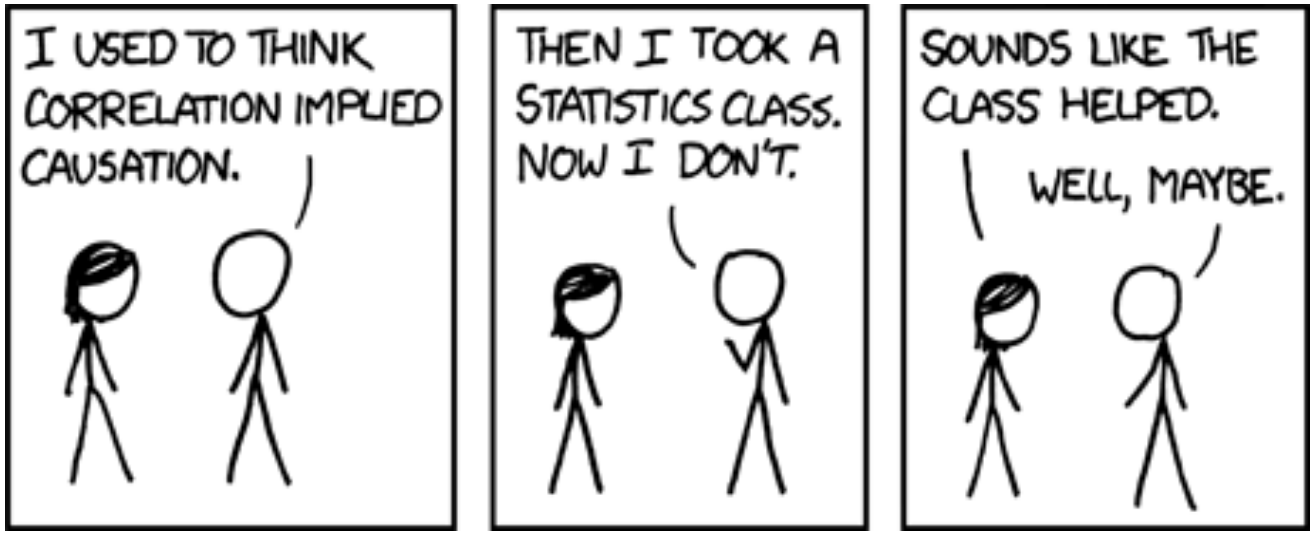


Statistical Analysis

How do we know if it works?

Group workbook:

Cartoon from XKCD.com. Subscribe! http://www.xkcd.com/552/

Significant Concepts

We **structure** the presentation and processing of data to suit the **function** of clear communication and interpretation

By understanding the **relationship** between raw and processed data and between mean and standard deviation we can determine the significance of **change**.

	Assessment Statement (You should know - You need to do)	**Obj.**
1.1.1	**State** that error bars are a graphical representation of the variability of data. • *These can show range, standard deviation or 95% confidence intervals.*	1
1.1.2	**Calculate** the mean and standard deviation of a set of values. • *Using your TI Inspire (for the exam)* • *Using Excel (for labs)*	2
1.1.3	**State** that the term standard deviation is used to summarize the spread of values around the mean, and that 68% of the values fall within one standard deviation of the mean	1
1.1.4	**Explain** how the standard deviation is useful for comparing the means and the spread of data between two or more samples. • *A greater standard deviation means a greater spread of data around the mean.* • *This can be used to infer variability.*	3
1.1.5	**Deduce** the significance of the difference between two sets of data using calculated values for *t* and the appropriate tables. • Use the value of t (given) from a table, comparing it to critical values • Use the =TTEST function in Excel to get to the P value	3
1.1.6	**Explain** that the existence of a correlation does not establish that there is a causal relationship between two variables.	3

Extension:

- SL/HL Maths: How is standard deviation calculated and why do we do it this way?

Statistical Analysis

Understand:

We **structure** the presentation and processing of data to suit the **function** of clear communication and interpretation

By understanding the **relationship** between raw and processed data and between mean and standard deviation we can determine the significance of **change**.

Know:

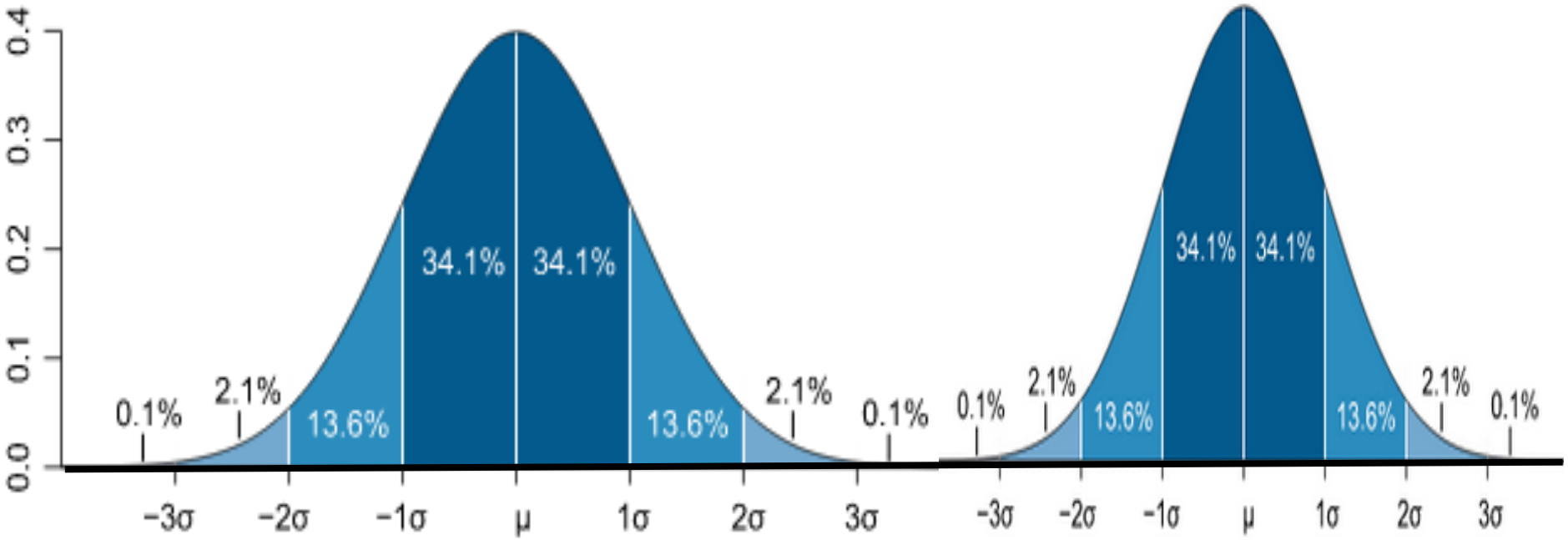
- How to calculate mean and standard deviation (s)
- s is a measure of the spread of a set of data
- Error bars can represent variability of data (STDEV, 95% confidence)
- t-tests determine the significance of a difference between means
- correlation does not imply causation

