# 9.3 Histograms and Box Plots

**Essential Question:** How can you interpret and compare data sets using data displays?

## 🧭 Explore  Understanding Histograms

A **histogram** is a bar graph that is used to display the frequency of data divided into equal intervals. The bars must be of equal width and should touch but not overlap. The heights of the bars indicate the frequency of data values within each interval.

(A) Look at the histogram of "Scores on a Math Test." Which axis indicates the frequency?

The _____ axis shows the frequency for each interval.

(B) What does the horizontal axis indicate?

The horizontal axis shows the _____.
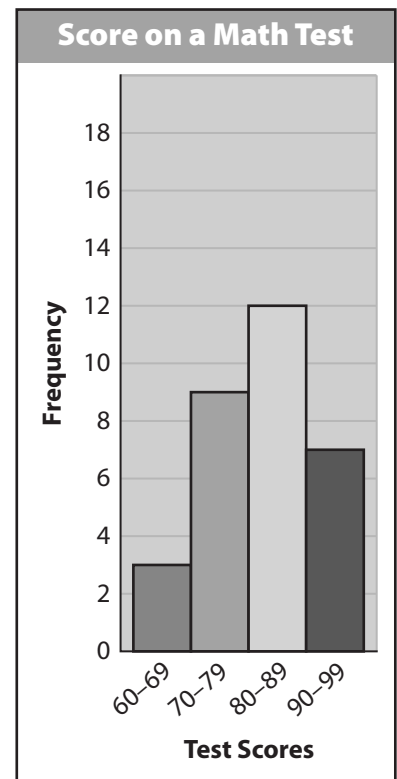
(C) How is the horizontal axis organized?

It is organized in groups of _____.

(D) How many had scores in the interval 60–69? ☐

(E) How many had scores in the interval 70–79? ☐

(F) How many had scores in the interval 80–89? ☐

(G) How many had scores in the interval 90–99? ☐

**Score on a Math Test**

*Frequency vs Test Scores*

### Reflect

**1.** What statistical information can you tell about a data set by looking at a histogram? What statistical information cannot be determined by looking at a histogram?

_____

_____

**2.** How many test scores were collected? How do you know?

_____

_____

## ⚙ Explain 1   Creating Histograms

When creating a histogram, make sure that the bars are of equal width and that they touch without overlapping. Create a frequency table to help organize the data before constructing the histogram. Consider the range of the data values when creating intervals.

**Example 1**    **Create a frequency table from the data. Then use the frequency table to create a histogram.**

(A)   Listed are the ages of the 100 U.S. senators at the start of the 112th Congress on January 3, 2011.

39, 39, 42, 44, 46, 47, 47, 47, 48, 49, 49, 49, 50, 50, 51, 51,
52, 52, 53, 53, 54, 54, 55, 55, 55, 55, 55, 55, 56, 56, 57, 57,
57, 58, 58, 58, 58, 58, 59, 59, 59, 59, 60, 60, 60, 60, 60, 60,
60, 61, 61, 62, 62, 62, 63, 63, 63, 63, 64, 64, 64, 64, 66, 66,
66, 67, 67, 67, 67, 67, 67, 67, 68, 68, 68, 68, 69, 69, 69, 70,
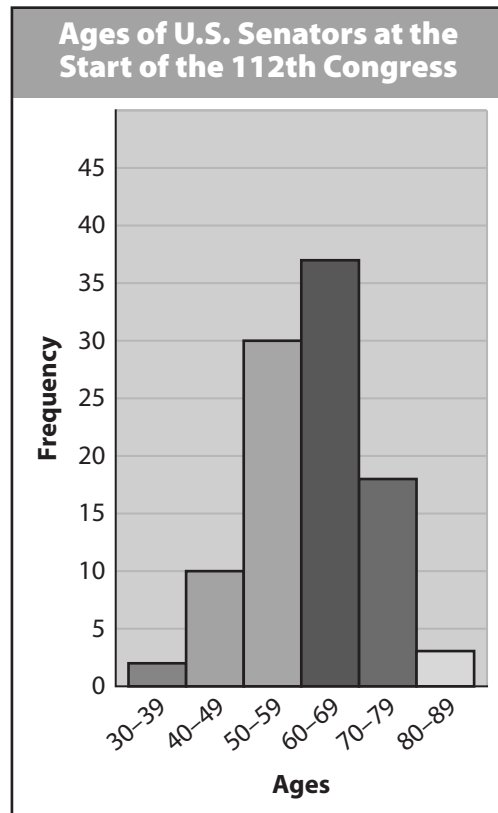70, 70, 71, 71, 73, 73, 74, 74, 74, 75, 76, 76, 76, 76, 77, 77,
78, 86, 86, 86

Create a frequency table. The data values range from 39 to 86, so use an interval width of 10 and start the first interval at 30.

| Age Interval | Frequency |
|:---:|:---:|
| 30–39 | 2 |
| 40–49 | 10 |
| 50–59 | 30 |
| 60–69 | 37 |
| 70–79 | 18 |
| 80–89 | 3 |

Check that the sum of the frequencies is 100.

