

*Solutions Manual for Introductory
Business Statistics 2nd Edition by
Holmes, Illowsky, Dean, Hadley*

ISBN: 9781711472591



Introductory Business

Statistics 2e.

CHAPTER 1: SAMPLING AND DATA

Exercise 1. *Identify: a. the population, b. the sample, c. the parameter, d. the statistic, e. the variable, and f. the data. Give examples where appropriate.*
A fitness center is interested in the mean amount of time a client exercises in the center each week.

Solution

- a. The population is all of the clients of the fitness center.
- b. A sample of the clients that use the fitness center for a given week.
- c. The average amount of time that all clients exercise in one week.
- d. The average amount of time that a sample of clients exercises in one week.
- e. The amount of time that a client exercises in one week.
- f. Examples are: 2 hours, 5 hours, and 7.5 hours

Exercise 2. *Identify: a. the population, b. the sample, c. the parameter, d. the statistic, e. the variable, and f. the data. Give examples where appropriate.*
Ski resorts are interested in the mean age that children take their first ski and snowboard lessons. They need this information to plan their ski classes optimally.

Solution

- a. all children who take ski or snowboard lessons
- b. a group of these children
- c. the population mean age of children who take their first snowboard lesson
- d. the sample mean age of children who take their first snowboard lesson
- e. X = the age of one child who takes their first ski or snowboard lesson
- f. values for X , such as 3, 7, and so on

Exercise 3. *Identify: a. the population, b. the sample, c. the parameter, d. the statistic, e. the variable, and f. the data. Give examples where appropriate.*
A cardiologist is interested in the mean recovery period of her patients who have had heart attacks.

Solution

- a. the cardiologist's patients
- b. a group of the cardiologist's patients
- c. the mean recovery period of all of the cardiologist's patients
- d. the mean recovery period of the group of the cardiologist's patients
- e. X = the mean recovery period of one patient
- f. values for X , such as 10 days, 14 days, 20 days, and so on

Exercise 4. *Identify: a. the population, b. the sample, c. the parameter, d. the statistic, e. the variable, and f. the data. Give examples where appropriate.*
Insurance companies are interested in the mean health costs each year of their clients, so that they can determine the costs of health insurance.

Solution

- a. the clients of the insurance companies
- b. a group of the clients

- c. the mean health costs of the clients
- d. the mean health costs of the sample
- e. X = the health costs of one client
- f. values for X , such as 34, 9, 82, and so on

Exercise 5. *Identify: a. the population, b. the sample, c. the parameter, d. the statistic, e. the variable, and f. the data. Give examples where appropriate.*
A politician is interested in the proportion of voters in his district who think he is doing a good job.

Solution

- a. all voters in the politician's district
- b. a random selection of voters in the politician's district
- c. the proportion of voters in this district who think this politician is doing a good job
- d. the proportion of voters in this district who think this politician is doing a good job in the sample
- e. X = the number of voters in the district who think this politician is doing a good job
- f. Yes, he is doing a good job. No, he is not doing a good job.

Exercise 6. *Identify: a. the population, b. the sample, c. the parameter, d. the statistic, e. the variable, and f. the data. Give examples where appropriate.*
A marriage counselor is interested in the proportion of clients she counsels who stay married.

Solution

- a. all the clients of this counselor
- b. a group of clients of this marriage counselor
- c. the proportion of all their clients who stay married
- d. the proportion of the sample of the counselor's clients who stay married
- e. X = the number of couples who stay married
- f. yes, no

Exercise 7. *Identify: a. the population, b. the sample, c. the parameter, d. the statistic, e. the variable, and f. the data. Give examples where appropriate.*
Political pollsters may be interested in the proportion of people who will vote for a particular cause.

Solution

- a. all voters (in a certain geographic area)
- b. a random selection of all the voters
- c. the proportion of voters who are interested in this particular cause
- d. the proportion of voters who are interested in this particular cause in the sample
- e. X = the number of voters who are interested in this particular cause
- f. yes, no

Exercise 8. *Identify: a. the population, b. the sample, c. the parameter, d. the statistic, e. the variable, and f. the data. Give examples where appropriate.*
A marketing company is interested in the proportion of people who will buy a particular product.

Solution

- a. all people (maybe in a certain geographic area, such as the United States)
- b. a group of the people
- c. the proportion of all people who will buy the product
- d. the proportion of the sample who will buy the product
- e. X = the number of people who will buy it
- f. buy, not buy

Exercise 9.