Do:

- Calculate mean & standard deviation using Excel & TI Inspire
- Plot and interpret custom error bars on a graph in Excel
- Determine the significance of the difference between two means using
  - Excel (P values)

Standard Deviation is used to summarise the **spread** of the values around the **mean**.

- 34% of data is one s above the mean, 34% below (68% within ± 1s)
- A large value for the standard deviation means that data are more spread
- This can suggest reliability of data
- It could be used to determine if a single data point is outside the 'normal' range

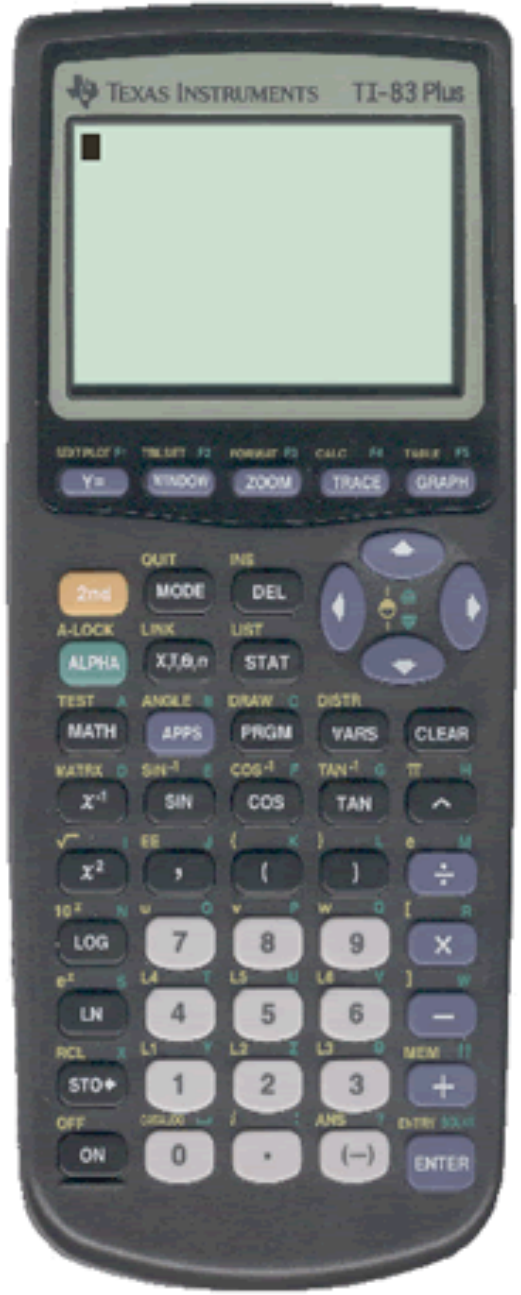


http://en.wikipedia.org/wiki/Standard_deviation

Compare the standard deviations for the class set of data on the Classification quiz before and after the classes. What inferences can you make? Are there any data outside the 'normal range?

Extension: if you want to know more, click here.

Simple Stats with the TI 83/84

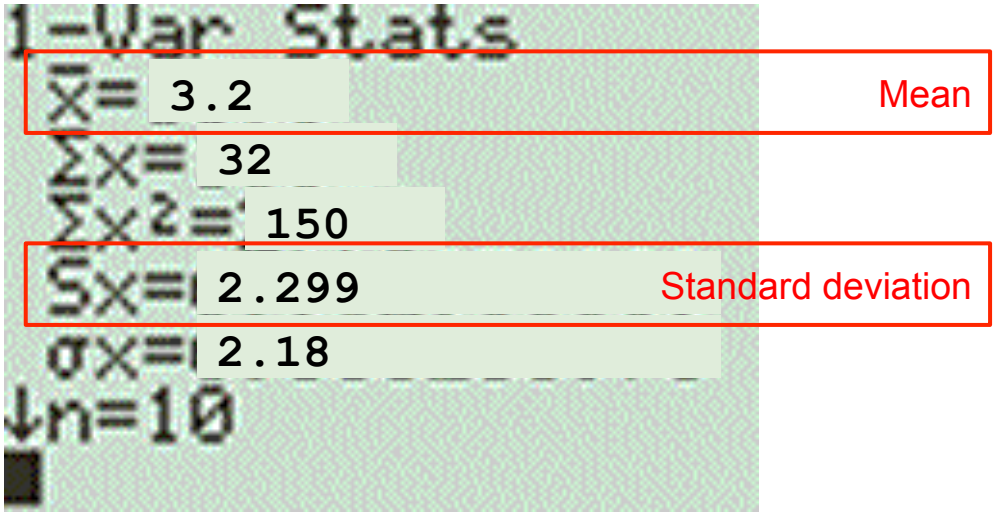


Follow these instructions:

- http://click4biology.info/c4b/1/gcStat.htm

Complete the steps for these two sets of data. Make sure your answers match with these Excel results (if in doubt, check on Excel):