$2 + 10 + 30 + 37 + 18 + 3 = 100$

Use the frequency table to create a histogram.

**Ages of U.S. Senators at the Start of the 112th Congress**

(B) Listed are the scores from a golf tournament.

68, 78, 76, 71, 69, 73, 72, 74, 76, 70, 77, 74, 75, 76, 71, 74

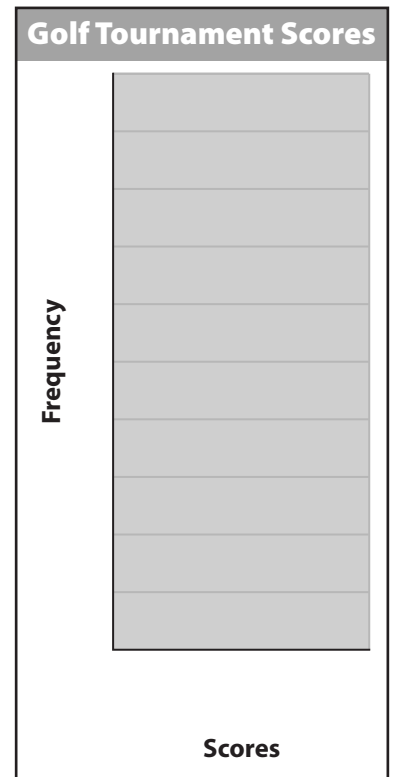Create a frequency table. The data values range from [   ] to [   ],

so use an interval width of 3, and start the first interval at [   ].

**Golf Tournament Scores**

| Score Interval | Frequency |
|---|---|
| 68–[   ] | [   ] |
| 71–[   ] | [   ] |
| 74–[   ] | [   ] |
| 77–[   ] | [   ] |

Check that the sum of the frequencies is [   ].

$3 + 4 + 7 + 2 =$ [   ]

Use the frequency table to create a histogram.
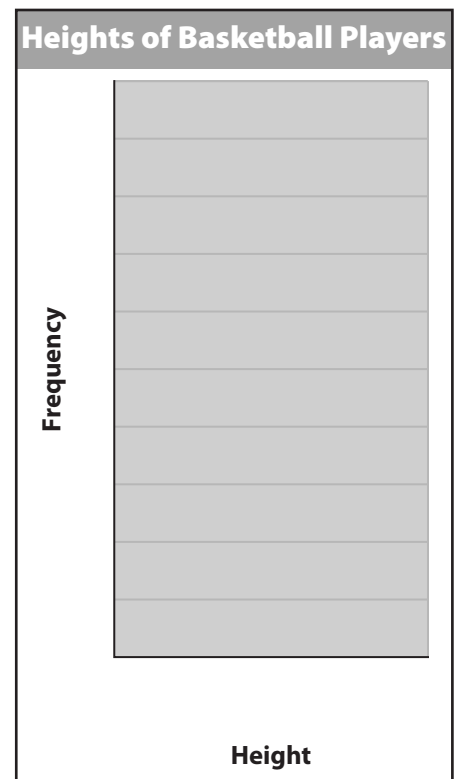


Frequency

Scores

---

**Reflect**

**3.** Describe the shape of the distribution of senators' ages. Interpret the meaning.

_____

_____

**Your Turn**

**4.** Listed are the heights of players, in inches, on a basketball team. Create a frequency table from the data. Then use the frequency table to create a histogram.

79, 75, 74, 68, 63, 76, 74, 73, 69, 65, 71, 68, 74, 73, 70

**Heights of Basketball Players**



Frequency

Height

## ⚙ Explain 2  **Estimating from Histograms**

You can estimate statistics by studying a histogram.

**Example 2**  **Estimate the mean of the data set displayed in each histogram.**

(A)  The histogram shows the ages of teachers in a high school.

To estimate the mean, first find the midpoint of each interval, and multiply by the frequency.

