ADVANCED PHARMACEUTICAL BIOSTATISTICS

Dr. Mohammed S. Al-Lami
PhD Pharmaceutics

18 Sep. 2023

Hypothesis Testing

- Scientific experiments typically involve postulating one or more hypotheses – statements that some fact concerning the population under study is **TRUE**
  - Men (on average) are taller than women
  - Weight training improves running speed
  - Aerobic exercise reduces blood pressure

- Statistical **hypothesis tests** provide an objective means of accepting or rejecting these hypotheses based upon the available experimental data.

Scientific Method

1. Pose a question (hypothesis) to be tested
2. Define experimental groups to be studied
    - E.g. males / females
    - Weight training / no training
    - Exercise / no exercise
3. Define variable(s) to be measured
    - Height
    - Weight
    - Blood pressure
4. Measure variable(s) of interest in a random sample of each experimental group from the population
5. Use hypothesis test to determine whether differences in descriptive statistics (mean, median, etc.) of sample groups justify acceptance or rejection of hypothesis.

Hypothesis Tests

Test		Comparison	Type	Applicable to
T Test (unpaired)	t	Difference between means of 2 independent sample groups	Parametric	Normally distr. ratio, interval data
T Test (paired)	t	Mean of differences between 2 repeated measures on same sample	Parametric	Normally distr. ratio, interval data
T Test (single)	t	Difference between mean of sample and theoretical value	Parametric	Normally distr. ratio, interval data
Mann-Whitney U Test	U	Difference in rank order of 2 independent groups (non-parametric equivalent of unpaired T test)	Non-parametric	Ordinal, ratio, interval data
Wilcoxon Test	W	Signed rank test (non-parametric equivalent of paired T test)	Non-parametric	Ordinal, ratio, interval data
Analysis of Variance	F	Ratio of between- and within-group variances of 3 or more independent sample groups	Parametric	Normally distr. ratio, interval data

T Test (Unpaired)

- Determines whether the difference between the means of two **independent** (i.e. composed of different subjects) sample groups is significant.

- **T Score**
  - A measure of the difference between group means relative to the standard error (i.e. accuracy) of the means.
  - The T score will be large when the difference between group means is large and/or the standard error is small.

$$T = \frac{Mean_1}{SE_1} - \frac{Mean_2}{SE_2}$$

$$T = \frac{Mean_1 - Mean_2}{SE_{pooled}}$$

(Samples assumed to have same standard deviation)

T Test: Hypotheses

- **Two mutually exclusive hypotheses are defined:**
  - **Null hypothesis (H_0)**
    There is no difference in the population mean values between our experimental groups
  - **Alternate hypothesis (H_1)**
    There is a difference between experimental groups

- **The significance probability (p)**
  - The T test calculates how probable it is that a T score at least as large as the one observed will occur, if the null hypothesis **(H_0)** is TRUE.

- If the significance probability is small enough ($p \leq 0.05$), we can reject the null hypothesis and conclude that there is a difference between the two groups.

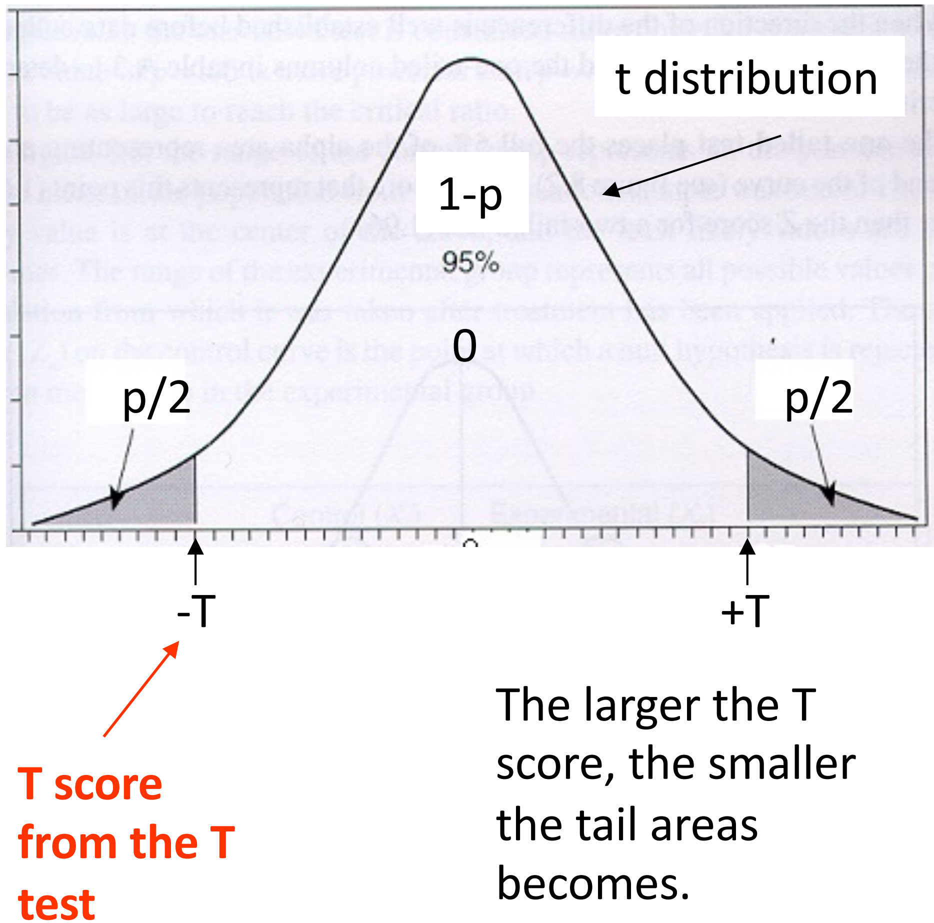
T Test: The significance probability (p)

- The T distribution is the range of T values that can occur when there is no difference between the two groups being tested (i.e. null hypothesis is TRUE).

  It can be calculated mathematically for a given sizes of samples for the two groups (Minitab does this for you).

- The significance probability for the test T value is obtained from the area under the tails of the **t-distribution** lying outside the region from –*T* to +T.

