



Introduction à SPSS

Premiers pas

Jonathan Godin

Département de mathématiques et de statistique

septembre 2017

Pourquoi utiliser SPSS ?

Méthode de travail

Étape 1 : importation

Étape 2 : Analyses statistiques

Syntaxe

Étape 3 : Exportation

Ressources

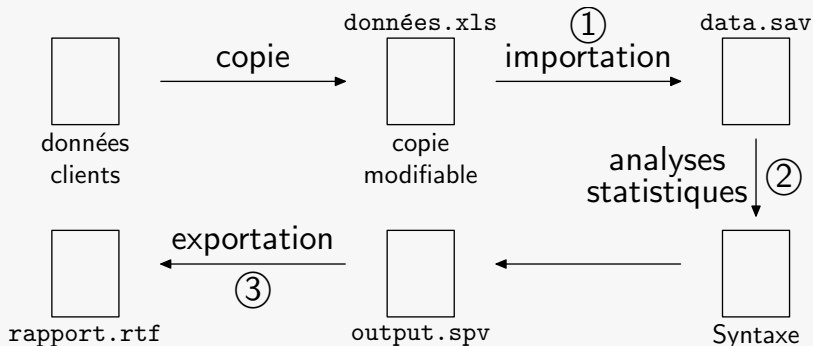
Pourquoi utiliser SPSS ?

Qu'est-ce que SPSS ?

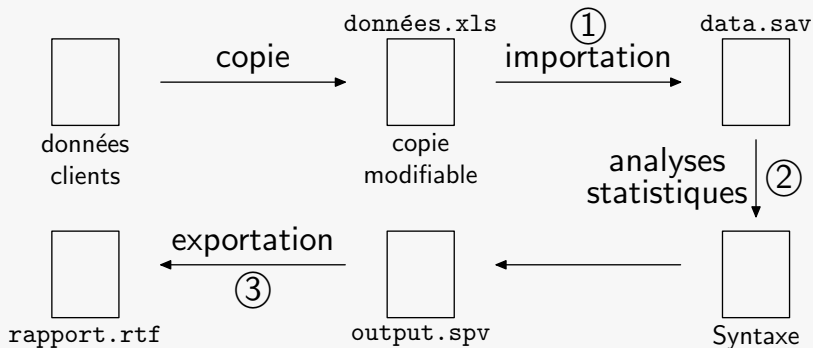
- ▶ SPSS : *Statistical Package for the Social Sciences*
- ▶ Logiciel pour l'analyse statistique.

- ▶ Interface facile à utiliser
 - ▶ Presque tout peut se faire en cliquant
 - ▶ Coller dans Syntaxe
- ▶ Très utilisé sur le campus
- ▶ Important de connaître plus d'un langage

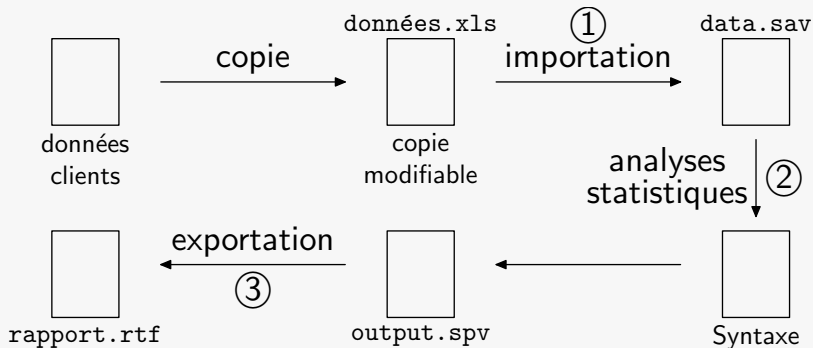
Méthode de travail



① Ajuster (le moins possible) la copie pour permettre l'importation.



② Analyser les données. La syntaxe contient la démarche du travail.



③ Exporter les résultats.