1st interval: $\left(\dfrac{20 + 29}{2}\right)(20) = (24.5)(20) = 490$

2nd interval: $\left(\dfrac{30 + 39}{2}\right)(25) = (34.5)(25) = 862.5$

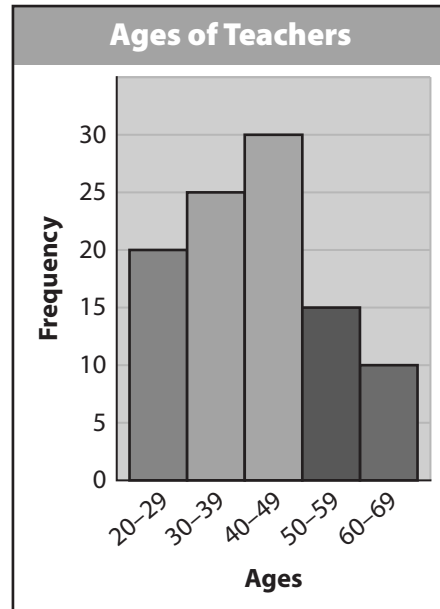3rd interval: $\left(\dfrac{40 + 49}{2}\right)(30) = (44.5)(30) = 1335$

4th interval: $\left(\dfrac{50 + 59}{2}\right)(15) = (54.5)(15) = 817.5$

5th interval: $\left(\dfrac{60 + 69}{2}\right)(10) = (64.5)(10) = 645$

Add the products and divide by the sum of the frequencies.

Mean: $\dfrac{490 + 862.5 + 1335 + 817.5 + 645}{20 + 25 + 30 + 15 + 10} = \dfrac{4150}{100} = 41.5$

A good estimate for the mean of this data set is an age of 41.5.

**Ages of Teachers**



(B)  The histogram shows the 2012 Olympic results for women's weightlifting.

To estimate the mean, first find the midpoint of each interval, and multiply by the frequency.

1st interval: $\left(\dfrac{160 + 179}{2}\right)\boxed{\phantom{x}} = \boxed{\phantom{xxx}}\,\boxed{\phantom{x}} = \boxed{\phantom{xxx}}$

2nd interval: $\left(\dfrac{180 + 199}{2}\right)\boxed{\phantom{x}} = \boxed{\phantom{xxx}}\,\boxed{\phantom{x}} = \boxed{\phantom{xxx}}$

3rd interval: $\left(\dfrac{200 + 219}{2}\right)\boxed{\phantom{x}} = \boxed{\phantom{xxx}}\,\boxed{\phantom{x}} = \boxed{\phantom{xxx}}$

4th interval: $\left(\dfrac{220 + 239}{2}\right)\boxed{\phantom{x}} = \boxed{\phantom{xxx}}\,\boxed{\phantom{x}} = \boxed{\phantom{xxx}}$

5th interval: $\left(\dfrac{240 + 259}{2}\right)\boxed{\phantom{x}} = \boxed{\phantom{xxx}}\,\boxed{\phantom{x}} = \boxed{\phantom{xxx}}$

Add the results and divide by the total number of values in the data set.

Mean: $\dfrac{169.5 + 379 + 838 + 2065.5 + 249.5}{1 + 2 + 4 + 9 + 1} = \dfrac{\boxed{\phantom{xxx}}}{17} \approx \boxed{\phantom{xxx}}$

A good estimate for the mean of this data set is $\boxed{\phantom{xxx}}$ kg.

**Women's Weightlifting**
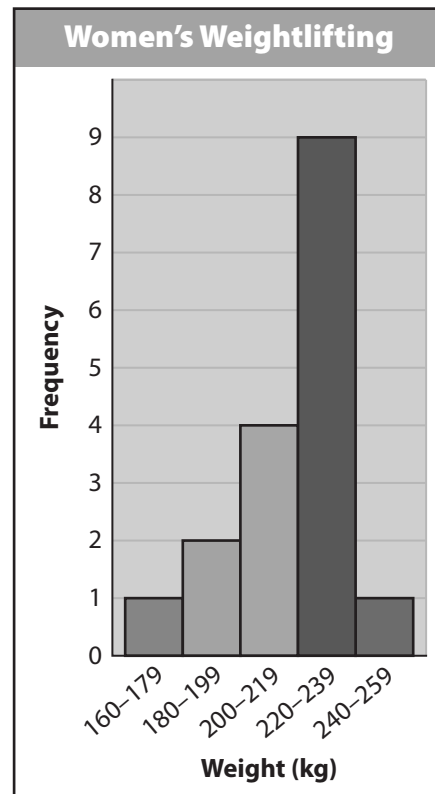
**5.** The histogram shows the length, in days, of Maria's last vacations. Estimate the mean of the data set displayed in the histogram.

