



---

## Points clés

- Identifier visuellement des profils, des segments ou des groupes à l'aide d'arbres de classification
  - Choisir parmi quatre algorithmes (CHAID, Exhaustive CHAID, C&RT et QUEST) celui qui est le mieux adapté à vos données
  - Présenter les résultats d'une façon intuitive idéale pour les utilisateurs non initiés
  - Enregistrer les informations des arbres en tant que nouvelles variables dans les données (des informations telles que le nombre de nœuds terminaux, la valeur prédite et les probabilités à posteriori)
- 

# IBM SPSS Decision Trees

*Identifiez facilement des groupes et prévoyez les résultats*

IBM SPSS Decision Trees crée des arbres de classification et de segmentation pour vous aider à mieux identifier des groupes, découvrir les relations qui existent entre eux et prévoir des événements futurs.

Les arbres de classification et de décision permettent de réaliser les opérations suivantes:

- Segmentation
- Stratification
- Prédiction
- Réduction des données et filtrage de variables
- Identification d'interactions
- Fusion de catégories
- Discrétisation de variables continues.

Des diagrammes très visuels vous permettent de présenter des résultats de manière intuitive, afin de les expliquer clairement à un public non initié. Ces arbres vous permettent d'explorer vos résultats et de déterminer visuellement comment fonctionne votre modèle. Les représentations graphiques peuvent vous aider à découvrir des sous-groupes spécifiques et des relations indétectables à l'aide des outils statistiques traditionnels. Comme les arbres de classification répartissent les données en branches et en nœuds, il est facile de voir où un groupe se divise ou prend fin.





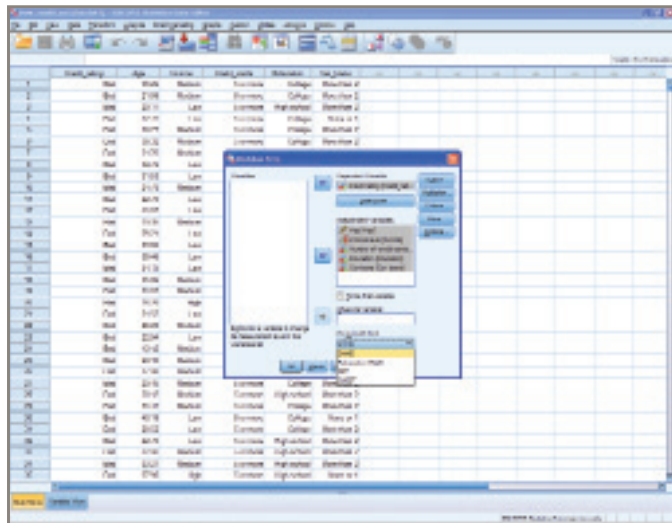


Figure 2 : Utilisez les résultats de la modélisation par arbre pour scorer les observations directement depuis IBM SPSS Statistics.

## Une plus grande valeur ajoutée grâce à la collaboration

Pour partager et réutiliser efficacement les ressources, pour les protéger selon les exigences de conformité internes et externes et publier les résultats afin qu'un plus grand nombre d'utilisateurs puisse les consulter et interagir avec ceux-ci, envisagez d'enrichir votre logiciel IBM SPSS Statistics avec IBM SPSS Collaboration and Deployment Services. Vous pouvez obtenir plus d'informations sur ces fonctionnalités sur le site [ibm.com/spss/cds](http://ibm.com/spss/cds)

## Quatre algorithmes d'arbre de décision disponibles

SPSS Decision Trees comprend quatre algorithmes reconnus pour la construction d'arbres :

- **CHAID** – un algorithme rapide, à séparation multiple pour explorer rapidement et efficacement les données et construire des segments et des profils en fonction du résultat désiré
- **Exhaustive CHAID** – une version modifiée de CHAID, qui examine toutes les divisions possibles de chaque valeur prédite
- **Arbres de classification et de régression (C&RT)** – un algorithme d'arborescence binaire complet qui partitionne les données et produit des sous-groupes homogènes précis
- **QUEST** – un algorithme qui sélectionne sans biais les variables explicatives et construit rapidement et efficacement des arbres binaires précis

Grâce à ces quatre algorithmes, vous avez la possibilité d'expérimenter différentes méthodes de développement d'arborescence afin de trouver celle qui convient le mieux à vos données.

Notre suite de logiciels statistiques est maintenant disponible en trois éditions: IBM SPSS Statistics Standard, IBM SPSS Statistics Professional et IBM SPSS Statistics Premium. En regroupant les fonctionnalités essentielles, ces éditions offrent un moyen efficace de garantir que toute votre équipe ou département dispose des fonctions nécessaires pour effectuer les analyses qui contribuent au succès de votre organisation.

