


Adding html tags using SAS

Marco

Hing-Poon Chan
Canada Revenue Agency




Background

- Income Statistics Reports on PDF format
- Raw data from IBM mainframe (flat files)
- Download to PC and  convert to dbf format
- 19 different report templates
- Over 600 reports in both languages (Canada and provinces)
- Reports created with R&R Report Writer





 Canada Customs and Revenue Agency / Agence des douanes et du revenu du Canada

Income Statistics 2003 2001 tax year
Final Basic Table 2 - Sample Data

All returns by total income class (all money figures in thousands of dollars)

Item	Loss and nil		\$1 - \$10,000		\$10,000 - \$15,000	
	Number	Amount \$	Number	Amount \$	Number	Amount \$
Number of taxable returns	1	1,380	556,440		1,406,800	
Number of non-taxable returns	2	860,450	3,994,090		1,535,050	
Total number of returns	3	861,820	4,550,530		2,941,850	
Sources of income						
Employment income	4	13,510	99,832	2,403,190	10,755,624	1,319,500
Commissions (from employment)	5	380	5,870	53,810	38,776	36,430
Other employment income	6	1,990	7,139	221,420	560,499	123,550
Old Age Security pension	7	2,360	10,269	376,330	1,524,623	1,027,220
CPP or QPP benefits	8	4,970	21,275	614,510	1,853,819	1,034,090
Other pensions or superannuation	9	1,290	10,564	78,650	216,256	181,580
Employment insurance benefits	10	1,990	7,195	277,800	853,142	341,840
Taxable amount of dividends	11	7,020	8,492	245,190	90,963	217,930
Investment income	12	21,160	37,777	1,059,980	1,070,591	817,750
Annuity income	13	430	1,212	61,230	109,085	131,270
Net rental income	14	13,930	(121,943)	106,470	62,895	62,750
Taxable capital gains	15	4,780	13,715	166,520	133,143	110,110
RRSP income	16	5,560	23,117	132,920	325,972	96,430
Net business income	17	51,400	(563,091)	335,870	1,099,055	246,710
Net professional income	18	2,760	(37,827)	36,150	125,925	23,630
Net commission income	19	3,190	(21,514)	24,500	67,025	18,990
Net farming income	20	12,920	(218,493)	60,030	26,190	60,800
Net fishing income	21	190	(2,476)	3,780	1,501	4,590
Tax-exempt income	22	2,540	10,677	1,049,080	4,883,429	1,206,340
Other income	23	10,760	(113,298)	475,590	902,240	299,350
Total income assessed	24	77,630	(821,509)	4,550,530	24,700,752	2,941,850
Deductions						
RPP contributions	25	2,760	6,277	63,210	20,485	81,530
RRSP contributions	26	3,110	2,889	112,090	115,554	228,250
Union and professional dues	27	5,920	1,256	392,800	42,662	278,680
Child care expenses	28	880	988	93,290	107,170	103,070
Carrying charges and interest expenses	29	5,030	27,259	106,880	78,662	157,600
CPP or QPP contributions on self-employment or other ear	30	940	130	148,540	22,420	192,140
Other employment expenses	31	930	3,134	16,610	26,415	26,510
Cleric's residence deduction	32			200	1,559	1,550
Other deductions (from total income)	33	4,930	44,143	106,970	130,406	83,290
Total deduction before adjustments	34	20,670	86,077	874,820	545,331	868,680

Treasury Board Guideline

- Treasury Board new accessibility guideline for the visually impaired
- All government department and Agency websites to be complied with the guideline by January 2003
- Original PDF format do not comply with Treasury Board Guideline
- CCRA Internet Content Group set up templates for html representation of reports in table format for all groups to follow



CCRA Template

- All table tags (e.g. `<th>`, `<td>`) have to have unique identifiers within the same html page
- Allows person with impaired vision to access data on each cell (with the identifiers) with screen reader software (audio driven)





