INSTRUCTOR'S SOLUTIONS MANUAL

BUSINESS STATISTICS
A DECISION-MAKING APPROACH

TENTH EDITION

David F. Groebner
Boise State University

Patrick W. Shannon
Boise State University

Phillip C. Fry
Boise State University

Pearson

Reproduced by Pearson from electronic files supplied by the author.

Publishing as Pearson, 330 Hudson Street, NY NY 10013

Pearson

ISBN-13: 978-0-13-449642-9
ISBN-10: 0-13-449642-6

Contents

Chapter 1: The Where, Why, and How of Data Collection

Section 1.1

1.1. This application is primarily descriptive in nature. The owner wishes to develop a presentation. She will most likely use charts, graphs, tables and numerical measures to describe her data.

1.2. The graph is a bar chart. A bar chart displays values associated with categories. In this case the categories are the departments at the food store. The values are the total monthly sales (in dollars) in each department. A bar chart also typically has gaps between the bars. A histogram has no gaps and the horizontal axis represents the possible values for a numerical variable.

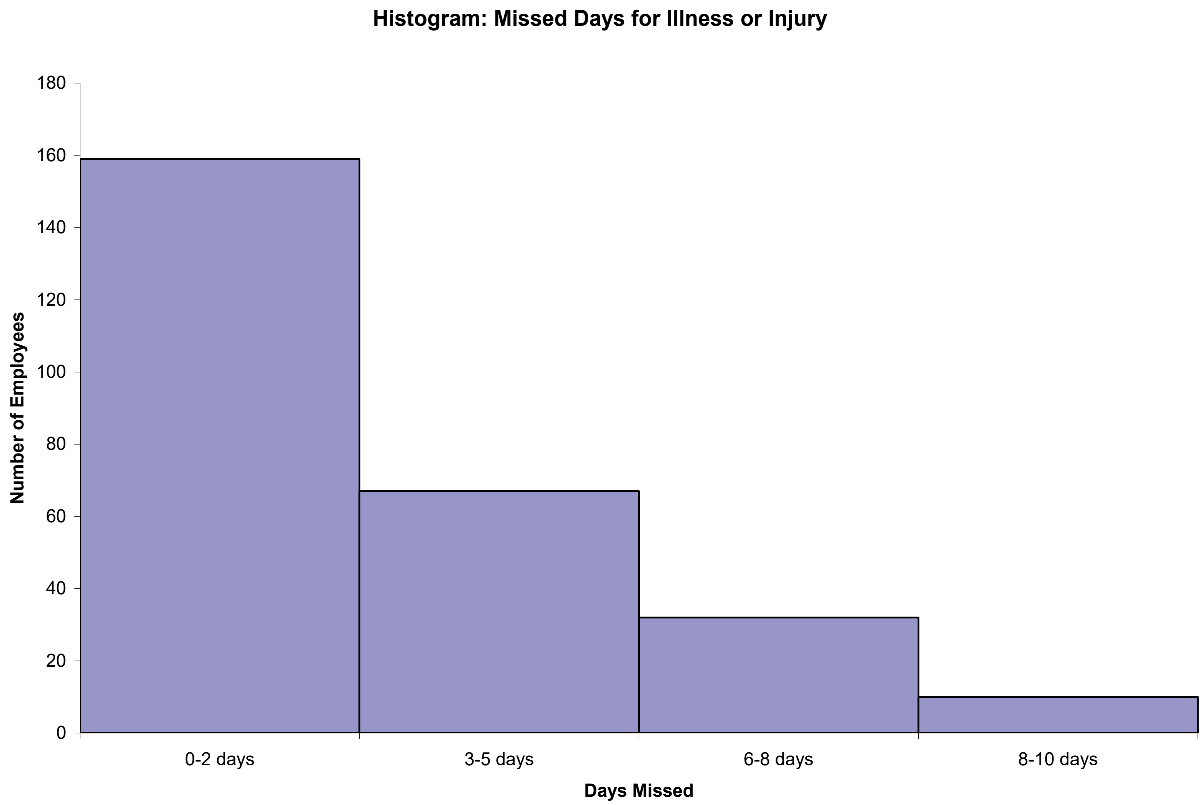
1.3. A bar chart is used whenever you want to display data that has already been categorized while a histogram is used to display data over a range of values for the factor under consideration. Another fundamental difference is that there typically are gaps between the bars on a bar chart but there are no gaps between the bars of a histogram.

1.4. Businesses often make claims about their products that can be tested using hypothesis testing. For example, it is not enough for a pharmaceutical company to claim that its new drug is effective in treating a disease. In order for the drug to be approved by the Food and Drug Administration the company must present sufficient evidence that the drug first does no harm and that it also provides an effective treatment against the disease. The claims that the drug does no harm and is an effective treatment can be tested using hypothesis testing.

1.5. The company could use statistical inference to determine if its parts last longer. Because it is not possible to examine every part that could be produced the company could examine a randomly chosen subset of its parts and compare the average life of the subset to the average life of a randomly chosen subset of the competitor's parts. By using statistical inference procedures the company could reach a conclusion about whether its parts last longer or not.

1.6. Student answers will vary depending on the periodical selected and the periodical's issue date, but should all address the three parts of the question.

1.7. The appropriate chart in this case is a histogram where the horizontal axis contains the number of missed days and the height of the bars represent the number of employees who missed each number of days



Note, there are no gaps between the bars.

1.8. Because it would be too costly, too time consuming, or practically impossible to contact every subscriber to ascertain the desired information, the decision makers at *Fortune* might decide to use statistical inference, particularly estimation, to answer its questions. By looking at a subset of the data and using the procedures of estimation it would be possible for the decision makers to arrive at values for average age and average income that are within tolerable limits of the actual values.

