Chapter 2 Graphical Methods for Describing Data Distributions

Section 2.1 Exercise Set 1

2.1: (a) numerical, discrete (b) categorical (c) numerical, continuous (d) numerical, continuous (e) categorical

2.2: (a) discrete (b) continuous (c) discrete (d) discrete

2.3:

Data Set 1:

Question 1: There is one variable in the data set.

Question 2: The data set is categorical.

Question 3: The purpose of the graphical display is to summarize the data distribution.

Appropriate Graphical Display: Bar chart

Data Set 2:

Question 1: There is one variable in the data set.

Question 2: The data set is numerical.

Question 3: The purpose of the graphical display is to compare groups (full-time students or part-time students).

Appropriate Graphical Display: Comparative dotplot, comparative stem-and-leaf displays, and comparative histograms are all appropriate.

Data Set 3:

Question 1: There are two variables in the data set.

Question 2: The data set is numerical.

Question 3: The purpose of the graphical display is to investigate the relationship between two numerical variables.

Appropriate Graphical Display: Scatterplot

Data Set 4:

Question 1: There is one variable in the data set.

Question 2: The data set is categorical.

Question 3: The purpose of a graphical display is to compare groups (faculty, students)

Appropriate Graphical Display: Comparative bar chart

Data Set 5:

Question 1: There is one variable in the data set.

Question 2: The data set is numerical.

Question 3: The purpose of a graphical display is to summarize a data distribution.

Appropriate Graphical Display: Dotplot, stem-and-leaf, and histogram are all appropriate graphical displays.

Section 2.1 Exercise Set 2

2.4: (a) categorical

(b) categorical

(c) numerical - discrete

(d) numerical - continuous

(e) categorical

(f) numerical - continuous

2.5: (a) continuous (b) continuous (c) continuous (d) discrete

2.6: Data Set 1:

Question 1: There is one variable in the data set.

Question 2: The data set is numerical.

Question 3: The purpose of a graphical display is to summarize a data distribution.

Appropriate Graphical Display: Dotplot, stem-and-leaf, and histogram are all appropriate graphical displays.

Data Set 2:

Question 1: There is one variable in the data set.

Question 2: The data set is categorical.

Question 3: The purpose of a graphical display is to compare two groups (male and female color choice).

Appropriate Graphical Display: Comparative bar chart

Data Set 3:

Question 1: There are two variables in the data set.

Question 2: The data set is numerical.

Question 3: The purpose of a graphical display is to investigate the relationship between two numerical variables.

Appropriate Graphical Display: Scatterplot

Data Set 4:

Question 1: There is one variable in the data set.

Question 2: The data set is numerical.

Question 3: The purpose of a graphical display is to compare two groups.

Appropriate Graphical Display: Comparative dotplot, comparative stem-and-leaf displays, and comparative histograms are all appropriate.

Data Set 5:

Question 1: There is one variable in the data set.

Question 2: The data set is categorical.

Question 3: The purpose of a graphical display is to summarize a data distribution.

Appropriate Graphical Display: Bar chart.

Additional Exercises for Section 2.1

2.7: (a) numerical (b) numerical (c) categorical (d) numerical (e) categorical

2.8: Discrete: (b); Continuous: (a) and (d)

2.9: (a) categorical (b) numerical (c) numerical (d) categorical

2.10: Discrete: (b); Continuous: (c)

2.11: (a) numerical (b) numerical (c) numerical (d) categorical (e) categorical (f) numerical (g) categorical

2.12:

Question 1: There is one variable in the data set.

Question 2: The data set is categorical.

Question 3: The purpose is to summarize the data distribution.

Appropriate graphical display: Bar chart

2.13:

Question 1: There is one variable in the data set.

Question 2: The data set is numerical.

Question 3: The purpose is to compare groups (male students, female students).

Appropriate graphical display: Comparative dotplot, comparative stem-and-leaf, or histograms are all appropriate.

2.14:

Question 1: There are two variables in the data set.

Question 2: The data set is numerical.

Question 3: The purpose is to investigate the relationship between two numerical variables.

Appropriate graphical display: Scatterplot

2.15:

Question 1: There is one variable in the data set.

