

---

Subject: Re: How to read data with format codes?

Posted by [Adam Solomon](#) on Sun, 14 Mar 2010 00:39:41 GMT

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

---

Ah, I've found another problem with this: it doesn't read the first character (at least sometimes) in the third column, so if that column has a value of 1.0234, it'll return 0.0234.

On Mar 13, 7:38 pm, Adam Solomon <rampa...@gmail.com> wrote:

> Hi all, I've been pulling my hair out over using readf with a format  
> code. So I have a data file with three columns of floats, all  
> separated by a |. Here's an example:

>  
> | 355.9559189095| -79.4625574877| 0.279817 |  
> | 355.5057979042| -79.4067330132| 0.189663 |  
> | 359.5215546187| -79.7976864503| 0.161308 |  
> | 357.7890741750| -79.6427829568| 0.026100 |  
> | 354.7163524970| -79.3099884939| 0.026578 |  
> | 354.6329916471| -79.3061286761| 0.025486 |  
> | 0.0561720211 | -79.8234642170| 0.025670 |  
> | 357.3786228505| -79.5728335169| 0.000030 |  
> | 356.7569554568| -79.5138546776| 0.199416 |  
> | 355.0388858964| -79.3264538270| 0.190393 |  
> | 354.8099356150| -79.2912157538| 0.026006 |  
> | 352.5968356937| -79.0184276231| 0.207497 |  
> | 352.4362834602| -79.0094360366| 0.217226 |  
> | 0.1800716026 | -79.8081762850| 0.091767 |  
> | 359.0525340400| -79.7227001135| 0.205676 |  
> | 354.7045386126| -79.2652363648| 0.119250 |

>  
> So every format code I've used which has even come close to working  
> runs into issues at one point or another, probably because of the  
> variable spacing. I want to read this file into three arrays, or a  
> 3xlots array, or whatever - just any way to read in this data!! It's  
> unbelievable that such a simple task is proving to be so difficult.

>  
> I won't bore you with all the format codes I've tried (none of which  
> have worked) but here's the most recent example:

>  
> data=fltarr(3,rows-100);the 100 is so i don't run into EOF issues yet,  
> just for testing  
> readf,lun,data,format='(3(2x,F0.0))'

>  
> The problems with this particular one are that when the first column  
> is longer (as in most of these examples), the minus sign doesn't get  
> read in the second column, and there may be issues reading the third  
> column (not sure; will look into it some more).

>

> What's the appropriate way to read these data in?

---

---

Subject: Re: How to read data with format codes?

Posted by [David Fanning](#) on Sun, 14 Mar 2010 01:04:26 GMT

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

---

Adam Solomon writes:

> Hi all, I've been pulling my hair out over using readf with a format  
> code. So I have a data file with three columns of floats, all  
> separated by a |. Here's an example:

>  
> | 355.9559189095| -79.4625574877| 0.279817 |  
> | 355.5057979042| -79.4067330132| 0.189663 |  
> | 359.5215546187| -79.7976864503| 0.161308 |  
> | 357.7890741750| -79.6427829568| 0.026100 |  
> | 354.7163524970| -79.3099884939| 0.026578 |  
> | 354.6329916471| -79.3061286761| 0.025486 |  
> | 0.0561720211 | -79.8234642170| 0.025670 |  
> | 357.3786228505| -79.5728335169| 0.000030 |  
> | 356.7569554568| -79.5138546776| 0.199416 |  
> | 355.0388858964| -79.3264538270| 0.190393 |  
> | 354.8099356150| -79.2912157538| 0.026006 |  
> | 352.5968356937| -79.0184276231| 0.207497 |  
> | 352.4362834602| -79.0094360366| 0.217226 |  
> | 0.1800716026 | -79.8081762850| 0.091767 |  
> | 359.0525340400| -79.7227001135| 0.205676 |  
> | 354.7045386126| -79.2652363648| 0.119250 |

>  
> So every format code I've used which has even come close to working  
> runs into issues at one point or another, probably because of the  
> variable spacing. I want to read this file into three arrays, or a  
> 3xlots array, or whatever - just any way to read in this data!! It's  
> unbelievable that such a simple task is proving to be so difficult.

>  
> I won't bore you with all the format codes I've tried (none of which  
> have worked) but here's the most recent example:

>  
> data=fltarr(3,rows-100);the 100 is so i don't run into EOF issues yet,  
> just for testing  
> readf,lun,data,format='(3(2x,F0.0))'

>  
> The problems with this particular one are that when the first column  
> is longer (as in most of these examples), the minus sign doesn't get  
> read in the second column, and there may be issues reading the third  
> column (not sure; will look into it some more).

>

> What's the appropriate way to read these data in?

