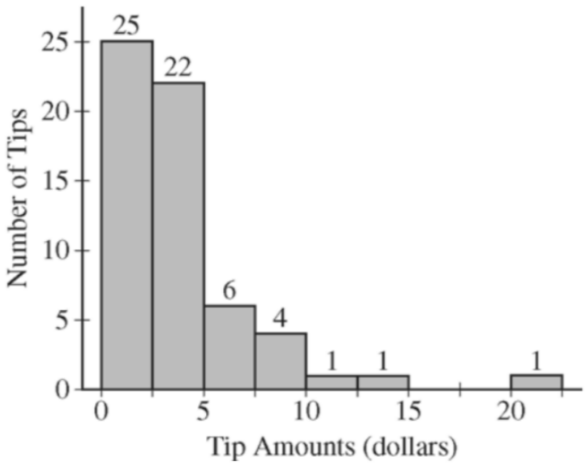
Unit 2 Day 9: FRAPPYs

2016 #1

1. Robin works as a server in a small restaurant, where she can earn a tip (extra money) from each customer she serves. The histogram below shows the distribution of her 60 tip amounts for one day of work.



(a) Write a few sentences to describe the distribution of tip amounts for the day shown.

(b) One of the tip amounts was $8. If the $8 tip had been $18, what effect would the increase have had on the following statistics? Justify your answers.

The mean:

The median:

Intent of Question

The primary goals of this question were to assess a student's ability to (1) describe the distribution of a quantitative variable based on a histogram and (2) determine the effect of changing one data value on the mean and the median.

Solution

Part (a):

The distribution of Robin's tip amounts is skewed to the right. There is a gap between the largest tip amount (in the \$20 to \$22.50 interval) and the second largest tip amount (in the \$12.50 to \$15 interval), and the largest tip amount appears to be an outlier. The median tip amount is between \$2.50 and \$5.00. Robin's tip amounts vary from a minimum of between \$0 and \$2.50 to a maximum of between \$20.00 and \$22.50. About 78 percent of the tip amounts are between \$0 and \$5.

Part (b):

The mean: If the \$8 tip had been \$18, the mean would increase by \$10 divided by 60, or $\$\frac{1}{6}$, or about 17 cents.

The median: If the \$8 tip had been \$18, the median would not change because the current median is between \$2.50 and \$5.00, and both \$8 and \$18 are greater than that.

Scoring

Parts (a) and (b) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as follows:

Essentially correct (E) if the response includes reasonable comments on the following five components:

1. Shape (skewed right)
2. Outlier (at least one) *OR* gap (one tip amount greater than $20, next highest at most $15)
3. Center between $2.50 and $5.00 (median) or between $2.62 and $5.13 (mean)
4. Variability, by noting that the tip amounts vary from about $0 to at most $22.50, or that a majority of tip amounts are between $0 and a value greater than or equal to $5, or by providing a correct numerical approximation of a measure of variability
5. Context (tip amounts)

Partially correct (P) if the response includes only three or four of the five components.

Incorrect (I) if the response includes at most two of the five components.

Part (b) is scored as follows:

Essentially correct (E) if the response includes the following four components:

1. Comments that the mean will increase
2. Correctly justifies why the mean will increase
3. Comments that the median will not change
4. Correctly justifies why the median will not change

Partially correct (P) if the response includes only two or three of the four components.

Incorrect (I) if the response includes at most one of the four components.

4 **Complete Response**

Both parts essentially correct

3 **Substantial Response**

One part essentially correct and one part partially correct

2 **Developing Response**

One part essentially correct and one part incorrect

OR

Both parts partially correct

1 **Minimal Response**

One part partially correct and one part incorrect

2015 #1

1. Two large corporations, A and B, hire many new college graduates as accountants at entry-level positions. In 2009 the starting salary for an entry-level accountant position was $36,000 a year at both corporations. At each corporation, data were collected from 30 employees who were hired in 2009 as entry-level accountants and were still employed at the corporation five years later. The yearly salaries of the 60 employees in 2014 are summarized in the boxplots below.

(a) Write a few sentences comparing the distributions of the yearly salaries at the two corporations.

(b) Suppose both corporations offered you a job for $36,000 a year as an entry-level accountant.

(i) Based on the boxplots, give one reason why you might choose to accept the job at corporation A.

(ii) Based on the boxplots, give one reason why you might choose to accept the job at corporation B.

Intent of Question

The primary goals of this question were to assess a student's ability to (1) compare features of two distributions of data displayed in boxplots and (2) identify statistical measures that are important in making decisions based on data sets.

Solution

Part (a):

The median salary is approximately the same for both corporations. The range and interquartile range of the salaries are greater for Corporation A than for Corporation B. The two highest salaries at Corporation A are outliers while Corporation B has no outliers.

