph representation should be used to present the data/results.

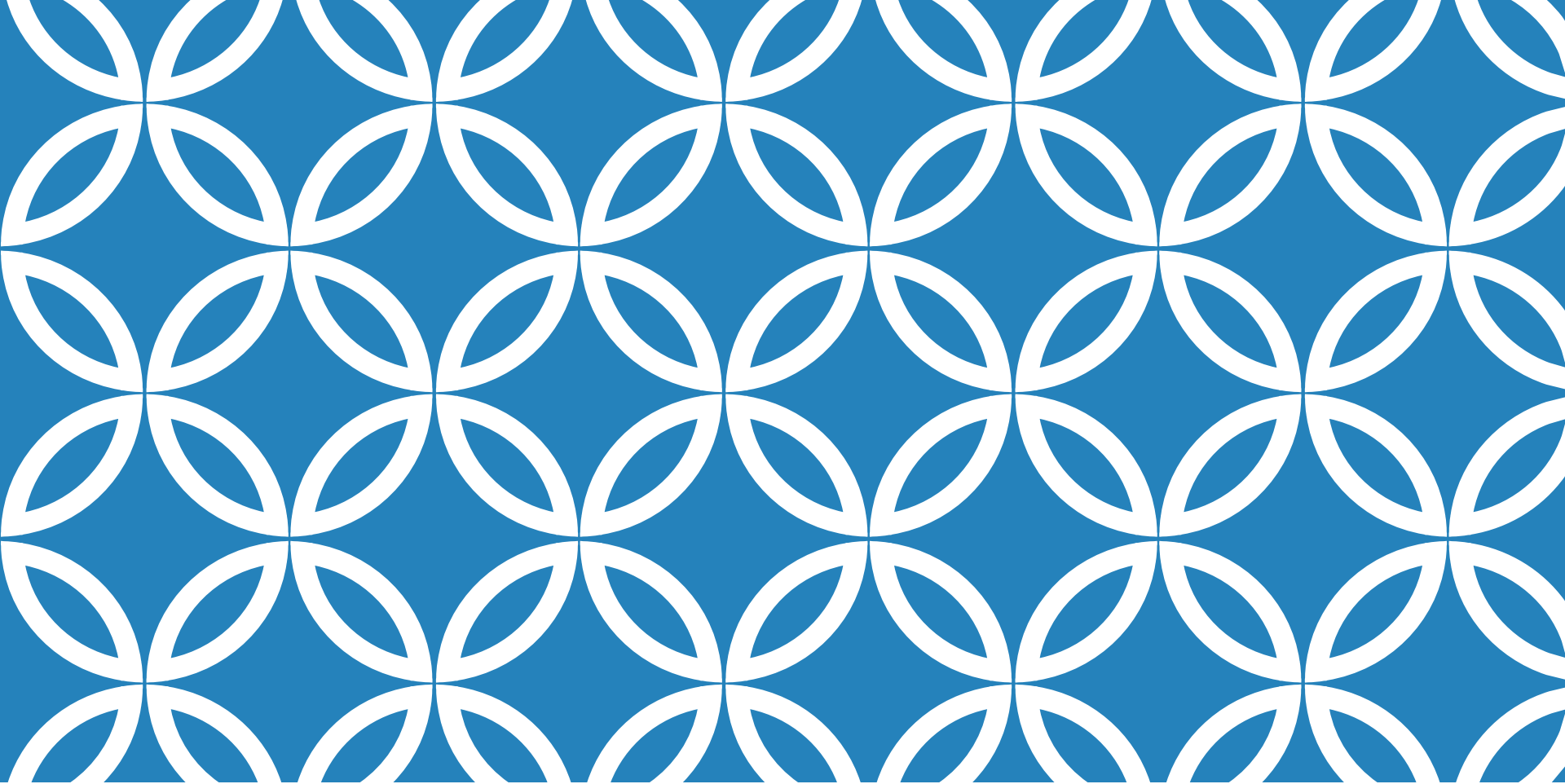
Levels/Scales of Measurement

Ratio

Interval

Ordinal

Nominal



* LET'S TAKE A CLOSER LOOK AT
EACH OF THE FOUR LEVELS

	Properties/ Characteristics	Examples	Descriptive Statistics	Graphs/chart
Nominal data				
Ordinal data				
Interval data				
Ratio data				

Visit [Levels of Measurement | Nominal, Ordinal, Interval and Ratio \(scribbr.com\)](https://www.scribbr.com/levels-of-measurement/) to help you complete the table

NOMINAL LEVEL DATA

Properties/Characteristics

Nominal data does not have a numerical value. Can be categorised into mutually exclusive groups. The response option/groups cannot be put into a hierarchy or order.

Examples

Hair colour, ethnicity, sex

What statistics can we find?

Mode

What charts/graphs can we use?

Bar chart, pie chart

ORDINAL LEVEL DATA

Properties/Characteristics

Ordinal data does not have a numeric value. The data can be categorised and the use of numerical codes can be assigned to textual responses. The response options can be put into a hierarchy or order, but the intervals between them cannot be commented upon – they cannot be presumed to be equal

Examples

Competency of swimming ability. Ratings of feelings/satisfaction towards something

What statistics can we find?

Mode response. Comment on frequencies. Calculate proportions (percentages/fractions).

What charts/graphs can we use?

Pie charts, bar chart

INTERVAL LEVEL DATA

Properties/Characteristics

Differences between the values are equal and have meaning. Values exist below zero. Zero is an arbitrary value.

Examples

Standardised test/exam scores. Temperature scale in C or F.

What statistics can we find?

Mean, Mode, Median. Standard deviation. Percentiles and Interquartile range.

What charts/graphs can we use?

Bar chart, Frequency diagram (grouped data), Histogram, Box Plot, linegraph

RATIO LEVEL DATA

Properties/Characteristics

Differences between the values are equal and have meaning.
Continuous data. True zero starting point – ratios can be calculated

Examples

Age, weight, height, ...

What statistics can we find?

Mean, Mode, Median. Standard deviation. Percentiles and Interquartile range.

What charts/graphs can we use?

Bar chart, Frequency diagram (grouped data), Histogram, Box Plot, linegraph

QUESTIONNAIRES

Questionnaires can be classified as both **quantitative** and **qualitative** methods depending on the nature of questions.

DICHOTOMOUS QUESTIONS

These type of questions gives two options to respondents –

Yes or No,

True or False

This is **Nominal data**

Properties/Characteristics

Discrete data. Qualitative. Order is not important

What statistics can we find?

Mode response. Comment on frequencies. Calculate proportions (percentages/fractions)

What charts/graphs can we use?

Frequencies – Bar chart

Proportions – Pie chart

SCALING QUESTIONS

Also referred to as ranking questions, they present an option for respondents to rank the available answers to the questions on the scale of given range of values (for example from 1 to 10).

This is **Ordinal data. Example: Likert Scale**

Properties/Characteristics

Ranked. Satisfaction scales. Qualitative. Ordered categories. Meaningful order, but the intervals may not be equal – the space between ranks is subjective

What statistics can we find?

Mode or Median response. Comment on frequencies. Calculate proportions (percentages/fractions).

What charts/graphs can we use?

Frequencies – Bar chart

Agreement

Strongly agree

Agree

Undecided

Disagree

Strongly disagree

Frequency

Very frequently

Frequently

Occasionally

Rarely

Never

Examples of Likert Scales

Why is Mean not an appropriate statistical analysis?