If you really want a format code, I'm not your man.  
They always totally confuse me. :-)

I would read this "irregular" file like this.

```
rows = File_Lines('example.dat')
data = DblArr(3,rows)
oneLine = ""
OpenR, lun, 'example.dat', /GET_LUN
FOR j=0,rows-1 DO BEGIN
  ReadF, lun, oneLine
  parts = StrSplit(oneLine, '|', /EXTRACT)
  data[* ,j] = Double(parts[1:3])
ENDFOR
Free_Lun, lun
```

Cheers,

David

--

David Fanning, Ph.D.

Fanning Software Consulting, Inc.

Coyote's Guide to IDL Programming: <http://www.dfanning.com/>

Sepore ma de ni thui. ("Perhaps thou speakest truth.")

---

Subject: Re: How to read data with format codes?

Posted by [penteado](#) on Sun, 14 Mar 2010 01:19:50 GMT

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

---

The easiest way I see is not with format codes, but with `read_ascii` with a template:

```
templ=ascii_template(file)
a=read_ascii(file,template=templ)
```

If you want it to be non-interactive, you can either use `ascii_template` once, then save the resulting template, or generate one yourself, as in:

```
template={version:1.0,datastart:0L,delimiter:byte('|'),missingvalue:!  
values.d_nan,$  
commentssymbol:',fieldcount:[4L],fieldtypes:[0L,5L,5L,5L],$  
fieldnames:['FIELD1','FIELD2','FIELD3','FIELD4'],$  
fieldlocations:[0L,0L,0L,0L],fieldgroups:[0L,1L,2L,3L]}
```

The most relevant here being delimiter, fieldcount and fieldtype.

If you had commas instead of "|" separating the columns, none of that would be needed, you would be able to just use read\_csv.

Another way would be to read the file into a string array, then use strsplit to separate each line at the "|":

```
nlines=file_lines(file)
a=strarr(nlines)
openr,unit,file,/get_lun
readf,unit,a
free_lun,unit
a=strtrim(a,2)
b=dblarr(3,nlines)
for i=0,nlines-1 do b[* ,i]=strsplit(a[i],'|',/extract)
```

---

---

Subject: Re: How to read data with format codes?

Posted by [Adam Solomon](#) on Sun, 14 Mar 2010 01:20:44 GMT

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

---

On Mar 13, 8:04 pm, David Fanning <n...@dfanning.com> wrote:

> Adam Solomon writes:

>> Hi all, I've been pulling my hair out over using readf with a format  
>> code. So I have a data file with three columns of floats, all  
>> separated by a |. Here's an example:

```
>
>> | 355.9559189095| -79.4625574877| 0.279817 |
>> | 355.5057979042| -79.4067330132| 0.189663 |
>> | 359.5215546187| -79.7976864503| 0.161308 |
>> | 357.7890741750| -79.6427829568| 0.026100 |
>> | 354.7163524970| -79.3099884939| 0.026578 |
>> | 354.6329916471| -79.3061286761| 0.025486 |
>> | 0.0561720211 | -79.8234642170| 0.025670 |
>> | 357.3786228505| -79.5728335169| 0.000030 |
>> | 356.7569554568| -79.5138546776| 0.199416 |
>> | 355.0388858964| -79.3264538270| 0.190393 |
>> | 354.8099356150| -79.2912157538| 0.026006 |
>> | 352.5968356937| -79.0184276231| 0.207497 |
>> | 352.4362834602| -79.0094360366| 0.217226 |
>> | 0.1800716026 | -79.8081762850| 0.091767 |
>> | 359.0525340400| -79.7227001135| 0.205676 |
>> | 354.7045386126| -79.2652363648| 0.119250 |
```

>
>> So every format code I've used which has even come close to working  
>> runs into issues at one point or another, probably because of the

```

>> variable spacing. I want to read this file into three arrays, or a
>> 3xlots array, or whatever - just any way to read in this data!! It's
>> unbelievable that such a simple task is proving to be so difficult.
>
>> I won't bore you with all the format codes I've tried (none of which
>> have worked) but here's the most recent example:
>
>> data=fltarr(3,rows-100);the 100 is so i don't run into EOF issues yet,
>> just for testing
>> readf,lun,data,format='(3(2x,F0.0))'
>
>> The problems with this particular one are that when the first column
>> is longer (as in most of these examples), the minus sign doesn't get
>> read in the second column, and there may be issues reading the third
>> column (not sure; will look into it some more).
>
>> What's the appropriate way to read these data in?
>
> If you really want a format code, I'm not your man.
> They always totally confuse me. :-)
>
> I would read this "irregular" file like this.
>
> rows = File_Lines('example.dat')
> data = DblArr(3,rows)
> oneLine = ""
> OpenR, lun, 'example.dat', /GET_LUN
> FOR j=0,rows-1 DO BEGIN
>     ReadF, lun, oneLine
>     parts = StrSplit(oneLine, '|', /EXTRACT)
>     data[*,j] = Double(parts[1:3])
> ENDFOR
> Free_Lun, lun
>
> Cheers,
>
> David
>
> --
> David Fanning, Ph.D.
> Fanning Software Consulting, Inc.
> Coyote's Guide to IDL Programming:http://www.dfanning.com/
> Sepore ma de ni thui. ("Perhaps thou speakest truth.")

