Subject: Re: Duplicates - a new twist
Posted by Bruce Bowler on Tue, 18 May 2004 19:10:28 GMT
View Forum Message <> Reply to Message

On Tue, 18 May 2004 08:32:45 -0400, Ben Tupper put fingers to keyboard and
said:

> Martin Doyle wrote:
>
>> I have a dataset which consists of 3 columns: longitude, latitude and
>> a value for an emission of an air pollutant. European countries report
>> the emission of this pollutant for the latitude longitude coordinates
>> which are within their countries. However, some of the latitude,
>> longitude coordinates lie on the borders of countries and therefore an
>> emission is sometimes reported by 2 or more countries for the same
>> coordinate (i,e. There are multiple instances of the same coordinate
>> within the dataset).
>>
>> What I need to do is to look through the dataset and sum the emissions
>> when the coordinate is the same, resulting in a dataset with unique
>> coordinates and a total emission for each grid point.
>>
>> Does anyone have any ideas about how to go about this? I've seen posts
>> on this newsgroup which have had problems with duplicate values in one
>> column of data, but I'm unsure about how to go about it when there are
>> 2 columns which need to be examined.
>>
>
> Hello,
>
> You should consider using GRID_INPUT. This is from the docs...
>
>
>   The GRID_INPUT procedure preprocesses and sorts two-dimensional
> scattered data points, and removes duplicate values.
>
> Ben

But Ben, he doesn't want to remove dup's, he wants to sum them...
(personally, I would have thought that average was better based on the
description, but what the heck...)

Bruce

--
 +------------------+-------------------------------------- ------------+
Bruce Bowler | What garlic is to salad, insanity is to art. -
1.207.633.9600 | Augustus Saint-Gaudens

bbowler@bigelow.org |
  +------------------+------------------------------------- -----------+