1.9. Student answers will vary depending on the business periodical or newspaper selected and the article referenced. Some representative examples might include estimates of the number of CEO's who will vote for a particular candidate, estimates of the percentage increase in wages for factory workers, estimates of the average dollar advertising expenditures for pharmaceutical companies in a specific year, and the expected increase in R&D expenditures for the coming quarter.

1.10. Student answers will vary. However, the examples should illustrate how statistics has been used and should clearly indicate the type of statistical analysis employed.

Section 1.2

1.11. As discussed in this section, the pet store would most likely use a written survey or a telephone survey to collect the customer satisfaction data.

1.12. A leading question is one that is designed to elicit a specific response, or one that might influence the respondent's answer by its wording. The question is posed so that the respondent believes the researcher has a specific answer in mind when the question is asked, or worded in such a way that the respondent feels obliged to provide an answer consistent with the question. For example, a question such as "Do you agree with the experts who recommend that more tax dollars be given to clean up dangerous and unhealthy pollution?" could cause respondents to provide the answer that they think will be consistent with the "experts" with whom they do not want to disagree. Leading question should be avoided in surveys because they may introduce bias.

1.13. An experiment is any process that generates data as its outcome. The plan for performing the experiment in which the variable of interest is defined is referred to as an experimental design. In the experimental design one or more factors are identified to be changed so that the impact on the variable of interest can be observed or measured.

1.14. There will likely by a high rate of nonresponse bias since many people who work days will not be home during the 9–11 AM time slot. Also, the data collectors need to be careful where they get the phone number list as some people do not have listed phones in phone books and others have no phone or only a cell phone. This may result in selection bias.

1.15. a. Observation would be the most likely method. Observers could be located at various bike routes and observe the number of riders with and without helmets. This would likely be better than asking people if they wear a helmet since the popular response might be to say yes even when they don't always do so.

b. A telephone survey to gas stations in the state. This could be a cost effective way of getting data from across the state. The respondent would have the information and be able to provide the correct price.

c. A written survey of passengers. This could be given out on the plane before the plane lands and passengers could drop the surveys in a box as they de-plane. This method would likely garner higher response rates compared to sending the survey to passengers' mailing address and asking them to return the completed survey by mail.

1.16. The two types of validity mentioned in the section are internal validity and external validity. For this problem external validity is easiest to address. It simply means the sampling method chosen will be sufficient to insure the results based on the sample will be able to be generalized to the population of all students. Internal validity would involve making sure the data gathering method, for instance a questionnaire, accurately determines the respondent's attitude toward the registration process.

1.17. This data could have been collected through a survey. Employees of the USDA could provide periodic reports of fire ant activity in their region. Also, medical reports could be used to collect data assuming people with bites had required medical attention.

1.18. There are many potential sources of bias associated with data collection. If data is to be collected using personal interviews it will be important that the interviewer be trained so that interviewer bias, arising from the way survey questions are asked, is not injected into the survey. If the survey is conducted using either a mail survey or a telephone survey then it is important to be aware of nonresponse bias from those who do not respond to the mailing or refuse to answer your calls. You must also be careful when selecting your survey subjects so that selection bias is not a problem. In order to have useful, reliable data that is representative of the true student opinions regarding campus food service, it is necessary that the data collection process be conducted in a manner that reduces or eliminates the potential for these and other sources of potential bias.

1.19. For retailers technology that scans the product UPC code at checkout makes the collection of data fast and accurate. Retailers that use such technology can automatically update their inventory records and develop an extensive collection of customer buying habits. By applying advanced statistical techniques to the data the retailer can identify relationships among purchases that might otherwise go unnoticed. Such information could enable retailers to target their advertising or even rearrange the placement of products in the store to increase sales. Manufacturing firms use bar code scanning to collect information concerning product availability and product quality. Credit card purchases are automatically tracked by the retailer and the bankcard company. In this way the credit card company is able to track your purchases and even alert you to potential fraud if purchases on your card appear to be unusual. Finally, some companies are using radio frequency identification (RFID) to track products through their supply chain, so that product delays and inventory problems can be minimized.

1.20. One advantage of this form of data gathering is the same as for mail questionnaires. That is low cost. Additional factors being speed of delivery and, with current software, with closed- ended questions, instant updating of data analysis. Disadvantages are also similar, in particular low response and potential confusion about questions. An additional factor might be the ability of competitors to "hack" into the database and analysis program.

1.21. Student answers will vary. Look for clarity of questions and to see that the issue questions are designed to gather useful data. Look for appropriate demographic questions.

1.22. Students should select some form of personal observation as the data-gathering technique. In addition, there should be a discussion of a sampling procedure with an effort made to ensure the sample randomly selected both days of the week unless daily observations are made, and randomly selected times of the day since 24 hour observation would likely be impossible. A complete answer would also address efforts to reduce the potential bias of having an observer standing in an obvious manner by the displays.

1.23. Student answers will vary. However, the issue questions should be designed to gather the desired data regarding customers' preferences for the use of the space. Demographic questions should provide data so that the responses can be broken down appropriately so that United Fitness Center managers can determine which subset of customers have what opinion about this issue. Regarding questionnaire layout, look at neatness and answer location space. Make sure questions are properly worded, used reasonable vocabulary, and are not leading questions.

1.24. The results of the survey are based on telephone interviews with 744 adults, aged 18 and older. Students may also answer that the survey could have been conducted using a written survey via mail questionnaire or internet survey. Because telephone interviews were used to collect the survey data nonresponse biases associated with sampled adults who are not at home when phoned, or adults who refuse to participate in the survey. There is also the problem that some adults do not have a landline phone. If written surveys are used to collect the data then it is important to guard against nonresponse bias from those sampled adults who do not complete the survey There is also the problem of selection bias. In phone interviews we may miss the people who work evenings and nights. If written surveys are used we must be careful to select a representative sample of the adult population.