A Lake Tahoe Community College instructor is interested in the mean number of days Lake Tahoe Community College math students are absent from class during a quarter. What is the population she is interested in?

- a. all Lake Tahoe Community College students
- b. all Lake Tahoe Community College English students
- c. all Lake Tahoe Community College students in the instructor's classes
- d. all Lake Tahoe Community College math students

Solution D

Exercise 10.

A Lake Tahoe Community College instructor is interested in the mean number of days Lake Tahoe Community College math students are absent from class during a quarter.
Consider the following:
 X = number of days a Lake Tahoe Community College math student is absent
In this case, X is an example of a:

- a. variable.
- b. population.
- c. statistic.
- d. data.

Solution A

Exercise 11.

A Lake Tahoe Community College instructor is interested in the mean number of days Lake Tahoe Community College math students are absent from class during a quarter.
The instructor's sample produces a mean number of days absent of 3.5 days.
This value is an example of a:

- a. parameter.
- b. data.
- c. statistic.
- d. variable.

Solution C

Exercise 12. *Identify the type of data that would be used to describe a response (quantitative discrete, quantitative continuous, or qualitative), and give an example of the data.*
number of tickets sold to a concert

Solution quantitative discrete, 150

Exercise 13. *Identify the type of data that would be used to describe a response (quantitative discrete, quantitative continuous, or qualitative), and give an example of the data.*
percent of body fat

Solution quantitative continuous, 19.2%

Exercise 14. *Identify the type of data that would be used to describe a response (quantitative discrete, quantitative continuous, or qualitative), and give an example of the data.*
favorite baseball team

Solution qualitative, Oakland A's

Exercise 15. *Identify the type of data that would be used to describe a response (quantitative discrete, quantitative continuous, or qualitative), and give an example of the data.*
time in line to buy groceries

Solution quantitative continuous, 7.2 minutes

Exercise 16. *Identify the type of data that would be used to describe a response (quantitative discrete, quantitative continuous, or qualitative), and give an example of the data.*
number of students enrolled at Evergreen Valley College

Solution quantitative discrete, 11,234 students

Exercise 17. *Identify the type of data that would be used to describe a response (quantitative discrete, quantitative continuous, or qualitative), and give an example of the data.*
most-watched television show

Solution qualitative, Dancing with the Stars

Exercise 18. *Identify the type of data that would be used to describe a response (quantitative discrete, quantitative continuous, or qualitative), and give an example of the data.*
brand of toothpaste

Solution qualitative, Crest

Exercise 19. *Identify the type of data that would be used to describe a response (quantitative discrete, quantitative continuous, or qualitative), and give an example of the data.*
distance to the closest movie theatre

Solution quantitative continuous, 8.32 miles

Exercise 20. *Identify the type of data that would be used to describe a response (quantitative discrete, quantitative continuous, or qualitative), and give an example of the data.*
age of executives in Fortune 500 companies

Solution quantitative continuous, 47.3 years

Exercise 21. *Identify the type of data that would be used to describe a response (quantitative discrete, quantitative continuous, or qualitative), and give an example of the data.*
number of competing computer spreadsheet software packages

Solution quantitative discrete, three

Exercise 22. *A study was done to determine the age, number of times per week, and the duration (amount of time) of resident use of a local park in San Jose. The first house in the neighborhood around the park was selected randomly and then every 8th house in the neighborhood around the park was interviewed.*
“Number of times per week” is what type of data?
a. qualitative
b. quantitative discrete
c. quantitative continuous

Solution B

Exercise 23. *A study was done to determine the age, number of times per week, and the duration (amount of time) of resident use of a local park in San Jose. The first house in the neighborhood around the park was selected randomly and then every 8th house in the neighborhood around the park was interviewed.*
“Duration (amount of time)” is what type of data?
a. qualitative
b. quantitative discrete

c. quantitative continuous

Solution C

Exercise 24. *Airline companies are interested in the consistency of the number of babies on each flight, so that they have adequate safety equipment. Suppose an airline conducts a survey. Over Thanksgiving weekend, it surveys six flights from Boston to Salt Lake City to determine the number of babies on the flights. It determines the amount of safety equipment needed by the result of that study.*
a. Using complete sentences, list three things wrong with the way the survey was conducted.
b. Using complete sentences, list three ways that you would improve the survey if it were to be repeated.

