

Feedback – Lecture 02B: univariate data analysis

This quiz is based on the 4 video new video lectures:

- Univariate data analysis (2): Histograms
- Univariate data analysis (3): Basic terminology
- Univariate data analysis (4): Outliers, medians and MAD
- Univariate data analysis (5): The central limit theorem

Prior video content can be tested too.

You have 1 attempt for the quiz. Please read the instructions carefully. Please double check your answers before submitting.

Solutions will be released when the quiz closes, at 09:25am, Friday, 17 January 2015.

Question 1

What does the acronym CSV stand for (as used in this course)?

Your Answer	Score	Explanation
<input type="radio"/> Central standard value		
<input type="radio"/> Confidence standard value		
<input type="radio"/> Computerized system validation		
<input checked="" type="radio"/> Comma-separated values	✓ 1.00	
Total	1.00 / 1.00	

Question Explanation

Comma-separated values, as in a CSV file that contains many values separated by commas.

Question 2

Which of the following are equivalent, or nearly equivalent, words for the concept of the **average**, as

used in the videos this week.

Please check all that apply.

Your Answer	Score	Explanation
<input checked="" type="checkbox"/> mean	✓ 0.20	
<input checked="" type="checkbox"/> measure of location	✓ 0.20	
<input checked="" type="checkbox"/> measure of central tendency	✓ 0.20	
<input checked="" type="checkbox"/> dispersion	✗ 0.00	
<input checked="" type="checkbox"/> location	✓ 0.20	
Total	0.80 / 1.00	

Question 3

What is the median absolute deviation of this sequence of data (use the definition as shown in the class videos)

70, 121, 218, 162, 65, 198, 67, 12, 19, 104, 115

Report your answer correct to one decimal place.

You entered:

57.8

Your Answer	Score	Explanation
57.8	✓ 1.00	
Total	1.00 / 1.00	

Question Explanation

```
x <- c(70, 121, 218, 162, 65, 198, 67, 12, 19, 104, 115)
mad(x)
# returns the answer
```

[1] 57.8214

which is 57.8 when rounded to one decimal place. It is important for numerical-based answers to follow instructions carefully, as these are auto-graded.

Question 4

Regarding the central limit theorem: *(please check all options that correctly apply)*

Your Answer	Score	Explanation
<input checked="" type="checkbox"/> it can apply to the median as well.	✗ 0.00	Not true; the CLT is defined for taking the average.
<input checked="" type="checkbox"/> the sampled data could have come from an t-distribution .	✓ 0.25	Correct, that distribution is finite, so it would apply.
<input checked="" type="checkbox"/> the sampled data should be independent.	✓ 0.25	Yes, this is a key requirement.
<input checked="" type="checkbox"/> the sampled data used to calculate the average should be uniformly distributed.	✗ 0.00	No, the data can be from any distribution.
Total	0.50 / 1.00	

Question 5

The standard deviation is variance squared

Your Answer	Score	Explanation
<input checked="" type="radio"/> False	✓ 1.00	
<input type="radio"/> True		
Total	1.00 / 1.00	

Question 6

If this class wrote a really easy test then we would expect to get a histogram, that when plotted in the usual way, would be ...

(check all options that correctly apply)

Your Answer	Score	Explanation
<input checked="" type="checkbox"/> normally distributed, because of the central limit theorem.	✗ 0.00	Not so, the data will be "bunched up" near the upper bound of 100%.
<input checked="" type="checkbox"/> uniformly distributed.	✗ 0.00	Definitely not.
<input checked="" type="checkbox"/> skewed to the right.	✗ 0.00	It is tempting to think this, but the "skew" refers to the location of the tail.
<input checked="" type="checkbox"/> skewed to the left.	✓ 0.25	Yes, this is correct.
Total	0.25 / 1.00	

Question Explanation

We would not get a normal distribution, because the upper limit of the test is 100, and many students will be in the bins that fall at that upper bound, with a tail going off to the left.

Question 7

You want to test a new paint for wood that your company is producing. If successful you can beat your competitor's product, which requires 2 layers of paint; your product only requires applying a single layer. You go to the hardware store, randomly buy a **2x4 wooden board** and bring it back to your lab.

You cut the board into 8 pieces and paint them. You send these 8 pieces to your quality group for analysis.

Select all options that correctly apply

Your Answer	Score	Explanation
<input checked="" type="checkbox"/> If you send the samples to the quality control group on 8 different days, spread apart in time, and don't tell where they are from you, then you will have 8 independent measurements.	✗ 0.00	
<input checked="" type="checkbox"/> You will get 8 independent measurements back from the quality	✗ 0.00	

control group.

These **are not** 8 independent measurements that you get back from the quality control group. ✔ 0.25

If you had painted the board first, and then cut it into 8 pieces, only then would you have 8 independent samples. ✘ 0.00

Total 0.25 /
1.00

Question Explanation

The key issue in this question is whether the samples are all independent. They are all cut from the same board, which is the same tree originally. Therefore, the samples have a relationship with each other.

If you truly wanted 8 independent samples to test your product, you would have to get boards from a variety of suppliers (e.g Home Depot, Rona, Lowes, etc) and buy them on different days, so they are from different batches.

As said in the videos, you will violate the assumption of independence in engineering work at some point in your analysis. We have to be mindful of it.

Question 8

Based on the [2006 Census of Population in Canada](#) if 65% of the population had responded they speak a language other than English, then

Your Answer	Score	Explanation
-------------	-------	-------------

the value of 65% is, practically speaking, a parameter. ✔ 1.00

the value of 65% is a statistic.

the census is biased towards English speakers.

Total 1.00 / 1.00

Question Explanation

Reading the link and the information posted there, it is apparent that the Census is mandatory. Therefore a substantially large number of people will answer the question, and it can be deemed, from a practical perspective, to be from the population.