Section 1.3

1.25. a. Because the population is spread over a large geographical area, a cluster random sample could be selected to reduce travel costs.

b. A stratified random sample would probably be used to keep sample size as small as possible.

c. Most likely a convenience sample would be used since doing a statistical sample would be too difficult.

1.26. To determine the range of employee numbers for the first employee selected in a systematic random sample use the following:

$$\text{Part Range} = \frac{\text{Population Size}}{\text{Sample Size}} = \frac{18{,}000}{100} = 180$$

Thus, the first person selected will come from employees 1–180. Once that person is randomly selected, the second person will be the one numbered 180 higher than the first, and so on.

1.27. Whenever a descriptive numerical measure such as an average is calculated from the entire population it is a parameter. The corresponding measure calculated from a subset of the population, that is to say a sample, is a statistic.

1.28. Statistical sampling techniques consist of those sampling methods that select samples based on chance. Nonstatistical sampling techniques consist of those methods of selecting samples using convenience, judgment, or other nonchance processes. In convenience sampling, samples are chosen because they are easy or convenient to sample. There is no attempt to randomize the selection of the selected items. In convenience sampling not every item in the population has a random chance of being selected. Rather, items are sampled based on their convenience alone. Thus, convenience sampling is not a statistical sampling method.

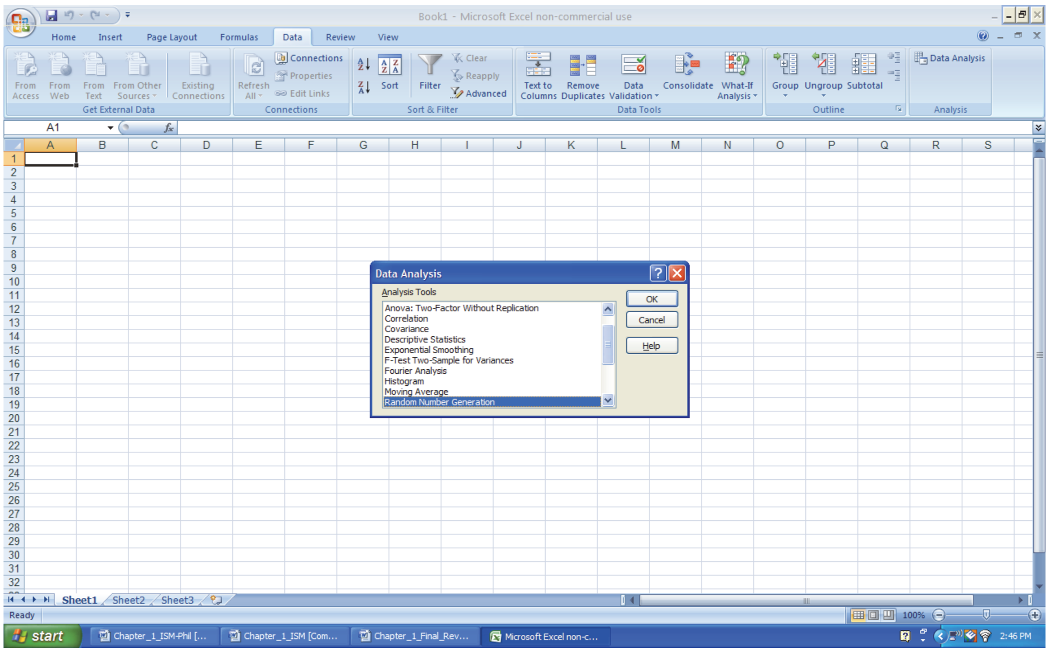
1.29. From a numbered list of all customers who own a certificate of deposit the bank would need to randomly determine a starting point between 1 and k, where k would be equal to 25000/1000 = 25. This could be done using a random number table or by having a statistical package or a spreadsheet generate a random number between 1 and 25. Once this value is determined the bank would select that numbered customer as the first sampled customer and then select every 25th customer after that until 100 customers are sampled.

1.30. A census is an enumeration of the entire set of measurements taken from the population as a whole. While in some cases, the items of interest are obtained from people such as through a survey, in many instances the items of interest come from a product or other inanimate object. For example, a study could be conducted to determine the defect rate for items made on a production line. The census would consist of all items produced on the line in a defined period of time.

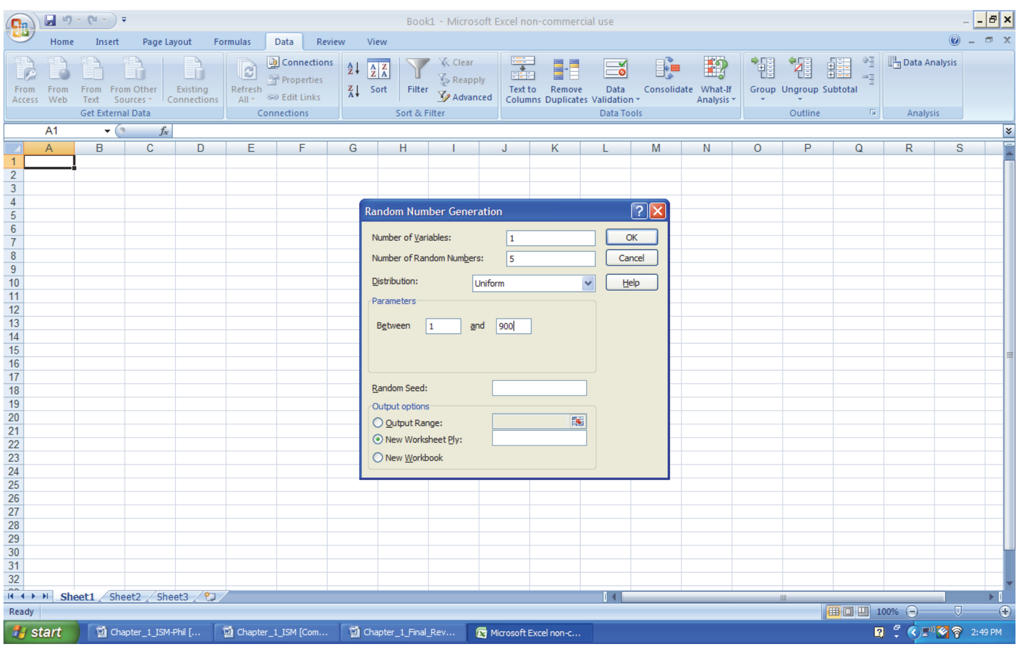
1.31. Values computed from a sample are always considered statistics. In order for a value, such as an average, to be considered a parameter it must be computed from all items in the population.

