

# SAS GRAPH

## (Cours STA111 2006 par O.Wolber)

Le module SAS GRAPH permet de tracer des graphes dont la qualité est bien supérieure à celle des procédures précédentes (PLOT, CHART, ...), à conditions de bien maîtriser les options des procédures permettant de modifier tous les aspects des graphiques : dimensions, couleurs, titres, légendes, étiquettes, axes, échelles, polices et dimensions des caractères....

Les procédures de base (**GCHART**, **GPlot**) possèdent la même syntaxe que les procédures correspondantes en basse résolution mais sont complétées par des options spécifiques. D'autres procédures permettent de tracer :

- des fonds de cartes (**GMAP**) ;
- des surfaces ou nuages de points en trois dimensions (**G3D**).

Les graphiques ainsi créés peuvent être stockés dans des catalogues graphiques que l'on pourra consulter à nouveau en utilisant la procédure **GREPLAY**.

Tout texte ou figure géométrique complémentaires peuvent être rajoutés sur un graphe en les décrivant dans une table SAS spéciale dite d'annotation : **ANNOTATE DATA SET**.

### V. 1. Options générales de SAS GRAPH

#### Polices de caractères

Voir en annexe 1 la liste des principales polices spécifiques de SAS. Il est également possible de récupérer de nouvelles polices de caractères à partir du menu HELP de SAS (en particulier, les polices utilisées par WORD ou EXCEL).

#### Couleurs

Les possibilités de l'option color sont très nombreuses, on compte jusqu'à 256 couleurs. À partir d'une couleur de base, on peut décliner toute la gamme d'intensité, du plus clair au plus foncé<sup>1</sup>. Si l'option color n'est pas utilisée, SAS prendra la liste des couleurs par défaut.

### V. 2. Instructions générales pour la construction des graphiques

Tous les graphiques, quelles que soient les variables représentées et donc les procédures utilisées ont en commun certaines instructions.

Structure du programme :

```
goptions .... ; /*orientation du graphique, bordures, dimensions, couleurs, police de caractère*/
title1 .... ; /*édition de titres dans les graphiques*/
titlei.... ;
footnote1... ; /*notes de bas de graphique*/
footnotei... ;
legend1... ; /*légendes*/
legendi... ;
axis1...; /*options sur les axes*/
axisi...;
symbol1...; /*uniquement dans gplot*/
```

---

<sup>1</sup> Voir annexe 2.

```

symboli... ;
pattern1... ; /*instructions sur les surfaces*/
patterni... ;

data annotate ; /*pour créer les instructions qui permettent*/
...          /*d'écrire à l'intérieur du graphique*/
run ;

proc gplot data= ....gout=... ;
    plot y*x=n/<options>;
run ;

```

*gout* : envoi dans un catalogue graphique  
*y* : axe des ordonnées  
*x* : axe des abscisses  
*n* : renvoi à un numéro de symbole

### Instructions options

**goptions** liste des options ;

Il est souhaitable lorsqu'on fait des séries de graphiques, d'initialiser chacune des séries qui n'auraient pas les mêmes options par l'option **reset**.

goptions reset=all ;

On dispose : d'options d'orientation (rotate), de bordure (border), d'unité générale à utiliser par défaut (gunit), de dimensions du graphique (hsize et vsize). On peut dans goptions donner les valeurs de police de caractères et couleurs à utiliser par défaut pour les titres, les axes, les labels et les symboles.

Exemple :

**goptions reset=all rotate=landscape border hsize=15 cm vsize=10 cm  
gunit=pct ftext=hwdmx021 htext=3 ctext=red;**

Ces options fourniront un graphique de format paysage (rotate=landscape), de dimension horizontale 15 cm (hsize) et verticale 10 cm (vsize), bordé d'un trait (border), la police de caractères est Times New Roman (ftext), sa hauteur 3% du graphique (gunit et htext), de couleur rouge (ctext).

### Titres et notes

Les commandes **TITLE** et **FOOTNOTE** définissent des lignes de texte autour du graphique. Des options sont disponibles concernant la taille, la couleur, la police des caractères, les rotations de la ligne de texte et des caractères, les tracés de lignes.

On peut fournir jusqu'à 10 titres. Par défaut, ils sont situés horizontalement en haut et au centre du graphique.