Solution a. The survey was conducted using six similar flights.
The survey would not be a true representation of the entire population of air travelers.
Conducting the survey on a holiday weekend will not produce representative results.
b. Conduct the survey during different times of the year.
Conduct the survey using flights to and from various locations.
Conduct the survey on different days of the week.

Exercise 25. *Suppose you want to determine the mean number of students per statistics class in your state. Describe a possible sampling method in three to five complete sentences. Make the description detailed.*

Solution Answers will vary. Sample Answer: Randomly choose 25 colleges in the state. Use all statistics classes from each of the chosen colleges in the sample. This can be done by listing all the colleges together with a two-digit number starting with 00 then 01, etc. The list of colleges can be found on Wikipedia. http://en.wikipedia.org/wiki/List_of_colleges_and_universities_in_California Use a random number generator to pick 25 colleges.

Exercise 26. *Suppose you want to determine the mean number of cans of soda drunk each month by students in their twenties at your school. Describe a possible sampling method in three to five complete sentences. Make the description detailed.*

Solution Answers will vary. Sample Answer: You could use a systematic sampling method. Stop the tenth person as they leave one of the buildings on campus at 9:50 in the morning. Then stop the tenth person as they leave a different building on campus at 1:50 in the afternoon.

Exercise 27. *List some practical difficulties involved in getting accurate results from a telephone survey.*

Solution Answers will vary. Sample Answer: Not all people have a listed phone number. Many people hang up or do not respond to phone surveys.

Exercise 28. *List some practical difficulties involved in getting accurate results from a mailed survey.*

Solution Answers will vary. Sample Answer: Many people will not respond to mail surveys. If they do respond to the surveys, you can't be sure who is responding. In addition, mailing lists can be incomplete.

Exercise 29. *With your classmates, brainstorm some ways you could overcome these problems if you needed to conduct a phone or mail survey.*

Solution Ask everyone to include their age then take a random sample from the data. Include in the report how the survey was conducted and why the results may not be accurate.

Exercise 30. *The instructor takes a sample by gathering data on five randomly selected students from each Lake Tahoe Community College math class. The type of sampling used is*
a. cluster sampling
b. stratified sampling
c. simple random sampling
d. convenience sampling

Solution B

Exercise 31. *A study was done to determine the age, number of times per week, and the duration (amount of time) of residents using a local park in San Jose. The first house in the neighborhood around the park was selected randomly and then every eighth house in the neighborhood around the park was interviewed. The sampling method was:*
a. simple random
b. systematic
c. stratified
d. cluster

Solution b

Exercise 32. *Name the sampling method used in each of the following situations:*
a. A person in the airport is handing out questionnaires to travelers asking them to evaluate the airport's service. They do not ask travelers who are hurrying through the airport with their hands full of luggage but instead ask all travelers who are sitting near gates and not taking naps while they wait.
b. A teacher wants to know if students are doing homework, so they randomly select rows two and five and then call on all students in row two and all

students in row five to present the solutions to homework problems to the class.

c. The marketing manager for an electronics chain store wants information about the ages of its customers. Over the next two weeks, at each store location, 100 randomly selected customers are given questionnaires to fill out asking for information about age, as well as about other variables of interest.

d. The librarian at a public library wants to determine what proportion of the library users are children. The librarian has a tally sheet on which they mark whether books are checked out by an adult or a child. The librarian records this data for every fourth patron who checks out books.

e. A political party wants to know the reaction of voters to a debate between the candidates. The day after the debate, the party's polling staff calls 1,200 randomly selected phone numbers. If a registered voter answers the phone or is available to come to the phone, that registered voter is asked whom they intend to vote for and whether the debate changed their opinion of the candidates.

Solution

- a. convenience
- b. cluster
- c. stratified
- d. systematic
- e. simple random

Exercise 33. *A "random survey" was conducted of 3,274 people of the "microprocessor generation" (people born since 1971, the year the microprocessor was invented). It was reported that 48% of those individuals surveyed stated that if they had \$2,000 to spend, they would use it for computer equipment. Also, 66% of those surveyed considered themselves relatively savvy computer users.*

a. Do you consider the sample size large enough for a study of this type? Why or why not?

b. Based on your "gut feeling," do you believe the percents accurately reflect the U.S. population for those individuals born since 1971? If not, do you think the percents of the population are actually higher or lower than the sample statistics? Why?

