## Measures of Position

In this section, we will learn to use fractiles. **Fractiles** are numbers that partition, or divide, an ordered data set into equal parts (each part has the same number of data entries). The median is a fractile because it divides the data into two equal parts.

The three **quartiles,** $Q_1$, $Q_2$, and $Q_3$, divide an ordered data set into four equal parts. About one-quarter of the data fall on or below the **first quartile $Q_1$,** About one-half of the data fall on or below the **second quartile $Q_2$** (the second quartile is the same as the median of the data set). About three-quarters of the data fall on or below the **third quartile $Q_3$.**

The number of nuclear power plants in the top 15 nuclear power-producing countries are listed Find the first, second, and third quartiles of the data set. What do you observe?

7 20 16 6 58 9 20 50 23 33 8 10 15 16 104

First order the data from smallest to largest and find the median.

6 7 8 9 10 15 16 16 20 20 23 33 50 58 104

first quartile (9) — median (16) — second quartile (33)

Then find the median of the left side and the median of the right side.

Interpretation: about 25% of the countries have fewer than 9 nuclear power plants; 50% have fewer than 16 nuclear power plants; and 75% have fewer than 33 nuclear power plants.

Using a TI - 84:

STAT Enter Clear the list then enter new list

STAT CALC 1 - Var Stats Enter 2nd List $L_1$ ENTER ENTER ENTER

Scroll down to n = This is the information for the quartiles

Try Example 2 at top of P. 103. See if your information matches what the book has.

The **interquartile range (IQR)** of a data set is a measure of variation that gives the range of the middle portion (about half) of the data. The IQR is the difference between the third and first quartiles.

$$IQR = Q_3 - Q_1$$

Remember that an outlier is any data that is far removed from the other entries of a data set. Use the guidelines below to identify any outliers.

**Using the Interquartile Range to Identify Outliers**

**1.** Find the first ($Q_1$) and third ($Q_3$) quartiles of the data set.

**2.** Find the Interquartile range: IQR = Q3 - Q1

**3.** Multiply IQR by 1.5: 1.5(IQR)

**4.** Subtract 1.5(IQR) from Q1. Any data entry less than $Q_1$ - 1.5(IQR) is an outlier.

**5.** Add 1.5(IQR) to $Q_3$. Any data entry greater than Q3 + 1.5(IQR) is an outlier.

Looking back at example 1 (NUCLEAR POWER PLANTS), we can determine if theree are any ouliers. $Q_1$ = 9 and $Q_3$ = 33. The IQR = 33 - 9 = 24. 1.5 x 24 = 36. so

9 - 36 = -27 anything less than this is an outlier

33 + 36 = 69 anything greater than this is an outlier.

There are no lower outliers, but 104 is greater than 69, so 104 is an outlier.

---

Another important application of quartiles is to represent data sets using a **box-and-whisker plot (**or **boxplot).** Thee box-and-whisker plot is an exploratory data analysis tool that highlights important features of a data set. To graph a box-and-whisker plot, you must know the following values.

**1. The minimum entry**

**2. The first quartile $Q_1$**

**3. The median $Q_2$**

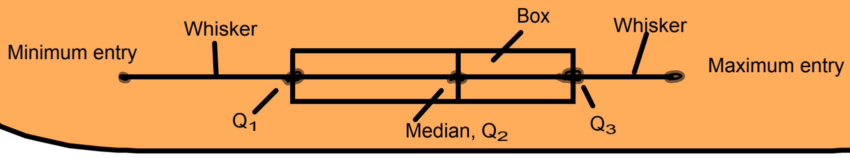
**4. The third quartile $Q_3$**

**5. The maximum entry.**

**This is called the five-number summary.**

**Drawing a Box-and-Whisker Plot**

1. Find the five-number summary of the data.
2. Construct a horizontal scale that spans the range of the data.
3. Plot the five numbers above the horizontal scale.
4. Draw a box above the horizontal scale from $Q_1$ to $Q_3$ and draw a vertical line in the box at $Q_2$.
5. Draw whiskers from the box to the minimum and maximum entries.




Using the nuclear power plant example, we can draw a box-and-whisker plot.

