

Cross Validated is a question and answer site for people interested in statistics, machine learning, data analysis, data mining, and data visualization. It's 100% free, no registration required.

Here's how it works:

Sign up

Anybody can ask a question

Anybody can answer

The best answers are voted up and rise to the top

## Find interquartile range of these data

Take these data:

5, 8, 9, 14

R says the interquartile range is 3 :

```
IQR(c(5, 8, 9, 14))
# 3
```

...but I make it 5 . What am I doing wrong? Here are the steps I've taken:

1. Find median, which is 8.5
2. Split the data around the median into two groups, like this: (5, 8), (9, 14)
3. Find the median of these two groups, which are 6.5 and 11.5 , respectively
4. Subtract 6.5 from 11.5 , which yields 5

r descriptive-statistics median interquartile

edited Aug 11 '15 at 19:26



Placidia

7,127 13 38

asked Sep 6 '14 at 19:07



luciano

2,641 3 31 57

- 3 You have computed something called the "H-spread" (Tukey, *Exploratory Data Analysis*, p. 44). – [whuber](#) ♦ Sep 6 '14 at 19:38
- 3 The hinge spread (or H-spread) is not identical to the most usual definitions of the IQR in small samples. If you want R to compute the hinge spread specifically (say for a box-plot), you can use `diff(fivenum(x)[c(2,4)])` . On your data ( `diff(fivenum(c(5,8,9,14))[c(2,4)])` ) that gives 5. – [Glen\\_b](#) ♦ Sep 6 '14 at 23:55
- 1 The R `IQR` function calls the `quantile` function, which in turn has 9 different algorithms (with one of them used as default if you don't specify your own choice). Results from these algorithms can differ with small data sets such as yours. The R Help pages for `IQR` and `quantile` provide the details. – [EdM](#) Sep 8 '14 at 15:09

## 1 Answer

Your difficulty is that in order to find IQR you must first find the two quartiles. And there are many different formulas for quartiles (including quartiles) in common use.

In particular, major statistical software packages disagree on which methods to implement as their default: (a) SAS, (b) Minitab and SPSS, and (c) R (and its parent S) use three different methods. Furthermore, these methods differ from methods found in reputable elementary texts. (Adding to the confusion: Tukey's 'fourths', sometimes used in making boxplots and often considered essentially the same as quartiles, use yet other criteria.)

Generally speaking the differences among these methods become negligible for large sample sizes. However, there can be marked differences for small samples. Fortunately, it is for large samples that quartiles make the most sense. (Roughly, quartiles are intended to divide a sample into four 'chunks' of equal size: how do you do that with a sample of size 10?)

In R, you can type `?quantile` to see the nine different types of quantiles supported by R (using an extra argument), mentioned just now by [@EdM](#). The default result from `quantile` is min, Q1, med, Q3, max, so once you have selected a `type` you could define your own IQR function based on the idea in the [@Glen\\_b](#) Comment and code like `as.numeric(diff(quantile(x, type=5)[c(2,4)]))` .

edited Aug 11 '15 at 18:37

answered Aug 11 '15 at 18:14



BruceET

151 5