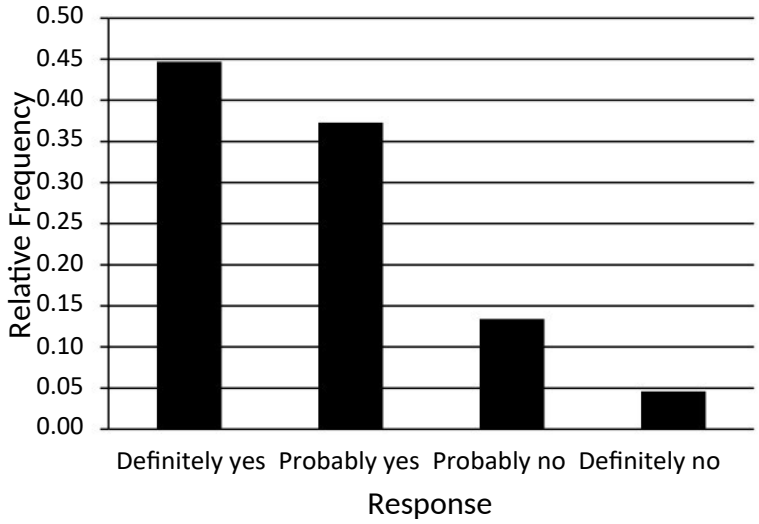
Question 2: The data set is numerical.

Question 3: The purpose of a graphical display is to summarize the data distribution.

Appropriate graphical display: Dotplots, stem-and-leaf plots, and histograms are all appropriate graphical displays.

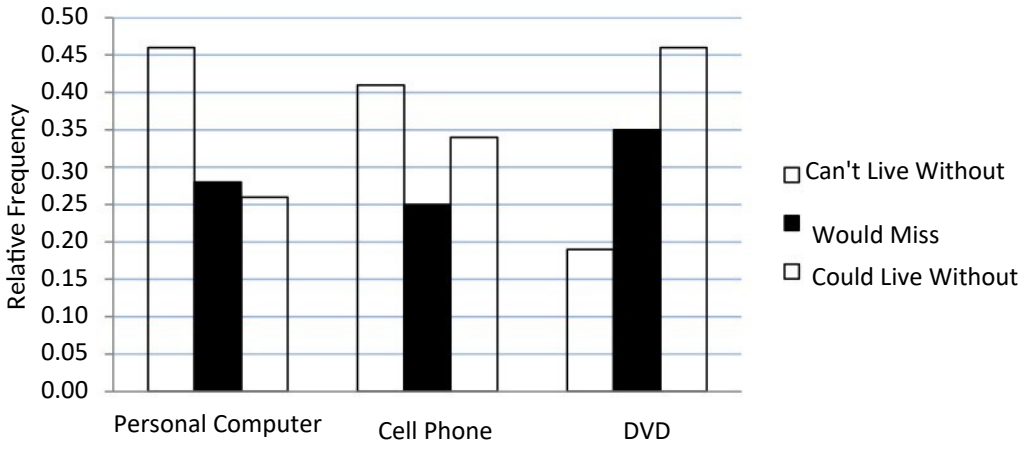
Section 2.2 Exercise Set 1

2.16: (a)



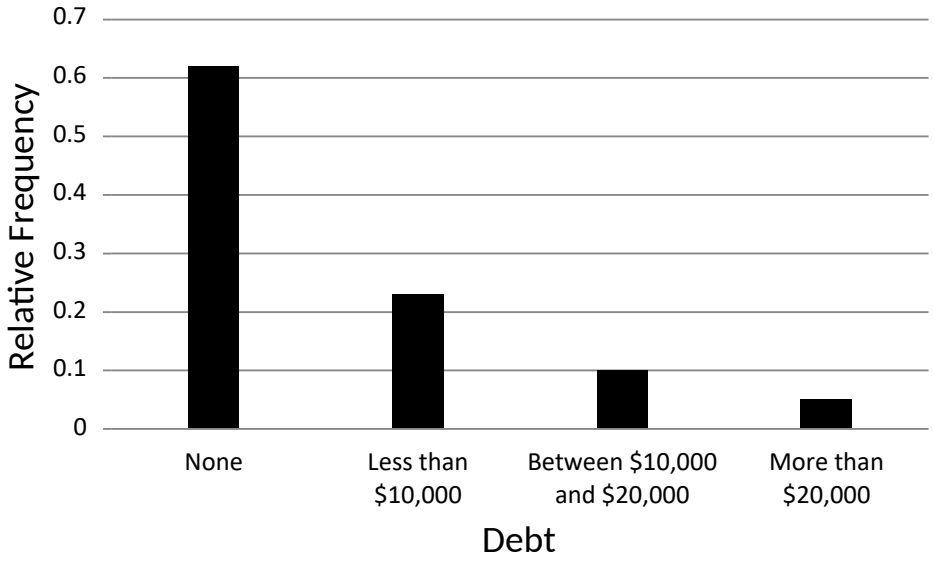
(b) "Senior Satisfaction! Over 80% say they would enroll again."

