

STATISTICA

Version 10

Disponibilité : Novembre 2010

Nouvelles Fonctionnalités & Améliorations

Table des Matières

PERFORMANCES	1
INTÉGRATION et CONNEXION	1
SharePoint (Toute la Gamme)	1
Office 2010 (Toute la Gamme).....	2
OLAP (Toute la Gamme).....	2
<i>STATISTICA</i> Connecteur PI (Produit Complémentaire).....	3
REPRÉSENTATION DES DONNÉES	3
Introduction.....	3
Échelle Interactive	4
Navigation Interactive.....	4
Transparence	4
Droites de Référence.....	5
Édition Interactive du Texte.....	6
INTERFACE-UTILISATEUR, ERGONOMIE	6
Introduction.....	6
Rubans (Toute la Gamme)	6
Amélioration des Espaces de Travail (<i>STATISTICA</i> Data Miner et <i>STATISTICA</i> Text Miner)	6
Produits <i>STATISTICA</i> Complémentaires dans le Ruban (Toute la Gamme).....	7
STATISTIQUES	7
Simulation de Modèles (Toute la Gamme sauf <i>STATISTICA</i> Base).....	7
Modèles à Risques Proportionnels de Cox (Toute la Gamme sauf <i>STATISTICA</i> Base)	7

DATA MINING	8
Courbe de Profit (<i>STATISTICA</i> Data Miner)	8
Courbe ROC (<i>STATISTICA</i> Data Miner).....	8
Text Mining (<i>STATISTICA</i> Text Miner).....	8
Déploiement en Java et C# (<i>STATISTICA</i> Data Miner – Interface Directe avec les Bases de Données).....	9
STATISTICA ENTREPRISE.....	9
Ruban	9
Données Source	10
Migration de la Base de Données	10
Publier des Macros dans <i>STATISTICA Entreprise</i>	10
Noms dans la Base de Configuration <i>STATISTICA Entreprise</i>	10
<i>STATISTICA Entreprise Serveur</i> :	
Mise à Jour Automatique des Analyses de Contrôle Qualité	10
<i>STATISTICA Entreprise Serveur</i> : Balayage de Contrôle Qualité.....	11
STATISTICA MSP Multidimensionnelle en Temps Réel (Produit <i>STATISTICA MSPC Online</i>)	11
STATISTICA Saisie des Données par le Web (Produit <i>STATISTICA Saisie des Données par le Web</i>).....	12
STATISTICA Live Score (Produit <i>STATISTICA Live Score</i>)	12
STATISTICA Scorecard	13
AIDE et DOCUMENTATION (Toute la Gamme).....	13
Exemples du Modèle-Objet de <i>STATISTICA</i>	13

STATISTICA 10

PERFORMANCES

(Toute la Gamme)

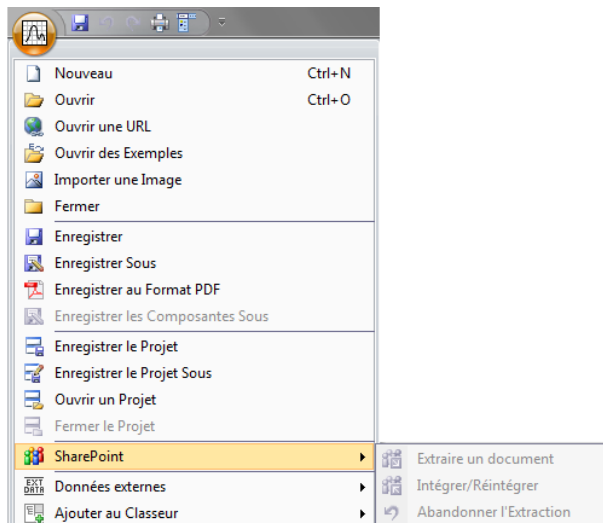
STATISTICA 10 apporte de nouvelles améliorations significatives en termes de performances, en tirant automatiquement parti de la technologie 64-bit (si elle est disponible sur l'ordinateur utilisé), ainsi que des fonctionnalités ultra-optimisées de multithreading. La plupart des fonctions analytiques et de gestion des données de *STATISTICA* (*C&RT*, *CHAID*, *Modèle Linéaire Général*, etc...), qui fonctionnaient en mode mono-threading dans la version 9, sont désormais optimisées par rapport à la technologie avancée du multithreading ce qui leur permet d'exploiter toute la puissance des processeurs ou cœurs multiples.

INTÉGRATION et CONNEXION

SharePoint

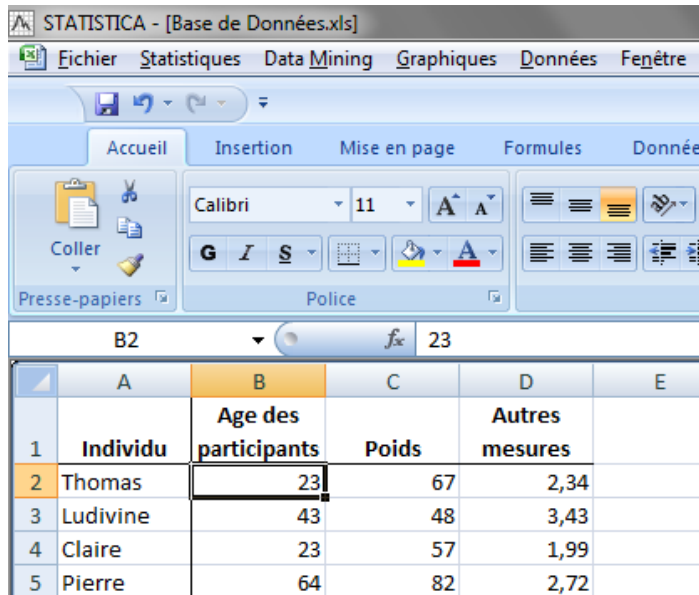
(Toute la Gamme)