**Maria's Vacations**
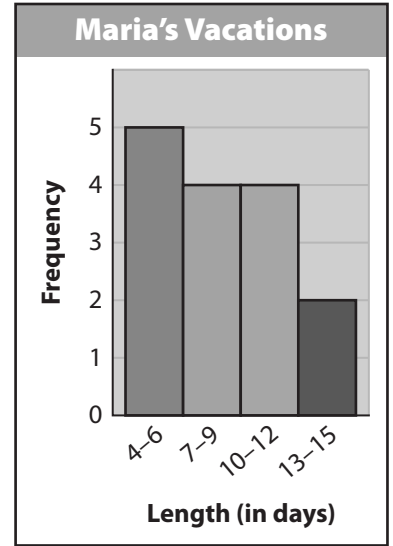


## Explain 3   Constructing Box Plots

A box plot can be used to show how the values in a data set are distributed. You need 5 values to make a box plot: the minimum (or least value), first quartile, median, third quartile, and maximum (or greatest value).

**Example 3**   **Use the data to make a box plot.**

Ⓐ  The numbers of runs scored by a softball team in 20 games are given.

3, 4, 8, 12, 7, 5, 4, 12, 3, 9, 11, 4, 14, 8, 2, 10, 3, 10, 9, 7

Order the data from least to greatest.

2, 3, 3, 3, 4, 4, 4, 5, 7, 7, 8, 8, 9, 9, 10, 10, 11, 12, 12, 14

Identify the 5 needed values. Those values are the minimum, first quartile, median, third quartile, and maximum.

| 2, | 3, 3, 3, | 4, 4, | 4, 5, 7, | 7, 8, | 8, 9, 9, | 10, 10, | 11, 12, 12, | 14 |

| Minimum | $Q_1$ | Median | $Q_3$ | Maximum |
|---|---|---|---|---|
| 2 | 4 | 7.5 | 10 | 14 |

Draw a number line and plot a point above each of the 5 needed values. Draw a box whose ends go through the first and third quartiles, and draw a vertical line segment through the median. Draw horizontal line segments from the box to the minimum and maximum.
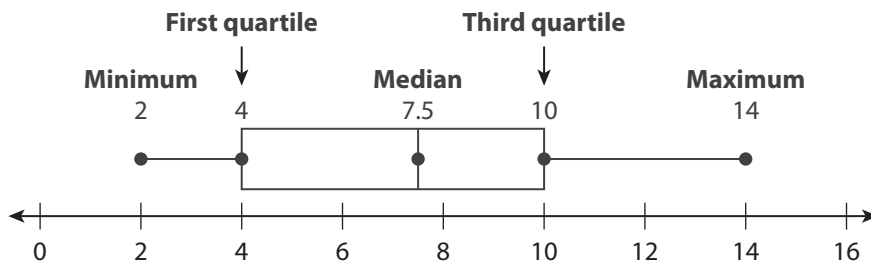
(B)  13, 14, 18, 13, 12, 17, 15, 12, 13, 19, 11, 14, 14, 18, 22, 23

Order the data from least to greatest.

☐ , 12, 12, ☐ , ☐ , 13, 14, ☐ , ☐ , 15, 17, ☐ , ☐ , 19, 22 ☐

Identify the 5 needed values. Those values are the minimum, first quartile, median, third quartile, and maximum.

☐ , 12, 12, ☐ 13, 14, ☐ 15, 17, ☐ 19, 22, ☐

↑              ↑              ↑              ↑              ↑

Minimum       $Q_1$        Median        $Q_3$        Maximum

☐              ☐              ☐              ☐              ☐

Draw a number line and plot a point above each of the 5 needed values. Draw a box whose ends go through the first and third quartiles, and draw a vertical line through the median. Draw horizontal lines from the box to the minimum and maximum.



10      15      20      25
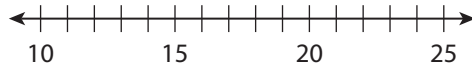
### Reflect

6.  The lines that extend from the box in a box plot are sometimes called "whiskers." What part (lower, middle, or upper) and about what percent of the data does the box represent? What part and about what percent does each "whisker" represent?

_____

_____

7.  Which measures of spread can be determined from the box plot, and how are they found? Calculate each measure.
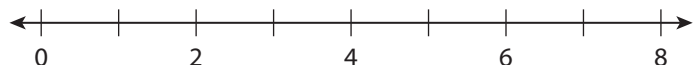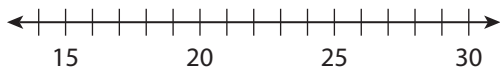
_____

_____

### Your Turn

**Use the data to make a box plot.**

8.  25, 28, 26, 16, 18, 15, 25, 28, 26, 16

9.  The numbers of goals scored by Lisa's soccer team in 13 games are listed below.

    2, 3, 4, 1, 1, 3, 4, 2, 6, 2, 2, 3, 2



15      20      25      30



0      2      4      6      8

## ⚙ Explain 4   Comparing Box Plots

You can plot two box plots above a single number line to compare two data sets.

**Example 4**   Construct two box plots, one for each data set. Compare the medians and measures of variation for each distribution.

