

Ajouts et améliorations dans STATGRAPHICS Centurion 19

Statgraphics 19 contient de nombreux ajouts et améliorations. Des modifications importantes ont été apportées au logiciel dont une nouvelle interface pour l'utilisateur, 16 nouvelles procédures statistiques, des améliorations à 40 procédures statistiques déjà présentes et une interface avec Python. La plupart de ces modifications sont illustrées ci-dessous.

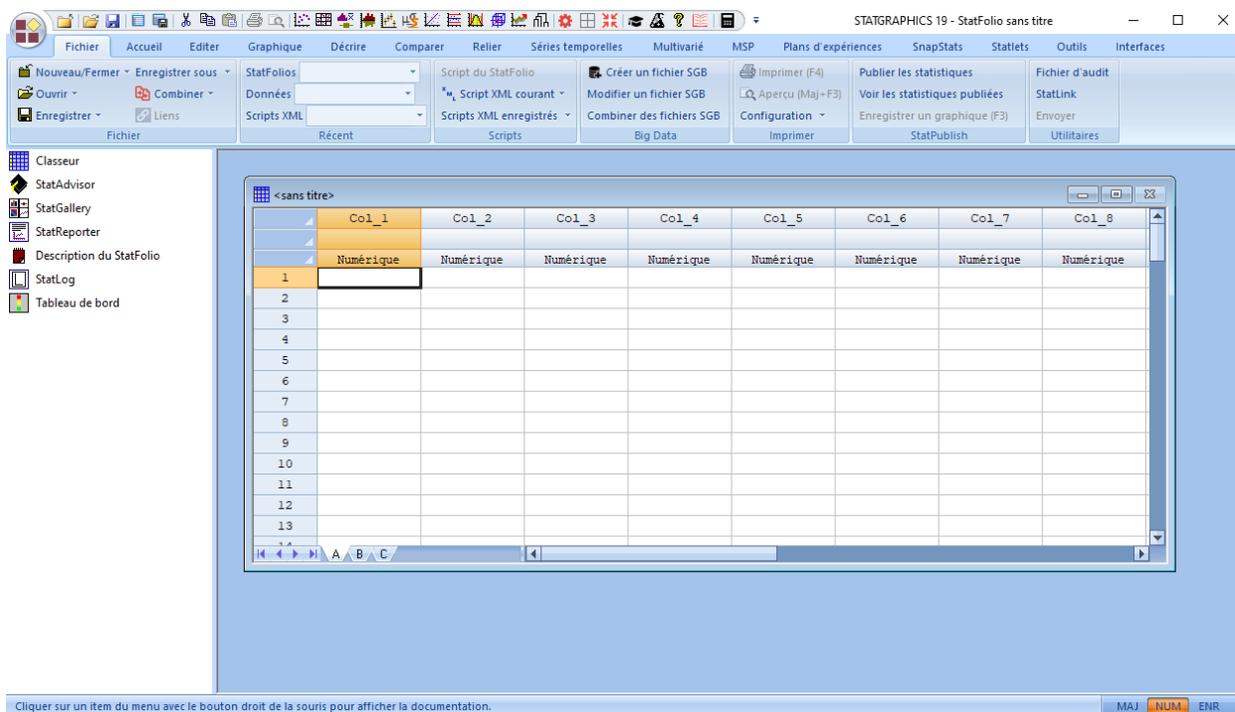
Ajouts et améliorations dans STATGRAPHICS Centurion 19.....	1
Interface Utilisateur	3
Rubans et barre d'accès rapide.....	3
Ruban des outils d'analyse.....	3
Affichage dans une unique fenêtre de la fenêtre d'analyse.....	4
Info-bulles pour les menus contextuels.....	5
Nouvelles boîtes de dialogue avec barres de navigation.....	6
Identification d'un point	7
Tableau de bord	8
En-têtes des lignes et des colonnes	9
Interface Python	10
Installation et configuration.....	10
Echange des données	11
Exécution de scripts Python.....	12
Plans d'expériences.....	13
Plans alias-optimaux	13
Augmentation optimale de plans existants	15
Régression et analyse de variance	17
Modèle linéaire général – Spécification du modèle et sélection pas à pas.....	17
Régression linéaire par morceaux.....	19
Régression quantile.....	20
Graphique de probabilités des résidus	21
Etudes de stabilité.....	22
Graphique des contributions en analyse de variance à facteurs imbriqués.....	23
Régressions de Poisson et binomiale négative à inflation de 0.....	24

