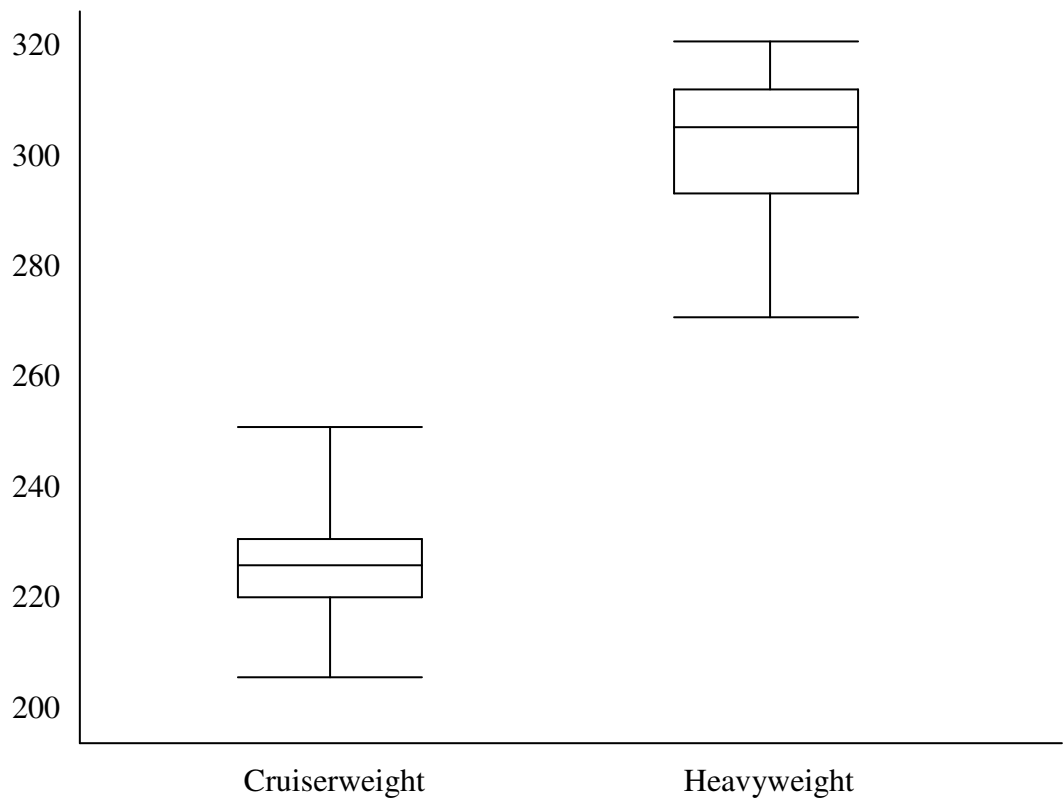


STAT 2220 – Sample Midterm Test 1A

Part A

The first **two** questions (**1 and 2**) refer to the following:

The sport of boxing divides its athletes into different weight classes in order to make the competition fairer. The side-by-side basic (quantile) boxplots shown below display the weights (in pounds) of a random sample of 50 Cruiserweight boxers and 49 Heavyweight boxers.



The five number summaries for the two weight classes are shown below:

	Minimum	Q1	Median	Q3	Maximum
Cruiserweight	204	220	226	230	250
Heavyweight	270	295	304	312	320