Ⓐ   The tables show the total gross earnings, in dollars, for the top 10 movies of 2013 and 2012.

| Total Gross Earnings by the Top 25 Movies of 2013 & 2012 | | | | |
|---|---|---|---|---|
| **Rank** | **Total Gross in 2013** | | **Rank** | **Total Gross in 2012** |
| 1 | $420,468,544 | | 1 | $623,357,910 |
| 2 | $409,013,994 | | 2 | $448,139,099 |
| 3 | $368,061,265 | | 3 | $408,010,692 |
| 4 | $352,946,000 | | 4 | $304,360,277 |
| 5 | $291,045,518 | | 5 | $303,003,568 |
| 6 | $268,492,764 | | 6 | $292,324,737 |
| 7 | $262,547,000 | | 7 | $262,030,663 |
| 8 | $253,029,814 | | 8 | $237,283,207 |
| 9 | $238,679,850 | | 9 | $218,815,487 |
| 10 | $234,911,825 | | 10 | $216,391,482 |

### 🧩 Analyze Information

For each set of data, identify the five values you need to make a box plot: the

minimum, _____. In this

case, the data is from least to _____ reading from the bottom to the top.
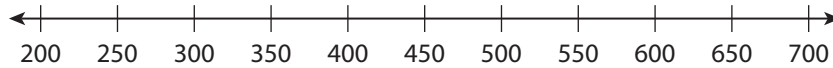
### 🧩 Formulate a Plan

With the _____ needed values for each data set, construct 2 _____

on the same number line. The number line for both plots can go from $ _____

million to $ _____ million. Interpret the box plots to compare the gross
earnings for the top 10 movies in 2012 and 2013.

Using the statistics feature of a graphing calculator, find the five needed values, rounded to the nearest hundred million.

| | 2013 Top 10 Movies Gross Earnings ($100,000,000s) | 2012 Top 10 Movies Gross Earnings ($100,000,000s) |
|---|---|---|
| Minimum | | |
| $Q_1$ | | |
| Median | | |
| $Q_3$ | | |
| Maximum | | |

```
  200  250  300  350  400  450  500  550  600  650  700
```

The medians for both years are very close together, but gross earnings for _____

did not vary as much as in _____. The range and interquartile range were

greater for the _____ data than for the _____ data. The maximum

value in the 2012 data appears to be an outlier, but because _____ is not greater

than _____, it is not an outlier.

Considering the little difference between minimum values but the great difference between maximum values the data sets, it makes sense that their measures of variation would / would not be alike.
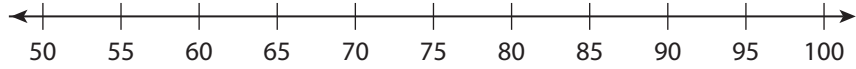
**Your Turn**

**Construct two box plots, one for each data set. Compare the medians and measures of variation for each distribution.**

**10.** The net worth of the 10 richest people in the world for 2012 and 2013 (in billions) are:
**2012:** 69, 61, 44, 41, 37.5, 36, 30, 26, 25.5, 25.4   **2013:** 73, 67, 57, 53.5, 43, 34, 34, 31, 30, 29

```
  20    25    30    35    40    45    50    55    60    65    70    75    80
```
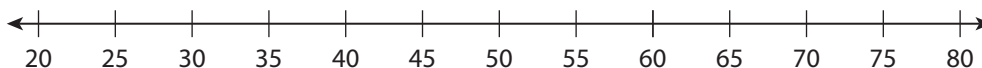
**11.** The ages of the 10 richest people in the world for 2012 and 2013 (in years) are:

**2012**: 72, 56, 81, 63, 75, 67, 55, 64, 83, 92          **2013**: 72, 57, 76, 82, 68, 77, 72, 84, 90, 63

```
←——┼————┼————┼————┼————┼————┼————┼————┼————┼————┼————┼——→
   50    55    60    65    70    75    80    85    90    95   100
```

