## Example 3.1.5

**Creating a Cumulative Relative Frequency Distribution of Survey Results**

Determine the cumulative relative frequency for the July customer satisfaction survey data from Example 3.1.1.

| July Customer Satisfaction Survey | | | |
|---|---|---|---|
| **Response** | **Frequency** | **Cumulative Frequency** | **Cumulative Relative Frequency** |
| **1–Very Dissatisfied** | 0 | 0 | $\frac{0}{20}=0$ |
| **2–Somewhat Dissatisfied** | 5 | 5 | $\frac{5}{20}=0.25=25\%$ |
| **3–Neutral** | 7 | 12 | $\frac{12}{20}=0.6=60\%$ |
| **4–Somewhat Satisfied** | 6 | 18 | $\frac{18}{20}=0.9=90\%$ |
| **5–Very Satisfied** | 2 | 20 | $\frac{20}{20}=1=100\%$ |

From the cumulative relative frequency distribution it is apparent that 60% of customers from the July survey do not have a favorable opinion of the product.

## Example 3.1.6

**Creating a Cumulative Relative Frequency Distribution for the Heart Rate Data**

Determine the cumulative relative frequency distribution for the heart rate data.

### Solution

| Heart Rate Cumulative Relative Frequency | | |
|---|---|---|
| **Heart Rate** | **Relative Frequency** | **Cumulative Relative Frequency** |
| **57–66** | 0.04 | 0.04 |
| **67–76** | 0.20 | 0.24 |
| **77–86** | 0.64 | 0.88 |
| **87–96** | 0.10 | 0.98 |
| **97–106** | 0.02 | 1.00 |

From the cumulative relative frequency distribution it is easy to see that 88% of the students had heart rates less than or equal to 86 beats per minute.

# 3.1 Exercises

## Basic Concepts

1. What does it mean for a set of data to have a distribution?
2. Describe the purpose of a frequency distribution.

3. What are the basic questions to ask when examining the structure of a data set?
4. What are the three steps to constructing a frequency distribution?
5. In the construction of a frequency distribution, what are the two requirements that the classification categories must meet?
6. What are the fundamental decisions in constructing frequency distributions for quantitative data?
7. Describe the general guidelines for selecting the number of classes for a quantitative frequency distribution.
8. What is a good starting point for determining the class width?
9. What is a relative frequency distribution? How do you calculate relative frequencies from raw frequencies?
10. What is a cumulative frequency distribution? How do you calculate a cumulative frequency distribution?
11. What is a cumulative relative frequency distribution? What are two ways to determine a cumulative relative frequency distribution?

## Exercises

12. Parkinsonism is an affliction of the aged and is frequently caused by Parkinson's disease, Alzheimer's disease, or other illnesses. The results from a recent study on Parkinsonism were reported in "Prevalence of Parkinsonian Signs and Associated Mortality in a Community Population of Older People," *New England Journal of Medicine*.[2] A sample of 467 people, all 65 years of age or older, was selected from East Boston, Massachusetts. Each person was clinically evaluated and various signs of Parkinsonism, if any, were noted. The following table is a frequency distribution for some of the signs of Parkinsonism.

| Signs of Parkinsonism | |
|---|---|
| **Sign** | **Frequency** |
| Reduced arm swing | 210 |
| Prolonged turning | 153 |
| Right leg rigidity | 141 |
| Left leg rigidity | 154 |
| Slow finger taps | 197 |
| Shuffling gait | 83 |

**a.** What level of measurement does the data possess?

**b.** What percent of the sample suffered from left leg rigidity? Round your answer to two decimal places.

**c.** Add up the frequencies. Why does the sum of the frequencies exceed the total sample size of 467?

**d.** Suppose 30 people suffer from both left leg rigidity and right leg rigidity. How many people in the sample suffer from rigidity in at least one of their legs?

**e.** Would it be reasonable to create a cumulative distribution for this data? Explain.

**13.** A small commuter airline in the West keeps records of complaints received from its customers. Complaints for March and July are listed in the following table.

| Customer Complaints | | |
|---|---|---|
| **Type of Complaint** | **March** | **July** |
| Tickets cost too much | 11 | 15 |
| Stewardess did not provide blankets | 8 | 3 |
| Schedules not convenient | 12 | 17 |
| Plane often late | 17 | 16 |
| Seats too stiff | 3 | 3 |
| Airplane too hot | 6 | 20 |
| Airplane too cold | 8 | 5 |
| Poor reservation system | 5 | 5 |
| Plane interior looks shabby | 5 | 6 |

**a.** Classify the items by the following categories: comfort, price, service, and schedule, and develop a qualitative frequency distribution.

**b.** Classify the items by the following categories: plane, personnel, building/equipment, and other, and develop a qualitative frequency distribution.

**c.** Would another person necessarily assign the same items to the same categories as you have?

**d.** Do the categories chosen in parts **a.** and **b.** meet the requirement that categories be mutually exclusive and exhaustive? Discuss.