Les entrées et résultats de *STATISTICA 10* sont désormais intégrés avec la plate-forme Microsoft SharePoint qui est en train de s'imposer comme le standard pour l'échange et l'intégration des données. Vous pouvez désormais facilement modifier et réintégrer tous vos documents *STATISTICA* dans SharePoint depuis l'interface-utilisateur de *STATISTICA*. À notre connaissance, *STATISTICA 10* est actuellement la seule application analytique ou de data mining offrant cette fonctionnalité (totalement intégrée).



Office 2010 (Toute la Gamme)

STATISTICA permet d'importer directement des fichiers natifs Office 2007 et 2010, en préservant la mise en forme. Cette nouvelle technologie a permis d'améliorer à la fois la rapidité et la tolérance aux fautes pour l'importation de fichiers Excel 2007 et 2010 dans des feuilles de données *STATISTICA* ; l'importation/exportation de fichiers Excel 2007/2010 permet dorénavant de gérer des cellules de texte mis en forme.

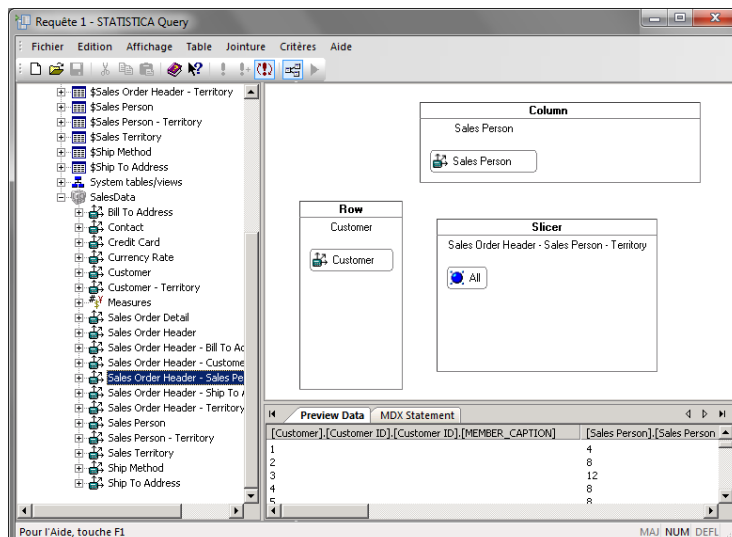


The screenshot shows the STATISTICA software interface with a spreadsheet containing the following data:

	A	B	C	D	E
		Age des participants		Autres mesures	
1	Individu		Poids		
2	Thomas	23	67	2,34	
3	Ludivine	43	48	3,43	
4	Claire	23	57	1,99	
5	Pierre	64	82	2,72	

OLAP (Toute la Gamme)

STATISTICA Query peut à présent récupérer des données issues de fournisseurs de cubes OLAP comme Microsoft OLE DB Provider for Analysis Services ou SAP Business Warehouse. Vous pouvez générer des requêtes MDX en environnement glisser-déplacer, ou saisir directement du code MDX.



The screenshot shows the STATISTICA Query interface with the following configuration:

- Column:** Sales Person
- Row:** Customer
- Slicer:** Sales Order Header - Sales Person - Territory

The Preview Data table shows the following data:

[Customer]	[Customer ID]	[Customer ID]	[MEMBER_CAPTION]	[Sales Person]	[Sales Person]
1				4	
2				8	
3				12	
4				8	
5				8	

STATISTICA Connecteur PI (Produit Complémentaire)

Il est maintenant plus simple d'installer et de gérer le Connecteur PI de *STATISTICA* dans la version 10 ; l'installateur du Connecteur PI est désormais totalement intégré, et ne requiert plus un installateur séparé.

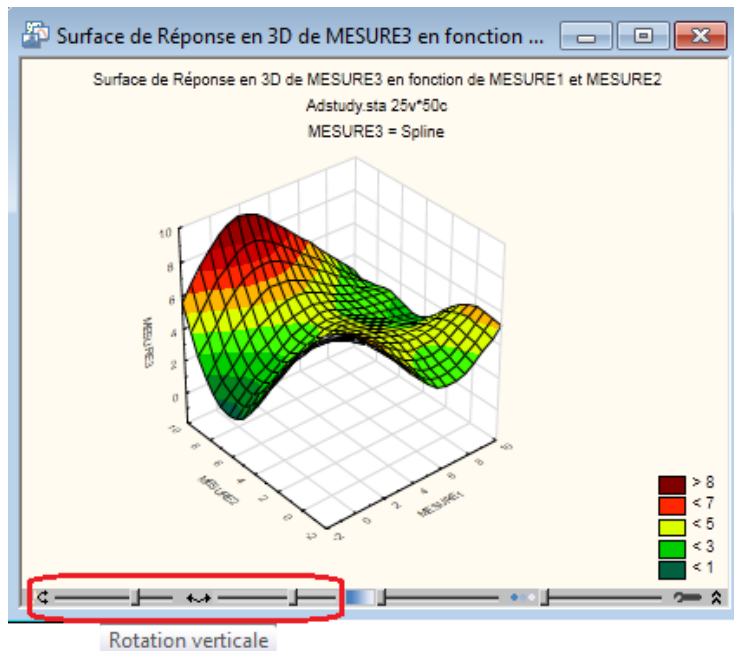
REPRÉSENTATION DES DONNÉES (Toute la Gamme)

Introduction

La technologie d'affichage des Graphiques *STATISTICA* a été substantiellement améliorée afin de détecter automatiquement et tirer parti des accélérations matérielles qui existent désormais dans les contrôleurs vidéo les plus performants, mais également dans la plupart des contrôleurs vidéo présents sur les ordinateurs de bureau et les ordinateurs portables.