2-tail significance probability

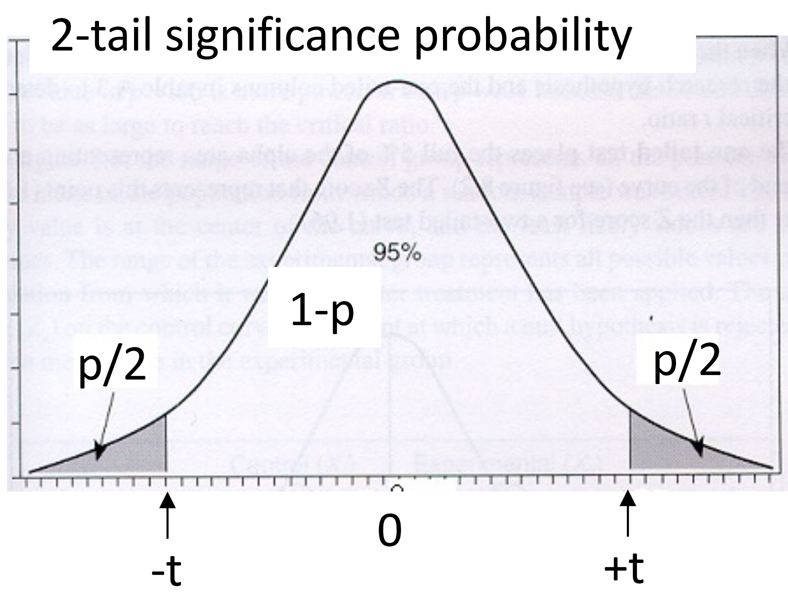


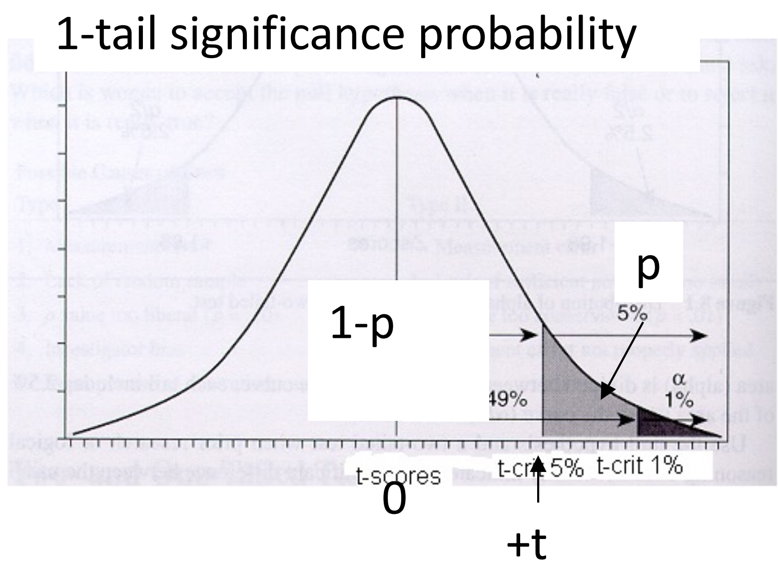
The larger the T score, the smaller the tail areas becomes.

Why include both tails of the distribution?

- Since we have made no distinction between positive or negative differences, we need to accept the possibility of either occurring when calculating the probability of obtaining the observed *t* score
- Therefore, we add both the negative (≤ -t) and positive tails ≥ +t) of the t distribution, to get the significance probability, p – **this is known as a 2-tail test.**

- **1-tail test**
  - If we could be <u>absolutely sure</u> that only a positive or negative difference between means were possible we could ignore one tail and thereby reduce the value p (making it possible to reach significance with a smaller t score)
  - However, it rarely possible to be absolutely sure, so it is safer to always use the 2-tail test.



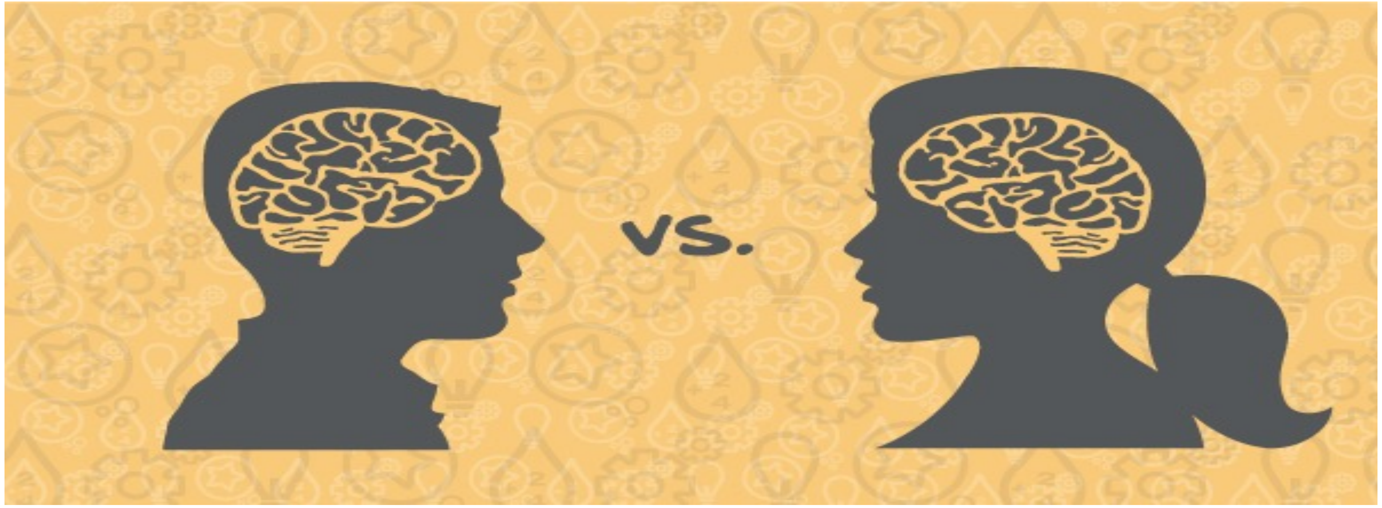


T-tests

- The *t* test is one type of inferential, parametric statistic
- Determine whether there is a significant difference between the means of two groups / conditions
- There are three main types

