Feedback — Lecture 02A: data visualization and introduction to univariate data analysis

This quiz is based on the 2 video new video lectures:

- Visualization (4): scatter plots
- Univariate data analysis (1): Introduction

Prior video content can be tested too.

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

Solutions will be released when the quiz closes, at 09:25am, Tuesday, 13 January 2015.

Question 1

Have you installed the R software and RStudio on your laptop?

If not, please follow the 9 step tutorial (link will open in a new tab) from this website. We will use the software in Tuesday's class. You must click "Yes" to get full grade for this question.

Your Answer		Score	Explanation
○ No			
◉ Yes	✔	1.00	
Total		1.00 / 1.00	

Question 2

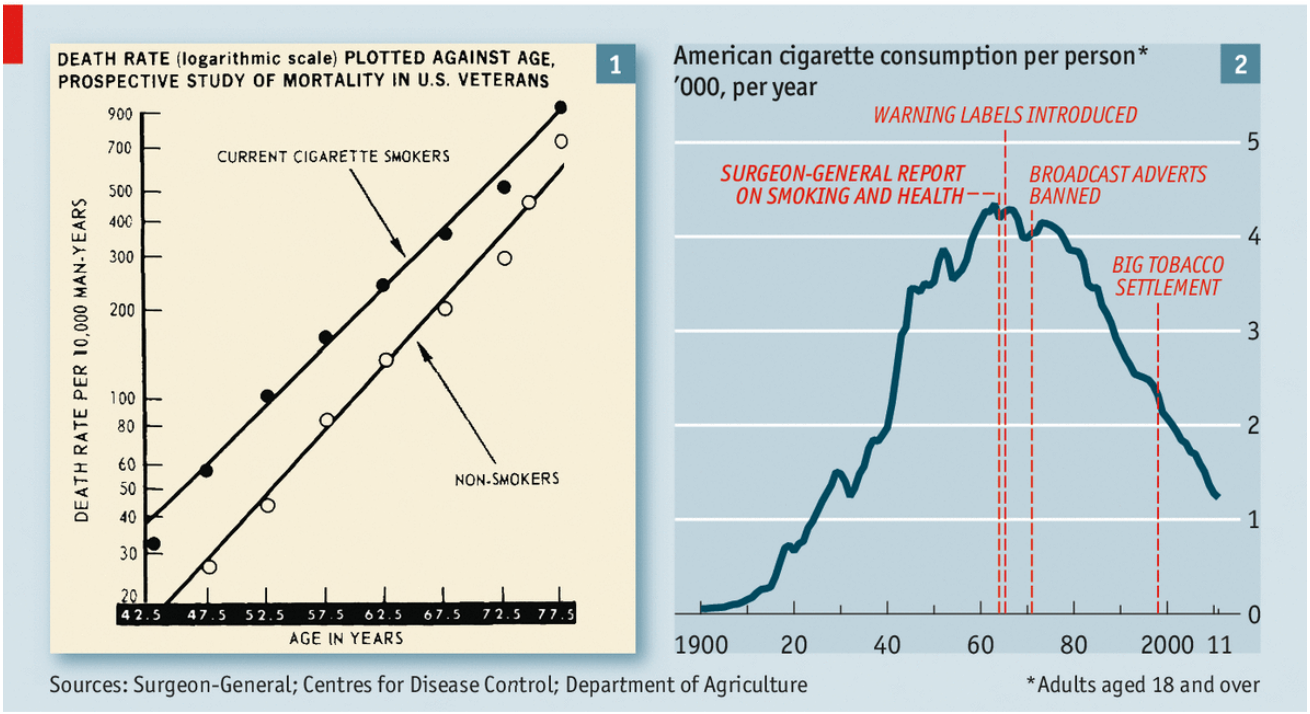
A scatter plot (*check all that correctly apply*):

Your Answer	Score	Explanation

☑ can have box plots added to the axes to help the user understand the spread of the data; box plots should therefore be shown.	✔	0.20	Yes, these are helpful, and a reasonable use of data ink.
☑ **should not** be used in an environment where people must make decisions from the data since they could be misread.	✖	0.00	Scatter plots are perfectly acceptable for decision making. For example, a certain region on a 2D scatter plot could be related to bad quality, allowing operators to stop the process, and fix the problem(s).
☑ requires cause-and-effect between variables being plotted.	✖	0.00	Certainly not; you can plot any two variables you like on a scatter plot. While cause-and-effect is implied, it does not mean it is requirement that it exist.
☑ can, with good use of marker size, colour, and animation, show up to 5 variables in a data set.	✔	0.20	Yes.
☑ wastes data ink if a histogram is added, because it shows redundant information; histograms should therefore not be shown.	✖	0.00	Not true: histograms help inform the user of the data density. For example, overlapping points on a scatter plot cannot be seen. A histogram will show a higher density of points.
Total		0.40 / 1.00	

Question 3

Study this plot, then *check all options that correctly apply*.



