⚠ Lagunita is retiring and will shut down at 12 noon Pacific Time on March 31, 2020. A few courses may be open for self-enrollment for a limited time. We will continue to offer courses on other online learning platforms; visit http://online.stanford.edu.

Course > EDA: Examining Distributions > One Quantitative Variable: Measures of Spread - Range, IQR, & Outliers > Calculating the IQR

🔖 Bookmark this page

Calculating the IQR

Learning Objective: Relate measures of center and spread to the shape of the distribution, and choose the appropriate measures in different contexts.

Example: Best Actress Oscar Winners

To find the IQR of the Best Actress Oscar winners distribution, it will be convenient to use the stemplot.

```
2|12
2|56678999                    Bottom Half
3|012233333444
3|5566789
4|111224
4|5599
5|
5|
6|112                         Top Half
6|
7|4
7|
8|0
```

Q1 is the median of the bottom half of the data. Since there are 22 observations in that half, Q1 is the mean of the 11th and 12th ranked observations in that half:

$$Q1 = \frac{(30 + 31)}{2} = 30.5$$

Similarly, Q3 is the median of the top half of the data, and since there are 22 observations in that half, Q3 is the mean of the 11th and 12th ranked observations in that half:

$$Q3 \; = \; \frac{(42 + 42)}{2} \; = \; 42$$

$$IQR \; = \; (42 - 30.5) \; = \; 11.5$$

Note that in this example, the range covered by all the ages is 59 years, while the range covered by the middle 50% of the ages is only 11.5 years. While the whole dataset is spread over a range of 59 years, the middle 50% of the data is packed into only 11.5 years. Looking again at the histogram will illustrate this:
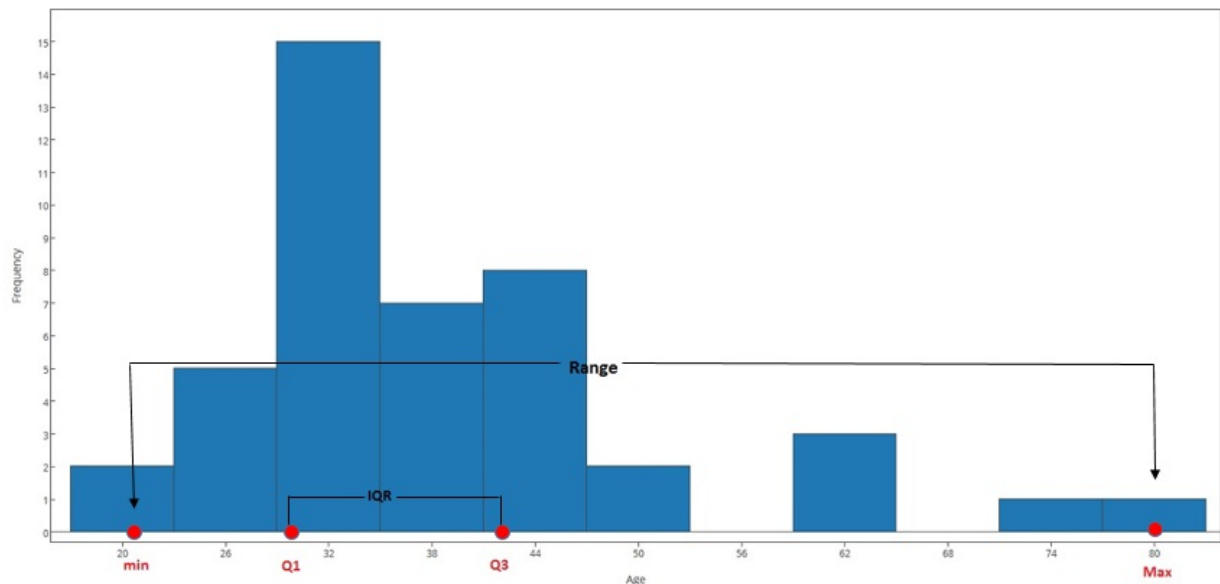
Comment

Software packages use different formulas to calculate the quartiles Q1 and Q3. This should not worry you, as long as you understand the idea behind these concepts. **Note** that Q1 and Q3 as reported by the various software packages differ from each other and are also slightly different from the ones we found here. There are different acceptable ways to find the median and the quartiles. These can give different results occasionally, especially for datasets where n (the number of observations) is fairly small. As long as you know what the numbers mean, and how to interpret them in context, it doesn't really matter much what method you use to find them, since the differences are really negligible.

Open Learning Initiative ⧉