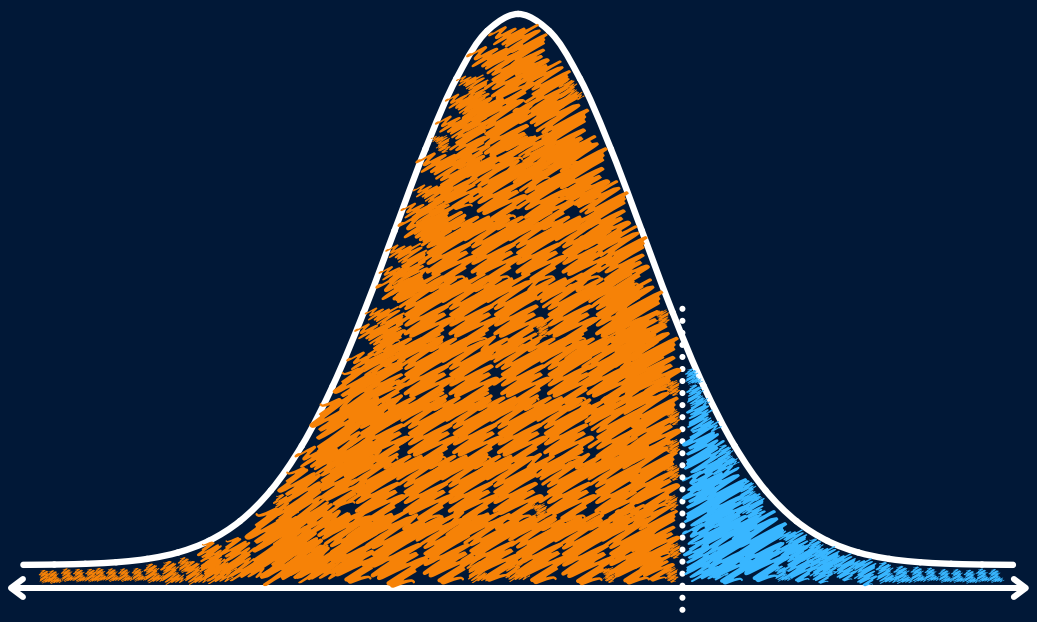
DATA SCIENCE INFINITY

COMMON HYPOTHESIS TESTS



AN OVERVIEW

What?

In Statistics, a **Hypothesis Test** is used to assess & understand the plausibility, or likelihood of some assumed viewpoint (*a hypothesis*) - based upon data. In other words, we are using Statistics to test or investigate ideas.

Let's say we're the coach of an NBA Basketball team - we might find ourselves with the below dilemma...

The New Sensation >>

Games Played: 2
Shooting Rate: 60%

The Current Star >>

Games Played: 102
Shooting Rate: 49%

Coach

My new player seems to be a better shooter...but she's only played 2 games. I need more confidence before making a big change!

Let's test it!

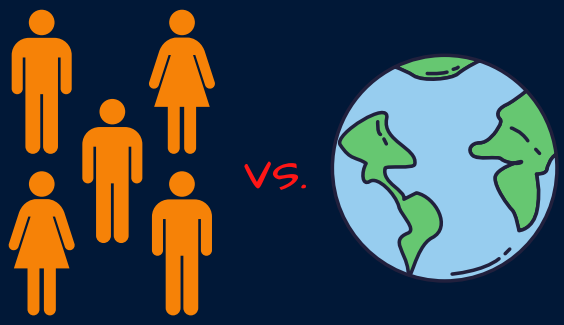
DATA SCIENCE INFINITY

Common Tests...

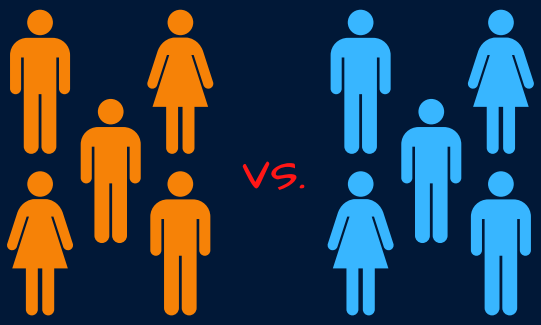
There are many, many different types of hypothesis test - each of which are appropriate for different types of data, and/or dealing with different scenarios & comparisons.

