Exam In-class activity (chapter 1 and 2)

Name Komal Gogi, Siva Kalyan Akella

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

Fill in the blank.

1) The three main aspects of statistics are Design, Description, and Inference. 1) _____

Provide an appropriate response.

2) What is statistics? is the art and science of learning data 2) _____

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Select the most appropriate answer.

3) The following statement refers to which aspect of a statistical study: "A meteorologist constructs a graph showing the total precipitation in Phoenix, Arizona in each of the months of a given year"? 3) C

A) Inference B) Description C) Design

4) The following statement refers to which aspect of a statistical study: "The average age of the students in a statistics class is 25 years"? 4) B

A) Inference B) Design C) Description

5) Summarizing the data that are obtained refers to which aspect of statistics? 5) B

A) Sampling
B) Description
C) Design
D) None of these
E) Inference

6) Making decisions and predictions based on the data refers to which aspect of statistics? 6) D

A) Design
B) Description
C) None of these
D) Inference
E) Sampling

7) The following statement refers to which aspect of a statistical study: "Based on a study of 25 hospitals nationwide, researchers have concluded that there is a relationship between smoking cigarettes and contracting emphysema"? 7) B

A) Description B) Inference C) Design

8) The following statement refers to which aspect of a statistical study: "Based on previous clients, a marriage counselor concludes that the majority of marriages that begin with cohabitation before marriage will result in divorce"? 8) A

A) Inference B) Description C) Design

9) Planning the methods for data collection to study the effects of Vitamin E on athletic strength would be classified as which aspect of statistics: design, inference or description? 9) C

A) Inference B) Description C) Design

The owners of a coffee shop conducted a taste test to determine whether its customers preferred a new coffee brand to the current one sold by the shop. Customers who were willing to participate were given small samples of each of the two brands in random order and were asked to select which one they preferred without knowing the brand. Of the 100 participating customers, 90% chose the new brand. Based on these results, the owners determined that a majority of their customers preferred the new brand and therefore switched their coffee supplier.

10) Predicting the preference of all of the coffee shop customers based on the taste test results refers to which aspect of statistics? 10) E

A) Design
B) Description
C) None of these
D) Investigation
E) Inference

11) Randomizing the order in which the samples of each brand were given to each customer refers to which aspect of statistics? 11) E

A) Investigation
B) Inference
C) Description
D) None of these
E) Design

12) Stating that 90% of the taste testers preferred the new brand is an example of which type of statistics? 12) B

A) Investigation
B) Inference
C) None of these
D) Description
E) Design

Select the most appropriate answer.

13) The following statement refers to which aspect of a statistical study: "From past figures, it is predicted that 47% of the registered voters in Virginia will vote in the June primary"? 13) A

A) Inference B) Description C) Design

14) Planning how to obtain data to answer the questions of interest refers to which aspect of statistics? 14) A

A) Sampling
B) Inference
C) Design
D) Description
E) None of these

15) A survey of American adults asks "would you like to see more or less government spending on natural disasters?" Of the 1496 respondents, 723 responded "more" or "much more". The population of interest consists of 15) D

A) the proportion of respondents who responded "more" or "much more"
B) the 1496 respondents
C) the 723 respondents who responded "more" or "much more"
D) all American adults
E) the proportion of American adults who would respond "more" or "much more"

Determine whether the summary measure is better described as a parameter or a statistic.

16) The proportion of teenagers in a nationwide survey who stated that they consumed alcohol on a regular basis — 16) A
A) Statistic B) Parameter

17) The proportion of regional flights for that were considered on–time during the month of December — 17) A
A) Statistic B) Parameter

Answer true or false.

18) Descriptive statistics refers to methods of making decisions or predictions about a population, based on data obtained from a sample of that population; while inferential statistics refers to methods for summarizing the data. — 18) A
A) False B) True

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

Provide an appropriate response.

