Outcomes:

- Explain what standard deviation is and how it is calculated
- Interpret data in terms of means and the overlap of standard deviation bars and apply knowledge of, to draw and explain conclusions
- Evaluate the quality of results and reliability of conclusions. Explain how to ensure that estimates and conclusions are reliable.

This lesson will focus on the maths skills that are relevant to your Biology course. These maths skills are also very common in Biodiversity and Population questions.

Definition:

Mean: _____

Mode: _____

Median: _____

Standard deviations: _____

_____

Mean, modes and medians

Exam note: NEVER use the word "average" in your exam, this will lose you the mark.

Apply the definitions to the questions below.

Family	No. of kids
A	0
B	1
C	1
D	1
E	2
F	2
G	3
H	4
I	6
J	6
K	7

1. Work out the mean (show your working):

2. Work out the mode:

3. Work out the median (show your working):

Cow Name	Weight (kg)
Daisy	62
Dot	77
Fleur	67
Dandelion	71
Buttercup	63
Cowslip	92

4. Work out the mean (show your working):

5. Work out the mode:

6. Work out the median (show your working):

Standard deviation

What is it and how should it be used?

It takes into account **ALL the values**, not just the two most extreme ones - which is the range. E.g. If one of the extreme values is an outlier, then **the range will be misleading**.

This means standard deviation is **less influenced by anomalies.**

Standard deviation is a statistical test allows you to compare the means of 2 sets of data to see if the difference between the data sets is significant or not.

A standard deviation close to 0 indicates that the data points tend to be very close to the mean of the set, while a high standard deviation indicates that the data points are spread out over a wider range of values**.**

The **smaller** the standard deviation the **more accurate the mean**, the more reliable the results.

- If the standard deviations of the 2 sets of **data DO overlap then the DIFFERENCE between the 2 DATA SETS is NOT significant so NO conclusion** can be formed from the data.
- If the standard deviations of the 2 sets of **data DON'T overlap then the DIFFERENCE between the 2 DATA SETS IS significant so A conclusion** can be formed from the data.

Worked example:

The St. Kilda field mouse lives only on one island off the coast of Scotland. It is very similar in appearance to the long-tailed field mouse but is larger and has lighter coloured fur. Biologists wanted to find out if the St. Kilda field mouse and the long-tailed field mouse populations belonged to different species. They measured the length of the same features of a large number of individuals from the two populations. The results are shown in **Table 2**.

Population	**Mean length (±SD) / mm**			
	Head and body	Values	Tail	Values
St. Kilda field mouse	112.3 (±9.3)		105.5 (±8.4)	
Long-tailed field mouse	95.2 (±8.2)		90.2 (±7.3)	

(a) Do the data in **Table 2** provide evidence that the two populations belong to different species? Use calculations of ratios to support your answer.

_____

_____

_____

_____

_____

_____

_____

_____**(3)**

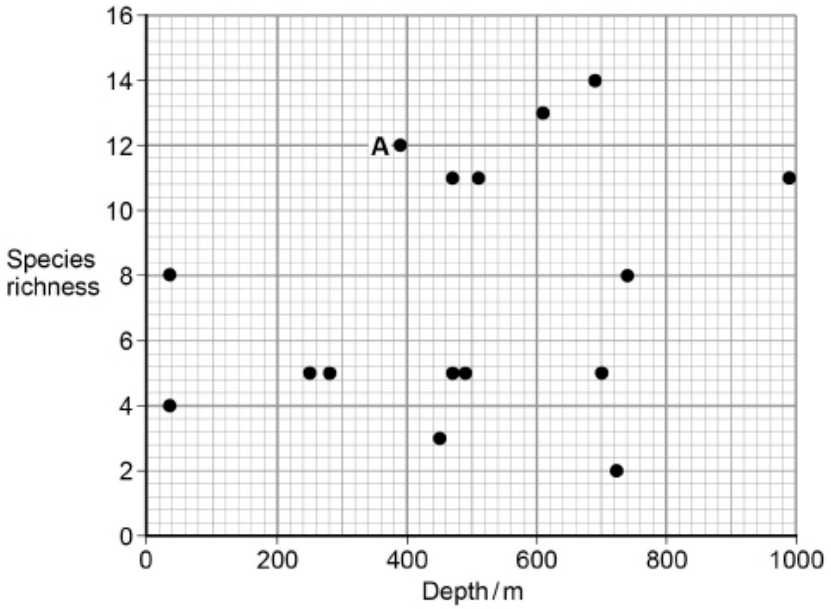
Exam Questions:

Q1. (a) Define each of the following terms.

