

Name Solution

Section \_\_\_\_\_

STA 291 Midterm 1  
February 17th, 2009

	Pb.	Points
1	/	4
2	/	4
3	/	4
4	/	4
5	/	4
total	/	20
6 (a)	/	5
6 (b)	/	5
6 (c)	/	5
6 (d)	/	5
total	/	20
7 (a)	/	12
7 (b)	/	18
7 (c)	/	5
7 (d)	/	5
8	/	10
9	/	10
total	/	100

Question 1 through Question 5 are multiple choice questions. You have to choose one of the items.

Question 1:

A company has developed a new light bulbs, but the average lifetime is unknown. In order to estimate this average, a sample of 1000 light bulbs is tested and the average lifetime of this sample is found to be 225 hours. The 225 hours is the value of a:

- (a) sample
- (b) population
- (c) statistics
- (d) parameter

Question 2:

A researcher at Michigan State University (MSU) wants to estimate the average number of credits earned by students last semester at MSU. She randomly selects 500 students from last semester and finds that they averaged 14.85 credits per student. The population of interest to the researcher is:

- (a) all MSU students.
- (b) all college students.
- (c) all MSU students enrolled last semester.
- (d) the 500 MSU students selected at random.

Question 3:

In a histogram, the proportion of the total area which must be to the left of the median is:

- (a) exactly 0.50.
- (b) less than 0.50 if the distribution is negatively skewed.
- (c) more than 0.50 if the distribution is positively skewed.
- (d) unknown.

Questions 4:

In a right-skewed distribution:

- (a) the median equals the mean.
- (b) the median is less than the mean.
- (c) the median is larger than the mean.
- (d) the mean, median, and mode are equal.

Question 5:

The best way to collect data for a poll is:

- (a) to use observational data
- (b) to use simple random sampling
- (c) to use volunteer sampling
- (d) to use convenience sampling

Question 6:

For each of the following examples of data, determine the type (the scale of the variable):

- (a) the number of miles joggers run per week

Interval

- (b) the starting salaries of graduates of MBA programs

Interval

- (c) the months in which a firm's employees choose to take their vacations

Nominal

- (d) the final letter grades received by students in a statistics course

Ordinal

**Question 7:**

The midterm test for a statistics course has a time limit of 1 hour. However, like most statistics exams this one was quite easy. To assess how easy, the professor recorded the amount of time taken by a sample of nine students to hand in their test papers. The times (rounded to the nearest minute) are

33 29 45 60 42 19 52 38 36

(a) Compute the mean, median, and mode.

mean = 39.333

median = 38

mode = all

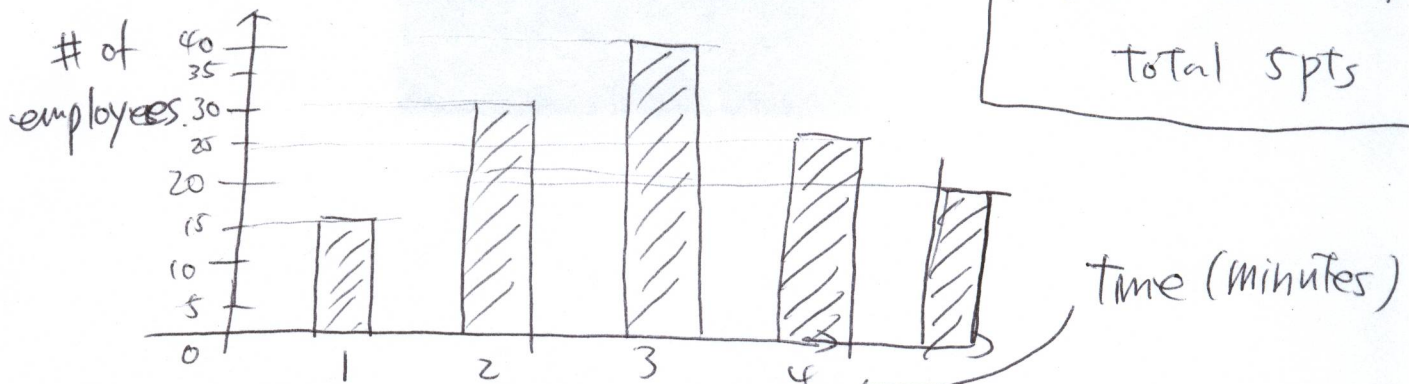
not discussing  
dist (-5 pts)

(b) What have you learned from the three statistics calculated in Part a?

mean and median are very close to each other. So it is likely we have a symmetric distribution also (or slightly right skewed) since they are very similar to each other so likely we do not have extreme outliers in one tail (too large or too small). Mode does not tell us anything.