## ⊙ Elaborate

**12.** How can you create a histogram from a data set?

_____

_____

_____

**13.** How can you create a box plot from a data set?

_____

_____

_____

_____

**14.** **Essential Question Check-In**  How can you use histograms and box plots to interpret and compare data sets?

_____

_____

## ✪ Evaluate: Homework and Practice

**Use the histogram to answer the following questions.**

**1.** What does each axis indicate?

**2.** How is the horizontal axis organized?

**3.** How many bowlers competed?

**4.** Describe the general shape of the distribution.
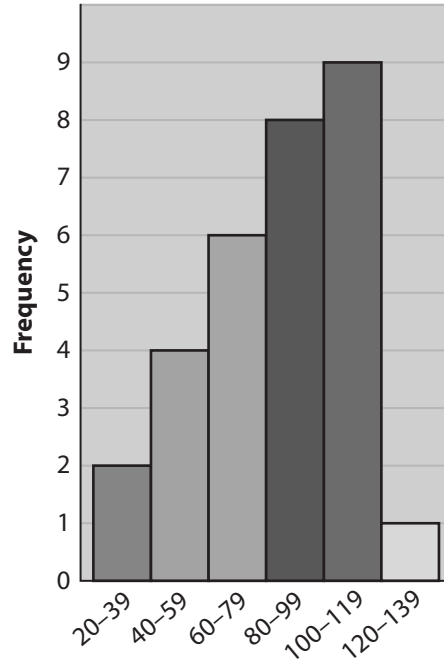


**Bowlers' Scores**

**Create a histogram for the given data.**

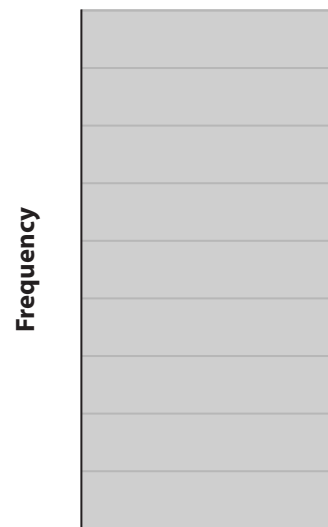**5.** Listed are the ages of the first 44 U.S. presidents on the date of their first inauguration.

57, 61, 57, 57, 58, 57, 61, 54, 68, 51, 49, 64, 50, 48, 65, 52, 56, 46, 54, 49, 51, 47, 55, 55, 54, 42, 51, 56, 55, 51, 54, 51, 60, 62, 43, 55, 56, 61, 52, 69, 64, 46, 54, 47



**Ages**

**6.** Listed are the breathing intervals, in minutes, of gray whales.

8, 5, 13, 7 ,16, 9, 15, 11, 8, 6, 10, 9, 9, 11, 14, 12, 13, 15, 16, 11, 14, 9, 15, 6, 14



**Time (min)**

**7.** Listed are the heights, in inches, of the students in Marci's karate class.

42, 44, 47, 50, 51, 53, 53, 55, 56, 57, 57, 58, 59, 60, 66



**Height (in.)**
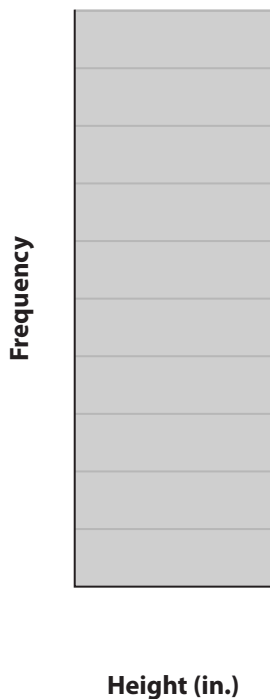
**8.** Listed are the starting salaries, in thousands of dollars, for college graduates.

34, 20, 32, 45, 32, 48, 34, 32, 20, 35, 34, 32, 40, 47, 21, 37, 21, 47, 30, 31, 40, 31, 21, 22, 30, 22, 34, 48, 35, 37, 22, 46, 38, 39, 45, 37, 52, 25, 26, 26, 27, 43, 34, 28, 55, 29, 31, 42, 24, 21, 42, 42, 31, 30, 20, 39, 23, 41, 24, 33, 49, 24, 36, 36, 23, 38, 33, 33, 54
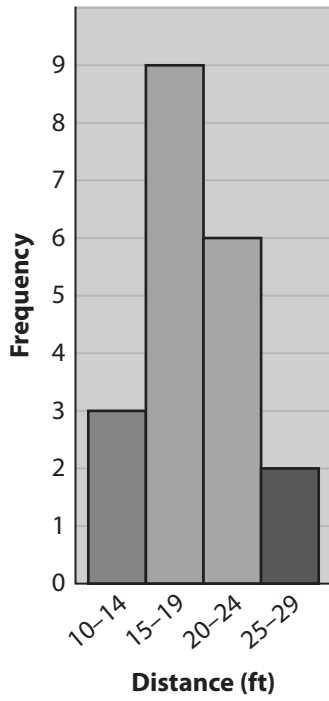


