Table of Contents

CHAPTER 1

1.1 The type of beverage sold yields categorical or "qualitative" responses.
The type of beverage sold yields distinct categories in which no ordering is implied.

1.2 Three sizes of U.S. businesses are classified into distinct categories—small, medium, and large—in which order is implied.

1.3 The time it takes to download a video from the Internet is a continuous numerical or "quantitative" variable because time can have any value from 0 to any reasonable unit of time.

1.4
(a) The number of cellphones is a numerical variable that is discrete because the outcome is a count.
(b) Monthly data usage is a numerical variable that is continuous because any value within a range of values can occur.
(c) Number of text messages exchanged per month is a numerical variable that is discrete because the outcome is a count.
(d) Voice usage per month is a numerical variable that is continuous because any value within a range of values can occur.
(e) Whether a cellphone is used for email is a categorical variable because the answer can be only yes or no.

1.5
(a) numerical, continuous
(b) numerical, discrete
(c) categorical
(d) categorical

1.6
(a) Categorical
(b) Numerical, continuous
(c) Categorical
(d) Numerical, discrete
(e) Categorical

1.7
(a) numerical, continuous *
(b) categorical
(c) categorical
(d) numerical, discrete
*Some researchers consider money as a discrete numerical variable because it can be "counted."

1.8
(a) numerical, continuous *
(b) numerical, discrete
(c) numerical, continuous *
(d) categorical
*Some researchers consider money as a discrete numerical variable because it can be "counted."

1.9 (a) Income may be considered discrete if we "count" our money. It may be considered continuous if we "measure" our money; we are only limited by the way a country's monetary system treats its currency.

(b) The first format is preferred because the responses represent data measured on a higher scale.

1.10 The underlying variable, ability of the students, may be continuous, but the measuring device, the test, does not have enough precision to distinguish between the two students.

1.11 (a) The population is "all working women from the metropolitan area." A systematic or random sample could be taken of women from the metropolitan area. The director might wish to collect both numerical and categorical data.

(b) Three categorical questions might be occupation, marital status, type of clothing. Numerical questions might be age, average monthly hours shopping for clothing, income.

1.12 The answer depends on the chosen data set.

1.13 The answer depends on the specific story.

1.14 The answer depends on the specific story.

1.15 The transportation engineers and planners should use primary data collected through an observational study of the driving characteristics of drivers over the course of a month.

1.16 The information presented there is based mainly on a mixture of data distributed by an organization and data collected by ongoing business activities.

1.17 (a) 001 (b) 040 (c) 902

1.18 Sample without replacement: Read from left to right in 3-digit sequences and continue unfinished sequences from end of row to beginning of next row.

Row 05: 338 505 855 551 438 855 077 186 579 488 767 833 170
Rows 05-06: 897
Row 06: 340 033 648 847 204 334 639 193 639 411 095 924
Rows 06-07: 707
Row 07: 054 329 776 100 871 007 255 980 646 886 823 920 461
Row 08: 893 829 380 900 796 959 453 410 181 277 660 908 887
Rows 08-09: 237
Row 09: 818 721 426 714 050 785 223 801 670 353 362 449
Rows 09-10: 406
Note: All sequences above 902 and duplicates are discarded.

1.19 (a) Row 29: 12 47 83 76 22 99 65 93 10 65 83 61 36 98 89 58 86 92 71
Note: All sequences above 93 and all repeating sequences are discarded.

(b) Row 29: 12 47 83 76 22 99 65 93 10 65 83 61 36 98 89 58 86
Note: All sequences above 93 are discarded. Elements 65 and 83 are repeated.

1.20 A simple random sample would be less practical for personal interviews because of travel costs (unless interviewees are paid to attend a central interviewing location).

1.21 This is a probability sample because the selection is based on chance. It is not a simple random sample because A is more likely to be selected than B or C.

1.22 Here all members of the population are equally likely to be selected and the sample selection mechanism is based on chance. But not every sample of size 2 has the same chance of being selected. For example the sample "B and C" is impossible.

1.23 (a) Since a complete roster of full-time students exists, a simple random sample of 200 students could be taken. If student satisfaction with the quality of campus life randomly fluctuates across the student body, a systematic 1-in-20 sample could also be taken from the population frame. If student satisfaction with the quality of life may differ by gender and by experience/class level, a stratified sample using eight strata, female freshmen through female seniors and male freshmen through male seniors, could be selected. If student satisfaction with the quality of life is thought to fluctuate as much within clusters as between them, a cluster sample could be taken.

(b) A simple random sample is one of the simplest to select. The population frame is the registrar's file of 4,000 student names.

(c) A systematic sample is easier to select by hand from the registrar's records than a simple random sample, since an initial person at random is selected and then every 20th person thereafter would be sampled. The systematic sample would have the additional benefit that the alphabetic distribution of sampled students' names would be more comparable to the alphabetic distribution of student names in the campus population.

(d) If rosters by gender and class designations are readily available, a stratified sample should be taken. Since student satisfaction with the quality of life may indeed differ by gender and class level, the use of a stratified sampling design will not only ensure all strata are represented in the sample, it will also generate a more representative sample and produce estimates of the population parameter that have greater precision.

(e) If all 4,000 full-time students reside in one of 10 on-campus residence halls which fully integrate students by gender and by class, a cluster sample should be taken. A cluster could be defined as an entire residence hall, and the students of a single randomly selected residence hall could be sampled. Since each dormitory has 400 students, a systematic sample of 200 students can then be selected from the chosen cluster of 400 students. Alternately, a cluster could be defined as a floor of one of the 10 dormitories. Suppose there are four floors in each dormitory with 100 students on each floor. Two floors could be randomly sampled to produce the required 200 student sample. Selection of an entire dormitory may make distribution and collection of the survey easier to accomplish. In contrast, if there is some variable other than gender or class that differs across dormitories, sampling by floor may produce a more representative sample.