Importance

Very important

Important

Moderately important

Of little importance

Unimportant

Likelihood

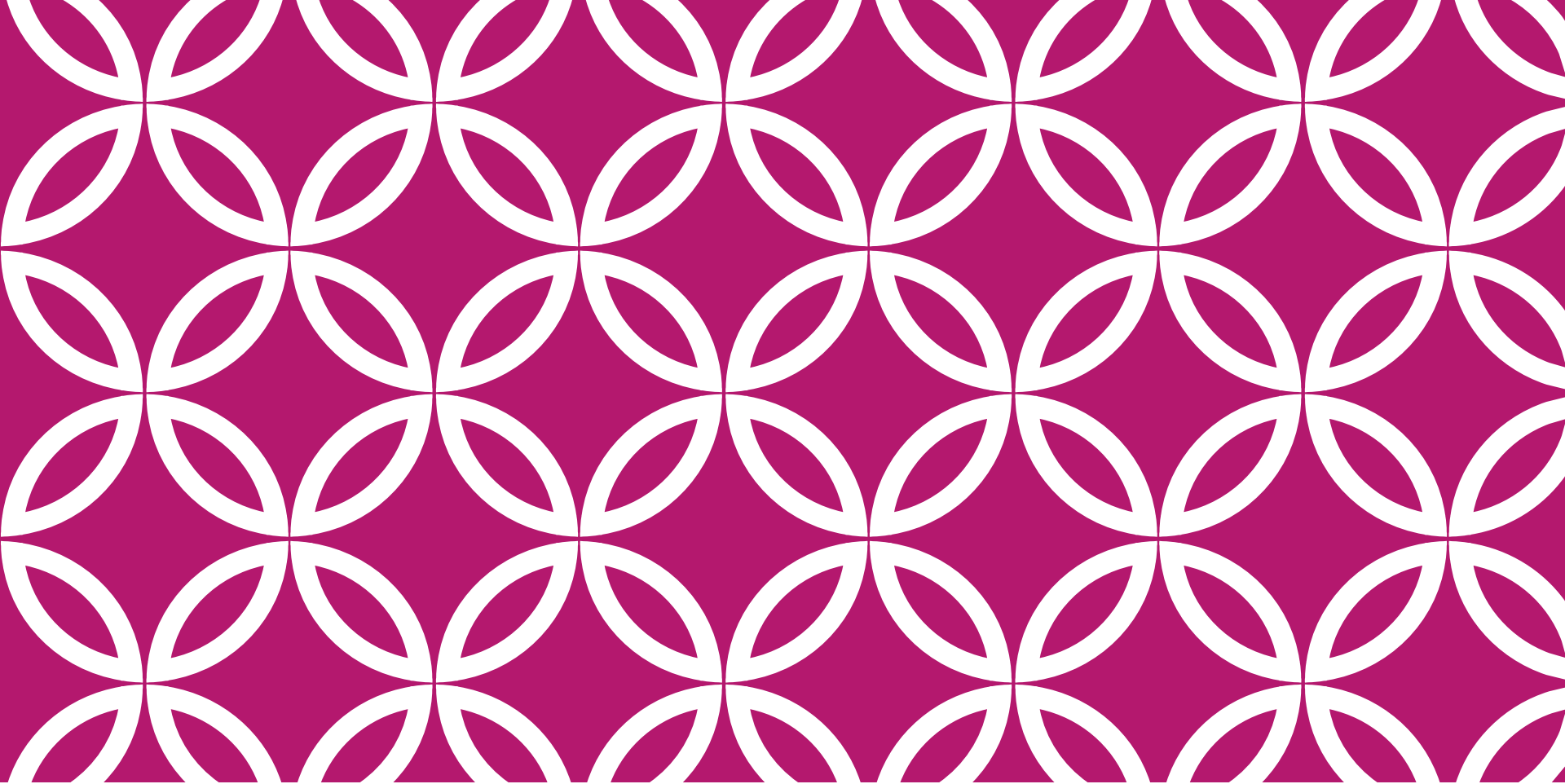
Almost always true

Usually true

Occasionally true

Usually not true

Almost never true



MEASURES OF CENTRAL TENDENCY

VISIT [CENTRAL TENDENCY | UNDERSTANDING THE MEAN, MEDIAN & MODE \(SCRIBBR.COM\)](https://www.scribbr.com/understanding-the-mean-median-mode/)

Types of calculations.

Measures of central tendency help you find the central or 'average' value. The most common measures of central tendency are the arithmetic mean, the median, and the mode.

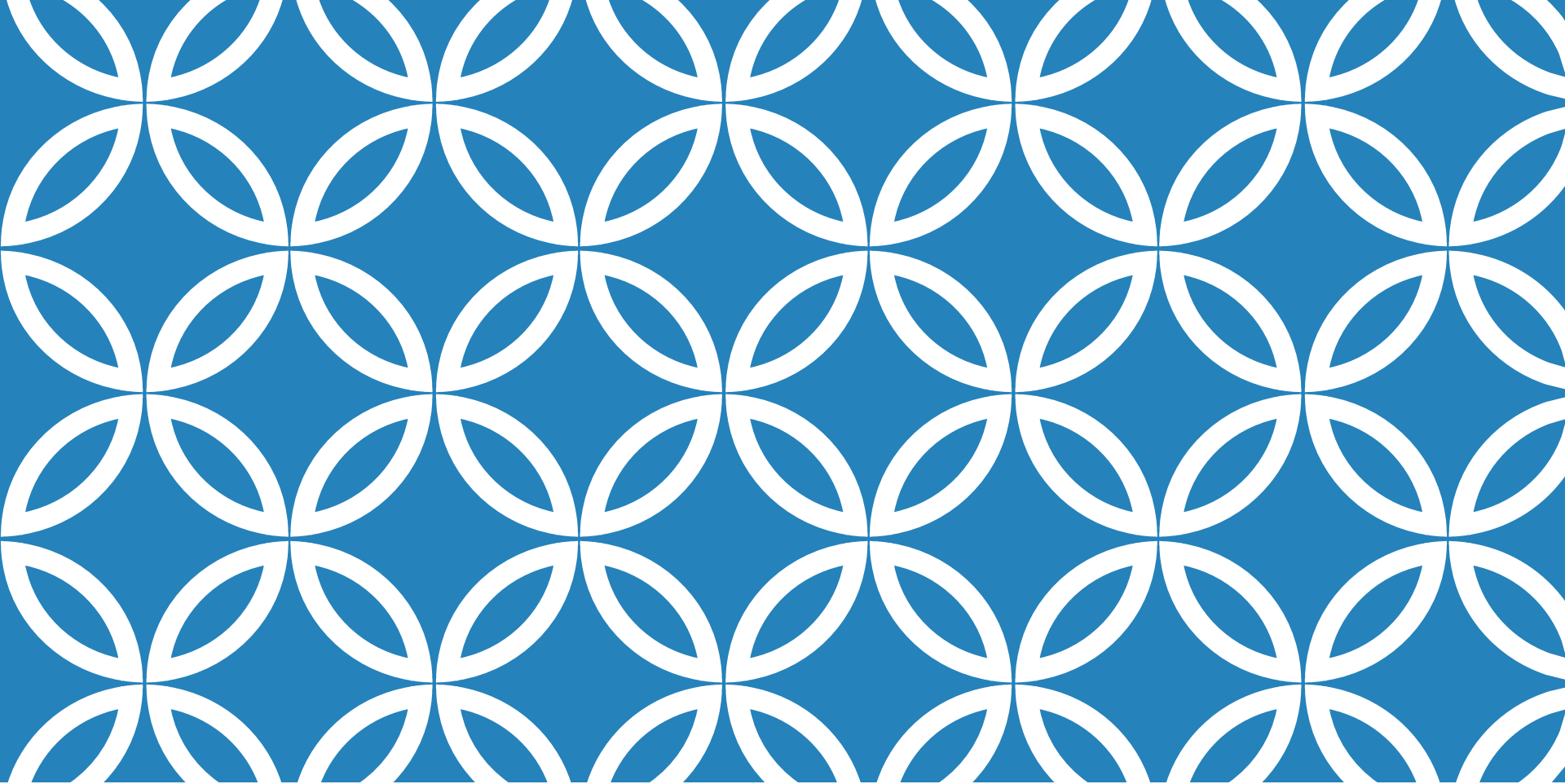
TWO GROUPS OF 5 CHILDREN ARE TIMED COMPLETING AN ACTIVITY