**14.** Before purchasing a specific product, a consumer reviews the ratings provided by previous customers. The following table displays the product ratings given by customers on a scale of 1 star (worst) to 5 stars (best).

| Number of Stars | Frequency |
|---|---|
| 1 | 9 |
| 2 | 11 |
| 3 | 30 |
| 4 | 17 |
| 5 | 8 |

**a.** What level of measurement does the data possess?

**b.** How many total customers rated the product?

**c.** What percentage of customers rated the product with 4 stars?

**Data**

stat.hawkeslearning.com under **Discovering Statistics and Data, Fourth Edition > Data Sets > Mount Pleasant Real Estate Data**

**15.** Using the *House Style* variable from the Mount Pleasant Real Estate data set from the web resource, consider the following.

**a.** What level of measurement does the data possess?

**b.** Is the data qualitative or quantitative?

**c.** Use technology to construct a frequency distribution for the *House Style* variable.

**16.** The following data represents the area, in square feet, of thirty homes listed for sale in Washington County, Wisconsin.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1152 | 1200 | 1216 | 1331 | 1408 | 1474 | 1479 | 1492 |
| 1508 | 1560 | 1647 | 1652 | 1654 | 1654 | 1670 | 1843 |
| 1905 | 1920 | 1924 | 2030 | 2165 | 2169 | 2202 | 2207 |
| | 2314 | 2498 | 2710 | 3234 | 3251 | 3440 | |

**a.** What level of measurement does the data possess?

**b.** Determine the minimum number of classes necessary if a class width of 400 is used.

**c.** Construct a frequency distribution with five classes.

**d.** Construct a frequency distribution with eight classes.

**e.** Which frequency distribution do you think best displays the data? Explain.

**17.** A business magazine was conducting a study into the amount of travel required for mid-level managers across the U.S. Seventy-five managers were surveyed for the number of days they spent traveling each year.

| Mid-Level Manager Travel | | | | | | |
|---|---|---|---|---|---|---|
| **Days Traveling** | 0–6 | 7–13 | 14–20 | 21–27 | 28–34 | 35 and above |
| **Frequency** | 15 | 21 | 27 | 9 | 2 | 1 |

**a.** Construct a relative frequency distribution.

**b.** Construct a cumulative frequency distribution.

**18.** Every year, the average temperatures of 100 selected US cities are published by the National Oceanic and Atmospheric Administration (NOAA). The average temperature (°F) for the month of October for 15 randomly selected cities from the list of 100 are listed in the following table.[3]

| Average Temperatures (°F) | | | | |
|---|---|---|---|---|
| 68.5 | 50.9 | 67.5 | 57.5 | 56.0 |
| 47.1 | 50.1 | 65.8 | 51.5 | 49.5 |
| 75.2 | 56.0 | 62.3 | 53.0 | 46.1 |

**a.** Construct a frequency distribution for the average temperatures for the month of October.

**b.** Construct a relative frequency distribution for the average temperatures for the month of October.

**c.** Construct a cumulative frequency distribution for the average temperatures for the month of October.

**19.** The average temperatures (°F) for the month of January for forty randomly selected cities are listed in the following table.

| Average January Temperature ($^0$F) | Number of Cities |
|---|---|
| 14-22 | 1 |
| 23-31 | 4 |
| 32-40 | 11 |
| 41-49 | 10 |
| 50-58 | 8 |
| 59-67 | 4 |
| 68-76 | 2 |

**a.** What level of measurement does the data possess?

**b.** Choose the interval that contains the average January temperature for Boston, Massachusetts which is 36°F.

**c.** What percentage of cities have an average January temperature that is less than 32°F?

**d.** What percentage of cities have an average January temperature of 50°F or greater?

**e.** Explain how to use the relative frequency to determine the answer in part **c**.

**Data**

stat.hawkeslearning.com
**Discovering Statistics and Data, Fourth Edition > Data Sets > California DDS Expenditures**

**20.** Using the California DDS Expenditures data set from the web resource, perform the following.

**a.** Construct a frequency distribution with 8 classes for the *Expenditures* variable.

**b.** Construct a relative frequency distribution for the expenditures using the frequency distribution from part **a**.

**c.** Construct a cumulative frequency distribution for the expenditures.

## 3.2 Displaying Qualitative Data Graphically

Graphical analysis presents a trade-off: although we lose sight of the individual observations (the raw data), it allows us to create a representation of the data distribution that "speaks to the eyes." The trade is almost always beneficial since a well-designed graph gives our visual processing system the kind of image it processes best, a picture.

Because a set of data can be graphically represented in many different ways, selecting and creating graphical displays requires a certain amount of artistic judgment.

Several types of graphs and tabular displays will be discussed in this chapter. Bar charts, stacked bar charts, and pie charts are effective, visually appealing methods of graphically displaying qualitative data. A quick look at publications such as *Time*, *USA Today*, *The Wall Street Journal*, *Scientific American*, and *Forbes* provides convincing evidence of the frequent and beneficial usage of these data visualization techniques.