Étape 1 : importation

- ▶ On peut importer les formats suivants :
 - ▶ feuille excel
 - ▶ texte (.txt ou .csv)
 - ▶ SQL
 - ▶ SAS
- ▶ SPSS interprète les données
 - ▶ Il lit la première ligne comme le nom des variables (si disponible)
 - ▶ Il associe un **type** à chaque variable (colonne).
Si la colonne :
 - ▶ ne contient que des nombre : type numérique
 - ▶ ne contient que des dates **du même format** : type date
 - ▶ sinon : type chaîne de caractères (*string*)

Données originales

0	Id	Sexe	Distance
1	121	H	10.3
2	122	H	7,6
3	123	F	-1
4	124	3	11,2
5	125	F	9.4

Copie modifiée des données

0	Id	Sexe	Distance
1	121	0	10,3
2	122	0	7,6
3	123	1	99
4	124	99	11,2
5	125	1	9,4

0 = « Homme »

1 = « Femme »

99 = « Valeur manquante »

- ▶ Décimaux : toujours la virgule ou toujours le point !
- ▶ Identifiez les valeurs manquantes pas une valeur insensée

- ▶ Rester uniforme dans tout le fichier
 - ▶ Soit toutes les chaînes de caractères sont entre " " ou aucune ne l'est
 - ▶ Chaque colonne est délimitée par le même caractère (espace, virgule, etc.)
 - ▶ Les nombres sont sans espaces (10210 au lieu de 10 210)

ID, Sexe, scolarite, salaire

1, 0, 2, 10000.00

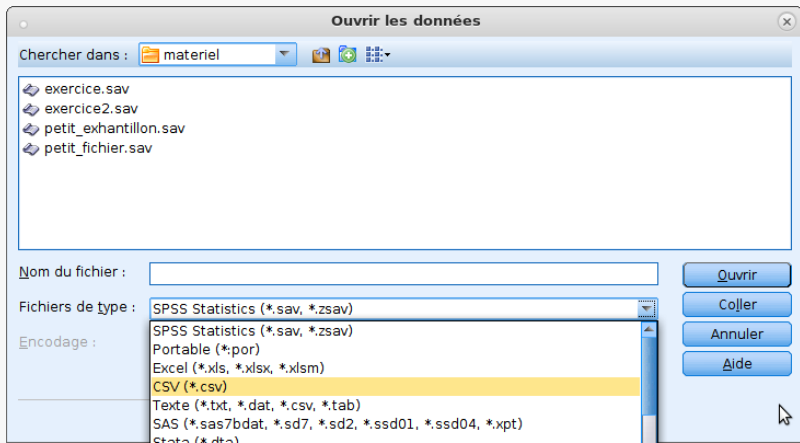
2, 1, 2, 12250.12

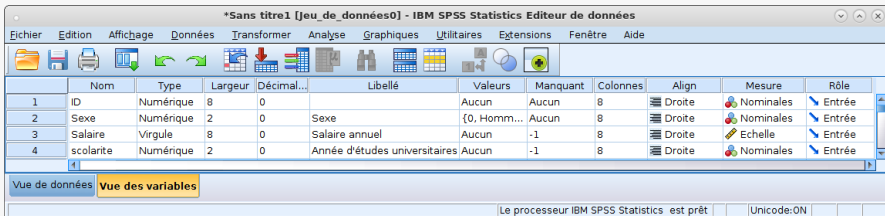
3, 1, 5, 34125.25

4, 1, 1, 10400.85

- ▶ **Attention** à la ponctuation pour les décimaux (un point dans ce cas)

fichier -> ouvrir -> données





The screenshot shows the 'Edit Data' window in IBM SPSS Statistics. The title bar reads '*Sans titre1 [Jeu_de_données0] - IBM SPSS Statistics Éditeur de données'. The menu bar includes 'Fichier', 'Édition', 'Affichage', 'Données', 'Transformer', 'Analyse', 'Graphiques', 'Utilitaires', 'Extensions', 'Fenêtre', and 'Aide'. The toolbar contains icons for file operations, data manipulation, and analysis. The main table lists variables with the following columns: Nom, Type, Largeur, Décimal..., Libellé, Valeurs, Manquant, Colonnes, Align, Mesure, and Rôle. The table contains four rows of data.