Species _____

_____(1)

Scientists investigated the species richness of fish caught at various depths in the Pacific Ocean close to the western coast of Chile. The graph shows the scientists' results. 68% of all the fish caught in this investigation came from sample **A**.



(b) What is the modal value of species richness?

_____**(1)**

Q2. Mayflies are insects which lay their eggs in streams and rivers. The nymphs which hatch from the eggs live in the water for several years. Mayfly nymphs were collected by disturbing the gravel of a stream bed. A net placed immediately downstream caught any animals which were washed out of the gravel. Eight samples were collected from shallow, fast-flowing parts of the stream and eight from deeper, slow-flowing parts. Nymphs from two different families of mayfly were found. The results are given in the table.

	Family Caenidae		**Family Baetidae**	
	Shallow water	**Deep water**	**Shallow water**	**Deep water**
Mean number of nymphs	2.38	12.88	24.50	6.00
Standard deviation	1.51	7.92	6.72	1.51

(a) Which **one** of the four samples showed the greatest variation within the sample? Give evidence from the table for your answer.

_____

_____**(1)**

Exam note: You do not need to have studied the topic yet to be able to comment on the standard deviations.

Q3. The scientists' results are shown in the table below.

Treatment	**Mean dry mass / g per plant (standard deviation)**	
	Root	**Shoot**
A – heat-treated fungus and nitrate fertiliser	0.40 (±0.05)	1.01 (±0.12)
B – fungus and nitrate fertiliser	1.61 (±0.28)	9.81 (±0.33)
C – heat-treated fungus and ammonium fertiliser	0.34 (±0.03)	0.96 (±0.26)
D – fungus and ammonium fertiliser	0.96 (±0.18)	4.01 (±0.47)

Heat treating the fungus kills it so means the fungus can have no effect on the results.

(e) What conclusions can be drawn from the data in the table about the following?

The effects of the fungus on growth of the pea plants.

_____

_____

_____

The effects of nitrate fertiliser and ammonium fertiliser on growth of the pea plants.

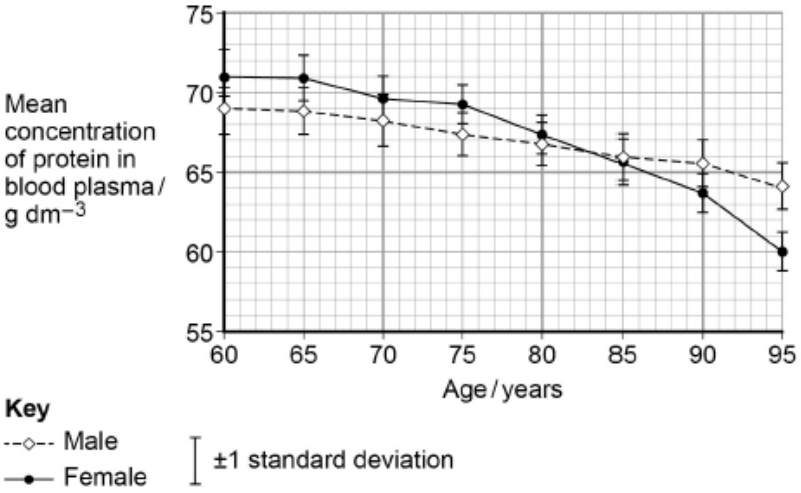
_____

_____

_____

_____ **(4)**

Q4. Scientists investigated how the concentration of protein in blood plasma changes in people between the ages of 60 and 95. The graph shows the scientists' results. The bars show ±1 standard deviation.



(a) What is the difference between males and females in the fall in mean concentration of protein in blood plasma between 60 and 95 years?