19) A recent poll asked 1,245 registered voters nationwide, "Which party do you think can do a better job of handling immigration issues?" 31% of the respondents answered "Republicans." With a margin of error of ±3%, it is estimated that between 28 and 34 percent of registered voters nationwide feel that the Republican party can do a better job of handling immigration issues. Identify which part of this example is inferential. — 19) 31% respondents answered "Republicans"

20) A recent poll asked 1,245 registered voters nationwide, "Which party do you think can do a better job of handling immigration issues?" 31% of the respondents answered "Republicans." With a margin of error of ±3%, it is estimated that between 28 and 34 percent of registered voters nationwide feel that the Republican party can do a better job of handling immigration issues. Identify which part of this example is descriptive. — 20) It is estimated that between 28 and 34..

Fill in the blank.

21) A parameter is a numerical summary of the population; while a Statistics is a numerical summary of the sample. — 21) _____

22) The Population is the set of all subjects of interest; while the Sample is a subset of the entire set of interest. — 22) _____

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Answer true or false.

23) Inferential statistics are used when data are available only for a sample; however, descriptive statistics are used when data are available for either a sample or a population. — 23) A
A) True B) False

24) Random sampling enables the sample to be a good reflection of the population. — 24) B
A) False B) True

25) A postal inspector has the following data set for a sample of three packages. 25) C

Package	Length	Width	Weight	Domestic or International
1	10 in.	20 cm	2.7 lb	Domestic
2	27 cm	22 cm	1.3 lb	International
3	18 cm	14 cm	0.5 lb	Domestic

What must be done to clean up these data?

A) Change to use 0 or 1 to represent "Domestic" and the other to represent "International"

B) Convert the lengths, widths, and weights to all use standard units or all use metric units

C) Convert the lengths to all be in inches or all in centimeters

D) Delete the "Package" column

26) Which of these could be a data file for two characteristics for a sample of four people, where one characteristic takes numerical values and the other takes values that are categories? 26) A

A)

Movie Genre	Number that Enjoy
Comedy	4
Drama	3
Horror	2
Sci-Fi	2

B)

Class	Daily Screen Time (hours)
Fresh	5.5
Soph	4.75
Fresh	9.25
Jr	7.0

C)

Age	Daily Screen Time (hours)
18	7.25
19	5.25
18	3.75
18	9.5

D)

Class	Grade
Fresh	A-
Fresh	B
Soph	C+
Jr	B+

27) A researcher wants to create a device to detect the presence of explosive materials in closed luggage as they pass by on a conveyor belt. Which of these data sets would be useful as training data? 27) B

A) A data set that contains whether the device detected explosive materials in each piece of luggage and whether the luggage contained explosive materials

B) A data set that contains whether the device detected explosive materials in randomly selected pieces of luggage that may or may not contain explosive materials

C) A data set that contains whether the device detected explosive materials in each piece of luggage

D) A data set that contains whether each piece of luggage contained explosive materials

28) A center for the performing arts wants to identify which types of performances would maximize revenue. They use a data set that includes, for each of 100 randomly selected sustaining members for each of last 137 performances, whether the member attended, the individual ticket price, the date, the genre, and the local weather conditions. For which of these reasons could the data lead to algorithmic bias? 28) D

A) The data do not include the number of times each was performed.
B) The data include only the local weather conditions.
C) The sample includes only sustaining members.
D) The performances are not randomly selected.

Classify as categorical or qualitative data.

29) A survey of automobiles parked in the student and staff lots at a large college recorded the make and model of the automobiles. The variable "make" is: 29) Categorical

A) Categorical B) Quantitative

30) The amount of time spent watching television or playing video games is considered a significant factor on predicting childhood obesity. 290 parents of school-aged children were asked to estimate the number of hours per week that their child spent watching television or playing video games. This is an example of what type of variable? 30) Quantitative

A) Categorical B) Quantitative

Classify the variable as either discrete or continuous.

31) The following table shows the heights of the five tallest mountains in North America. 31) Discrete

Mountain	Height (ft)	Rank
McKinley	20,320	1
Logan	19,850	2
Citlaltepec	18,700	3
St. Elias	18,008	4
Popocatepetl	17,930	5

The heights given in the second column represent what type of data?

A) Discrete B) Continuous

Select the most appropriate answer.