1.24 (a) Row 16: 2323 6737 5131 8888 1718 0654 6832 4647 6510 4877
Row 17: 4579 4269 2615 1308 2455 7830 5550 5852 5514 7182
Row 18: 0989 3205 0514 2256 8514 4642 7567 8896 2977 8822
Row 19: 5438 2745 9891 4991 4523 6847 9276 8646 1628 3554
Row 20: 9475 0899 2337 0892 0048 8033 6945 9826 9403 6858
Row 21: 7029 7341 3553 1403 3340 4205 0823 4144 1048 2949
Row 22: 8515 7479 5432 9792 6575 5760 0408 8112 2507 3742
Row 23: 1110 0023 4012 8607 4697 9664 4894 3928 7072 5815
Row 24: 3687 1507 7530 5925 7143 1738 1688 5625 8533 5041
Row 25: 2391 3483 5763 3081 6090 5169 0546
Note: All sequences above 5000 are discarded. There were no repeating sequences.

(b) 089 189 289 389 489 589 689 789 889 989
1089 1189 1289 1389 1489 1589 1689 1789 1889 1989
2089 2189 2289 2389 2489 2589 2689 2789 2889 2989
3089 3189 3289 3389 3489 3589 3689 3789 3889 3989
4089 4189 4289 4389 4489 4589 4689 4789 4889 4989

(c) With the single exception of invoice #0989, the invoices selected in the simple random sample are not the same as those selected in the systematic sample. It would be highly unlikely that a random process would select the same units as a systematic process.

1.25 (a) A stratified sample should be taken so that each of the three strata will be proportionately represented.

(b) The number of observations in each of the three strata out of the total of 1,000 should reflect the proportion of the three categories in the customer database. For example, 3,500/10,000 = 35% so 35% of 1,000 = 350 customers should be selected from the prospective buyers; similarly 4,500/10,000 = 45% so 450 customers should be selected from the first time buyers, and 2,000/10,000 = 20% so 200 customers from the repeat buyers.

(c) It is not simple random sampling because, unlike the simple random sampling, it ensures proportionate representation across the entire population.

1.26 Before accepting the results of a survey of college students, you might want to know, for example:
Who funded the survey? Why was it conducted? What was the population from which the sample was selected? What sampling design was used? What mode of response was used: a personal interview, a telephone interview, or a mail survey? Were interviewers trained? Were survey questions field-tested? What questions were asked? Were they clear, accurate, unbiased, valid? What operational definition of "vast majority" was used? What was the response rate? What was the sample size?

1.27 (a) Possible coverage error: Only employees in a specific division of the company were sampled.

(b) Possible nonresponse error: No attempt is made to contact nonrespondents to urge them to complete the evaluation of job satisfaction.

(c) Possible sampling error: The sample statistics obtained from the sample will not be equal to the parameters of interest in the population.

(d) Possible measurement error: Ambiguous wording in questions asked on the questionnaire.

1.28 The results are based on an online survey. If the frame is supposed to be small business owners, how is the population defined? This is a self-selecting sample of people who responded online, so there is an undefined nonresponse error. Sampling error cannot be determined since this is not a random sample.

1.29 Before accepting the results of the survey, you might want to know, for example:
Who funded the study? Why was it conducted? What was the population from which the sample was selected? What was the frame being used? What sampling design was used?
What mode of response was used: a personal interview, a telephone interview, or a mail survey? Were interviewers trained? Were survey questions field-tested? What other questions were asked? Were they clear, accurate, unbiased, and valid? What was the response rate? What was the margin of error? What was the sample size?

1.30 Before accepting the results of the survey, you might want to know, for example: Who funded the study? Why was it conducted? What was the population from which the sample was selected? What sampling design was used? What mode of response was used: a personal interview, a telephone interview, or a mail survey? Were interviewers trained? Were survey questions field-tested? What other questions were asked? Were the questions clear, accurate, unbiased, and valid? What was the response rate? What was the margin of error? What was the sample size? What frame was used?

1.31 A population contains all the items of interest whereas a sample contains only a portion of the items in the population.

1.32 A statistic is a summary measure describing a sample whereas a parameter is a summary measure describing an entire population.

1.33 Categorical random variables yield categorical responses such as yes or no answers. Numerical random variables yield numerical responses such as your height in inches.

1.34 Discrete random variables produce numerical responses that arise from a counting process. Continuous random variables produce numerical responses that arise from a measuring process.

1.35 Items or individuals in a probability sampling are selected based on known probabilities while items or individuals in a nonprobability samplings are selected without knowing their probabilities of selection.

1.36 Microsoft Excel could be used to perform various statistical computations that were possible only with a slide-rule or hand-held calculator in the old days.

1.37 (a) The population of interest was 18-54 year olds who currently own a smartphone and/or tablet, and who use and do not use these devices to shop.
(b) The sample was the 1,003 18-54 year olds who currently own a smartphone and/or tablet, who use and do not use these devices to shop, and who responded to the study.
(c) A parameter of interest is the proportion of all tablet users in the population who use their device to purchase product and services.
(d) A statistic used to estimate the parameter of interest in (c) is the proportion of tablet users in the sample who use their device to purchase product and services.

1.38 The answers to this question depend on which article and its corresponding data set is being selected.

1.39 (a) The population of interest was supply chain executives in a wide range of industries representing a mix of company sizes from across three global regions: Asia, Europe, and the Americas.
(b) The sample was the 503 supply chain executives in a wide range of industries representing a mix of company sizes from across three global regions: Asia, Europe, and the Americas surveyed by PwC from May to July 2012.
(c) A parameter of interest is the proportion of supply chain executives in the population who acknowledge that supply chain is seen as a strategic asset in their company.
(d) A statistic used to estimate the parameter of interest in (c) is the proportion of supply chain executives in the sample who acknowledge that supply chain is seen as a strategic asset in their company.

1.40 The answers to this question depend on which data set is being selected.

1.41 (a) Categorical variable: Which of the following best describes this firm's primary business?
(b) Numerical variable: On average, what percent of total monthly revenues are e-commerce revenues?

1.42 (a) The population of interest was the collection of all the 10,000 benefitted employees at the University of Utah when the study was conducted.
(b) The sample consisted of the 3,095 benefitted employees participated in the study.
(c) gender: categorical; age: numerical; education level: numerical; marital status: categorical; household income: numerical; employment category: categorical

1.43 (a) (i) categorical (iii) numerical, discrete
(ii) categorical (iv) categorical
(b) The answers will vary.
(c) The answers will vary.

