

Basic Concepts for Documenting SAS® Projects

Documentation Styles for SAS Projects, Programs, and Variables

Potential
of One

Power
of
All

Basic Concepts for Documenting SAS® Projects

Documentation Styles for SAS Projects, Programs, and Variables

Peter Timusk
Statistics Canada

Abstract

This paper kicks off a project to write a comprehensive book of best practices for documenting SAS® projects. The presenter's existing documentation styles are explained. The presenter wants to discuss and gather current best practices used by the SAS user community. The presenter shows documentation styles at three different levels of scope. The first is a style used for project documentation, the second a style for program documentation, and the third a style for variable documentation. This third style enables researchers to repeat the modeling in SAS research, in an alternative language, or conceptually.

Objective

- To introduce a book project on documentation
- To share effective documentation
- To start/continue a conversation
- Advance the fields of information management and knowledge management

Method

- A production officer's project summary template is shared.
- A programmer's header file for a program is shared.
- A demographer with 50 years of experience developed a template for documenting SAS variables. This is shared.

Examples

- [Project Summary Template](#)
- [Program Header File](#)
- [Variable Documentation Template](#)

Conclusions

- A number of SAS Global Forum papers have been written about SAS programming style and documentation.
- It may be possible to present some best practices in a book form.
- This presentation has presented some examples of documentation.

References

- Peter Timusk, 2011 Internal Document, Statistics Canada, Ottawa, Ontario, Canada.
- Francois Mainville, 2009 Internal Document, Statistics Canada, Ottawa, Ontario, Canada.
- Leroy Stone, 2008 Internal Document, Statistics Canada, Ottawa, Ontario, Canada.

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies.



Washington, D.C.
March 23–26, 2014

Basic Concepts for Documenting SAS® Projects

Documentation Styles for SAS Projects, Programs, and Variables

Peter Timusk
Statistics Canada

Example 1 Project Summary Template

Project Summary Template (This is typically a 2 page Microsoft Word document stored in the project folder.)

Author: Peter Timusk

Date: July 8th, 2011.

Title: Documentation of Table Production work on the Federal Project X Business Survey (FPXBS) 2008 - Ontario files

Dates of Project: 17th of February 2011 to the 4th of March 2011.

Approximate duration: 10 days at 1 Programmer-04 level resource.

The major task was tabulation of new survey data: High level view: Place estimate values in the Excel sheet and check for confidentiality with new survey data.

Database and input files: Survey files: FPXBS2008_sample_20080121_v1, FPXBS2008_weights_20080121_v1

Specifications: An Excel template was provided by the Survey manager. This template lists all row and column variables and variance values to be calculated for each estimate.

Details:

Edits: The standard estimation program was edited to fit the template's specs and use the survey file and weights.

Confidentiality details: The robustness of the sample was analysed with a PROC FREQ and where the counts were too low the estimates were suppressed. An in-house Macro was used to check for residual suppression and the pattern from this macro was then applied. The suppression pattern was then reviewed by the survey manager.

SAS programs and other tools used in the project: Table programs used the in-house standard table production template. This program called the estimation macro, and the residual suppression macro. Finally a SAS Dynamic Data Exchange (DDE) program placed the SAS datasets in an Excel workbook template.

First set of programs:

The estimation and residual suppression programs are located here: \\networkhost\pl1\Data\ProjectFPXBS\2008\Tables\Table_1\Code

The tabulation programs are located here \\networkhost\pl1\Data\ProjectFPXBS\2008\DDE_Tabulation

Two final steps:

1. A PROC TABULATE program was run independently to produce the same estimates without survey specific variances and also tabulated with DDE and then the two Excel spreadsheets were compared with Visual Basic Macro and estimation programs were adjusted until all estimates matched.

2. The survey manager examined the results and compared this to previous year's estimates and other sources for coherency.

