3.4

\[ Z = \frac{X - \mu}{\sigma} \quad \text{or} \quad Z = \frac{X - \bar{X}}{s} \]

Allows for comparison since Z-score is unitless with mean 0 and standard deviation 1.

Suppose babies born after a gestation period of 32 to 35 weeks have a mean weight of 2800 grams and a standard deviation of 600 grams while babies born after a gestation period of 40 weeks have a mean weight of 3200 grams and a standard deviation of 440 grams. If a 32-week gestation period baby weighs 2400 grams and a 40-week gestation period baby weighs 2800 grams, which baby weighs less relative to the gestation period?

Which baby weighs relatively less?

\[ Z_{32} = \frac{X - \mu}{\sigma} = \frac{2400 - 2800}{600} = -0.67 \quad \text{0.67 standard deviations below the mean.} \]

\[ Z_{40} = \frac{X - \bar{X}}{s} = \frac{2800 - 3200}{440} = -0.91 \quad \text{0.91 standard deviations below the mean.} \]
The *kth percentile*, denoted, $P_k$, of a set of data is a value such that $k$ percent of the observations are less than or equal to the value.

$P_{20} = 20^{th}$ percentile

The value is higher than 20\% of all the values.

<table>
<thead>
<tr>
<th>Age</th>
<th>10th</th>
<th>25th</th>
<th>50th</th>
<th>75th</th>
<th>90th</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 − 29</td>
<td>166.8</td>
<td>171.5</td>
<td>176.7</td>
<td>181.4</td>
<td>186.8</td>
</tr>
<tr>
<td>30 − 39</td>
<td>166.9</td>
<td>171.3</td>
<td>176.0</td>
<td>181.9</td>
<td>186.2</td>
</tr>
<tr>
<td>40 − 49</td>
<td>167.9</td>
<td>172.1</td>
<td>176.9</td>
<td>182.1</td>
<td>186.0</td>
</tr>
<tr>
<td>50 − 59</td>
<td>166.0</td>
<td>170.8</td>
<td>176.0</td>
<td>181.2</td>
<td>185.4</td>
</tr>
<tr>
<td>60 − 69</td>
<td>165.3</td>
<td>170.1</td>
<td>175.1</td>
<td>179.5</td>
<td>183.7</td>
</tr>
<tr>
<td>70 − 79</td>
<td>163.2</td>
<td>167.5</td>
<td>172.9</td>
<td>178.1</td>
<td>181.7</td>
</tr>
<tr>
<td>80 or older</td>
<td>161.7</td>
<td>166.1</td>
<td>170.5</td>
<td>175.3</td>
<td>179.4</td>
</tr>
</tbody>
</table>

**Heights of Males (cm)**

172.1 cm is at 25\% percentile

You’re taller than 25\% of all 40-49 year old males, and shorter than 75\%.

**Quartiles** - Divide data into quarters

$P_{25} = Q_1$ (25\% percentile is the first quartile)

$P_{50} = Q_2 = M$ (50\% percentile = 2nd quartile = median)

$P_{75} = Q_3$ (75\% percentile = 3rd quartile)
Finding Quartiles

Step 1: Arrange the data in ascending order.
Step 2: Determine the median, \( M \), or second quartile, \( Q_2 \).
Step 3: Determine the first and third quartiles, \( Q_1 \) and \( Q_3 \), by dividing the data set into two halves; the bottom half will be the observations below (to the left of) the location of the median and the top half will be the observations above (to the right of) the location of the median. The first quartile is the median of the bottom half and the third quartile is the median of the top half.

Watch video to get quartiles by hand.


\[
\text{Interquartile Range} = Q_3 - Q_1
\]

Wealth Data:
\[
\text{IQR} = 10.5 - 4 = 6.5 \text{ billion}
\]

This is a measure of dispersion that represents the range of the middle 50% of observations. It is resistant.
3.4 Outliers

Outliers are extreme observations.

Checking for Outliers by Using Quartiles

**Step 1:** Determine the first and third quartiles of the data.
**Step 2:** Compute the interquartile range.
**Step 3:** Determine the fences. Fences serve as cutoff points for determining outliers.

Lower fence = $Q_1 - 1.5(IQR)$
Upper fence = $Q_3 + 1.5(IQR)$

**Step 4:** If a data value is less than the lower fence or greater than the upper fence, it is considered an outlier.

Using Wealth Data

Lower Fence = $Q_1 - 1.5(IQR)$

$= 4 - 1.5(6.5)$

$= -5.75$

Upper Fence = $Q_3 + 1.5(IQR)$

$= 10.5 + 1.5(6.5)$

$= 20.25$

Forbes Wealthiest 100

<table>
<thead>
<tr>
<th>Row</th>
<th>Name</th>
<th>Net worth (B$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bill Gates</td>
<td>54</td>
</tr>
<tr>
<td>2</td>
<td>Warren Buffett</td>
<td>45</td>
</tr>
<tr>
<td>3</td>
<td>Larry Ellison</td>
<td>27</td>
</tr>
<tr>
<td>4</td>
<td>Christy Walto</td>
<td>24</td>
</tr>
<tr>
<td>5</td>
<td>Charles Koch</td>
<td>21.5</td>
</tr>
<tr>
<td>6</td>
<td>David Koch</td>
<td>21.5</td>
</tr>
<tr>
<td>7</td>
<td>Jim Walton</td>
<td>20.1</td>
</tr>
<tr>
<td>8</td>
<td>Alice Walton</td>
<td>20</td>
</tr>
<tr>
<td>9</td>
<td>S. Robson W</td>
<td>19.7</td>
</tr>
<tr>
<td>10</td>
<td>Michael Bloor</td>
<td>18</td>
</tr>
<tr>
<td>11</td>
<td>Larry Page</td>
<td>15</td>
</tr>
<tr>
<td>12</td>
<td>Sergey Brin</td>
<td>15</td>
</tr>
<tr>
<td>13</td>
<td>Sheldon Adels</td>
<td>14.7</td>
</tr>
</tbody>
</table>