**Salary Range (thousands $)**
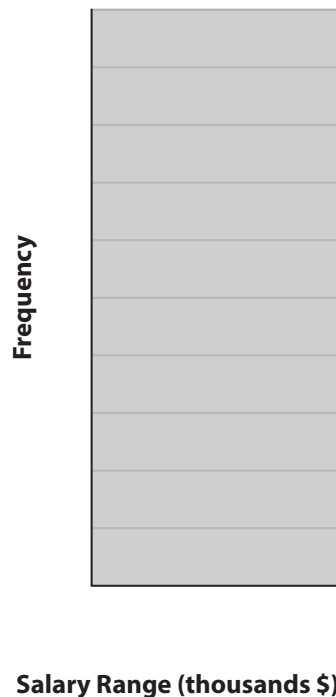
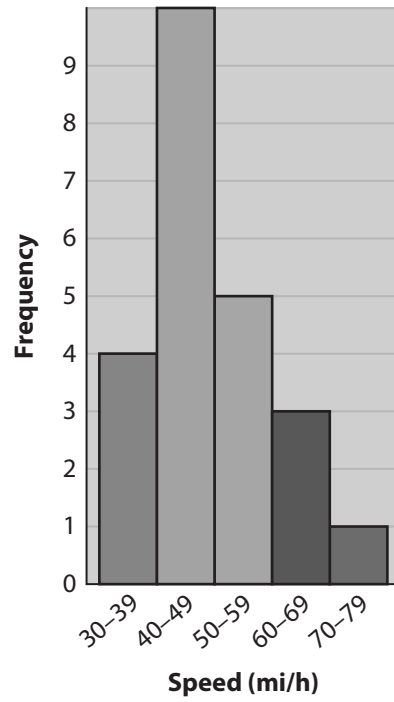**Estimate the mean of the data set displayed in each histogram.**

**9.** The histogram shows the GPAs of the students in George's class.
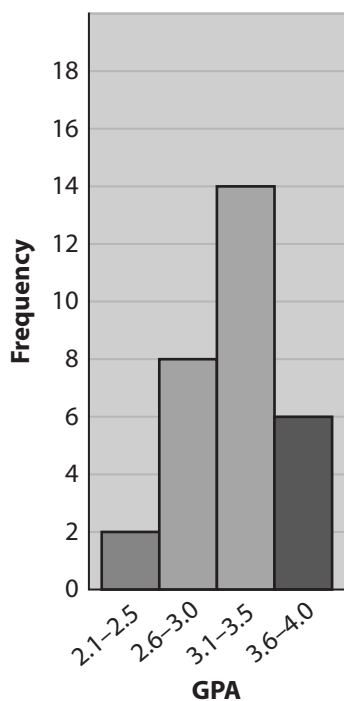
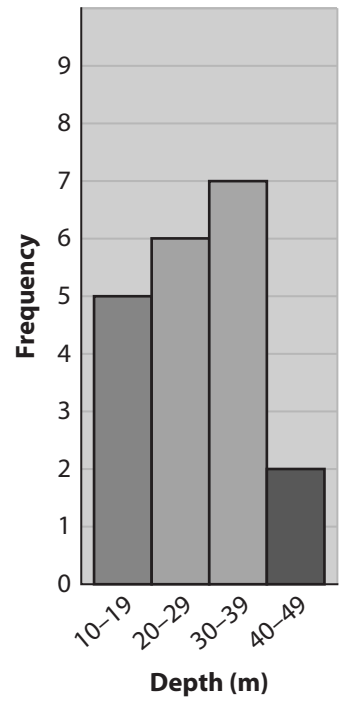**10.** The histogram shows long jump distances, in feet, for a track and field team.



**Distance (ft)**

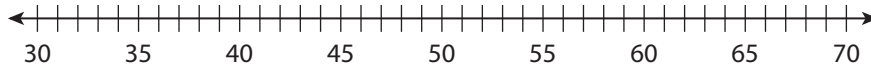**11.** The histogram shows the speeds of downhill skiers, in miles per hour, during a competition.



**Speed (mi/h)**

**12.** The histogram shows the depths a diver has been to, in meters.



**Depth (m)**
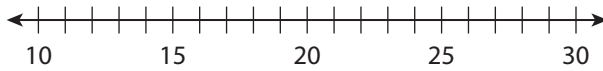
**Use the data to make a box plot.**

**13.** The numbers of points Julia's basketball team scored in 11 games are listed below.

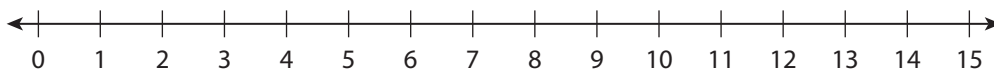50, 62, 37, 36, 34, 44, 44, 36, 37, 42, 36



**14.** The numbers of baskets Kelly's team scored in 9 games are listed below.
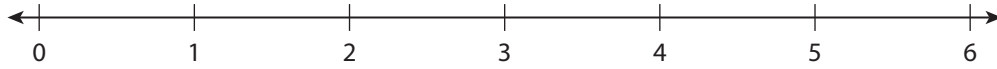
14, 24, 29, 15, 16, 20, 24, 15, 15