Student's T-test

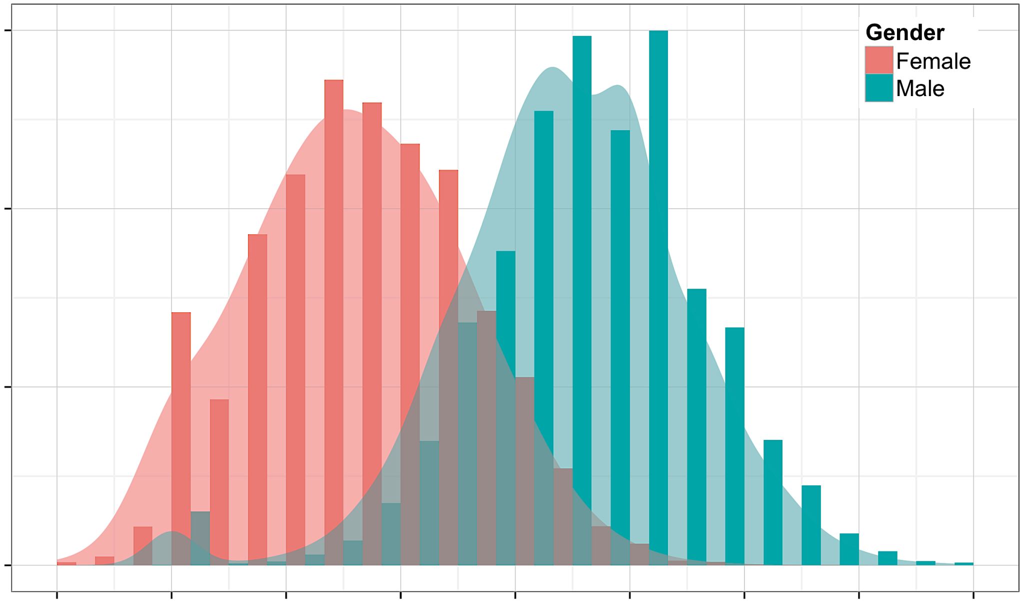
How can we tell whether one type of barley produces more than another with a believable level of scientific certainty?



- Determines whether there is a statistically significant difference between the means in two unrelated groups.

- **It is also known as independent samples t-test, two sample t-tests, between samples t-test and unpaired samples t-test.**

Student's T-test

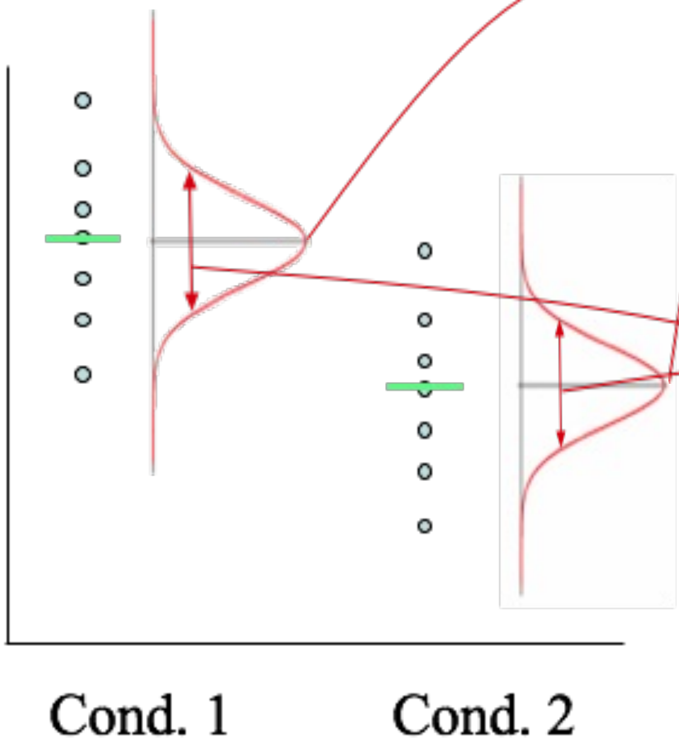


- **Independent groups**
- **Independent measurements**
- **One independent, categorical variable that has two levels/groups**
- **One continuous dependent variable**

FORMULA

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}}$$

difference between the mean's divided by the pooled standard error of the mean.



$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}}$$

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

-2.2762709

t-value $= \dfrac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\dfrac{{S_1}^2}{n_1} + \dfrac{{S_2}^2}{n_2}}}$$

t = 2.276

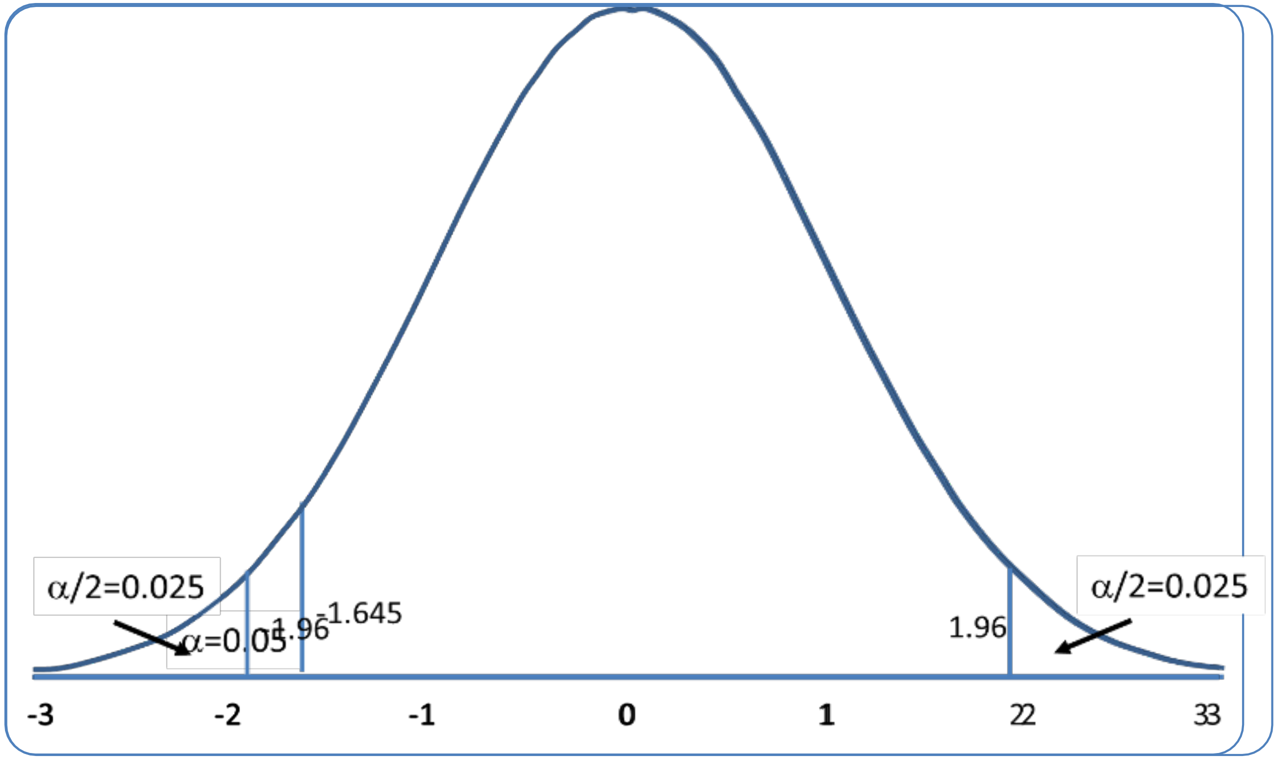
	Field 1	Field 2
1	16	15.2
2	15.8	16.6
3	15.6	15.2
4	14.9	15.8
5	15	16.2
6	15.2	15.6
7	15.6	15.6
8	15.7	15.8
9	15.5	16.2
10	15.2	15.6
11	15.5	15.8
12	15.2	15.5
13	15.5	15.5
14	15.1	15.5
15	15.3	14.9
16	15	15.9
Mean	15.38125	15.68125
SD	0.3166886	0.42145581
SE	0.07917215	0.10536395

