



## La proc TABULATE, sa vie, son œuvre

Dotée d'une syntaxe joueuse, dont la ponctuation est un vrai casse-tête chinois, on ne peut pas dire que la procédure TABULATE génère spontanément la sympathie. Cela dit, dans tous les casse-têtes, on trouve une logique implacable – le but de ces pages est de décortiquer celle de TABULATE. Petit itinéraire de vacances au pays des tableaux sous SAS.

### Les préparatifs : syntaxe de base

```
PROC TABULATE DATA = tableLue ;
  CLASS variables1 ;
  VAR variables2 ;
  TABLE ... ;
RUN ;
```



Comme souvent dans SAS, tout commence avec une option DATA= pour indiquer quelle est la table sur laquelle on souhaite travailler. Puis viennent deux instructions, VAR et CLASS, qui permettent de planter le décor, et d'indiquer les rôles des différentes variables :

- Toutes les variables de CLASS servent à définir des groupes d'observations sur lesquels les statistiques du tableau seront calculées. Ce seront les variables qui définiront les lignes et les colonnes du tableau. A ce stade, SAS ne distingue pas encore de notion de lignes et de colonnes : on peut donc mettre tous les noms de variables sans se soucier de leur ordre.
- Toutes les variables de VAR doivent être numériques ; ce sont sur ses variables que les statistiques du tableau seront calculées.

Il n'est pas obligatoire d'avoir défini les deux instructions : pour un tableau ne contenant que des tableaux et des pourcentages, par exemple, il n'y aura pas d'instruction VAR. En revanche, on ne peut pas faire fonctionner TABULATE sans avoir au moins une de ces instructions.

Dans l'instruction TABLE, on définira plus précisément l'architecture du tableau. La logique de cette instruction est détaillée au paragraphe suivant.

### Navigation au GPS : construire des tableaux croisés



C'est au niveau de l'instruction TABLE que se joue la mise en forme du tableau. Elle est extrêmement sensible à la ponctuation :

- la virgule marque la limite entre les lignes et les colonnes ;
- l'espace indique une juxtaposition de deux éléments dans la même dimension (ligne ou colonne) ;

- l'étoile signifie que l'on imbrique deux éléments dans une même dimension.

La syntaxe de base est souvent :

```
TABLE (lignes ) ,
      (colonnes) *
      (cellules)
;
```

Entre les parenthèses de la dimension ligne, on trouve des noms de variables citées dans CLASS, ainsi que le mot-clé ALL qui indique un récapitulatif. Il en va de même entre les parenthèses de la dimension colonne.

Pour les cellules, on retrouve soit une statistique comme une fréquence ou un pourcentage, soit une variable de calcul (déjà citée dans VAR) suivie d'une étoile et de sa statistique.

Principales statistiques disponibles dans la procédure TABULATE			
<b>N</b>	Comptage d'observations	<b>MEAN</b>	Moyenne arithmétique
<b>PCTN</b>	Pourcentage du nombre de total d'observations	<b>SUM</b>	Total
<b>ROWPCTN</b>	Pourcentage du nombre d'observations dans la ligne	<b>MIN et MAX</b>	Minimum et maximum
<b>COLPCTN</b>	Pourcentage du nombre d'observations dans la colonne	<b>P1, P5, P10, P25 ou Q1, P50 ou MEDIAN, P75 ou Q3, P90, P95 et P99</b>	Centiles à x%, dont quartiles et médiane
<b>NMISS</b>	Nombre de valeurs non renseignées	<b>STD et VAR</b>	Ecart-type et variance
<b>PCTSUM</b>	Pourcentage du total d'une variable	<b>SKEWNESS et KURTOSIS</b>	Coefficients d'asymétrie et d'aplatissement
<b>ROWPCTSUM</b>	Pourcentage du total-ligne d'une variable	<b>LCLM et UCLM</b>	Intervalle de confiance (bornes inférieure et supérieure) à 95%
<b>COLPCTSUM</b>	Pourcentage du total-colonne d'une variable	<b>CV</b>	Coefficient de variation

Quelques exemples d'instructions TABLE vont vous aider à mieux comprendre l'intrication des différents composants. Voici les données de départ (table SASHELP.PRDSALE)