CHAPTER 2

2.1 (a)

Category	Frequency	Percentage
A	13	26%
B	28	56
C	9	18

(b) Category "B" is the majority.

2.2 (a) Table frequencies for all student responses

	Student Major Categories			
Gender	A	C	M	Totals
Male	14	9	2	25
Female	6	6	3	15
Totals	20	15	5	40

(b) Table percentages based on overall student responses

	Student Major Categories			
Gender	A	C	M	Totals
Male	35.0%	22.5%	5.0%	62.5%
Female	15.0%	15.0%	7.5%	37.5%
Totals	50.0%	37.5%	12.5%	100.0%

Table based on row percentages

	Student Major Categories			
Gender	A	C	M	Totals
Male	56.0%	36.0%	8.0%	100.0%
Female	40.0%	40.0%	20.0%	100.0%
Totals	50.0%	37.5%	12.5%	100.0%

Table based on column percentages

	Student Major Categories			
Gender	A	C	M	Totals
Male	70.0%	60.0%	40.0%	62.5%
Female	30.0%	40.0%	60.0%	37.5%
Totals	100.0%	100.0%	100.0%	100.0%

2.3 (a) You can conclude that Android smartphones have seen steady increase in market shares while Blackberry and Other OS smartphones have seen steady decrease in market shares since 2011. Android smartphones dominated the market in all those three years.

(b) The iOS smartphones have overtaken Other OS smartphones and owned the second largest market share since 2012. The Microsoft smartphones have arisen to the third place in terms of market share in 2013 from the fifth place position in 2011 while the Other OS smartphones have dropped from the second place in 2011 to the last place in 2013.

2.4 (a) The percentage of complaints for each automaker:

Automaker	Frequency	Percentage	Cumulative Pct.
General Motors	551	18.91%	18.91%
Other	516	17.71%	36.62%
Nissan Motors Corporation	467	16.03%	52.64%
Ford Motor Company	440	15.10%	67.74%
Chrysler LLC	439	15.07%	82.81%
Toyota Motor Sales	332	11.39%	94.20%
American Honda	169	5.80%	100.00%

(b) General Motors has the most complaints, followed by Other, Nissan Motors Corporation, Ford Motor Company, Chryler LLC, Toyota Motor Sales and American Honda.

(c) The percentage of complaints for each category:

Category	Frequency	Percentage	Cumulative Pct.
Powertrain	1148	42.82%	42.82%
Steering	397	14.81%	57.63%
Interior Electronics/Hardware	279	10.41%	68.03%
Fuel/Emission/Exhaust System	240	8.95%	76.99%
Airbags and Seatbelts	201	7.50%	84.48%
Body and Glass	182	6.79%	91.27%
Brakes	163	6.08%	97.35%
Tires and Wheels	71	2.65%	100.00%

(d) Powertrain has the most complaints, followed by steering, interior electronics/hardware, fuel/emission/exhaust system, airbags and seatbelts, body and glass, brakes, and, finally, tires and wheels.

2.5 (a) The percentage of values for each factor:

Most Important Factor	Frequency	Percentage	Cumulative Pct.
Product	464	35.80%	35.80%
Leadership	400	30.86%	66.67%
Marketing	346	26.70%	93.36%
Technology	86	6.64%	100.00%

(b) Product is the most influencing factor in successful start-ups, followed by Leardership, Marketing and Technology.

2.6 (a)

Region	Oil Production	Percentag
	(millions of barrels a day)	
Iran	2.69	3.27%
Saudi Arabia	9.58	11.66%
Other OPEC countries	17.93	21.82%
Non-OPEC countries	51.99	63.26%
Total	82.19	100.00%

(b) More than half the oil produced is from non-OPEC countries. About 22% is produced by OPEC countries other than Iran and Saudi Arabia.

2.7 (a) The percentage of values for each response need:

Needs	Frequency	Percentage	Cumulative Pct.
Easier-to-use analytic tools	127	30.98%	30.98%
Improved ability to present and interpret data	123	30.00%	60.98%
Improved ability to predict impacts of my actions	49	11.95%	72.93%
Faster access to data	41	10.00%	82.93%
Improved relationships to the business line organizations	37	9.02%	91.95%
Improved ability to plan actions	33	8.05%	100.00%

(b) "Easier-to-use analytic tools" is the most frequently mentioned need, followed by "Improved ability to present and interpret data", "Improved ability to predict impacts of my actions", "Faster access to data", "Improved relationships to the business line organizations" and "Improved ability to plan actions".

2.8 (a) Table of total percentages

ENJOY SHOPPING FOR CLOTHING FOR YOURSELF	GENDER		
	Male	Female	Total
Yes	22%	25%	47%
No	28%	25%	53%
Total	50%	50%	100%

Table of row percentages

ENJOY SHOPPING FOR CLOTHING FOR YOURSELF	GENDER		
	Male	Female	Total
Yes	46%	54%	100%
No	53%	47%	100%
Total	50%	50%	100%

Table of column percentages

ENJOY SHOPPING FOR CLOTHING FOR YOURSELF	GENDER		
	Male	Female	Total
Yes	44%	51%	47%
No	56%	49%	53%
Total	100%	100%	100%

(b) A higher percentage of females enjoy shopping for clothing for themselves.

2.9 (a)

Table of total percentages:

	Project Owner's Backing History		
Project Outcomes	Backing History	No Backing History	Total
Successful	26%	28%	54%
Not successful	16%	30%	46%
Total	42%	58%	100%

Table of row percentages:

	Project Owner's Backing History		
Project Outcomes	Backing History	No Backing History	Total
Successful	48%	52%	100%
Not successful	35%	65%	100%
Total	42%	58%	100%

Table of column percentages:

	Project Owner's Backing History		
Project Outcomes	Backing History	No Backing History	Total
Successful	62%	49%	54%
Not successful	38%	51%	46%
Total	100%	100%	100%

(b) The column percentages is most informative for these data as they show that among those owners with backing history, 62% are successful while only 49% are successful among those with no backing history.

(c) The percentage of success among owners with backing history is higher than those with no backing history.