32) Which of the following is a continuous variable? 32) D

A) type of fish caught
B) brand of tennis shoe
C) number of homeruns in a professional baseball player's career
D) daily high temperature in New York City
E) number of pars in a round of golf

33) Which of the following is a discrete variable? 33) A

A) number of phones per household
B) none of these
C) weight of a newborn baby
D) amount of coffee in an 8-ounce cup
E) time it takes to drive to work

Classify the variable as either discrete or continuous.

34) The time it takes an athlete to run 100 meters. 34) Continuous

A) Discrete B) Continuous

Provide an appropriate response.

35) A safety engineer wishes to use the following data to show the number of deaths in a year from the collision of passenger cars with trucks on a particular highway. 35) E

Year	Number of Deaths
1	12
2	17
3	22
4	21
5	16
6	13
7	11
8	12

What is the mode of the number of deaths?

A) 15.5 B) 16 C) 22 D) 13 E) 12

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

36) A stock broker has been following different stocks over the last month and has recorded whether the various stock values are up, unchanged, or down at the end of the month. The results were 36) _____

Stock performance	up	same	down
Count	21	7	12

a. What is the variable of interest? No of stocks
b. Is the variable categorical or quantitative? Quantitative
c. Which response is the mode? No Mode
d. Add proportions to this frequency table. 13.3

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

37) Parking at a large university has become a major issue. University administrators would like to determine the average time it takes a student to find a parking spot in a university lot. Students who are willing to participate in the study were asked to record the time between entering campus and pulling into a parking spot. Which of the following would not be appropriate for displaying the parking time data? 37) C

A) Box plot
B) None of these should be used.
C) Histogram
D) Pie chart
E) Stem-and-leaf plot

38) Each year advertisers spend billions of dollars purchasing commercial time on network sports television. A recent article listed the top 10 leading spenders (in millions of dollars) over a 6 month period: 38) A

Company A	$72.0	Company F	$26.9
Company B	63.1	Company G	25.0
Company C	54.7	Company H	23.9
Company D	54.3	Company I	23.0
Company E	29.0	Company J	20.0

Which of the following graphs would not be appropriate for displaying this data?

A) Pie chart
B) Stem-and-leaf plot
C) None of these should be used.
D) Histogram
E) Dot plot

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

39) The enrollment for fall semester at University X is as follows. 39) _____

Enrollment	Count
Undergraduate	24,814
Graduate/Professional	8386
Independent Study	20

a. Construct a bar graph for these data.
b. Would a dot plot or a stem-and-leaf plot make sense for these data? Explain.

40) The following data represent the number of grams of fat in various breakfast foods. 40) _____

Breakfast Food	Fat (in grams)
Muffin and egg sandwich	12
Muffin, egg, and ham sandwich	22
Muffin, egg, and bacon sandwich	27
Muffin and sausage sandwich	22
Bagel, egg, and ham sandwich	25
Bagel, egg, and bacon sandwich	30
Bagel, egg, and sausage sandwich	32
Bagel, egg, sausage, and cheese sandwich	37
Bagel, egg, ham, and cheese sandwich	27
Bagel, egg, bacon, and cheese sandwich	31
Bagel	11
Pancakes platter	16
Pancakes and eggs platter	21
Pancakes, eggs, and bacon platter	32
Yogurt	2

Construct a dot plot for these data.

MULTIPLE CHOICE. **Choose the one alternative that best completes the statement or answers the question.**

A sporting goods retailer conducted a customer survey to determine its customers primary reason for shopping at their store. The results are shown in the graph below.

41) What proportion of the customers responded that the merchandise was the reason they shopped at the store? 41) B

A) 30
B) 0.43
C) 0.50
D) none of these
E) 0.30

SHORT ANSWER. **Write the word or phrase that best completes each statement or answers the question.**

Provide an appropriate response.

42) A survey investigated exposure to tobacco use in a series of G-rated animated films. Data on the total tobacco exposure time (in seconds) is below. 42) _____

223	176	548	37	158	51	299	37	11
165	74	9	2	9	23	206	9	

Construct a dot plot for these data. Comment on the shape of the distribution.

43) The scores for a statistics test are as follows: 43) _____

87 76 94 77 95 96 88 85 66 89
79 98 54 90 83 88 82 55 14 69

Create a stem-and-leaf display for the data. The stem should consist of the tens digit and range from 1 to 9. The leaves should be drawn aside the appropriate stem based on the data values.