Syntaxe :

**title**< 1 à 10 ><options 'texte' ;

Les options concernent :

soit l'apparence :

**color**=couleur (c=...)

**font**=police de caractère (f=...)

**height**=hauteur (h=...)

soit la position :

**justify**=left ou right ou center (j=...)

**rotate**=degrés (r=...) pour les lettres du texte

**angle**=degrés (a=...) pour le texte

soit l'encadrement :

**box**= 1 à 4 (selon l'épaisseur du trait)

**bspace**=distance du texte à la bordure en in ou cm

Exemple :

```
title1 color=red font=hwdmx024 height=1 cm 'Présentation d'un graphique';
```

L'instruction **FOOTNOTE** permet de fournir des informations au bas des graphiques (jusqu'à 10 notes de bas de page). Elle a une syntaxe et des options identiques à l'instruction **title**.

Notons une différence avec l'instruction **title** : on peut mettre sur une même ligne (donc dans un même footnote) des textes ayant des options d'apparence ou de positions différentes. Ainsi dans l'exemple suivant, on aura sur une même ligne un texte à gauche (date) un texte au centre (Présentation d'un graphique) et un texte à droite (source).

Exemple :

```
footnote1 c=red f=hwdmx024 h=1 cm 'Présentation d'un graphique'  
c=blue f=hwdmx020 h=0.5 cm j=right 'Source'  
c=black f=hwdmx023 h=0.4 cm j=left 'date';
```

De même que pour les titres, les options des notes de bas de graphiques pourront être utilisées et modifiées aisément avec l'éditeur graphique.

## Légendes

Comme pour les axes, différents types de légendes (de 1 à 99) sont définissables. Ils spécifient positions et textes des libellés qui identifient les différents graphismes et symboles utilisés.

Syntaxe :

```
legend <1 à 99>  
  <options d'apparence>  
  <options de position>  
  <options de texte>;
```

options d'apparence :

```
frame (entoure la légende d'un cadre)  
across=n (légende en ligne n=nombre d'éléments)  
down=n (légende en colonne n=nombre d'éléments)
```

options de position :

```
position= (<bottom|middle|top>  
  <left|center|right>  
  <outside|inside>);
```

ou

```
origin=(<x><pct|in|cm><y><pct|in|cm>);
```

x et y sont les coordonnées du bas gauche de la légende, les unités sont soit in, cm ou % du graphique.

options de texte :

```
label=(<options de description du texte> 'texte');
```

**label** concerne le label de la variable représentée, par défaut c'est le nom de la variable si elle ne possède pas de label.

```
value=(<options de description du texte> 'texte'  
  <options de description du texte> 'texte');
```

Il y a autant d'options de description de texte et de texte que de modalités de la variable représentée pour **gchart** ou **gmap**.

## Instruction pattern

L'instruction **pattern** définit les caractéristiques des surfaces que l'on utilise dans les représentations (procédures **gchart**, **gmap**, **gplot**...).

Syntaxe :

```
pattern<1...99>  color=(couleur de la surface c=...)  
                  value=(type de remplissage v=...)  
                  repeat=(nombre de fois à répéter r=...)
```

L'ordre des « patterns » définit l'ordre dans lesquels ils vont être utilisés. On peut jouer sur la combinaison des options repeat et sur les options de couleurs définies par goptions, pour ne pas avoir à réécrire toutes les caractéristiques de chacune des surfaces.

Ainsi : **goptions (color=blue red) ;**

**pattern1 v=s;**

**pattern2 v=x3 color=green r=2; /\*r=2 : il va prendre ces valeurs 2 fois\*/**

fournira 4 patterns et l'ordre suivant :

<i>séquence numéro</i>	<i>modèle de pattern</i>	<i>couleur</i>	<i>type de surface</i>
1	pattern1	blue	solid
2	pattern1	red	solid
3	pattern2	green	x3
4	pattern2	green	x3

### Instruction symbol

Cette instruction n'est utilisée que dans la procédure gplot. Elle définit les options de représentation des points et des tracés des courbes.

Syntaxe :

```

Symbol <1 à 99> <color= >
<mode=exclude|include>
<repeat=nombre de répétition>
<options d'apparences>
<options d'interpolation> ;

```

La numérotation de **symbol** a le même effet que la numérotation de **pattern** et le rôle de l'option repeat aussi. L'option **mode**, par défaut **mode=exclude**, s'utilise avec les interpolations et spécifie si les valeurs hors de l'échelle des axes doivent être exclues ou incluses.

### Axes

Des types d'axes, numérotés de 1 à 99, sont définis avant de pouvoir être utilisés dans les différents graphiques. Ils précisent l'échelle (liste de valeurs, logarithmique), l'apparence (longueur, couleur, épaisseur, style de ligne, origine), les marques d'échelle (nombre, couleur, épaisseur, hauteur), le libellé (police...).

