Subject: Re: I would like to average the first n columns based on duplicate values of the n+1th column
Posted by belkaraza on Tue, 04 Oct 2016 11:32:29 GMT
View Forum Message <> Reply to Message

Am Dienstag, 4. Oktober 2016 13:23:59 UTC+2 schrieb belk...@web.de:
> Am Dienstag, 4. Oktober 2016 13:17:24 UTC+2 schrieb belk...@web.de:
>> Am Dienstag, 4. Oktober 2016 12:32:48 UTC+2 schrieb Markus Schmassmann:
>>> On 10/03/2016 11:05 PM, belkaraza@web.de wrote:
>>>> 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
>>>
>>> if isa(A,/integer) then begin
>>> h=histogram(A[2,*],reverse_indices=ri)
>>> idx=where(h ne 0,n)
>>> B=fltarr(3,n)
>>> for i=0,n-1 do begin
>>> if ri[idx[i]] eq ri[idx[i]+1]-1 then $
>>> B[0,i]=A[*,ri[ri[idx[i]]:ri[idx[i]+1]-1]] else $
>>> B[0,i]=mean(A[*,ri[ri[idx[i]]:ri[idx[i]+1]-1]],dim=2)
>>> endfor
>>> endif else
>>> values=A[2,uniq(A[2,*],sort(A[2,*]))]
>>> ; if A[2,*] is already sorted, A[2,uniq(A[2,*])] is sufficient there

```
>>>       n=n_elements(values)
>>>       B=fltarr(3,n)
>>>       for i=0,n-1 do begin
>>>          w=where(A[2,*] eq values[i],cnt)
>>>          if w cnt 1 then B[0,i]=A[*,where(A[2,*] eq values[i])] else $
>>>             B[0,i]=mean(A[*,where(A[2,*] eq values[i])],dim=2,/nan)
>>>       endfor
>>>  endelse
>>>
>>>
>>>  hope that does it, Markus
>>
>>
>>  Hey, thanks for the answer. The last if loop is bugged. if w cnt 1 then B[0,i]
>>  Can't see how to fix that
>  Ok fixed it with "if w[cnt] eq 1 then B[0,i]"
>  Again thanks alot for your help ;)

In case someone wants to use it as a function:
FUNCTION tsm,A,columntotal,column

if isa(A,/integer) then begin
  h=histogram(A[column,*],reverse_indices=ri)
  idx=where(h ne 0,n)
  B=fltarr(columntotal,n)
  for i=0,n-1 do begin
    if ri[idx[i]] eq ri[idx[i]+1]-1 then $
      B[0,i]=A[*,ri[ri[idx[i]]:ri[idx[i]+1]-1]] else $
      B[0,i]=mean(A[*,ri[ri[idx[i]]:ri[idx[i]+1]-1]],dim=2)
  endfor
endif else begin
values=A[column,uniq(A[column,*],sort(A[column,*]))]
; if A[2,*] is already sorted,   A[2,uniq(A[2,*])]   is sufficient there
n=n_elements(values)
B=fltarr(columntotal,n)
for i=0,n-1 do begin
  w=where(A[column,*] eq values[i],cnt)
  if w[cnt] eq 1 then B[0,i]=A[*,where(A[column,*] eq values[i])] else $
    B[0,i]=mean(A[*,where(A[column,*] eq values[i])],dim=2,/nan)
endfor
endelse
return,B

end
```

Credits to Mr. Schmassmann