	Nom	Type	Largeur	Décimal...	Libellé	Valeurs	Manquant	Colonnes	Align	Mesure	Rôle
1	ID	Numérique	8	0		Aucun	Aucun	8	Droite	Nominales	Entrée
2	Sexe	Numérique	2	0	Sexe	{0, Homm...	Aucun	8	Droite	Nominales	Entrée
3	Salaire	Virgule	8	0	Salaire annuel	Aucun	-1	8	Droite	Echelle	Entrée
4	scolarité	Numérique	2	0	Année d'études universitaires	Aucun	-1	8	Droite	Nominales	Entrée

At the bottom, there are buttons for 'Vue de données' and 'Vue des variables' (highlighted in orange). The status bar at the bottom right indicates 'Le processeur IBM SPSS Statistics est prêt' and 'Unicode:ON'.

Nom : Nom de la variable, sans espace (64 caractères max)

Type : S'assurer que les variables numériques soient numériques et non des chaînes !

Largeur : Nombre maximal de chiffres (numérique) ou de caractères (chaîne) affichés.

Libellé : Nom complet des variables (apparaît dans les tableaux !)

Valeurs : Nom associé à une valeur (p.ex. {0,Homme}, {1, Femme}).

*Sans titre1 [Jeu_de_données0] - IBM SPSS Statistics Éditeur de données

Fichier Edition Affichage Données Transformer Analyse Graphiques Utilitaires Extensions Fenêtre Aide

	Nom	Type	Largeur	Décimal...	Libellé	Valeurs	Manquant	Colonnes	Align	Mesure	Rôle
1	ID	Numérique	8	0		Aucun	Aucun	8	Droite	Nominales	Entrée
2	Sexe	Numérique	2	0	Sexe	{0, Homm...	Aucun	8	Droite	Nominales	Entrée
3	Salaire	Virgule	8	0	Salaire annuel	Aucun	-1	8	Droite	Echelle	Entrée
4	scolarité	Numérique	2	0	Année d'études universitaires	Aucun	-1	8	Droite	Nominales	Entrée

1

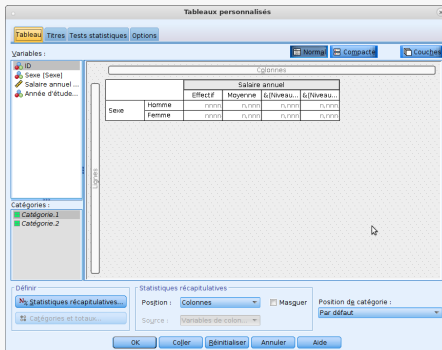
Vue de données **Vue des variables**

Le processeur IBM SPSS Statistics est prêt Unicode:ON

- Manquant :** Valeurs spéciales pour indiquer une valeur manquante.
- Colonnes :** Largeurs des colonnes (esthétique).
- Mesure :** Échelle : variable continue ;
Ordinale : variable ayant un ordre (p.ex. l'intensité) ;
Nominal : variable ayant des groupes (p.ex. homme, femme).
- Rôle :** Rôle que joue les données. (Sans importance.)