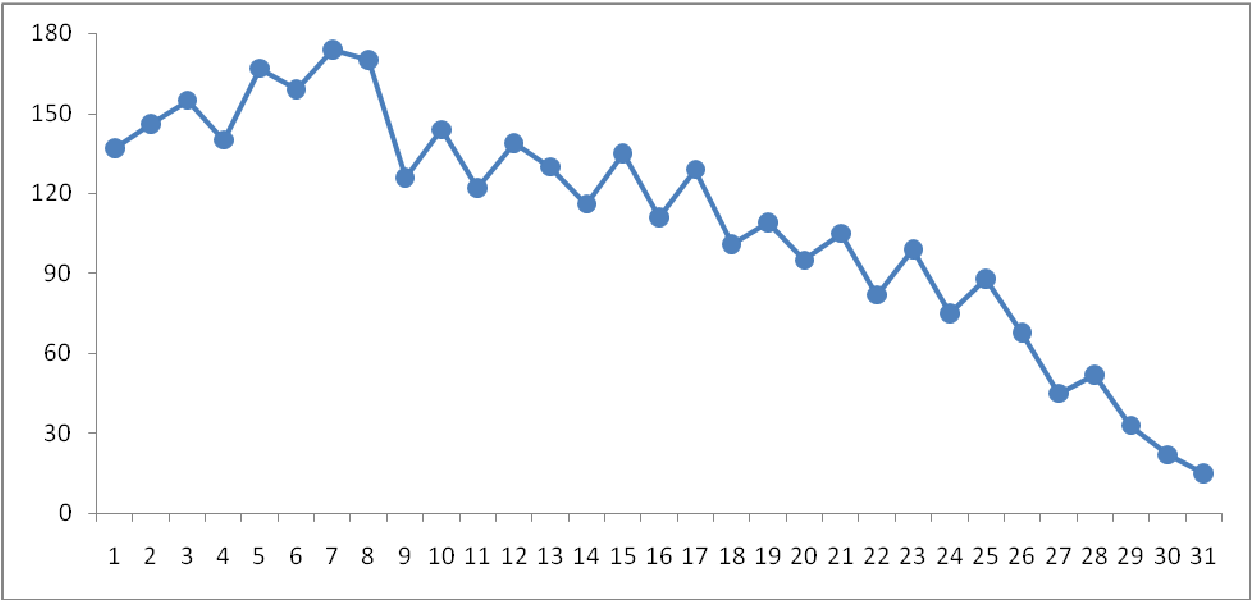
1. Which of the following statements is/are true?

I. Any Cruiserweight with a weight less than 205 pounds or greater than 245 pounds would be considered an outlier.
II. There are about 37 Cruiserweights in the sample who weigh more than 220 pounds.
III. The distribution of weights for the Heavyweights is skewed to the right.

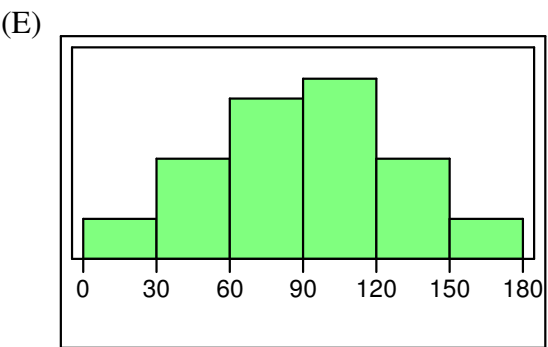
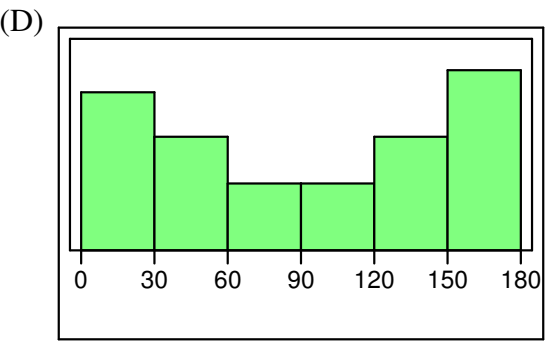
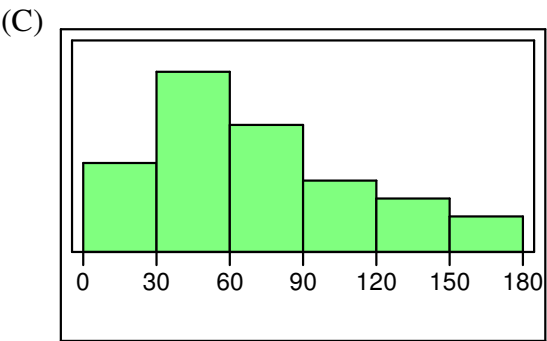
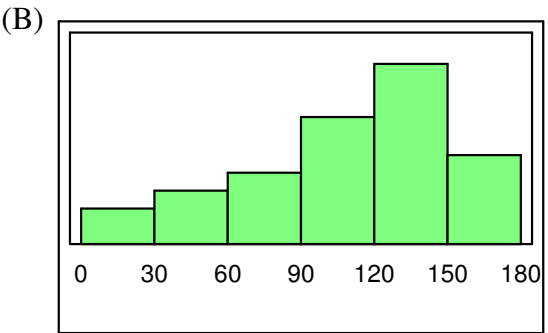
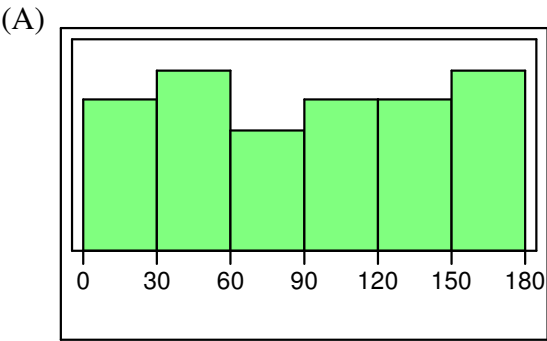
(A) I only (B) III only (C) I and II only (D) II and III only (E) I, II and III
2. What is the median weight of all 99 boxers in the sample (Cruiserweights and Heavyweights) combined?

