Subject: Re: I would like to average the first n columns based on duplicate values of the n+1th column
Posted by Helder Marchetto on Tue, 04 Oct 2016 12:34:00 GMT
View Forum Message <> Reply to Message

On Monday, October 3, 2016 at 11:05:31 PM UTC+2, belk...@web.de wrote:
> Hey,
> Can Someone help me solve this problem in IDL:
> "I have a matrix with duplicate numbers in one of the columns. I would like to average the rows with duplicate numbers. For example, I have duplicate values in a matrix A in column 3:
>
> A =
> 1 2 1
> 4 4 2
> 5 4 2
> 4 5 2
> 5 5 3
> 10 3 3
>
>
> B =
> 1 2 1
> 4.3333 4.3333 2.0000
> 7.5000 4.0000 3.0000
>
> where each row is the average values of the duplicate rows of column 3.
>
> Can anyone help?"
>
> found here:
> http://stackoverflow.com/questions/15270019/i-would-like-to-average-the-first-n-columns-based-on-duplicate-values-of-the -n1
>
> Cheers,
> B.R.

Ok, this might not be instructive. But it was fun to look into.
I basically shortened the whole thing into two instructions:

```
u = [uniq(a[2,*],sort(a[2,*])),n_elements(a[2,*])-1]
for i=0,n_elements(u)-2 do print,  [total(reform(a[0:1,lindgen(u[i+1]-u[i]+1,start=u[i])],2,u[i+1]-u[i]+1),2)/float(u[i+1]-u[i]+1),a[2,u[i]]]
```

This works if a is defined as:

```
a = [[ 1, 2, 1],$
    [ 4, 4, 2],$
    [ 5, 4, 2],$
    [ 4, 5, 2],$
```

```
    [ 5, 5, 3],$
    [10, 3, 3]]
```

This is what I get:

```
     2.50000      3.00000      1.00000
     4.50000      4.50000      2.00000
     7.50000      4.00000      3.00000
```

Similar to Markus version, but it does not use the where().

Anyway, this was already solved, so it was a just for fun thing to do.

Cheers,
Helder
