

Dictionnaire de données de SAS Partie 2

Conférencier : Patrice Bourdages

2022-06-07



Ordre du jour

- A qui s'adresse cette conférence
- Objectifs de la formation
- Problème à résoudre
 - Préalable
 - Structure de répertoires (exemple)
 - **Étape 1 – Extraction des répertoires**
 - Étape 2 – Nettoyage et préparation
 - Étape 3 – Décompte des répertoires
 - **Partie 4 – Création de « libname » incrémental**
 - Partie 5 – Extraction du SDD
 - Partie 6 – Analyse des résultats
- Références
- Conclusion

A qui s'adresse cette conférence

- Programmeur SAS
 - Analyste SAS
 - Analyste-programmeur SAS
 - Actuaire
 - Administrateur de base de données
 - Statisticien
 - Etc...
-
- Tout le monde qui utilise SAS quoi ... !!!
-
- Mais surtout à ceux qui ont vu la première présentation sur le dictionnaire de données de SAS (**SAS Data Dictionary – SDD**)

Objectifs de la formation

- Permettre à l'usager ... :
 - D'approfondir ses connaissances du dictionnaire de données SAS
 - D'acquérir des connaissances sur une utilisation spécifique du dictionnaire
 - Découvrir une vieille macro qui fait encore des merveilles.

Problème à résoudre

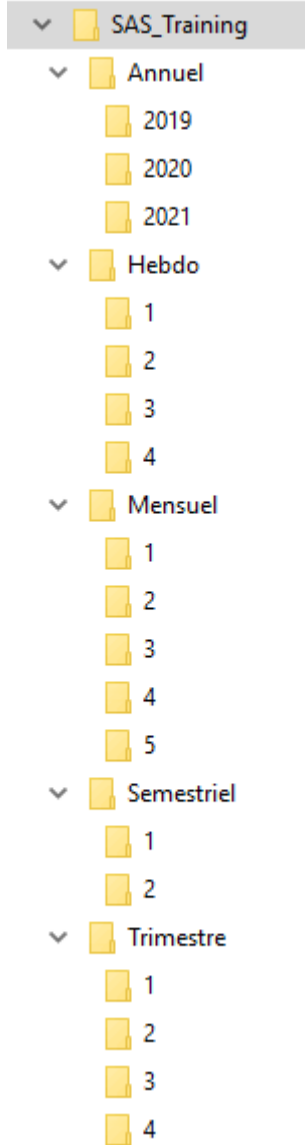
- Vous avez 532 répertoires et sous-répertoires qui contiennent des datasets de SAS
- Vous voulez utiliser le dictionnaire de données de SAS pour connaître l'existence des tables SAS, leurs structures (variables), etc.
- Vous ne voulez pas écrire 532 énoncés LIBNAME manuellement

■ **Que faire ???**

Préalable

- Macro %FindFiles de Richard A. Devenezia (voir références)
 - Télécharger et exécuter le code
 - Corriger pour adapter à votre environnement s'il y a lieu
- Travailler dans un environnement Windows

Structure de répertoires (exemple)









Étape 1 – Extraction des répertoires

- Extraire la liste des répertoires désirés


```
%findfiles(path      = h:\sasusers\bop003\SAS_Training\  
           , filespec = *  
           , out      = work.extraire_01 (where = (type = "Directory"))  
           , recurse  = Y);
```

- Utiliser la macro %FindFiles pour extraire cette liste selon vos spécifications

	 path	 name	 type	 altname	 size	 modified
1	h:\sasusers\bop003\SAS_Training\	Annuel	Directory	Annuel	0	30MAY22:10:40:24
2	h:\sasusers\bop003\SAS_Training\	Hebdo	Directory	Hebdo	0	30MAY22:09:01:48
3	h:\sasusers\bop003\SAS_Training\	Mensuel	Directory	Mensuel	0	30MAY22:09:02:16
4	h:\sasusers\bop003\SAS_Training\	Semestriel	Directory	Semestriel	0	30MAY22:09:02:28
5	h:\sasusers\bop003\SAS_Training\	Trimestre	Directory	Trimestre	0	30MAY22:09:04:18

Etape 2 – Nettoyage et préparation

```
data work.repertoires (keep = path);  
    set work.extrait_01;  
    path = compress(path) || compress(name);  
run;
```

	 path
1	h:\sasusers\bop003\SAS_Training\Annuel
2	h:\sasusers\bop003\SAS_Training\Hebdo
3	h:\sasusers\bop003\SAS_Training\Mensuel
4	h:\sasusers\bop003\SAS_Training\Semestriel
5	h:\sasusers\bop003\SAS_Training\Trimestre
6	h:\sasusers\bop003\SAS_Training\Annuel\2019
7	h:\sasusers\bop003\SAS_Training\Annuel\2020
8	h:\sasusers\bop003\SAS_Training\Annuel\2021
9	h:\sasusers\bop003\SAS_Training\Hebdo\1
10	h:\sasusers\bop003\SAS_Training\Hebdo\2