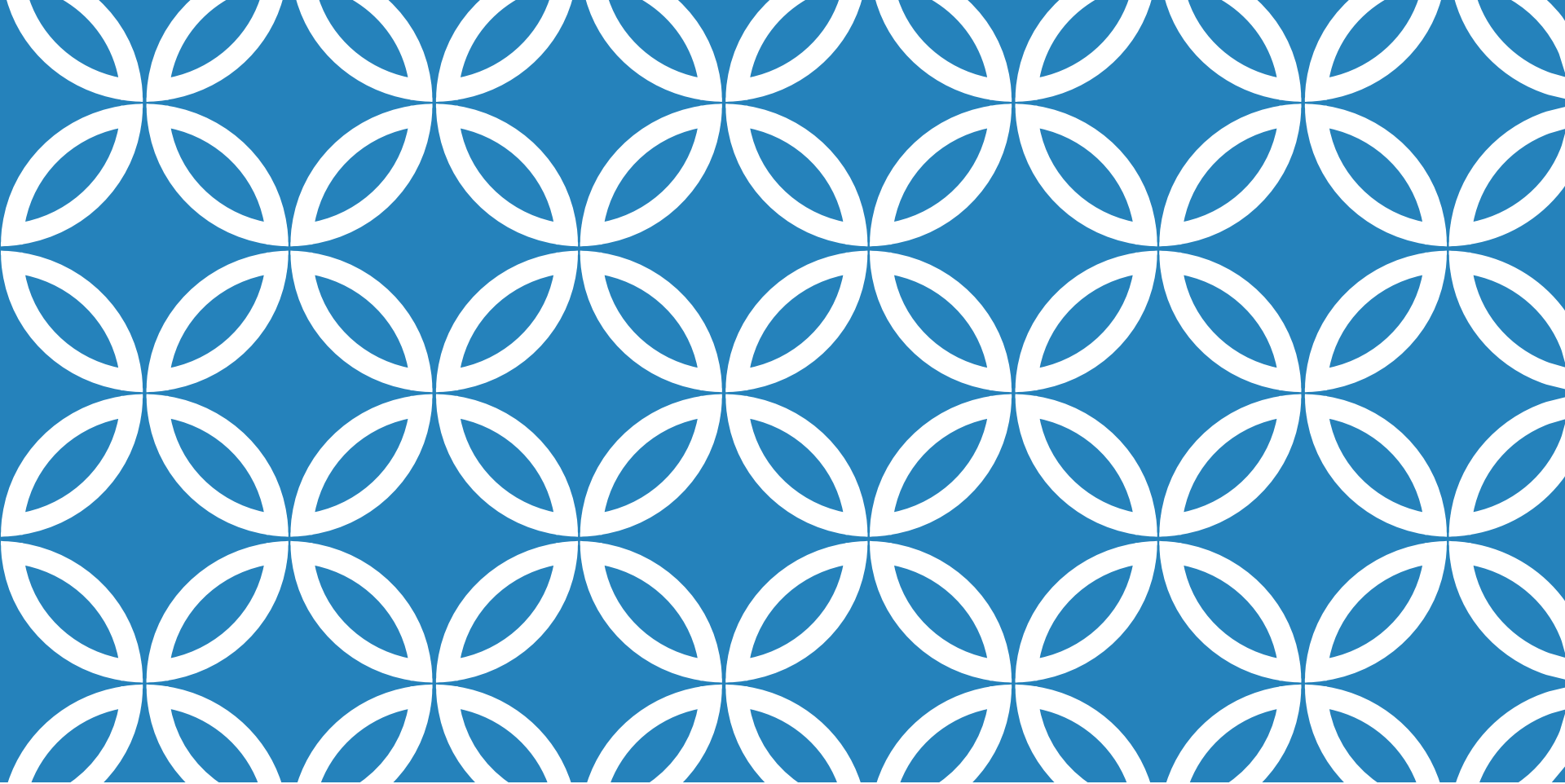
Group 1 Times (mins)	Group 2 Times (mins)
15	2
15	7
14	14
15	22
16	30

What does the data tell us?

We need to do some number crunching



MEAN AND STANDARD DEVIATION



USING EXCEL

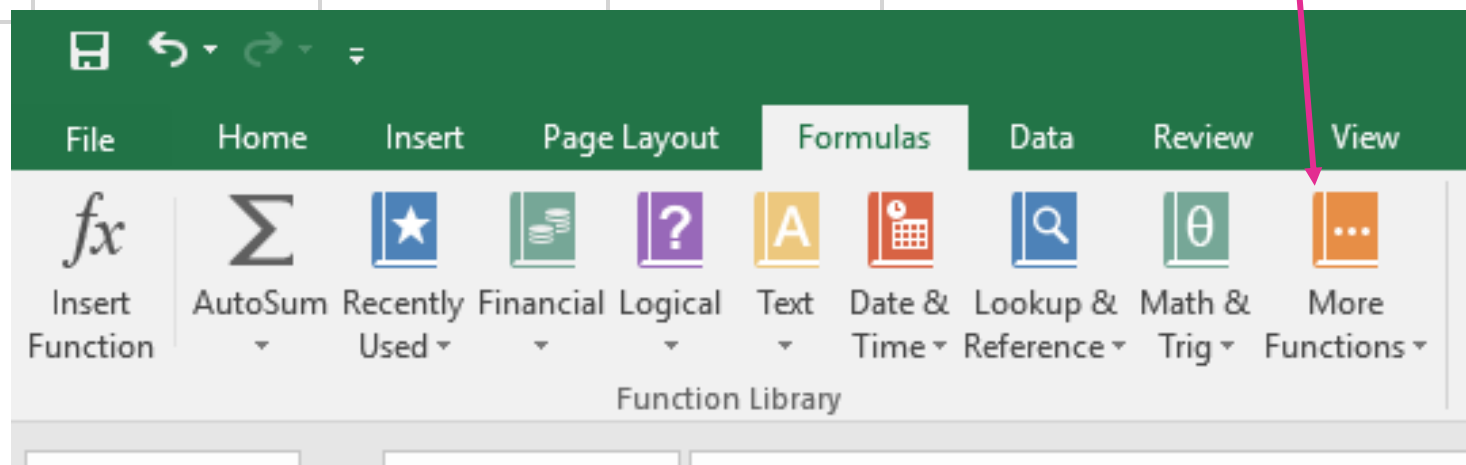
Descriptive Statistics.

mean and standard deviation examples

Group 1	Group 2
15	2
15	7
15	14
14	22
16	30
15	15

Average

Std Dev.