**15.** The numbers of runs Jane's baseball team scored in 9 games are listed below.
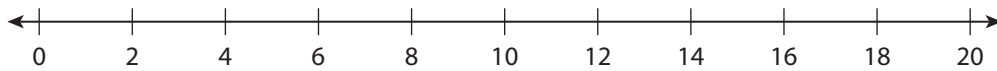
0, 2, 3, 2, 4, 11, 3, 4, 3

**16.** The numbers of field goals James' football team scored in 12 games are listed below.
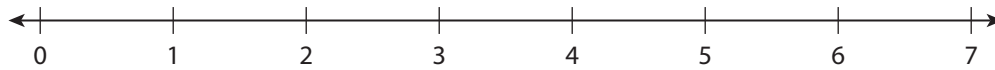
3, 2, 5, 3, 5, 2, 1, 0, 5, 4, 4, 2

```
◄─┼────────┼────────┼────────┼────────┼────────┼────────┼─►
  0        1        2        3        4        5        6
```

**17.** The numbers of points scored by Jane in 8 basketball games are listed below.

0, 13, 12, 2, 0, 16, 3, 14

```
◄─┼────┼────┼────┼────┼────┼────┼────┼────┼────┼────►
  0    2    4    6    8    10   12   14   16   18   20
```

**18.** The numbers of goals Claudia's soccer team scored in 21 games are shown below.

0, 5, 4, 3, 3, 2, 1, 1, 6, 2, 2, 1, 1, 1, 2, 2, 0, 1, 0, 1, 4

```
◄─┼────────┼────────┼────────┼────────┼────────┼────────┼────────┼─►
  0        1        2        3        4        5        6        7
```

**19.** Mario and Carlos, two brothers, play for the same basketball team. Here are the points they scored in 10 games:

| Game | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|------|---|---|---|---|---|---|---|---|---|----|
| Mario | 15 | X | X | 12 | 7 | 11 | 12 | 11 | 10 | 11 |
| Carlos | 10 | X | 9 | 12 | 15 | X | 19 | 11 | 12 | 12 |

(Xs mark games each one missed.) Which brother had the highest-scoring game? How many more points did he score in that game than the other brother did in his highest-scoring game?

```
◄─┼────┼────┼────┼────┼────┼────┼────┼────┼────┼────┼────┼────┼────┼────►
  6    7    8    9    10   11   12   13   14   15   16   17   18   19   20
```
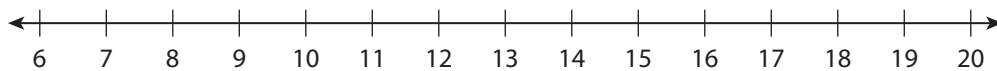
**20. Communicate Mathematical Ideas** Describe how you could estimate the IQR of a data set from a histogram.

**21. Critical Thinking** Suppose the minimum in a data set is the same as the first quartile. How would this affect a box plot of the data? Explain.

**22. Draw Conclusions** Dolly and Willie's scores are shown. Dolly claims that she is the better student, but Willie claims that he is the better student. What statistics make either Dolly or Willie seem like the better student? Explain.
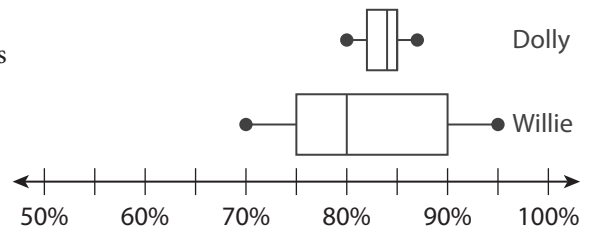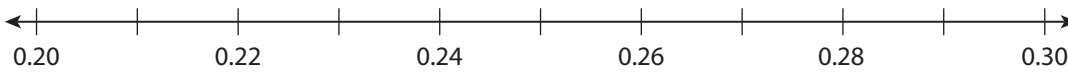
# Lesson Performance Task

The batting averages for the starting lineup of two competing baseball teams are given in the table.

| Batting Averages for Team A | 0.270 | 0.260 | 0.300 | 0.290 | 0.260 | 0.280 | 0.280 | 0.240 | 0.270 | 0.280 |
|---|---|---|---|---|---|---|---|---|---|---|
| Batting Averages for Team B | 0.290 | 0.270 | 0.240 | 0.280 | 0.230 | 0.280 | 0.230 | 0.270 | 0.250 | 0.270 |

**a.** Create a box plot for each data set on the same number line.

**Batting Averages**



**b.** Compare the distributions of the batting averages for each team.

**c.** Which team has a better starting lineup? Explain.