Additional information: The survey, reported by Intel Corporation, was filled out by individuals who visited the Los Angeles Convention Center to see the Smithsonian Institute's road show called "America's Smithsonian."

c. With this additional information, do you feel that all demographic and ethnic groups were equally represented at the event? Why or why not?

d. With the additional information, comment on how accurately you think the sample statistics reflect the population parameters.

Solution

a. Yes, in polling, samples that are from 1,200 to 1,500 observations are considered large enough and good enough if the survey is random and is well done.

- b. We do not have enough information to decide if this is a random sample from the U.S. population.
- c. No, this is a convenience sample taken from individuals who visited an exhibition in the Angeles Convention Center. This sample is not representative of the U.S. population.
- d. It is possible that the two sample statistics, 48% and 66% are larger than the true parameters in the population at large. In any event, no conclusion about the population proportions can be inferred from this convenience sample.

- Exercise 34. *The Well-Being Index is a survey that follows trends of U.S. residents on a regular basis. There are six areas of health and wellness covered in the survey: Life Evaluation, Emotional Health, Physical Health, Healthy Behavior, Work Environment, and Basic Access. Identify the type of data obtained from each question used in this survey: qualitative, quantitative discrete, or quantitative continuous.*
- a. *Do you have any health problems that prevent you from doing any of the things people your age can normally do?*
 - b. *During the past 30 days, for about how many days did poor health keep you from doing your usual activities?*
 - c. *In the last seven days, on how many days did you exercise for 30 minutes or more?*
 - d. *Do you have health insurance coverage?*

- Solution
- a. qualitative
 - b. quantitative discrete
 - c. quantitative discrete
 - d. qualitative

- Exercise 35. *In advance of the 1936 Presidential Election, a magazine titled Literary Digest released the results of an opinion poll predicting that the republican candidate Alf Landon would win by a large margin. The magazine sent post cards to approximately 10,000,000 prospective voters. These prospective voters were selected from the subscription list of the magazine, from automobile registration lists, from phone lists, and from club membership lists. Approximately 2,300,000 people returned the postcards.*
- a. *Think about the state of the United States in 1936. Explain why a sample chosen from magazine subscription lists, automobile registration lists, phone books, and club membership lists was not representative of the population of the United States at that time.*
 - b. *What effect does the low response rate have on the reliability of the sample?*
 - c. *Are these problems examples of sampling error or nonsampling error?*
 - d. *During the same year, George Gallup conducted his own poll of 30,000 prospective voters. His researchers used a method they called "quota sampling" to obtain survey answers from specific subsets of the population.*

Quota sampling is an example of which sampling method described in this module?

- Solution
- a. The country was in the middle of the Great Depression and many people could not afford these “luxury” items and therefore not able to be included in the survey.
 - b. Samples that are too small can lead to sampling bias.
 - c. sampling error
 - d. stratified

Exercise 36. *Crime-related and demographic statistics for 47 US states in 1960 were collected from government agencies, including the FBI's Uniform Crime Report. One analysis of this data found a strong connection between education and crime indicating that higher levels of education in a community correspond to higher crime rates. Which of the potential problems with samples discussed in Section 1.2 could explain this connection?*

Solution

Causality: The fact that two variables are related does not guarantee that one variable is influencing the other. We cannot assume that crime rate impacts education level or that education level impacts crime rate.

Confounding: There are many factors that define a community other than education level and crime rate. Communities with high crime rates and high education levels may have other lurking variables that distinguish them from communities with lower crime rates and lower education levels. Because we cannot isolate these variables of interest, we cannot draw valid conclusions about the connection between education and crime. Possible lurking variables include police expenditures, unemployment levels, region, average age, and size.

Exercise 37. *Imagine you work for a polling company and a member of your team has proposed the following survey question:*

“Do you feel happy paying your taxes while some politicians are allowed to use loopholes and avoid paying their fair share of taxes?”

As part of preliminary data collection, 11 people responded to this question. Each participant answered “NO!” Which of the potential problems with samples discussed in this module could explain this connection?

Solution

Self-Selected Samples: Only people who are interested in the topic are choosing to respond. Sample Size Issues: A sample with only 11 participants will not accurately represent the opinions of a nation.

Undue Influence: The question is worded in a specific way to generate a specific response. Self-Funded or Self-Interest Studies: This question was

generated to support one person's claim and it was designed to get the answer that the person desires.