CCRA Template example

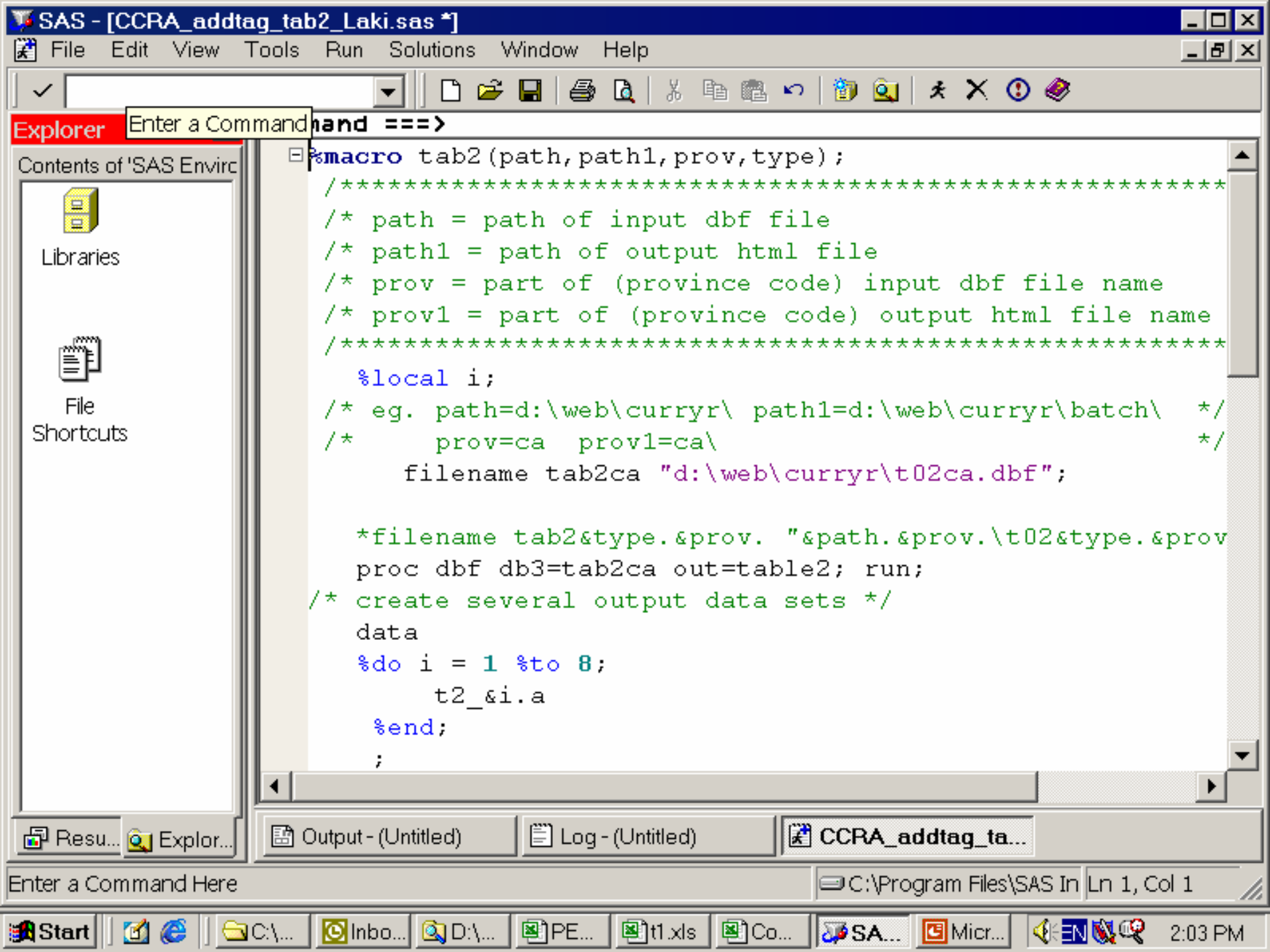
```
<table...><caption>...</caption>  
<tr> <th id="header1">..cell content..</th>  
<th id="header2">.. cell content ..</th>  
...  
<th id="headern">.. cell content ..</th> <tr>  
<tr><th id="row_header1" headers="header1"> .. cell  
content ..</th>  
<th id="row_header2" headers="header2"> .. cell content  
..</th>  
...  
<th id="row_headern" headers="headern"> .. cell content  
..</th>  
</tr></table>
```



