

---

Subject: Re: Selecting sample within one standard deviation in R

Posted by [Bridgette-DHS](#) on Tue, 15 Feb 2022 21:52:45 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

You have an extremely skewed distribution. The "68%" rule works for normally distributed variables, and the normal approximation doesn't work for your variable. I can think of two options. One would be to take the log of the frequency, which will have a distribution that is more nearly normal, but there's the problem that you can't take the log of 0. Another option would be to calculate the percentiles of the distribution. If you identify the 25th and 75th percentiles, then you have the boundaries for the middle 50%. Or identify the 16th and 84th percentiles, which enclose the middle 68%.

---