Name Answer Key

# Chapter 2 Organizing Data

## Section 2.1 Frequency Distributions, Histograms, and Related Topics

Objective: In this lesson you learned to organize raw data, construct graphs, and recognize distribution shapes.

| **Important Vocabulary** | | |
|---|---|---|
| Frequency Table | Class Width | Lower/Upper Class Limit |
| Lower/Upper Class Boundaries | Cumulative Frequency | |

What are three criteria a graphical display should have?

- show the data
- induce the viewer to thing about the substance of the graphic
- avoid distorting what the data have to say

### I. Frequency Table

Focus Point:
- Organize raw data using a frequency table

A **frequency table:** partitions data into classes or intervals of equal width and shows how many data values are in each class.

How to find the class width (interger data):

1. $\frac{\text{largest data} - \text{smallest data}}{\text{\# of classes}}$

2. Increase the value to the next highest whole number

The **lower class limit:** is the lowest data value that can fit in a class.

The **upper class limit:** is the highest data value that can fit in a class.

The **class width:** is the difference between the lower class limit of one class and the lower class limit of the next class

What is class frequency?

the number of tally marks corresponding to that class.

What is the midpoint of a class?

the center of each class.

How to find class boundaries (integer data):

- Add 0.5 to the upper class limit
- Subtract 0.5 from the lower class limit

What is relative frequency?

the proportion of all data values that fall into that class

How do you calculate relative frequency?

$$\text{Relative frequency} = \frac{f}{n} = \frac{\text{Class frequency}}{\text{Total frequencies}}$$

The total of the relative frequencies should be 1 ≈ .

Due to rounding, results might be above/below 1.

How to make a frequency table:

1. Determine the number of classes and class width.
2. Create the classes.
   - start with lower class limits
3. Fill in upper class limits
4. Tally the data into classes.
5. Compute the midpoint for each class (class mark)
6. Determine the class boundaries

How to make a relative frequency table:

After making a frequency table, compute the relative frequency (f/n) for each class.

## II. Histograms and Relative-Frequency Histograms

Focus Point:
- Construct histograms, relative-frequency histograms, and ogives.

How to make a (Relative-Frequency) Histogram:

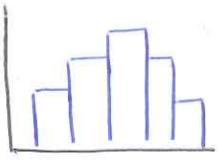
1. Make a frequency table (including relative frequencies) with the designated number of classes
2. Place class boundaries on the horizontal axis and (relative) frequencies on the vertical axis
3. Draw a bar the width of the class boundaries and the height of the (relative) frequency.

Why use class boundaries in histograms?

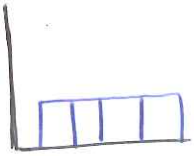
This assures us that the bars of the histogram touch and that no data fall on the boundaries.

## III. Distribution of Shapes

Focus Point:
- Recognize basic distribution shapes: uniform, symmetric, skewed, and bimodal

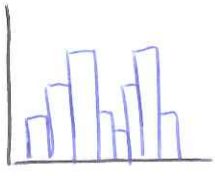


**Mound-shaped Symmetrical:** Both sides are (roughly) the same when the graph is folded vertically down the middle



**Uniform or Rectangular:** every class has equal frequency.


**Skewed left/right:** one tail (side) is stretched out longer than the other.



**Bimodal:** the two classes with the largest frequencies are separated by at least one class.

**Outliers:** data values that are very different from other measurements in the data set.

What do (relative-frequency) histograms tell us?

- if the data distribution is more symmetric, skewed, or bimodal.
- if there are possible outliers
- which data intervals contain the most data.
- how spread out the data are.

## IV. Cumulative-Frequency Tables and Ogives

Focus Point:
- Interpret graphs in the context of the data setting

The **cumulative frequency:** for a class is the sum of the frequencies for that class and all previous classes

What is an ogive?

a graph that displays cumulative frequencies

How to make an ogive:

1. Make a frequency table showing class boundaries and cumulative frequencies
2. For each class, make a dot over the upper class boundary at the height of the cumulative class frequency. Connect these dots with line segments.
3. By convention, an ogive begins on the horizontal axis at the lower class boundary of the first class.

What does an ogive tell us?

- How many data are less than the indicated value on the horizontal axis
- How slowly or rapidly the data values accumulate over the range of the data.

## Section 2.1 Examples – Frequency Distributions, Histograms, and Related Topics

( 1 ) An irate customer called Dollar Day Mail Order Company 40 times during the last two weeks to see why his order had not arrived. Each time he called, he recorded the length of time he was put "on hold" before being allowed to talk to a customer service representative. See table above.

**Length of Time on Hold, in Minutes**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 5 | 5 | 6 | 7 | 4 | 8 | 7 | 6 | 5 |
| 5 | 6 | 7 | 6 | 6 | 5 | 8 | 9 | 9 | 10 |
| 7 | 8 | 11 | 2 | 4 | 6 | 5 | 12 | 13 | 6 |
| 3 | 7 | 8 | 8 | 9 | 9 | 10 | 9 | 8 | 9 |

a. What are the largest and smallest values in the table? If we want five classes in a frequency table, what should the class width be?

The largest value is 13.

The smallest value is 1.

$\text{class width} = \frac{13-1}{5} = 2.4$

↑ Round up!

class width = 3

b. Complete the following frequency table.

