
Subject: Re: Intersection of 2 sets--Beginner IDL question
Posted by [Craig Markwardt](#) on Wed, 17 Oct 2001 15:05:25 GMT
[View Forum Message](#) <> [Reply to Message](#)

John-David Smith <jdsmith@astro.cornell.edu> writes:

> David Fanning wrote:
>>
>> Ted Cary (tedcary@yahoo.com) writes:
>>
>>> Is there any programming technique for finding the intersection of two sets
>>> (arrays) of numbers without using WHERE in a loop to search the larger array
>>> for every element in the smaller array? It seems like a very clumsy way to
>>> find values shared by both arrays, especially with integer sets/arrays.
>>
>> How about the very tiny program, SetIntersection,
>> which uses--what else--a Histogram! :-)
>>
>> http://www.dfanning.com/tips/set_operations.html
>
>
> It's amazing how much recycled information flows through the newsgroup, if you
> watch it long enough. I remember just like it was 4 years ago the detailed
> discussions with which we whiled away our days, concerning value-based
> intersection vs index-based intersection, order N vs. unknown order operations,
> etc.
>
...
> In a classic example posed by Mark Fardal, you are matching up social security
> numbers in two lists containing age and income. The set_intersection style
> solution fails miserably here, and to a lesser degree for any arrays which are
> somewhat sparse (where *somewhat* seems to be about 1 in 10, depending on lots
> of factors).

Hi JD--

Thanks for beating me to the punch. The HISTOGRAM method is indeed very cool for a new learner, but it definitely starts to suck air (and memory) when the data sets become sparse.

Long ago (1 year?) I tried to collect all the various algorithms that were being discussed, and some that weren't yet, to do set operations. CMSET_OP has the dreaded "CM" prefix, but it also knows how to do intersections, unions, and exclusive or's. It can do X AND NOT Y type intersections as well, in one self contained function.

The syntax is:

```
x_and_y = cmset_op(X, 'AND', y)
```

It can return by value or index.

Craig

--

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