2.17:



Section 2.2 Exercise Set 2

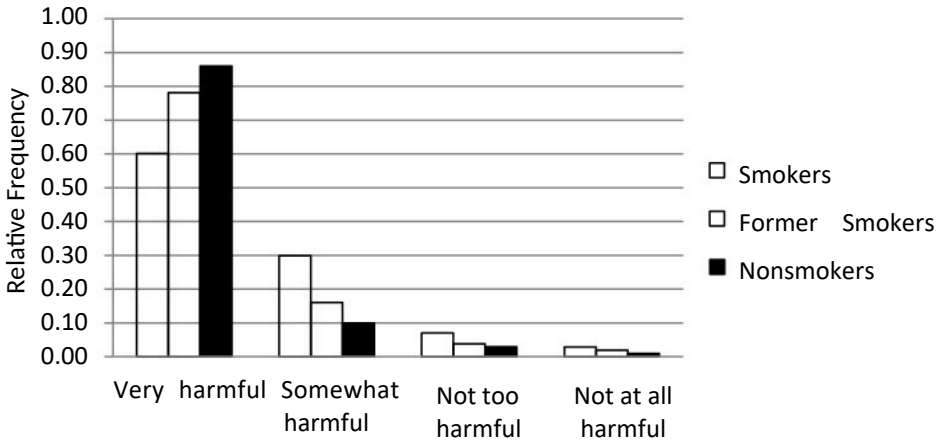
2.18: (a)



(b) Over 60% of students graduating with an AA degree from a public community college in 2008 graduate with no debt. As the amount of debt increases, fewer students reach that debt level. Twenty-three percent of students have less than $10,000 in debt, 10% have between $10,000 and $20,000 in debt, and only 5% have over $20,000 in debt.

2.19: The relative frequencies needed are shown in the table below, followed by the comparative bar chart.

Perceived Risk of Smoking	Smokers	Former Smokers	Nonsmokers
Very harmful	0.60	0.78	0.86
Somewhat harmful	0.30	0.16	0.10
Not too harmful	0.07	0.04	0.03
Not at all harmful	0.03	0.02	0.01

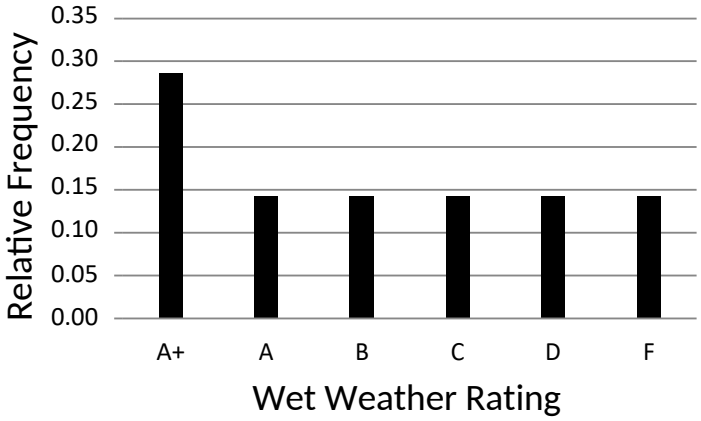


Within each category (smokers, former smokers, and nonsmokers), the ordering based on relative frequency of the perceived risks (very harmful, somewhat harmful, not too harmful, and not at all harmful) is the same. In the somewhat harmful, not too harmful, and not at all harmful categories, smokers had the highest relative frequencies, followed by former smokers, and then nonsmokers. The very harmful perceived risk category is different, in that the nonsmokers had the highest relative frequency, followed by the former smokers, and then the smokers.

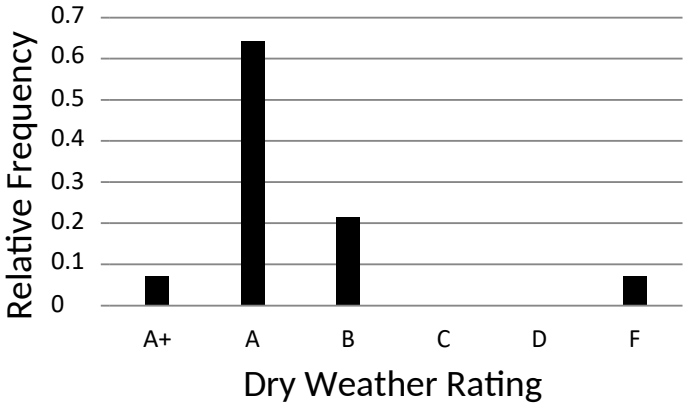
Additional Exercises for Section 2.2

2.20: (a)