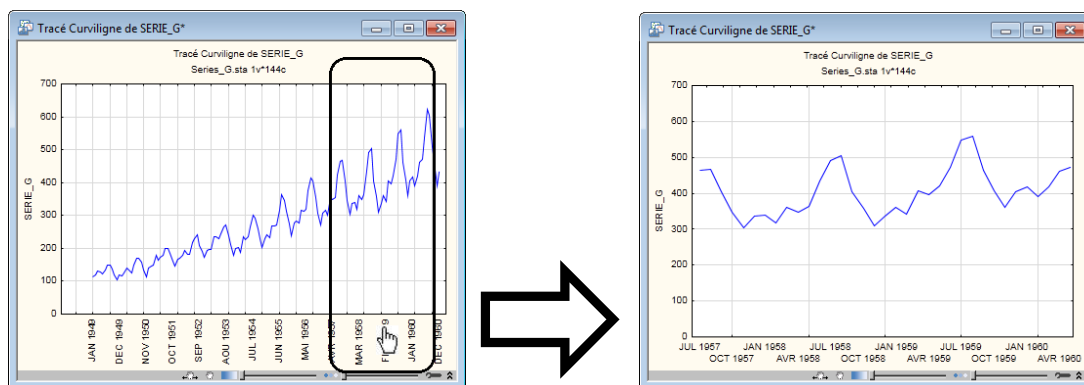
Les graphiques sont non seulement générés plus rapidement, mais permettent en outre des options d'affichage plus lisses et nuancées. Tous les Graphiques *STATISTICA* ont été améliorés avec un meilleur rendu, grâce aux nouvelles nuances/couleurs de remplissage et à des courbes (ou surfaces) plus lisses.

En outre, toutes les fenêtres de Graphiques *STATISTICA* (qu'il s'agisse de fenêtres individuelles ou de fenêtres intégrées dans des classeurs) offrent désormais des contrôles graphiques interactifs (une barre avec des ascenseurs et d'autres contrôles situés dans la partie inférieure de la fenêtre graphique), qui vous permettent d'ajuster de façon interactive ces nouvelles fonctionnalités d'affichage. Ceci se traduit par une représentation largement améliorée de tous les graphiques, mais également par de nouvelles options analytiques et exploratoires, comme la possibilité de révéler des tendances cachées par désaturation graduelle des affichages denses ou la rotation horizontale et verticale des graphiques en 3D.



Échelle Interactive

Vous pouvez désormais interagir directement sur l'échelle d'un graphique en positionnant le pointeur de la souris sur les libellés de l'axe à l'extrémité de l'axe, et en faisant glisser le pointeur vers la gauche ou vers la droite pour modifier l'échelle. L'Échelle Interactive constitue une technique exploratoire graphique puissante qui vous permet de mettre en évidence des tendances cachées en étirant ou en comprimant certaines parties de l'affichage.

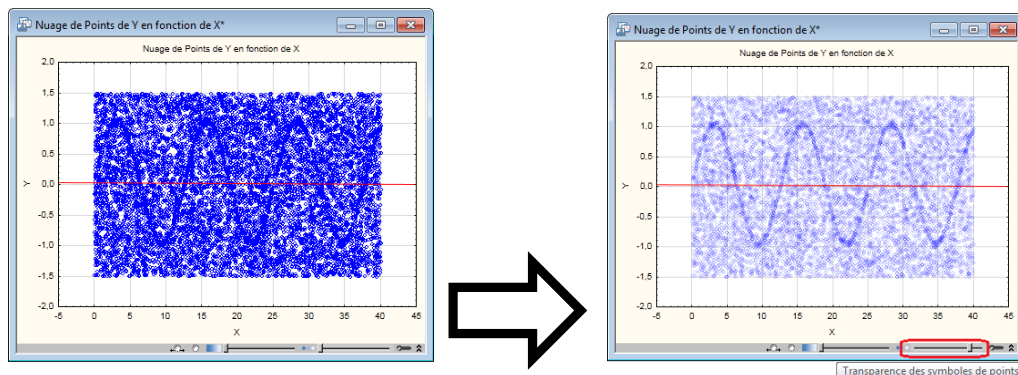


Navigation Interactive

Vous pouvez désormais interagir directement avec l'axe d'un graphique avec de naviguer vers la droite ou vers la gauche en positionnant le pointeur de la souris sur les libellés de l'axe vers le milieu de l'axe. La Navigation Interactive est une technique exploratoire graphique puissante qui vous permet d'explorer des tendances cachées dans les données.