Étape 3 – Décompte des répertoires

```
%* Capture le nombre de répertoire à traiter *;  
%let _NombreRepertoire = &sysnobs.;  
%put _NombreRepertoire = &_NombreRepertoire.;
```

- On utilise la variable macro SYSNOBS pour déterminer le nombre d'observations dans la dernière table SAS traitée

```
23      data work.repertoires (keep = path);  
24          set work.extrait_01;  
25          path = compress(path) || compress(name);  
26      run;  
  
NOTE: There were 23 observations read from the data set WORK.EXTRAIT_01.  
NOTE: The data set WORK.REPERTOIRES has 23 observations and 1 variables.  
NOTE: Compressing data set WORK.REPERTOIRES increased size by 100.00 percent.  
       Compressed is 2 pages; un-compressed would require 1 pages.  
NOTE: DATA statement used (Total process time):  
       real time          0.00 seconds  
       cpu time           0.00 seconds  
  
27  
28      %* Capture le nombre de répertoire à traiter *;  
29      %let _NombreRepertoire = &sysnobs.;  
30      %put _NombreRepertoire = &_NombreRepertoire.;
```

`_NombreRepertoire = 23`

Partie 4 – Création de « libname » incrémental

```
%macro creer_libname;
  /* Pour chacun des répertoires, assigner un libname bidon (incrémental) */
  %do i = 1 %to &_NombreRepertoire.;
    /* Extraire le "path" du répertoire à traiter */
    data _null_;
      set repertoires (obs = &i.);
      call symput("m_Chemin", compress(path));
    run;
    /* On crée un libname dont les premiers caractères sont L_ */
    libname L_&i. "&m_Chemin.";
  %end;
%mend;
%creer_libname;
```

Voici un des résultats:

```
NOTE: Libref L_22 was successfully assigned as follows:
      Engine:          V9
      Physical Name: h:\sasusers\bop003\SAS_Training\Trimestre\3
```

Partie 4 (suite) – Constat des librairies créées

▼ Libraries

>  L_1

>  L_2

>  L_3

>  L_4

>  L_5

 L_6

>  L_7

>  L_8

>  L_9

>  L_10

>  L_11

>  L_12

>  L_13

Première partie des librairies SAS créées
dynamiquement par notre routine

Partie 5 – Extraction du SDD

```
proc sql;
    create table work.tables as
        select *
        from dictionary.tables
        where upcase(substr(libname, 1, 2)) = "L_";

    create table work.colonnes as
        select *
        from dictionary.columns
        where upcase(substr(libname, 1, 2)) = "L_";

    create table work.libnames as
        select *
        from dictionary.libnames
        where upcase(substr(libname, 1, 2)) = "L_" and
              sysname = "Filename";
quit;
```

Partie 6 – Analyse des résultats

< Where | Query Builder Tasks ▾

TABLES

COLONNES

LIBNAMES

	libname	memname	memtype	dbms_memtype	memlabel	typemem	crdate	modate	
1	L_23	TABLE1	DATA			DATA	30MAY22:10:40:24	30MAY22:10:40:24	
2	L_22	TABLE1	DATA			DATA	30MAY22:10:40:24	30MAY22:10:40:24	
3	L_21	TABLE1	DATA			DATA	30MAY22:10:40:24	30MAY22:10:40:24	
4	L_20	TABLE1	DATA			DATA	30MAY22:10:40:24	30MAY22:10:40:24	
5	L_19	TABLE1	DATA			DATA	30MAY22:10:40:24	30MAY22:10:40:24	
6	L_18	TABLE1	DATA			DATA	30MAY22:10:40:24	30MAY22:10:40:24	
7	L_17	TABLE1	DATA			DATA	30MAY22:10:40:24	30MAY22:10:40:24	
8	L_16	TABLE1	DATA			DATA	30MAY22:10:40:24	30MAY22:10:40:24	
9	L_15	TABLE1	DATA			DATA	30MAY22:10:40:24	30MAY22:10:40:24	
10	L_14	TABLE1	DATA			DATA	30MAY22:10:40:24	30MAY22:10:40:24	
11	L_13	TABLE1	DATA			DATA	30MAY22:10:40:24	30MAY22:10:40:24	
12	L_12	TABLE1	DATA			DATA	30MAY22:10:40:24	30MAY22:10:40:24	
13	L_11	TABLE1	DATA			DATA	30MAY22:10:40:24	30MAY22:10:40:24	
14	L_10	TABLE1	DATA			DATA	30MAY22:10:40:24	30MAY22:10:40:24	
15	L_9	TABLE1	DATA			DATA	30MAY22:10:40:24	30MAY22:10:40:24	
16	L_1	TABLE1	DATA			DATA	30MAY22:10:40:24	30MAY22:10:40:24	