Ajustement de lois	25
Mélanges gaussiens bivariés.....	25
Lois de Johnson.....	26
Mélanges gaussiens univariés.....	27
Lois de Poisson et binomiale négative à inflation de 0.....	28
Tests statistiques.....	29
Analyse de conformité.....	29
Tests d'équivalence et de non-infériorité – Comparer deux variances.....	30
Tests d'équivalence et de non-infériorité – Comparer une variance à une valeur cible.....	31
Test de Mann-Kendall.....	32
Test de Levene modifié.....	32
Etudes R&R non équilibrées avec facteurs additionnels optionnels.....	33
Test de Wald-Wolfowitz.....	33
Apprentissage machine.....	34
Forêts aléatoires décisionnelles.....	34
Classification par les K-moyennes.....	35
Visualisation de données.....	36
Diagramme en bâtons avec ligne avec ligne ajoutée.....	36
Graphique des données manquantes.....	37
Diagrammes de Venn et Euler.....	38
Diagrammes en cascade.....	39
Autres modifications.....	41
Modèles d'étalonnage.....	41
Classification.....	41
Cartes de contrôle.....	42
Classeur de données – Tri en ordre inversé.....	43
Classeur de données - Défaire.....	43
Importation des données.....	43
Graphique radar.....	44
Enregistrement des graphiques.....	45
StatGallery.....	45
Fond transparent dans les graphiques.....	45

Interface Utilisateur

Rubans et barre d'accès rapide

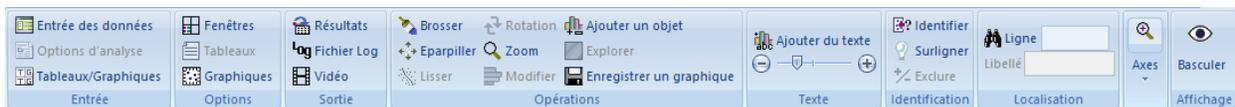
La modification la plus visible est la modernisation de l'interface Utilisateur.



Les anciens menus déroulants ont été remplacés par des rubans constitués de 16 catégories si le menu Standard est utilisé et de 17 catégories si le menu Six-Sigma est utilisé. Chaque catégorie contient plusieurs panneaux donnant accès aux options du logiciel et aux procédures statistiques. La barre d'accès rapide a été déplacée en haut de la fenêtre et donne un accès aisé aux procédures les plus courantes.