COUNTRY	REGION	PRODUCT	QUARTER	ACTUAL	PREDICT
CANADA	EAST	SOFA	1	\$925.00	\$850.00
CANADA	EAST	SOFA	1	\$999.00	\$297.00
CANADA	EAST	SOFA	1	\$608.00	\$846.00
CANADA	EAST	SOFA	2	\$642.00	\$533.00
CANADA	EAST	SOFA	2	\$656.00	\$646.00
CANADA	EAST	SOFA	2	\$948.00	\$486.00
CANADA	EAST	SOFA	3	\$612.00	\$717.00
CANADA	EAST	SOFA	3	\$114.00	\$564.00
CANADA	EAST	SOFA	3	\$685.00	\$230.00
CANADA	EAST	SOFA	4	\$657.00	\$494.00

```
PROC TABULATE DATA = sashelp.prdsale ;
  CLASS product country ;
  TABLE (country),
         (product)*
         (N) ;
RUN ;
```

	Product	
	CHAIR	SOFA
	N	N
Country		
CANADA	96	96
GERMANY	96	96
U.S.A.	96	96

On a ici la version « simple » du tableau : une variable en ligne, une en colonne, et des comptages dans les cellules.

```
CLASS product country ;
VAR actual ;
TABLE (country),
      (product)*
      (actual * MEAN) ;
```

	Product	
	CHAIR	SOFA
	Actual Sales	Actual Sales
	Mean	Mean
Country		
CANADA	523.32	522.24
GERMANY	490.68	573.54
U.S.A.	530.58	452.01

Seconde version, avec une variable de calcul (ACTUAL) dont on calcule la moyenne.

```
TABLE (country ALL),
      (product)*
      (actual * MEAN) ;
```

	Product	
	CHAIR	SOFA
	Actual Sales	Actual Sales
	Mean	Mean
Country		
CANADA	523.32	522.24
GERMANY	490.68	573.54
U.S.A.	530.58	452.01
All	514.86	515.93

Variante du tableau précédent avec en sus une ligne récapitulative (ALL). A noter que cette ligne supplémentaire n'est pas un *total*, mais bien un *récapitulatif* de la variable en ligne : ici, on aura des moyennes générales, et pas une somme des moyennes, en bas de chaque colonne. L'espace qui sépare le ALL de COUNTRY indique que le détail par pays précède la ligne récapitulative. On aurait écrit ALL COUNTRY pour avoir le récapitulatif en première ligne du tableau.

```
CLASS product country ;
VAR actual predict ;
TABLE (country),
      (product)*
      ((actual predict)
       * MEAN) ;
```

On notera dans cet exemple la factorisation dans la dimension cellule : là où on écrirait ACTUAL\*MEAN PREDICT\*MEAN, on peut raccourcir avec une mise en facteur commun et une paire de parenthèses. La logique est toute mathématique.

	Product			
	CHAIR		SOFA	
	Actual Sales	Predicted Sales	Actual Sales	Predicted Sales
	Mean	Mean	Mean	Mean
Country				
CANADA	523.32	487.46	522.24	476.31
GERMANY	490.68	459.05	573.54	515.80
U.S.A.	530.58	471.30	452.01	470.92

```
CLASS product country region ;
VAR actual ;
TABLE (region * (country ALL)),
      (product) *
      (actual * MEAN) ;
```

On rajoute, dans la dimension ligne, une seconde variable de classification, REGION. Avec une étoile qui la sépare de COUNTRY et ALL, on va avoir le détail par pays et le récapitulatif tous pays confondus pour chaque valeur de REGION.

		Product	
		CHAIR	SOFA
		Actual Sales	Actual Sales
		Mean	Mean
Region	Country		
EAST	CANADA	525.00	502.58
	GERMANY	484.94	546.13
	U.S.A.	570.38	463.81
	All	526.77	504.17
WEST	Country		
	CANADA	521.65	541.90
	GERMANY	496.42	600.96
	U.S.A.	490.79	440.21
	All	502.95	527.69

```
PROC TABULATE DATA = sashelp.prdsale ;
CLASS product country region ;
VAR actual ;
TABLE (region country ALL),
      (product)*
      (actual * MEAN) ;
RUN ;
```

Si, par rapport au tableau précédent, on remplace l'étoile dans la dimension ligne par un espace, alors on a d'abord le détail par REGION, puis par COUNTRY, puis le récapitulatif. Avec l'étoile, le détail par COUNTRY était imbriqué dans celui par REGION.