2.10 Social recommendations had very little impact on correct recall. Those who arrived at the link from a recommendation had a correct recall of 73.07% as compared to those who arrived at the link from browsing who had a correct recall of 67.96%.

2.11 Ordered array: 63 64 68 71 75 88 94

2.12 Ordered array: 73 78 78 78 85 88 91

2.13 (a) $(17+7)/70 = 34.29\%$ of small businesses pay less than 21% of the employee monthly health-care premium.

(b) $(7+4)/70 = 15.71\%$ of small businesses pay between 21% and 75% of the employee monthly health-care premium.

(c) $(35)/70 = 50.00\%$ of small businesses pay more than 75% of the employee monthly health-care premium.

2.14 (a) 0 but less than 5 million, 5 million but less than 10 million, 10 million but less than 15 million, 15 million but less than 20 million, 20 million but less than 25 million, 25 million but less than 30 million.

(b) 5 million

(c) 2.5 million, 7.5 million, 12.5 million, 17.5 million, 22.5 million, and 27.5 million.

2.15 (a) Ordered array: Cost($) 203.06, 208.48, 212.16, 227.36, 240.04, 249.22, 262.40, 263.10, 266.40, 268.28, 271.74, 273.98, 280.98, 295.40, 308.18, 309.30, 319.10, 321.18, 321.63, 324.08, 336.05, 338.00, 344.92, 382.00, 395.20, 434.96, 456.60, 472.20, 542.00, 659.92,

(b) PHStat output:

Bin Cell	Frequency	Percentage	Cumulative Pctage.
200 but less than 270	10	0.3333333	33.33%
270 but less than 340	12	0.4	73.33%
340 but less than 410	3	0.1	83.33%
410 but less than 480	3	0.1	93.33%
480 but less than 550	1	0.0333333	96.67%
550 but less than 620	0	0	96.67%
620 but less than 690	1	0.0333333	100.00%

(c) The costs of attending a basketball game is concentrating around $305 for twelve of the teams have costs between $270 and $340.

2.16 (a)

Electricity Costs	Frequency	Percentage
$80 to $99	4	8%
$100 to $119	7	14
$120 to $139	9	18
$140 to $159	13	26
$160 to $179	9	18
$180 to $199	5	10
$200 to $219	3	6

(b)

Electricity Costs	*Frequency*	*Percentage*	*Cumulative %*
$99	4	8%	8%
$119	7	14%	22%
$139	9	18%	40%
$159	13	26%	66%
$179	9	18%	84%
$199	5	10%	94%
$219	3	6%	100%

(c) The majority of utility charges are clustered between $120 and $180.

2.17 (a), (b) Annual Time Sitting in Traffic (hours)

Bin Cell	Frequency	Percentage	Cumulative Pctage.
15 but less than 20	1	3.23%	3.23%
20 but less than 25	4	12.90%	16.13%
25 but less than 30	4	12.90%	29.03%
30 but less than 35	2	6.45%	35.48%
35 but less than 40	7	22.58%	58.06%
40 but less than 45	3	9.68%	67.74%
45 but less than 50	4	12.90%	80.65%
50 but less than 55	2	6.45%	87.10%
55 but less than 60	1	3.23%	90.32%
60 but less than 65	1	3.23%	93.55%
65 but less than 70	0	0.00%	93.55%
70 but less than 75	2	6.45%	100.00%

Cost of Sitting in Traffic($)

Bin Cell	Frequency	Percentage	Cumulative Pctage.
300 but less than 450	4	12.90%	12.90%
450 but less than 600	6	19.35%	32.26%
600 but less than 750	6	19.35%	51.61%
750 but less than 900	5	16.13%	67.74%
900 but less than 1050	6	19.35%	87.10%
1050 but less than 1200	2	6.45%	93.55%
1200 but less than 1350	1	3.23%	96.77%
1350 but less than 1550	0	0.00%	96.77%
1550 but less than 1650	1	3.23%	100.00%

(c) The annual time sitting in traffic is concentrated around 37.5 hours with a few spending as much as around 72.5 hours.

(d) The cost of sitting in traffic per year is concentrated around $675 with one costing as much as $1,575.

2.18 (a), (b)

Bin Cell	Frequency	Percentage	Cumulative Pctage.
695 but less than 705	3	2.10%	2.10%
705 but less than 715	12	8.39%	10.49%
715 but less than 725	12	8.39%	18.88%
715 but less than 735	19	13.29%	32.17%
735 but less than 745	18	12.59%	44.76%
745 but less than 755	24	16.78%	61.54%
755 but less than 765	22	15.38%	76.92%
765 but less than 775	20	13.99%	90.91%
775 but less than 785	10	6.99%	97.90%
795 but less than 795	3	2.10%	100.00%

(c) The average credit scores are concentrated around 750.

2.19 (a), (b)

Bin	*Frequency*	*Percentage*	*Cumulative %*
-0.00350 but less than -0.00201	13	13.00%	13.00%
-0.00200 but less than -0.00051	26	26.00%	39.00%
-0.00050 but less than 0.00099	32	32.00%	71.00%
0.00100 but less than 0.00249	20	20.00%	91.00%
0.00250 but less than 0.00399	8	8.00%	99.00%
0.004 but less than 0.00549	1	1.00%	100.00%

(c) Yes, the steel mill is doing a good job at meeting the requirement as there is only one steel part out of a sample of 100 that is as much as 0.005 inches longer than the specified requirement.

2.20 (a), (b)

Bin	*Frequency*	*Percentage*	*Cumulative %*
8.310 -- 8.329	3	6.12%	6.12%
8.330 -- 8.349	2	4.08%	10.20%
8.350 -- 8.369	1	2.04%	12.24%
8.370 -- 8.389	4	8.16%	20.41%
8.390 -- 8.409	4	8.16%	28.57%
8.410 -- 8.429	15	30.61%	59.18%
8.430 -- 8.449	7	14.29%	73.47%
8.450 -- 8.469	5	10.20%	83.67%
8.470 -- 8.489	5	10.20%	93.88%
8.490 -- 8.509	3	6.12%	100.00%

(c) All the troughs will meet the company's requirements of between 8.31 and 8.61 inches wide.

