Subject: Incorrect behavior of /NAN
Posted by K. Bowman on Thu, 04 Apr 2002 15:40:11 GMT
View Forum Message <> Reply to Message

I tried to incite some interest in this last week, but had no takers,
so I'll try again. ;-)

I believe that the behavior of the TOTAL function is incorrect in
the case where all of the data are NaN and the /NAN keyword is set.

Here are several possible cases:

```
IDL> a = !values.f_nan
IDL> print, total([a, 1.0])
         NaN
IDL> print, total([a, 1.0], /nan)
     1.00000
IDL> print, total([a, a])
         NaN
IDL> print, total([a, a], /nan)
     0.00000
```

I believe that the last case is incorrect.

The documentation for /NAN says "Elements with the value NaN are treated
as missing data." In the last case there are no valid data, so how can
their sum be zero?

I think this is probably a simple error in the TOTAL algorithm.
Rather than computing the "sum of all the non-NaN values", it is
probably doing "set sum to zero and then add all non-Nan values".

For comparison, the MEAN function does not behave this way.

```
IDL> print, mean([a, 1.0])
         NaN
IDL> print, mean([a, 1.0], /nan)
     1.00000
IDL> print, mean([a, a])
         NaN
IDL> print, mean([a, a], /nan)
         NaN
```

Before I submit something (at least a question) to RSI, does anyone
have comments?

Ken Bowman

Subject: Re: Incorrect behavior of /NAN
Posted by Paul van Delst on Thu, 04 Apr 2002 21:52:02 GMT
View Forum Message <> Reply to Message

Mark Hadfield wrote:
>
> "Paul van Delst" <paul.vandelst@noaa.gov> wrote in message
> news:3CAC9914.B52D1669@noaa.gov...
>> Seems reasonable to me. It only sums the data not flagged as missing. If
>> all the data is missing, the sum of nothing is, well, nothing.
>
> And IDL normally uses NaN ("Not a Number") to represent nothing.

Huh. I though that's what zero was for. Or is a throwback from Roman times:

... -V -IV -III -II -I NaN I II III IV V .....

paulv

--
Paul van Delst Religious and cultural
CIMSS @ NOAA/NCEP purity is a fundamentalist
Ph: (301)763-8000 x7274 fantasy
Fax:(301)763-8545 V.S.Naipaul

Subject: Re: Incorrect behavior of /NAN
Posted by Paul van Delst on Fri, 05 Apr 2002 15:35:22 GMT
View Forum Message <> Reply to Message

Craig Markwardt wrote:
>
> Paul van Delst <paul.vandelst@noaa.gov> writes:
>>
>> Maybe a holistic division operator is required in these sorts of cases. :o)
>
> Or should IDL apply L'Hopital's rule? :-)

That would be good - I'd never have to fret over calculating sinc functions again.....

paulv

--
Paul van Delst Religious and cultural

CIMSS @ NOAA/NCEP
Ph: (301)763-8000 x7274
Fax:(301)763-8545

purity is a fundamentalist
fantasy
V.S.Naipaul

Subject: Re: Incorrect behavior of /NAN
Posted by Craig Markwardt on Fri, 05 Apr 2002 21:53:00 GMT
View Forum Message <> Reply to Message

Kenneth Bowman <k-bowman@null.tamu.edu> writes:

> In article <onofgz74jo.fsf@cow.physics.wisc.edu>,
> Craig Markwardt <craigmnet@cow.physics.wisc.edu> wrote:
>
>> If all the data are missing, how can the sum be NaN? Neither a return
>> value of 0, nor a return value of NaN, seems to be appropriate to
>> me. This appears to be an undefined case.
>
> If NaN represents missing data, and all of the data are missing, then the
> total is missing also, so it seems reasonable to me to return NaN.

:-)

I could easily argue the other way, in terms of reasonableness. The
sum of "no" numbers is zero.

Craig

--

Craig B. Markwardt, Ph.D. EMAIL: craigmnet@cow.physics.wisc.edu
Astrophysics, IDL, Finance, Derivatives | Remove "net" for better response
---

Subject: Re: Incorrect behavior of /NAN
Posted by K. Bowman on Fri, 05 Apr 2002 22:25:04 GMT
View Forum Message <> Reply to Message

In article <onu1qpzubn.fsf@cow.physics.wisc.edu>,
Craig Markwardt <craigmnet@cow.physics.wisc.edu> wrote:

> I could easily argue the other way, in terms of reasonableness. The
> sum of "no" numbers is zero.

Zero is a number. How can the sum of no numbers be a number?
If there are no numbers how can there even be a sum?