	Product	
	CHAIR	SOFA
	Actual Sales	Actual Sales
	Mean	Mean
<b>Region</b>		
<b>EAST</b>	526.77	504.17
<b>WEST</b>	502.95	527.69
<b>Country</b>		
<b>CANADA</b>	523.32	522.24
<b>GERMANY</b>	490.68	573.54
<b>U.S.A.</b>	530.58	452.01
<b>All</b>	514.86	515.93

```
TABLE ((region ALL)
      (country ALL)),
      (product)*
      (actual * MEAN) ;
```

Ici, à cause des espaces entre les éléments de la dimension ligne, les deux lignes ALL de récapitulatifs sont identiques.

	Product	
	CHAIR	SOFA
	Actual Sales	Actual Sales
	Mean	Mean
<b>Region</b>		
<b>EAST</b>	526.77	504.17
<b>WEST</b>	502.95	527.69
<b>All</b>	514.86	515.93
<b>Country</b>		
<b>CANADA</b>	523.32	522.24
<b>GERMANY</b>	490.68	573.54
<b>U.S.A.</b>	530.58	452.01
<b>All</b>	514.86	515.93

```
TABLE ((region ALL)*
      (country ALL)),
      (product)*
      (actual * MEAN) ;
```

Même programme, hormis une étoile à la place de l'espace central dans la dimension ligne. Là, les deux ALL ont des portées différentes. Le second récapitule les valeurs de COUNTRY à l'intérieur de l'imbrication REGION. Le premier correspond à un récapitulatif de toutes les valeurs de REGION.

		Product	
		CHAIR	SOFA
		Actual Sales	Actual Sales
		Mean	Mean
Region	Country		
EAST	CANADA	525.00	502.58
	GERMANY	484.94	546.13
	U.S.A.	570.38	463.81
	All	526.77	504.17
WEST	Country		
	CANADA	521.65	541.90
	GERMANY	496.42	600.96
	U.S.A.	490.79	440.21
	All	502.95	527.69
All	Country		
	CANADA	523.32	522.24
	GERMANY	490.68	573.54
	U.S.A.	530.58	452.01
	All	514.86	515.93

## Itinéraire bis : quelques mises en forme plus complexes

Maintenant que vous voyez bien la logique des espaces, des virgules et des étoiles, déstructurons un petit peu par rapport au schéma initial (lignes),(colonnes)\*(cellules).

On peut en effet déplacer les éléments constitutifs des cellules à l'intérieur des dimensions lignes et colonnes, afin de changer leur emplacement dans le tableau. Cependant, il faut qu'au final l'étoile se retrouve entre les variables de calcul (celles de VAR) et leur statistique.



```
PROC TABULATE
  DATA = sashelp.prdsale ;
  CLASS country product ;
  VAR actual predict ;
  TABLE (country),
         (MEAN*product) *
         (actual) ;
RUN ;
```

Ici, on « remonte » la statistique MEAN de manière à ne la voir apparaître qu'une fois, avant la variable en colonne PRODUCT.

		Mean	
		Product	
		CHAIR	SOFA
		Actual Sales	Actual Sales
Country			
CANADA	523.32	522.24	
GERMANY	490.68	573.54	
U.S.A.	530.58	452.01	

```
TABLE (country *
      (actual predict)),
      ((actual * MEAN) *
       product) ;
```

Ici, on remonte carrément le couple variable de calcul / statistique, de manière à éviter au maximum les répétitions. On mesure alors, pour SAS, tout l'intérêt d'avoir défini au préalable les rôles des variables avec VAR et CLASS, pour comprendre l'instruction TABLE correctement.

	Actual Sales	
	Mean	
	Product	
	CHAIR	SOFA
Country		
CANADA	523.32	522.24
GERMANY	490.68	573.54
U.S.A.	530.58	452.01

```
TABLE (country *
      (actual predict)),
      (product) *
      (MEAN) ;
```

Avec deux variables de calcul (ou plus), il peut être intéressant de dupliquer des lignes et non des colonnes. Il suffit alors de déplacer les variables en question (ici, ACTUAL et PREDICT) dans la dimension ligne, en les imbriquant dans les variables de classification avec une étoile.

	Country	Product	
		CHAIR	SOFA
		Mean	Mean
CANADA	Actual Sales	523.32	522.24
	Predicted Sales	487.46	476.31
GERMANY	Actual Sales	490.68	573.54
	Predicted Sales	459.05	515.80
U.S.A.	Actual Sales	530.58	452.01
	Predicted Sales	471.30	470.92

```
TABLE (country *
      (actual predict)),
      (MEAN) *
      (product) ;
```

Une variante du tableau précédent pour éviter les répétitions consiste, là encore, à remonter la statistique au-dessus de la variable de classification de la dimension colonne.

