Interquartile range and box plots

Study Development Factsheet

Interquartile range is similar to range, but is for data sets with extreme outliers. For example, in a survey about number of study hours that students do in a week, an answer of 0 hours might be an outlier, and a value of 100 hours would definitely be an outlier. Calculating range from these values would give an inaccurate picture of the spread of the data. In this case, it would be more appropriate to calculate the interquartile range.

## Method

There are 5 ‘quartile’ values we can calculate, and the method is very similar to calculating a median (in fact, one of the quartiles *is* the median).

1. Place all the values in order from smallest to largest.
2. The lowest value is called $Q\_{0}$ (quartile 0). This is often not really needed, but can be found if useful.
3. The highest value is called $Q\_{4}$ (4th quartile). The range can be calculated by finding $Q\_{4}-Q\_{0}.$
4. Find the median. For $n $of data points, find the value in position $\frac{n}{2}+0.5$ (for example, for 5 data points, find the 3rd data point. This is the median). If $n$ is even, we find the values in positions $\frac{n}{2}$ and $\frac{n}{2}+1$ and find the point between them by adding them together and dividing by 2 (for example, for 10 data points, add together the 5th and 6th values and divide that by 2).
The median is called $Q\_{2}$ (2nd quartile).
5. We are now left with two sets of numbers: those above the median, and those below. We find the median of each of these sets (make sure not to put $Q\_{2}$ in either set before finding the median).
6. The median of the numbers lower than $Q\_{2}$ is called $Q\_{1}$ (1st quartile) and the median of the numbers higher than $Q\_{2}$ is called $Q\_{3}$ (3rd quartile).
7. The interquartile range (IQR) is given by $Q\_{3}-Q\_{1}$.

Interquartile range and box plots

Study Development Factsheet

## Example

A bookshop records how many books each customer bought one day. They get the following results:

2, 1, 1, 3, 95, 1, 4, 5, 1, 0, 1, 0, 1, 0, 0, 0

Find the interquartile range and justify your reason for finding the IQR instead of the range.

## Answer

1. We put the values in order from smallest to largest:
0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 2, 3, 4, 5, 95
2. $Q\_{0}$ = 0
3. $Q\_{4}=95$
4. There are 16 values, so we find the values in position 8 and 9 and average them:
8th = 1, 9th = 1, average = 1.

$Q\_{2}$ = 1

1. We now have two sets of numbers: those lower than $Q\_{2}$: 0, 0, 0, 0, 0, 1, 1, 1, and those higher than $Q\_{2}:$ 1, 1, 1, 2, 3, 4, 5, 95.
2. The median of those lower than $Q\_{2}$ is found by averaging those in the 4th and 5th positions: 4th = 0, 5th = 0, average = 0.
$Q\_{1}$ = 0
3. The median of those higher than $Q\_{2}$ is found by averaging those in the 12th and 13th positions: 12th = 2, 13th = 3, average = 2.5.
$Q\_{3}$ = 2.5
4. The IQR is $Q\_{3}-Q\_{1}$ = 2.5 – 0 = 2.5

We find the IQR rather than the range since 95 books is obviously an outlier for the shop (perhaps a large order was placed for an event- but it seems strange next to all the other values).

Interquartile range and box plots

Study Development Factsheet

## Drawing a box plot

We can use the values for $Q\_{0},Q\_{1},Q\_{2},Q\_{3}$and $Q\_{4}$ to draw a box plot (also called a box and whisker plot). These are useful for quickly seeing if data is skewed.

A box plot is drawn as follows:

 

The lines that show $Q\_{0}$ to $Q\_{1}$ and $Q\_{3}$ to $Q\_{4}$ are called the ‘whiskers’ of the box and whisker plot, and the box between $Q\_{1}$ and $Q\_{3}$ is the ‘box’.

We can tell the skew by looking at $Q\_{2}$, if it is closer to $Q\_{1}$ than $Q\_{3}$, this is called ‘left skew’ or ‘positive skew’. If $Q\_{2}$ is closer to $Q\_{3}$ than $Q\_{1}$, this is called ‘right skew’ or ‘negative skew’.

Interquartile range and box plots

Study Development Factsheet

Sometimes these can look quite strange. For example, the box plot for the earlier bookshop example looks like this:



With $Q\_{4}$ so far away that it isn’t actually shown on this diagram. This is still a box and whisker plot, and shows very strong negative skew.

**Support**: Study Development offers workshops, short courses, 1 to 1 and small group tutorials.

* Join a tutorial or workshop on the [Study Development tutorial and workshop webpage](https://www.yorksj.ac.uk/students/study-skills/study-development-tutorials/) or search ‘YSJ study development tutorials.’
* Access our Study Success resources on the [Study Success webpage](https://www.yorksj.ac.uk/students/study-skills/study-success/) or search ‘YSJ study success.’