- Exercise 38. *A scholarly article about response rates begins with the following quote: "Declining contact and cooperation rates in random digit dial (RDD) national telephone surveys raise serious concerns about the validity of estimates drawn from such research." The Pew Research Center for People and the Press admits: "The percentage of people we interview – out of all we try to interview – has been declining over the past decade or more."*
- What are some reasons for the decline in response rate over the past decade?*
 - Explain why researchers are concerned with the impact of the declining response rate on public opinion polls.*

Solution

- Possible reasons: increased use of caller id, decreased use of landlines, increased use of private numbers, voice mail, privacy managers, hectic nature of personal schedules, decreased willingness to be interviewed.
- When a large number of people refuse to participate, then the sample may not have the same characteristics of the population. Perhaps the majority of people willing to participate are doing so because they feel strongly about the subject of the survey.

- Exercise 39. *Fifty part-time students were asked how many courses they were taking this term. The (incomplete) results are shown below:*

# of Courses	Frequency	Relative Frequency	Cumulative Relative Frequency
1	30	0.6	
2	15		
3			

Table 1.15 Part-time Student Course Loads

- Fill in the blanks in Table 1.15.*
- What percent of students take exactly two courses?*
- What percent of students take one or two courses?*

Solution

a.

# of Courses	Frequency	Relative Frequency	Cumulative Relative Frequency
1	30	0.6	0.6
2	15	0.3	0.9
3	5	0.1	1.0

- 30%
- 90%

Exercise 40. *Sixty adults with gum disease were asked the number of times per week they used to floss before their diagnosis. The (incomplete) results are shown in Table 1.16.*

# Flossing per Week	Frequency	Relative Frequency	Cumulative Relative Freq.
0	27	0.4500	
1	18		
3			0.9333
6	3	0.0500	
7	1	0.0167	

Table 1.16 Flossing Frequency for Adults with Gum Disease

- Fill in the blanks in Table 1.16.
- What percent of adults flossed six times per week?
- What percent flossed at most three times per week?

Solution

a.

# Flossing per Week	Frequency	Relative Frequency	Cumulative Relative Freq.
0	27	0.4500	0.4500
1	18	0.3000	0.7500
3	11	0.1833	0.9333
6	3	0.0500	0.9833
7	1	0.0167	1

Table 1.40

- 5.00%
- 93.33%

Exercise 41. *Nineteen immigrants to the U.S were asked how many years, to the nearest year, they have lived in the U.S. The data are as follows: 2; 5; 7; 2; 2; 10; 20; 15; 0; 7; 0; 20; 5; 12; 15; 12; 4; 5; 10.*

Table 1.17 was produced.

Data	Frequency	Relative Frequency	Cumulative Relative Frequency
0	2	$\frac{2}{19}$	0.1053
2	3	$\frac{3}{19}$	0.2632
4	1	$\frac{1}{19}$	0.3158
5	3	$\frac{3}{19}$	0.4737

7	2	$\frac{2}{19}$	0.5789
10	2	$\frac{2}{19}$	0.6842
12	2	$\frac{2}{19}$	0.7895
15	1	$\frac{1}{19}$	0.8421
20	1	$\frac{1}{19}$	1.0000

Table 1.17 Frequency of Immigrant Survey Responses

- Fix the errors in Table 1.17. Also, explain how someone might have arrived at the incorrect number(s).
- Explain what is wrong with this statement: “47 percent of the people surveyed have lived in the U.S. for 5 years.”
- Fix the statement in **b** to make it correct.
- What fraction of the people surveyed have lived in the U.S. five or seven years?
- What fraction of the people surveyed have lived in the U.S. at most 12 years?
- What fraction of the people surveyed have lived in the U.S. fewer than 12 years?
- What fraction of the people surveyed have lived in the U.S. from five to 20 years, inclusive?