	DV (unit, ±uncertainty)											
repeat	1	2	3	4	5	6	7	8	9	10	mean	sd
Population A	1.0	1.0	2.0	3.0	2.0	1.0	3.0	6.0	7.0	6.0	3.2	2.3
Population B	1.0	1.0	1.0	1.0	1.0	1.0	2.0	1.0	2.0	1.0	1.2	0.4



http://click4biology.info/c4b/1/gcStat.htm

Simple Stats with the TI Nspire

Follow these instructions:

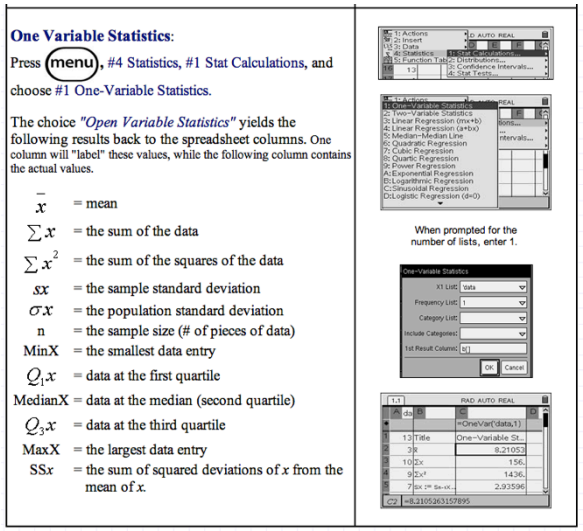
- http://mathbits.com/MathBits/TINSection/Statistics1/Spreadsheet.html

Complete the steps for these two sets of data. Make sure your answers match with these Excel results (if in doubt, check on Excel):

	DV (unit, ±uncertainty)											
repeat	1	2	3	4	5	6	7	8	9	10	mean	sd
Population A	1.0	1.0	2.0	3.0	2.0	1.0	3.0	6.0	7.0	6.0	3.2	2.3
Population B	1.0	1.0	1.0	1.0	1.0	1.0	2.0	1.0	2.0	1.0	1.2	0.4

$\bar{x}$ = mean (3.2)

sx = the sample standard deviation (2.299)



One Variable Statistics:

Press (menu), #4 Statistics, #1 Stat Calculations, and choose #1 One-Variable Statistics.

The choice *"Open Variable Statistics"* yields the following results back to the spreadsheet columns. One column will "label" these values, while the following column contains the actual values.

$\bar{x}$	= mean
$\sum x$	= the sum of the data
$\sum x^2$	= the sum of the squares of the data
sx	= the sample standard deviation
σx	= the population standard deviation
n	= the sample size (# of pieces of data)
MinX	= the smallest data entry
Q_1x	= data at the first quartile
MedianX	= data at the median (second quartile)
Q_3x	= data at the third quartile
MaxX	= the largest data entry
SSx	= the sum of squared deviations of x from the mean of x.

When prompted for the number of lists, enter 1.

http://mathbits.com/MathBits/TINSection/Statistics1/Spreadsheet.html

Table 1: Class scores on the Classification Quia Quiz

	Before Lessons	After Lessons
	7	25
	0	23
	17	22
	15	19
	21	22
	3	23
	17	24
	19	23
	20	25
	13	21
	18	24
	13	22
	0	20
	13	19
	1	24
	0	18
	9	23
	1	14
	11	14
	14	24
	17	21
	0	16
n=	22	22
Mean=	**10.4**	**21.2**
Standard deviation=	7.56	3.29
95% Confidence	3.16	1.34
T-Test P=	0.0	

Excel Formulas to Remember

Look in the cells for the descriptive statistics used in this class task.

State the formulas for:

Mean: “=

Standard deviation: “=

To calculate 95% confidence:

=CONFIDENCE.NORM(**0.05**,C28,22)

nearly always this in Bio

calculated value (or cell) of STDEV

n (the sample size)

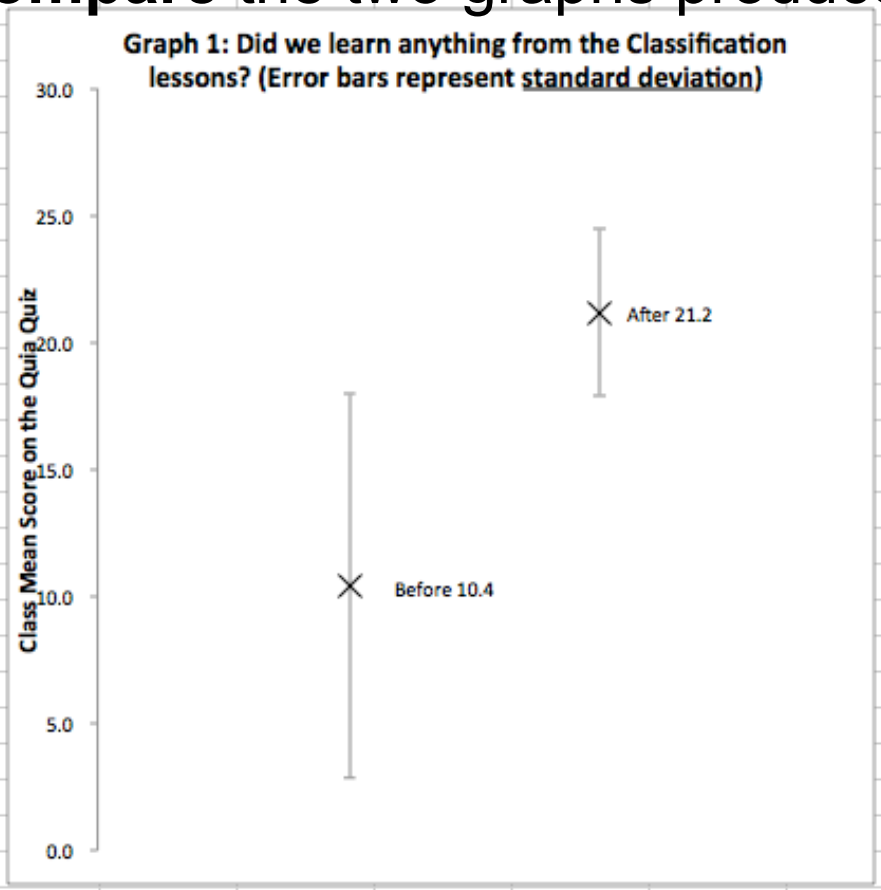
To calculate the P value for a t-test:

