1 Review Questions to Sections 1.1 - 1.3

Exercise 1
In each of these statements, tell whether descriptive or inferential statistics have been used.
(a) Expenditures for the cable industry were $5.66 billion in 1996.
(b) The median household income for people aged 25-34 is $35,888.
(c) Drinking decaffeinated coffee can raise cholesterol levels by 7%.
(d) The national average annual medicine expenditure per person in USA is $1,052.
(f) Experts say that mortgage rates may soon hit bottom.
Ans: (a) descriptive (b) descriptive (c) inferential (d) descriptive (f) inferential.

Exercise 2
Each of the following processes involves sampling from a population. Define the population, and state whether it is tangible or conceptual.
(a) 8 welds are made with the same process, and the strength of each is measured.
(b) A quality engineer needs to estimate the percentage of parts manufactured on a certain day that are defective. At 2.30 in the afternoon he samples the last 100 parts to be manufactured.
Ans: (a) conceptual: all welds that could be made by that process (b) tangible: all parts manufactured that day

Exercise 3
For each following questions, identify the type of sampling obtained (simple random, stratified, or cluster).
(a) Suppose you wanted to study dance club and bar employees in NYC with a sample of n = 600. Yet there is no list of these employees from which to draw a simple random sample. Suppose you obtained a list of all bars/clubs in NYC. One way to get this would be to randomly sample 300 bars and then randomly sample 2 employees within each bars/club.
(b) Suppose a farmer wishes to work out the average milk yield of each cow type in his herd which consists of Ayrshire, Friesian, Galloway and Jersey cows. He could divide up his herd into the four sub-groups and take samples from these.
(c) To select a sample of 25 people who live in your college dorm, make a list of all the 250 people who live in the dorm. Assign each person a unique
number, between 1 and 250. Then refer to a table of random numbers.
Ans: (a) cluster (b) stratified (c)

Exercise 4
Politicians want to determine the attitudes of those they represent when they report on the number of letters voluntarily sent to them by their constituents. Would this be considered a sample of convenience?
Ans: Yes.

Exercise 5
A certain process for manufacturing integrated circuits has been in use for a period of time, and it is known that 12% of the circuits it produces are defective. A new process that is supposed to reduce the proportion of defectives is being tested. In a simple random sample of 100 circuits produced by the new process, 12 were defective.
(a) One of the engineers suggests that the test proves that the new process is no better than the old process, since the proportion of defectives in the sample is the same. Is this conclusion justified? Explain.
(b) Assume that there had been only 11 defective circuits in the sample of 100. Would this have proven that the new process is better? Explain.
(c) Which outcome represents stronger evidence that the new process is better: finding 11 defective circuits in the sample, or finding 2 defective circuits in the sample?
Ans: (a) No. What is important is the population proportion of defectives; the sample proportion is only an approximation. The population proportion for the new process may in fact be greater or less than that of the old process.
(b) No. The population proportion for the new process may be 0.12 or more, even though the sample proportion was only 0.11.
(c) Finding 2 defective circuits in the sample.

Exercise 6
An urn contains five balls numbered 1 through 5. I pick two balls and write down their numbers. Then I pick another two balls of the remaining three and write down their numbers. Are the two samples dependent or independent?
Ans: Dependent (sampling without replacement)

Exercise 7
Identify each study as being either observational or experimental.
(a) Subjects are randomly assigned to four groups. Each group is placed on one of four special diets: low-fat diet, a high-fish diet, a combination of low-fat diet and high-fish diet, and a regular diet. After 6 months, the blood pressures of the groups are compared to see if diet has any effect on blood pressure.

(b) A researcher finds that people who are more hostile have higher total cholesterol levels than those who are less hostile.

Ans: (a) Experimental (b) Observational

Exercise 8
Classify each data as qualitative or quantitative.
(a) Classification of children in a day care center (infant, toddler, preschool).
(b) Weights of fish caught in Lake Dardanelle.
(c) Number of pages in statistics textbooks.

Ans: (a) Qualitative (b) Quantitative (c) Quantitative.

Exercise 9
A statistics class of 40 students took a quiz. The highest possible score was 4 points. 10 students scored 4 points, 12 students scored 3 points, 8 students scored 2 point, 6 students scored 1 point, and 4 students scored 0. Compute the mean, median, and standard deviation of the population scores.

Ans: Mean = 2.5, median = 3, standard deviation = 1.3.

Exercise 10
As part of a class project, Dilan examined a questionnaire sent to 50 universities’ head basketball coaches. 25 responded. The following data shows the number of years each coach had been coaching at their university.

7, 6, 8, 5, 9, 4, 7, 7, 6, 6, 13, 2, 1, 20, 1, 1, 7, 6, 7, 5, 8, 4, 8, 3, 6.