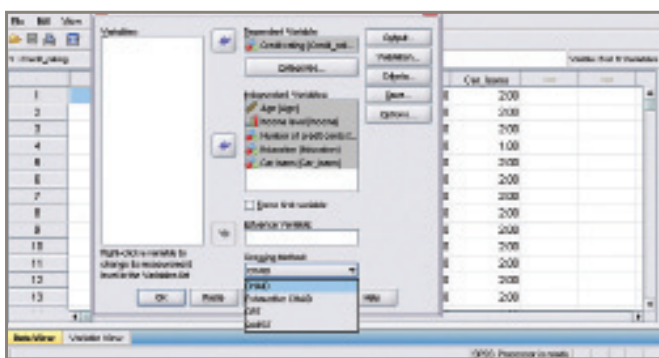


Figure 3 : Créez des modèles d'arbres dans IBM SPSS Statistics avec CHAID, Exhaustive CHAID, C&RT ou QUEST.

## Développez vos résultats en approfondissant l'analyse grâce à IBM SPSS Statistics Base

Si vous utilisez IBM SPSS Decision Trees avec IBM SPSS Statistics Base, vous pouvez facilement créer des arbres de classification et utiliser aisément les résultats pour segmenter et grouper des observations, directement au sein des données. De plus, vous pouvez générer des règles de sélection ou de classification/prédiction sous la forme de syntaxe IBM SPSS Statistics, de commandes SQL ou de texte simple (via la syntaxe). Vous pouvez afficher ces règles dans l'Afficheur et les enregistrer dans un fichier externe afin de les réutiliser ultérieurement, pour réaliser des prédictions sur les observations individuelles actuelles ou sur de nouvelles observations. Si vous souhaitez utiliser vos résultats pour évaluer d'autres jeux de données, vous pouvez inscrire les informations du modèle directement dans vos données ou créer des modèles XML utilisables dans IBM SPSS Statistics Base Server.

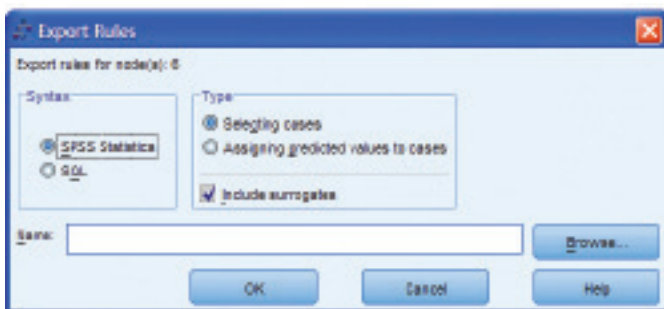


Figure 4 : Utilisez les résultats de modélisation directement dans les données pour sélectionner des observations ou affecter les classes, ou exportez des règles pour un usage ultérieur.

## Fonctionnalités

### Arbres

- Affichez des diagrammes en arbres, des schémas d'arborescence, des histogrammes et des tables de données
- Construisez facilement des arbres en utilisant l'interface de paramétrage détaillé qui permet de définir les paramètres suivants:
  - Le niveau de mesure (nominal, ordinal, et continu)
  - Les variables indépendantes
  - Les variables dépendantes
  - Les variables d'influence
  - La méthode de construction de l'arbre
  - Les résultats à afficher: arbres, statistiques, graphiques et règles
  - Validation par partition ou validation croisée
  - Les critères d'arrêt
  - L'enregistrement des valeurs prédites, la probabilité et les modèles XML
- Choisissez parmi quatre méthodes de construction d'arbre
- Affichez la répartition de la variable cible pour chacun des nœuds selon la méthode de votre choix: histogrammes et/ou des tableaux de fréquence et %
- Refermez ou déployez des branches, et modifiez des aspects cosmétiques tels que les polices et les couleurs employées
- Affichez et imprimez les arbres
- Spécifiez le pourcentage d'agrandissement pour l'affichage de l'arbre dans l'interface
- Automatisez la construction d'arbres à l'aide du mode de production
  - Générez automatiquement une syntaxe à partir de l'interface
- Forcez l'introduction d'une variable dans le modèle
- Spécifiez des probabilités à priori, les coûts de mauvaise classification, et des scores de revenus, de dépenses et d'échelle

### Algorithmes de développement d'arbres

- Construisez votre modèle à l'aide de l'un de ces quatre algorithmes reconnus:
  - CHAID de Kass (1980)
  - Exhaustive CHAID de Biggs, de Ville et Suen (1991)
  - Arbres de classification et de régression (C&RT) de Breiman, Friedman, Olshen et Stone (1984)
  - QUEST de Loh et Shih (1997)
- Gérez les données des variables manquantes à l'aide de l'une de ces deux méthodes: Affectation à une catégorie ou imputation à l'aide d'un substitut
- Réalisez la discrétisation des variables continues en fonction du nombre de classes spécifié
- Fonctions d'élagage pour les méthodes C&RT et QUEST
- Pour la validation par partition, utilisez un échantillonnage aléatoire ou spécifiez une variable de scission