Answer = _____ g dm^{-3}

(1)

(b) Use the graph above to calculate the rate of change of the mean concentration of protein in the blood plasma of males between the ages of 60 and 95. Show your working.

Answer = _____ g dm^{-3} $year^{-1}$

(2)

(c) What can you conclude from the graph above about the effect of ageing on the mean concentration of protein in the blood plasma in males and females?

_____

_____

_____

_____

_____ **(2)**

Probability and chance:

Definitions

Chi-squared test: _____

_____

Student t- test: _____

_____

Correlation co-efficient: _____

_____

What do statistical tests do? _____

_____

If the p-value given in the question **is less than 0.05 or 5%:**

_____

_____

_____

If the p-value given in the question is **more than 0.05 or 5%:**

_____

_____

_____

If the p-value in a **student t-test** is more than 0.05 or 5%:

_____

_____

_____

Common mistake, write notes for yourself.:

_____

_____

Probability and chance- Example:

- Identify the statistical test.
- Check the p-values
- Identify which are more or less than 0.05 or 5%

Species of butterfly	Mean number of butterflies		P value
	In canopy	In understorey	
Prepona laertes	15	0	< 0.001
Archaeoprepona demophon	14	37	< 0.001
Zaretis itys	25	11	> 0.05
Memphis arachne	89	23	< 0.001
Memphis offa	21	3	< 0.001
Memphis xenocles	32	8	< 0.001

(a) The scientists carried out a statistical test to see if the difference in the distribution of each species between the canopy and understorey was due to chance. The P values obtained are shown in the table. Explain what the results of these statistical tests show.

_____

_____

_____

_____**(3)**

Exam questions:

Q1. A student investigated the duration of afterimages using red squares and purple squares. His results are shown in the table.

Colour of square	Cone cell(s) stimulated	Colour of afterimage	Mean duration of afterimage / seconds
Red	Red	Blue-green	15
Purple	Red and blue	Green	12

(a) Suggest a statistical test that would be appropriate for this investigation. Give a reason for your answer.

Statistical test _____

Reason for choice _____**(1)**

Q2. The scientists then calculated the mean mass of each species of alga eaten by the consumers. They used a statistical test to determine the P value. **Table 2** shows some of the results they obtained.

Species of alga	Mean mass eaten / g		P value
	Sea slug	Shore crab	
Laurencia pacifica	4.42	0.22	<0.01
Egregia leavigata	0.12	0.08	>0.05
Microcystis pyrifera	0.19	0.14	>0.05
Cystoseira osmondacea	0.17	0.04	<0.05

(a) Suggest what conclusions the scientists could have made from this investigation when using the probability values in **Table 2**.

_____

_____

_____

_____

_____

_____ **(3)**

Q3. Several years later, scientists repeated their study on this population. They found that the frequency of the recessive allele had decreased. A statistical test showed that the difference between the two frequencies of the **t** allele was significant at the P = 0.05 level. Use the terms **probability** and **chance** to help explain what this means.

_____

_____

_____

_____

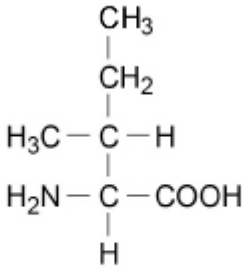
_____ **(2)**

Q4. Scientists measured the mean amino acid concentration in white wines made from grapes grown organically and white wines made from grapes that were not grown organically.

(a) Which test could the scientists have used to identify that there are amino acids in white wine?

_____ **(1)**

(b) All amino acids have the same general structure. The image below shows the structure of the amino acid isoleucine. Draw a box around the part of the molecule that would be the same in all amino acids.



Isoleucine

(1)

(c) Name the chemical element found in all amino acids that is **not** found in triglycerides.

_____ **(1)**

(d) The scientists used a statistical test to determine whether there was a significant difference in the amino acid concentration in the two types of white wine. They obtained a value for P of 0.04.
Name the statistical test the scientists used and give a reason for your answer.

Was the difference significant? Give a reason for your answer.

Name of statistical test _____

Reason for choice _____

_____

Explanation of test result _____

_____

_____

_____ **(3)**