	Country	Mean	
		Product	
		CHAIR	SOFA
CANADA	Actual Sales	523.32	522.24
	Predicted Sales	487.46	476.31
GERMANY	Actual Sales	490.68	573.54
	Predicted Sales	459.05	515.80
U.S.A.	Actual Sales	530.58	452.01
	Predicted Sales	471.30	470.92

Comment jouer sur l'ordre d'affichage des valeurs dans les variables de classification ? Par défaut, c'est l'ordre des valeurs non formatées qui sera suivi. Le fait d'appliquer seulement un format à une variable n'influe donc en rien sur l'ordre d'affichage, comme le prouve l'exemple suivant.

```
PROC FORMAT ;
    VALUE $pays
        "CANADA" = "Canada"
        "U.S.A." = "Etats-Unis"
        "GERMANY" = "Allemagne"
    ;
RUN ;
PROC TABULATE DATA = sashelp.prdsale ;
    CLASS country product ;
    FORMAT country $pays. ;
    VAR actual ;
    TABLE (ALL country),
        ((MEAN * actual) *
        product) ;
RUN ;
```

	Mean	
	Actual Sales	
	Product	
	CHAIR	SOFA
All	514.86	515.93
Country		
Canada	523.32	522.24
Allemagne	490.68	573.54
Etats-Unis	530.58	452.01

L'option ORDER= (au niveau de l'instruction PROC TABULATE) permet de jouer réellement sur l'ordre d'affichage des valeurs des variables de CLASS. On peut donner à cette option trois valeurs : INTERNAL (valeurs telles qu'elles sont stockées dans la table, par défaut), FORMATTED (ordonne selon les valeurs formatées) et DATA (valeurs dans l'ordre où elles apparaissent dans la table, sans considérations d'ordre alphabétique).

```
PROC TABULATE DATA = sashelp.prdsale ORDER = FORMATTED ;
    CLASS country product ;
    FORMAT country $pays. ;
    VAR actual ;
    TABLE (ALL country),((MEAN * actual) * product) ;
RUN ;
```

<table border="1"> <thead> <tr> <th rowspan="4"></th> <th colspan="2">Mean</th> </tr> <tr> <th colspan="2">Actual Sales</th> </tr> <tr> <th colspan="2">Product</th> </tr> <tr> <th>CHAIR</th> <th>SOFA</th> </tr> </thead> <tbody> <tr> <th>All</th> <td>514.86</td> <td>515.93</td> </tr> <tr> <th>Country</th> <td></td> <td></td> </tr> <tr> <th>Canada</th> <td>523.32</td> <td>522.24</td> </tr> <tr> <th>Allemagne</th> <td>490.68</td> <td>573.54</td> </tr> <tr> <th>Etats-Unis</th> <td>530.58</td> <td>452.01</td> </tr> </tbody> </table> <p>avec ORDER=INTERNAL (par défaut)</p>		Mean		Actual Sales		Product		CHAIR	SOFA	All	514.86	515.93	Country			Canada	523.32	522.24	Allemagne	490.68	573.54	Etats-Unis	530.58	452.01	<table border="1"> <thead> <tr> <th rowspan="4"></th> <th colspan="2">Mean</th> </tr> <tr> <th colspan="2">Actual Sales</th> </tr> <tr> <th colspan="2">Product</th> </tr> <tr> <th>CHAIR</th> <th>SOFA</th> </tr> </thead> <tbody> <tr> <th>All</th> <td>514.86</td> <td>515.93</td> </tr> <tr> <th>Country</th> <td></td> <td></td> </tr> <tr> <th>Allemagne</th> <td>490.68</td> <td>573.54</td> </tr> <tr> <th>Canada</th> <td>523.32</td> <td>522.24</td> </tr> <tr> <th>Etats-Unis</th> <td>530.58</td> <td>452.01</td> </tr> </tbody> </table> <p>avec ORDER = FORMATTED et le format \$pays. (exemple ci-dessus)</p>		Mean		Actual Sales		Product		CHAIR	SOFA	All	514.86	515.93	Country			Allemagne	490.68	573.54	Canada	523.32	522.24	Etats-Unis	530.58	452.01	<table border="1"> <thead> <tr> <th rowspan="4"></th> <th colspan="2">Mean</th> </tr> <tr> <th colspan="2">Actual Sales</th> </tr> <tr> <th colspan="2">Product</th> </tr> <tr> <th>SOFA</th> <th>CHAIR</th> </tr> </thead> <tbody> <tr> <th>All</th> <td>515.93</td> <td>514.86</td> </tr> <tr> <th>Country</th> <td></td> <td></td> </tr> <tr> <th>Canada</th> <td>522.24</td> <td>523.32</td> </tr> <tr> <th>Etats-Unis</th> <td>452.01</td> <td>530.58</td> </tr> <tr> <th>Allemagne</th> <td>573.54</td> <td>490.68</td> </tr> </tbody> </table> <p>avec ORDER = DATA et une proc SORT préalable selon ACTUAL</p>		Mean		Actual Sales		Product		SOFA	CHAIR	All	515.93	514.86	Country			Canada	522.24	523.32	Etats-Unis	452.01	530.58	Allemagne	573.54	490.68
		Mean																																																																								
		Actual Sales																																																																								
		Product																																																																								
	CHAIR	SOFA																																																																								
All	514.86	515.93																																																																								
Country																																																																										
Canada	523.32	522.24																																																																								
Allemagne	490.68	573.54																																																																								
Etats-Unis	530.58	452.01																																																																								
	Mean																																																																									
	Actual Sales																																																																									
	Product																																																																									
	CHAIR	SOFA																																																																								
All	514.86	515.93																																																																								
Country																																																																										
Allemagne	490.68	573.54																																																																								
Canada	523.32	522.24																																																																								
Etats-Unis	530.58	452.01																																																																								
	Mean																																																																									
	Actual Sales																																																																									
	Product																																																																									
	SOFA	CHAIR																																																																								
All	515.93	514.86																																																																								
Country																																																																										
Canada	522.24	523.32																																																																								
Etats-Unis	452.01	530.58																																																																								
Allemagne	573.54	490.68																																																																								