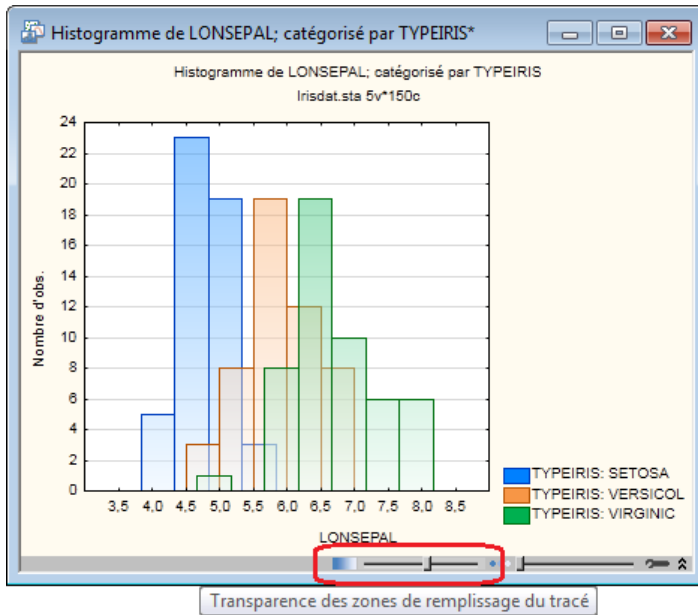
Transparence

STATISTICA 10 permet de contrôler le degré de transparence des symboles de points superposés grâce à des contrôles interactifs sur le graphique (cette fonctionnalité n'est compatible qu'à partir des systèmes d'exploitation Windows Vista SP 2 ou Windows 7). Le contrôle de la transparence est une technique exploratoire graphique puissante qui vous permet de mettre en évidence des tendances cachées par désaturation des zones fortement concentrées en symboles de points (en particulier, dans les nuages de points et dans les nuages de points matriciels réalisés sur de grosses volumétries de données).



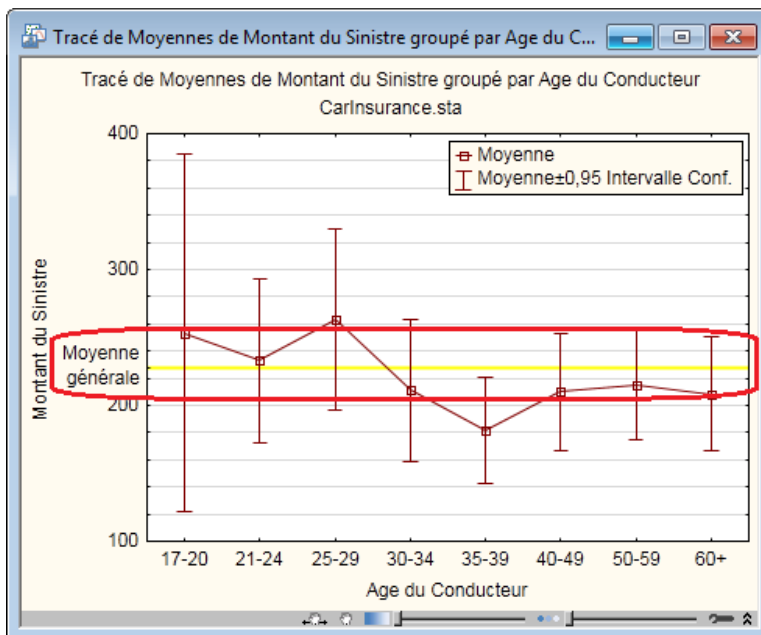
L'objectif est d'obtenir un niveau de densité optimal permettant de mettre en évidence les phénomènes structurels présents dans les données, mais cachés par un trop grand nombre de points aléatoires (bruit blanc).

En outre, le fait de rendre certaines zones du graphique transparentes permet à des portions du tracé de se chevaucher tout en restant visibles.



Droites de Référence

Vous pouvez ajouter des droites de référence dans les graphiques, beaucoup plus facilement dans *STATISTICA 10* grâce aux options dédiées de *Droites de Référence*, accessibles par la boîte de dialogue *Options du Graphique*.



Édition Interactive du Texte

Vous pouvez désormais modifier directement du texte à l'écran (en cliquant simplement et en saisissant les modifications), sans avoir à ouvrir la fenêtre d'édition. Les contrôles disponibles dans l'éditeur de texte du graphique restent accessibles et offrent des options d'édition avancées.

INTERFACE-UTILISATEUR, ERGONOMIE

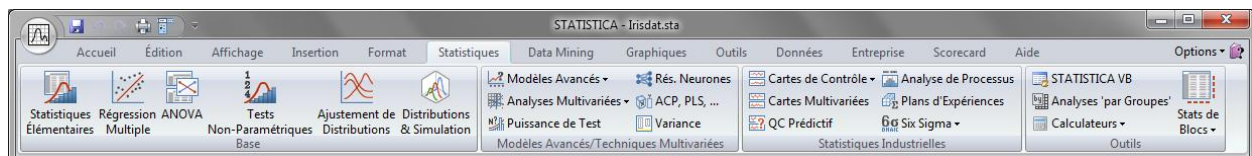
Introduction

De nombreuses améliorations pratiques ont été introduites dans *STATISTICA 10* pour renforcer l'intuitivité et le confort d'utilisation de l'application, mais aussi tirer parti des dernières avancées ergonomiques et scientifiques (1) dans le domaine de la réduction de la fatigue visuelle, et (2) pour une meilleure efficacité de l'interaction entre l'homme et la machine. *STATISTICA 10* offre une interface-utilisateur plus efficace grâce à une technologie d'affichage entièrement repensée et de nouvelles icônes.