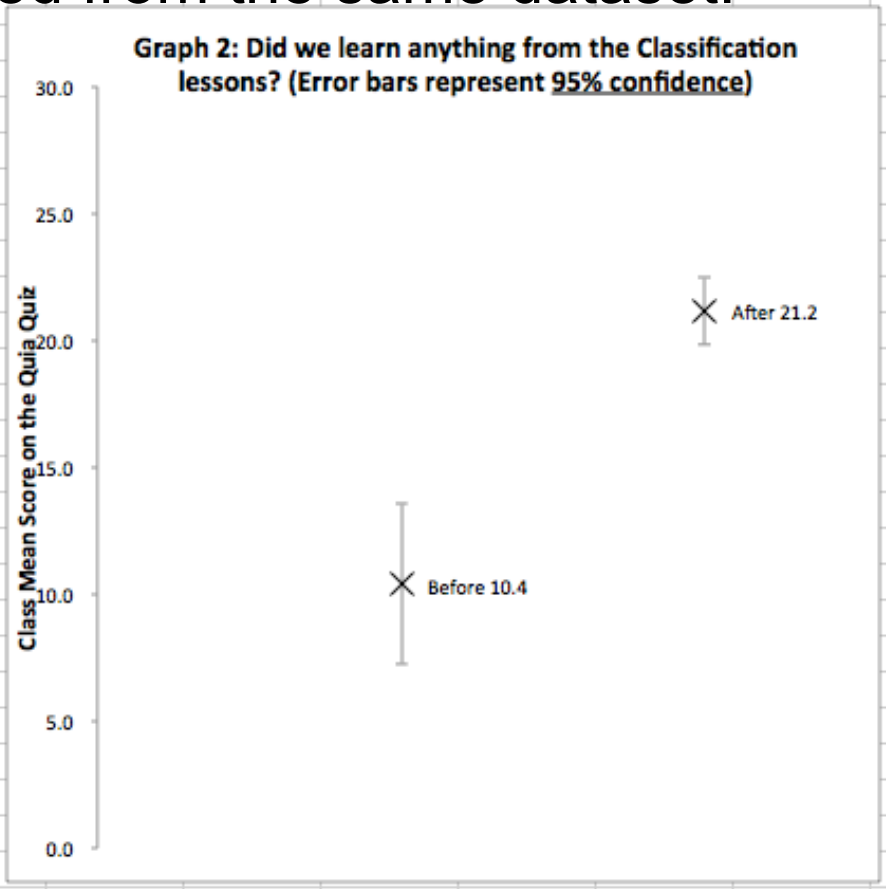
=TTEST(column1raw,column2raw,2,2)

always 2,2 for these **unpaired** t-tests

Extension: what is the difference between a ‘paired’ and an ‘unpaired’ t-test? When might you use each?

Compare the two graphs produced from the same dataset:





Compare the size of the error bars (Stdev vs 95% confidence):

-

Compare the overlap of the error bars (Stdev vs 95% confidence):

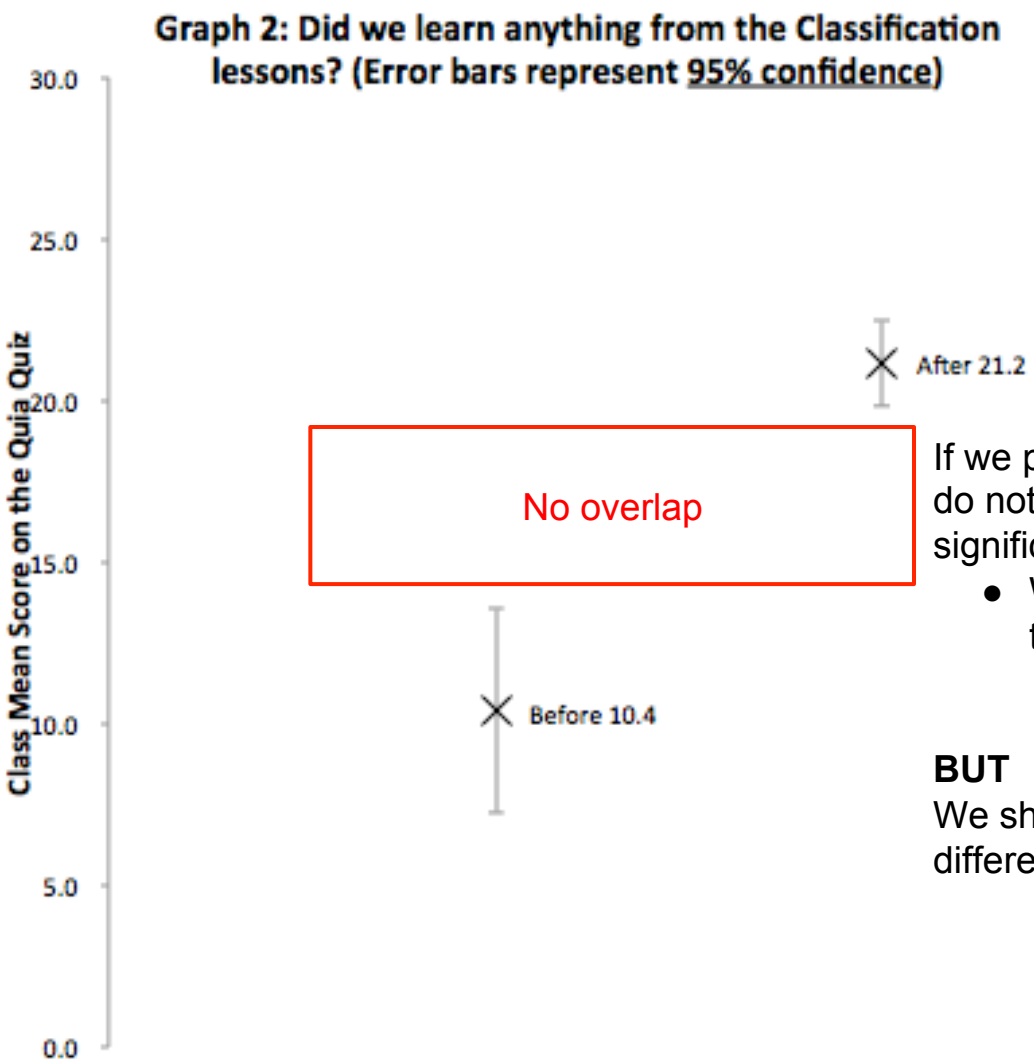
-

Do these results suggest a significant difference between Before/After class?