Part (b):

(i) Five years after starting, at least 3 out of 30 (10%) of the salaries at Corporation A are greater than the maximum salary at Corporation B. If I accept the offer from Corporation A, I might be able to make a higher salary at Corporation A than at Corporation B.

(ii) Five years after starting, the minimum salary at Corporation B is greater than at Corporation A. In fact, at Corporation A it looks like some people are still making the starting salary of $36,000 and never received a raise in the five years since they were hired. So if I work at Corporation A, I might never receive a raise in salary.

Scoring

Parts (a) and (b) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as follows:

Essentially correct (E) if the response includes the following four components:

1. A correct comparison of center.
2. A correct comparison of spread.
3. A discussion of the outliers for Corporation A.
4. The response is in context.

Partially correct (P) if the response includes only three of the four components.

Incorrect (I) if the response includes at most two of the four components.

Note: Any mention of shape should be ignored because complete shape information cannot be determined from a boxplot.

Part (b) is scored as follows:

Essentially correct (E) if the response includes the following four components:

1. In part (b-i) a relevant statistical measure is identified (or described) or a relevant statistical comparison is provided that supports the choice of Corporation A.
2. In part (b-i) an explanation is provided for why the measure or comparison is relevant.
3. In part (b-ii) a relevant statistical measure is identified (or described) or a relevant statistical comparison is provided that supports the choice of Corporation B.
4. In part (b-ii) an explanation is provided for why the measure or comparison is relevant.

Partially correct (P) if the response includes only two or three of the four components.

Incorrect (I) if the response includes none or one of the four components.

Note: If a response does not provide a statistical measure or comparison in part (b-i) or (b-ii), the second and fourth components can still be satisfied if an acceptable explanation is provided that would follow from a relevant statistical measure or comparison. For example, if the response in part (b-i) only states "At Corporation A, I have the potential to earn a higher salary," the second component is satisfied.

4 **Complete Response**

Both parts essentially correct

3 **Substantial Response**

One part essentially correct and one part partially correct

2 **Developing Response**

One part essentially correct and one part incorrect

OR

Both parts partially correct

1 **Minimal Response**

One part partially correct and one part incorrect

2002 EXAM - MULTIPLE CHOICE PRACTICE

20. A small town employs 34 salaried, nonunion employees. Each employee receives an annual salary increase of between $500 and $2,000 based on a performance review by the mayor's staff. Some employees are members of the mayor's political party, and the rest are not.

Students at the local high school form two lists, A and B, one for the raises granted to employees who are in the mayor's party, and the other for raises granted to employees who are not. They want to display a graph (or graphs) of the salary increases in the student newspaper that readers can use to judge whether the two groups of employees have been treated in a reasonably equitable manner.

Which of the following displays is least likely to be useful to readers for this purpose?

(A) Back-to-back stemplots of A and B

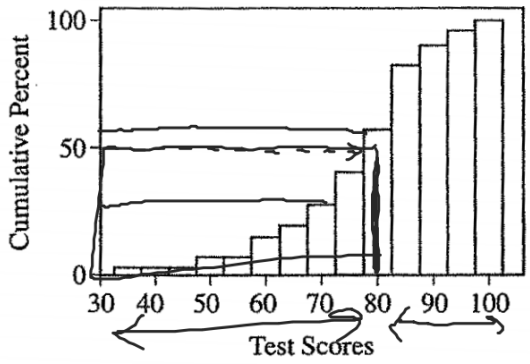
(B) Scatterplot of B *versus* A

(C) Parallel boxplots of A and B

(D) Histograms of A and B that are drawn to the same scale

(E) Dotplots of A and B that are drawn to the same scale

AP STATISTICS
TEST SCORES



27. The figure above shows a cumulative relative frequency histogram of 40 scores on a test given in an AP Statistics class. Which of the following conclusions can be made from the graph?

(A) There is greater variability in the lower 20 test scores than in the higher 20 test scores.

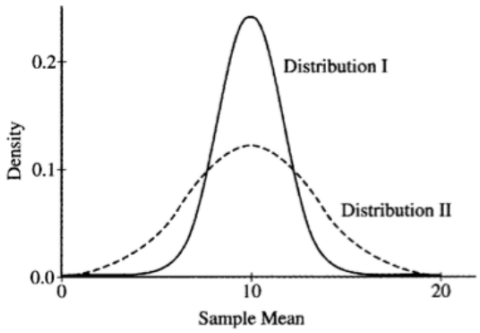
(B) The median test score is less than 50.

(C) Sixty percent of the students had test scores above 80.

(D) If the passing score is 70, most students did not pass the test.

(E) The horizontal nature of the graph for test scores of 60 and below indicates that those scores occurred most frequently.