T test

- Significance level: 0.05
- Degrees of freedom: (n1 + n2) - 2
- Degrees of freedom: (16 + 16) - 2 = 30
- Critical Value: 2.048
- T-Value: 2.27
- We reject H_0

P-value: 0.0306749

P= ≤ 0.05

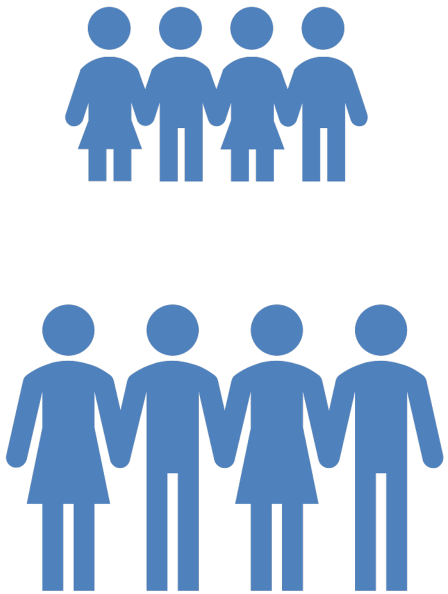


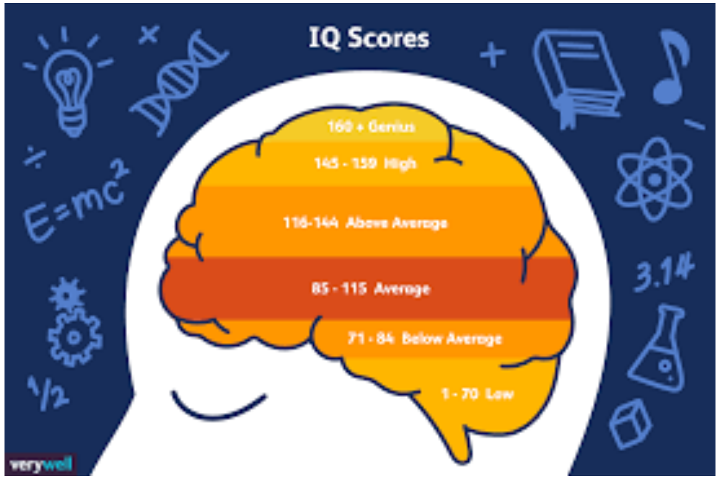
t-Test: Two-Sample Assuming Unequal Variances			
	Variable 1	Variable 2	
Mean	15.38125	15.68125	
Variance	0.10029167	0.177625	
Observations	16	16	
Hypothesized Mean Difference	0		
df	28		
t Stat	-2.2762709		
P(T<=t) one-tail	0.01533745		
t Critical one-tail	1.70113093		
P(T<=t) two-tail	0.0306749		
t Critical two-tail	2.04840714		

Paired T-test

- Estimate whether the means of two related measurements are significantly different from one another
- Used when two continuous variables are related
  - Same participant at different times
  - Different sites on the same person
  - Cases and their matched controls.
- Also known as within-subjects, repeated-measures and dependent-samples.

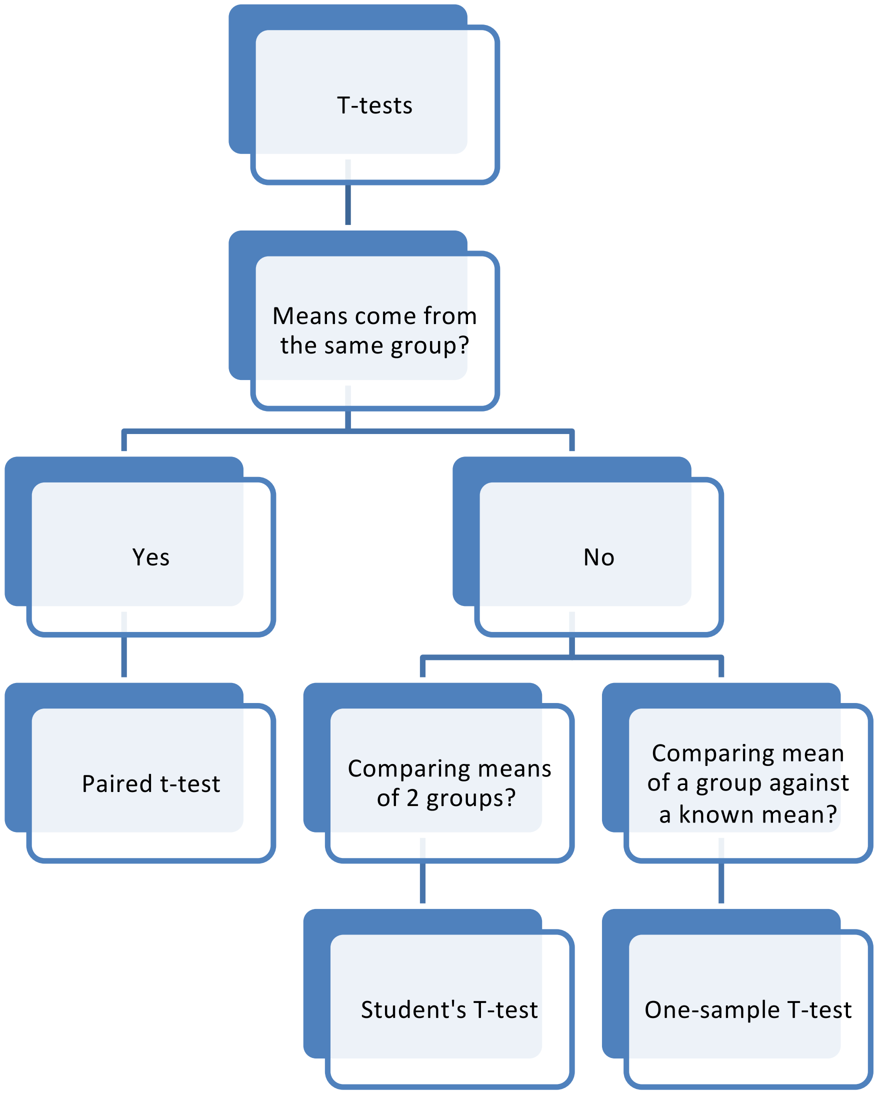
Paired T-test





One sample T-test

- **There is only one group which is to be compared to a set value or a known population mean.**