2.21 (a),(b)

Strength	*Frequency*	*Percentage*	*Cumulative Percentage*
1500 -- 1549	1	3.33%	3.33%
1550 -- 1599	2	6.67%	10.00%
1600 -- 1649	2	6.67%	16.67%
1650 -- 1699	7	23.33%	40.00%
1700 -- 1749	5	16.67%	56.67%
1750 -- 1799	7	23.33%	80.00%
1800 -- 1849	3	10.00%	90.00%
1850 -- 1899	3	10.00%	100.00%

(c) The strength of all the insulators meets the company's requirement of at least 1500 lbs.

2.22 (a), (b) Manufacturer A:

Bin Cell	Frequency	Percentage	Cumulative Pctage.
6,500 but less than 7,500	3	7.50%	7.50%
7,500 but less than 8,500	5	12.50%	20.00%
8,500 but less than 9,500	20	50.00%	70.00%
9,500 but less than 10,500	9	22.50%	92.50%
10,500 but less than 11,500	3	7.50%	100.00%

Manufacturer B:

Bin Cell	Frequency	Percentage	Cumulative Pctage.
7,500 but less than 8,500	2	5.00%	5.00%
9,500 but less than 9,500	8	20.00%	25.00%
9,500 but less than 10,500	16	40.00%	65.00%
10,500 but less than 11,500	9	22.50%	87.50%
11,500 but less than 12,500	5	12.50%	100.00%

(c) Manufacturer B produces bulbs with longer lives than Manufacturer A. The cumulative percentage for Manufacturer B shows 65% of its bulbs lasted less than 10,500 hours, contrasted with 70% of Manufacturer A's bulbs, which lasted less than 9,500 hours. None of Manufacturer A's bulbs lasted more than 11,499 hours, but 12.5% of Manufacturer B's bulbs lasted between 11,500 and 12,499 hours. At the same time, 7.5% of Manufacturer A's bulbs lasted less than 7,500 hours, whereas all of Manufacturer B's bulbs lasted at least 7,500 hours

2.23 (a)

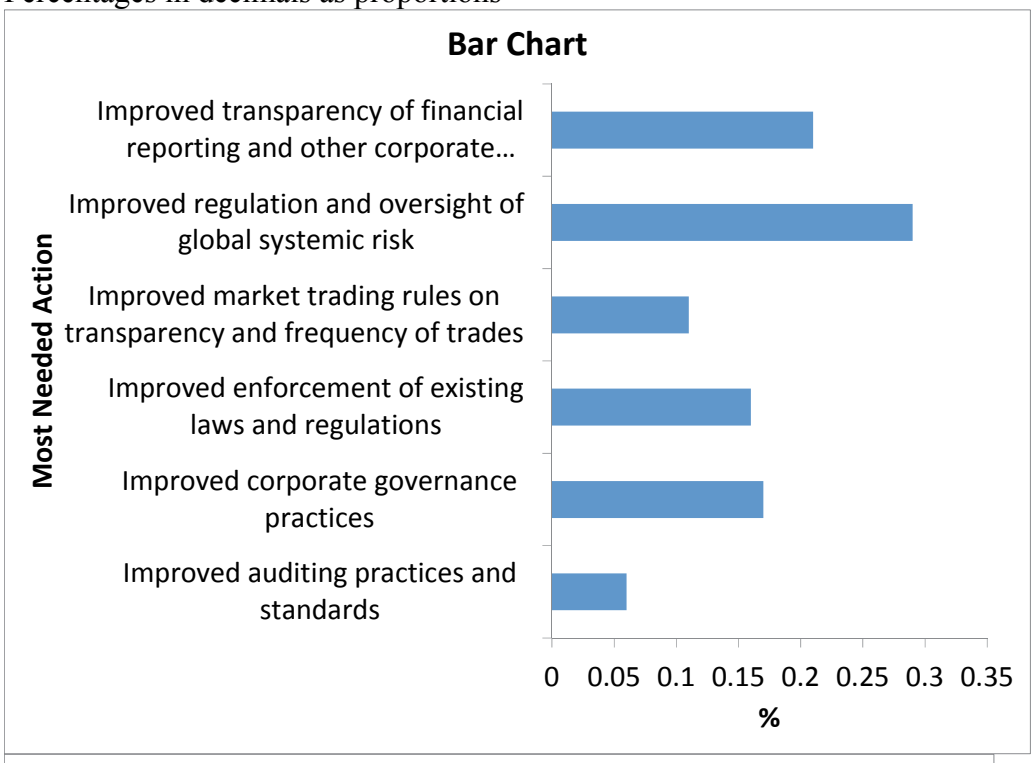
Amount of Soft Drink	Frequency	Percentage
1.850 – 1.899	1	2%
1.900 – 1.949	5	10
1.950 – 1.999	18	36
2.000 – 2.049	19	38
2.050 – 2.099	6	12
2.100 – 2.149	1	2