1.32. In stratified random sampling, the population is divided into homogeneous groups called strata. The idea is to make all items in a stratum as much alike as possible with respect to the variable of interest thereby reducing the number of items that will need to be sampled from each stratum. In cluster sampling, the idea is to break the population into heterogeneous groups called clusters (usually on a geographical basis) such that each cluster looks as much like the original population as possible. Then clusters are randomly selected and from the cluster, individual items are selected using a statistical sampling method.

1.33. Using Excel, choose the Data tab, select Data Analysis from the Analysis Group , then Random Number Generation—shown as follows:



The next step is to complete the random number generation dialog as follows:



The resulting random numbers generated are:

344.4182
91.51183
537.2394
809.2961
796.264

Note, the students' answers may differ since Excel generates different streams of random numbers each time it is used. Also, if the application requires integer numbers, the Decrease Decimal option can be used.

1.34. If these percentages were based on all students attending college in those years they would be parameters, if the percentages were based on a sample they would be statistics.

1.35. This is a statistic. A poll would be a sample of eligible voters rather than all eligible voters.

1.36. Solution

a. Stratified random sampling

b. Simple random sampling or possibly cluster random sampling

c. Systematic random sampling

d. Stratified random sampling

1.37. This is a statistical sample. Every employee has an equal chance of being selected using this method. In fact, this is an example of a simple random sample because every possible sample of size 50 has an equal chance of being selected.

1.38. a. Student answers will vary

b. Cluster sampling could be used to ensure that you get all types of cereal. Make each cluster the area where certain cereals are located (i.e., isle, row, shelf, etc.)

c. Cluster sampling would give you a better idea of the inventory of all types of cereal. Simple random sampling could possibly end up with only looking at 2 or 3 cereal types.

1.39. Students should choose the Data tab, select Data Analysis from the Analysis group—Random Number Generation process. Students' answers will differ since Excel generates different streams of random numbers each time it is used, but 40 random numbers should be generated from a uniform distribution with values ranging from 1 to 578. Since the application requires integer numbers, the Decrease Decimal option should be used.

1.40. a. The population should be all users of cross-country ski lots and trailheads in Colorado.

b. Several sampling techniques could be selected. Be sure that some method of ensuring randomness is discussed. In addition, some students might give greater weight to frequent users of the lots. In which case the population would really be user days rather than individual users.

c. Students using Excel should choose the Data tab, select Data Analysis from the Analysis group—Random Number Generation process. Students' answers may differ since Excel generates different streams of random numbers each time it is used. Since the application requires integer numbers, the Decrease Decimal option should be used.

1.41. a. Since there are 4,000 patient files we could give each file a unique identification number consisting of 4 digits. The first file would be given the identification number "0001." The last file would be given the identification number of "4000." By assigning each patient a number and randomly selecting the 100 numbers allows each possible sample of 100 an equal chance of being selected.

b. Either use a random number table (randomly select the starting row and column), or use a computer program, such as Microsoft Excel, which has a random number generator.

c. Since each patient is assigned a 4-digit identification number, we would need a 4-digit random number for each random number selected.

d. Answers will vary.

Section 1.4

1.42. a. Time-series
b. Cross-sectional
c. Time-series
d. Cross-sectional

1.43. Qualitative data are categories or numerical values that represent categories. Quantitative data is data that is purely numerical.

1.44. a. Ordinal—categories with defined order
b. Nominal—categories with no defined order
c. Ratio
d. Nominal—categories with no defined order

1.45. Nominal data involves placing observations in separate categories according to some measurable characteristic. Ordinal data also involves placing observations into separate categories, but the categories can be rank-ordered.

1.46. Since the circles involve a ranking from best to worst, this would be ordinal data.

1.47. a. The data are cross-sectional. The data are collected from 2,300 customers at approximately the same point in time
b. This is a ratio level, quantitative variable. The data represent a measurement of time.
c. Ordinal with a numerical value representing customers rating of level of service

1.48. a. Nominal Data
b. Ratio Data
c. Ratio Data
d. Ratio Data
e. Nominal

1.49. a. Cross-sectional
b. Time-series
c. Cross-sectional
d. Cross-sectional
e. Time-Series

1.50. Columns A–G are nominal—they are all codes
Columns H–L are ratio level.

End of Chapter Exercises

1.51. Answers will vary with the student. But a good discussion should include the following factors:
Sampling techniques and possible problems selecting a representative sample.
Determining how to develop questions to measure approval.
Structuring questions to avoid bias.
The measurement scale associated with the questions.
The fact these polls tend to develop time-series data.

1.52. Nominal data or ordinal data.

1.53. Interval or ratio data.

1.54. Ratings are typical uses of ordinal scale data. And since ratings are based on personal opinion, even though people are using the same scale, a direct comparison between the two ratings is not possible. This is a common problem when people are asked to rate an object using an ordinal scale.