Using SAS Macro to insert table tags with unique identifiers

- There are 19 different reports
- Each report for Canada and every province in both official languages
- Amount to over 600 reports
- A lot of repetitions
- Good candidate for SAS Macro





Enter a Command

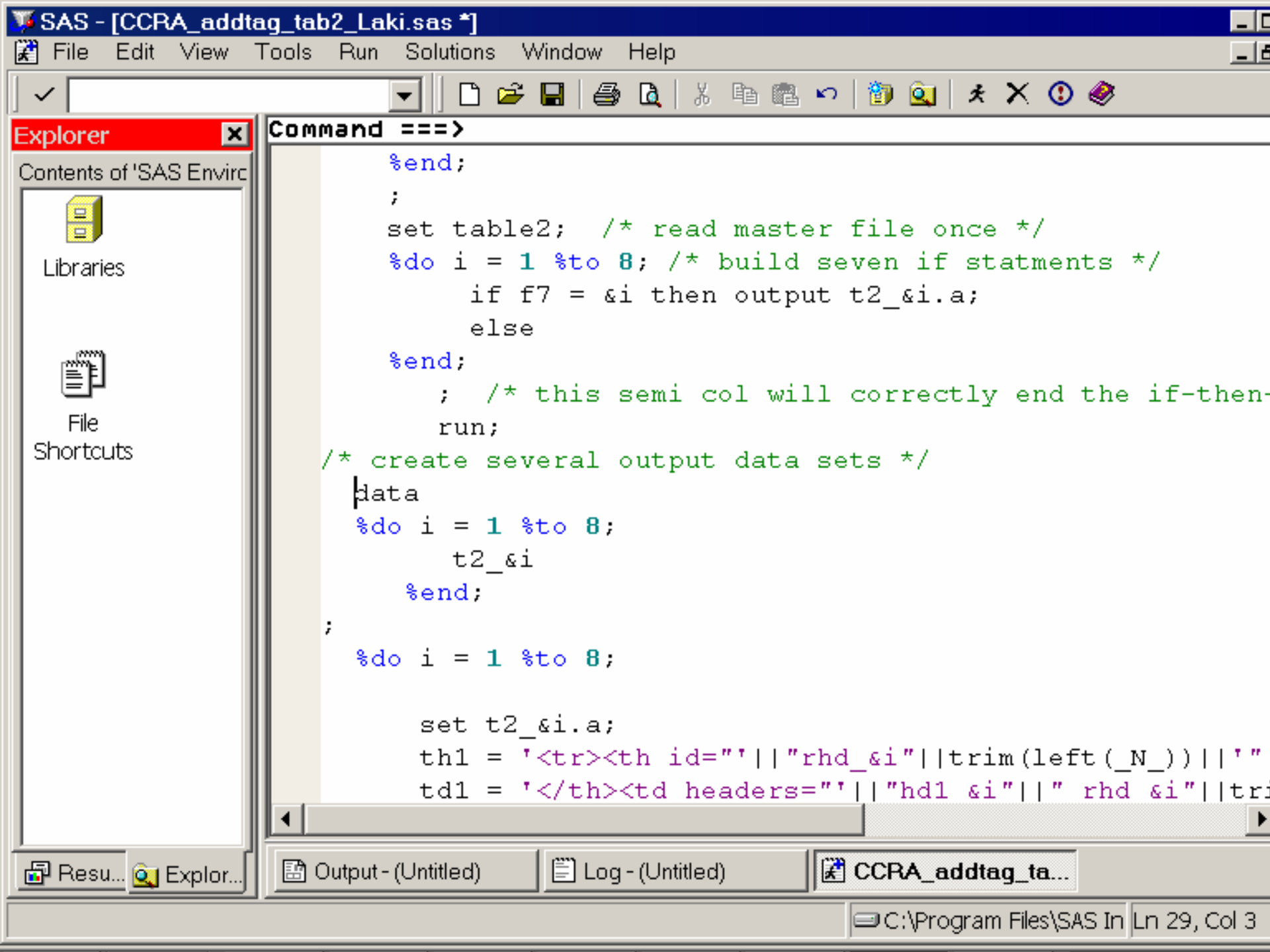
Command ==>

Contents of 'SAS Envir...

Libraries

File Shortcuts

```
%macro tab2 (path, path1, prov, type);  
/*****  
/* path = path of input dbf file  
/* path1 = path of output html file  
/* prov = part of (province code) input dbf file name  
/* prov1 = part of (province code) output html file name  
/*****  
  %local i;  
/* eg. path=d:\web\curryr\ path1=d:\web\curryr\batch\ */  
/*   prov=ca prov1=ca\ */  
  filename tab2ca "d:\web\curryr\t02ca.dbf";  
  
  *filename tab2&type.&prov. "&path.&prov.\t02&type.&prov  
  proc dbf db3=tab2ca out=table2; run;  
/* create several output data sets */  
data  
  %do i = 1 %to 8;  
    t2_&i.a  
  %end;  
;
```

Explorer

Contents of 'SAS Envir...