Rubans

(Toute la Gamme)

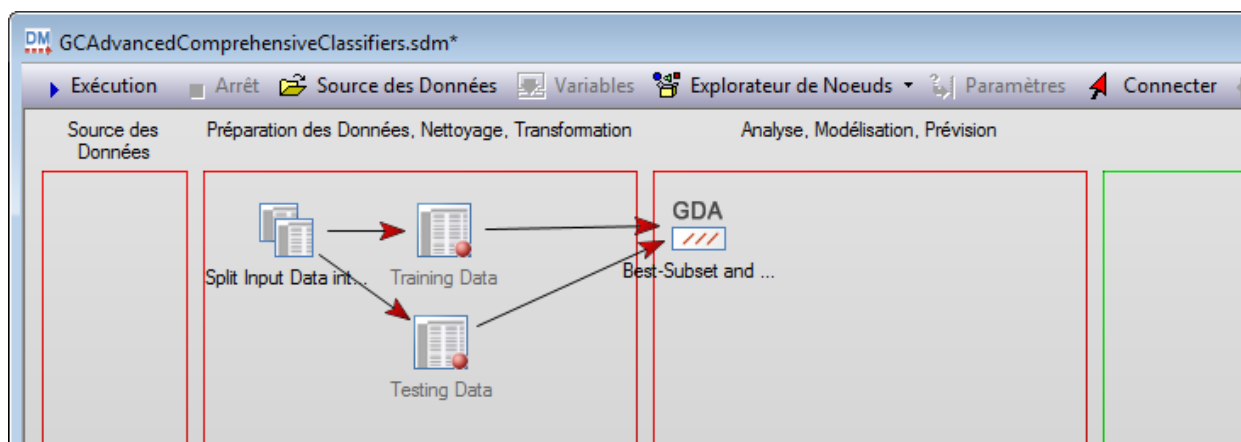
Tous les rubans ont été mis à jour et intègrent désormais des symboles entièrement recréés [l'interface-utilisateur traditionnelle avec des menus déroulants (menus classiques) reste proposée dans un souci de compatibilité]. Vous pouvez désormais ajouter des macros *STATISTICA Visual Basic* dans vos rubans *STATISTICA*.



Amélioration des Espaces de Travail

(STATISTICA Data Miner/STATISTICA Text Miner)

L'espace de travail de *STATISTICA Data Miner* offre désormais des icônes plus grandes (avec un aspect plus moderne). D'autres fonctionnalités ont également été ajoutées pour améliorer l'interface-utilisateur.



Produits **STATISTICA** Complémentaires dans le Ruban (Toute la Gamme)

Vous pouvez désormais contrôler le ruban *STATISTICA* par programmation. Les développeurs peuvent désormais personnaliser le ruban grâce à des appels API (Application Programming Interface). Cette fonctionnalité présente un intérêt majeur pour la création d'*Add-Ins* à *STATISTICA*.

STATISTIQUES

Simulation de Modèles

(Toute la Gamme sauf **STATISTICA Base**)

STATISTICA 10 vous permet de simuler plus facilement des données suivant une distribution spécifique grâce à l'option *Simulation de Modèles* du module *Distributions & Simulation*.

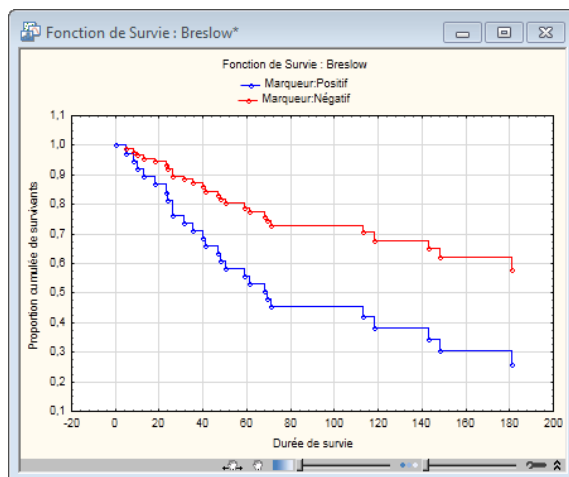
Modèles à Risques Proportionnels de Cox

(Toute la Gamme sauf **STATISTICA Base**)

Un nouveau module complet et entièrement paramétrable de *Modèles à Risques Proportionnels de Cox* (une technique puissante de modélisation des données de survie) a été introduit dans *STATISTICA 10*. Parmi les applications de ce nouveau module :

- Analyse des données de survie sur des patients dans les études médicales
- Analyse de l'attrition (perte de client) ou de *churn*
- Modélisation et défaillance de pièces mécaniques (fiabilité)

Le module *Modèles à Risques Proportionnels de Cox* vous permet de gérer facilement les données censurées, les prédicteurs catégoriels et des modèles comportant des interactions et/ou des effets imbriqués. Il utilise des techniques de construction de modèles tels que la régression pas-à-pas ou par recherche exhaustive du meilleur modèle. Vous pouvez déployer les fonctions de survie sur de nouvelles données grâce au module *STATISTICA Déploiement Rapide*.