- YES/NO
- Reason:

How do we know if a difference is significant?

When we collect sets of data, there may be overlap in the values.
How do we know if this overlap is too much for the difference to be considered 'significant?'



If we plot **95% significance** on the error bars and they do not overlap, this suggests that the data are significantly different:

- We are **95% sure** that the data are different and that this difference is **not caused by chance**.

BUT
We should really perform a t-test to show whether the difference is significant or not.

Carrying out the t-test in Excel

Edit the formula in the merged cell.
To calculate the P value for a t-test:

=TTEST(column1raw,column2raw,2,2)

Always format your answers like this:

Null Hypothesis (H_0):
"There is no significant difference."

Alternate Hypothesis (H_1):
"There is a significant difference."

Determine P P =

If P > 0.05, accept (H_0)
If P > 0.05, reject (H_0)

Conclusion: accept/reject H_0:

-

There is/is not a significant difference in the results of the class before and after the lessons.

Table 1: Class scores on the Classification Quia Quiz

	Before Lessons	After Lessons
	7	25
	0	23
	17	22
	15	19
	21	22
	3	23
	17	24
	19	23
	20	25
	13	21
	18	24
	13	22
	0	20
	13	19
	1	24
	0	18
	9	23
	1	14
	11	14
	14	24
	17	21
	0	16
n=	22	22
Mean=	**10.4**	**21.2**
Standard deviation=	7.56	3.29
95% Confidence	3.16	1.34
T-Test P=	0.0	

Extension: what is the difference between a 'paired' and an 'unpaired' t-test? When might you use each?

Correlation does not imply causation.

Did the lessons really cause the class to learn more?
Discuss all the things we did over the week to learn about Classification.

Are we able to use the data we generated to attribute cause to any single factor? Why? Why not?

-

What would we need to do to be more sure of the true cause of the difference?

-

Extension: What lies behind the standard deviation results?

- **Find out more about the mathematics of standard deviation**
  - *What is the formula?*
  - *What do the components of the formula represent?*
  - *Why does this give us measure of the spread of the data?*

Steps involved:

- *Calculate population mean*
- *Calculate difference of each point from the population mean*
  - *square that result*
- *Calculate the mean of all these values (the squares of the differences)*
- *Find the square root of that value.*

SD = the root of the mean of the squares of the differences from the mean

Carrying out t-tests using t-tables.

Work through the problem set in pairs on this sheet:
https://docs.google.com/document/d/1gzcREFikPkZTAh2KM6e9Bq95YoCp68eAPeJVDfZYMmc/edit?usp=sharing

The table the right shows the critical values for values of t.

In Biology, we will usually use a **P-value of 0.05**. This means that we are 95% sure that differences in results are real, not due to chance (there is 5% or less overlap in data).

- ❑ **Identify** the column that shows the values for P=0.05. **Highlight** this column.
- ❑ **Draw** an arrow above the table to show the direction in which we become 'more confident' in our results.

"DF" means "**degrees of freedom**" - the **total sample size** (adding both groups together) **minus two**.

- ❑ **State** the degrees of freedom in these tests:

A. There are 20 samples in total. DF= _____
B. Group A has 16 samples, Group B has 10. DF= _____
C. 11 *before* and 11 *after* samples are taken. DF= _____

The critical value is found by cross-referencing the DF with the P=0.05 column. Find the critical values for the samples above.

A. cv = _____ B. cv = _____ C. cv = _____

A **t-value** is calculated from the sample data. We do not need to know how to do this. We do need to know how to determine significance by comparing 't' to the critical value.

Table from: http://www.medcalc.org/manual/t-distribution.php

DF	A P	0.80 0.20	0.90 0.10	0.95 0.05	0.98 0.02
1		3.078	6.314	12.706	31.820
2		1.886	2.920	4.303	6.965
3		1.638	2.353	3.182	4.541
4		1.533	2.132	2.776	3.747
5		1.476	2.015	2.571	3.365
6		1.440	1.943	2.447	3.143
7		1.415	1.895	2.365	2.998
8		1.397	1.860	2.306	2.897
9		1.383	1.833	2.262	2.821
10		1.372	1.812	2.228	2.764
11		1.363	1.796	2.201	2.718
12		1.356	1.782	2.179	2.681
13		1.350	1.771	2.160	2.650
14		1.345	1.761	2.145	2.625
15		1.341	1.753	2.131	2.602
16		1.337	1.746	2.120	2.584
17		1.333	1.740	2.110	2.567
18		1.330	1.734	2.101	2.552
19		1.328	1.729	2.093	2.539
20		1.325	1.725	2.086	2.528
21		1.323	1.721	2.080	2.518
22		1.321	1.717	2.074	2.508
23		1.319	1.714	2.069	2.500
24		1.318	1.711	2.064	2.492
25		1.316	1.708	2.060	2.485
26		1.315	1.706	2.056	2.479
27		1.314	1.703	2.052	2.473
28		1.313	1.701	2.048	2.467
29		1.311	1.699	2.045	2.462
30		1.310	1.697	2.042	2.457
31		1.309	1.695	2.040	2.453
32		1.309	1.694	2.037	2.449

Deducing significance by comparing calculated values of t with the t-table.

- *In a test for significance we set out our working like this.*
- ***Use** this as an example for your work.*

Degrees of freedom = _____

H_0 (Null hypothesis):
"There is no significant difference."