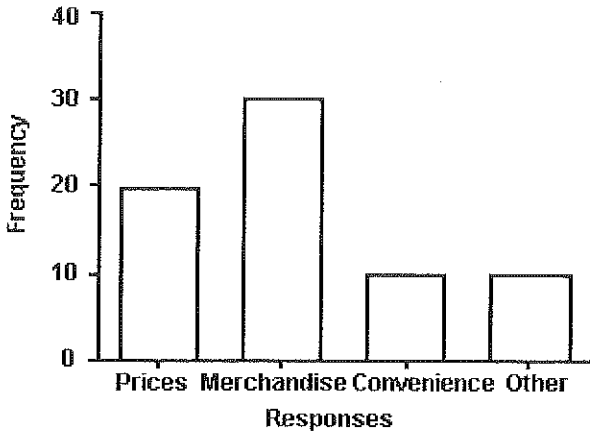
44) The table below shows the unemployment rate in one city from 2003 to 2012. 44) _____

Year	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Unemployment Rate (Percent)	5.90	5.78	5.45	5.28	5.06	4.88	4.80	4.63	4.44	4.24

a. Construct a time plot for these data.
b. Is there a trend? If so, what kind? Yes, decreasing
c. Would a histogram more clearly describe the above dataset? Explain. yes

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

A sporting goods retailer conducted a customer survey to determine its customers primary reason for shopping at their store. The results are shown in the graph below.

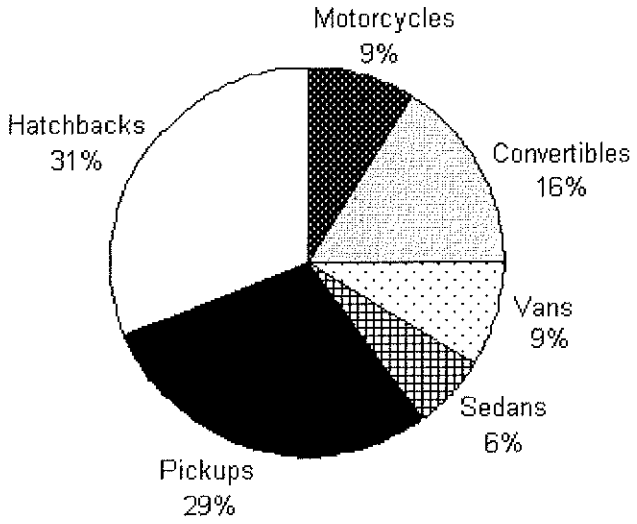


45) What percentage of the customers gave "prices" or "merchandise" as their answer? 45) B

A) 0.30 B) 0.71 C) 0.14 D) 0.10 E) 0.20

Provide an appropriate response.

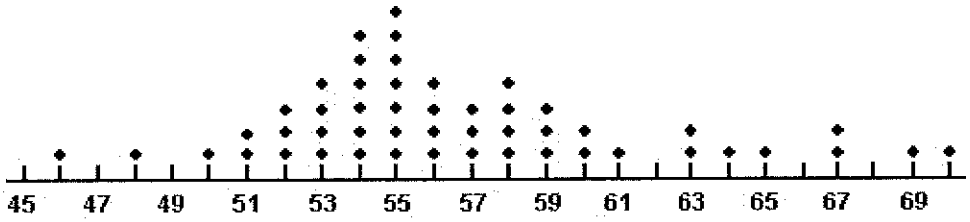
46) Results from a survey of 7116 vehicle types on the campus of State College are summarized in the following pie chart. 46) B



How many of the vehicles were sedans? Give your answer to the nearest whole number.

A) 4270 B) 427 C) 60 D) 600 E) 6

A sample of fifty motorists was taken on a Federal highway where the speed limit was 60 miles per hour. A dot plot of their speeds is shown below.



47) What proportion of the motorists were speeding? 47) A

A) 0.22 B) 0.72 C) 0.18 D) 2 E) 0.04

48) What is the variable of interest? 48) D

A) whether or not a motorist was speeding
B) number of motorists on the Federal highway
C) motorist's speed
D) number of speeding motorists

A survey was conducted to determine how people rated the quality of programming available on television. Respondents were asked to rate the overall quality from 0 (no quality at all) to 100 (extremely good quality). The stem-and-leaf display of the data is shown below.