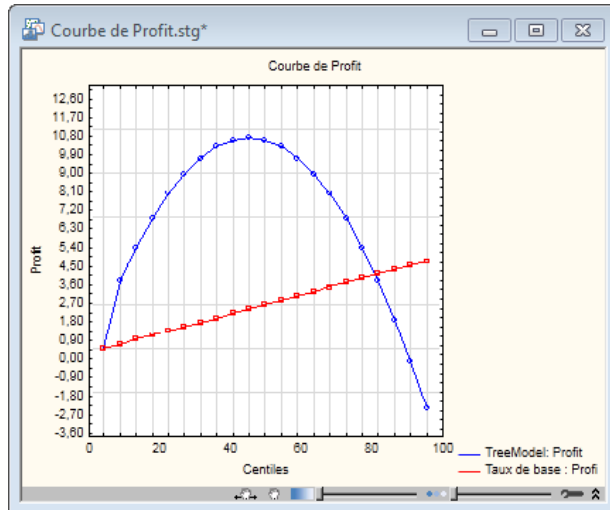


DATA MINING

Courbe de Profit

(*STATISTICA Data Miner*)

Vous pouvez produire des courbes de profit dans le module *Déploiement Rapide de Modèles Prédicatifs* de *STATISTICA*. La courbe de profit est une synthèse des coûts et bénéfices attendus du modèle courant, et constitue un outil supplémentaire dans de nombreuses applications de data mining pour comparer des modèles.



Courbe ROC

(*STATISTICA Data Miner*)

Vous pouvez produire des courbes ROC dans le module *Déploiement Rapide de Modèles Prédicatifs* de *STATISTICA*. Cet outil supplémentaire vous permet d'évaluer la qualité des modèles en représentant le taux de "vrais" positifs en fonction du taux de "faux" positifs. Les courbes ROC sont utilisées dans de nombreux domaines comme la médecine, le contrôle qualité ou la psychologie.

Remarque : Il est intéressant de rappeler que la méthode des courbes ROC trouve ses origines dans les applications militaires de détection de signaux radar, durant la Seconde Guerre Mondiale. Les opérateurs radar étaient alors évalués en fonction de leur capacité à distinguer les "vrais" signaux (avions) des "faux" signaux (oiseaux). Les courbes ROC sont aujourd'hui utilisées dans des applications de data mining pour des raisons similaires.

Text Mining

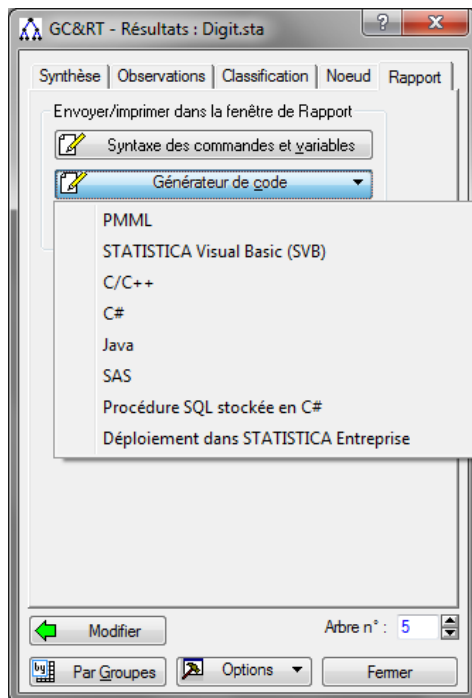
(*STATISTICA Text Miner*)

En réponse aux évolutions récentes dans le domaine du text mining, qui voit des volumétries considérables de données être analysées et modélisées, le moteur de calcul de *STATISTICA Text Miner* a été largement amélioré et optimisé pour lui permettre gérer ces grosses volumétries sans sacrifier la performance. La base de données interne qui gère les procédures a été remodelée et le module peut désormais gérer des jeux de données gigantesques de façon très efficace grâce à l'utilisation intensive du multithreading.

Déploiement en Java et C#

(STATISTICA Data Miner – Interface Directe avec les Bases de Données)

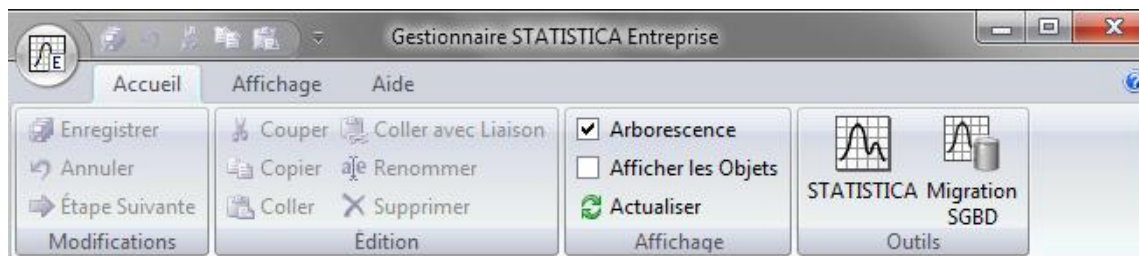
STATISTICA 10 offre deux nouvelles options de déploiement : Java et C#. Cette dernière option offre également la possibilité de générer du code C# sous une forme qui peut être directement intégrée dans une fonction personnalisée SQL Server, que vous pouvez ensuite utiliser dans une procédure stockée pour scorer le modèle directement à l'intérieur de la base de données. Le code Java peut être utilisé de la même manière dans des fonctions personnalisées Oracle. Remarque : ces options sont payantes. Le principal avantage de cette méthode de déploiement réside dans les gains de performances ; le déploiement interne sur la base de données peut être exécuté environ dix fois plus vite que le traitement externe.



