

## Cours 5

### ➤ Une brève présentation de logiciel d'analyse de données statistiques SPSS

---

#### 1. Introduction

Le logiciel SPSS est un instrument particulièrement adapté à la mise en œuvre de techniques **d'analyse des données statistiques**. Il favorise la gestion des données dans un **environnement graphique** convivial associant menu descriptif et boîtes de dialogue. En plus cet environnement offre un langage de commande qui permet d'écrire des programmes afin d'optimiser des tâches de production. Il permet de traiter les données avec efficacité et d'effectuer des analyses sur de grandes bases de données. Il offre plusieurs possibilités pour organiser et synthétiser les informations statistiques.

L'exploitation et le traitement statistique des données d'enquête de terrain supposent toujours un travail préalable d'organisation et de restructuration des informations. SPSS présente toutes les fonctionnalités requises pour effectuer ce travail. Ce cours d'initiation porte sur la dernière version du logiciel : PASW Statistics 18.

#### - **Qu'est-ce que SPSS ?**

SPSS signifie "Statistical package for the Social Sciences". Il s'agit d'un logiciel dont les premières versions datent des années soixante (sans doute l'un des plus anciens). À l'origine, il s'agissait d'un programme "open source", c'est à dire que n'importe qui pouvait ajouter de nouvelles commandes, et les possesseurs du logiciel recevaient un pamphlet supplémentaire décrivant cette commande. Dans les années 80, le logiciel a cessé d'être "open source" et est maintenant la propriété exclusive de SPSS inc. Néanmoins, les propriétaires sont très agressifs, sortant de nouvelles versions régulièrement (en moins de 8 ans, nous sommes passés de SPSS v. 6 à SPSS v.12). À partir de la version 7, SPSS est devenu un produit pour Windows. SPSS est un produit très dispendieux, et la licence dure généralement une année seulement.

#### 2. L'environnement graphique SPSS Windows

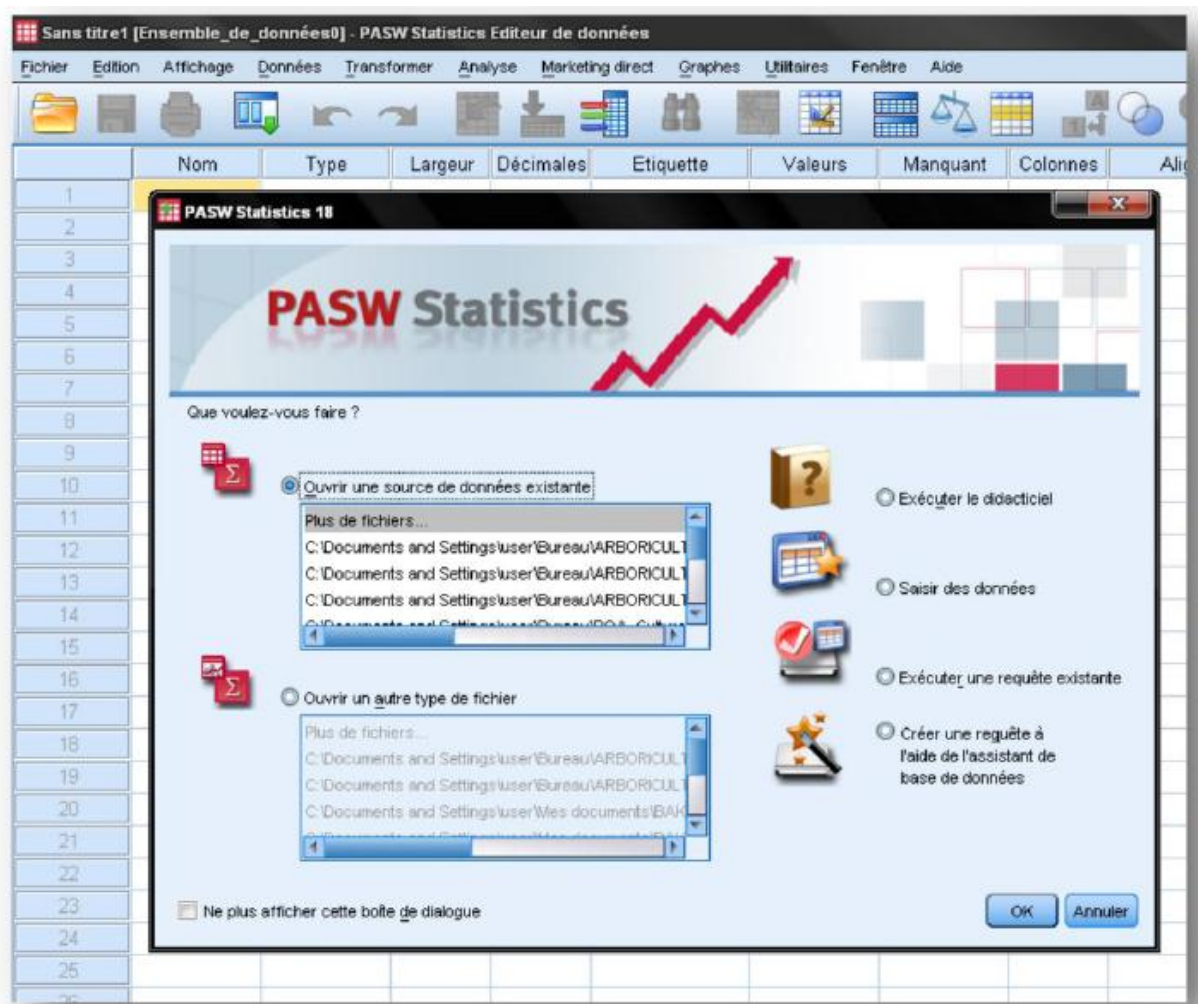
Le logiciel SPSS fonctionne dans un environnement graphique Windows, son utilisation repose largement sur la manipulation **de fenêtres et de menus déroulants**. Le mode