```

Thanks! I was going to do that but I read on this website - dfanning something or other - to avoid for loops in IDL :) Figured a format code would save time since there's on the order of  $10^6$  data points here, but I guess that's nothing compared to all the time I'm spending

writing the damn code!

Adam

---

---

Subject: Re: How to read data with format codes?

Posted by [penteado](#) on Sun, 14 Mar 2010 01:45:52 GMT

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

---

On Mar 13, 10:20 pm, Adam Solomon <rampa...@gmail.com> wrote:

> Thanks! I was going to do that but I read on this website - dfanning  
> something or other - to avoid for loops in IDL :) Figured a format  
> code would save time since there's on the order of  $10^6$  data points  
> here, but I guess that's nothing compared to all the time I'm spending  
> writing the damn code!

In this case, it is too awkward to avoid a loop. It would be easy if  
strpos accepted an array as its pos argument (which it really should,  
to work nicely with strmid):

```
nlines=file_lines(file)
a=strarr(nlines)
openr,unit,file,/get_lun
readf,unit,a
free_lun,unit
;Get rid of the initial |
pos=transpose(strpos(a,'|'))
a=strmid(a,pos+1)
;Extract the first column
pos=transpose(strpos(a,'|'))
col1=double(strmid(a,0,pos))
a=strmid(a,pos+1)
;Extract the second column
pos=transpose(strpos(a,'|'))
col2=double(strmid(a,0,pos))
a=strmid(a,pos+1)
;Extract the third column
pos=transpose(strpos(a,'|'))
col3=double(strmid(a,0,pos))
```

---

---

Subject: Re: How to read data with format codes?

Posted by [R.G.Stockwell](#) on Sun, 14 Mar 2010 02:03:02 GMT

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

---

"Adam Solomon" <ramparts@gmail.com> wrote in message  
news:e5bdbfda-121a-41fb-8204-57764e0cc70d@g10g2000yqh.google groups.com...

On Mar 13, 8:04 pm, David Fanning <n...@dfanning.com> wrote:  
> Adam Solomon writes:  
>> Hi all, I've been pulling my hair out over using readf with a format  
>> code. So I have a data file with three columns of floats, all  
>> separated by a |. Here's an example:  
>  
>> | 355.9559189095| -79.4625574877| 0.279817 |  
...  
> rows = File\_Lines('example.dat')  
> data = DblArr(3,rows)  
> oneLine = ""  
> OpenR, lun, 'example.dat', /GET\_LUN  
> FOR j=0,rows-1 DO BEGIN  
> ReadF, lun, oneLine  
> parts = StrSplit(oneLine, '|', /EXTRACT)  
> data[:,j] = Double(parts[1:3])  
> ENDFOR  
> Free\_Lun, lun

Yep. Always use the same rules to read, as you do to write.  
If they wrote it with a format, you read it with the same format.  
If they wrote "stuff" and then "|", then that is how you read it  
(as David has shown).

btw, one suggestion. perhaps you can write a Convert program  
to read this huge ascii data set once, and write it in a quick binary  
format (perhaps as a structure with all pertinent information).

cheers,  
bob

---

Subject: Re: How to read data with format codes?  
Posted by [David Fanning](#) on Sun, 14 Mar 2010 02:10:49 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Adam Solomon writes:

> Thanks! I was going to do that but I read on this website - dfanning  
> something or other - to avoid for loops in IDL :) Figured a format  
> code would save time since there's on the order of 10^6 data points  
> here, but I guess that's nothing compared to all the time I'm spending  
> writing the damn code!

Well, 30 seconds to write the code I sent you,  
10 minutes to read the damn file the first time,

and write it out the way it should have been written the first time, then \*all\* of your format code versions should have worked. :-)

Cheers,

David

--

David Fanning, Ph.D.  
Fanning Software Consulting, Inc.  
Coyote's Guide to IDL Programming: <http://www.dfanning.com/>  
Sepore ma de ni thui. ("Perhaps thou speakest truth.")

---

---

Subject: Re: How to read data with format codes?  
Posted by [Kenneth P. Bowman](#) on Sun, 14 Mar 2010 17:23:59 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

In article  
<e5bdbfda-121a-41fb-8204-57764e0cc70d@g10g2000yqh.googlegroups.com>,  
Adam Solomon <ramparts@gmail.com> wrote:

> Thanks! I was going to do that but I read on this website - dfanning  
> something or other - to avoid for loops in IDL :) Figured a format  
> code would save time since there's on the order of 10^6 data points  
> here, but I guess that's nothing compared to all the time I'm spending  
> writing the damn code!

When you are doing I/O, the slow parts of the operation are most likely to be getting the data from the disk and converting ASCII numerals to their internal binary representation. Using a loop should have little impact on the time required.

Ken Bowman

---