Here we will cover several commonly used tests...

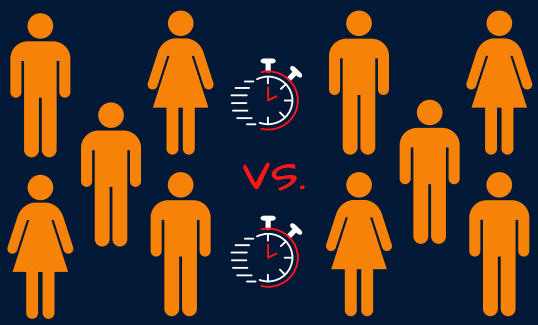
One Sample T-Test



Independent Samples T-Test



Paired Samples T-Test



Chi-Square Test Of Independence

DATA SCIENCE INFINITY

The Setup...

It is important to lay down our hypotheses/assumptions and a required level of confidence before running the test itself.

Generally, we specify 3 things; the **Null Hypothesis**, the **Alternate Hypothesis**, and our **Acceptance Criteria**...

The **Null Hypothesis** is where we state our initial viewpoint. In statistics, and specifically hypothesis testing, our initial viewpoint is always that the **result is purely by chance** or that there is **no relationship or association between two outcomes or groups**

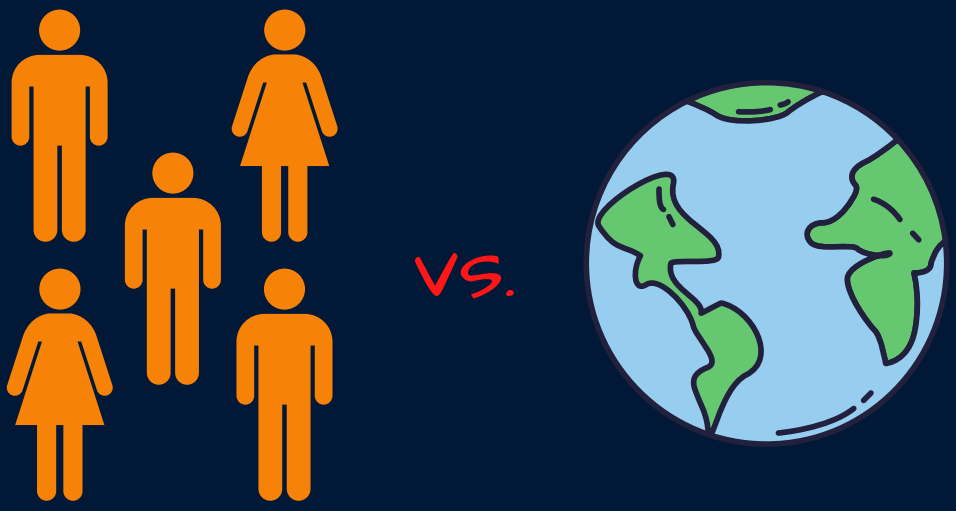
The **Alternate Hypothesis** is is essentially the opposite viewpoint to the Null Hypothesis - that **the result is not by chance**, or that **there is a relationship between two outcomes or groups**

The **Acceptance Criteria** is the threshold we set that will determine if there is **enough evidence** to support the Null Hypothesis. This is often set to a p-value of 0.05 but it **does not have to be**. If we want more certainty, we can set this to a lower value

DATA SCIENCE INFINITY

One Sample T-Test

A **One Sample T-Test** looks to assess differences between a sample, and the entire population from which that sample resides.



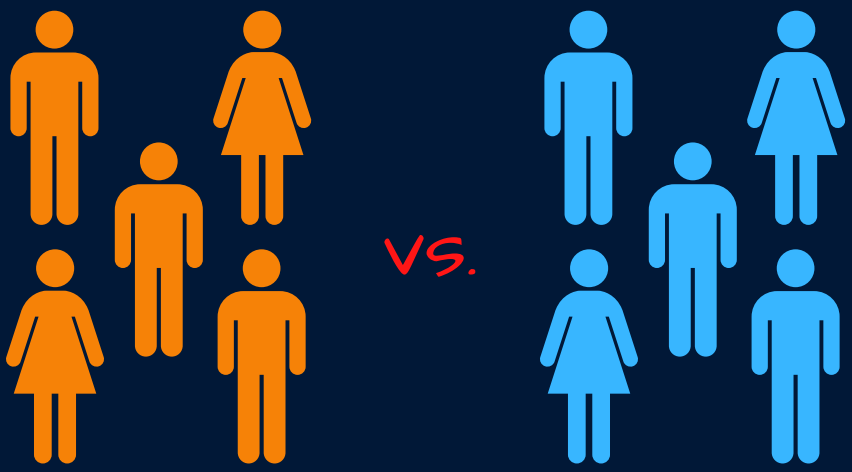
If we were the coach of an NBA Basketball team, this might help us with the following question...