(A) 250 (B) 265 (C) 270 (D) 285 (E) 260

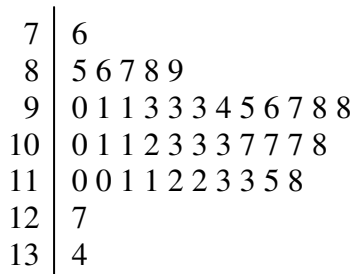
3. A university student records the time (in minutes) she spent studying each day for a month. The data are displayed in the time plot shown below:



Which of the following is a histogram for this data set?

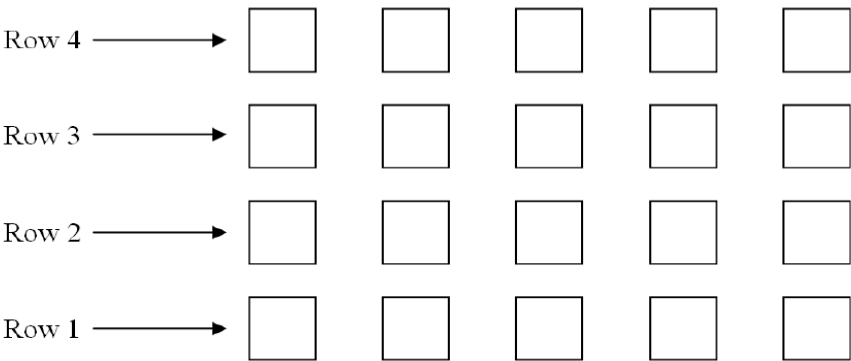


4. The stemplot below displays the number of points scored by the Toronto Raptors for each of their 41 homes games in the 2008/2009 NBA season.



What is the interquartile range for the number of points scored by the Raptors in these games?

- (A) 14.5 (B) 16 (C) 16.5 (D) 17 (E) 17.5
5. A small graduate class of three students writes a math test. The student who finished writing the fastest got the highest score in the class. The student who finished second got the second highest score, and the student who took the longest to write the test got the lowest score. Let X = the time it takes for a student to write the test and let Y = the student's test score. What can be said about the correlation r between X and Y for this class?
- (A) There is a perfect negative linear relationship between X and Y , and so $r = -1$.
 (B) The correlation between X and Y is negative, but not necessarily equal to -1 .
 (C) There is no linear relationship between X and Y , and so $r = 0$.
 (D) There is a perfect positive linear relationship between X and Y , and so $r = 1$.
 (E) The correlation between X and Y is positive, but not necessarily equal to 1.
6. Which of the following statements about the least squares regression line is (are) **true**?
- (I) The slope of the least squares regression line always has the same sign as the correlation.
 (II) The least squares regression line is the line that minimizes the sum of residuals.
 (III) The least squares regression line is the line that maximizes the value of the correlation.
- (A) I only (B) II only (C) I and II only (D) I and III only (E) I, II and III
7. The twenty students in Professor Smith's graduate class have handed in their term projects and oral presentations of the projects will begin today. There will only be time for five students to present their reports today. The seating plan of the class is shown below, where each square represents a desk occupied by a student:

