



Macro-free parsing of multiple text files with SAS

Adam Amborski

2016-10-12



Agenda

- Aim
- Potential Pitfall
- Process
- Success

Aim: Modify lots of files exactly where needed

Why not add a comment giving their current location into all of our programs:

```
/* program_a.sas is stored on h: drive. */
```

Potential Pitfall

No changes like this one:

```
diff ...line 2 found changed as line 2 of the other file...
2c2
<      a="1"; /* some blanks follow here */
---
>      a="1"; /* some blanks follow here */
```

It's the (removed) trailing blanks that make the unwanted change!

<pre>1 data.pies;␣ -2a="1"; /* 3 blanks follow here */␣ 3 run;␣</pre>	<pre>1 data.pies;␣ +2a="1"; /* 3 blanks follow here */␣ 3 run;␣</pre>
---	---

Process

Get list of files:

```
data files (keep=filename full_path label="All SAS programs in a directory");  
  length filename full_path $1000;  
  rc=filename("dir", "&dir.");  
  did=dopen("dir");  
  if did ne 0 then do;  
    do i=1 to dnum(did);  
      filename=dread(did,i);  
      full_path="&dir"||filename;  
      if lowercase(scan(filename, -1, "."))="sas" then output;  
    end;  
  end;  
  else do;  
    put "ERROR: Failed to open the directory &dir..";  
    stop;  
  end;  
run;
```

Process

Read in file contents:

```
option lrecl=&lrecl;

data contents (label="Original files' full contents"
               rename=(full_path_tmp=full_path));

  set files end=last;
  length full_line $&lrecl;
  infile codes filevar=full_path length=line_length_tmp end=all;
  full_path_tmp=full_path;
  line_number=0;
  do while(not all);
    input;
    full_line=_infile_;
    line_length=line_length_tmp;
    line_number+1;
    output;
  end;
run;
```

Process

Manipulation: simple example

```
data processed (label="Manipulate file contents");  
  set contents;  
  by full_path;  
  output;  
  if last.full_path and index(filename,"dont_modify")=0 then  
    do;  
      full_line=catx(" ", "/*", filename, "is stored on",  
                    scan(full_path, 1, "\""), "drive.  
*/");  
      line_length=length(full_line);  
      line_number=line_number+1;  
      output;  
    end;  
run;
```

Process

Write the files:

```
data _NULL_;  
  set tosave;   
  file sascode filevar=full_path;  
  put full_line $varying. line_length;  
run;
```

Success

Diff output:

...after line 3 of the first file, added line 4 in the other file...

3a4

> /* */

...after line 4 of the first file, added line 5 in the other file...

4a5

> /* program_b.sas is stored on h: drive. */

... and no other differences!



Thank You!

Adam Amborski
inVentiv Health Clinical Poland
ul. Nowogrodzka 68
02-014 Warszawa
Adam.Amborski@inventivhealth.com
www.inventivhealth.com