Libraries

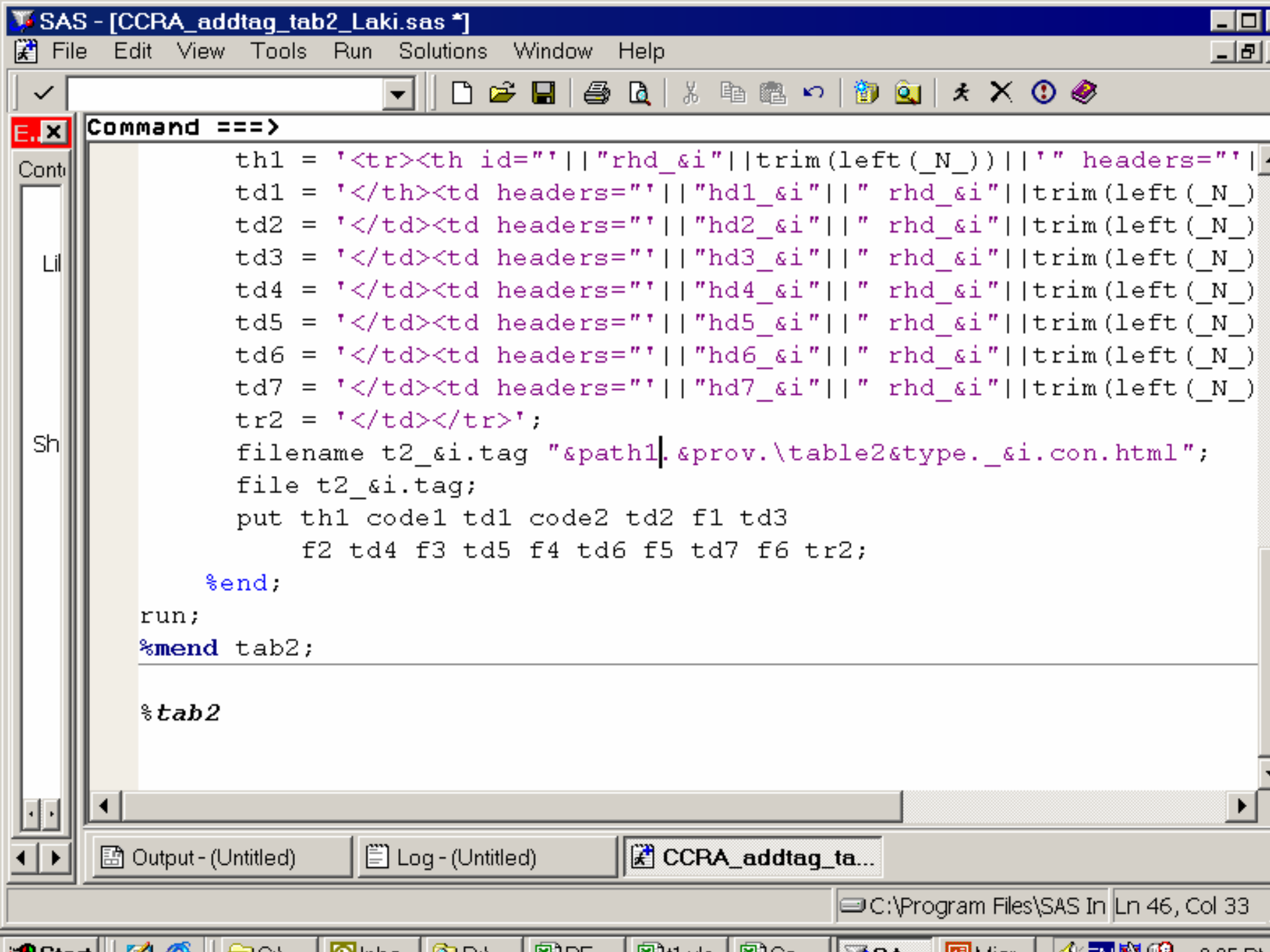


File Shortcuts

Command ==>

```
%end;
;
set table2; /* read master file once */
%do i = 1 %to 8; /* build seven if statments */
    if f7 = &i then output t2_&i.a;
    else
%end;
    ; /* this semi col will correctly end the if-then-
run;
/* create several output data sets */
data
%do i = 1 %to 8;
    t2_&i
%end;
;
%do i = 1 %to 8;

    set t2_&i.a;
    th1 = '<tr><th id="" || "rhd_&i" || trim(left(_N_)) || '
    td1 = '</th><td headers="" || "hd1_&i" || " rhd_&i" || tri
```



Command ==>

```

th1 = '<tr><th id="" || "rhd_&i" || trim(left(_N_)) || "' headers="" |
td1 = '</th><td headers="" || "hd1_&i" || " rhd_&i" || trim(left(_N_
td2 = '</td><td headers="" || "hd2_&i" || " rhd_&i" || trim(left(_N_
td3 = '</td><td headers="" || "hd3_&i" || " rhd_&i" || trim(left(_N_
td4 = '</td><td headers="" || "hd4_&i" || " rhd_&i" || trim(left(_N_
td5 = '</td><td headers="" || "hd5_&i" || " rhd_&i" || trim(left(_N_
td6 = '</td><td headers="" || "hd6_&i" || " rhd_&i" || trim(left(_N_
td7 = '</td><td headers="" || "hd7_&i" || " rhd_&i" || trim(left(_N_
tr2 = '</td></tr>';
filename t2_&i.tag "&path1|. &prov.\table2&type._&i.con.html";
file t2_&i.tag;
put th1 code1 td1 code2 td2 f1 td3
    f2 td4 f3 td5 f4 td6 f5 td7 f6 tr2;

%end;
run;
%mend tab2;

%tab2

```

SAS - [Editor - Untitled1]