Excel ribbon showing the **Formulas** tab, **More Functions** dropdown menu, and the **AVERAGE** function selection.

The **More Functions** dropdown menu is open, showing the **Statistical** category selected. The **AVERAGE** function is highlighted in the list.

The **AVERAGE** function description is displayed:

AVERAGE(number1,number2,)

Returns the average (arithmetic mean) of its arguments, which can be numbers or names, arrays, or references that contain numbers.

[Tell me more](#)

The spreadsheet data is shown below:

	D	E	F
		Group 1	Group 2
		15	2
		15	7
		15	14
		14	22
		16	30
		15	15
Average			
Std. Deviation			

fx =AVERAGE(E4:E9)		
D	E	F
	Group 1	Group 2
	15	2
	15	7
	15	14
	14	22
	16	30
	15	15
Average	=(E4:E9)	
Std. Deviation		

?

×

Function Arguments

AVERAGE

Number1

E4:E9

= {15;15;15;14;16;15}

Number2

= number

= 15

Returns the average (arithmetic mean) of its arguments, which can be numbers or names, arrays, or references that contain numbers.

Number1: number1,number2,... are 1 to 255 numeric arguments for which you want the average.

Formula result = 15

[Help on this function](#)

OK

Cancel

fx			=AVERAGE(F4:F9)		
	D	E	F		
		Group 1	Group 2		
		15	2		
		15	7		
		15	14		
		14	22		
		16	30		
		15	15		
Average		15	= (F4:F9)		
Std. Deviation					

Function Arguments

AVERAGE

Number1

F4:F9



= {2;7;14;22;30;15}

Number2



= number

= 15

Returns the average (arithmetic mean) of its arguments, which can be numbers or names, that contain numbers.

Number1: number1,number2,... are 1 to 255 numeric arguments to be averaged.

Formula result = 15

[Help on this function](#)

		Group 1	Group 2
		15	2
		15	7
		15	14
		14	22
		16	30
		15	15
Average		15	15
Std. Deviation			

Both groups of children, on average, took 15 minutes to complete the set task

File Home Layout Formulas Data Review View ACROBAT Tell me what you want to do...

Function Library

More Functions

- Statistical
- Engineering
- Cube
- Information
- Compatibility
- Web

STDEV.P

STDEV.S

STDEVA

STDEVPA

STEYX

T.DIST

T.DIST.2T

T.DIST.RT

T.INV

T.INV.2T

T.TEST

TREND

TRIMMEAN

VAR.P

VAR.S

STDEV.S(number1,number2,)

Estimates standard deviation based on a sample (ignores logical values and text in the sample).

[Tell me more](#)

	D	E	F
		Group 1	Group 2
		15	2
		15	7
		15	14
		14	22
		16	30
		15	15
Average		15	15
Std. Deviation			

	Group 1	Group 2
	15	2
	15	7
	15	14
	14	22
	16	30
	15	15
Average	15	15
Std. Deviation	0.632456	10.07968
Std. Deviation	0.6	10.1

Standard deviation shows the variation in data – how spread out the data is around the mean value

MEASURE OF DISPERSION

In statistics, the **standard deviation** is a measure of spread or dispersion. It tells us how far, on average, the results are from the mean. A low standard deviation indicates that the values tend to be close to the mean of the set, while a high standard deviation indicates that the values are spread out over a wider range.

We can also consider the **variance**, which is the spread between numbers in a data set. Variance is a statistical measurement used to determine how far each number is from the mean and from every other number in the set.

The more spread the data, the larger the variance is in relation to the mean

	Group 1	Group 2
Mean	15	15
Standard deviation	0.632456	10.07968
Variance	0.4	101.6

VARIABILITY

Variability helps us to understand more about the spread/scatter/dispersion of results a dataset. It describes how far apart data points lie from each other and from the centre of a distribution.

When variability is calculated **in addition** to measures of central tendency you are provided with descriptive statistics to help you summarise your data.

Variability is commonly measured with the following:

Range: calculated by subtracting the lowest value from the highest value

Interquartile range: is the range of the 'middle half' of the dataset/distribution

Standard deviation: average distance from the mean

Variance: average of squared distances from the mean

QUARTILES

Quartiles are the values that divide a list of ordered numbers into quarters:

The 1st quartile – 25% of the way through the data set

The 2nd quartile – 50% of the way through the data set

The 3rd quartile – 75% of the way through the data set

Quartiles are the values that divide a list of ordered number into quarters:

Lower Quartile– 25% of the way through the data set

Median– 50% of the way through the data set

Upper quartile– 75% of the way through the data set

QUARTILES - EXAMPLE

The test scores from 13 pupils is given below (marked out of 20)

9, 10, 12, 14, 14, 15, 17, 17, 17, 18, 19, 20, 20

Quartiles are the values that divide a list of ordered number into quarters:

Lower Quartile— 25% of the way through the data set 13 points

Median— 50% of the way through the data set 17 points

Upper quartile— 75% of the way through the data set 18.5 points

We can also find the [Interquartile range](#). This is the Upper Quartile – Lower Quartile

In this case $IQR = 18.5 - 13 = 4.5$

[Visit Variability | Calculating Range, IQR, Variance, Standard Deviation \(scribbr.com\)](#)

BOX PLOT

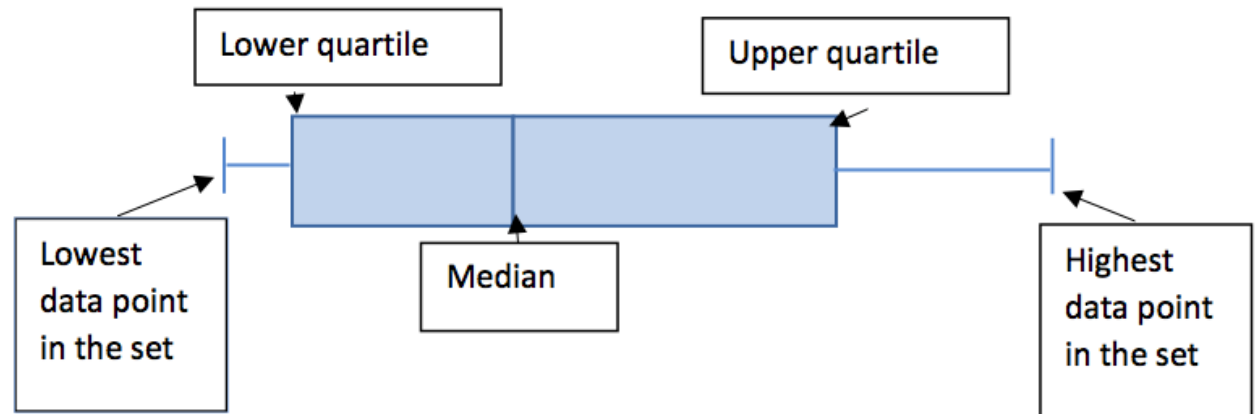
Quartiles are the values that divide a list of ordered number into quarters:

Lower Quartile— 25% of the way through the data set 13 points

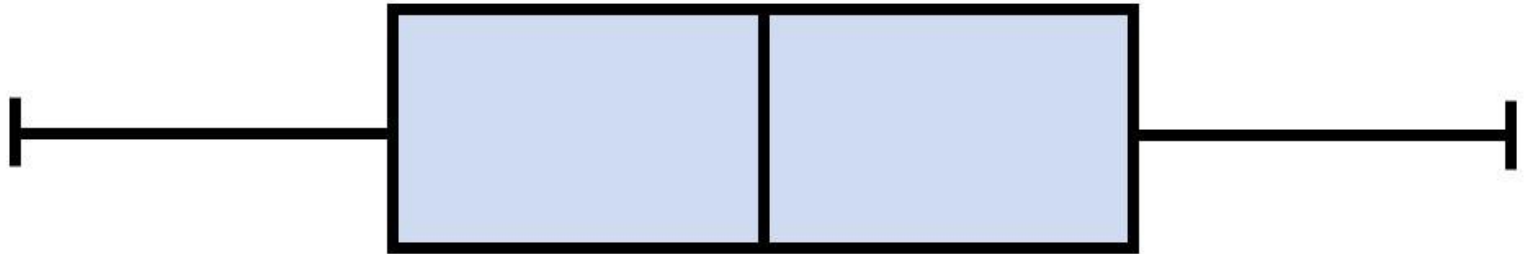
Median— 50% of the way through the data set 17 points

Upper quartile— 75% of the way through the data set 18.5 points

This information can be represented using a Box Plot (also known as a Box and Whisker diagram)



Normal Distribution



Positive Skew



Negative Skew



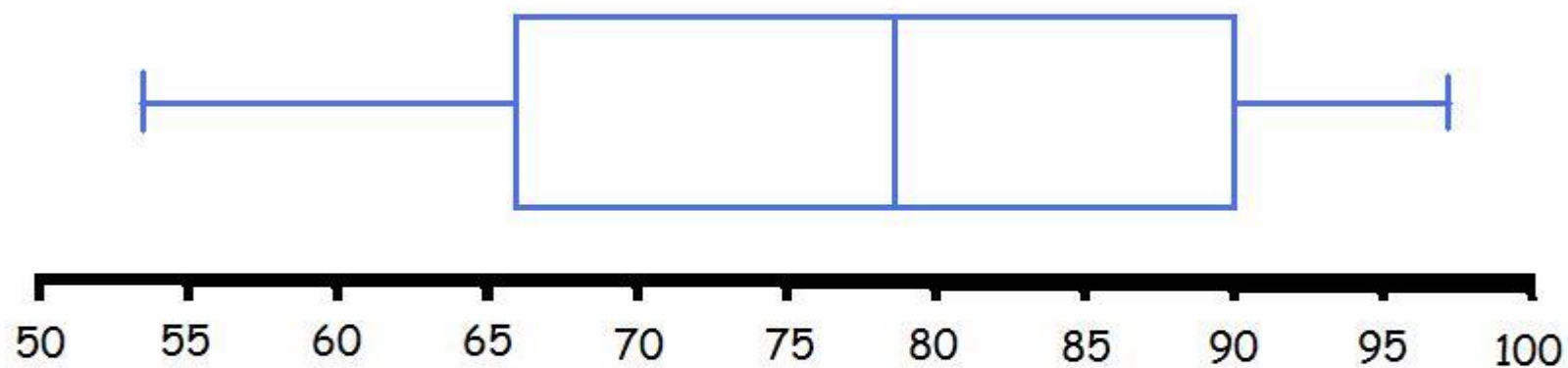
Skewness	What does the box plot look like	What does it mean
Normal Distribution	Median in the centre	The data is evenly spread
Positive Skew	Median to the left of the box.	The data above the median is more widely spread
Negative Skew	Median to the right of the box.	The data below the median is more widely spread

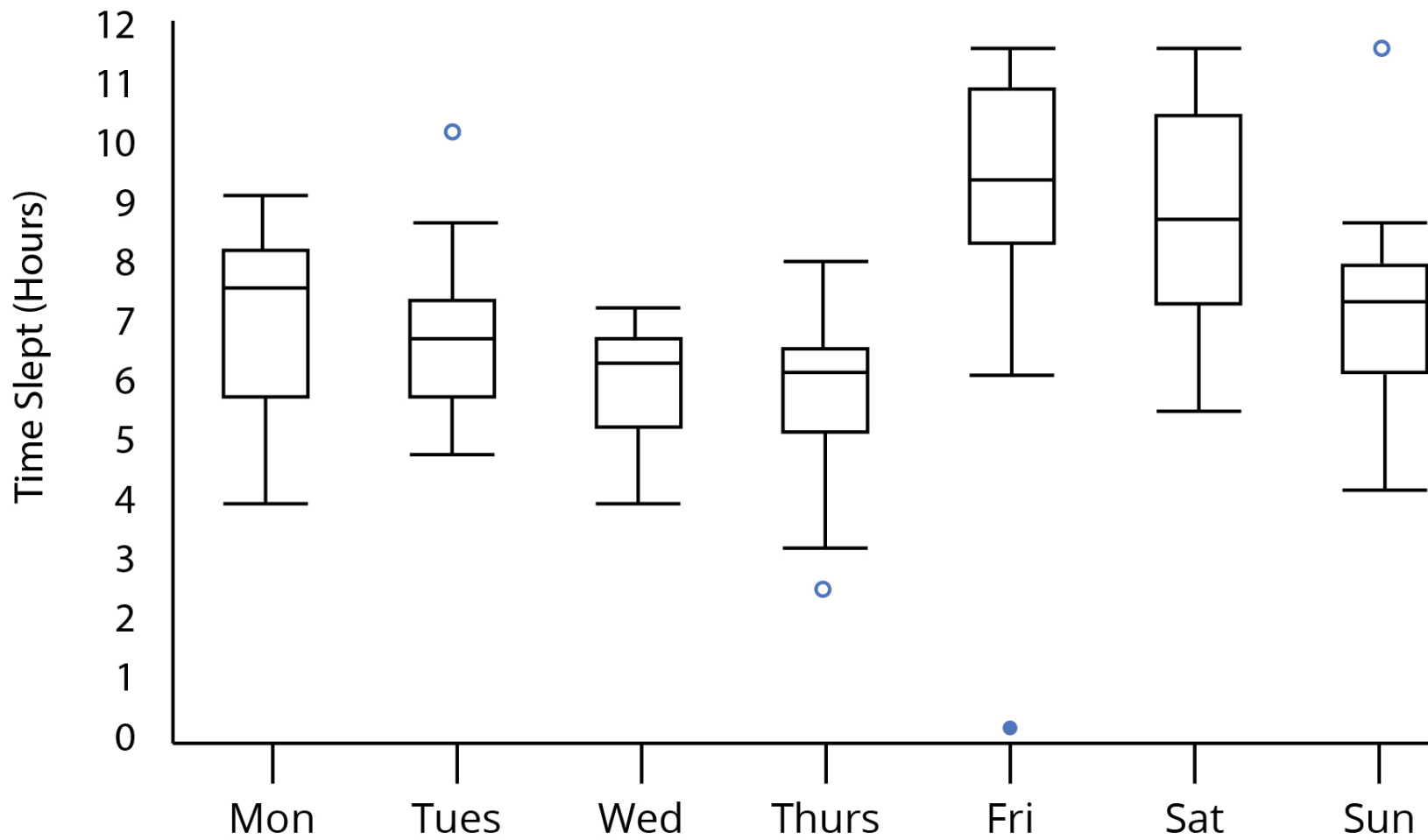
Remember each section of the box plot contains 25% of the data.

A smaller section indicates the data is more condensed (close together)

Variance (or variability) is described by the Interquartile Range (IQR). The larger the IQR, the more varied the data is. The IQR ignores outliers, therefore, if the data is skewed it is a more appropriate measure of variability than the standard deviation.