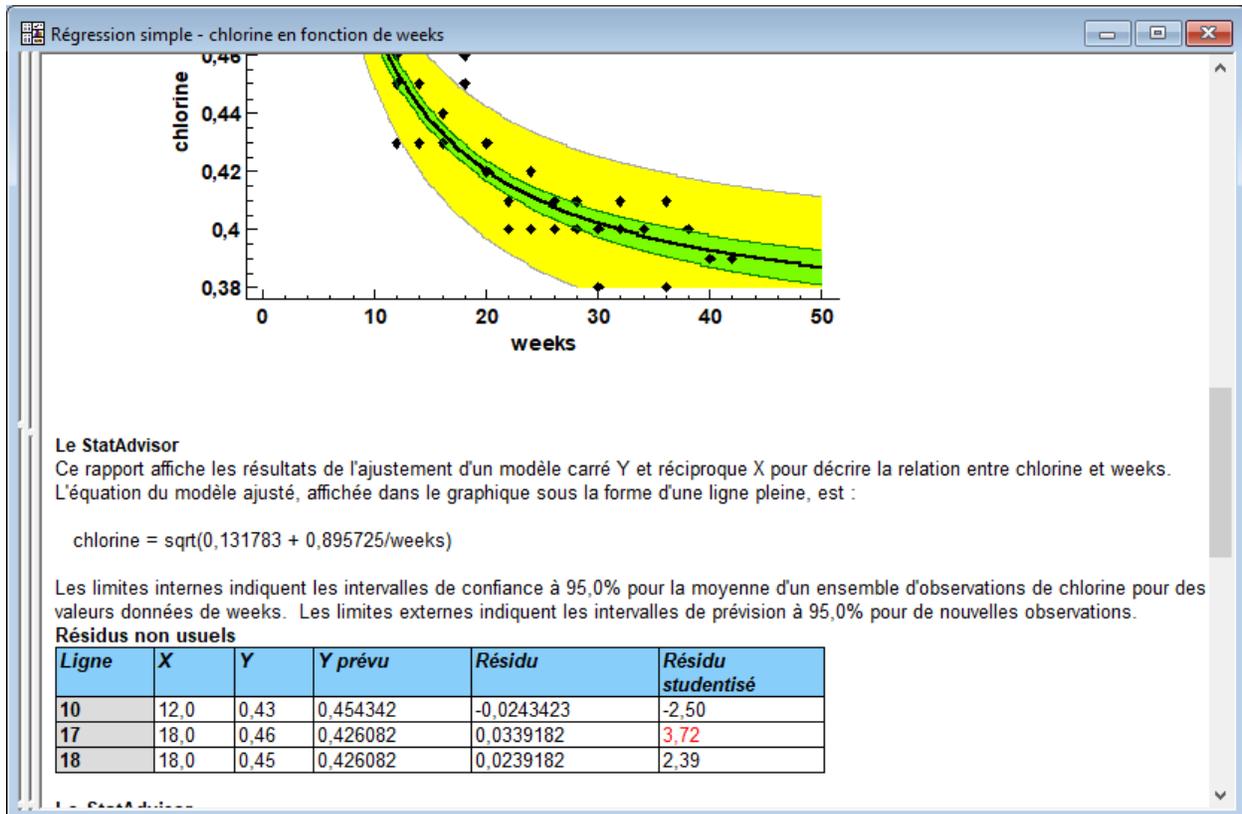
Ruban des outils d'analyse

Une fois une analyse exécutée, un ruban affiche des panneaux qui remplacent l'ancienne barre des outils d'analyse. Cela donne plus d'espace pour afficher la fenêtre d'analyse.



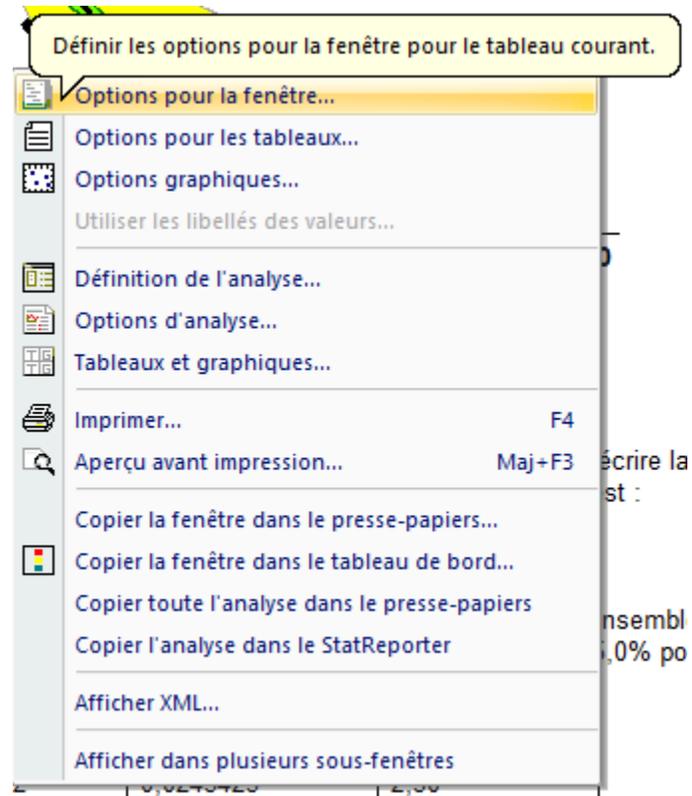
Affichage dans une unique fenêtre de la fenêtre d'analyse

La fenêtre d'analyse peut maintenant afficher de deux façons les tableaux et les graphiques : plusieurs sous-fenêtres comme dans les versions précédentes ou une unique fenêtre contenant tous les tableaux et tous les graphiques dans un format de type rapport. Le bouton *Basculer* dans *Affichage* permet de basculer d'un affichage à un autre.