STATISTICA ENTREPRISE

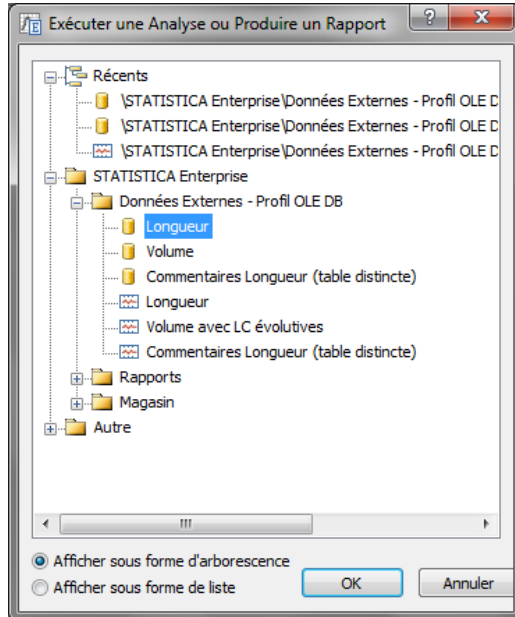
Ruban

La navigation dans le *Gestionnaire STATISTICA Entreprise* est désormais plus simple et rapide dans la version 10 grâce au nouveau ruban.



Données Source

Vous pouvez désormais sélectionner des Données Source dans l'Arborescence de *STATISTICA*, ce qui permet à l'utilisateur d'explorer les données sources dans l'interface-utilisateur de *STATISTICA*, sans avoir à utiliser le *Gestionnaire STATISTICA Entreprise*.



Migration de la Base de Données

L'outil de *Migration de Base de Données* a été mis à jour pour refléter le schéma de la base de données *STATISTICA Entreprise 10*. Il est à présent accessible directement depuis *STATISTICA Entreprise*. Il peut être exécuté par un administrateur pour copier des éléments de configuration d'une base de données vers une autre.

Publier des Macros dans *STATISTICA Entreprise*

STATISTICA 10 rend la publication de macros plus simple dans *STATISTICA Entreprise*. Cette méthode plus simple pour créer des *Analyses SVB*, fonctionne non seulement avec des programmes SVB mais également avec des scripts R. Pour accéder à cette nouvelle option, créez votre macro dans *STATISTICA*, puis ouvrez l'onglet *Entreprise* et cliquez sur *Déployer une Macro*.

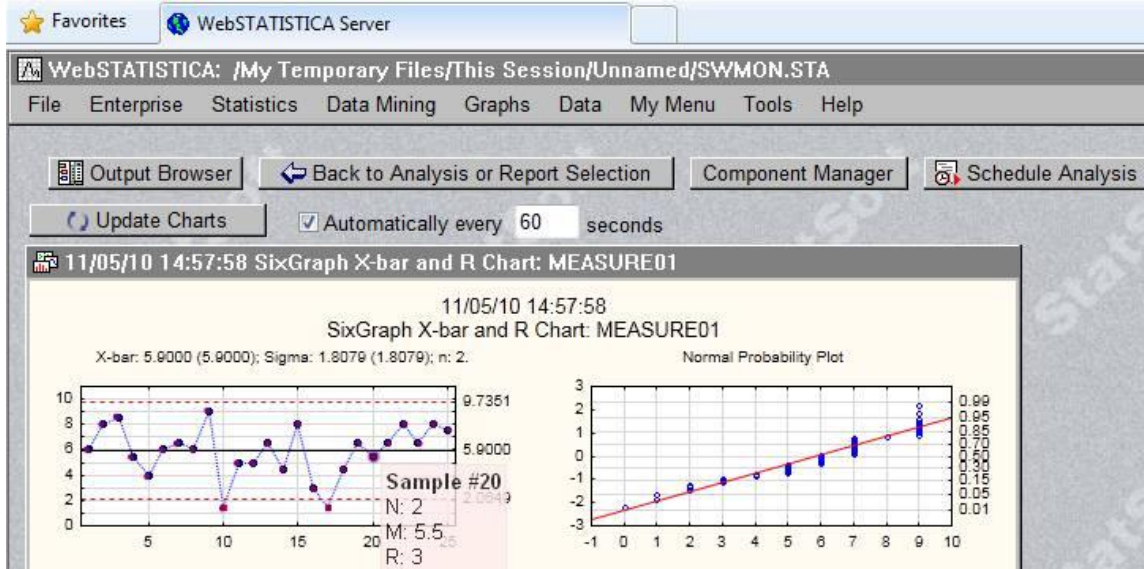
Noms dans la Base de Configuration *STATISTICA Entreprise*

Le *Gestionnaire STATISTICA Entreprise* vous offre davantage de souplesse pour définir les noms des objets de configuration dans *STATISTICA Entreprise*. Il faut juste que les noms soient différents à l'intérieur d'un même dossier de l'Arborescence.

STATISTICA Entreprise Serveur :

Mise à Jour Automatique des Analyses de Contrôle Qualité

Les Analyses de Contrôle Qualité peuvent se mettre à jour automatiquement dans un navigateur ; vous pouvez définir la fréquence de mise à jour, ou demander une mise à jour manuelle. Grâce aux dernières technologies web, il est possible de mettre à jour l'image du graphique sans avoir à recharger la page web (c'est-à-dire, sans "rafraîchissement" de la page web).