Scores on a math test





BOX PLOT AND OUTLIERS

Consider this data,

12, 23, 34, 37, 45, 49, 50, 60, 79, 120

Lower Quartile 34

Upper Quartile 60

IQR $60 - 34 = 22$

To decide **if this data has any outliers**, use the following calculations to have the **range** you will accept. IQR - 1.5 TIMES THE FIRST AND IQR + 1.5 TIMES THE THIRD QUARTILE

IQR - 1.5 x Lower Quartile,

$22 - 1.5 \times 34,$

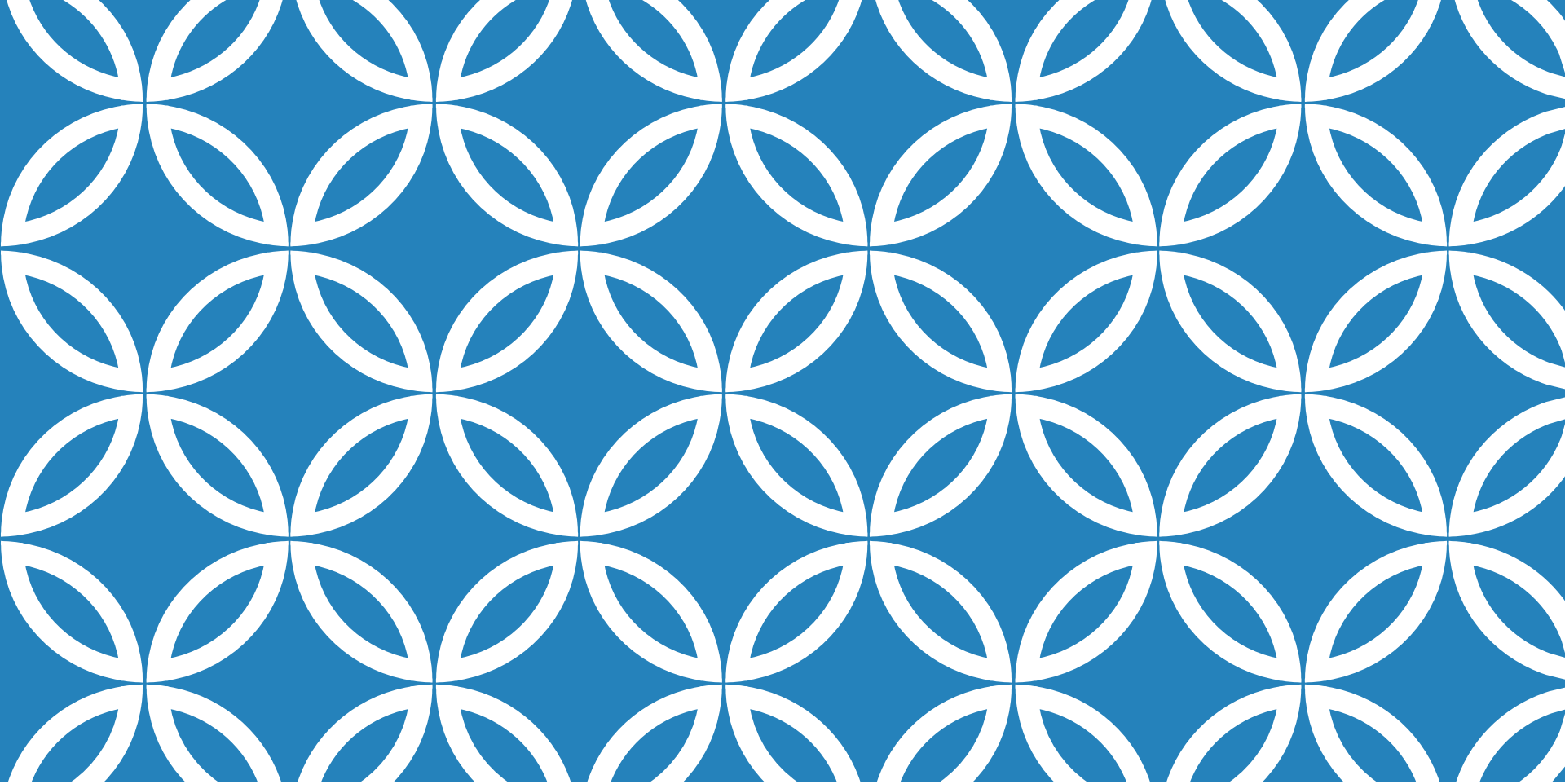
- 29,

IQR + 1.5 x Upper Quartile

$22 + 1.5 \times 60$

112

120 does not fall in this range, therefore 120 as an outlier



HYPOTHESES TESTING

VISIT [HYPOTHESIS TESTING | A STEP-BY-STEP
GUIDE WITH EASY EXAMPLES \(SCRIBBR.COM\)](https://www.scribbr.com/hypothesis-testing/)

STATISTICAL HYPOTHESIS TESTS

In order to test a hypothesis statistically we need two hypotheses –

a null hypothesis (called H_0)

And

an experimental or alternative hypothesis (called H_1)

STATISTICAL HYPOTHESIS TESTS

A null hypothesis (called H_0) – states there is NO STATISTICAL DIFFERENCE between sets of data or NO STATISTICAL RELATIONSHIP between data.

In other words, any observed difference is most likely by chance alone.

An experimental or alternative hypothesis (called H_1) – states there IS A STATISTICAL DIFFERENCE between sets of data or A STATISTICAL RELATIONSHIP between data.

In other words, it is unlikely that the difference occurred by chance alone.

When writing your hypothesis, you need to be specific depending on your context and what you are finding out

Testing hypotheses:

“I want to find out if there is a difference in mental arithmetic ability between boys and girls”

What would your H_0 and H_1 be?

In this case

H_0 is:

“There is no difference between boys’ and girls’ mental arithmetic ability”

H_1 is:

“There is a difference in mental arithmetic ability between boys and girls”

(2 tailed t- test)

STATISTICAL TESTS

There are a variety of statistical tests which are suited to different sample types.

Recall the 4 levels of measurement earlier, some statistical tests can only be used with specific types of data sets.

Good news – there is only some number crunching that you will have to do. The tests can be carried out in EXCEL, using a statistical package or downloadable specific statistical test calculator!

When you have carried out the statistical test, you will find a p value (probability value)

You will need to decide if this p value is significant...there are rules we follow



We need to compare
the value we get to
something

UNDERSTANDING THE P VALUE

The **P Value** stands for Probability Value and states **how likely** it is the results (from your research) **occurred by chance alone**.

If a p value is **large** it indicates that it is **likely** that these results have **occurred by chance alone**

If a p value is **small** it indicates that it is **unlikely** that these results have **occurred by chance alone**. Note: these results will then be known as **being statistically significant**.

How small is small, how large is large?

UNDERSTANDING THE P VALUE

In educational research the tradition is that

If the p value is greater than 0.05 researchers decide the results are not significant. *In other words any observed difference is most likely by chance alone*

$P > 0.05$ means probability value is greater than 5%

If the p value is less than 0.05 researchers decide the results are significant. *In other words any observed difference is unlikely to be by chance alone*

$P < 0.05$ means probability value is less than 5%

ALPHA

The significance level, α , is **the probability of rejecting the null hypothesis when it is actually true**. For example, a significance level of 0.05 indicates a 5% risk of the researcher concluding that a difference exists when there is no actual difference. *In other words, the chance, written as a percentage that you have made a mistake and chosen the wrong hypothesis to accept.*

The researcher needs to decide the level of alpha being used for the testing.

