
Subject: Re: Function names (was mean and sdev)
Posted by [J.D. Smith](#) on Tue, 03 Mar 1998 08:00:00 GMT
[View Forum Message](#) <> [Reply to Message](#)

Brian Jackel wrote:

>
> Dr. G. Scott Lett wrote:
>>
>> The MOMENT function is obscure and slow when all you need is
>> the mean and standard deviation, so we're adding the more obvious
>> MEAN, STDDEV, VARIANCE and other statistics functions for IDL 5.1.
>>
>> Best wishes,
>> Scott Lett
>
> I'm not sure who the "we" above refers to, but I'd like to
> make a general comment. IMHO IDL is already dangerously
> cluttered with lots of special purpose functions that could
> perhaps be more usefully bundled together. In the case of
> statistics it would be better to have a function like MOMENTS
> that could return any or all required moments. Even better
> would be something (perhaps called STATISTICS) that could
> provide a number of useful quantities. Although the use
> of square brackets should eliminate the confusion between
> functions and arrays, the proliferation of built-in functions
> (especially with common names) may cause problems if the
> name has been used for existing code.
>
> Does anyone else have an opinion on this subject?
>
> --
> Brian Jackel

I agree wholeheartedly. It seems to me that making multipurpose routines fast was the reason (among others) for introducing the function `arg_present`. One can use it to make a version of `MOMENT` for instance, that only calculates variance, skew and kurtosis if the user wants them. I think in general a balance has to be struck between ease of use and over-cluttering. In this case, since `mean` and `stdev` are likely the most common calculations performed, they may merit their own functions, but I do agree that restraint is in order (or a more robust shadowing mechanism... but I won't hold my breath).

JD

--
J.D. Smith |*| WORK: (607) 255-5842

Subject: Re: Function names (was mean and sdev)
Posted by [Kevin Spencer](#) on Wed, 04 Mar 1998 08:00:00 GMT
[View Forum Message](#) <> [Reply to Message](#)

Brian Jackel <jackel@danlon.physics.uwo.ca> wrote in article
<6dhu2f\$eoi@ds2.acs.ucalgary.ca>...

> Dr. G. Scott Lett wrote:

>>

>> The MOMENT function is obscure and slow when all you need is
>> the mean and standard deviation, so we're adding the more obvious
>> MEAN, STDDEV, VARIANCE and other statistics functions for IDL 5.1.

>>

>> Best wishes,

>> Scott Lett

>

> I'm not sure who the "we" above refers to, but I'd like to
> make a general comment. IMHO IDL is already dangerously
> cluttered with lots of special purpose functions that could
> perhaps be more usefully bundled together. In the case of
> statistics it would be better to have a function like MOMENTS
> that could return any or all required moments. Even better
> would be something (perhaps called STATISTICS) that could
> provide a number of useful quantities. Although the use
> of square brackets should eliminate the confusion between
> functions and arrays, the proliferation of built-in functions
> (especially with common names) may cause problems if the
> name has been used for existing code.

>

> Does anyone else have an opinion on this subject?

I've been very annoyed at RSI's inconsistent support for these simple statistical functions. There used to be a "stdev" procedure, then (if I recall correctly) one like it with a similar name. In IDL 5, they got rid of these functions and replaced them with "moment". Now according to Scott Lett, who I assume works for RSI, they are going to have another standard deviation procedure, plus one for computing the mean and another for the variance. Why did they get rid of stdev in the first place? And calculating the mean is trivial; just use total(x)/n_elements(x). A separate function for computing variance is likewise trivial and unnecessary, if there's a function for computing the standard deviation.

Kevin

Kevin Spencer
Cognitive Psychophysiology Laboratory
Dept. of Psychology and Beckman Institute
University of Illinois at Urbana-Champaign
kspence1@uiuc.edu

Subject: Re: Function names (was mean and sdev)
Posted by [Martin Schultz](#) on Wed, 04 Mar 1998 08:00:00 GMT
[View Forum Message](#) <> [Reply to Message](#)

Brian Jackel wrote:

>
> Dr. G. Scott Lett wrote:
>>
>> The MOMENT function is obscure and slow when all you need is
>> the mean and standard deviation, so we're adding the more obvious
>> MEAN, STDDEV, VARIANCE and other statistics functions for IDL 5.1.
>>
>> Best wishes,
>> Scott Lett
>
> I'm not sure who the "we" above refers to, but I'd like to
> make a general comment. IMHO IDL is already dangerously
> cluttered with lots of special purpose functions that could
> perhaps be more usefully bundled together. In the case of
> statistics it would be better to have a function like MOMENTS
> that could return any or all required moments. Even better
> would be something (perhaps called STATISTICS) that could
> provide a number of useful quantities. Although the use
> of square brackets should eliminate the confusion between
> functions and arrays, the proliferation of built-in functions
> (especially with common names) may cause problems if the
> name has been used for existing code.
>
> Does anyone else have an opinion on this subject?
>
> --
> Brian Jackel

Oh yes, I do! In my opinion it gives much more readable code if you can call a function named MEAN or STDDEV to get these values. On the other hand, I see your point that there are far too many functions in IDL that nobody can remember. But bundling always raises issues of maintenance and completeness. My suggestion would be to think of a number of "wrapper" routines (like your suggestion for statistics)

which can then take keyword parameters like /MEAN, /STDDEV, etc. and which would call the appropriate internal IDL functions. But those would not necessarily have to be developed by the IDL people ;-)
What I find much more bothering are inconsistencies within "families" of functions, e.g. different ordering of parameters or naming of keywords in say a linear fit and a polyfit function (of course, I cannot find a suitable example right now, but I come across these things once a while).

Just one additional word: the capability of finding the appropriate IDL function depends strongly on the quality of the links in the online help. After I had complained about a missing link to IDL once, they replied and said they would be grateful for specific suggestions to improve these links. At least there is some hierarchical sorting in the Navigate-Search Index menu - this could probably be improved (e.g. "curve fitting" -- but why not "fitting" ?; how about an option to search only the "topic" indices ?)

regards,
Martin.

Dr. Martin Schultz
Department for Earth&Planetary Sciences, Harvard University
186 Pierce Hall, 29 Oxford St., Cambridge, MA-02138, USA

phone: (617)-496-8318
fax : (617)-495-4551

e-mail: mgs@io.harvard.edu
IDL-homepage: <http://www-as.harvard.edu/people/staff/mgs/idl/>

Subject: Re: Function names (was mean and sdev)
Posted by [wmc](#) on Thu, 05 Mar 1998 08:00:00 GMT
[View Forum Message](#) <> [Reply to Message](#)

In article 2748ae80@p4.psych.uiuc.edu, "Kevin Spencer" <kspence1@uiuc.edu> writes:
> Brian Jackel <jackel@danlon.physics.uwo.ca> wrote in article
> <6dhu2f\$eoi@ds2.acs.ualgary.ca>...
>> Dr. G. Scott Lett wrote:
>>> The MOMENT function is obscure and slow when all you need is
>>> the mean and standard deviation, so we're adding the more obvious
>>> MEAN, STDDEV, VARIANCE and other statistics functions for IDL 5.1.
>> I'm not sure who the "we" above refers to, but I'd like to
>> make a general comment. IMHO IDL is already dangerously

>> cluttered with lots of special purpose functions that could
>> perhaps be more usefully bundled together.
> I've been very annoyed at RSI's inconsistent support for these
> simple statistical functions. There used to be a "stdev" procedure,
> then (if I recall correctly) one like it with a similar name. In IDL 5,
> they got rid of these functions and replaced them with "moment".
> Now according to Scott Lett, who I assume works for RSI, they
> are going to have another standard deviation procedure, plus one
> for computing the mean and another for the variance. Why did
> they get rid of stdev in the first place? And calculating the mean
> is trivial; just use `total(x)/n_elements(x)`.

If they do introduce a function called `stddev`, instead of recalling `stdev` from "obsolete", they will be very silly indeed (its still useable in IDL 5 BTW, just not in the help).

There is disagreement between those who like things bundled together (the "momentists") who are presumably happy to write `av=total(x)/n_elements(x)` in order to avoid having a "mean" or "avg" function defined, and those who don't really mind having piles of extra functions lying around in order to write shorter and more comprehensible code (I'm in the latter camp).

To some extent this could be fixed by putting all the extra functions in a separate directory whose name you could remove from the `!PATH` if you wanted to. Presumably, if you don't want to see the obsolete functions you can remove "obsolete" from the `!PATH`.

- William

William M Connolley | wmc@bas.ac.uk | <http://www.nbs.ac.uk/public/icd/wmc/>
Climate Modeller, British Antarctic Survey | Disclaimer: I speak for myself