Étape 2 : Analyses statistiques

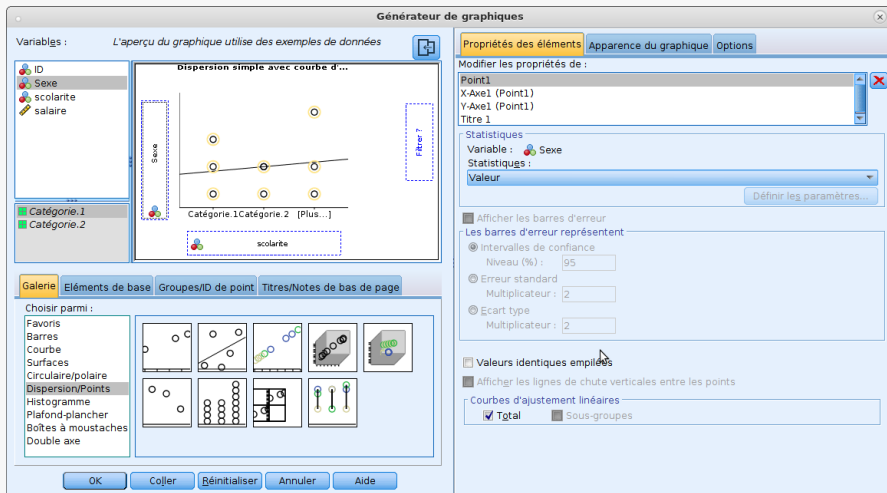
Analyse -> Tableaux -> Tableaux personnalisés...



Mise en garde ! Il
est important que les
attributs des variables
soient bien définis

Choisir dans Statistiques récapitulatives...
moyenne, médiane, maximum et bien d'autres

Graphiques -> Générateur de graphiques...



Générateur de graphiques

Variables : *L'aperçu du graphique utilise des exemples de données*

Propriétés des éléments | Apparence du graphique | Options

Modifier les propriétés de :

Point1
X-Axe1 (Point1)
Y-Axe1 (Point1)
Titre 1

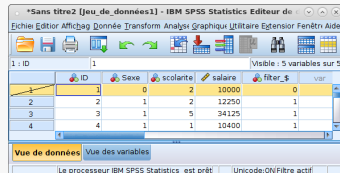
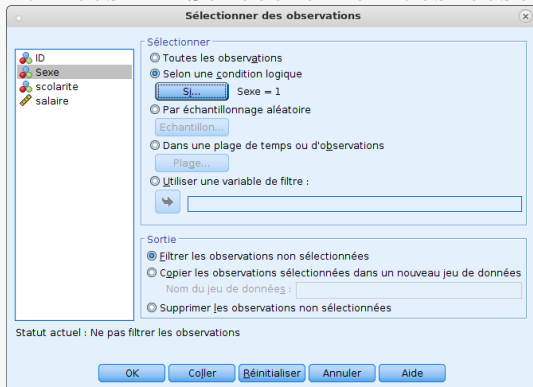
Statistiques
Variable : Sexe
Statistiques : Valeur
Définir les paramètres...

☐ Afficher les barres d'erreur
Les barres d'erreur représentent
☒ Intervalles de confiance
Niveau (%) : 95
☐ Erreur standard
Multiplicateur : 2
☐ Écart type
Multiplicateur : 2

☒ Valeurs identiques empiées
☐ Afficher les lignes de chute verticales entre les points
Courbes d'ajustement linéaires
☒ Total ☐ Sous-groupes

OK Coller Réinitialiser Annuler Aide

Données -> Sélectionner des observations



*Sans titre2 [Jeu_de_données1] - IBM SPSS Statistics Editeur de données

Fichier Editeur Affichage Données Transform Analyse Graphique Utilitaire Extension Fenêtre Aide

1 : ID 1 Sexe 0 2 10000 0 1

2 2 1 2 12250 1

3 3 1 5 34125 1

4 4 1 1 10400 1

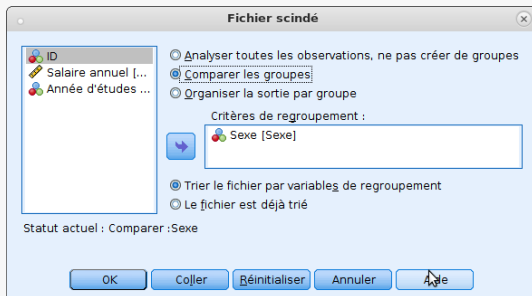
Visible : 5 variables sur 5

Vue de données Vue des variables

Le processeur IBM SPSS Statistics est prêt Unicode:ON/Filtre actif

- ▶ Permet de faire des analyses sur **un sous-ensemble de données**.
- ▶ Les autres données sont ignorées jusqu'à ce que le filtre soit levé.
(Revenir à cette fenêtre et choisir «Toutes les observations»)