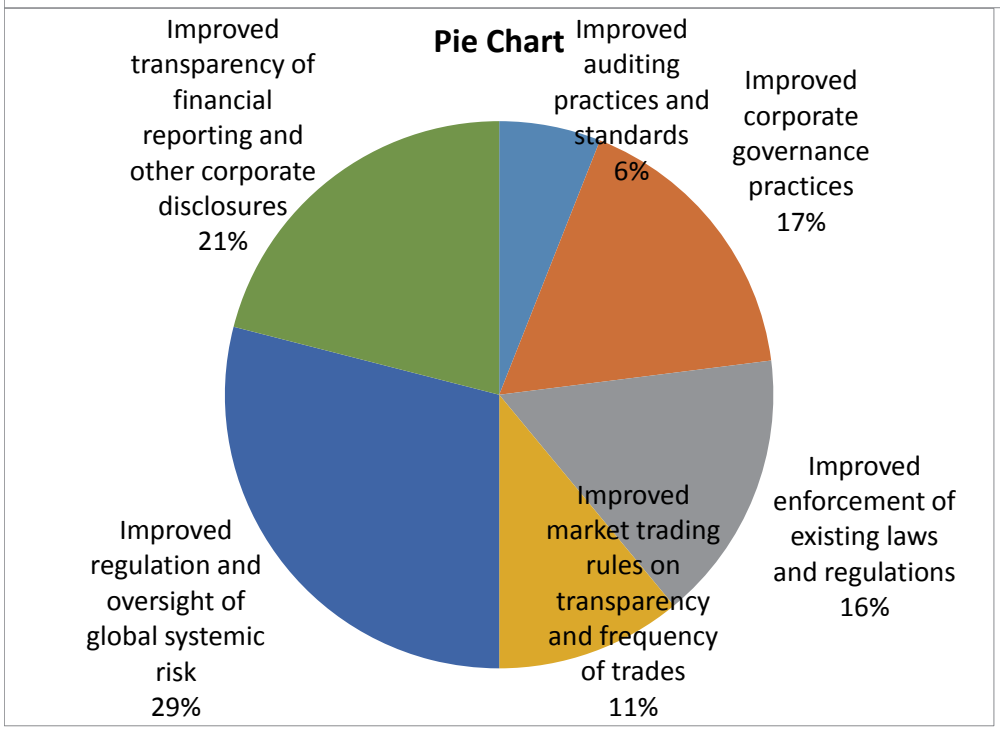
Amount of Soft Drink	Frequency Less Than	Percentage Less Than
1.899	1	2%
1.949	6	12
1.999	24	48
2.049	43	86
2.099	49	98
2.149	50	100

(b) The amount of soft drink filled in the two liter bottles is most concentrated in two intervals on either side of the two-liter mark, from 1.950 to 1.999 and from 2.000 to 2.049 liters. Almost three-fourths of the 50 bottles sampled contained between 1.950 liters and 2.049 liters.

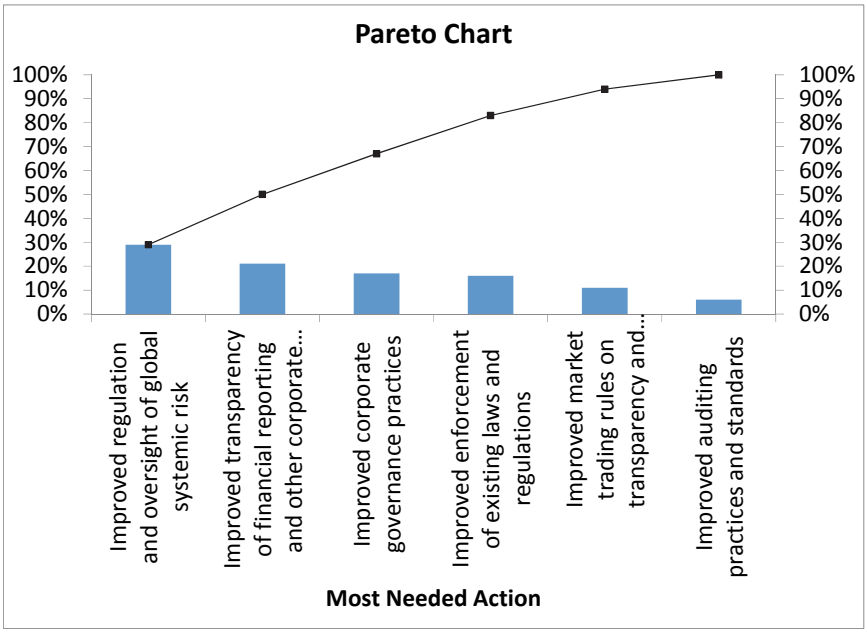
2.24 (a)

Percentages in decimals as proportions





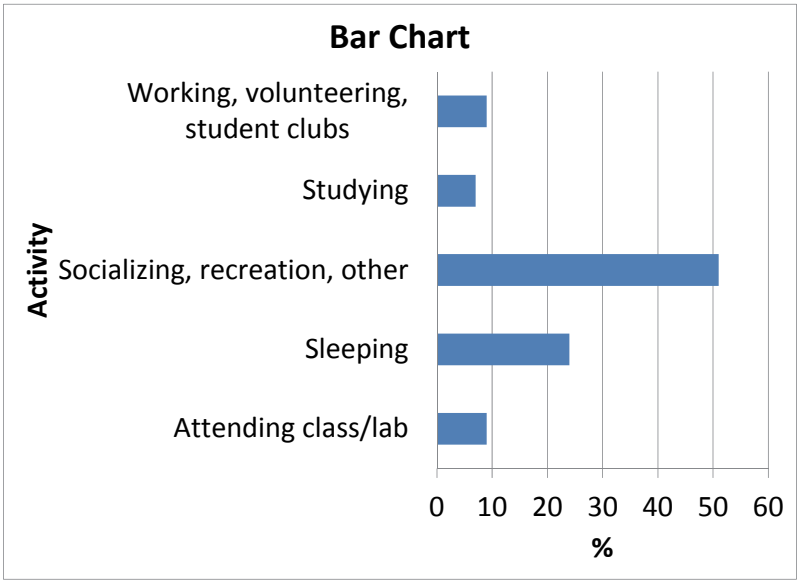
2.24 (a) cont.



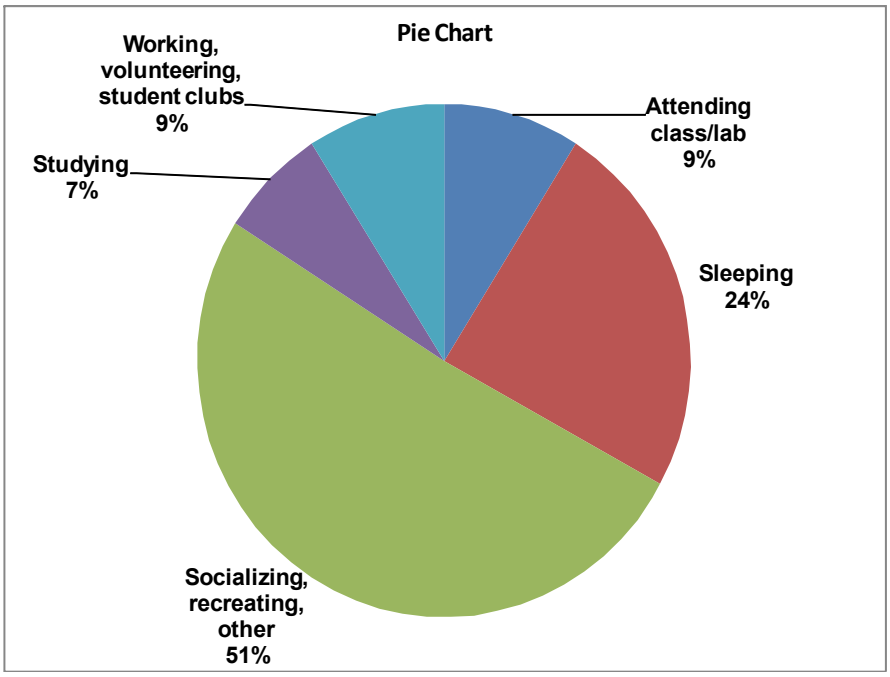
(b) The Pareto diagram is better than the pie chart to portray these data because it not only sorts the frequencies in descending order, it also provides the cumulative polygon on the same scale.