Source: The Economist.

Your Answer		Score	Explanation
☑ The plot on the left conveys a cause-and-effect message.	✔	0.50	Yes it does, but not in the same way as described in the video lectures. How does it convey cause-and-effect?
☑ The plot on the right is a histogram.	✖	0.00	Not quite. It is a time-series plot. Take a look at the title of the plot: it shows the time-based trend of cigarette consumption.
☑ The plot on the left is a scatter plot.	✔	0.50	Yes, that's right, with two series shown (one with open circles, the other with closed circles).
☑ The use of log-axes distorts the message of the plot.	✖	0.00	Not true; the message is still very apparent in the plot: at all ages ranges, the difference in death rate between smokers and non-smokers is significantly higher. A log-plot would still have shown the same message.
Total		1.00 / 2.00	

Question 4

Having high variability in the product produced from your process ... (*check all that correctly*

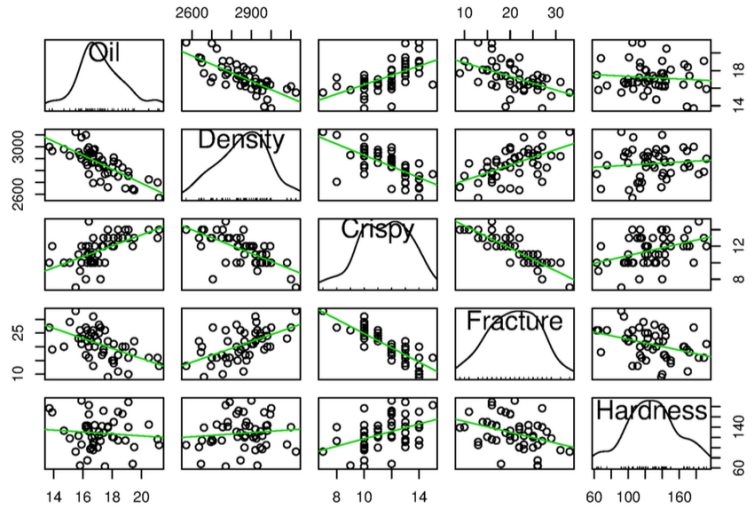
apply)

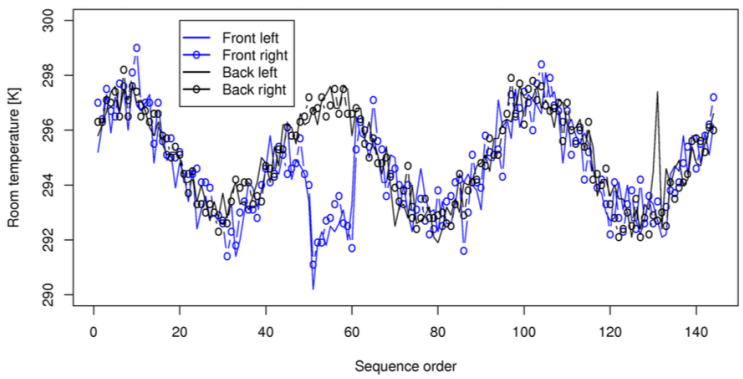
Your Answer		Score	Explanation
☑ will get you a bad reputation among your customers.	✔	0.20	Yes, you will quickly lose reputation if you make variable products.
☑ is only acceptable when that product is going into the next step on your flowsheet, and not to the customer.	✖	0.00	Not true; the next step in your process has to undo this variation and that will cost you energy, resources and/or money. This thinking should definitely be avoided.
☑ can lead to a situation where your product is totally unusable for the customer; so this is not a desirable condition.	✔	0.20	Yes, that's correct.
☑ might cost your customer money, because they will have to rework your product to make it suitable for their process.	✔	0.20	True.
☑ is manageable, since a generally acceptable quality control strategy is to throw out badly produced product before shipping it to the customer.	✖	0.00	This is not a good strategy in general. 1. It costs money to inspect every product. 2. What will you do with all the product that you throw out?
Total		0.60 / 1.00	

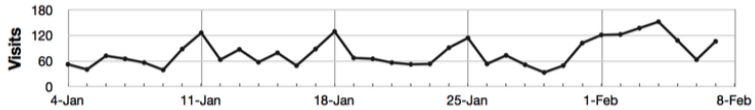
Question 5

Which one of the following shows scatter plots?