Partie 6 (suite) – Analyse des résultats

< Where | Query Builder | Tasks ▾

TABLES

COLONNES

LIBNAMES

	libname	memname	memtype	name	type	length	npos	varnum	label	format	informat	id
1	L_23	TABLE1	DATA	x	num	8	0	1				
2	L_23	TABLE1	DATA	y	num	8	8	2				
3	L_23	TABLE1	DATA	z	num	8	16	3				
4	L_22	TABLE1	DATA	x	num	8	0	1				
5	L_22	TABLE1	DATA	y	num	8	8	2				
6	L_22	TABLE1	DATA	z	num	8	16	3				
7	L_21	TABLE1	DATA	x	num	8	0	1				
8	L_21	TABLE1	DATA	y	num	8	8	2				
9	L_21	TABLE1	DATA	z	num	8	16	3				
10	L_20	TABLE1	DATA	x	num	8	0	1				
11	L_20	TABLE1	DATA	y	num	8	8	2				
12	L_20	TABLE1	DATA	z	num	8	16	3				
13	L_19	TABLE1	DATA	x	num	8	0	1				
14	L_19	TABLE1	DATA	y	num	8	8	2				
15	L_19	TABLE1	DATA	z	num	8	16	3				
16	L_18	TABLE1	DATA	x	num	8	0	1				

Partie 6 (suite) – Analyse des résultats

< Where | Query Builder Tasks ▾

TABLES

COLONNES

LIBNAMES

	libname	engine	path	level	fileformat	readonly	sequential	
1	L_23	V9	h:\sasusers\bop003\SAS_Training\Trimestre\4	0	7	no	no	Host depe
2	L_22	V9	h:\sasusers\bop003\SAS_Training\Trimestre\3	0	7	no	no	Host depe
3	L_21	V9	h:\sasusers\bop003\SAS_Training\Trimestre\2	0	7	no	no	Host depe
4	L_20	V9	h:\sasusers\bop003\SAS_Training\Trimestre\1	0	7	no	no	Host depe
5	L_19	V9	h:\sasusers\bop003\SAS_Training\Semestriel\2	0	7	no	no	Host depe
6	L_18	V9	h:\sasusers\bop003\SAS_Training\Semestriel\1	0	7	no	no	Host depe
7	L_17	V9	h:\sasusers\bop003\SAS_Training\Mensuel\5	0	7	no	no	Host depe
8	L_16	V9	h:\sasusers\bop003\SAS_Training\Mensuel\4	0	7	no	no	Host depe
9	L_15	V9	h:\sasusers\bop003\SAS_Training\Mensuel\3	0	7	no	no	Host depe
10	L_14	V9	h:\sasusers\bop003\SAS_Training\Mensuel\2	0	7	no	no	Host depe
11	L_13	V9	h:\sasusers\bop003\SAS_Training\Mensuel\1	0	7	no	no	Host depe
12	L_12	V9	h:\sasusers\bop003\SAS_Training\Hebdo\4	0	7	no	no	Host depe
13	L_11	V9	h:\sasusers\bop003\SAS_Training\Hebdo\3	0	7	no	no	Host depe
14	L_10	V9	h:\sasusers\bop003\SAS_Training\Hebdo\2	0	7	no	no	Host depe
15	L_9	V9	h:\sasusers\bop003\SAS_Training\Hebdo\1	0	7	no	no	Host depe
16	L_8	V9	h:\sasusers\bop003\SAS_Training\Annuel\2021	0	7	no	no	Host depe

Conclusion

- Structure complexe de répertoires
 - 2, 100, 1000 et plus, ce n'est plus un problème pour vous
 - P.S. : je l'ai fait sur 27 000 répertoires en 7 minutes
- **Utile pour vous?** J'en suis convaincu à un moment donné.
Peut-être plus vite que vous ne le pensez.

Références

- **Richard A Devenezia macro %FindFiles**
 - <https://www.devenezia.com/downloads/sas/macros/index.php?m=findfiles>
- **Accessing SAS System Information by Using DICTIONARY Tables**
 - <http://support.sas.com/documentation/cdl/en/sqlproc/62086/HTML/default/viewer.htm#a001385596.htm>
- **Summary of SAS Dictionary Tables and Views**
 - DictionaryTables.pdf
 - DictionaryTablesRefCard.pdf
 - Seront déposés avec la présentation



Période de question