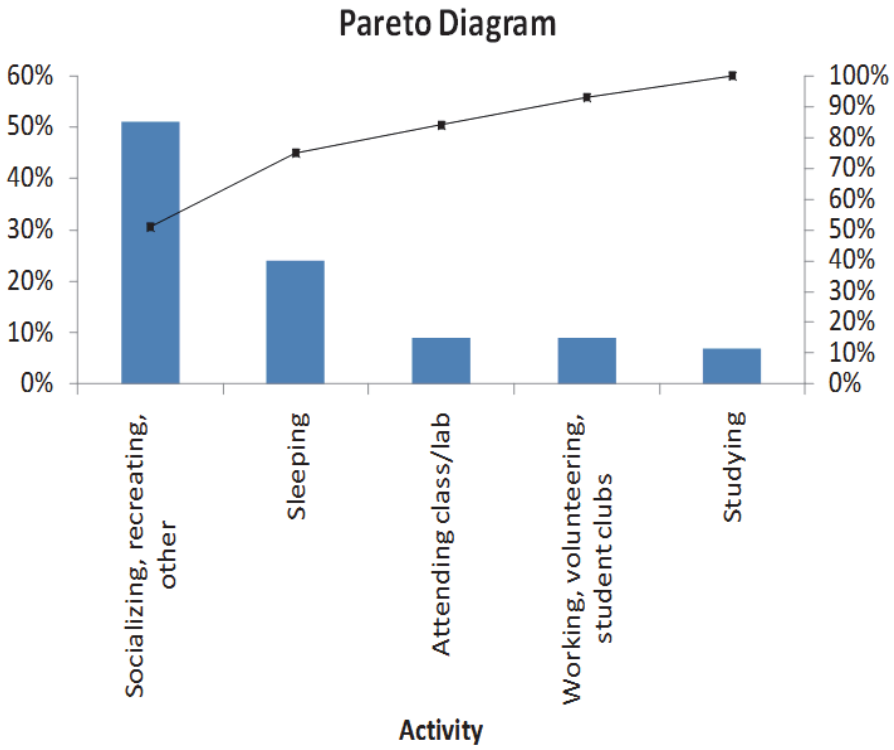
(c) You can conclude that "improved regulation and oversight of global systemic risk" accounts for the largest percentage (29%) of the most needed action to improve investor trust and market integrity.

2.25 (a)



2.25 cont. (a)

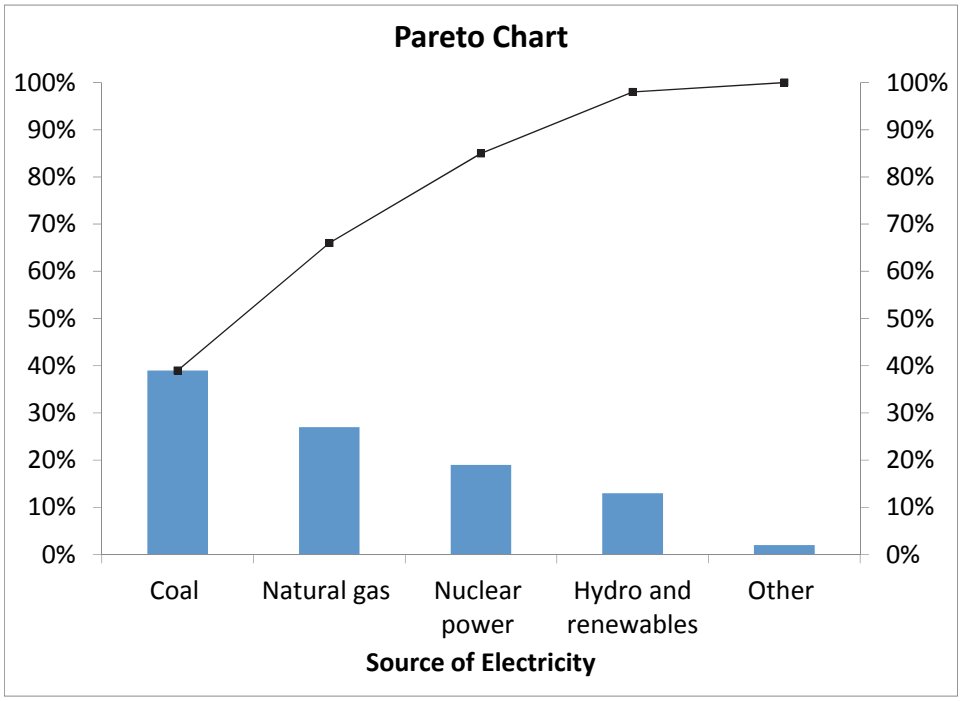




(b) The Pareto diagram is better than the pie chart or the bar chart because it not only sorts the frequencies in descending order, it also provides the cumulative polygon on the same scale.