de fonctionnement de ces menus et la gestion des fenêtres répond donc aux principes généraux du système d'exploitation Windows.

### 3. Démarrage

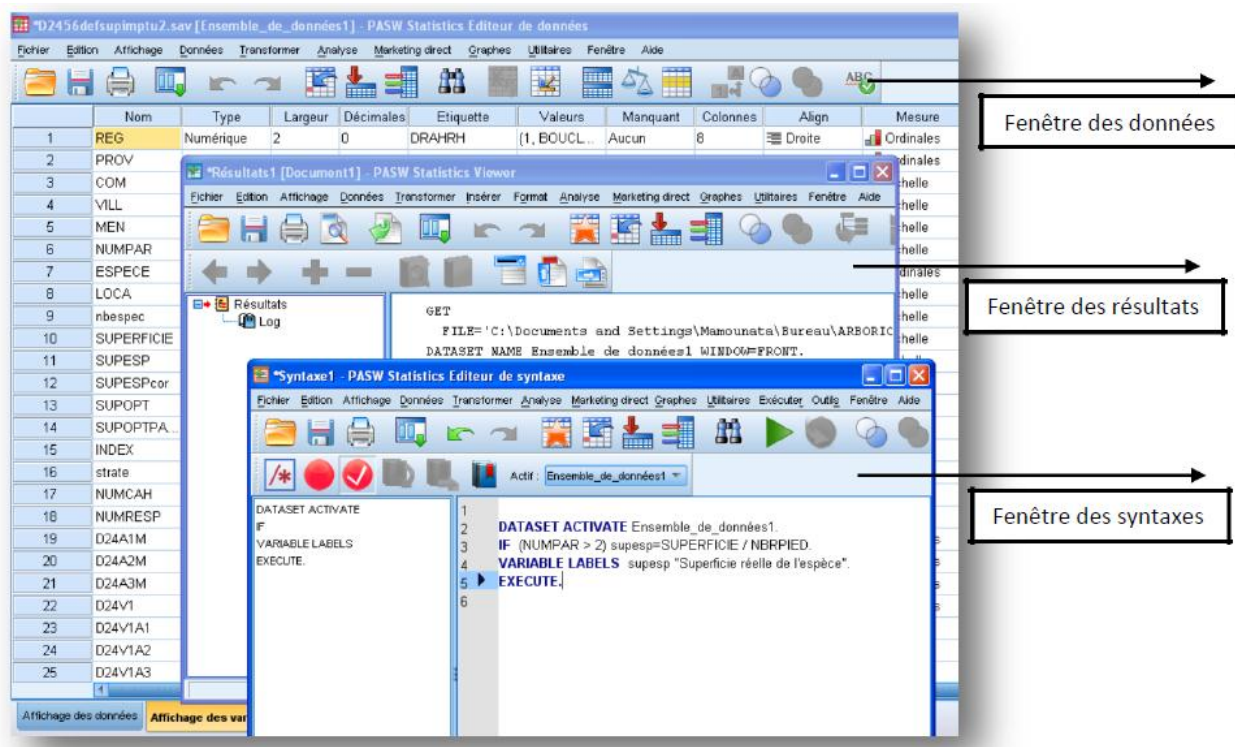
Après l'**installation** du logiciel, pour le **démarrage** il faut aller dans *Démarrer + Tous les programmes + SPSS Inc + PASW Statistics 18* et cliquer sur *PASW Statistics 18*. Toutefois il est conseillé de faire un clic droit et créer un **raccourci sur le bureau** pour ne pas reprendre chaque fois ce long processus.