1.55. Answers will vary with the student. But a good discussion should include the following factors:

Sampling techniques and possible problems selecting a representative sample.

Determining how to measure confidence.

Structuring questions to avoid bias.

The measurement scale associated with the questions.

The fact this poll is specifically intended to develop time-series data.

1.56. Answers will vary with the student.

1.57. Answers will vary with the student.

1.58. a. No because a random sample means that every item in the population has an equal chance of being selected. Individuals who do not have or use email do not have an equal chance of being included in this survey. Also, volunteer emails would not be random.

b. In this survey the biggest drawback is that only individuals with strong feelings one way or the other are apt to respond to this survey. This could lead to a great deal of bias in the results of the survey. Another big problem with a survey is nonresponse bais. Again because they are requesting viewers to write in there will be a great deal of nonresponse to this survey. I would also include in the answer that the question being asked is somewhat leading. The phrase "using too much force in routine traffic stops" implies that, in fact, force is being used which one would not expect in a routine traffic stop.

1.59. a. They would probably want to sample the salsa jars as they come off the assembly line at the plant for a specified time period. They would want to use a random sample. One method would be to take a systematic random sample. They could then calculate the percentage of the sample that had an unacceptable thickness.

b. The product is going to be ruined after testing it. You would not want to ruin the entire product that comes off the assembly line.

1.60. a. Student answers will vary but one method would be personal observation at grocery stores or another method would be to simply look at their sales. Are buyers of the energy drinks purchasing bottles or cans?

b. If using personal observation just have people at grocery stores observe people over a specified period of time and note which are selecting cans and which are selecting bottles and look at the percentages of each.

c. You would be looking at ratio data because you could have a true 0 if, for example, no one purchased bottles.

d. Depends on the way the data are collected. Sales data would be quantitative.

1.61. a. The fact that the friend has selected his favorite players means that all players did not have a chance of being selected in the sample. The sample would be biased toward the type of players the friend favors.

b. One method would be to obtain a list of all NBA players. Then assign each player a number. Then you could use Excel's random number generator to obtain a random sample of 40 players from the list.

1.62. The appropriate design would be a stratified random sampling method. Start by dividing the students into class standing (Freshman, Sophomore, Junior, and Senior). Then randomly select students from each strata.

Chapter 2: Graphs, Charts, and Tables—Describing Your Data

When applicable, the first few problems in each section will be done following the appropriate step by step procedures outlined in the corresponding sections of the chapter. Following problems will provide key points and the answers to the questions, but all answers can be arrived at using the appropriate steps.

Section 2.1

2.1. Step 1: List the possible values.

The possible values for the discrete variable are 0 through 12.

Step 2: Count the number of occurrences at each value.

The resulting frequency distribution is shown as follows:

x	Frequency
0	1
1	0
2	2
3	4
4	1
5	2
6	5
7	6
8	1
9	1
10	1
11	0
12	1
Total =	25

2.2. Given $n = 2{,}000$, the minimum number of groups for a grouped data frequency distribution determined using the $2^k \geq n$ guideline is:

$$2^k \geq n \text{ or } 2^{11} = 2{,}048 \geq 2{,}000. \text{ Thus, use } k = 11 \text{ groups.}$$

2.3. a. Given $n = 1{,}000$, the minimum number of classes for a grouped data frequency distribution determined using the $2^k \geq n$ guideline is:

$$2^k \geq n \text{ or } 2^{10} = 1{,}024 \geq 1{,}000. \text{ Thus, use } k = 10 \text{ classes.}$$

b. Assuming that the number of classes that will be used is 10, the class width is determined as follows:

$$w = \frac{\text{High} - \text{Low}}{\text{Classes}} = \frac{2{,}900 - 300}{10} = \frac{2{,}600}{10} = 260$$

Then we round to the nearest 100 points giving a class width of 300.

2.4. Recall that the Ogive is produced by plotting the cumulative relative frequency against the upper limit of each class. Thus, the first class upper limit is 100 and has a relative frequency of $0.2 - 0.0 = 0.2$. The second class upper limit is 200 and has a relative frequency of $0.4 - 0.2 = 0.2$. Of course, the frequencies are obtained by multiplying the relative frequency by the sample size. As an example, the first class has a frequency of $(0.2)50 = 10$. The others follow similarly to produce the following distribution

Class	Frequency	Relative Frequency	Cumulative Relative Frequency
0 – < 100	10	0.20	0.20
100 – < 200	10	0.20	0.40
200 – < 300	5	0.10	0.50
300 – < 400	5	0.10	0.60
400 – < 500	20	0.40	1.00
500 – < 600	0	0.00	1.00

2.5. a. There are $n = 60$ observations in the data set. Using the $2^k > n$ guideline, the number of classes, *k*, would be 6. The maximum and minimum values in the data set are 17 and 0, respectively. The class width is computed to be: $w = (17-0)/66 = 2.833,$ which is rounded to 3. The frequency distribution is

Class	Frequency
0–2	6
3–5	13
6–8	20
9–11	14
12–14	5
15–17	2
	Total = 60

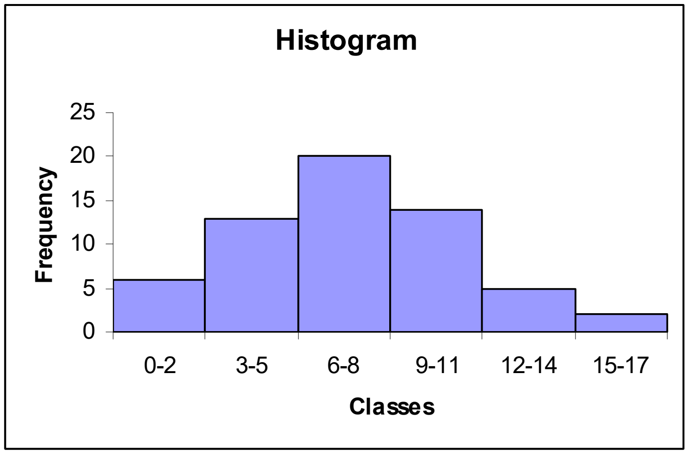
b. To construct the relative frequency distribution divide the number of occurrences (frequency) in each class by the total number of occurrences. The relative frequency distribution is shown below.