## Route ou autoroute ? Blocs BY et dimension page

On peut automatiquement produire toute une série de tableaux avec la procédure TABULATE, sans recourir aux macros ni aux copier/coller. Pour cela, deux solutions : l'instruction BY, qui nécessite un tri préalable des données, et la complexification (encore !) de l'instruction TABLE avec l'ajout d'une

dimension page, qui, elle, ne nécessite pas de tri préalable. Si la variable de la dimension page doit apparaître impérativement dans CLASS, celle d'un BY n'a pas besoin d'y figurer. La différence entre les deux solutions ne se mesure qu'en termes de disposition du titre indiquant de quelle sous-population on consulte le tableau.

```
PROC SORT DATA = sashelp.prdsale OUT = work.ventes ;
  BY region ;
RUN ;
/* solution avec un BY : tri préalable obligatoire */
PROC TABULATE DATA = work.ventes ;
  CLASS country product ;
  FORMAT country $pays. ;
  BY region ;
  VAR actual ;
  TABLE (country), (MEAN*actual*product) ;
RUN ;
/* solution avec une dimension PAGE */
PROC TABULATE DATA = sashelp.prdsale ;
  CLASS region country product ;
  FORMAT country $pays. ;
  VAR actual ;
  TABLE region, (country), (MEAN*actual*product) ;
RUN ;
```

Region=EAST		
	Mean	
	Actual Sales	
	Product	
	CHAIR	SOFA
Country		
Canada	525.00	502.58
Allemagne	484.94	546.13
Etats-Unis	570.38	463.81

  

Region=WEST		
	Mean	
	Actual Sales	
	Product	
	CHAIR	SOFA
Country		
Canada	521.65	541.90
Allemagne	496.42	600.96
Etats-Unis	490.79	440.21

  

Region EAST		
	Mean	
	Actual Sales	
	Product	
	CHAIR	SOFA
Country		
Canada	525.00	502.58
Allemagne	484.94	546.13
Etats-Unis	570.38	463.81

  

Region WEST		
	Mean	
	Actual Sales	
	Product	
	CHAIR	SOFA
Country		
Canada	521.65	541.90
Allemagne	496.42	600.96
Etats-Unis	490.79	440.21

avec une instruction BY

avec une dimension page

## Echangeur autoroutier : gestion des croisements

Que se passe-t-il si les croisements de deux variables de classes dans une même dimension (ligne ou colonne) n'existent pas tous ?