La fenêtre de démarrage se présente comme l'indique l'image ci-dessous. Cette fenêtre donne la liste des fichiers SPSS récemment ouverts. Pour ouvrir un fichier, il suffit de faire un **double clic sur son nom**.



### 4. Les fenêtres et leurs contenus

Le logiciel SPSS est composé de trois fenêtres : la fenêtre **des données** (ou fenêtre de l'éditeur), celle **des résultats** (souvent appelée viewer) et celle **des syntaxes**.



#### 4.1. La fenêtre des données

La fenêtre des données permet d'entrer les données. Cette fenêtre présente deux onglets (en bas à gauche) : affichage des variables, et affichage des données. En cliquant sur affichage des variables, vous aurez l'ensemble de vos variables si vous avez déjà ouvert une base. Vous pouvez aussi ajouter ou supprimer des variables dans cette fenêtre.

#### 4.2. La fenêtre des résultats

Elle affiche les résultats des commandes exécutées par SPSS. C'est dans cette fenêtre que nous observerons les sortis de tableaux, de graphiques, etc. Cette fenêtre s'ouvre automatiquement dès que vous lancez une procédure qui doit afficher des résultats.

#### 4.3. La fenêtre des syntaxes

Elle permet de saisir les commandes SPSS. En effet, SPSS présente une configuration Windows avec menu déroulant (on peut donc travailler avec la souris) et une possibilité de saisir les programmes en langage SPSS. SPSS permet en effet d'obtenir la syntaxe de toute opération avec le bouton *coller*.

## 5. Comment organiser les données dans SPSS

Les données doivent être organisées de la façon suivante:

- Chaque **ligne** représente **un participant**; si le participant a été observé plusieurs fois, on doit retrouver plusieurs **colonnes**, une par **observation**.
- Chaque colonne représente une variable, soit:
  - La ou les variables indépendantes identifiant la condition expérimentale, le numéro de sujets, et possiblement des informations sur le sujet (tel le sexe, l'âge, etc.) selon les besoins;
  - La (ou les) variable dépendante, c'est à dire le résultat de la mesure. Si vous mesurez la taille, il y aura une colonne pour la taille; si vous cherchez à mesurer le Q.I, une colonne contiendra le résultat du sujet.
- Vous pouvez utiliser autant de colonnes que désiré. De plus, il peut y avoir des colonnes avec de l'information qui ne servira pas mais que vous avez récolté quand même.
- L'ordre des colonnes n'est pas important.

Dans SPSS, vous pouvez donner les noms que vous voulez à vos colonnes. Utiliser des noms significatifs (tel **sexe**, **age**, **qi**, etc.) de préférence à des noms sans signification (tel **foto**, **patate**, ou encore **v1**). Par contre, SPSS ne permet pas l'utilisation de symboles spéciaux dans le nom d'une colonne (le point, la virgule, le dièse, etc.), ne permet pas d'utiliser plus que huit lettres, et finalement, un nom de colonne doit commencer par une lettre.