### Évaluation du modèle

- Générez des tableaux de risque et de classification
- Résumez les performances des nœuds à l'aide de graphiques et de tableaux d'évaluation qui vous permettront d'identifier les meilleurs segments:
  - Gains
  - Index (lift)
  - Réponse
  - Moyenne
  - Profit moyen
  - Retour sur investissements
- Partitionnez les données en jeu d'apprentissage et jeu de test pour vérifier la précision
- Affichez des graphiques synthétiques ou les règles de classification des nœuds sélectionnés, à l'aide de la fenêtre de résumé des nœuds

### Déploiement

- Exportez:
  - Les représentations des arbres, les graphiques au format: HTML, texte, Word/RTF, Excel et PDF
- Enregistrez les détails du modèle sous forme de variables dans le fichier de données
- Exportez sous la forme desyntaxe de commande Statistics, au format SQL ou sous forme de texte simple, les règles de décision qui définissent les affectations
- Exportez les arbres en tant que modèles XML pour les utiliser avec IBM SPSS Statistics Server et le portail Statistics
- Publiez les arbres sous forme d'images, et les tables sous forme statique ou interactive dans SPSS Statistics Portal
- Poursuivez l'analyse en sélectionnant dans votre ensemble de données les segments intéressants directement à partir des nœuds de l'arbre

### Configuration système requise

La configuration varie selon la plate-forme. Pour plus d'informations, consultez [ibm.com/spss/requirements](http://ibm.com/spss/requirements)

## À propos d'IBM Business Analytics

Les logiciels IBM Business Analytics aident les entreprises à mesurer, comprendre et anticiper leur performance financière et opérationnelle en transformant les données en décisions et les décisions en action. IBM propose à cet effet une gamme complète et unifiée d'applications d'aide à la décision, d'analyse prédictive avancée, de pilotage de la stratégie et des performances financières, de gouvernance, de gestion du risque et de la conformité.

Avec les logiciels IBM, les entreprises peuvent détecter les tendances, les schémas récurrents et les anomalies, comparer des scénarios de simulation, prédire les menaces et opportunités potentielles, identifier et gérer les risques principaux, et planifier, budgétiser et prévoir les ressources. Grâce aux puissantes fonctions analytiques dont ils disposent, nos clients à travers le monde sont à même de mieux comprendre, anticiper et maîtriser leurs résultats métier.

### Pour plus d'informations:

Pour plus d'informations ou pour contacter un interlocuteur IBM, visitez le site [ibm.com/software/fr/analytics/spss](https://ibm.com/software/fr/analytics/spss)

### Demander à être rappelé

Pour demander à être appelé ou pour poser une question, accédez au site [ibm.com/software/fr/analytics](https://ibm.com/software/fr/analytics)

Un interlocuteur IBM vous répondra dans les deux jours ouvrés.



---

**Compagnie IBM France**  
17 Avenue de l'Europe  
92275 Bois-Colombes Cedex  
France

Adresse de la page d'accueil IBM:  
**ibm.com**

IBM, le logo IBM, ibm.com et SPSS sont des marques d'International Business Machines Corp. déposées dans de nombreuses juridictions réparties dans le monde entier. Les autres noms de produit et de service peuvent être des marques d'IBM ou d'autres sociétés. La liste actualisée de toutes les marques d'IBM est disponible sur la page Web « Droits d'auteur et marques de commerce », à l'adresse : [ibm.com/legal/copytrade.shtml](http://ibm.com/legal/copytrade.shtml)

Microsoft, Windows, Windows NT et le logo Windows sont des marques de Microsoft Corporation aux États-Unis et/ou dans certains autres pays.

Le présent document (y compris les références en matière de devises OU de tarification qui excluent les taxes applicables) contient des informations qui étaient en vigueur et valides à la date de la première publication et qui peuvent être modifiées par IBM à tout moment. Toutes les offres mentionnées ne sont pas distribuées dans tous les pays où IBM exerce son activité.

LES INFORMATIONS DU PRÉSENT DOCUMENT SONT FOURNIES "EN L'ÉTAT" ET SANS GARANTIE EXPLICITE OU IMPLICITE D'AUCUNE SORTE. IBM DÉCLINE NOTAMMENT TOUTE RESPONSABILITÉ RELATIVE À CES INFORMATIONS EN CAS DE CONTREFAÇON AINSI QU'EN CAS DE DÉFAUT D'APTITUDE À L'EXÉCUTION D'UN TRAVAIL DONNÉ. Les produits IBM sont garantis conformément aux dispositions des contrats avec lesquels ils sont fournis.

© Copyright IBM Corporation 2012



Merci de recycler ce document