P = 0.05

t = _____

H_1 (Alternate hypothesis)
"There is a significant difference."

critical value = _____

If t < cv, we accept H_0.
If t > cv, we reject H_0.

Conclusion:

There is / is not a significant difference between the results.

Table from: http://www.medcalc.org/manual/t-distribution.php

DF	A P	0.80 0.20	0.90 0.10	0.95 0.05	0.98 0.02
1		3.078	6.314	12.706	31.820
2		1.886	2.920	4.303	6.965
3		1.638	2.353	3.182	4.541
4		1.533	2.132	2.776	3.747
5		1.476	2.015	2.571	3.365
6		1.440	1.943	2.447	3.143
7		1.415	1.895	2.365	2.998
8		1.397	1.860	2.306	2.897
9		1.383	1.833	2.262	2.821
10		1.372	1.812	2.228	2.764
11		1.363	1.796	2.201	2.718
12		1.356	1.782	2.179	2.681
13		1.350	1.771	2.160	2.650
14		1.345	1.761	2.145	2.625
15		1.341	1.753	2.131	2.602
16		1.337	1.746	2.120	2.584
17		1.333	1.740	2.110	2.567
18		1.330	1.734	2.101	2.552
19		1.328	1.729	2.093	2.539
20		1.325	1.725	2.086	2.528
21		1.323	1.721	2.080	2.518
22		1.321	1.717	2.074	2.508
23		1.319	1.714	2.069	2.500
24		1.318	1.711	2.064	2.492
25		1.316	1.708	2.060	2.485
26		1.315	1.706	2.056	2.479
27		1.314	1.703	2.052	2.473
28		1.313	1.701	2.048	2.467
29		1.311	1.699	2.045	2.462
30		1.310	1.697	2.042	2.457
31		1.309	1.695	2.040	2.453
32		1.309	1.694	2.037	2.449

1. A researcher measures the wing-spans of **12** red-throat and **13** broad-billed hummingbirds. The **t-value is calculated as 2.15**. **Deduce** the significance of the difference between the two types of birds.

Degrees of freedom = _____

H_0 (Null hypothesis):
"There is no significant difference."

P = 0.05

t = _____

H_1 (Alternate hypothesis)
"There is a significant difference."

critical value = _____

If t < cv, we accept H_0.
If t > cv, we reject H_0.

Conclusion:
There <u>is / is not</u> a significant difference between the results.

DF	A P	0.80 0.20	0.90 0.10	0.95 0.05	0.98 0.02
1		3.078	6.314	12.706	31.820
2		1.886	2.920	4.303	6.965
3		1.638	2.353	3.182	4.541
4		1.533	2.132	2.776	3.747
5		1.476	2.015	2.571	3.365
6		1.440	1.943	2.447	3.143
7		1.415	1.895	2.365	2.998
8		1.397	1.860	2.306	2.897
9		1.383	1.833	2.262	2.821
10		1.372	1.812	2.228	2.764
11		1.363	1.796	2.201	2.718
12		1.356	1.782	2.179	2.681
13		1.350	1.771	2.160	2.650
14		1.345	1.761	2.145	2.625
15		1.341	1.753	2.131	2.602
16		1.337	1.746	2.120	2.584
17		1.333	1.740	2.110	2.567
18		1.330	1.734	2.101	2.552
19		1.328	1.729	2.093	2.539
20		1.325	1.725	2.086	2.528
21		1.323	1.721	2.080	2.518
22		1.321	1.717	2.074	2.508
23		1.319	1.714	2.069	2.500
24		1.318	1.711	2.064	2.492
25		1.316	1.708	2.060	2.485
26		1.315	1.706	2.056	2.479
27		1.314	1.703	2.052	2.473
28		1.313	1.701	2.048	2.467
29		1.311	1.699	2.045	2.462
30		1.310	1.697	2.042	2.457
31		1.309	1.695	2.040	2.453
32		1.309	1.694	2.037	2.449

Table from: http://www.medcalc.org/manual/t-distribution.php

1. A researcher measures the wing-spans of **12** red-throat and **13** broad-billed hummingbirds. The **t-value is calculated as 2.15**. **Deduce** the significance of the difference between the two types of birds.

DF = 12 + 13 -2 = 23

H_0 (Null hypothesis):
"There is no significant difference."

P = 0.05

t = 2.15

H_1 (Alternate hypothesis)
"There is a significant difference."

critical value = 2.069

If t < cv, we accept H_0.
If t > cv, we reject H_0.

Conclusion:
There is a significant difference between the results.

Table from: http://www.medcalc.org/manual/t-distribution.php

DF	A P	0.80 0.20	0.90 0.10	0.95 0.05	0.98 0.02
1		3.078	6.314	12.706	31.820
2		1.886	2.920	4.303	6.965
3		1.638	2.353	3.182	4.541
4		1.533	2.132	2.776	3.747
5		1.476	2.015	2.571	3.365
6		1.440	1.943	2.447	3.143
7		1.415	1.895	2.365	2.998
8		1.397	1.860	2.306	2.897
9		1.383	1.833	2.262	2.821
10		1.372	1.812	2.228	2.764
11		1.363	1.796	2.201	2.718
12		1.356	1.782	2.179	2.681
13		1.350	1.771	2.160	2.650
14		1.345	1.761	2.145	2.625
15		1.341	1.753	2.131	2.602
16		1.337	1.746	2.120	2.584
17		1.333	1.740	2.110	2.567
18		1.330	1.734	2.101	2.552
19		1.328	1.729	2.093	2.539
20		1.325	1.725	2.086	2.528
21		1.323	1.721	2.080	2.518
22		1.321	1.717	2.074	2.508
23		1.319	1.714	2.069	2.500
24		1.318	1.711	2.064	2.492
25		1.316	1.708	2.060	2.485
26		1.315	1.706	2.056	2.479
27		1.314	1.703	2.052	2.473
28		1.313	1.701	2.048	2.467
29		1.311	1.699	2.045	2.462
30		1.310	1.697	2.042	2.457
31		1.309	1.695	2.040	2.453
32		1.309	1.694	2.037	2.449

2. A student measures the shells of 16 snails on the north side of an island and 15 on the south. The t-value is calculated as 1.61. **Deduce** the significance of the difference between the two types of snails.