Données -> Scinder un fichier

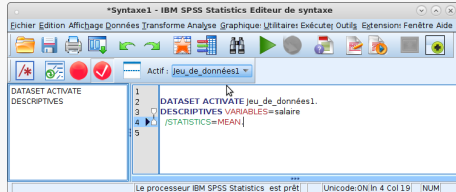
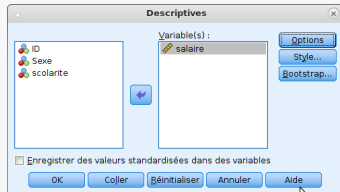


Permet d'effectuer la même analyse sur chacun des groupes

Syntaxe

Exemple : on veut la moyenne d'un échantillon

- Analyse -> Statistiques Descriptives -> Descriptives



Génère la syntaxe qui produit exactement l'opération effectué

L'option coller fonctionne pour toutes les opérations faites « en cliquant »

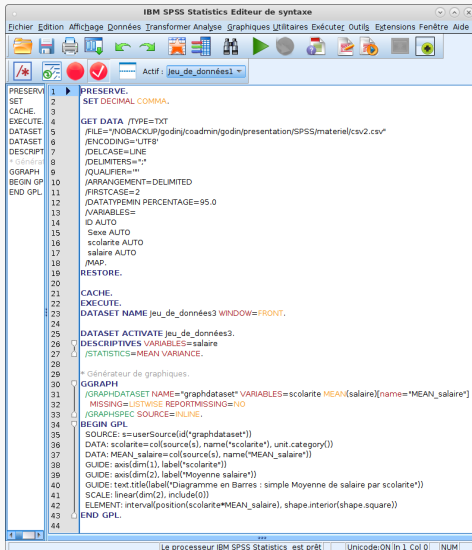
La Syntaxe est :

- ▶ le langage de programmation de SPSS
- ▶ une façon de reproduire et de partager ses procédures et ses analyses

Principe : On cherche et on expérimente pour trouver les analyses que l'on veut dans l'interface. Après avoir trouver *la bonne*, on la **colle** dans la syntaxe.

La syntaxe devrait contenir :