2007 AP EXAM



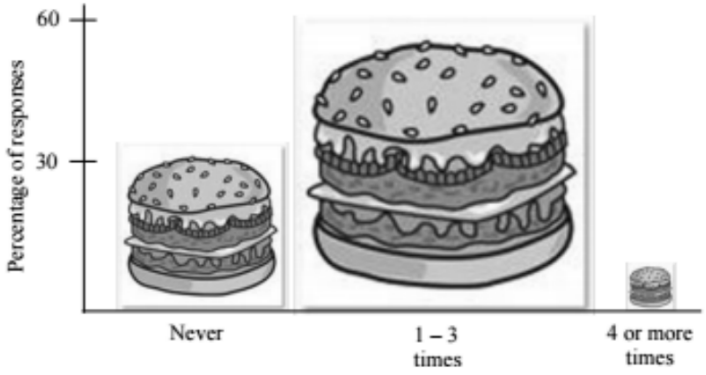
23. The graphs of the sampling distributions, I and II, of the sample mean of the same random variable for samples of two different sizes are shown below. Which of the following statements must be true about the sample sizes?

(A) The sample size of I is less than the sample size of II.
(B) The sample size of I is greater than the sample size of II.
(C) The sample size of I is equal to the sample size of II.
(D) The sample size does not affect the sampling distribution.
(E) The sample sizes cannot be compared based on these graphs.

1. At the beginning of the school year, a high-school teacher asks every student in her classes to fill out a survey that asks for their age, gender, the number of years they have lived at their current address, their favorite school subject, and whether they plan to go to college after high school. Which of the following best describes the variables that are being measured?

(a) four quantitative variables
(b) five quantitative variables
(c) two categorical variables and two quantitative variables
(d) two categorical variables and three quantitative variables
(e) three categorical variables and two quantitative variables

2. The graph below shows how mothers of young children respond to the question, "How many times a week do you choose fast food as a dining option for your family?"

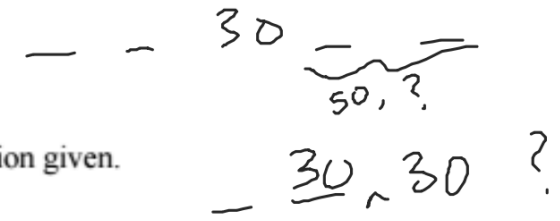


What's wrong with this method of presenting information?

(a) This kind of data should always be presented in a pie chart.
(b) The vertical axis should be "number of responses," not "percentage of responses."
(c) The horizontal axis should be divided into more than three categories.
(d) Using proportionally-sized hamburgers exaggerates differences between responses.
(e) We don't know if the mothers who responded were thinking about dinner, or both lunch and dinner.

3. The median age of five people in a meeting is 30 years. One of the people—a 50-year-old—leaves the room. The median age of the remaining four people in the room is

(a) 40 years.
(b) 30 years.
(c) 25 years.
(d) less than 30 years.
(e) Cannot be determined from the information given.

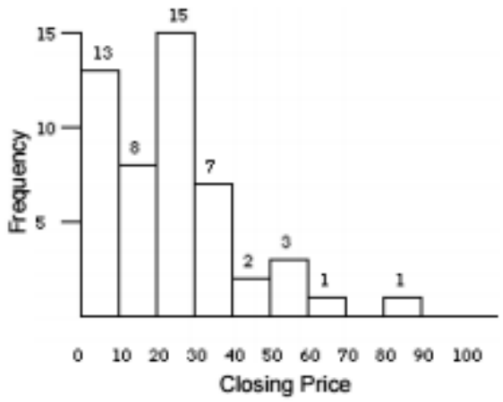


5. A researcher reports that the participants in his study lost a mean of 10.4 pounds after two months on his new diet. A friend of yours comments that she tried the diet for two months and lost no weight, so clearly the report was a fraud. Which of the following statements is correct?

(a) Your friend must not have followed the diet correctly, since she did not lose weight.
(b) Since your friend did not lose weight, the report must not be correct.
(c) The report gives only the mean. This does not imply that all participants in the study lost 10.4 pounds or even that all lost weight. Your friend's experience does not necessarily contradict the study results.
(d) In order for the study to be correct, we must now add your friend's results to those of the study and recalculate the new average.
(e) Your friend is an outlier.

6. The following is a histogram showing the actual frequency of the closing prices of a particular stock on the New York Stock Exchange over a 50-day period. The class that contains the third quartile is

(a) 10–20
(b) 20–30
(c) 30–40
(d) 40–50
(e) 50–60



7. For the data in the previous problem, which measures of center and spread would be most appropriate to use?

(a) Mean and standard deviation
(b) Mean and interquartile range
(c) Mean and range
(d) Median and interquartile range
(e) Median and standard deviation