Rating	Relative Frequency
A+	4/14 = 0.286
A	2/14 = 0.143
B	2/14 = 0.143
C	2/14 = 0.143
D	2/14 = 0.143
F	2/14 = 0.143

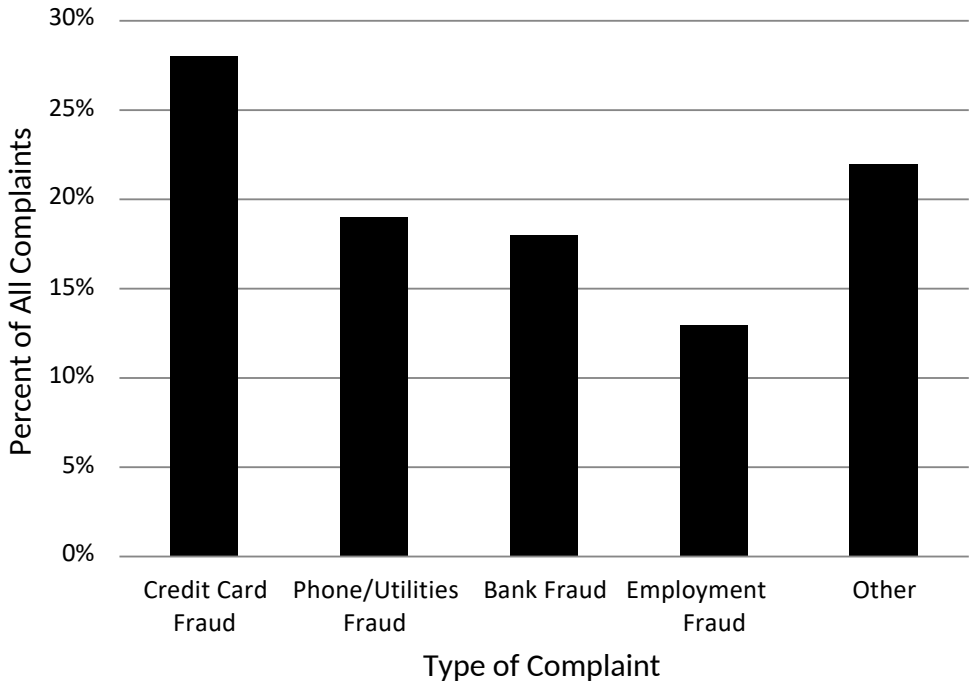


(b)



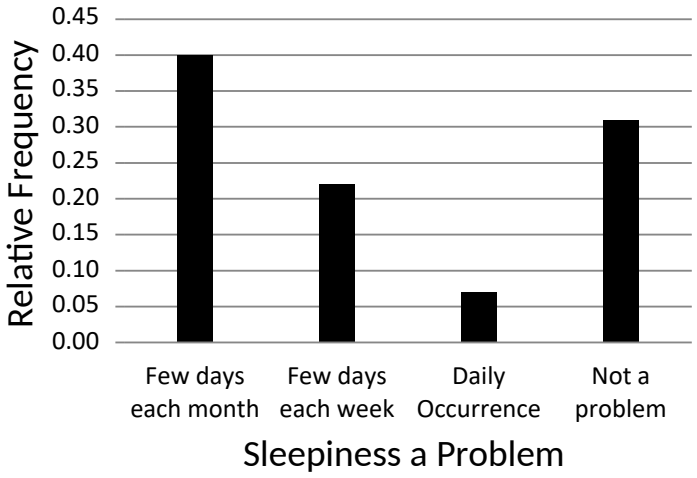
(c) Yes, the bar charts from (a) and (b) support the statement that beach water quality tends to be better in dry weather conditions. The bar charts show that in dry weather condtions, approximately 93% of the beaches have a rating of "B" or higher, whereas in wet weather, only approximately 57% of the beaches have a rating of "B" or higher.

2.21:



Credit card fraud is the most commonly occurring identity theft type. Although phone/utility, bank, and employment fraud each constitute a relatively large portion of overall type of identity theft, the collective "other fraud" category is greater than any one of these other three.

2.22: To construct a relative frequency bar chart, the sum of the relative frequencies for all the categories must add to 1 (or 100%). As such, the category "sleepiness on the job was not a problem" must account for 100 - (40 + 22 + 7) = 31% of those surveyed.



2.23: The relative frequency distribution is:

Type of Household	Relative Frequency
Nonfamilies	0.29
Married with Children	0.27
Married without Children	0.29
Single Parent	0.15