Class	Frequency	Relative Frequency
0–2	6	0.100
3–5	13	0.217
6–8	20	0.333
9–11	14	0.233
12–14	5	0.083
15–17	2	0.033
	Total = 60	

c. To develop the cumulative frequency distribution, compute a running sum for each class by adding the frequency for that class to the frequencies for all classes above it. The cumulative relative frequencies are computed by dividing the cumulative frequency for each class by the total number of observations. The cumulative frequency and the cumulative relative frequency distributions are shown below.

Class	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
0–2	6	0.100	6	0.100
3–5	13	0.217	19	0.317
6–8	20	0.333	39	0.650
9–11	14	0.233	53	0.883
12–14	5	0.083	58	0.967
15–17	2	0.033	60	1.000
	Total = 60			

d. To develop the histogram, first construct a frequency distribution (see part a). The classes form the horizontal axis and the frequency forms the vertical axis. Bars corresponding to the frequency of each class are developed. The histogram based on the frequency distribution from part (a) is shown below.



2.6. a. Proportion of days in which no shortages occurred = 1 – proportion of days in which shortages occurred $= 1 - 0.24 = 0.76$.

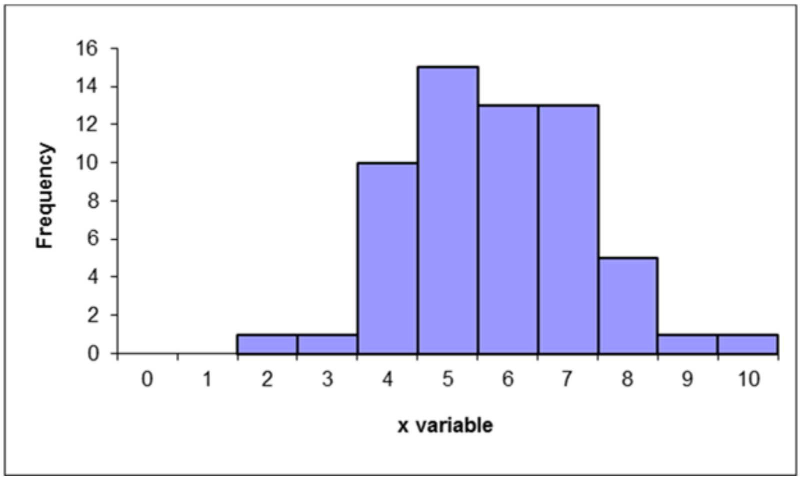
b. Less than \$20 off implies that overage was less than \$20 and the shortage was less than \$20 = (proportion of overages less \$20) – (proportion of shortages at most \$20) $= 0.56 - 0.08 = 0.48$.

c. Proportion of days with less than \$40 over or at most \$20 short = Proportion of days with less than \$40 over – proportion of days with more than \$20 short $= 0.96 - 0.08 = 0.88$.

2.7. a. The data do not require grouping. The following frequency distribution is given:

x	Frequency
0	0
1	0
2	1
3	1
4	10
5	15
6	13
7	13
8	5
9	1
10	1

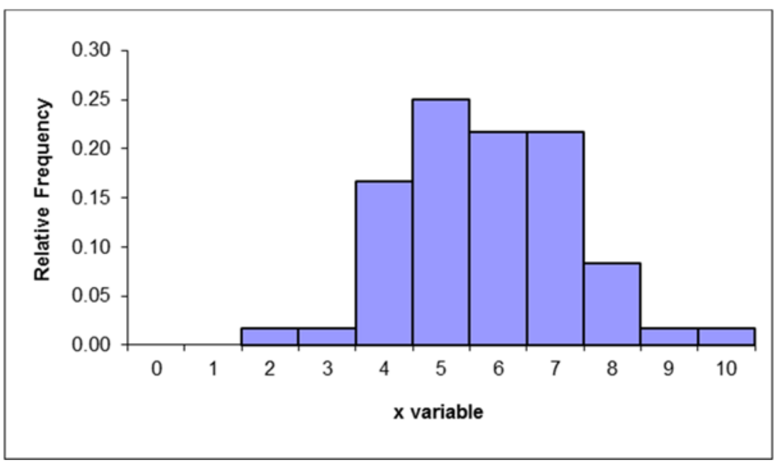
b. The following histogram could be developed.



c. The relative frequency distribution shows the fraction of values falling at each value of *x*.

x	Frequency	Relative Frequency
0	0	0.00
1	0	0.00
2	1	0.02
3	1	0.02
4	10	0.17
5	15	0.25
6	13	0.22
7	13	0.22
8	5	0.08
9	1	0.02
10	1	0.02
	60	

d. The relative frequency histogram is shown below.



e. The two histograms look exactly alike since the same data are being graphed. The bars represent either the frequency or relative frequency.