Exemple :

```

axis1 order=(1973 to 1981 by 2)
      label=(h=1 f=hwdmx021 'Année')
      minor=(h=0.2 w=1 number=1)
      width=3;
axis2 order=(0 to 10000 by 1000)
      label=(h=1 f=hwdmx021 'revenu en francs')
      minor=none
      width=3;

```

### V. 3. Annotate data set

Une *table d'annotations*, définie lors d'une étape data, est une table SAS contenant les descriptifs d'un ensemble de graphiques qui viendront se superposer aux résultats des procédures précédemment décrites (GCHART, GPLOT,...). Elle contient ainsi un certain nombre de variables de positions, de descriptions et de texte.

Les variables de la table d'annotation ont des noms prédéfinis ; les plus usuelles sont :

- les variables qui définissent les éléments de dessins : **function**.  
**function** indique ce qu'il faut tracer : *bar, draw, frame, pie, symbol, label...*
- les variables de position qui spécifient les coordonnées des points où les tracés doivent être faits et le système de coordonnées :
  - x** abscisse du point
  - y** ordonnée du point
  - z** si 3 dimensions
- système de coordonnées pour le tracé (se combine avec size) :
  - hsys** unité à utiliser pour la taille du texte
  - xsys** système de coordonnées pour x
  - ysys** système de coordonnées pour y
  - zsys** système de coordonnées pour z

Ces variables dont les noms sont prédéfinis sont des variables caractères ; leurs valeurs doivent être rentrées comme telles.

- les variables d'attribut :
  - color** couleur du texte
  - style** police de caractère
  - text** contient le texte à écrire ou le nom de la variable du fichier qui contient les différents textes comme autant de valeurs
  - size** détermine la hauteur du texte
  - position** d'un texte par rapport au point (calé à gauche, centré...)
  - line** type de ligne (par défaut, continue).

Exemple :

Représentation des resultants d'une analyse des correspondances des données de parc3 à partir d'un fichier de sortie de **proc corresp**. Dans ce fichier de sortie, on ne garde que les variables :

**\_TYPE\_** (=OBS ou VAR)

**\_NAME\_** (contient les noms des variables et des observations). Si un format a été utilisé, c'est le libellé qui apparaîtra)

**dim1** et **dim2** (coordonnées des points sur les axes 1 et 2).

Le programme suivant créera un fichier **annotate** (table SAS) :

```
data anno;
set corr1(keep=_TYPE_ _NAME_ dim1 dim2);
x=dim1;
y=dim2;
xsys='2';
ysys='2' ;
if _TYPE_='OBS' then do;
text=_NAME_;
size=0.7;
color='b1';
style='hwdmx010';
output;
end;
if _TYPE_='VAR' then do;
text=_NAME_;
size=0.7;
color='b1';
style='hwdmx011';
output;
end;
label x='dimension 1'
      y='dimension 2';
keep x y text xsys ysys size color style ;
run;
```

Obs	x	y	xsys	ysys	text	size	color	style
1	-0.15047	-0.00000	2	2	1	0.7	b1	hwdmx010
2	0.82759	0.00000	2	2	2	0.7	b1	hwdmx010
3	0.93777	0.00000	2	2	3	0.7	b1	hwdmx010
.....	.....	.....	.....	.....	.....	.....	.....	.....
14	-0.28133	0.00000	2	2	HERB	0.7	b1	hwdmx011
15	0.89370	0.00000	2	2	BOSQUET	0.7	b1	hwdmx011
16	-0.55856	0.96462	2	2	EAU	0.7	b1	hwdmx011
17	-0.55856	-0.32154	2	2	HIPPO	0.7	b1	hwdmx011

Dans les versions antérieures de SAS, l'enrichissement des graphiques était difficile : tout devait être programmé, en particulier à partir des fichiers *annotate*.

Maintenant, la présentation peut être très soignée en utilisant de façon conviviale l'éditeur graphique de SAS. Il est possible de travailler directement sur un graphique en l'enrichissant de textes, en choisissant des polices de caractères adéquates et variées, en modifiant les aspects des surfaces ou des courbes...