Delivery: The location of the final Excel tables is presently \\networkhost\pl1\Data\ProjectFPXBS\2008\Tables\Table_1\Output\Final\FPXBS2008_Table_1_v1.xls



Washington, D.C.
March 23–26, 2014

[Beginning](#)

[Main Page](#)

[End](#)

Basic Concepts for Documenting SAS® Projects

Documentation Styles for SAS Projects, Programs, and Variables

Peter Timusk
Statistics Canada

Example 2 Program Header File

Program Header File (This is header SAS comments file placed at the top of a SAS program.)

```
• /*****
• *   Prog Name : FPXBS2008_Table1_v1.sas
• *   Prog Path : \\networkhost\pl1\Data\ProjectFPXBS\2008\Tables\Table_1\Code
• *   SAS Ver#  : 9.2 - Template Version 25
• *
• *   Prog Version : 1.1
• *
• *   Author      : Peter Timusk
• *   Revised By  : Peter Timusk
• *   Proofed By   : TEAM
• *   Date Created : February / 19 / 2011   Last Revised : March / 2 / 2011
• *
• *
• *   Prog Desc   : This program prepares a series of survey variable specifically for this table
• *               : The program then calls the estimation macro and the suppression macro.
• *
• *   Comments    : This is the generic table program to be
• *               : changed and customized for each table.
• *
• *
• *   Project     : Federal Project X Business Survey (FPXBS) 2008 - Ontario files
• *
• *****/
```




Washington, D.C.
March 23–26, 2014

[Beginning](#)

[Main Page](#)

[End](#)

Basic Concepts for Documenting SAS® Projects

Documentation Styles for SAS Projects, Programs, and Variables

Peter Timusk
Statistics Canada

Example 3 Variable Documentation Template Slide 1 of 5

Variable Documentation Template (This is text file and every new variable in the modeling program has a file like this.) (slide1)

file = Docum var educ3.txt

```
/*Documentation for the variable "educ3"  
Date 08-11-03, Author: Peter Timusk  
Modified 08-11-03 Peter Timusk  
*/  
Variable = educ3  
Code from impute & model KnowRisk in GSS.sas*/  
/* end header file*/
```

SAS CODE

==== in SOM ====

educ3 levels:

- 1 HS diploma and lower
- 2 Comm College + Some Univ
- 3 Univ degree
- 0 Others

definition code in SOM

```
educ3 = 0;  
IF educ2 LT 3 THEN educ3 = 1; / HS diploma and lower /  
ELSE IF educ2 EQ 3 THEN educ3 = 2; / Comm College + Some Univ /  
ELSE IF educ2 EQ 4 THEN educ3 = 3; / Univ degree /  
*/
```

Continued on slide 2

[Beginning](#)

[Main Page](#)

[End](#)

Basic Concepts for Documenting SAS® Projects

Documentation Styles for SAS Projects, Programs, and Variables

Peter Timusk
Statistics Canada

Example 3 Variable Documentation Template Slide 2 of 5

Variable Documentation Template (This is text file and every new variable in the modeling program has a file like this.) (slide2)

Continued from slide 1

DERIVED VARIABLE:

```
educ2 = 0;  //  
IF Z2G EQ 1 OR Z2G EQ 3 THEN Educ2 = 1;  No HS diploma  
ELSE IF Z2G EQ 2 THEN Educ2 = 2;  / HS diploma /  
ELSE IF Z2G LT 6 THEN Educ2 = 3;  / Comm College + Some Univ /  
ELSE IF Z2G LT 90 THEN Educ2 = 4;  / Univ degree /
```

DETAILED DOCUMENTATION

```
* IF educ2 LT 3 THEN educ3 = 1;  / HS diploma and lower /  
[ educ2: Schooling
```

- 1 None
- 2 High school diploma or equivalent
- 3 Certificate or diploma from a vocational school

```
]
```

```
* ELSE IF educ2 EQ 3 THEN educ3 = 2;  / Comm College + Some Univ /  
[ educ2: Schooling
```

- 4 Community College or Cegep certificate or diploma
- 5 University certificate or diploma, pre-bachelor's degree

```
]
```