File Edit View Tools Run Solutions Window Help

Exp... Contents

```
data t2_1;
  set t2_1a;
  th1 = '<tr><th id=""||"rhd_1"||trim(left(_N_))||" headers=""||
        "hd0_1"||">';
  td1 = '</th><td headers=""||"hd1_1"||" rhd_1"||trim(left(_N_))||
        " align="right">';
  td2 = '</td><td headers=""||"hd2_1"||" rhd_1"||trim(left(_N_))||
        " align="right">';
  td3 = '</td><td headers=""||"hd3_1"||" rhd_1"||trim(left(_N_))||
        " align="right">';
  td4 = '</td><td headers=""||"hd4_1"||" rhd_1"||trim(left(_N_))||
        " align="right">';
  td5 = '</td><td headers=""||"hd5_1"||" rhd_1"||trim(left(_N_))||
        " align="right">';
  td6 = '</td><td headers=""||"hd6_1"||" rhd_1"||trim(left(_N_))||
        " align="right">';
  td7 = '</td><td headers=""||"hd7_1"||" rhd_1"||trim(left(_N_))||
        " align="right">';
  tr2 = '</td></tr>';
  filename t2_1tag "d:\web\curryr\out\ca\table2a_1con.html";
  file t2_1tag;
```

Output - (Untitled) Log - (Untitled) Editor - Untitled1 *

Autosave complete C:\Program Files\SAS In Ln 22, Col 1

Start C:\Documents... Inbox - Microso... SAS - [Edito... Microsoft Pow... 8:41 AM

SAS - [Editor - Untitled1]

File Edit View Tools Run Solutions Window Help

Exp... Contents

```
td1 = '</th><td headers="" || "hd1_1" || " rhd_1" || trim(left(_N_)) ||  
      " align="right">'  
td2 = '</td><td headers="" || "hd2_1" || " rhd_1" || trim(left(_N_)) ||  
      " align="right">'  
td3 = '</td><td headers="" || "hd3_1" || " rhd_1" || trim(left(_N_)) ||  
      " align="right">'  
td4 = '</td><td headers="" || "hd4_1" || " rhd_1" || trim(left(_N_)) ||  
      " align="right">'  
td5 = '</td><td headers="" || "hd5_1" || " rhd_1" || trim(left(_N_)) ||  
      " align="right">'  
td6 = '</td><td headers="" || "hd6_1" || " rhd_1" || trim(left(_N_)) ||  
      " align="right">'  
td7 = '</td><td headers="" || "hd7_1" || " rhd_1" || trim(left(_N_)) ||  
      " align="right">|  
tr2 = '</td></tr>';  
filename t2_ltag "d:\web\curryr\out\ca\table2a_1con.html";  
file t2_ltag;  
put th1 code1 td1 code2 td2 f1 td3 f2 td4 f3 td5 f4 td6 f5 td7 f6
```

Output - (Untitled) Log - (Untitled) Editor - Untitled1 *

Autosave complete C:\Program Files\SAS In Ln 18, Col 26

Start C:\Documents ... Inbox - Microso... SAS - [Edito... Microsoft Pow... 8:42 AM