The two most common ones are $\alpha = 0.05$ and $\alpha = 0.01$

When interpreting if a p value is significant (or not) the researcher refers to the α value they have chosen in order to make a direct comparison

α

In our earlier case of two hypothesis.

H_0 is:

“There is no difference between boys’ and girls’ mental arithmetic ability”

H_1 is:

“There is a difference in mental arithmetic ability between boys and girls”

(2 tailed t- test)

What if we carried out a test and p was

If $p > 0.05$

This is traditionally reported as:

We cannot reject H_0

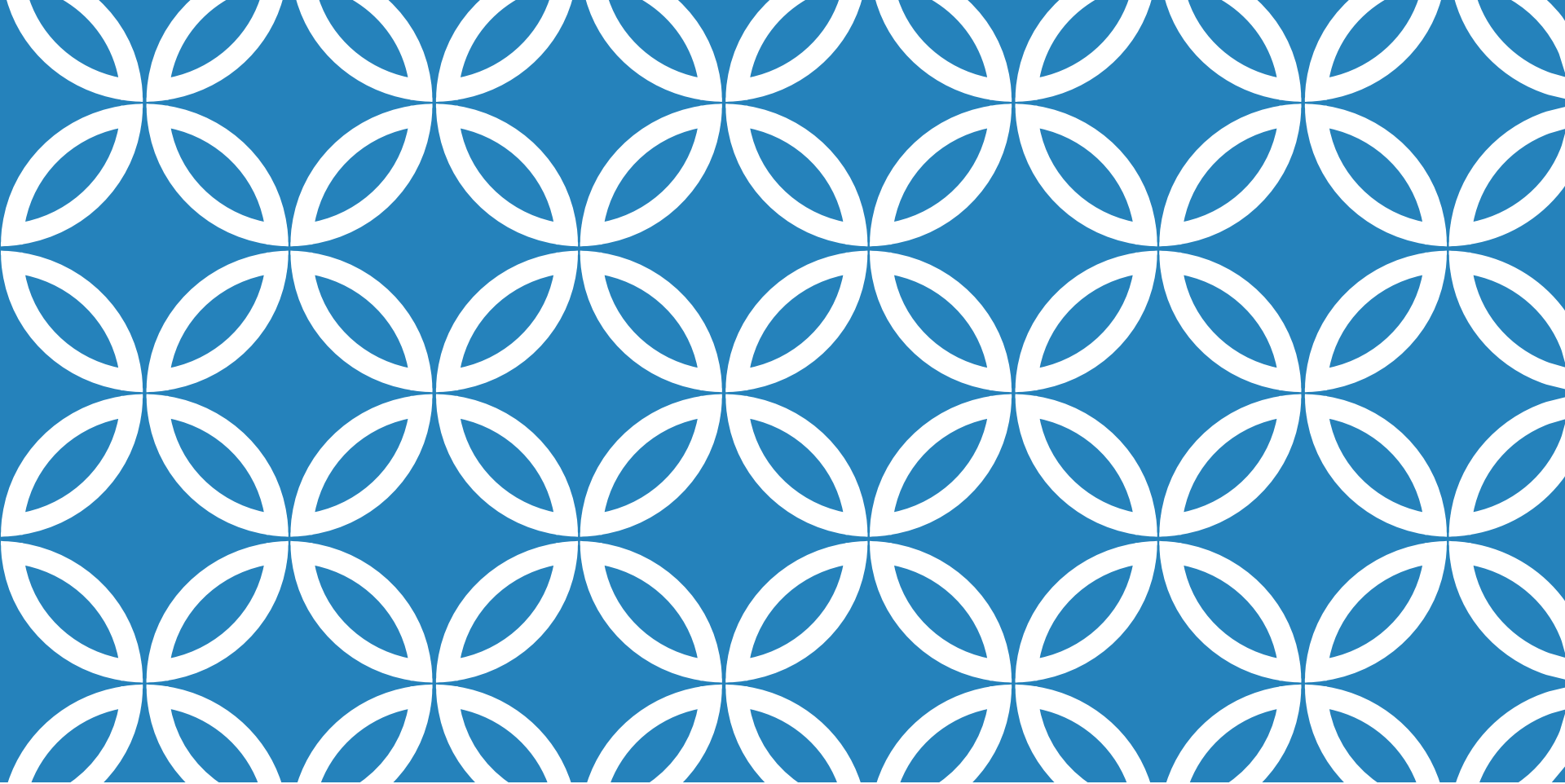
OR

If $p < 0.05$

This is traditionally reported as:

We can reject H_0 and accept H_1 .

Note: we would state the actual p value in our writing for the reader to see for themselves what the calculated statistical test provided.



USING A CHI-SQUARE TEST OF INDEPENDENCE

Use a Chi-Square
Calculator

<https://www.mathsisfun.com/data/chi-square-calculator.html>

CHI-SQUARE TEST

Suppose you want to know whether or not employment status is associated with a preference for a pet cat or dog

Carry out a survey

Collate the results from the survey in a table

	CAT	DOG
Employed	207	282
Not employed	231	242

Note: Categorical data

Does employment status affect pet preference?

If employment status **does** affect preferred pet we say they are **dependent**.
By doing some special calculations (explained later), we come up with a "p" value:

STATISTICAL HYPOTHESIS TESTS

A null hypothesis (called H_0) – states there is NO STATISTICAL RELATIONSHIP between the data.

In other words any observed relationship is most likely by chance alone

An experimental or alternative hypothesis (called H_1) – states there IS A STATISTICAL RELATIONSHIP.

In other words, that it is unlikely that the relationship occurred by chance alone

When writing your hypothesis, you need to be specific depending on your context and what you are finding out

Define the hypotheses

'I will perform the Chi Square test of independence using the following hypotheses

H_0 : Employment status and pet preference are independent.

H_1 : Employment status and pet preference are *not* independent

Carry out the test to calculate p

Understanding "p" Value

"p" is the probability the variables are **independent**.

If the **p** value is more than **0.05** there is no significant relationship

If the **p** value is less than **0.05** there is a significant relationship

Let's find out

<https://www.mathsisfun.com/data/chi-square-calculator.html>

<https://www.mathsisfun.com/data/chi-square-test.html>

Chi-Square Calculator

The results are in! And the groups have different numbers

- But is that just random chance?
- Or have you found something significant?

The Chi-Square Test gives us a "p" value to help us decide.

Chi-Square Calculator

or

207	282
231	242

Chi-Square: 4.10353

Degrees of Freedom: 1

p: 0.04279

p is calculated as 0.042782

In this case **$p < 0.05$** ,

So, this result is thought of as being "significant"

This means, we think the variables are **not** independent.

