
Subject: Re: edit a large dataset

Posted by [Vince Hradil](#) on Wed, 03 Oct 2007 16:40:58 GMT

[View Forum Message](#) <> [Reply to Message](#)

On Oct 3, 9:48 am, queiny <quein...@yahoo.com> wrote:

> Hi, I have a large dataset (>1e7 records). One field of the record is
> a string 'landuse'. Now I want to add one more field based on the
> 'landuse', i.e., if it is 'forest', then add 'aa', if it is 'urban',
> then add 'bb', etc.. Since there are more than 20 different landuse,
> use 'if-then', or 'case' for each record will be quite inefficient.
>
> I am thinking of 'sort' the records based on 'landuse', and then use
> 'uniq' to locate the starting point and number of records for each
> 'landuse' type and add the new fields accordingly. But how could I go
> back to the original order afterward?
>
> a=sort(data_orig)
> data_sorted=data_orig(a)
> ;processing based on 'data_sorted'
>
> but how could I go back the order of 'data_orig'? Is there a inversion
> of 'sort'.
>
> Thanks,

```
a=sort(data_orig)
data_sorted=data_orig[a]
data_orig=data_sorted[sort(a)]
```

Cheers,
Vince