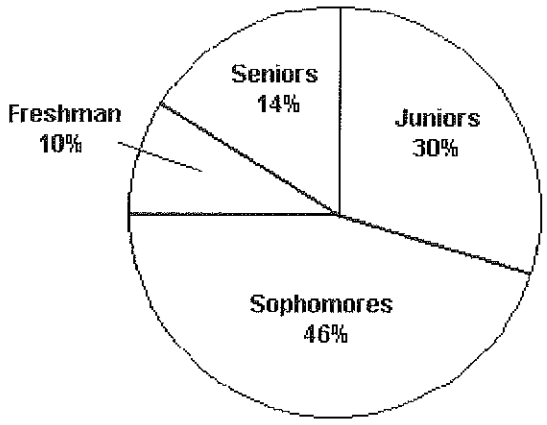
Stem	Leaves
3	2 6
4	0 3 4 7 8 9 9 9
5	0 1 1 2 3 4 5
6	1 2 5 6 6
7	1 7
8	
9	3

49) What percentage of the respondents rated overall television quality as very good (regarded as ratings of 80 and above)? 49) A

A) 1% B) 12% C) 32% D) 4% E) 3%

Provide an appropriate response.

50) The professor of economics at a small Texas University wanted to determine what year in school students were taking his tough economics course. Shown below is a pie chart of the results. 50) B



What percentage of the class took the course prior to reaching their senior year?

A) 44% B) 86% C) 14% D) 54% E) 30%

For the following variable, indicate whether you would expect its histogram to be symmetric, skewed to the right, or skewed to the left.

51) The scores of students (out of 100 points) on a very difficult exam in which most score poorly, but a few score very well 51) B

A) symmetric B) skewed to the right C) skewed to the left

52) Time needed to complete an easy exam (maximum time is 1 hour) 52) C

A) skewed to the left B) skewed to the right C) symmetric

53) The age of death for the general public 53) B

A) symmetric B) skewed to the left C) skewed to the right

Answer true or false.

54) Bar graphs and pie charts are graphical methods that are often used in summarizing quantitative data. 54) A

A) False B) True

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

Fill in the blank.

55) A _____ is a graph that uses bars to portray the frequencies or the relative frequencies of the possible outcomes for a quantitative variable. 55) Histogram

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Select the most appropriate answer.

56) A common pattern observed over time is called a/an 56) A

A) trend.
B) None of these.
C) mode
D) time series.
E) time plot.

Provide an appropriate response.

57) The distribution of salaries of professional basketball players is skewed to the right. Which measure of central tendency would be the best measure to determine the location of the center of the distribution? 57) B

A) Mode
B) Median
C) Standard Deviation
D) Range
E) Mean

Find the median for the given sample data.

58) A store manager kept track of the number of newspapers sold each week over a seven-week period. The results are shown below. 58) B

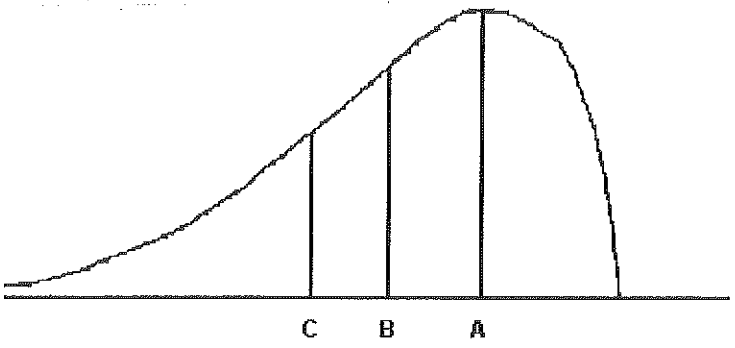
95, 38, 221, 122, 258, 237, 233

Find the median number of newspapers sold.

A) 258 newspapers
B) 221 newspapers
C) 233 newspapers
D) 122 newspapers
E) 172 newspapers

Provide an appropriate response.