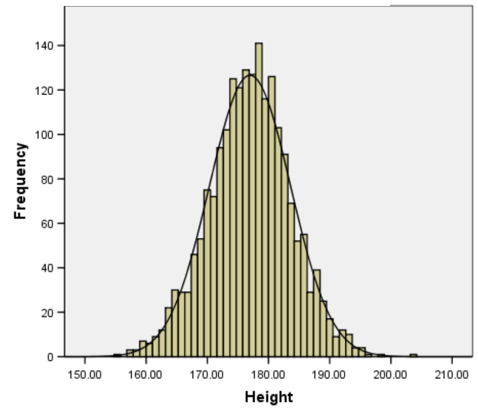


T-tests
Means come from the same group?
Yes
No
Paired t-test
Comparing means of 2 groups?
Comparing mean of a group against a known mean?
Student's T-test
One-sample T-test

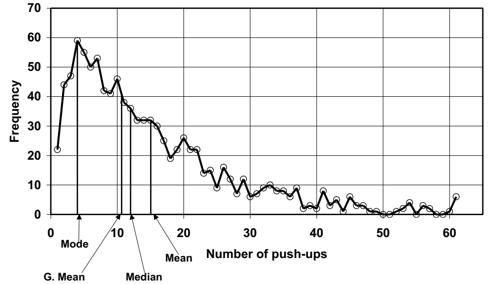
Assumptions when using an unpaired T test

- Data must be parametric
  - i.e. with interval or ratio scales
  - Not ordinal.
- Sample groups must be independent (i.e. composed of different subjects)
  - When experimental groups are repeatedly measured from same subjects a paired T test should be used.
- Samples must be drawn from a normally distributed population
  - This can be difficult to prove and normality is often just assumed. However, if clear evidence exists that population distribution is skewed non-parametric tests (e.g. Mann-Whitney) should be used instead.
- Each sample group has a similar standard deviation
  - If sample standard deviations differ by more a factor of two, the "equal variances not assumed" version of the T test should be used.

Normal Dist.



Non-normal Dist.



Example: Male and female average height?

- **Population: first year undergraduates**
  - 2000 Male
  - 2100 Female
- **Hypothesis**
  - Do women and men have different average heights?
- **Variable**
  - Height
- **Sample groups**
  - Male
  - Female
- **Hypothesis Test**
  - Unpaired T Test of 2 independent sample groups

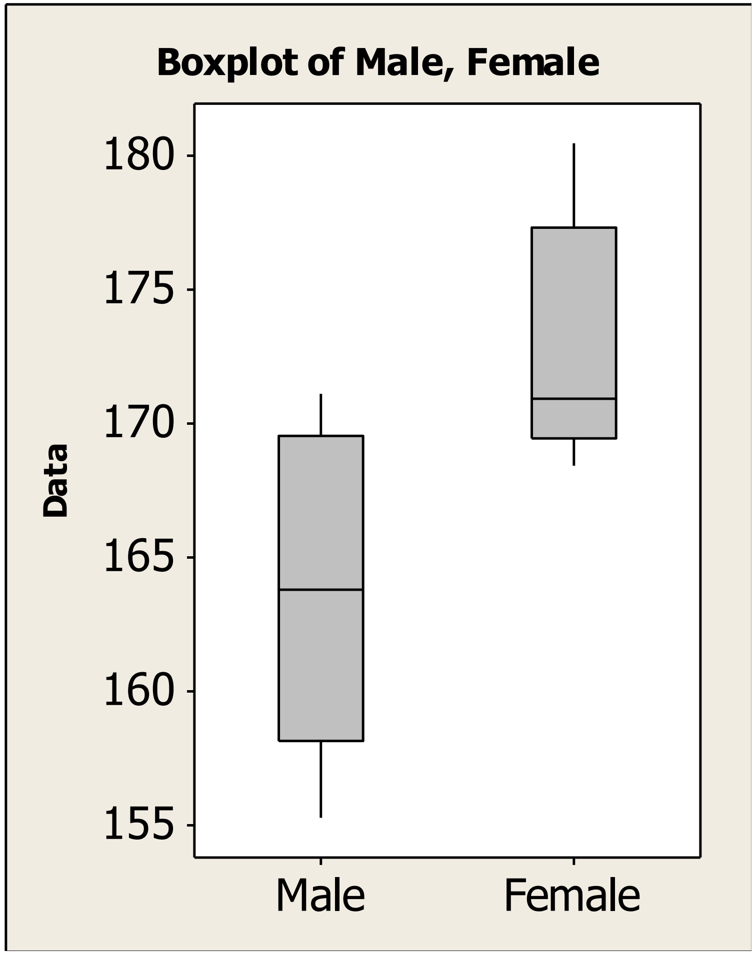
Example: Experimental Data

- Random samples of heights of 5 men and 5 women from population.
- **Results**
  - Male = 172.8 cm
  - Female = 163.8 cm
- **Question**
  - Does the difference between the **sample** means justify the conclusion that the **population** means for male & female heights differ?