Par défaut, la ligne ou la colonne correspondante n'est tout simplement pas générée.



```
PROC TABULATE DATA = sashelp.prdsale ;
  WHERE NOT (country = "CANADA"
            AND year = 1994) ;
  CLASS year country product ;
  VAR actual ;
  TABLE (year*country), (actual*MEAN*product) ;
RUN ;
```

Pour illustrer le comportement par défaut de la procédure TABULATE, on élimine du traitement, par un WHERE, le croisement de l'année 1994 et du Canada. La ligne correspondante n'est pas affichée dans le tableau.

		Actual Sales	
		Mean	
		Product	
		CHAIR	SOFA
Year	Country		
1993	CANADA	509.63	556.67
	GERMANY	482.19	640.17
	U.S.A.	529.44	458.00
1994	GERMANY	499.17	506.92
	U.S.A.	531.73	446.02

L'option PRINTMISS à la fin de l'instruction TABLE, après un slash /, permet de faire figurer tous les croisements dans le tableau, même ceux pour lesquels il n'y a pas d'observations valides à traiter.

```
TABLE (year*country),
      (actual*MEAN*product) /
      PRINTMISS ;
```

		Actual Sales	
		Mean	
		Product	
		CHAIR	SOFA
Year	Country		
1993	CANADA	509.63	556.67
	GERMANY	482.19	640.17
	U.S.A.	529.44	458.00
1994	CANADA	.	.
	GERMANY	499.17	506.92
	U.S.A.	531.73	446.02

Une autre option permet de faire figurer dans un tableau des valeurs qui ne font pas l'objet d'observations dans la table lue : PRELOADFMT, qui s'ajoute à l'instruction CLASS après un slash. En présence de cette option, la procédure TABULATE se base sur les formats utilisateurs associés aux variables de classification pour connaître les valeurs, et non sur le

contenu de la table SAS. Cette option doit être associée à PRINTMISS pour pouvoir en constater les effets dans le tableau.

```
PROC FORMAT ;
  VALUE $pays
    "CANADA" = "Canada"
    "U.S.A." = "Etats-Unis"
    "GERMANY" = "Allemagne"
    "SPAIN" = "Espagne"
;
RUN ;
PROC TABULATE DATA = sashelp.prdsale ;
  CLASS country product
    / PRELOADFMT ;
  FORMAT country $pays. ;
  VAR actual ;
  TABLE (country),
    (actual*MEAN*product)
    / PRINTMISS ;
RUN ;
```

Dans cet exemple, on ajoute un pays au format, qui n'existe pas dans les données.

Grâce à PRELOADFMT et PRINTMISS, il est néanmoins présent dans le tableau.

	Actual Sales	
	Mean	
	Product	
	CHAIR	SOFA
Country		
Canada	523.32	522.24
Allemagne	490.68	573.54
Espagne	.	.
Etats-Unis	530.58	452.01



## Dos d'âne : modification de labels et de formats

Jusqu'à présent, hormis l'application d'un format à une variable de classification, nous avons accepté tels quels les choix de mise en forme de la procédure TABULATE. Cela ne peut plus durer ! Prenons en main nos labels, en-têtes, et formats de cellules !

### Application d'un format aux cellules

Il est possible de changer l'affichage par défaut des nombres dans les cellules calculées du tableau ; pour cela, il faut ajouter derrière la statistique concernée une étoile, puis F= et le nom du format. Cela charge encore un peu la définition de l'instruction TABLE, mais le résultat est nettement plus élégant, surtout avec l'utilisation de formats comme NUMX (séparateur de décimales = virgule) et NLNUM en version 9 (séparateur de milliers = espace et séparateur de décimales = virgule, si on a spécifié OPTION LOCALE=FRENCH ; au préalable).

```
OPTION LOCALE = FRENCH ;
PROC TABULATE DATA = sashelp.prdsale ;
  CLASS country ;
  VAR actual ;
  TABLE (country),
    (actual*
      (SUM*F=NLNUM12.1
       MEAN*F=NUMX12.2)) ;
RUN ;
```

	Actual Sales	
	Sum	Mean
Country		
CANADA	100 374,0	522,78
GERMANY	102 165,0	532,11
U.S.A.	94 329,0	491,30

## Changements de libellés de cases

Pour changer l'intitulé d'une case contenant la valeur d'une variable de classification, il faut avoir recours à un format. En revanche, pour changer les autres intitulés (nom de statistique, labels de variables), il suffit de rajouter à côté de l'élément concerné, dans l'instruction TABLE, un signe = suivi, entre guillemets, du nouveau texte à afficher.

