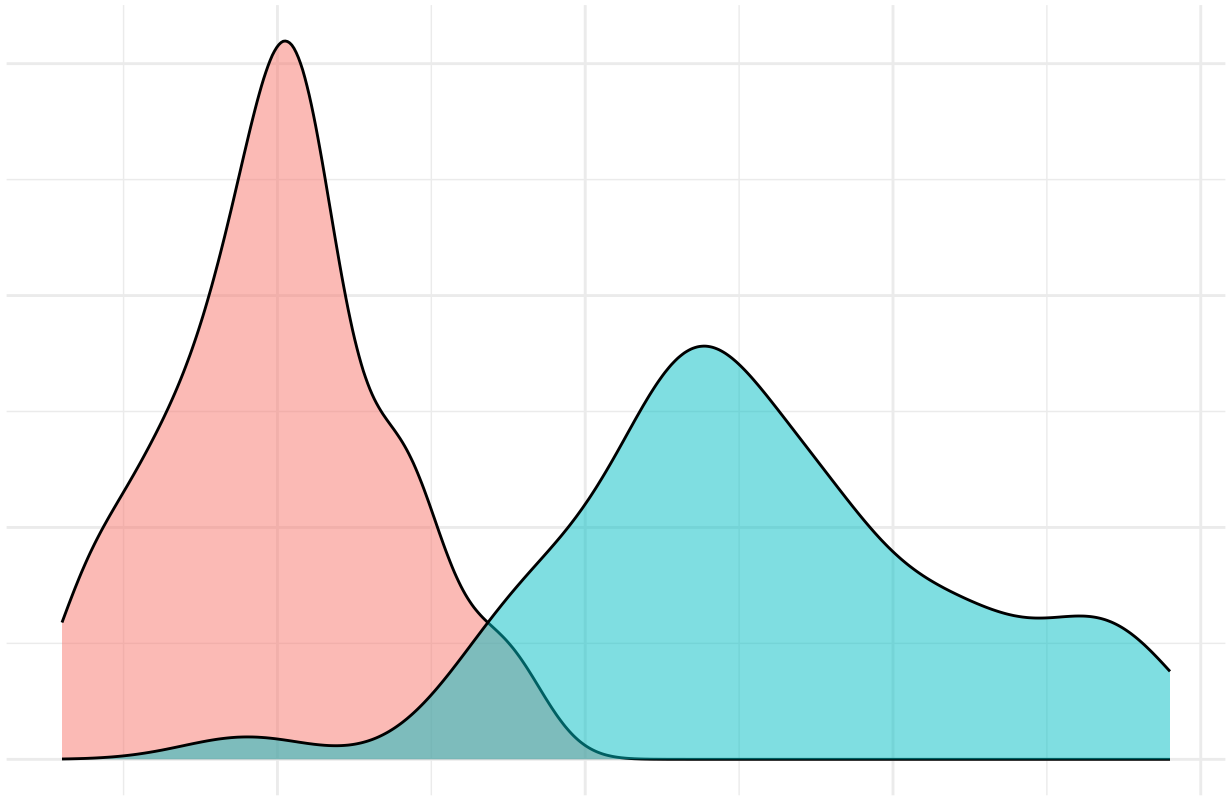
T-test

What is a T-test?

A t-test is one of the most commonly used statistical tests where it can be used to compare the means of **two** groups of data. It is very useful in answering the question - "Is the difference between these two groups real, or is it just down to chance?".



Are these populations different?

When to use a T-test?

When you are in a situation where you have one **continuous** variable and then a **categorical** variable with only two groups.

Examples could include; comparing enamel hardness between two restorative materials, comparing probe depth before and after treatment or comparing whitening effect of two brands.

Assumptions of a T-test

The assumptions of a T-test that must be met are as follows:

1) Continuous outcome
2) Normally distributed data (Close to a bell shaped curve)
3) Homogeneity of variance (Similar variation between group A and B)
4) Independence of observations

Different types of T-tests

Type of test: Independent T-test

Example: Plaque index between smokers and non-smokers

Type of test: Paired T-test

Example: Probing depth before or after treatment (When assumption 4 is not met)

Type of test: One-sample T-test

Example: Comparing fluoride concentration of a toothpaste to an industry standard

When assumption two is not met (normality of data) then a non-parametric equivalent can be used. Mann-Whitney U test (independent) and Wilcoxen signed-rank (paired).

How to implement in RStudio

```r
# Outcome = continous variable. Group = categorical variable
# df = the data frame
# changed paired to T for a paired T-test
t.test(outcome ~ group, data = df, paired = F)
```

Interpreting results

After you run your T-test you should be considering a few different things.

The p-value

The p-value tells you how likely it is to observe your results if there is actually no real difference between the groups. So a low p-value means that it is very unlikely to see your data when there is no real difference. A threshold of 0.05 is commonly used in research and this means there is a 5% chance the results you observe are due to chance and not another reason.

Effect size

Effect size tells you how big or meaningful the difference between your groups actually is. While the p-value tells you whether a difference is likely real, the effect size tells you how large that difference is in practical terms. Generally speaking; 0.2 = small effect, 0.5 = medium effect, 0.8+ = large effect.