## V. 4. Procédure gchart

La procédure proc gchart a de nombreuses options de représentation :

- block                construit des barres à 3 dimensions
- vbar                 histogramme barres verticales
- hbar                 histogramme barres horizontales
- pie                  camemberts
- star                 graphiques « étoilés »

Syntaxe :

```
proc gchart        <data=fichier des données>
                  <annotate=fichier annotate>
                  <gout=catalogue graphique> ;
```

Appel de l'option de tracé (exemple barres verticales) :

```
    vbar            variables à représenter
                  </annotate=fichier annotate>
                  <options d'apparence>
                  <options statistiques>
                  <options de classes>
                  <options d'axes>
                  <autres options descriptives>;
    <by <variable> ;
```

et toujours les instructions d'habillage :

```
    title ;
    footnote ;
    legend ;
    axis ;
```

Il est important de spécifier l'option **discrete** pour les variables qui ont été recodées par la procédure proc format, sinon le programme essaiera de créer des classes.

Des statistiques (fréquences et pourcentages) sont imprimées à côté du graphique. On peut ne pas les imprimer en utilisant l'option **nostat**.

Exemple :

On a un fichier d'enquête avec le nombre de volés par âge et par sexe :

OBS	AGE	SEX	COUNT	PERCENT
1	<25 ans	homme	21	5.6300
2	<25 ans	femme	27	7.2386
3	25-29 ans	homme	30	8.0429
4	25-29 ans	femme	31	8.3110
5	30-39 ans	homme	48	12.8686
6	30-39 ans	femme	66	17.6944
7	40-49 ans	homme	25	6.7024
8	40-49 ans	femme	35	9.3834
9	50-59 ans	homme	10	2.6810
10	50-59 ans	femme	32	8.5791
11	60&+	homme	27	7.2386
12	60&+	femme	21	5.6300

Le programme utilisera l'option `sumvar=` et l'option `sum`. Dans ce cas, les statistiques ne sont plus imprimées.

```
title1 f=hwdmx010 h=0.4 cm c=blue
'Repartition des voles en 1986 selon l"age et le sexe' ;
pattern1 c=gray v=s ;
pattern2 c=blue v=s;
axis1 w=3 c=red label=none length=8 cm;
axis2 w=3 label=none length=8 cm;

proc gchart data=chart;
hbar age/discrete
      sumvar=count
      subgroup=sex
      sum
      frame
      axis=axis2
      maxis=axis1
      space=1
      outline=bl;
run;
```

On obtient alors la représentation suivante :

{ EMBED PBrush }

Lorsque la variable principale est une variable numérique continue comme ici, l'âge, on peut laisser SAS créer les classes à représenter ; il faut spécifier alors combien de classes retenir, par l'option `levels=` . Les classes sont construites en divisant l'étendue de la distribution en classes d'intervalles égaux.

## V. 5. Procédure `gplot`

La procédure `gplot` produit des graphiques à deux dimensions dans un système d'axes perpendiculaires `haxis` et `vaxis`. Les points sont définis par leurs coordonnées sur les axes.

Syntaxe :

```
proc gplot data=<nom du fichier des données>
gout=<nom du catalogue de sortie du graphique>
annotate=<fichier annotate valable pour tous les graphiques dans cette procédure> par
exemple si on utilise by
<uniform> spécifie que les mêmes axes doivent être utilisés pour tous les graphiques de
cette procédure ;
```

Exemple :

```
proc gplot data=in.délits gout=fig.grf1 uniform ;
```

**plot**      **y\*x=z** ( **x** variable dont les valeurs se lisent sur l'axe horizontal  
                  **y** variable dont les valeurs se lisent sur l'axe vertical  
                  **z** désigne une 3<sup>ème</sup> variable de classification ; elle permet la construction  
                  d'autant de courbes que de valeurs et l'édition d'une légende automatique)

ou          **y\*x=n** ( **x** variable dont les valeurs se lisent sur l'axe horizontal  
                  **y** variable dont les valeurs se lisent sur l'axe vertical  
                  **n** désigne l'indice de l'instruction **symboli** avec **n=i**)

Les instructions placées après `goptions` (voir ci-dessus) contiennent les caractéristiques des éléments du graphique. Les instructions suivantes se situent après l'instruction `plot` :

Options d'apparence

**overlay** autorise le tracé de plusieurs courbes dans les mêmes axes

**areas=n** donne le nombre **n** de surfaces à dessiner en utilisant les options définies dans les instructions **pattern**

**nolegend** si on ne désire pas de légende ou

**legend=legend<1 à 99>** utilise la légende selon les options décrites dans les instructions legend.

**skipmiss** conserve les valeurs manquantes des données. Exemple dans les séries chronologiques ou des dates ne sont pas renseignées, la rupture de la série apparaît bien.