Ken

Subject: Re: Incorrect behavior of /NAN
Posted by James Kuyper on Fri, 05 Apr 2002 22:53:40 GMT
View Forum Message <> Reply to Message

Kenneth Bowman wrote:

> In article <onu1qpzubn.fsf@cow.physics.wisc.edu>,
> Craig Markwardt <craigmnet@cow.physics.wisc.edu> wrote:
>
>
>> I could easily argue the other way, in terms of reasonableness. The
>> sum of "no" numbers is zero.
>
>
> Zero is a number. How can the sum of no numbers be a number?
> If there are no numbers how can there even be a sum?
>
> Ken

The number of ball bearings produced in a given city during a specified time period can be expressed as a sum over the amounts produced by each ball bearing plant in that city. What should the value of that sum be, for a city that has no ball bearing plants? I'd say that it's definitely a number, and definitely 0. I think that this is the most reasonable value for the sum of almost any variable-length list, when that list happens to have a length of 0. It's also the natural result of the most obvious algorithm for calculating the sum.

Subject: Re: Incorrect behavior of /NAN
Posted by David Fanning on Sat, 06 Apr 2002 01:57:51 GMT
View Forum Message <> Reply to Message

James Kuyper (kuyper@gscmail.gsfc.nasa.gov) writes:

> The number of ball bearings produced in a given city during a specified
> time period can be expressed as a sum over the amounts produced by each
> ball bearing plant in that city. What should the value of that sum be,
> for a city that has no ball bearing plants? I'd say that it's definitely
> a number, and definitely 0. I think that this is the most reasonable
> value for the sum of almost any variable-length list, when that list
> happens to have a length of 0. It's also the natural result of the most
> obvious algorithm for calculating the sum.

I agree. I was thinking about this while I was
feeding balls to my tennis team. (Don't ask,
please.) A sum is a counting operation.

"How many marbles are in the box?"

"None. The box is full of cupcakes."

Then the sum of marbles in the box is 0.

Seems pretty straightforward to me. But then
I've never understood the whole Christian/Muslim
thing either. :-(

Cheers,

David

--
David W. Fanning, Ph.D.
Fanning Software Consulting
Phone: 970-221-0438, E-mail: david@dfanning.com
Coyote's Guide to IDL Programming: http://www.dfanning.com/
Toll-Free IDL Book Orders: 1-888-461-0155

Subject: Re: Incorrect behavior of /NAN
Posted by Kenneth P. Bowman on Sat, 06 Apr 2002 14:28:01 GMT
View Forum Message <> Reply to Message

In article <MPG.171804089a266d298986e@news.frii.com>,
David Fanning <david@dfanning.com> wrote:

> "How many marbles are in the box?"
>
> "None. The box is full of cupcakes."
>
> Then the sum of marbles in the box is 0.

You English language statement is almost correct, but it
should be "the number of marbles in the box is 0".

Translating to IDL

i = WHERE(thingInBox EQ marble, numberOfMarbles)

not

numberOfMarbles = TOTAL(thingInBox, /NaMarble)

;-)

Ken

Subject: Re: Incorrect behavior of /NAN
Posted by Craig Markwardt on Sat, 06 Apr 2002 15:46:10 GMT
View Forum Message <> Reply to Message

Kenneth Bowman <k-bowman@null.tamu.edu> writes:

> In article <onu1qpzubn.fsf@cow.physics.wisc.edu>,
> Craig Markwardt <craigmnet@cow.physics.wisc.edu> wrote:
>
>> I could easily argue the other way, in terms of reasonableness. The
>> sum of "no" numbers is zero.
>
> Zero is a number. How can the sum of no numbers be a number?
> If there are no numbers how can there even be a sum?

I think either value is a "reasonable" return from TOTAL, zero or NaN,
in the case you describe. The real issue is that it's not documented.

Now where did I put that box of cupcakes...

Craig

--
---------- ----------
Craig B. Markwardt, Ph.D. EMAIL: craigmnet@cow.physics.wisc.edu
Astrophysics, IDL, Finance, Derivatives | Remove "net" for better response
---------- ----------

Subject: Re: Incorrect behavior of /NAN
Posted by noymer on Mon, 15 Apr 2002 01:32:10 GMT
View Forum Message <> Reply to Message

NaN is an IEEE standard, is it not?

Does the IEEE define what to do in such cases?

It seems to me that IDL should conform to such a standard, if it exists,
even if it goes against our intuition.

That said, doing a google on IEEE 754 and NaN and sum didn't turn up a definitive answer (but I didn't look at all of the over 1,000 pages google found).

Just my $0.02

Andrew

Subject: Re: Incorrect behavior of /NAN
Posted by K. Bowman on Mon, 15 Apr 2002 13:47:45 GMT
View Forum Message <> Reply to Message

In article <yx6hvgatn3vp.fsf@socrates.Berkeley.EDU>,
 noymer@socrates.Berkeley.EDU wrote:

> NaN is an IEEE standard, is it not?
>
> Does the IEEE define what to do in such cases?
>
> It seems to me that IDL should conform to such a standard, if it exists,
> even if it goes against our intuition.

As far as I know, IDL correctly does arithmetic involving NaNs (i.e., it returns NaNs, etc.).

But IDL also uses NaNs as "missing values", which is not an IEEE issue, and the behavior here is not clearly defined in some cases.

Ken
