MA 111 Review for Exam 6

Exam 6 (given in class on Thursday, December 3) will cover Unit 6: Descriptive Statistics. Use Chapter 14 in the textbook as a study tool.

You should be familiar with the following key ideas:

- Understand how to make and/or interpret the following graphical representations of data:
  - frequency table
  - bar graph
  - histogram
  - pictogram
  - pie chart
- Remember the differences between continuous and discrete variables, and between numerical and categorical variables. Understand which of the above graph types are best suited for these different types of variables.
- Be able to critique a graph. What are some common ways that graphs can be made so they mislead the viewer?
- Know the definition of the mean (or average) and how to calculate it for a given data set.
- Know the definition of the pth percentile and how to find it for a data set.
- Understand the meanings of median, first quartile, and third quartile:
  - The first quartile Q_1 is the 25th percentile.
  - The median M is the 50th percentile.
  - The third quartile Q_3 is the 75th percentile.
- Be able to make and/or interpret a box-and-whisker plot.
- Know the definition of the range and the interquartile range.

Practice Problems

Use the odd exercises on pages 545-554 of the text to supplement these. You can check your answers in the back of the book.

1. Suppose you are given the following data table:

Favorite Colors of a Third Grade Class	
red	3
orange	1
yellow	4
green	3
blue	7
purple	2

   (a) Choose a graphical representation and use it to represent the data in the table.

   (b) Other than the graph you just chose, name another type of graph that would be appropriate for representing this data.

   (c) Name a type of graph that would be inappropriate for representing this type of data, and explain why.

2. Suppose you want to make a pie chart of the age of students in MA111. If 39% of the students are 19 years old, how big would the corresponding slice of the pie be? Calculate the size of the central angle of the slice.

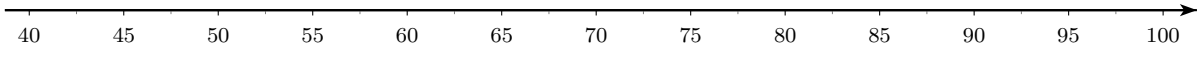
3. Consider the following two data sets of exam scores for two different sections of a course:

$$\begin{aligned} \text{Class A} &= \{40, 52, 65, 66, 68, 71, 72, 73, 77, 77, 78, 78, 80, 81, 81, 83, 86, 87, 92, 99\} \\ \text{Class B} &= \{55, 65, 66, 68, 72, 72, 73, 74, 76, 78, 79, 79, 80, 81, 85, 85, 85, 87, 90, 95\} \end{aligned}$$

   (a) Find the median, first quartile Q_1 and third quartile Q_3 for each class.

   (b) The professor wants to recommend the top 10% of each class for a scholarship. How many students will she recommend? (Remember, she will recommend students from *each* class).

   (c) Create box and whisker plots for both classes on a single axis.



4. Mike's average on the first five exams is 88. What must he earn on the next exam to raise his exam average to 90?

5. A professor is teaching two sections of MA111. Section 1 has 35 students, and Section 2 has 28 students. On the last exam, the average score for Section 1 was 73, while the average score for Section 2 was 78.

   For all of the students combined, find the average score for the exam.

6. Look at the hand-out of graphs I distributed at the beginning of this unit. Be able to identify the ways that each graph is misleading, and suggest a way to improve the graphs.

7. Consider the following data set:

$$\{7, 59, 25, 27, 64, 70, 68, 11, 45, 5, 17, 45, 52, 21, 26\}$$

   (a) Find the 30th percentile of the data set.

   (b) Find the 60th percentile of the data set.

   (c) Find the 95th percentile of the data set.

   (d) Find the 50th percentile of the data set.

8. Consider the data set $\{2, 18, 19, 22, 24, 24, 25, 78\}$.

   (a) Find the range.

   (b) Find the interquartile range.

   (c) Why might we consider the interquartile range a better measure of the spread of this particular set?

9. Give an example of a data set with $N = 5$ with the median less than the mean.

10. Try questions 45 and 46 on pages 549-550 of the textbook. Be able to interpret box-and-whisker plots to answer similar questions.

    Question 46 has two typos. The question should read:

    (a) Fill in the blank: Of the 612 engineering graduates, at most _____ had a starting salary greater than **$45,000**.

    (b) Fill in the blank: If there were 240 agriculture graduates with starting salaries of **$35,000** or less, the total number of agriculture graduates is approximately _____.

11. In 2006, the median SAT score was $d_{756,155}$, where $\{d_1, d_2, \ldots, d_N\}$ denotes the data set of all SAT scores ordered from lowest to highest. Determine the number of students N who took the SAT in 2006.