2.8. a. Step 1 and Step 2: Group the data into classes and determine the class width:
The problem asks you to group the data. Using the $2^k \geq n$ guideline we get:
$2^k \geq 60$ so $2^6 \geq 60$

Class width is:

$$W = \frac{\text{Maximum} - \text{Minumum}}{\text{Number of Classes}} = \frac{10-2}{6} = 1.33$$

which we round up to 2.0

Step 3: Define the class boundaries:

Since the data are discrete, the classes are:

Class
2–3
4–5
6–7
8–9
10–11

Step 4: Count the number of values in each class:

Class	Frequency	Relative Frequency
2–3	2	0.0333
4–5	25	0.4167
6–7	26	0.4333
8–9	6	0.1000
10–11	1	0.0167

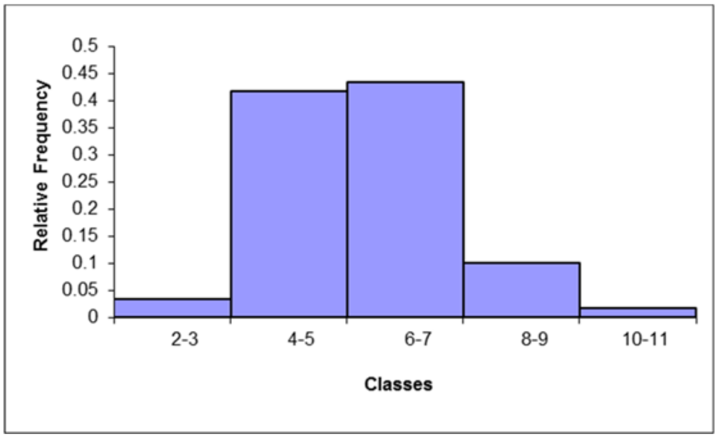
b. The cumulative frequency distribution is:

Class	Frequency	Cumulative Frequency
2–3	2	2
4–5	25	27
6–7	26	53
8–9	6	59
10–11	1	60

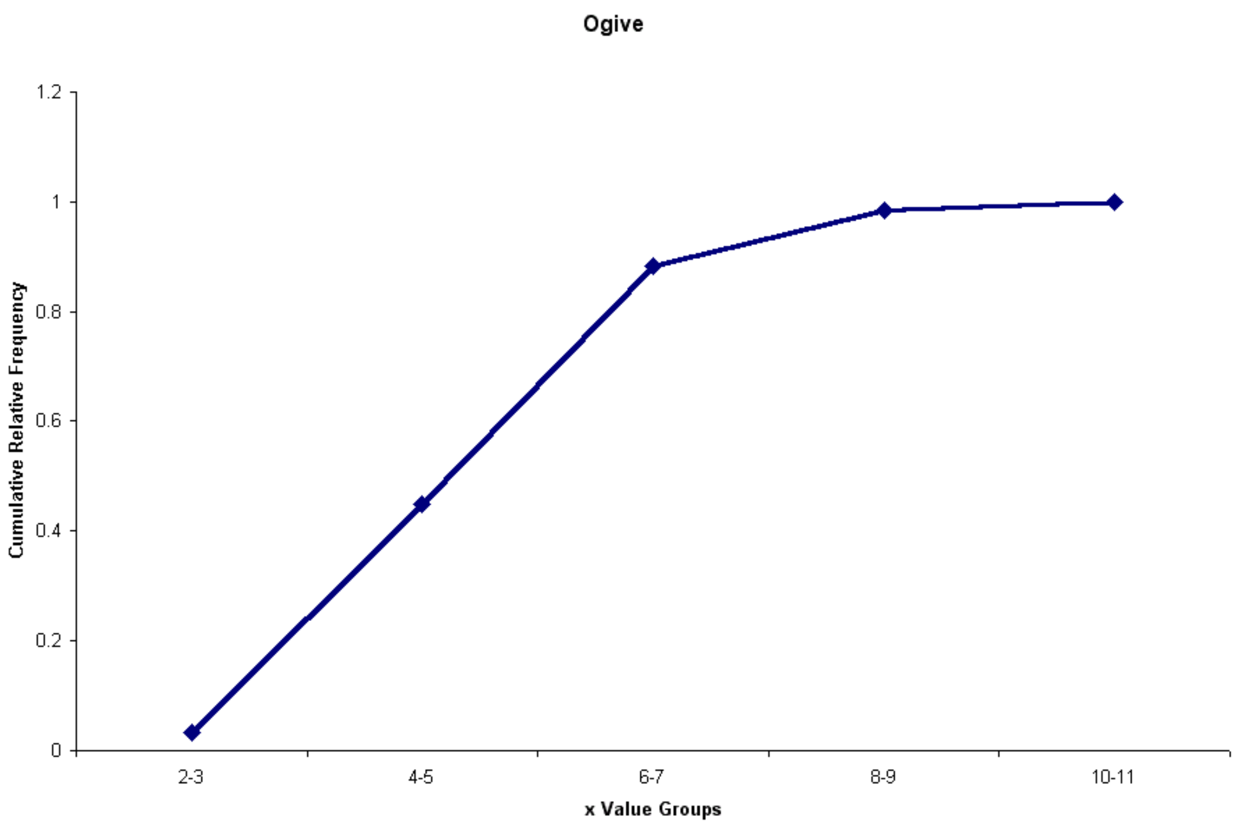
c.

Class	Frequency	Relative Frequency	Cumulative Relative Frequency
2–3	2	0.0333	0.0333
4–5	25	0.4167	0.4500
6–7	26	0.4333	0.8833
8–9	6	0.1000	0.9833
10–11	1	0.0167	1.000

The relative frequency histogram is:



d. The ogive is a graph of the cumulative relative frequency distribution.