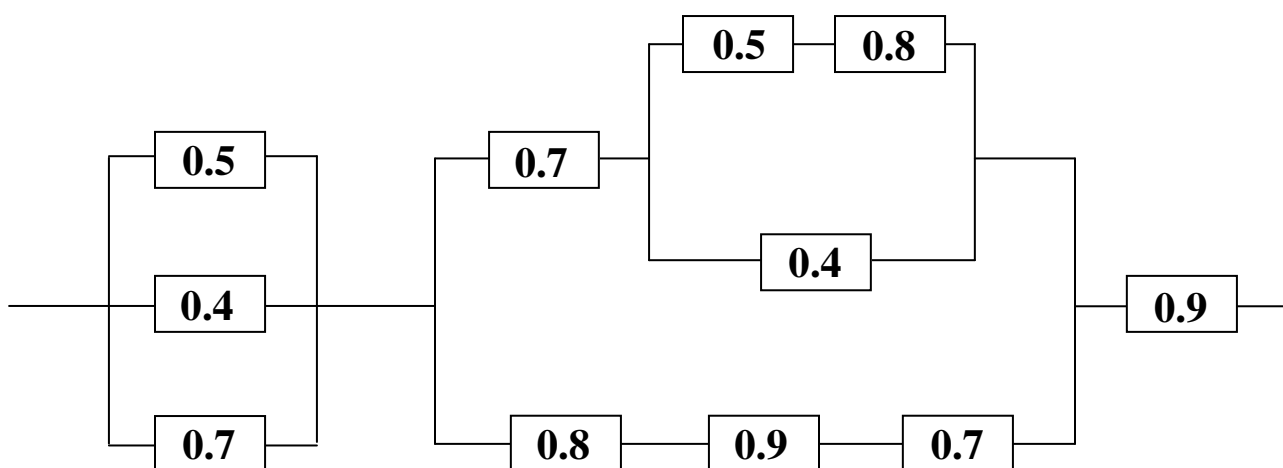


To select which students will present today, Professor Smith writes the numbers 1, 2, 3 and 4 on four pieces of paper and places them in a hat. She then selects one number randomly from the hat, and the five students in the corresponding row will present their reports today. The resulting sample is not a simple random sample because:

- (A) the students were not assigned to their desks randomly.
 (B) not all students had the same chance of being chosen.
 (C) not all possible combinations of five students had the same chance of being chosen.
 (D) the rows were not numbered randomly.
 (E) a simple random sample would not select five students all from the same row.

8. In a particular election, 40% of voters voted for the NDP, 35% voted for the Liberals and 25% voted for the Conservatives. If we take a random sample of two voters, what is the probability that they voted for different parties?
- (A) 0.345 (B) 0.465 (C) 0.575 (D) 0.655 (E) 0.715
9. Event A has a probability of 0.2 of occurring. Event B has a probability of 0.5 of occurring. If A and B are mutually exclusive, then $P(A \text{ and } B)$ is equal to:
- (A) 0.0 (B) 0.1 (C) 0.2 (D) 0.6 (E) 0.7
10. A hockey player scores a goal in 43% of his games. His team wins 65% of their games. In 31% of the team's games, the player scores a goal and the team wins. What is the probability that the team wins if the player **does not** score a goal?
- (A) 0.3595 (B) 0.4625 (C) 0.4985 (D) 0.5965 (E) 0.6345

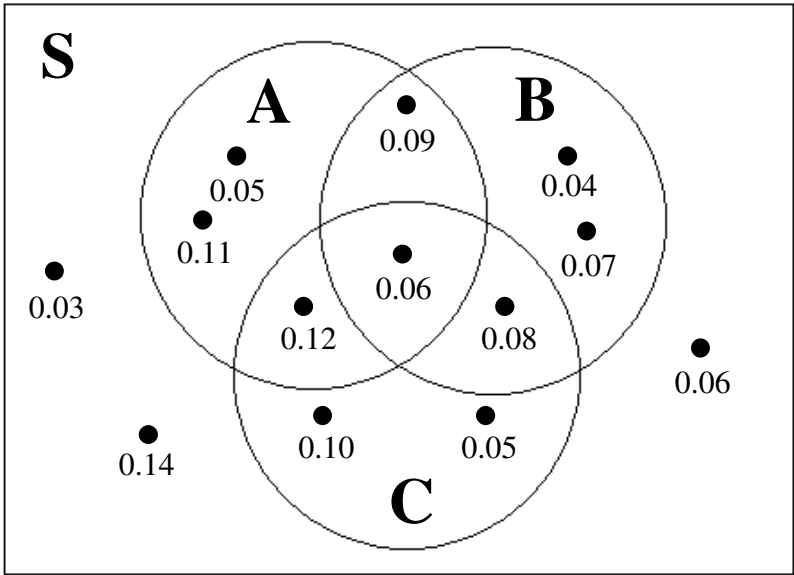
11. Consider the following system, where the value in each box represents the reliability of that component:



What is the reliability of this system?