(a) Calculate the sample mean years of coaching.
(b) What is the range value?
(c) Calculate the sample variance and standard deviation.
(d) Calculate the median and mode.
(e) What are the values of $Q_1$ and $Q_3$? (f) Find the outliers.

Ans: (a) 6.3 years (b) 19 years (c) $s^2 = 15.8$, $s = 4$ years (d) median = 6 years and mode = 6, 7 years (e) $Q_1 = 4$ years and $Q_3 = 7$ years (f) outliers = 13, 20.
Exercise 11
(a) Find 10% trimmed mean of 2, 4, 6, 7, 11, 21, 81, 90, 105, 121.
(b) Find the 15% trimmed mean.
Ans: (a) 40.625 (b) 36.2

Exercise 12
Determine whether the given value is statistic or parameter.
(a) In a study of all 2223 passengers aboard the Titanic, it is found that 706 survived when it sank.
(b) A sample of viewers is selected and the average amount of time watching television is 5.6 hours a day.
Ans: (a) Parameter (b) Statistic.

Exercise 13
The average temperatures for 10 randomly selected cities in the eastern United States on June 17 are listed below:

<table>
<thead>
<tr>
<th>City</th>
<th>Temperature</th>
<th>City</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlanta</td>
<td>89</td>
<td>Hartford</td>
<td>84</td>
</tr>
<tr>
<td>Baltimore</td>
<td>82</td>
<td>Miami Beach</td>
<td>88</td>
</tr>
<tr>
<td>Boston</td>
<td>76</td>
<td>New York</td>
<td>80</td>
</tr>
<tr>
<td>Burlington, Vt.</td>
<td>81</td>
<td>Orlando</td>
<td>91</td>
</tr>
<tr>
<td>Columbia, S.C.</td>
<td>87</td>
<td>Philadelphia</td>
<td>82</td>
</tr>
</tbody>
</table>

(a) What percentile is a value of 78 degrees?
(b) What is the value of the 20th percentile?
(c) What is the value of the 90th percentile?
(d) What is the median?
Ans: (a) 15 (b) 80 (c) 89 (d) 83

Exercise 14
In a certain company, every worker received a $50-per-week raise. How does this affect the mean salary? The standard deviation of the salaries?
Ans: The mean increases by $50; the standard deviation is unchanged.

Exercise 15
The tables give the number of days on which rain fail in 36 consecutive intervals of 30 days. Draw stem and leaf plot.
Exercise 16

The students in one social studies class were asked how many brothers and sisters (siblings) they each have. The dot plot here shows the results.

(a) How many of the students have six siblings?
(b) How many of the students have no siblings?
(c) How many of the students have three or more siblings?

Ans: (a) 1 (b) 0 (c) 11.

Exercise 17

The weather in Los Angeles is dry most of the time, but it can be quite rainy in the winter. The rainiest month of the year is February. The following table presents the annual rainfall in Los Angeles, in inches, for each February from 1965 to 2006.
Construct a relative frequency histogram with class width 2.

Ans:

Exercise 18
Determine which histogram is symmetric, positively skewed or negatively skewed.
Ans: (a) Positively skewed (b) Symmetric (c) Negatively skewed.

Exercise 19
Construct a box plot of the data in Exercise 17.
Ans:

The boxplot shows one outlier.

Exercise 20
Forty-five specimens of a certain type of powder were analyzed for sulfur trioxide content. Following are the results, in percent. The list has been sorted into numerical order.

14.1 14.4 14.7 14.8 15.3 15.6 16.1 16.6 17.3
14.2 14.4 14.7 14.9 15.3 15.7 16.2 17.2 17.3
14.3 14.4 14.8 15.0 15.4 15.7 16.4 17.2 17.8
14.3 14.4 14.8 15.0 15.4 15.9 16.4 17.2 21.9
14.3 14.6 14.8 15.2 15.5 15.9 16.5 17.2 22.4

(a) Construct a stem-and-leaf plot for these data.
(b) Construct a histogram for these data.
(c) Construct a dotplot for these data.
(d) Construct a boxplot for these data. Does the boxplot show any outliers?

Ans:

(a)  

<table>
<thead>
<tr>
<th>Stem</th>
<th>Leaf</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>1233344446778889</td>
</tr>
<tr>
<td>15</td>
<td>0023344567799</td>
</tr>
<tr>
<td>16</td>
<td>124456</td>
</tr>
<tr>
<td>17</td>
<td>2222338</td>
</tr>
<tr>
<td>18</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>9</td>
</tr>
<tr>
<td>22</td>
<td>4</td>
</tr>
</tbody>
</table>

(b)  

Sulfur Dioxide (percent)

(c)  

(d)  

The boxplot shows 2 outliers.