	Female Heights (cm)	Male Heights (cm)
	171.1	180.4
	161.0	168.4
	155.3	170.9
	167.9	174.1
	163.8	170.4
Mean	**163.8**	**172.8**
SD	6.12	4.69
SE	2.74	2.10

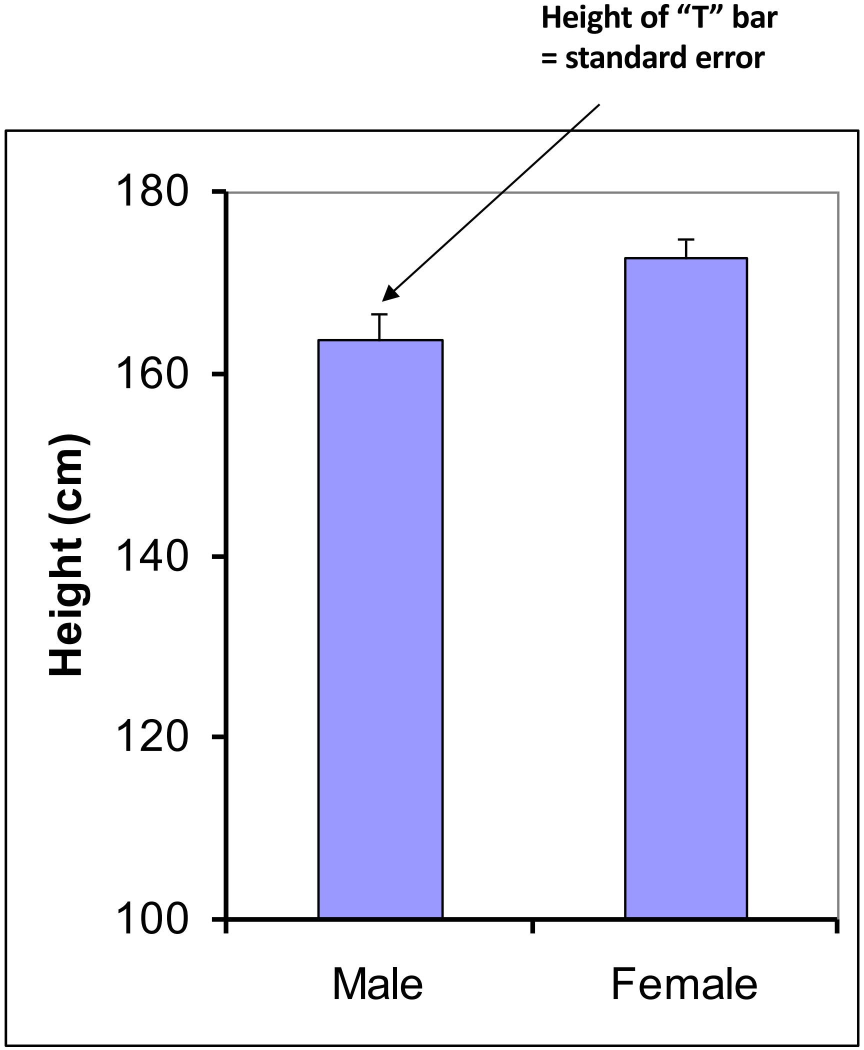
Represent the experimental data in a figure

Box plot of male and female heights



Minitab: Graph/Box Plot/Multipe Ys/Simple

Bar chart of mean male and female heights with error bars showing standard error of mean



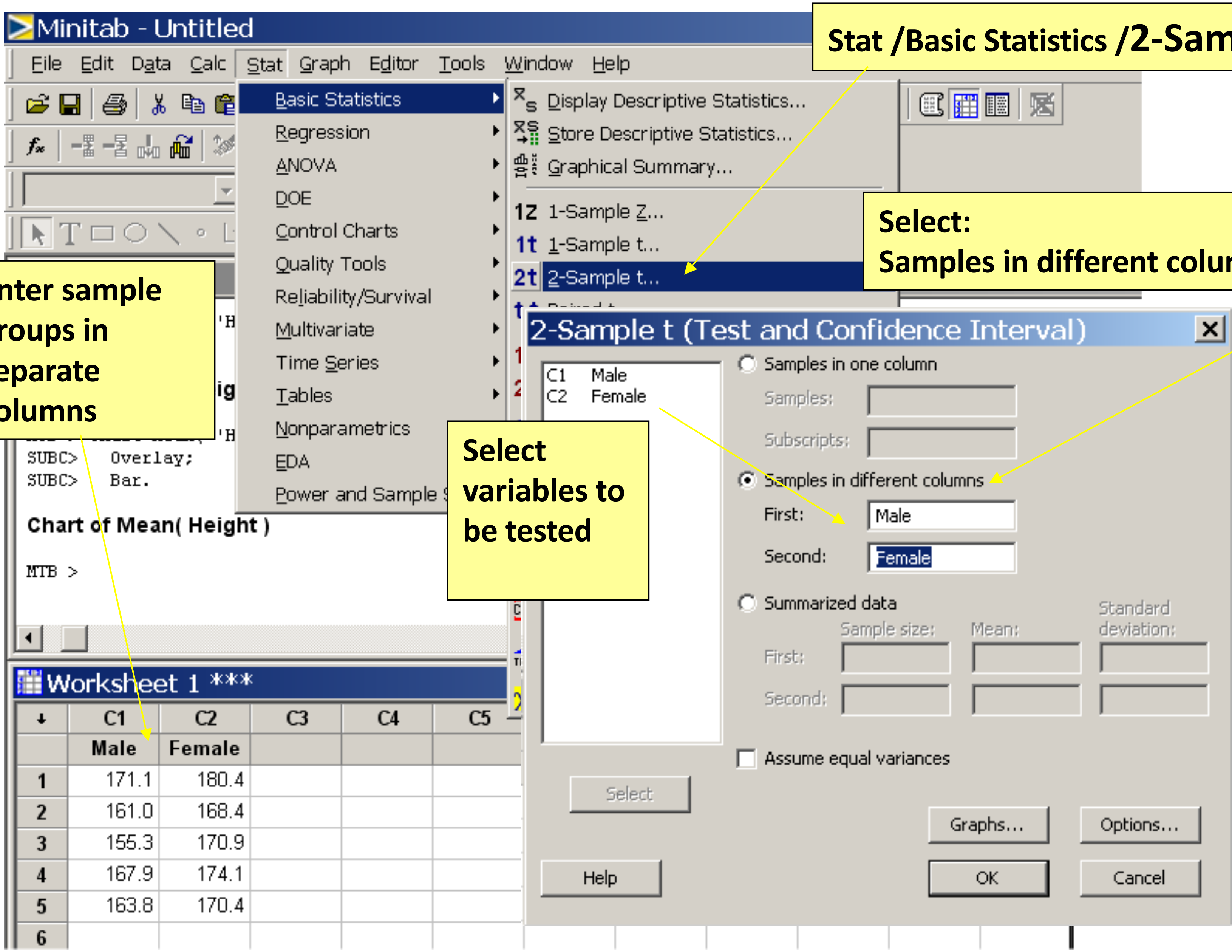
Note. Bar chart plotted using Excel

Example: Hypotheses

- **Null hypothesis**
  - Average height of men and women in population **do not** differ.
  - i.e. The observed differences in sample means are due only to random individual variation.