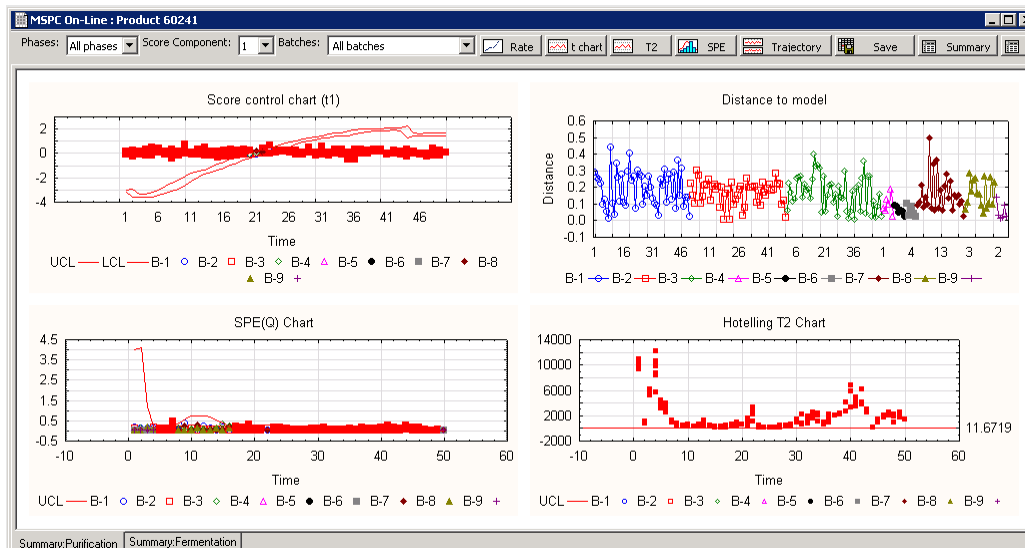


STATISTICA Entreprise Serveur : Balayage de Contrôle Qualité

Les Cartes de Contrôle Qualité permettent désormais le balayage interactif dans un navigateur Internet. Vous pouvez également affecter des Causes, Actions et Commentaires (et Inclure ou Exclure des points spécifiques) dans l'interface web. Grâce aux dernières technologies web mise en œuvre dans le programme, il est possible de mettre à jour l'image du graphique sans avoir à recharger la page web (c'est-à-dire, sans "rafraîchissement" de la page web).

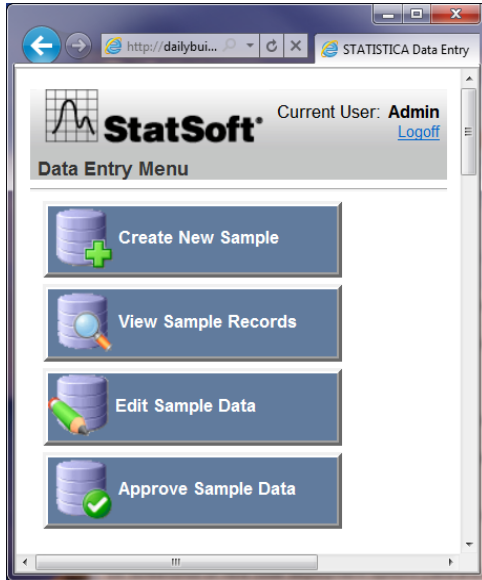
STATISTICA MSP Multidimensionnelle en Temps Réel (Produit STATISTICA MSPC Online)

Dans *STATISTICA 10*, l'option *STATISTICA MSPC Online* vous permet de déployer plus facilement des modèles d'analyse multidimensionnelle (*ACP*, *PLS*) dans *STATISTICA Entreprise* pour le suivi et la mise à jour en temps réel, le drill-down interactif des scores des composantes, le tracé des contributions et les cartes univariées.



STATISTICA Saisie des Données par le Web (Produit **STATISTICA Saisie des Données par le Web**)

STATISTICA Saisie des Données par le Web permet aux utilisateurs de définir leurs écrans de saisie de données dans une interface Web et de stocker/gérer ces données dans la base de données *STATISTICA Entreprise*.



STATISTICA 10 Saisie des Données par la Web offre de nombreuses améliorations, notamment :

- Facilité pour configurer les “champs requis”
- Navigation plus intuitive
- Organisation des champs en sections pour une saisie des données simplifiée
- Recherche des échantillons historiques grâce aux Libellés d’Échantillons
- Nouvelles options pour requêter les données à utiliser dans les analyses

STATISTICA Live Score (Produit **STATISTICA Live Score**)

Une version améliorée de *STATISTICA Live Score* accompagne la sortie de *STATISTICA 10*. *STATISTICA Live Score* intègre le logiciel *STATISTICA* Serveur dans la Plate-Forme d’Analyse des Données et de Data Mining de *STATISTICA*. Les données sont agrégées et nettoyées, puis les modèles sont entraînés et validés grâce à *STATISTICA Data Miner*. Lorsque les modèles sont validés, ils sont déployés sur le *Serveur STATISTICA Live Score*. *STATISTICA Live Score* constitue une application de scoring efficace, multi-threading et indépendante des systèmes d’exploitation pour une large gamme d’applications commerciales. Par exemple :

- Risque de crédit pour les dossiers de prêt (par exemple, systèmes de centres d’appel et dossiers en ligne)
- Segmentation de clientèle, vente additionnelle et montée en gamme, identification de l’attrition pour les départements marketing et de service-client
- Détection proactive de fraude

STATISTICA Scorecard (Produit **STATISTICA Scorecard**)

STATISTICA Scorecard est une solution dédiée pour développer, tester et suivre dans le temps des scores, avec toutes les étapes nécessaires, de la Sélection/Filtrage des Prédicteurs, à la Construction des Attributs, Construction des Modèles de Scoring, Sélection des Points de Rupture, Inférence de Rejet et jusqu'à la Stabilité de la Population.

AIDE et DOCUMENTATION (Toute la Gamme)

Exemples du Modèle-Objet de STATISTICA

Plusieurs centaines d'exemples *STATISTICA Visual Basic* ont été ajoutés à l'Aide de *STATISTICA 10*.