(c) Suppose you measure the number of minutes it takes an employee to complete a task, where the maximum allowed time is 5 minutes, and each time is rounded to the nearest minute. Data from 130 employees is summarized below. Construct a frequency bar chart from this data.

Time (minutes)	1	2	3	4	5
Frequency	15	30	40	25	20



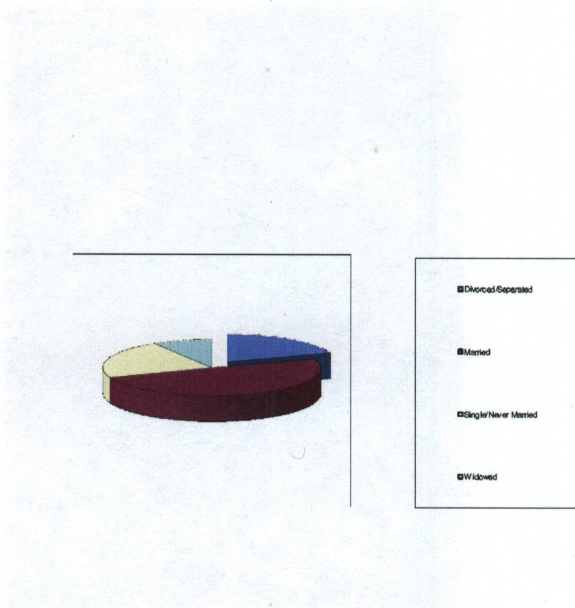
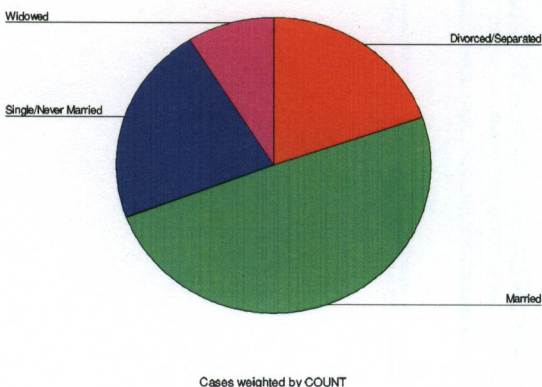
(d) Consider a survey that asks people to report their marital status. Which measure(s) of central tendency are appropriate to use?

Mode

**Question 8:**

The following pie chart shows the number of patients in a hospital in each marital status category. Data is taken from Polit, Denise F. (1996) Data Analysis and Statistics for Nursing Research, Appleton & Lange, New York NY, page 25.

We compare 2D graphics and 3D graphics.



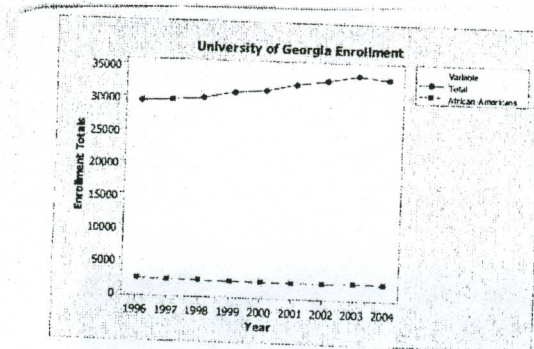
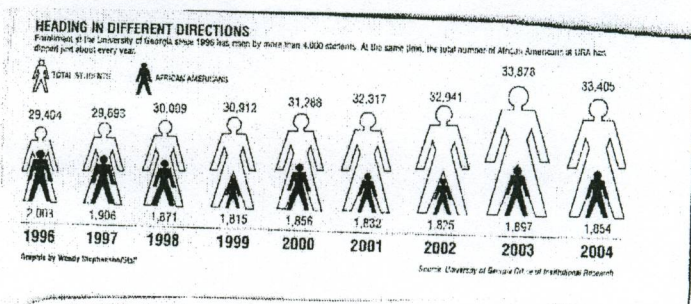
This is because

Which one (2D graphics or 3D graphics) is better? Why? Justify your answer.

**2D**: the ~~proportion of~~ % for each category is proportional to angle of each pie and area (as well as arc length). on the other hand, 3D graphic does NOT show proper angle of each pie and hard to compare ~~each~~ % of each category.

**Question 9:**

The following graphs show enrollment at the university of Georgia since 1996 to 2004 comparing the total number of students and the total number of African American students at UGA. Note that both graphs display the same data.



Compare the following two graphs. Which one is better (top one or bottom one)? Why? Justify your answer. (Again note that both graphs display the same data.)

The bottom one: the bottom one shows clear labels and ratio between the total # of students and # of American African students. On the other hand, the top one has different base for each category. This results the area of "person" ~~differ from~~ is not proportion to the actual freq of each category. This causes that it's very hard to see the ratio between the total # of students and # of American African students in each year.