Pour supprimer une case, il suffit de proposer comme nouveau texte un espace entre guillemets.

```
PROC TABULATE DATA = sashelp.prdsale ;
  WHERE product IN ("SOFA", "CHAIR") ;
  CLASS country year ;
  VAR actual ;
  TABLE (country = "" ALL = "Total"),
         (actual = "Ventes réelles (total annuel)" *
          SUM = "" * F = NLNUM10.2 *
          year = "") ;
RUN ;
```

Par rapport à la version brute du tableau, on supprime les cases correspondant à SUM et aux labels des variables YEAR et COUNTRY. On modifie le libellé des cases correspondant au label de ACTUAL et au mot-clé ALL.

Ces modifications de libellés sont totalement compatibles avec des changements de formats des cellules. Ici, on leur applique le format NLNUM avec 2 décimales.

	Ventes réelles (total annuel)	
	1993	1994
<b>CANADA</b>	51 182,00	49 192,00
<b>GERMANY</b>	53 873,00	48 292,00
<b>U.S.A.</b>	47 397,00	46 932,00
<b>Total</b>	152 452,00	144 416,00

## La case supérieure gauche

Pour le moment, la case supérieure gauche du tableau a été laissée en jachère. On contrôle son contenu avec l'option BOX, qui est à préciser à la fin de l'instruction TABLE, après un slash, comme l'option PRINTMISS déjà évoquée.

Derrière BOX, comme pour les autres éléments de l'instruction TABLE, on ajoute un signe = et, entre guillemets, le texte de remplacement. On peut également, après le signe =, donner sans guillemets le nom d'une variable : c'est alors son label qui sera affiché dans la case supérieure gauche. Enfin, on peut utiliser le mot-clé \_PAGE\_ qui correspondra à la valeur courante de la variable de page (voir plus haut).

```
TABLE (country = "" ALL = "Total"),
      (actual = "Ventes réelles ($US)" *
       SUM = "" * F = NLNUM10.2)
      / BOX = country ;
```

On intègre facilement, le label de la variable en ligne COUNTRY, non pas comme une 1ère ligne dans le tableau (éditée par défaut, mais supprimée par "="), mais dans la case supérieure gauche avec BOX.

Country	Ventes réelles (\$US)
<b>CANADA</b>	100 374,00
<b>GERMANY</b>	102 165,00
<b>U.S.A.</b>	94 329,00
<b>Total</b>	296 868,00

```

PROC TABULATE DATA = sashelp.prdsale ;
  CLASS product country ;
  VAR actual predict ;
  TABLE product,
         country=" ",
         (actual predict) * SUM = " " / BOX = _PAGE_ ;
RUN ;

```

Product BED	Actual Sales	Predicted Sales
CANADA	47729.00	44215.00
GERMANY	46134.00	43796.00
U.S.A.	48174.00	49856.00

Ici, c'est la valeur de la variable de page, PRODUCT, qui est intégrée à la case supérieure gauche.

### Utilisation de formats chevauchants

Comme dans la procédure MEANS, depuis la version 8 de SAS, il est possible d'utiliser dans TABULATE des formats chevauchants (*multilabel*) sur les variables de classification. Pour en appliquer aux variables de classification, il faut *impérativement* ajouter l'option MLF après un slash dans l'instruction CLASS, comme pour l'option PRELOADFMT déjà évoquée.

```

PROC FORMAT ;
  VALUE $paysmult (MULTILABEL)
    "CANADA","U.S.A." = "-- Amérique du Nord"
    "CANADA" = "Canada"
    "U.S.A." = "Etats-Unis"
    "GERMANY" = "-- Europe"
    "GERMANY" = "Allemagne" ;
RUN ;
PROC TABULATE DATA = sashelp.prdsale ;
  CLASS country / MLF ; FORMAT country $paysmult. ;
  VAR actual ;
  TABLE (country = " " ALL = "Total"),
         (actual = "Ventes réelles ($US)" *
          SUM = " " * F = NLNUM10.2)
         / BOX = "Pays/région" ;
RUN ;

```

On regroupe les pays en deux grandes zones (Amérique du Nord et Europe), qu'on met en tête des lignes en commençant leurs libellés par des tirets. Ces groupes sont ensuite subdivisés par pays.