Minimum = 6 $Q_1 = 9$ $Q_2 = 16$ $Q_3 = 33$ maximum = 104

## Percentiles and Other Fractiles

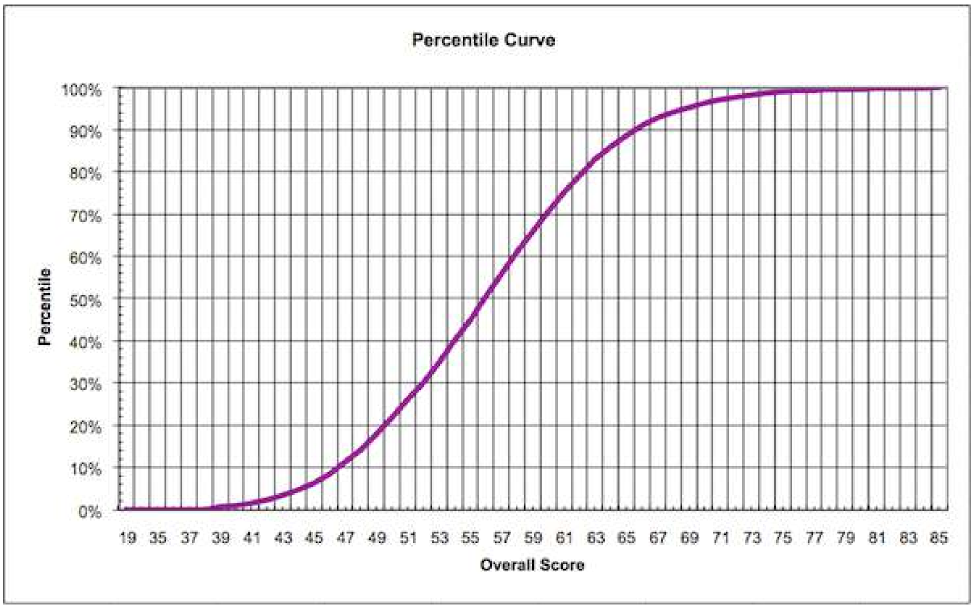
**Quartiles - divide a data set into 4 equal parts.**

**Deciles - divide a data set into 10 equal parts.**

**Percentiles - divide a data set into 100 equal parts.**

Percentiles are used in education and health-related fields to indicate how one individual compares with others in a group. Percentiles can also be used to identify unusually high or unusually low values. Measurements in the 95th percentile and above are unusually high, while those in the 5th percentile and below are unusually low.

A percentile score indicates a percent of individuals scored that value or below.




This is a graph that represents the percentile rankings of a medical school admissions test. Suppose someone scored a 61 on the admission test. Find 61 on the horizontal and see where it crosses the graph. It appears that the graph passes through 61 at the 78th percentile.

Turn to page 106 and do Try It Yourself 5.

To find the **percentile that corresponds to a specific data entry x,** use the formula

$$\text{Percentile of } x = \frac{\text{number of data entries less than } x}{\text{total number of data entries}} \cdot 100$$

and then round to the nearest whole number.

The following is the tuition costs of 25 liberal colleges in thousands of dollars. Find the percentile that corresponds to $30,000.

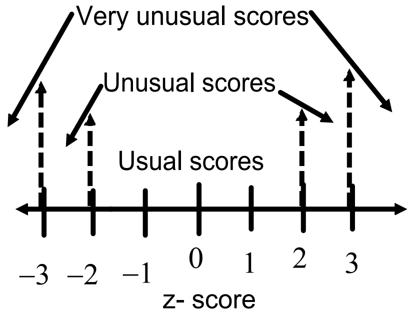
| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 19 | 22 | 23 | 26 | 26 | 27 | 27 | 27 | **30** | 31 | 32 | 32 |
| 33 | 34 | 34 | 38 | 40 | 41 | 42 | 44 | 45 | 45 | 46 | |

## The Standard Score

The **standard score,** or **z-score,** represents the number of standard deviations a value *x* lies from the mean μ. To find the *z*-score for a value, use the formula

$$z = \frac{\text{Value - Mean}}{\text{Standard Deviation}} = \frac{x - \mu}{\sigma}$$

Remember that the empirical rule stated that in a bell-shaped curve that about 95% of the data lies within 2 standard deviations of the mean.  A z-score outside of this range would be considered unusual.




A z-score gives a precise placement among the standard deviation of a data entry.   A z-score of 2.3 means 2.3 standard deviations to the right of the mean.  A z-score of  -3.5 means 3.5 standard deviations to the left of the mean.

The monthly utility bills in a city have a mean of $70 and a standard deviation of $8.  Find the z-scores that correspond to the utility bills of $60, $71, and $92.

$$\frac{x-\mu}{\sigma} = \frac{60-70}{8} = \frac{-10}{8} = -1.25$$

A $60 bill falls within 1.25 standard deviations of the mean.

| Men's height | Woman's height |
|---|---|
| $\mu = 69.9in$<br>$\sigma = 3.0in$ | $\mu = 64.3in$<br>$\sigma = 2.6in$ |

We can also use z-scores to compare different data sets. For example, men's and women's heights for a population of men and a population of women. Compare and interpret the data for a 6 foot tall man and a 6 foot tall woman.

z-score for a 6 foot tall man

$$z = \frac{x-\mu}{\sigma} = \frac{72-69.9}{3} = 0.7$$

z-score for a 6 foot tall woman

$$z = \frac{x-\mu}{\sigma} = \frac{72-64.3}{2.6} \approx 3.0$$

A 6 foot tall man falls within 1 standard deviation of the mean height of a man. So, 6 foot is normal for the height of a man. A 6 foot tall woman falls within 3 standard deviations of the mean height of a woman. So, this is an unusual height for a woman.

**P. 109 7 - 50 every 3rd item**