Solution

- The Frequencies for 15 and 20 should both be two and the Relative Frequencies should both be $\frac{2}{19}$. The mistake could be due to copying the data down wrong. The Cumulative Relative Frequency for five years should be 0.4737. The mistake is due to calculating the Relative Frequency instead of the Cumulative Relative Frequency. The Cumulative Relative Frequency for 15 years should be 0.8947
- The 47% is the Cumulative Relative Frequency, not the Relative Frequency.
- 47% of the people surveyed have lived in the U.S. for five years or less.
- $\frac{5}{19}$
- $\frac{15}{19}$
- $\frac{13}{19}$

g. $\frac{13}{19}$

Exercise 42. *How much time does it take to travel to work? Table 1.18 shows the mean commute time by state for workers at least 16 years old who are not working at home. Find the mean travel time, and round off the answer properly.*

24.0	24.3	25.9	18.9	27.5	17.9	21.8	20.9	16.7	27.3
18.2	24.7	20.0	22.6	23.9	18.0	31.4	22.3	24.0	25.5
24.7	24.6	28.1	24.9	22.6	23.6	23.4	25.7	24.8	25.5
21.2	25.7	23.1	23.0	23.9	26.0	16.3	23.1	21.4	21.5
27.0	27.0	18.6	31.7	23.3	30.1	22.9	23.3	21.7	18.6

Table 1.18

Solution The sum of the travel times is 1,173.1. Divide the sum by 50 to calculate the mean value: 23.462. Because each state's travel time was measured to the nearest tenth, round this calculation to the nearest hundredth: 23.46.

Exercise 43. *A leading business magazine published data on small businesses (defined as businesses that have been publicly traded for at least a year, have a stock price of at least \$5 per share, and have reported annual revenue between \$5 million and \$1 billion). Table 1.19 shows the ages of the chief executive officers for the first 60 ranked small businesses.*

Age	Frequency	Relative Frequency	Cumulative Relative Frequency
40–44	3		
45–49	11		
50–54	13		
55–59	16		
60–64	10		
65–69	6		
70–74	1		

Table 1.19

- What is the frequency for CEO ages between 54 and 65?
- What percentage of CEOs are 65 years or older?
- What is the relative frequency of ages under 50?
- What is the cumulative relative frequency for CEOs younger than 55?
- Which graph shows the relative frequency and which shows the cumulative relative frequency?

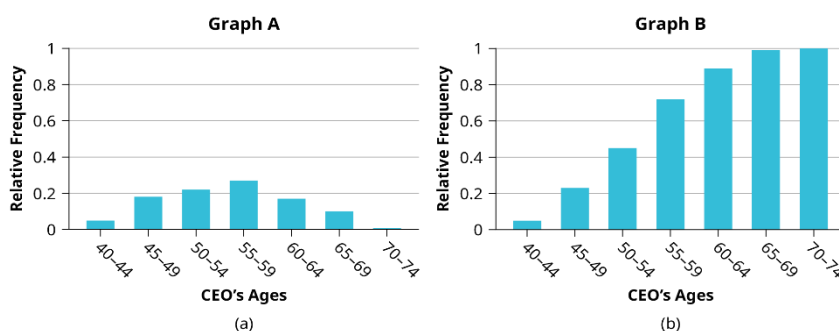


Figure 1.13

- Solution**
- 26 (This is the count of CEOs in the 55 to 59 and 60 to 64 categories.)
 - 12% (number of CEOs age 65 or older ÷ total number of CEOs)
 - 14/60; 0.23; 23%
 - 0.45
 - Graph B represents the cumulative relative frequency, and Graph A shows the relative frequency.

Exercise 44. *Table 1.20 contains data on hurricanes that have made direct hits on the U.S. Between 1851 and 2004. A hurricane is given a strength category rating based on the minimum wind speed generated by the storm.*

Category	Number of Direct Hits	Relative Frequency	Cumulative Frequency
1	109	0.3993	0.3993
2	72	0.2637	0.6630
3	71	0.2601	
4	18		0.9890
5	3	0.0110	1.0000
	Total = 273		

Table 1.20 Frequency of Hurricane Direct Hits

What is the relative frequency of direct hits that were category 4 hurricanes?

- 0.0768
- 0.0659
- 0.2601
- Not enough information to calculate

Solution B

Exercise 45. *Table 1.20 contains data on hurricanes that have made direct hits on the U.S. Between 1851 and 2004. A hurricane is given a strength category rating based on the minimum wind speed generated by the storm.*

Category	Number of Direct Hits	Relative Frequency	Cumulative Frequency
1	109	0.3993	0.3993

2	72	0.2637	0.6630
3	71	0.2601	
4	18		0.9890
5	3	0.0110	1.0000
	Total = 273		

Table 1.20 Frequency of Hurricane Direct Hits

What is the relative frequency of direct hits that were AT MOST a category 3 storm?

- a. 0.3480*
- b. 0.9231*
- c. 0.2601*
- d. 0.3370*

Solution B