5 (a) Clear statement of null hypothesis; 1
e.g. there is no difference in the masses of the magic and ordinary beans

5 (b) Standard error and (95% confidence limits); 1

5 (c) This test determines if there is a difference between the means of two samples; 1

5 (d) Test statistic correctly calculated; 1

Magic beans	**Ordinary beans**
Mean = 10.6	Mean = 8.2
SD = 0.48	SD = 0.66
SEM = 0.15	SEM = 0.21
Confidence limits	Confidence limits
10.3-10.9	7.78-8.62

5 (e) Correct statement concerning acceptance/rejection of null hypothesis;
Overlap of confidence limits used to determine if significant difference exists;
We must reject the null hypothesis as there is no overlap of confidence limits. There is more than a 95% probability that the difference in masses is not due to chance.
OR
We must reject the null hypothesis as there is no overlap of confidence limits. There is less than a 5% probability that the difference in masses is due to chance.

p37, 51, 54, 60, 91, 103

Cocaine

a) People used in the study have different masses/different blood volumes (see images below)

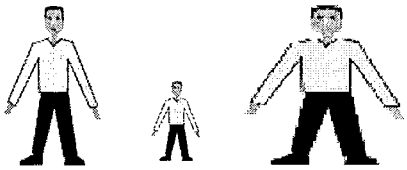
Discuss variation in body fluids between males and females: later having proportionally less fluid and more fat than males.

a) Continuous.

b) Grey area of the difference between 'occasional' and 'frequent'.

This is qualitative and would have been better to have a quantitative measure from the people's blood. Users may misinterpret the difference.

a) Different intervals used of the graphs. People may make errors in comparison because of this, if they do not read the independent scale.



boostprogress@hotmail.co.uk

Giardiasis

a) January to April, Difference: 1800 – 2100 = 3000. So 300/1800 = 16.7%
May to August, Difference 2500 – 5800 = 3300. So 3300/2500 = 132.0%
September to December, Difference 4400 – 2400 = -2000. So -2000/4400 = -45.5%

The last calculation is a negative value showing the decrease. The two others are positive.

b) Yes by by crudely representing the increases and decreases. However, No because they miss out some of the peaks and troughs in the data e.g. Feb/Mar/June.

c) Comparison of the trends from different months. Maybe useful for projections or predictions if assuming the trends continue into the future.

Giardiasis may always peak in summer months and fall in winter months – so these type of percentages may predict potential epidemics.

$$\frac{\text{Difference}}{\text{Original}} \times 100 =$$

boostprogress@hotmail.co.uk

CHD

a) Visually easy to determine proportions. Allows overlap of categories.

There will be individuals with two contributory factors and these are shown by the overlap.

b) Examples such as: age, gender, ethnicity, health...

c) Other contributing factors to CHD are not shown.

d) Examples: Smoking is one of the lowest contributing factors, less than lack of exercise. Smoking does solely cause CHD.

Allow humour in any marketing campaigns! "Lies, Damn Lies and Statistics...."?

Dairy Fat

a) Data is scattered but showing a general positive correlation.

b) Each country has a different sized population.

Students should think about comparing England and Wales or Scotland in terms of raw totals or per 1000 people and how this would skew the data.

c) Yes as a number of countries are included so yes and per 1000 people so suggesting it was drawn from a much larger samples of populations.

d) Names of countries to look at their distribution and compare them for GDP (developed or developing economies). Maybe how the data was gathered, particularly the very difficult task of measuring [mean] *Dairy fat eaten per day [per person]. Possibly some of the other control variables.*

PABA

a) These are not measured by S.I units.

b) They may be used for comparison purposes.

c) Graph A fast initial increase, then constant . Graph B slower initial increase (than Graph A) , then becoming constant later than Graph A.

The trends are less important that getting students to describe general and detailed trends, and ensure they cover all deviations in the trends: many questions like this have two or three obvious deviations. Encourage students to mark these deviations clearly on the graph .

Amylose

a) From the standard deviation of the sample mean based on the population mean.

Look at this example: Data of...

1 ,3, **5**, 7 and 9 has Standard Error of 1.62 so Mean of $\bar{5}$ ± 1.62. ± meaning plus or minus that 68% of the values should occur between 3.38 to 6.62. But.....

-100, -50, **5**, 50 and 100 has Standard Error of 31.64 so Mean of 5 ± 31.64. As above 68% of the values should now occur between -26.64 to 36.64.

The first set of numbers (data) are better represented by the mean value and hence the smaller Standard Error.

b) C as it has largest standard error.

c) Both A and C show a decrease as there is a clear gap between the standard error bars. B cannot be describes as showing a decrease as the standard error bars overlap (see below).

d) Rice grains were different sizes/had different masses so percentage allows comparison between them.