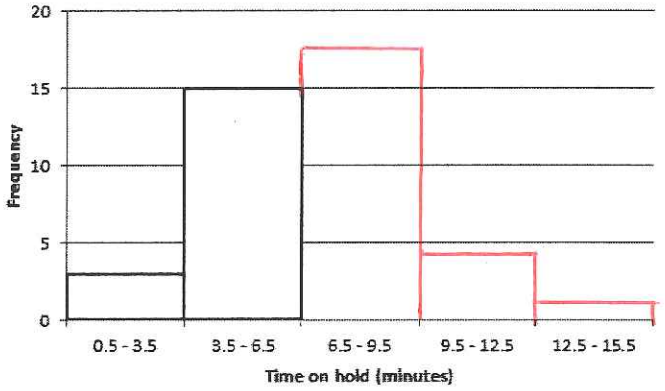
**Time on Hold**

| Class Limits Lower-Upper | Tally | Frequency | Midpoint |
|---|---|---|---|
| 1 - 3 | III | 3 | 2 |
| 4 - 6 | 𝍸 𝍸 | 15 | 5 |
| 7 - 9 | 𝍸 𝍸 II | 17 | 8 |
| 10 - 12 | IIII | 4 | 11 |
| 13 - 15 | I | 1 | 14 |

c. Recall that the class boundary is halfway between the upper limit of one class and the lower limit of the next. Use this fact to find the class boundaries in the table below (left) and to complete the partial histogram below (right).

**Class Boundaries**

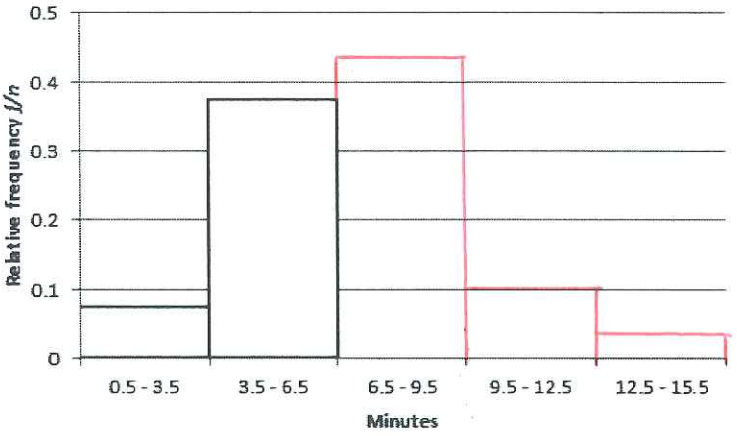
| Class Limits | Class Boundaries |
|---|---|
| 1 - 3 | 0.5 - 3.5 |
| 4 - 6 | 3.5 - 6.5 |
| 7 - 9 | 6.5 - 9.5 |
| 10 - 12 | 9.5 - 12.5 |
| 13 - 15 | 12.5 - 15.5 |



d. Compute the relative class frequency $f/n$ for each class in the table below (left) and complete the partial relative-frequency histogram below (right).

**Relative Class Frequency**

| Class | $f/n$ |
|---|---|
| 1 - 3 | $3/40 = 0.075$ |
| 4 - 6 | $15/40 = 0.375$ |
| 7 - 9 | $17/40 = 0.425$ |
| 10 - 12 | $4/40 = 0.100$ |
| 13 - 15 | $1/40 = 0.025$ |



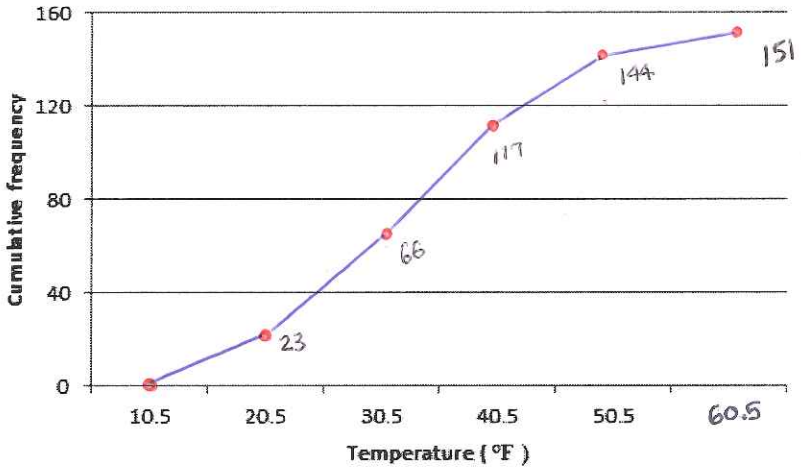

( 2 ) Aspen, Colorado, is a world-famous ski area. If the daily high temperature is above 40°F, the surface of the snow tends to melt. It then freezes again at night. This can result in a snow crust that is icy. It also can increase avalanche danger. The table below gives a summary of daily high temperatures (°F) in Aspen during the 151-day ski season.

a. The cumulative frequency for a class is computed by adding the frequency of that class to the frequencies of the previous classes. Complete the table below.

**High Temperatures During the Aspen Ski Season (°F)**

| Class Boundaries | | | |
|---|---|---|---|
| Lower | Upper | Frequency | Cumulative Frequency |
| 10.5 | 30.5 | 23 | 23 |
| 20.5 | 30.5 | 43 | 66 (sum 23 + 43) |
| 30.5 | 40.5 | 51 | 117 |
| 40.5 | 50.5 | 27 | 144 |
| 50.5 | 60.5 | 7 | 151 |

b. To draw the corresponding ogive, we place a dot a cumulative frequency 0 on the lower class boundary of the first class. Then we place dots over the *upper class boundaries* at the height of the cumulative class frequency for the corresponding class. Finally, we connect the dots. Complete the ogive below.




c. Looking at the ogive, estimate the total number of days with a high temperature lower than or equal to 40°F.

Approximately 117 days had high temperatures lower than or equal to 40°F