In other words, because **$0.043 < 0.05$** (0.043 is less than 0.05) we think that **employment status is linked to Pet Preference**

Refer back to the Hypothesis

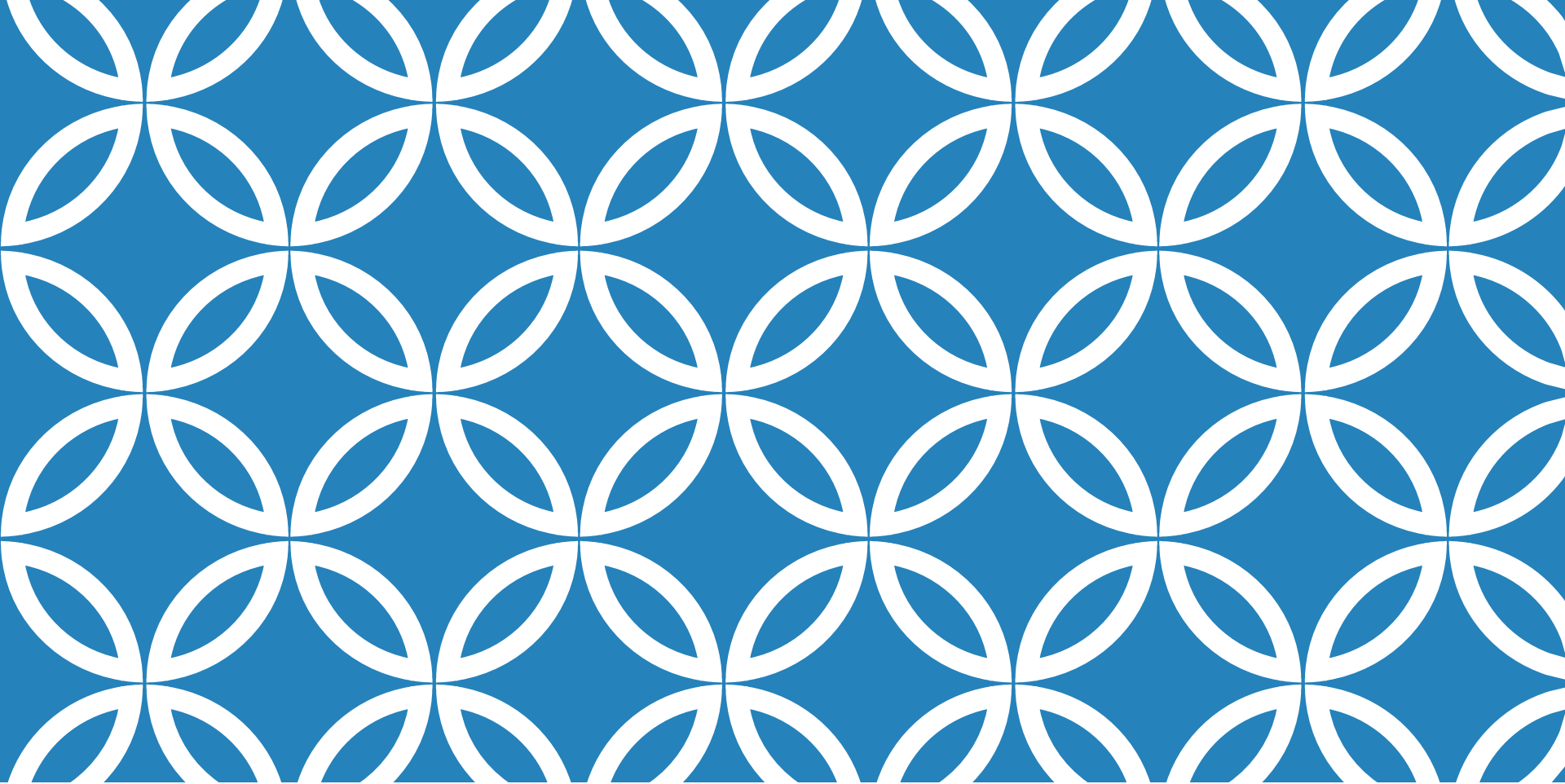
H_0 : Employment status and pet preference are independent.

H_1 : Employment status and pet preference are *not* independent

$$p < 0.05$$

We report this as:

We reject H_0 and accept H_1 .



USING A STUDENT T- TEST

Use Excel

T-TEST

Quantitative data can be investigated using a statistical test known as a t-test

A t-test helps the researcher to decide if there is a significant difference between two sets of data

Once again, it involves calculating a **p value** and comparing this value to the significant level, **$\alpha = 0.05$ or $\alpha = 0.01$**

It requires the researcher to state a null hypothesis and an alternative hypothesis

The result of the test eliminates researcher bias when making any conclusions and supports the strength of any argument the researcher maybe trying to make.

<https://www.youtube.com/watch?v=qvPWQ-e03tQ>

Pre- Scores	Post-Scores
34	63
56	89
21	57
15	68
19	69
28	75
27	89
32	67
67	98
45	87
81	99
45	76
23	76
53	98
42	73
49	75
37	69
48	78
24	76
36	56

T-TEST EXAMPLE

The table shows the scores obtained by 20 pupils who completed a test at the start of term and at the end of term. Here are the **paired** results.

Consider the data, what do we want to test?

Pre- Scores	Post-Scores
34	63
56	89
21	57
15	68
19	69
28	75
27	89
32	67
67	98
45	87
81	99
45	76
23	76
53	98
42	73
49	75
37	69
48	78
24	76
36	56

T-TEST EXAMPLE

When carrying out a t-test, the researcher needs to ask themselves two questions.

1. Are you conducting a one-tailed or two tailed test?
2. Is the data paired or unpaired?

Are you conducting a one-tailed or two tailed test?

This depends on how open-ended your question or hypothesis is.

Is the data paired or unpaired?

Does the order matter? Are the columns linked?

1 OR 2 TAILED T-TEST?

Are you conducting a one-tailed or two tailed test?

This depends on how open-ended your question or hypothesis is.

Example of Testing hypotheses:

“I want to find out if there is a difference in mental arithmetic ability between boys and girls” (Open question, therefore a **2 tailed t-test**)

OR

‘ I want to find out if boys do significantly better at mental arithmetic than girls’
(Closed question, therefore a **1 tailed t-test**)

Pre- Scores	Post-Scores
34	63
56	89
21	57
15	68
19	69
28	75
27	89
32	67
67	98
45	87
81	99
45	76
23	76
53	98
42	73
49	75
37	69
48	78
24	76
36	56

T-TEST EXAMPLE

Question

‘is there a significant difference between pupils’ scores at the start of term and at the end of term?’

When carrying out a t-test, the researcher needs to ask themselves two questions.

1. Are you conducting a one-tailed or two tailed test?
2. Is the data paired or unpaired?

1. I am conducting **two tailed test**
2. The data is **paired**

Function Arguments

T.TEST

Array1			= array
Array2			= array
Tails			= number
Type			= number

=

Returns the probability associated with a Student's t-Test.

Array1 is the first data set.

Formula result =

[Help on this function](#)

OK Cancel

$p = 1.70093E-11$

$p = 0.00000000000170093$

Visit

<https://youtu.be/qvPWQ-e03tQ?t=701>

Question

'is there a significant difference between pupils' scores at the start of term and at the end of term?

A null hypothesis (called H_0) – states there is NO STATISTICAL RELATIONSHIP between the data.

In other words any observed relationship is most likely by chance alone

H_0 = there is **no significant difference** between pupils' scores at the start of term and at the end of term?

An experimental or alternative hypothesis (called H_1) – states there IS A STATISTICAL RELATIONSHIP.

In other words, it is unlikely that the relationship occurred by chance alone

H_1 = there is **a significant difference** between pupils' scores at the start of term and at the end of term?

$P = 0.000000000007$

Therefore $p < 0.05$, infact $p < 0.01$

So I reject H_0 and accept H_1 .



Choosing which statistical test to use - statistics help.

Test for
Proportion



Difference of
two proportions



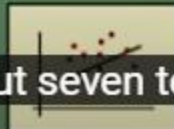
Difference of two means
(independent samples)



Chi-sq test for
independence



Regression analysis



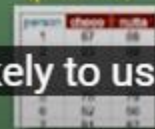
Test for
Watch later



Share

Info

Difference of two means
(paired)



MORE VIDEOS

This video will talk about seven tests you are likely to use.



0:15 / 9:32



YouTube