59) For the distribution shown below, identify the mean, median, and mode 59) C



A) A = mode, B = median, C = mean
B) A = median, B = mode, C = mean
C) A = median, B = mean, C = mode
D) A = mode, B = mean, C = median
E) A = mean, B = mode, C = median

60) The mean is less than the median 60) A

A) when the data is skewed to the left
B) when the data is symmetric
C) when the data is skewed to the right
D) never

61) Last year, batting averages in the National League averaged 0.257 with a high of 0.323 and a low of 0.250 (minimum 250 at bats). Based on this information, which measure of variation could be calculated? 61) A

A) mode
B) standard deviation
C) none of the above
D) range
E) variance

62) For the stem-and-leaf plot below, find the range of the data set. 62) D

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

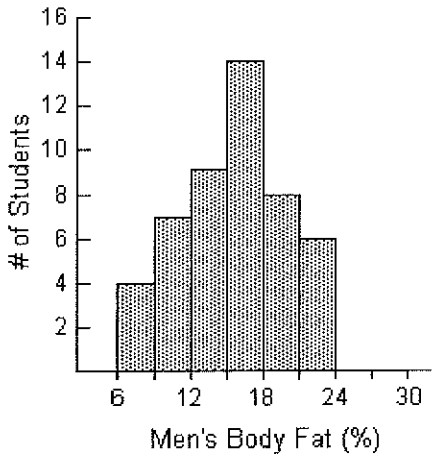
A) 14 B) 40 C) 36 D) 26 E) 34

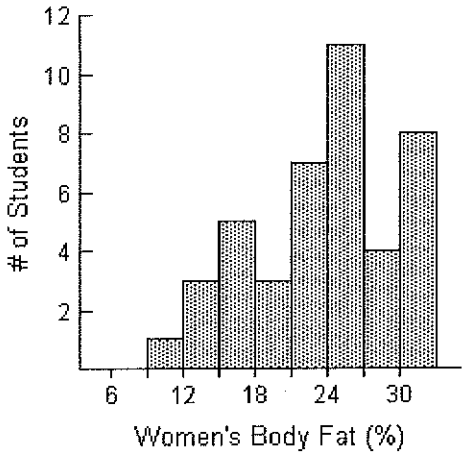
63) The cost for one semester's books (in dollars) are given below for a sample of five college students. Calculate the sample standard deviation, s of the book costs. Round to the nearest hundredth when necessary. 63) D

340, 170, 145, 420, 120

A) 17,680 B) 300 C) 132.97 D) 118.93

64) The histograms below display the body fat percentages of 42 female students and 48 male students taking a college health course. 64) A

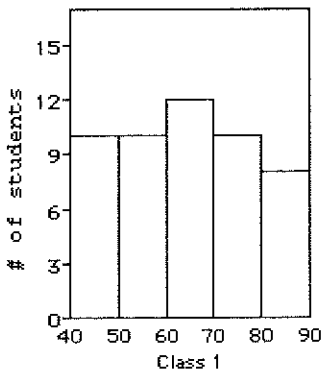


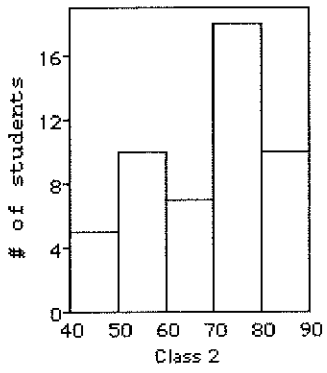


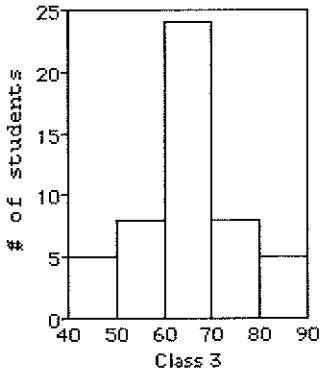
Do the female or male students have a larger standard deviation?

A) female students B) male students

65) Three statistics classes (each of 50 students) took the same test. Shown below are histograms of the scores for the classes. Which class had the smallest standard deviation? Which class had the largest standard deviation? 65) C







A) 3, 2 B) 2, 1 C) 1, 3 D) 3, 1