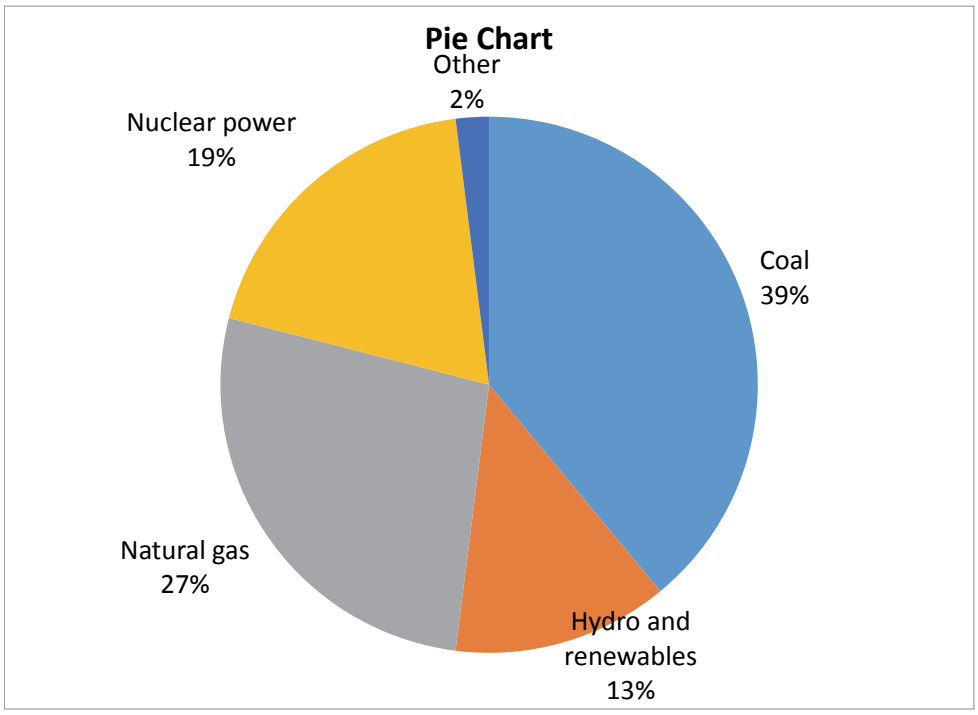
(c) From the Pareto diagram, it is obvious that slightly more than 50% of them were socializing, recreating or performing other activities.

2.26 (a)



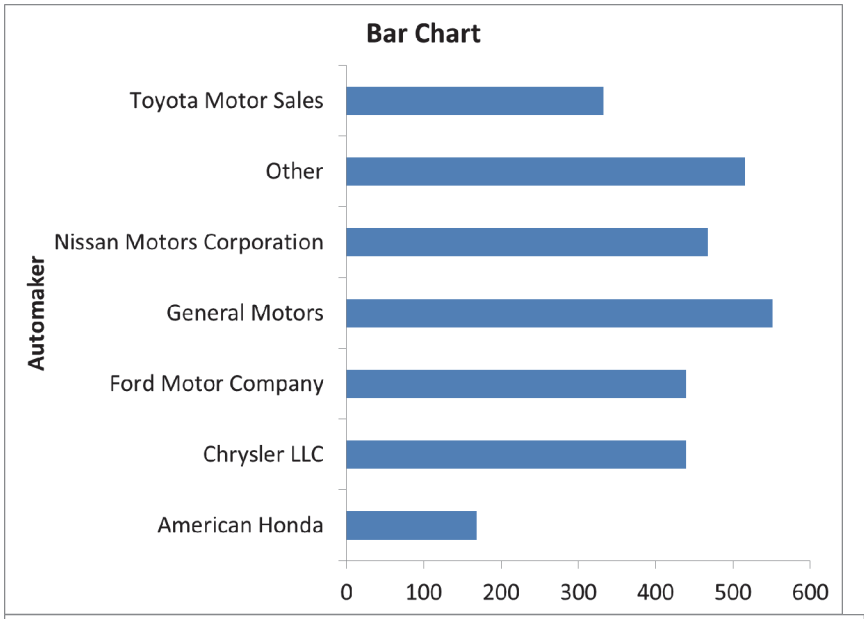
(b) Eighty-five percent of power is derived from coal, natural gas, or nuclear power.

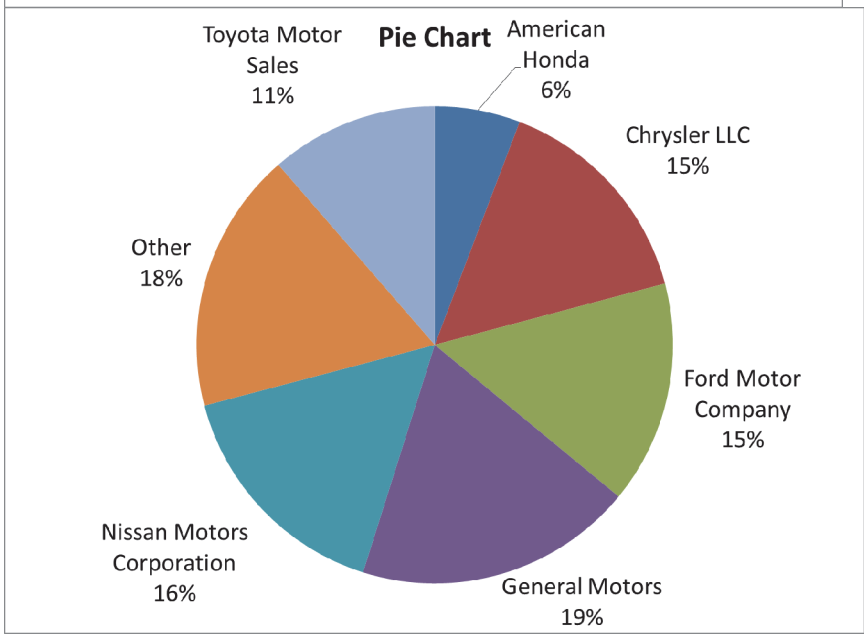
(c)



(d) The Pareto diagram is better than the pie chart because it not only sorts the frequencies in descending order, it also provides the cumulative polygon on the same scale.

2.27 (a)





(b) The bar chart is more suitable if the purpose is to compare the categories. The pie chart is more suitable if the main objective is to investigate the portion of the whole that is in a particular category. *

* Note: This is one of the many possible solutions for the question.