DF =

H_0 (Null hypothesis):
"There is no significant difference."

P = 0.05

t =

H_1 (Alternate hypothesis)
"There is a significant difference."

critical value =

If t < cv, we accept H_0.
If t > cv, we reject H_0.

Conclusion:
There <u>is / is not</u> a significant difference between the results.

DF	A P	0.80 0.20	0.90 0.10	0.95 0.05	0.98 0.02
1		3.078	6.314	12.706	31.820
2		1.886	2.920	4.303	6.965
3		1.638	2.353	3.182	4.541
4		1.533	2.132	2.776	3.747
5		1.476	2.015	2.571	3.365
6		1.440	1.943	2.447	3.143
7		1.415	1.895	2.365	2.998
8		1.397	1.860	2.306	2.897
9		1.383	1.833	2.262	2.821
10		1.372	1.812	2.228	2.764
11		1.363	1.796	2.201	2.718
12		1.356	1.782	2.179	2.681
13		1.350	1.771	2.160	2.650
14		1.345	1.761	2.145	2.625
15		1.341	1.753	2.131	2.602
16		1.337	1.746	2.120	2.584
17		1.333	1.740	2.110	2.567
18		1.330	1.734	2.101	2.552
19		1.328	1.729	2.093	2.539
20		1.325	1.725	2.086	2.528
21		1.323	1.721	2.080	2.518
22		1.321	1.717	2.074	2.508
23		1.319	1.714	2.069	2.500
24		1.318	1.711	2.064	2.492
25		1.316	1.708	2.060	2.485
26		1.315	1.706	2.056	2.479
27		1.314	1.703	2.052	2.473
28		1.313	1.701	2.048	2.467
29		1.311	1.699	2.045	2.462
30		1.310	1.697	2.042	2.457
31		1.309	1.695	2.040	2.453
32		1.309	1.694	2.037	2.449

Table from: http://www.medcalc.org/manual/t-distribution.php

	Group A	Group B	DF	t =	cv =	Accept or Reject H_0?	Significant difference?
3a	n = 5	n = 7		2.405			
3b	n = 12	n = 12		2.029			

DF	A P	0.80 0.20	0.90 0.10	0.95 0.05	0.98 0.02
1		3.078	6.314	12.706	31.820
2		1.886	2.920	4.303	6.965
3		1.638	2.353	3.182	4.541
4		1.533	2.132	2.776	3.747
5		1.476	2.015	2.571	3.365
6		1.440	1.943	2.447	3.143
7		1.415	1.895	2.365	2.998
8		1.397	1.860	2.306	2.897
9		1.383	1.833	2.262	2.821
10		1.372	1.812	2.228	2.764
11		1.363	1.796	2.201	2.718
12		1.356	1.782	2.179	2.681
13		1.350	1.771	2.160	2.650
14		1.345	1.761	2.145	2.625
15		1.341	1.753	2.131	2.602
16		1.337	1.746	2.120	2.584
17		1.333	1.740	2.110	2.567
18		1.330	1.734	2.101	2.552
19		1.328	1.729	2.093	2.539
20		1.325	1.725	2.086	2.528
21		1.323	1.721	2.080	2.518
22		1.321	1.717	2.074	2.508
23		1.319	1.714	2.069	2.500
24		1.318	1.711	2.064	2.492
25		1.316	1.708	2.060	2.485

Table from: http://www.medcalc.org/manual/t-distribution.php

4. 15 students take a quiz <u>before</u> and <u>after</u> a set of lessons and the results compared. The t-value is calculated as 1.94. **Deduce** the significance of the difference between the results before and after.

DF =

H_0 (Null hypothesis):
"There is no significant difference."

P = 0.05

t =

H_1 (Alternate hypothesis)
"There is a significant difference."

critical value =

If t < cv, we accept H_0.
If t > cv, we reject H_0.

Conclusion:
There <u>is / is not</u> a significant difference between the results.

Table from: http://www.medcalc.org/manual/t-distribution.php

DF	A P	0.80 0.20	0.90 0.10	0.95 0.05	0.98 0.02
1		3.078	6.314	12.706	31.820
2		1.886	2.920	4.303	6.965
3		1.638	2.353	3.182	4.541
4		1.533	2.132	2.776	3.747
5		1.476	2.015	2.571	3.365
6		1.440	1.943	2.447	3.143
7		1.415	1.895	2.365	2.998
8		1.397	1.860	2.306	2.897
9		1.383	1.833	2.262	2.821
10		1.372	1.812	2.228	2.764
11		1.363	1.796	2.201	2.718
12		1.356	1.782	2.179	2.681
13		1.350	1.771	2.160	2.650
14		1.345	1.761	2.145	2.625
15		1.341	1.753	2.131	2.602
16		1.337	1.746	2.120	2.584
17		1.333	1.740	2.110	2.567
18		1.330	1.734	2.101	2.552
19		1.328	1.729	2.093	2.539
20		1.325	1.725	2.086	2.528
21		1.323	1.721	2.080	2.518
22		1.321	1.717	2.074	2.508
23		1.319	1.714	2.069	2.500
24		1.318	1.711	2.064	2.492
25		1.316	1.708	2.060	2.485
26		1.315	1.706	2.056	2.479
27		1.314	1.703	2.052	2.473
28		1.313	1.701	2.048	2.467
29		1.311	1.699	2.045	2.462
30		1.310	1.697	2.042	2.457
31		1.309	1.695	2.040	2.453
32		1.309	1.694	2.037	2.449

5. In a test of a heart medication, a control group of 11 patients was compared to the test group of 14 patients. The t-value is calculated as 2.225. **Deduce** the significance of the difference between the results of the control and the test group.

DF =

H_0 (Null hypothesis):
"There is no significant difference."

P = 0.05

t =

H_1 (Alternate hypothesis)
"There is a significant difference."

critical value =

If t < cv, we accept H_0.
If t > cv, we reject H_0.