- **Alternate hypothesis**
  - Men and women **do** have different average heights

Minitab: Calculating a T Test (unpaired)



Stat /Basic Statistics /2-Sample t..

Select:
Samples in different columns

Enter sample groups in separate columns

Select variables to be tested

Minitab: T Test (unpaired) Results

Two-Sample T-Test and CI: Male, Female

```
Two-sample T for Male vs Female

        N    Mean  StDev  SE Mean
Male    5  163.82   6.13      2.7
Female  5  172.84   4.69      2.1


Difference = mu (Male) - mu (Female)
Estimate for difference:  -9.02
95% CI for difference:  (-17.18, -0.86)
T-Test of difference = 0 (vs not =): T-Value = -2.61  P-Value = 0.035  DF = 7
```

T-Value = -2.61

T score

P-Value = 0.035

Significance probability:
The probability of obtaining an absolute t value this large or larger given that the null hypothesis is true.

Interpreting the Unpaired T Test Results

- The mean height of the male sample is 9 cm larger than the female sample
- The T test reports a t value of 2.31 with a significance probability of p=**0.031**
- A t value of this magnitude would only occur in 3.1% of the samples if the null hypothesis H_0 (no difference between male/female mean height) were true.
- The standard criterion for rejecting the null hypothesis is **$p \leq 0.05$** (i.e. the t value must have less than a 5% chance of occurring when H_0 is true)
- We can therefore **reject** the null hypothesis and conclude that **there is a difference** between average male and female height.

Mean sample height (male)	172.8 cm
Mean sample height (female)	163.8 cm
Mean (male) – Mean (female)	9.0 cm
T	2.31
Significance Probability p	0.031

T Test (paired)

- Determine whether a significant difference exists between repeated measures of a variable from a single sample group before and after some kind of treatment.

$$D = X_1 - X_2$$

- **Null hypothesis**
  - The treatment has had no effect on the measured variable

- **Test statistic - T**
  - The mean of the differences in the measured variable before and after treatment for each subject divided by the standard error of the differences

$$t = \frac{\overline{D}}{SE_D}$$

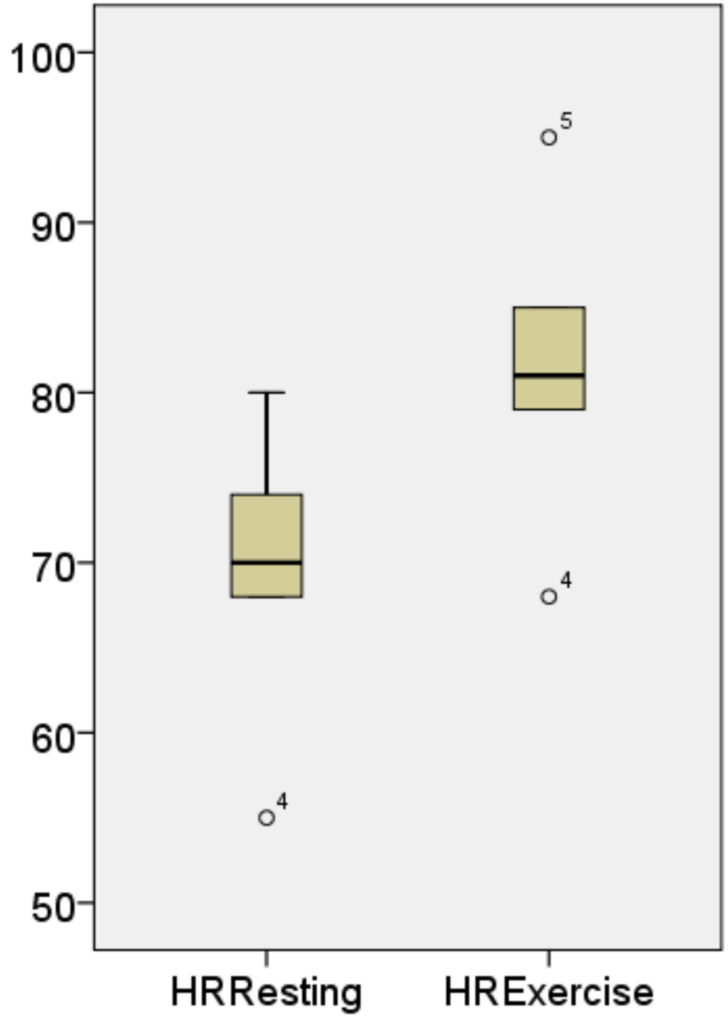
Example: Effect of exercise on heart rate?

- **Population**
  - First year undergraduates
- **Hypothesis**
  - Does exercise change heart rate?
- **Variables**
  - Heart Rate at rest
  - Heart Rate after exercise
- **Hypothesis Test**
  - Paired-sample T Test

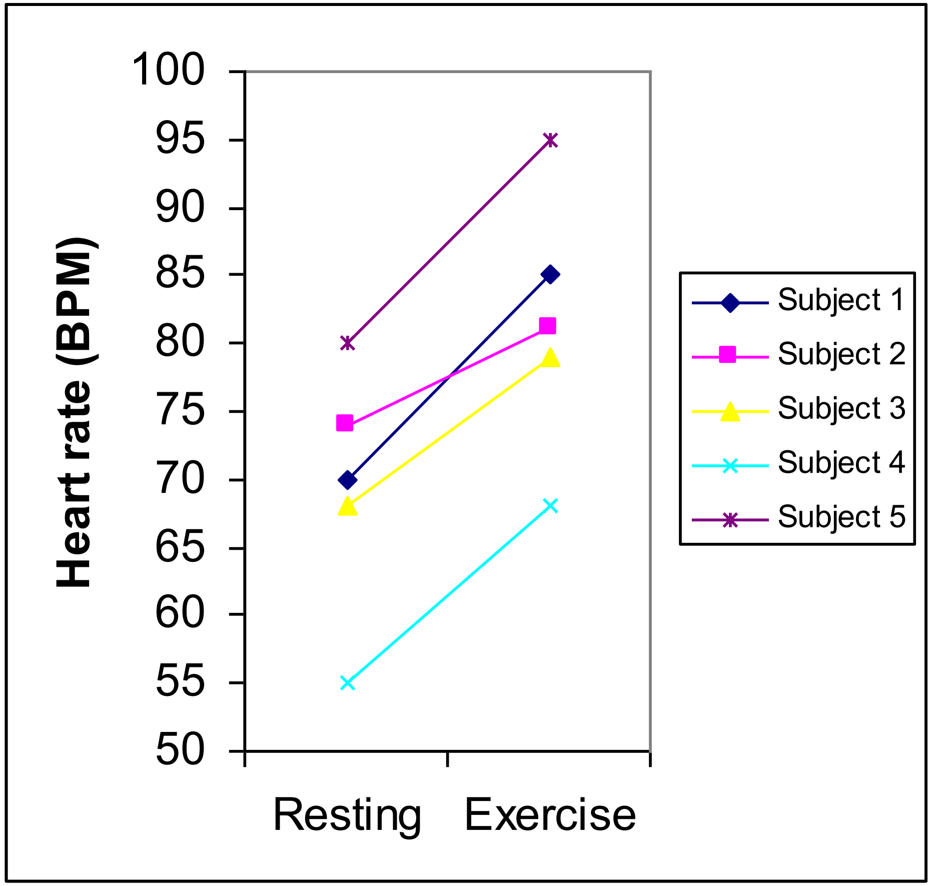
Subject #	Heart Rate (Resting) Beats/min	Heart Rate (After Exercise) Beats/min
1	70	85
2	74	81
3	68	79
4	55	68
5	80	95

