

---

Subject: Re: Cut down a big file into several files based on its a column.

Posted by [Maarten\[1\]](#) on Wed, 14 Mar 2007 09:46:41 GMT

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

---

On Mar 14, 1:22 am, "DirtyHarry" <kim20...@gmail.com> wrote:

```
> Good day, everyone! (I wrote 'Good morning', but maybe it is not
> morning for others. anyway) I would like to slice a big data file
> into several files based on its a column, but it was not successful
> yet. Please give me some suggestions.
>
> The contents of my original data file are...
>
> 90|2000|1|1|95|95|95|95|95|96|95|95|94|93|93|93|94|94|94|95| 95|93|91|
> 90|94|95|96|95|
> 90|2000|1|2|96|95|94|96|93|93|76|74|76|81|85|84|76|53|43|40| 39|41|33|
> 33|32|32|33|35|
> 90|2000|1|3|35|34|38|35|29|28|30|29|26|23|25|22|22|30|29|24| 23|24|36|
> 31|34|39|31|34|
> .
> .
> .
>
> This is a 2-dim array, I named this array as 'data(*, *)'.
>
> What I want to do now is to slice this file into 6 files based on the
> information in 2nd column, ( data(1,*) ). This column includes year
> info, and the values are from 2000 to 2006.
```

use your shell (bash) and awk:

```
for year in 2000 2001 2002 2003 2004 2005
do
awk -F'|' -v year=$year '{if($2==year){print $0}}' < hum.txt > HumNWS_
$year.txt
done
```

This assumes a unix-like system, and just copies the lines that match the year to a new file. Seems earlier to me than using IDL.

Maarten

---