Is the average height (cm) of my team higher than that of the entire NBA?

DATA SCIENCE INFINITY

Independent Samples T-Test

A **Independent Samples T-Test** looks to assess differences between a sample, and another sample.



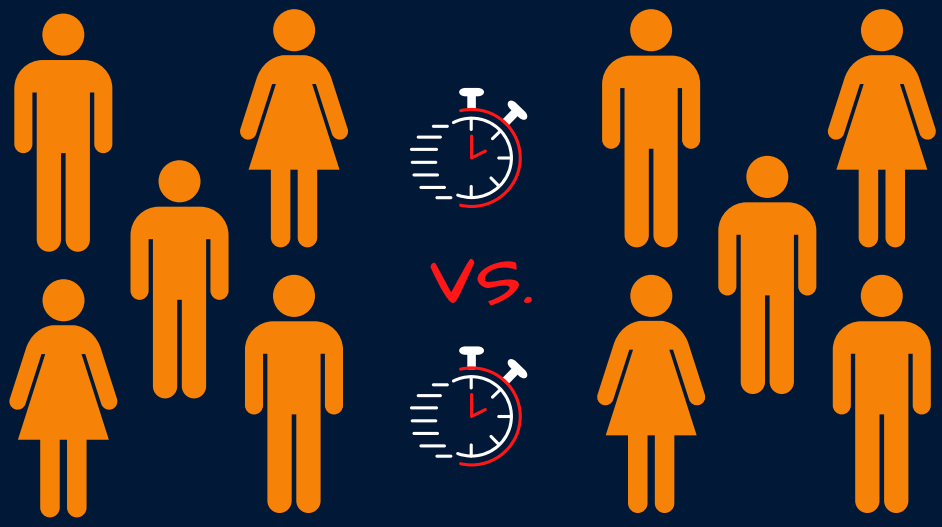
If we were the coach of an NBA Basketball team, this might help us with the following question...

Is the average height (cm) of my team higher than that of our rival team?

DATA SCIENCE INFINITY

Paired Samples T-Test

A **Paired Samples T-Test** looks to assess differences between a sample, and that same sample, at another point in time.



If we were the coach of an NBA Basketball team, this might help us with the following question...

Has the average jumping height (cm) for my team increased after our 4 week fitness programme?

DATA SCIENCE INFINITY

Chi-Square Test Of Independence

The Chi-Square Test For Independence is used to determine if there is a signficant relationship between two **categorical** variables. It examines the **actual and expected frequencies** to determine if there is a **dependence** between the two variables.

If we were the coach of an NBA Basketball team, this might help us with the following question...

Is the 3-point shooting % of my newly signed player, higher than that of my current star player?

DATA SCIENCE INFINITY

Let's talk about p-values

The outcome of a hypothesis test (in other words whether or not we support our initial assumption) is influenced by our **Acceptance Criteria**.

This Acceptance Criteria is often based upon a **p-value...**

...So let's talk about what a p-value **is**, and what a p-value **is not!**

*A **p-value** essentially helps us assess whether the results of some finding or test we have conducted are either, **likely to be ordinary, or likely to be strange.***

*In other words - these results we've got - do we think we'd get similar results if we ran **many more tests**, or do we think our results might be something of an rarity?*

*With an Acceptance Criteria of 0.05, we are essentially saying that we want a likelihood of 5% (or lower) that our result happened **by chance**.*

***A p-value is not a probability of an event occuring**, it is a probability or likelihood of seeing a different result if we were to sample many times*

***A p-value does not tell us how different two samples are**. Two samples with the same difference in means, but larger/smaller samples sizes will get different p-values. A p-value is instead telling us how likely it is that they are different (or in other words how confident we can be that they are different)*