Libraries

File Shortcuts

```
options mprint;
data pc;
  infile 'd:\web\curryr\pcode.txt';
  input provc $1-2;
  /* use symput to create 15 new macro variables */
  /* ie. pcode1 with value ca, etc */
  /* no need for do loop */
  call symput('pcode'||trim(left(_n_)), provc);
run;
%macro runtab2;
  %local i;
  %do i = 1 %to 15;
    %tab2(d:\web\curryr\,d:\web\curryr\batch\,&&pcode&i,);
    %tab2(d:\web\curryr\,d:\web\curryr\batch\,&&pcode&i,a);
  %end;
run;
%mend runtab2;
%runtab2
```

```
<tr>
<th id="rhd_11" headers="hd0_1"> A </th>
<td headers="hd1_1 rhd_11" align="right"> 1 </td>
<td headers="hd2_1 rhd_11" align="right"> 1380 </td>
<td headers="hd3_1 rhd_11" align="right"> 0 </td>
<td headers="hd4_1 rhd_11" align="right"> 556440</td>
<td headers="hd5_1 rhd_11" align="right"> 0 </td>
<td headers="hd6_1 rhd_11" align="right"> 1406800 </td>
<td headers="hd7_1 rhd_11" align="right"> 0 </td>
</tr>
<tr><th id="rhd_12" headers="hd0_1"> A </th>
<td headers="hd1_1 rhd_12" align="right"> 2 </td>
<td headers="hd2_1 rhd_12" align="right"> 860450 </td>
<td headers="hd3_1 rhd_12" align="right"> 0 </td>
<td headers="hd4_1 rhd_12" align="right"> 3994090</td>
<td headers="hd5_1 rhd_12" align="right"> 0 </td>
<td headers="hd6_1 rhd_12" align="right"> 1535050 </td>
<td headers="hd7_1 rhd_12" align="right"> 0 </td>
</tr>
<tr><th id="rhd_13" headers="hd0_1"> A </th>
<td headers="hd1_1 rhd_13" align="right"> 3 </td>
```

```
<tr>
<th id="rhd_163" headers="hd0_1"> F </th>
<td headers="hd1_1 rhd_163" align="right"> 63 </td>
<td headers="hd2_1 rhd_163" align="right"> 210 </td>
<td headers="hd3_1 rhd_163" align="right"> 9250 </td>
<td headers="hd4_1 rhd_163" align="right"> 94200</td>
<td headers="hd5_1 rhd_163" align="right"> 35054 </td>
<td headers="hd6_1 rhd_163" align="right"> 598530 </td>
<td headers="hd7_1 rhd_163" align="right"> 633163 </td>
</tr>
<tr>
<th id="rhd_164" headers="hd0_1"> F </th>
<td headers="hd1_1 rhd_164" align="right"> 64 </td>
<td headers="hd2_1 rhd_164" align="right"> 1380 </td>
<td headers="hd3_1 rhd_164" align="right"> 12884 </td>
<td headers="hd4_1 rhd_164" align="right"> 556440</td>
<td headers="hd5_1 rhd_164" align="right"> 207605 </td>
<td headers="hd6_1 rhd_164" align="right"> 1406800 </td>
<td headers="hd7_1 rhd_164" align="right"> 1414378 </td>
</tr>
```

Adding html tags using SAS

Marco

- Use DOS batch files to create the complete html files – use “copy” command to copy all the html files created by the macro





Français	Contact us	Help	Search	Canada Site
What's new	Media room	Electronic mailing lists	Electronic services	Forms and publications
Home	Site map			

[Tax](#) >

Final Basic Table 2: All returns by total income class, all Canada

- Refer to the [table below](#).
- The first field -- CODE1 -- is the code for the "Item" group. To identify the items group, see [t2item-e.html](#).
- The second field -- CODE2 -- is the code for the individual item. To identify the individual items, see [tab2-5-e.html](#).
- Fields F1 to F6 are the total number of returns and amount for the income class (grouped by F7). Each of them has 65 rows, for a total of 8 groups (F7 from 1 to 8). To identify the income classes, see [TAB2_F1-F7-e.html](#).

Income Statistics

- [Main page](#)

Final Statistics - 2003 Edition

- [Sample data](#)
- [Major changes for the 2001 tax year](#)
- [Major](#)

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites History Print Stop Home

Address <http://www.ccr-a-drc.gc.ca/tax/individuals/stats/gb01/pst/final/table2-e.html> Go**tables**

- .html format
- .pdf format
- .csv format

f7 = 1

code1	code2	f1	f2	f3	f4	f5	f6
A	1	1380	0	556440	0	1406800	0
A	2	860450	0	3994090	0	1535050	0
A	3	861820	0	4550530	0	2941850	0
B	4	13510	99832	2403190	10755624	1319500	13005910
B	5	380	5870	53810	38776	36430	60412
B	6	1990	7139	221420	560499	123550	425415
B	7	2360	10269	376330	1524623	1027220	4921094
B	8	4970	21275	614510	1853819	1034090	4210872
B	9	1290	10564	78650	216256	181580	576437
B	10	1990	7195	277800	853142	341840	1396311
B	11	7020	8492	245190	90963	217930	156049



Internet

Start Internet Explorer C:\ In... Ot... D... D... rc... rc... A... Mi... F...

Volume Network Security Time 8:08 AM