

1. Figure 1 is a box and whiskers plot of the number of vehicles (in 1000s) produced in the year 2000 by the 29 largest vehicle producing countries. (source: The World Almanac and Book of Facts, 2002). The five number summary values are listed on the plot. For each of the following statements, tell whether it is (i) supported by either the plot or the statements made above, (ii) contradicted by either the plot or the statements made above, or (iii) undecidable from the plot and statements made above. Justify your answers.

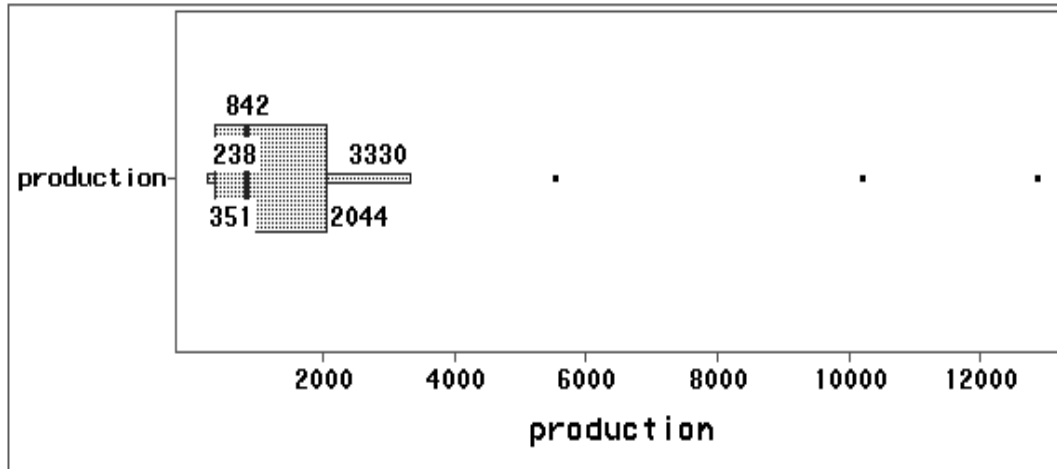


Figure 1: *Box and whiskers plot of vehicle production for the year 2000.*

- (a) (10 points) The smallest data value is 238.

ANS: *This is supported by the plot. Since (i) 238 is the lower adjacent value, it is one of the data values, and since there are no "outliers" indicated below the lower adjacent value, 238 must be the smallest data value.*

- (b) (10 points) One of the data values equals 842.

ANS: *This is supported by the plot and the statement that there are 29 data values. Since the number of observations is odd, the median (which equals 842) is the middle (i.e., 15th smallest) data value.*

- (c) (10 points) The the mean and standard deviation are more appropriate measures of location and spread than are median and interquartile range.

ANS: *This is contradicted by the plot, which indicates the distribution is skewed right. If the distribution is unimodal, this means the median and interquartile range are more appropriate measures than the mean and standard deviation. If the distribution is multimodal, neither set of measures is appropriate.*

- (d) (10 points) There is at least one data value greater than 2044 and less than 3330.

ANS: *This is undecidable from the plot and statements made above. In the data set, there are three values greater than 2044 and less than 3330. But if each of these three values is changed to 2044, the plot remains the same.*

2. Faster magnetic resonance imaging (MRI) scanning has made MRI a potential cost-effective replacement for radiographic imaging for patients with low back pain. However, whether rapid MRI scanning results in better patient outcomes than radiographic evaluation or a cost-effective alternative is unknown. An article in the June 4, 2003 issue of the Journal of the American Medical Association describes a study to determine the clinical and economic consequences of replacing spine radiographs with rapid MRI for primary care patients. The study looked at 380 patients aged 18 years or older whose primary physicians had ordered that their low back pain be evaluated.

Suppose the 380 patients were assigned by the investigators to have rapid MRI or radiographic evaluation.

(a) (10 points) What kind of study is this? Be as specific as you can. Justify your response.

ANS: *It is a controlled experiment, since treatments (radiographic or rapid MRI imaging) are imposed on experimental units (patients) to observe a response (clinical and economic outcomes).*

(b) (10 points) How should the subjects in the study be chosen so that the results of the study are to be applicable to all patients aged 18 years or older whose primary physicians order that their low back pain be evaluated?

ANS: *They should be randomly sampled from all such patients.*

(c) (10 points) How should the patients be assigned to the rapid MRI or radiographic evaluation groups? Justify your answer.

ANS: *They should be assigned randomly to prevent unsuspected biases.*

(d) (10 points) Give one example of how blocking could be used to improve the study results. Tell how your choice would improve the results.

ANS: *They could be blocked by severity of symptoms. This would reduce variation due to disease severity.*

(e) (10 points) If the results show a difference in patient outcomes between the rapid MRI and radiograph groups, can we conclude that the difference is due to which type of imaging was used? Justify your answer.

ANS: *Yes. Since this is a controlled experiment, we can conclude a causal effect for the treatment.*

3. (10 points) Figure 2 is a needle plot of the lifetimes, in days, of 70 switching relays in continuous use from their installation in 1970. Bob claims the plot demonstrates the stationarity of the process that produced them. Do you agree? Justify your answer.

ANS: *The stationarity of the process cannot be judged from these data, since there is no information on how these lifetimes are related to time of production. Even if such information were available, a needle plot does not incorporate it, and so cannot be used to judge stationarity.*

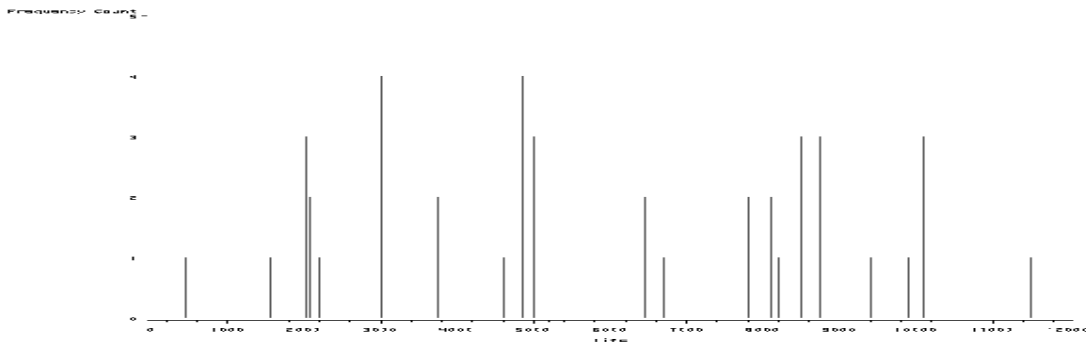


Figure 2: Needle plot of switching relay lifetimes.