
Subject: Re: How to ignore NaNs in the data array with function cgPercentile.pro or Percentile.pro?

Posted by [Helder Marchetto](#) on Mon, 18 May 2015 13:36:06 GMT

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

On Monday, May 18, 2015 at 3:18:53 PM UTC+2, Madhavan Bomidi wrote:

> Hello,

>
> I have two data arrays one with some missing data and the other without missing data as below:

>
> data = [Randomu(3L, 100) * 100, !Values.F_NAN, !Values.F_NaN]

>
> data1 = [Randomu(3L, 100) * 100]

>
> When I use the cgPercentiles, I don't want missing data to be included.

To avoid including NaN you could use the finite() function:

```
data = data[where(finite(data))]  
and then calculate percentiles.d
```

Here is an example

```
data = Randomu(3L, 100) * 100  
data1 = [data, !Values.F_NAN, !Values.F_NaN]  
data2 = data1[where(finite(data1))]  
Print, cgPercentiles(data, Percentiles=[0.25, 0.5, 0.75])  
Print, cgPercentiles(data1, Percentiles=[0.25, 0.5, 0.75])  
Print, cgPercentiles(data2, Percentiles=[0.25, 0.5, 0.75])
```

In my case, I got:

```
27.4920  45.3172  69.3138  
27.4920  45.4608  69.4824  
27.4920  45.3172  69.3138
```

First and last line are the same and that's what you want.

As David said, if you don't have a strategy to substitute NaNs with numbers, you can't deal with them.

Cheers,
Helder

ps: the above code will not work (or give you trouble) if your data is made out of *only* NaNs. When using the where() function you should check for how many finite numbers where found...