Conclusion:
There <u>is / is not</u> a significant difference between the results.

DF	A P	0.80 0.20	0.90 0.10	0.95 0.05	0.98 0.02
1		3.078	6.314	12.706	31.820
2		1.886	2.920	4.303	6.965
3		1.638	2.353	3.182	4.541
4		1.533	2.132	2.776	3.747
5		1.476	2.015	2.571	3.365
6		1.440	1.943	2.447	3.143
7		1.415	1.895	2.365	2.998
8		1.397	1.860	2.306	2.897
9		1.383	1.833	2.262	2.821
10		1.372	1.812	2.228	2.764
11		1.363	1.796	2.201	2.718
12		1.356	1.782	2.179	2.681
13		1.350	1.771	2.160	2.650
14		1.345	1.761	2.145	2.625
15		1.341	1.753	2.131	2.602
16		1.337	1.746	2.120	2.584
17		1.333	1.740	2.110	2.567
18		1.330	1.734	2.101	2.552
19		1.328	1.729	2.093	2.539
20		1.325	1.725	2.086	2.528
21		1.323	1.721	2.080	2.518
22		1.321	1.717	2.074	2.508
23		1.319	1.714	2.069	2.500
24		1.318	1.711	2.064	2.492
25		1.316	1.708	2.060	2.485
26		1.315	1.706	2.056	2.479
27		1.314	1.703	2.052	2.473
28		1.313	1.701	2.048	2.467
29		1.311	1.699	2.045	2.462
30		1.310	1.697	2.042	2.457
31		1.309	1.695	2.040	2.453
32		1.309	1.694	2.037	2.449

Table from: http://www.medcalc.org/manual/t-distribution.php

6. The weights of 12 babies were measured before and after a treatment of anti-parasite medication. The t-value is calculated as 3.112. **Deduce** the significance of the difference between the results of the control and the test group.

DF =

H_0 (Null hypothesis):
"There is no significant difference."

P = 0.05

t =

H_1 (Alternate hypothesis)
"There is a significant difference."

critical value =

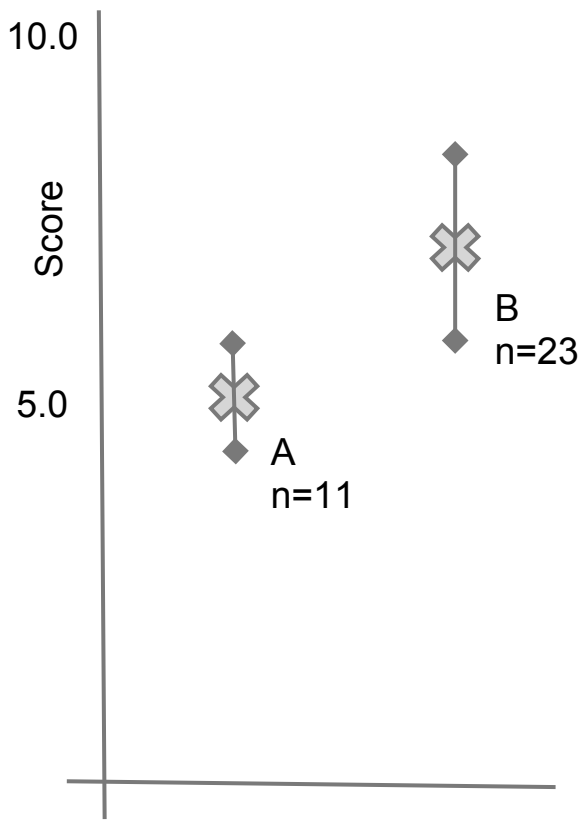
If t < cv, we accept H_0.
If t > cv, we reject H_0.

Conclusion:
There <u>is / is not</u> a significant difference between the results.

DF	A P	0.80 0.20	0.90 0.10	0.95 0.05	0.98 0.02
1		3.078	6.314	12.706	31.820
2		1.886	2.920	4.303	6.965
3		1.638	2.353	3.182	4.541
4		1.533	2.132	2.776	3.747
5		1.476	2.015	2.571	3.365
6		1.440	1.943	2.447	3.143
7		1.415	1.895	2.365	2.998
8		1.397	1.860	2.306	2.897
9		1.383	1.833	2.262	2.821
10		1.372	1.812	2.228	2.764
11		1.363	1.796	2.201	2.718
12		1.356	1.782	2.179	2.681
13		1.350	1.771	2.160	2.650
14		1.345	1.761	2.145	2.625
15		1.341	1.753	2.131	2.602
16		1.337	1.746	2.120	2.584
17		1.333	1.740	2.110	2.567
18		1.330	1.734	2.101	2.552
19		1.328	1.729	2.093	2.539
20		1.325	1.725	2.086	2.528
21		1.323	1.721	2.080	2.518
22		1.321	1.717	2.074	2.508
23		1.319	1.714	2.069	2.500
24		1.318	1.711	2.064	2.492
25		1.316	1.708	2.060	2.485
26		1.315	1.706	2.056	2.479
27		1.314	1.703	2.052	2.473
28		1.313	1.701	2.048	2.467
29		1.311	1.699	2.045	2.462
30		1.310	1.697	2.042	2.457
31		1.309	1.695	2.040	2.453
32		1.309	1.694	2.037	2.449

Table from: http://www.medcalc.org/manual/t-distribution.php

7. In this graph, two means are compared. Error bars show **standard deviation**.



State the mean of each population.

A = _____ B = _____

Compare the standard deviations of each group.

Can we tell from the error bars whether the two groups are significantly different?

The t-value of of the data is calculated as 2.025. What conclusion can be drawn from these data?

Experimental Design. How would you design an investigation into whether students are more alert in the morning or after lunch?

- ❑ What data might you collect?
- ❑ How would this be processed, graphed and analysed?
- ❑ What statistical test would you use?

EXIT TICKET: 2-minute Essay

Before you leave, **summarise** your learning on statistical analysis.

Include:

- ❑ Importance of standard deviation & uses of error bars
- ❑ Function of t-tests
- ❑ Correlation and causation

For more resources:

http://i-biology.net