Your Answer		Score	Explanation
◉	✔	1.00	This is a scatterplot matrix



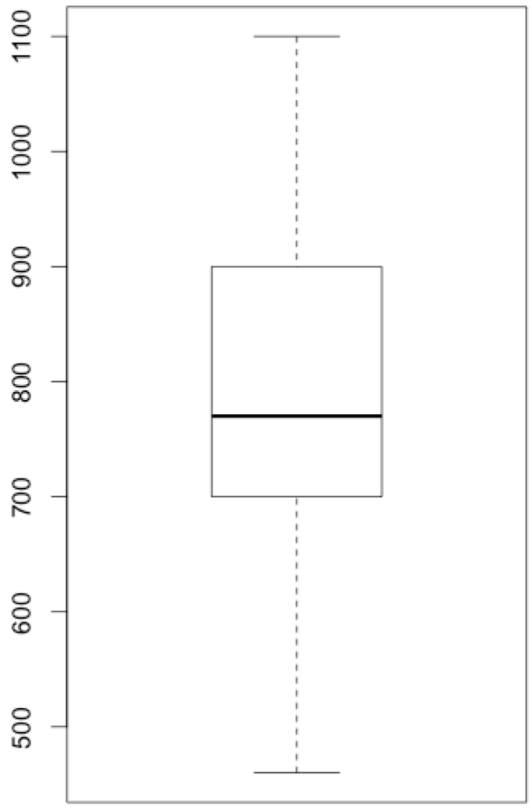
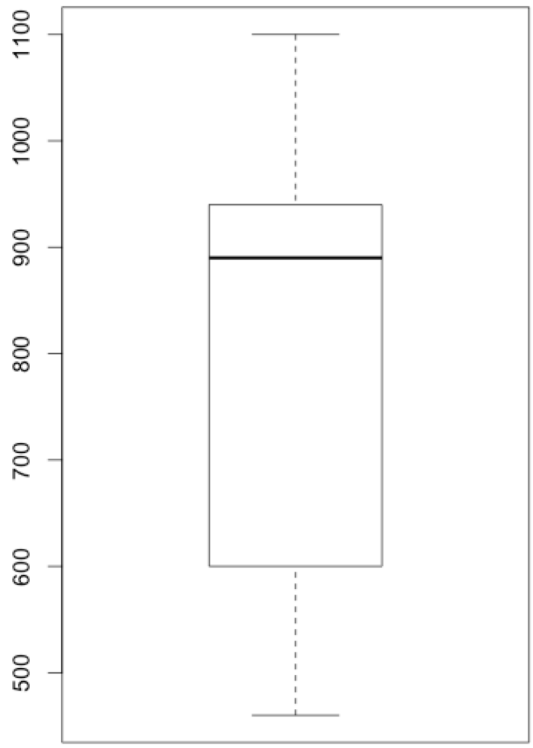


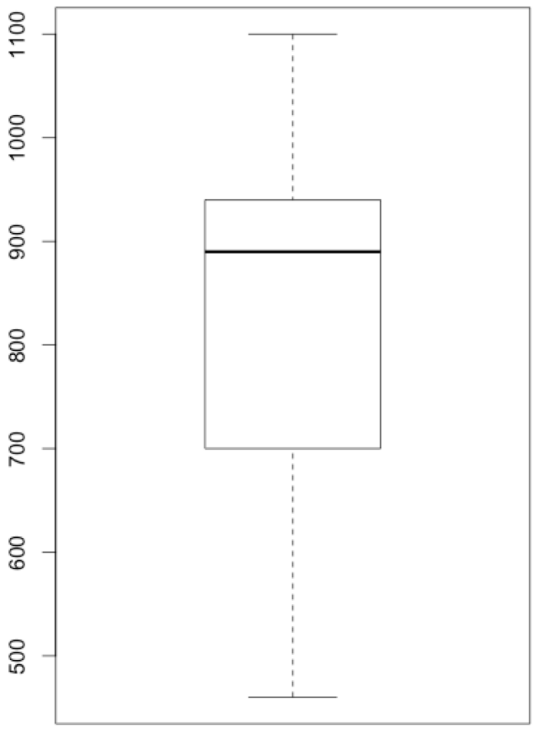


Total	1.00 / 1.00

Question 6

Use the "Salt" data from the prior class (01B) and draw a box plot from those data. Which of the following is a closest match to the true box plot?

Your Answer	Score	Explanation
		
		
	✔ 1.00	This is the box plot related to the salt data.



Total	1.00 / 1.00

Question Explanation

```
salt <- c(460, 520, 580, 700, 760, 770, 890, 910, 920, 940, 960, 1060, 1100)
boxplot(salt)
quantile(salt)
```