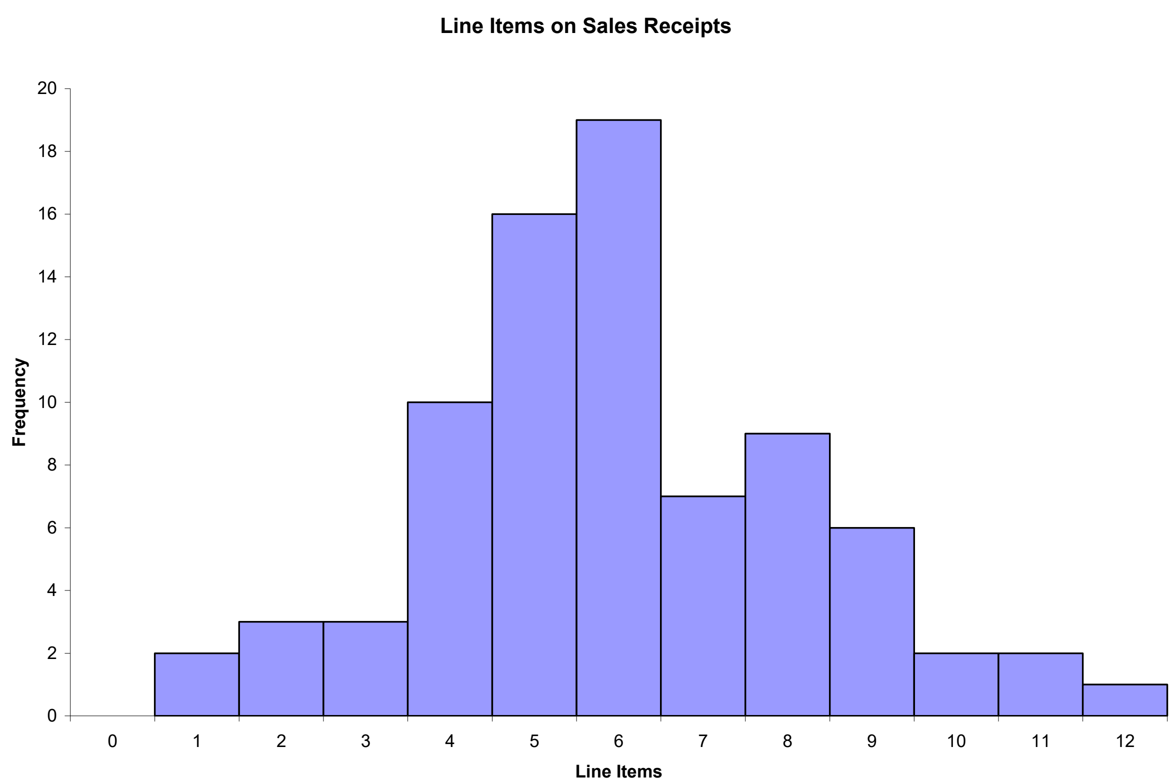


2.9. a. Because the number of possible values for the variable is relatively small, there is no need to group the data into classes. The resulting frequency distribution is:

x	Frequency
0	0
1	2
2	3
3	3
4	10
5	16
6	19
7	7
8	9
9	6
10	2
11	2
12	1
Total =	80

This frequency distribution shows the manager that most customer receipts have 4 to 8 line items.

b. A histogram is a graph of a frequency distribution for a quantitative variable. The resulting histogram is shown as follows.



2.10. a.

	Knowledge Level			
	Savvy	Experienced	Novice	Total
Online Investors	32	220	148	400
Traditional Investors	8	58	134	200
	40	278	282	600

b.

	Knowledge Level		
	Savvy	Experienced	Novice
Online Investors	0.0533	0.3667	0.2467
Traditional Investors	0.0133	0.0967	0.2233

c. The proportion that were both on-line and experienced is 0.3667.

d. The proportion of on-line investors is 0.6667

2.11. a. The following relative frequency distributions are developed for the two variables:

Rating	Frequency	Rel. Frequency
1	5	0.25
2	8	0.40
3	4	0.20
4	2	0.10
5	1	0.05
Total =	20	

Time Slot	Frequency	Rel. Frequency
1	9	0.45
2	3	0.15
3	5	0.25
4	3	0.15
Total =	20	

b. The joint frequency distribution is a two dimensional table showing responses to the rating on one dimension and time slot on the other dimension. This joint frequency distribution is shown as follows:

Rating	Morning	Afternoon	Evening	Various	Total
Very Good	5				5
Good	4	3		1	8
Fair			3	1	4
Poor			1	1	2
Very Poor			1		1
Total	9	3	5	3	20

c. The joint relative frequency distribution is determined by dividing each frequency by the sample size, 20. This is shown as follows:

Rating	Morning	Afternoon	Evening	Various	Total
Very Good	0.25	0.00	0.00	0.00	0.25
Good	0.20	0.15	0.00	0.05	0.40
Fair	0.00	0.00	0.15	0.05	0.20
Poor	0.00	0.00	0.05	0.05	0.10
Very Poor	0.00	0.00	0.05	0.00	0.05
Total	0.45	0.15	0.25	0.15	1.00

Based on the joint relative frequency distribution, we see that those who advertise in the morning tend to provide higher service ratings. Evening advertisers tend to provide lower ratings. The manager may wish to examine the situation further to see why this occurs.

2.12. a. The weights are sorted from smallest to largest to create the data array.

77	79	80	83	84	85	86
86	86	86	86	86	87	87
87	88	88	88	88	89	89
89	89	89	90	90	91	91
92	92	92	92	93	93	93
94	94	94	94	94	95	95
95	96	97	98	98	99	101

b. Five classes having equal widths are created by subtracting the smallest observed value (77) from the largest value (101) and dividing the difference by 5 to get the width for each class (4.8 rounded to 5). Five classes of width five are then constructed such that the classes are mutually exclusive and all inclusive. Identify the variable of interest. The weight of each crate is the variable of interest. The number of crates in each class is then counted. The frequency table is shown below.

Weight (Classes)	Frequency
77–81	3
82–86	9
87–91	16
92–96	16
97–101	5
	Total = 49

c. The histogram can be created from the frequency distribution. The classes are shown on the horizontal axis and the frequency on the vertical axis. The histogram is shown below.