- (A) 0.5948 (B) 0.6327 (C) 0.6715 (D) 0.7032 (E) 0.7446
12. We have a small deck of ten cards. Five of the cards are red, three are blue and two are green. We randomly select four cards from the deck with replacement. That is, after we select a card and record the colour, we replace the card in the deck and thoroughly shuffle them before we select another card.
- Let A be the event that the first selected card is the only red card in our four selections. Which of the following events is mutually exclusive (disjoint) from the event A?
- (A) Second card selected is blue.
 (B) No green cards are selected.
 (C) Third selected card is the only green.
 (D) Same number of red and blue cards are selected.
 (E) Same number of blue and green cards are selected.
13. You roll three fair dice. What is the probability that the total number of points facing up is at least 17?
- (A) 1/216 (B) 2/216 (C) 3/216 (D) 4/216 (E) 5/216

14. Consider the sample space and events shown in the Venn diagram below:



What is $P((A' \cap C') \cup (B \cap C))$?

- (A) 0.53 (B) 0.62 (C) 0.48 (D) 0.39 (E) 0.74

15. In a game of poker, you are dealt five cards from a deck of 52 cards. What is the probability that you get a flush (five cards of the same suit)?

- (A) 0.0005 (B) 0.0010 (C) 0.0020 (D) 0.0030 (E) 0.0040

Part B

1.
 - (a) If a variable Y can be written as a linear function of a variable X , i.e. if $y = a + bx$, then show that $\bar{y} = a + b\bar{x}$ and $s_y = |b|s_x$.
 - (b) A phone company sells prepaid long distance phone cards. For each call, the customer is charged a base fee of \$1.49, plus \$0.25 per minute. The mean time for a sample of calls is calculated to be 12 minutes, and the standard deviation is 7 minutes. What are the mean and standard deviation of charges for this sample of calls?
2. The time Y it takes (in seconds) for a chemical reaction to take place depends on (among other things) the temperature X (in $^{\circ}\text{C}$) to which the chemical solution is exposed. Ten trials are conducted at various temperatures and the reaction time is recorded for each trial. A scatterplot reveals that a linear relationship is a reasonable assumption. The least squares regression line is calculated to be $\hat{y} = 87.2 - 0.5x$. It is also observed that 90% of the variation in time is accounted for by its regression on temperature.
 - (a) What is the value of the correlation between temperature and time?
 - (b) Interpret the meaning of the slope of the least squares regression line.
 - (c) For one of the trials, a temperature of 70°C was used, and a reaction time of 54 seconds was observed. What is the value of the residual for this trial?
3. Gordon and his wife Michelle are avid gardeners who take pride in maintaining a healthy green lawn. Over the course of the last summer, Gordon was in charge of maintaining the grass in the back yard and Michelle cared for the lawn in the front yard. Gordon used one brand of fertilizer (Brand A) and watered the grass once a week. Michelle used a different fertilizer (Brand B) and watered the grass twice a week. At the end of the summer, it was apparent that the grass in the front yard looked much greener and healthier than the grass in the back yard. Michelle argues that this is clear evidence that Fertilizer B is better than Fertilizer A and that watering the grass twice a week is better than watering it just once a week.

There are many flaws with Michelle's reasoning. One problem with this conclusion is that the front yard gets a lot more sunlight than the back yard.

- (a) Explain some other problems with Michelle's conclusion.
- (b) You are asked to help Gordon and Michelle conduct an experiment together next summer to determine the optimum growing conditions for their grass. (The couple is willing to sacrifice the appearance of their lawn for one summer in order to conduct a proper experiment).
 - (i) What type of experimental design would you propose?
 - (ii) Identify the following for this experiment:

Experimental units:

Factors:

Factor levels:

Treatments:

Response variable:

- (c) Explain to the couple how the experiment should be conducted. Be sure to explain the role of randomization in your experiment.

4. There are three games scheduled in the National Hockey League one night. The games are shown below, together with the probabilities that each team will win their respective game, as determined by odds-makers. Note that tied games are not possible.

Game 1: Montreal Canadiens (0.4) vs. Boston Bruins (0.6)

Game 2: Toronto Maple Leafs (0.2) vs. Detroit Red Wings (0.8)

Game 3: Vancouver Canucks (0.7) vs. Los Angeles Kings (0.3)

- (a) The outcome of interest is the set of winners of the three games. List the complete sample space of possible outcomes, and calculate the probability of each outcome.
- (b) What is the probability that Boston or Los Angeles wins?
- (c) What is the probability that two Canadian teams win?