Experimental Data Figures

Box plot of heart rates before and after exercise



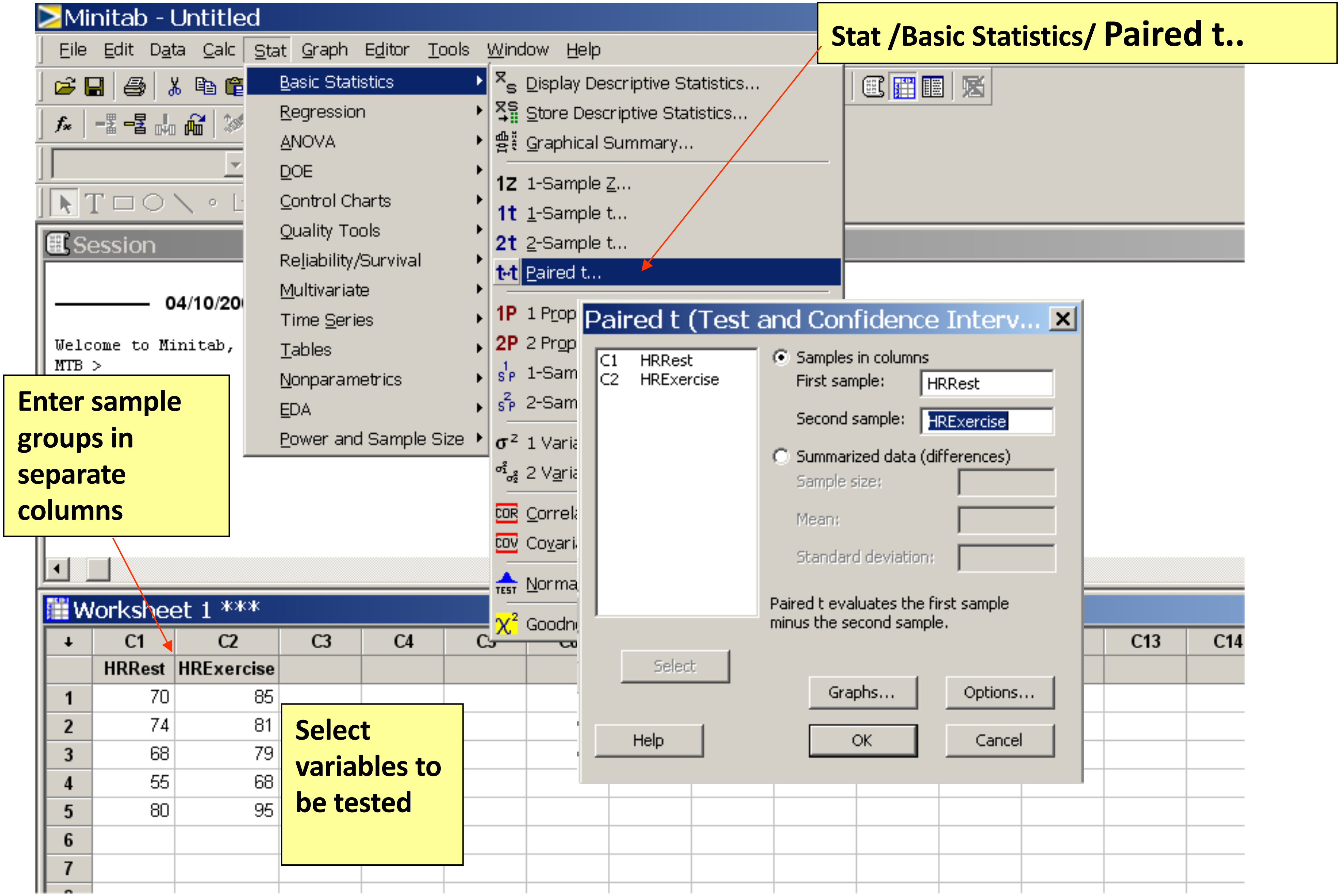
Line plot of changes in heart rate before and after exercise



An unpaired T test gives p = 0.083, Not significant!

Note. This graph produced using Excel

Minitab: Calculating a paired T Test



Minitab: Paired T Test Results

Paired T-Test and CI: HRRest, HRExercise

```
Paired T for HRRest - HRExercise

              N     Mean   StDev   SE Mean
HRRest        5    69.40    9.26      4.14
HRExercise    5    81.60    9.79      4.38
Difference    5   -12.20    3.35      1.50


95% CI for mean difference: (-16.36, -8.04)
T-Test of mean difference = 0 (vs not = 0): T-Value = -8.15  P-Value = 0.001
```

T score

Mean difference: -12.20

T-Value = -8.15

Since p = 0.001 is less than 0.05, we can **reject** the null hypothesis and conclude that exercise significantly increases heart rate.

P-Value = 0.001

Significance Probability (p)

Choosing 2 Sample Hypothesis Tests

- If possible, design your experiments to repeat measurements on the same group of subjects, to allow the use of paired tests (paired T test, Wilcoxon test) which are more powerful than unpaired test.

- If your data variables are of the ratio or interval type, and there is no evidence of a non-normal distribution use the T test.

- If your data variables are ordinal, or there is evidence of a non-normal distribution use a non-parametric test such as the Mann-Whitney U Test (unpaired data) or Wilcoxon Test (paired data)