## Options d'axes

### Lignes de quadrillage

**autoref** trace le quadrillage selon toutes les valeurs principales de l'axe horizontal

**autovref** trace le quadrillage selon toutes les valeurs principales de l'axe vertical

**chref=** référence de la couleur de l'axe horizontal

**cvref=** référence de la couleur de l'axe vertical

**href=** liste des valeurs de la variable x pour lesquelles on trace une droite perpendiculaire à l'axe des x

**lref=** code des lignes tracées suite à href.

**vref=** liste des valeurs de la variable y pour lesquelles on trace une droite perpendiculaire à l'axe des y

**lvref=** code des lignes tracées suite à vref

**frame** trace un cadre autour de la surface définie par les axes

**grid** trace le quadrillage selon toutes les valeurs principales de l'échelle des axes x et y

### Axe horizontal

**haxis=axis<1 à 99>** renvoie aux options définies dans l'instruction axis

**hminor=n** n est le nombre de marques entre deux divisions principales de l'axe

**hzero** spécifie que l'axe horizontal commence au point de coordonnée x=0.

### Mêmes options pour l'axe vertical (vaxis, vminor, vzero)

L'option **vreverse** spécifie que l'ordre des valeurs de l'échelle de l'axe vertical doit être inversé.

### Apparence pour l'ensemble des axes

**noaxis** pas de tracé des axes

**caxis=** couleur des axes

**ctext=** couleur des textes.

### Quelques exemples de tracés graphiques

Exemple avec une seule courbe :

```
goptions reset=all rotate=landscape border ftext=hwdmx013 htext=0.35;
title2 c=blue f=hwdmx020 h=1.4 j=center 'Evolution de la population
française de 1987 à 2002' ;
title3 c=bl f=hwdmx011 h=0.7 'Source : OCDE' ;
symbol1 v=none I=join l=1 c=blue w=2;
symbol2 v=none I=join l=8 c=green w=2;
symbol3 v=none I=join l=34 c=red w=2;
axis1 length=8 cm
    label=(h=1 f=hwdmx021 h=0.8)
    value=(f=hwdmx021 h=0.8)
    order=(1987 to 2002 by 2)
    minor=(h=0.2 w=1.0 n=4);
axis2 length=6 cm
    label=(h=1 f=hwdmx021 'population')
    value=(f=hwdmx021 h=0.8)
    order=(55500 to 59500 by 1000)
    major=(h=0.5 w=1.5)
    minor=(h=0.2 w=1.0 n=4);
proc gplot data=population gout=graf1;
```

```
plot France*Annee=1 / haxis=axis1 vaxis=axis2;
run;
```

{ EMBED PBrush }

Exemple de courbes multiples avec option overlay :

```
goptions reset=all rotate=landscape border ftext=hwdmx013 htext=0.35 cm;
title2 c=blue f=hwdmx020 h=1.4 j=center 'Evolution des populations de 1987
a 2002';
title3 c=bl f=hwdmx011 h=0.7 ' Source : OCDE';
symbol1 v=none i=j l=1 c=blue w=2;
symbol2 v=none i=j l=8 c=green w=2;
symbol3 v=none i=j l=34 c=red w=2;
axis1 length=10 cm
      label=(h=1 f=hwdmx021 'annees')
      value=(f=hwdmx021 h=0.8);
axis2 length=8 cm
      label=(h=0.8 f=hwdmx021 'Population')
      value=(f=hwdmx021 h=0.8)
      major=(h=0.5 w=1.5)
      minor=(h=0.2 w=1.0 n=4);

data anno;
set population;
length text $ 10;
if Annee=2000 then do;
  xsys='2';
  ysys='2';
  position='2';
  size=0.8 ;
  color='bl' ;
  style='hwdmx020';
  x=Annee;
  y=France;text='France';output;
  y=Germany;text='Allemagne';output;
end ;
keep x y text xsys ysys size color style position;
run;

proc gplot data=population gout=graf2;
plot France*Annee=3 Germany*Annee=2 / overlay haxis=axis1 vaxis=axis2
annotate=anno;
run;
```

{ EMBED PBrush }

## V. 6. Sauvegarde des graphiques

Les graphiques sont sauvegardés dans des catalogues graphiques qui peuvent être temporaires ou assignés par un libname.

La sauvegarde dans un catalogue créé se fait de deux façons :

- soit en utilisant le menu **FILE** et **save as** et en répondant aux demandes du sous menu qui apparaît.
- soit en utilisant l'option **gout=libname.catalogue** dans la procédure graphique.

L'accès aux catalogues se fait par la procédure **proc greplay**.

Exemple : on effectue la procédure suivante :

```
proc greplay data=graf1;  
run;
```

Dans la colonne « Sel » devant « Plot of France\* Annee », on tape « s » puis « enter » et on obtient alors le précédent graphique.