Continued on slide 3

[Beginning](#)

[Main Page](#)

[End](#)

Basic Concepts for Documenting SAS® Projects

Documentation Styles for SAS Projects, Programs, and Variables

Peter Timusk
Statistics Canada

Example 3 Variable Documentation Template Slide 3 of 5

Variable Documentation Template (This is text file and every new variable in the modeling program has a file like this.) (slide3)

Continued from slide 2

```
* ELSE IF educ2 EQ 4 THEN educ3 = 3; / Univ degree /
```

```
[ educ2: Schooling
```

```
    6 Bachelor's degree
```

```
    7 University certificate or diploma post-Bachelor's degree
```

```
    8 Master's degree
```

```
    9 Diploma in medicine, dentistry, veterinary medicine or optometry
```

```
   10 Doctor's degree
```

```
]
```

PSEUDOCODE:

Create a new variable educ3 and assign a value of 0 to this variable.

Assign a value of 1 to the variable if the respondent's highest level of education obtained is None or a High school diploma or equivalent or a Certificate or diploma from a vocational school.

Assign a value of 2 to the variable if the respondent's highest level of education obtained is a Community College or Cegep certificate or diploma or a University certificate or diploma, pre-bachelor's degree.

Assign a value of 3 to the variable if the respondent's highest level of education obtained is a Bachelor's degree or a University certificate or diploma post-Bachelor's degree or a Master's degree or a Diploma in medicine, dentistry, veterinary medicine or optometry or a Doctor's degree.

The value of 0 remains assigned to the variable if the response was "other" or refusal.

Continued on slide 4

[Beginning](#)

[Main Page](#)

[End](#)

Basic Concepts for Documenting SAS® Projects

Documentation Styles for SAS Projects, Programs, and Variables

Peter Timusk
Statistics Canada

Example 3 Variable Documentation Template Slide 4 of 5

Variable Documentation Template (This is text file and every new variable in the modeling program has a file like this.) (slide4)

Continued from slide 3

DOCUMENTATION IN ENGLISH:

A new variable educ3 is created to define the respondent's highest level of education obtained. A value is assigned to the variable according to the level of schooling of the respondent, starting with a value of 1 if the respondent's highest level of education obtained is None or a High school diploma or equivalent or a Certificate or diploma from a vocational school, a value of 2 if the respondent's highest level of education obtained is a Community College or Cegep certificate or diploma or a University certificate or diploma, pre-bachelor's degree, a value of 3 if the respondent's highest level of education obtained is a Bachelor's degree or a University certificate or diploma post-Bachelor's degree or a Master's degree or a Diploma in medicine, dentistry, veterinary medicine or optometry or a Doctor's degree. and a value of 0 if the response was "other" or refusal.

Continued on slide 5

[Beginning](#)

[Main Page](#)

[End](#)

Basic Concepts for Documenting SAS® Projects

Documentation Styles for SAS Projects, Programs, and Variables

Peter Timusk
Statistics Canada

Example 3 Variable Documentation Template Slide 5 of 5

Variable Documentation Template (This is text file and every new variable in the modeling program has a file like this.) (slide5)

Continued from slide 4

REFERENCE:

Z2G Schooling Z2G

MATRIX : Z2G

- 1 None
- 2 High school diploma or equivalent
- 3 Certificate or diploma from a vocational school
- 4 Community College or Cegep certificate or diploma
- 5 University certificate or diploma, pre-bachelor's degree
- 6 Bachelor's degree
- 7 University certificate or diploma post-Bachelor's degree
- 8 Master's degree
- 9 Diploma in medicine, dentistry, veterinary medicine or optometry
- 10 Doctor's degree
- 90 Other
- 99 Refusal
- 100 Special Code

End of text file.

[Beginning](#)

[Main Page](#)

[End](#)



Washington, D.C.
March 23–26, 2014

[Beginning](#)

[Main Page](#)

[End](#)