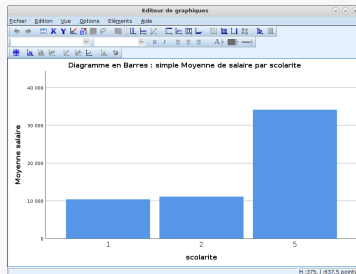
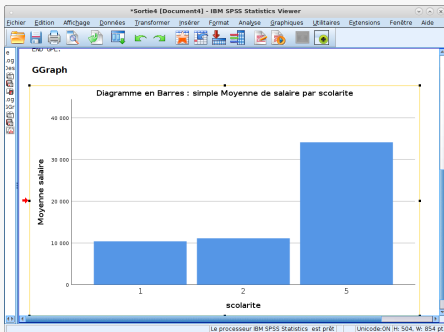
- ▶ l'importation des données
- ▶ une séquence d'instructions
 - ▶ qui manipule les données (p.ex. appliquer des filtres, scinder)
 - ▶ qui analyse les données
- ▶ des commentaires qui expliquent chaque étape
 - ▶ * commence une ligne de commentaire qui se termine par un point . ou une ligne blanche
 - ▶ /*...*/ contient du texte en commentaire



```
1 PRESERVE.  
2 SET DECIMAL COMMA.  
3  
4 GET DATA /TYPE=TEXT  
5 /FILE="\\NOBACKUP\\godin\\coadmin\\godin\\presentation\\SPSS\\materiel\\csv2.csv"  
6 /ENCODING="UTF8"  
7 /DELCASE=LINE  
8 /DELIMITERS="*" *  
9 /QUALIFIER=""  
10 /ARRANGEMENT=DELIMITED  
11 /FIRSTCASE=2  
12 /DATATYPEMIN PERCENTAGE=95.0  
13 /VARIABLES=  
14 ID AUTO  
15 Sexe AUTO  
16 scolaire AUTO  
17 salaire AUTO  
18 /MAP.  
19  
20 RESTORE.  
21  
22 CACHE.  
23 EXECUTE.  
24 DATASET NAME Jeu_de_données3 WINDOW=FRONT.  
25  
26 DATASET ACTIVATE Jeu_de_données3.  
27 DESCRIPTIVES VARIABLES=salaire  
28 /STATISTICS=MEAN VARIANCE.  
29  
30 * Générateur de graphiques.  
31  
32 GGRAPH  
33 /GRAPHDATASET NAME="graphdataset" VARIABLES=scolaire MEAN(salaire){name="MEAN_salaire"}  
34 MISSING=LISTWISE REPORTMISSING=NO  
35 /GRAPHSPEC SOURCE=INLINE.  
36  
37 BEGIN GPL  
38 SOURCE: s=userSource(id("graphdataset"))  
39 DATA: scolaire=col(source(s), name("scolaire"), unit.category())  
40 DATA: MEAN_salaire=col(source(s), name("MEAN_salaire"))  
41 GUIDE: axis(dim(1), label("scolaire"))  
42 GUIDE: axis(dim(2), label("Moyenne salaire"))  
43 GUIDE: text.title(label("Diagramme en Barres : simple Moyenne de salaire par scolaire"))  
44 SCALE: linear(dim(2), include(0))  
45 ELEMENT: interval(position(scolaire*MEAN_salaire), shape.interior(shape.square))  
46  
47 END GPL.
```

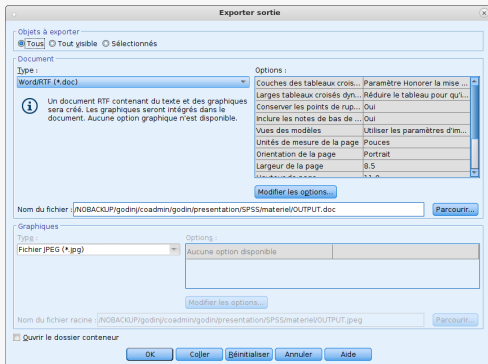
Étape 3 : Exportation

Dans la fenêtre de sortie, double-cliquer un graphique pour ouvrir l'éditeur.



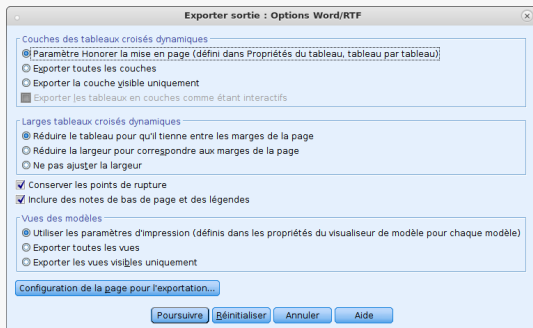
Ces changements sont surtout esthétiques.
(titre, bar de référence, etc.)

Sur la fenêtre de sortie :
fichier -> exporter...



Choisir le format. Changer les options au besoin dans « Modifier les options ».

Modifier les options...



Option parfois désirable pour exporter en Word/RTF :
Largeur des tableaux croisés dynamiques :
-> Ne pas ajuster la largeur

Consultez :

- ▶ Le wiki du département :
<http://dms.umontreal.ca/wiki/index.php/SPSS>
- ▶ Le manuel de référence pour la syntaxe

En cas de problème, cherchez d'abord sur Internet !

Pour des questions générales sur le logiciel

- ▶ Coadministrateur

Pour des questions plus spécifiques sur le programme ou sur les méthodes statistiques :

- ▶ le professeur
- ▶ le service de consultation