Le format MULTILABEL permet de faire cela sans nécessité de dupliquer la variable COUNTRY et de lui appliquer deux formats différents : sur une grosse table, on économise en temps de traitement et en mémoire requise.

Pays/région	Ventes réelles (\$US)
-- Amérique du Nord	194 703,00
-- Europe	102 165,00
Allemagne	102 165,00
Canada	100 374,00
Etats-Unis	94 329,00
Total	296 868,00

## Faire le niveau d'huile (de coude) : mise en forme pour l'ODS



Dernière coquetterie que l'on peut ajouter à la procédure TABULATE, si on prévoit de récupérer les tableaux résultats dans un document externe à SAS (Word, PDF, Excel, HTML) via l'ODS : on peut préciser toute une série d'options de mise en forme. Elles sont repérées par un mot-clé, STYLE, derrière lequel on retrouvera un signe = et, entre crochets (!) les options et leurs valeurs. On peut ainsi avoir barre sur les polices, les couleurs, les alignements, la prise en compte des espaces à gauche d'une valeur, etc.

Principales options de style disponibles		
Élément de style	Description	Valeur(s) possible(s)
FONT_WEIGHT	Épaisseur de la police de caractères utilisée (gras, normal ou maigre)	BOLD (pour gras), MEDIUM, LIGHT
FONT_STYLE	Mise en italique de la police de caractères utilisée	ITALIC, ROMAN (police droite)
FONT_SIZE	Taille de la police de caractères utilisée	Taille de 1 (le plus petit) à 7 (le plus grand), 12pt
FONT_FACE	Nom de la police de caractères utilisée	ARIAL, TAHOMA, TIMES, "TIMES NEW ROMAN", VERDANA, ...
JUST	Justification (alignement dans la ligne) du texte	LEFT, RIGHT ou CENTER
VJUST	Justification verticale (alignement dans la hauteur de la cellule du tableau) du texte	TOP, BOTTOM ou MIDDLE
BACKGROUND	Couleur de fond de la cellule	tous les codes couleurs valides dans le module SAS/GRAPH
FOREGROUND	Couleur de la police de caractères	
FLYOVER	Infobulle s'affichant quand la souris reste positionnée sur le contenu de la case (non valable vers la destination RTF)	"Texte de l'infobulle"
ASIS	Affichage des blancs à gauche dans la valeur d'un texte	ON, OFF (par défaut)
PRE_IMAGE	Image apparaissant avant le contenu d'une cellule	"Chemin d'une image à insérer"
POST_IMAGE	Image apparaissant après le contenu d'une cellule	

L'emplacement de ces instructions de mise en forme dépend étroitement de l'emplacement de la cellule à « décorer » :

- en option après un slash dans une instruction CLASS s'il s'agit d'une cellule contenant un label de variable de classification ;
- en option après un slash dans une instruction CLASSLEV avec le nom de la variable en question s'il s'agit d'une cellule contenant les valeurs d'une variable de classification ;
- en option après un slash dans une instruction VAR s'il s'agit d'une cellule contenant un label de variable de calcul ;

- en option après un slash dans une instruction KEYWORD suivie du mot-clé statistique s'il s'agit d'une cellule contenant un nom de statistique ;
- enfin, à la suite du mot-clé statistique dans une instruction TABLE, avec une étoile puis, entre crochets, S=[instructions de mise en forme].

La grande versatilité de cette syntaxe impose qu'on l'illustre par l'exemple...

```
PROC TABULATE DATA = sashelp.prdsale ;
  CLASS country ;
  CLASSLEV country / STYLE=[FONT_STYLE=ITALIC] ;
  VAR actual      / STYLE=[JUST=LEFT] ;
  TABLE (country = " "),
    (actual = "Ventes réelles" *
      SUM = " " * F=NLNUM10.2) *
      [STYLE = [foreground=WHITE
                background=BLACK]] ;
RUN ;
```

	<b>Ventes réelles</b>		<b>Ventes réelles</b>
<b>CANADA</b>	100 374,00	<i>CANADA</i>	100 374,00
<b>GERMANY</b>	102 165,00	<i>GERMANY</i>	102 165,00
<b>U.S.A.</b>	94 329,00	<i>U.S.A.</i>	94 329,00
Tableau de base		Avec les options de style du programme ci-dessus	