M203J Practice Quiz 2 October 2, 2009 Covering Chapters 8 and 9

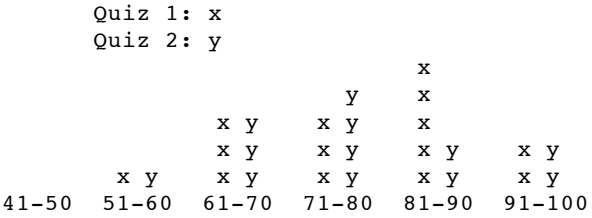
Instructions: Answer each question, and when required explain your answer. Your explanation must be clear and complete. You may refer to your book, your notes and your homework papers.

[1] A teacher wants to make a double stem and leaf plot using the scores on quizzes 1 and 2. Suppose the scores on quiz 1 are 77, 93, 75, 76, 84, 69, 65, 84, 83, 95, 85, 82, 65 and 55, and suppose the scores on quiz 2 are 55, 78, 65, 73, 98, 93, 82, 85, 65, 67, 40, 75 and 80. Construct the double stem and leaf plot for these two data sets.

Answer:

Quiz 1		Quiz 2
3 5 9		3 8
2 3 4 4 5 8		0 2 5
5 6 7 7		3 5 8
5 5 9 6		5 5 7
5 5		
4 0		

[2] Make a single comparison histogram using the two data sets given in Problem 1. Use bins of length 10 with each bin ending on an even multiple of 10 (so 31-40, 41-50, etc.).



[3] For each of the following kinds of data, choose an appropriate method for graphing the data and an inappropriate method. Indicate which is which, and explain what is inappropriate about the one which is not appropriate. Answers can vary.

(a) Annual profits of a company over a period of years.

Answer: A line graph would be appropriate, in order to emphasize the trend over time. A frequency histogram would not be appropriate; a frequency histogram would make sense in a situation in which we showed the number of years the profits were in various ranges.

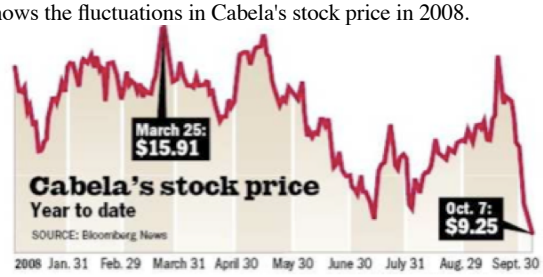
(b) The percentages of families below the poverty level in five different Nebraska cities.

Answer: A bar chart would make sense here, but not a pie chart since the percentages do not describe parts of a whole (they could add up to more than 100%, for example).

(c) The percentages of Nebraska families whose annual incomes are in the following ranges: up to $50,000, above $50,000 but below $100,000, and above $100,000.

Answer: A pie chart would make sense here, since the percentages do describe parts of a whole (they have to add up to 100% in this example). A dot plot would not make much sense here, since there would only be three dots.

[4] The following line graph shows the fluctuations in Cabela's stock price in 2008.



(a) Discuss what is misleading about this graph. Answers can vary.

Answer: The cropping makes the March 25 data point look much better compared to October 7 than it really is.

(b) Explain how you would redraw this graph so that it is not misleading.

Answer: Start the vertical scale at 0, and include labels on the vertical scale.

[5] Explain how to select a simple random sample of 4 elements from the whole numbers running from 1 to 100, using the table on page 570 (a copy is included with the quiz). What sample do you get? Explain in enough detail that I can verify that your sample is the one you should have gotten.

Answer: Randomly pick a starting entry in the table, say the entry in row 2 column 4 (which is 64569). Then read down and pick the

last two digits of each entry, skipping an entry if it gives a number already chosen. (If the two digits are 00 then that counts as 100.) Here is the simple random sample I get: 69, 7, 38, 8.

[6] Explain how to select a 30% independent sample from the whole numbers running from 1 to 10, using the table on page 570. What sample do you get? Explain in enough detail that I can verify that your sample is the one you should have gotten.

Answer: Randomly pick a starting entry in the table, say, as luck would have it, again the entry in row 2 column 4 (which is 64569). Then read down that column, counting out loud from 1 to 10 as you go. Every time the last two digits of the entry gives a number between 1 and 30 inclusive, the number you counted out loud is selected. The results are given in the following table, where the first column gives the count from 1 to 10, the second column gives the corresponding table entry, the third column gives its last two digits and the fourth column indicates whether we select the number in the first column or not:

1	64569	69	do not select 1
2	17707	07	do select 2
3	60638	38	do not select 3
4	93608	08	do select 4
5	78545	45	do not select 5
6	39445	45	do not select 6
7	50784	84	do not select 7
8	33358	58	do not select 8
9	36246	46	do not select 9
10	17068	68	do not select 10

Our 30% independent sample is thus {2, 4}.

[7] A farmer who wants to assess the level of pest infestation in her orchard is considering several different sampling methods. For each of (a) through (d), indicate what sampling technique it corresponds to (choose your answers from among independent sampling, simple random sampling, systematic sampling or cluster sampling).

(a) Randomly select a tree and then move row by row through the orchard, selecting every 15th tree for inspection.
Answer: systematic sampling

(b) For each tree flip a coin; if the coin lands heads, inspect the tree.
Answer: independent sampling

(c) Number the rows of trees, randomly select a certain number of rows, and inspect every tree in those rows.
Answer: cluster sampling

(d) Each tree is numbered and a random sample of those numbers is chosen to select the trees to be inspected.
Answer: simple random sampling

[8] Consider the data 7, 8, 8, 1, 5, 6, 6, 9, 11, 20.

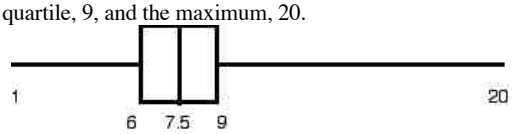
(a) Find the mean of this data.
$(7 + 8 + 8 + 1 + 5 + 6 + 6 + 9 + 11 + 20)/10 = 81/10 = 8.1$

(b) Find the median of this data.
First put the data in order: 1, 5, 6, 6, 7, 8, 8, 9, 11, 20. The median is $(7+8)/2 = 7.5$

(c) Find the mode(s) of this data.
6 and 8 are the modes (i.e., occur the most)

(d) Find the range of this data.
the range is $20 - 1 = 19$.

(e) Create and label a box and whisker plot of this data.
The five number summary is the minimum, 1, the first quartile, 6, the median, 7.5, the third quartile, 9, and the maximum, 20.



(f) Find the sample standard deviation of this data.
the variance is
$[(1-8.1)^2+(5-8.1)^2+(6-8.1)^2+(6-8.1)^2+(7-8.1)^2+(8-8.1)^2+(8-8.1)^2+(9-8.1)^2+(11-8.1)^2+(20-8.1)^2]/(10-1) =$
24.544, so the sample standard deviation is the square root of this, or about 4.95.