

---

## Introduction à IBM SPSS Modeler et au Data Mining (V16)

**Durée: 2 Jours**    **Réf de cours: 0A005G**    **Version: 16**

---

### Résumé:

Le cours "**A005G - Présentation de IBM SPSS Modeler et exploration de données (v16)**" est une formation en salle de classe animée sur deux jours par un formateur.

Cette formation présente les principes de base de l'utilisation de IBM SPSS Modeler.

Les théories et les méthodes pratiques de l'exploration de données sont illustrées à l'aide de la méthodologie CRISP-DM. Le cours se déroule selon les étapes d'un projet d'exploration de données standard, qui commence par la lecture des données et passe par leur transformation, la modélisation et l'interprétation efficace des résultats. Il expose les bases des méthodes de lecture, d'exploration et de manipulation des données avec IBM SPSS Modeler, puis la création et l'utilisation des modèles les plus performants.

**Formation** inscrite à l'inventaire de la CNCP, **éligible au [Compte Personnel de Formation](#)** (CPF), permettant de préparer la Certification IBM Certified Associate - SPSS Modeler Data Mining. Attention, différents Codes CPF existent selon le secteur métier de l'entreprise. Nous consulter.

---

### Public visé:

Toute personne devant acquérir les compétences dans l'utilisation de IBM SPSS Modeler et du data mining : analystes, utilisateurs responsables de la construction de modèles tels que les modèles de scoring, les modèles de construction de typologies ou les modèles d'associations.

---

### Objectifs:

- Introduction à l'exploration de données
  - Définition de l'unité d'analyse
  - Utilisation de SPSS Modeler
  - Intégration des données
  - Découverte de l'exploration de données
  - Calcul et remplacement des valeurs des champs
  - Collecte des données initiales
  - Recherche de relations
  - Compréhension des données
  - Introduction à la classification
- 

### Pré-requis:

- Connaissances générales en informatique
  - Aucune connaissance particulière en termes de statistiques n'est requise.
  - Il peut s'avérer utile de comprendre les données de votre entreprise, ainsi que les problématiques relatives à l'exploration de données effectuée au sein de votre entreprise.
- 

### Cours suivant(s):

Les analystes et experts données pourront suivre d'autres modules de formation afin de développer leurs compétences :

- Cours IBM 0A0U4G - Score et segmentation avec IBM SPSS Modeler
  - Cours IBM 0A045G - Construction de typologie et modèles d'association avec IBM SPSS Modeler (v16)
  - Cours IBM 0A054G - Préparation des données avec IBM SPSS Modeler - cours avancé
-

## Contenu:

### Introduction à l'exploration de données

- Présenter 2 exemples d'exploration de données
- Expliquer les étapes de la méthodologie CRISP-DM
- Décrire des projets d'exploration de données réussis et les raisons d'un échec
- Décrire les compétences nécessaires à l'exploration de données

### Utilisation de SPSS Modeler

- Décrire les différentes zones de l'interface utilisateur de Modeler
- Utiliser les noeuds et super noeuds
- Exécuter, ouvrir et enregistrer un flux
- Accéder à la fonction d'aide dans SPSS Modeler

### Découverte de l'exploration de données

- Expliquer les principaux concepts de l'exploration de données
- Créer un modèle
- Déployer un modèle

### Collecte des données initiales

- Expliquer les concepts « structure de données », « enregistrements », « unité d'analyse » et « stockage »
- Importer des fichiers Microsoft Excel
- Importer des fichiers IBM SPSS Statistics
- Importer des fichiers textes
- Importer à partir d'une base de données
- Exporter des données vers différents formats

### Compréhension des données

- Examiner les distributions des champs qualitatifs et continus
- Expliquer les méthodes les plus courantes de gestion des données manquantes
- Expliquer les méthodes les plus courantes de gestion des valeurs éloignées
- Expliquer comment vérifier la qualité des données et sélectionner des enregistrements valides à l'aide de Modeler

### Définition de l'unité d'analyse

- Supprimer les enregistrements en double
- Agréger des données
- Créer des champs booléens
- Restructurer des champs continus

### Intégration des données

- Ajouter des enregistrements de plusieurs ensembles de données dans un seul ensemble
- Ajouter des champs de plusieurs ensembles de données dans un seul ensemble
- Utiliser l'échantillonnage à des fins de test

### Calcul et remplacement des valeurs des champs

- Utiliser les expressions CLEM (Control Language for Expression Manipulation) pour transformer des données
- Utiliser le noeud Derive (Calculer) pour créer un nouveau champ et le noeud Filler (Remplacer) pour remplacer les valeurs d'un champ
- Expliquer comment générer automatiquement un noeud Derive (Calculer)
- Utiliser le noeud Reorder (Réorganiser) pour réorganiser les champs

### Recherche de relations

- Examiner la relation entre deux champs qualitatifs
- Examiner la relation entre deux champs quantitatifs
- Examiner la relation entre deux champs qualitatif et quantitatif

### Introduction à la classification

- Différencier la modélisation prédictive des autres types de modélisation
- Distinguer les différents types de modèles prédictifs
- Exécuter CHAID en mode interactif
- Exécuter CHAID et divers modèles en mode automatique

## Autre(s) Information(s):

Support de cours officiel IBM en anglais remis aux participants

## Plus d'informations:

Pour plus de renseignements ou pour vous inscrire, merci de nous contacter au 01 78 15 34 00

[info@globalknowledge.fr](mailto:info@globalknowledge.fr)

[www.globalknowledge.fr](http://www.globalknowledge.fr)

100, avenue Albert 1er - 92500 RUEIL-MALMAISON