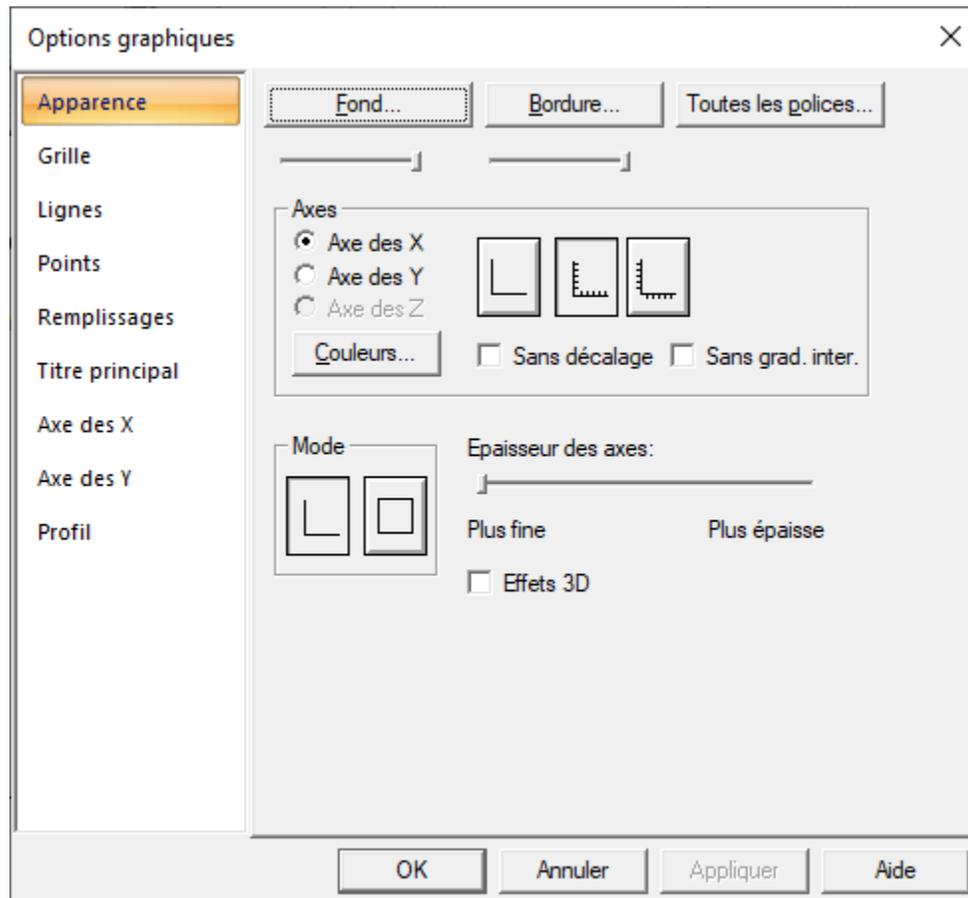
Info-bulles pour les menus contextuels

Les menus contextuels affichent maintenant des icônes et des info-bulles.



Nouvelles boîtes de dialogue avec barres de navigation

Les boîtes de dialogue à onglets utilisées pour définir les préférences du logiciel, les options graphiques et les alertes pour le StatFolio ont été remplacées par des boîtes de dialogue contenant une barre de navigation le long du côté gauche.



Identification d'un point

Lorsque l'utilisateur clique sur un point d'un graphique avec la souris, les coordonnées de ce point sont affichées ainsi que d'autres informations.

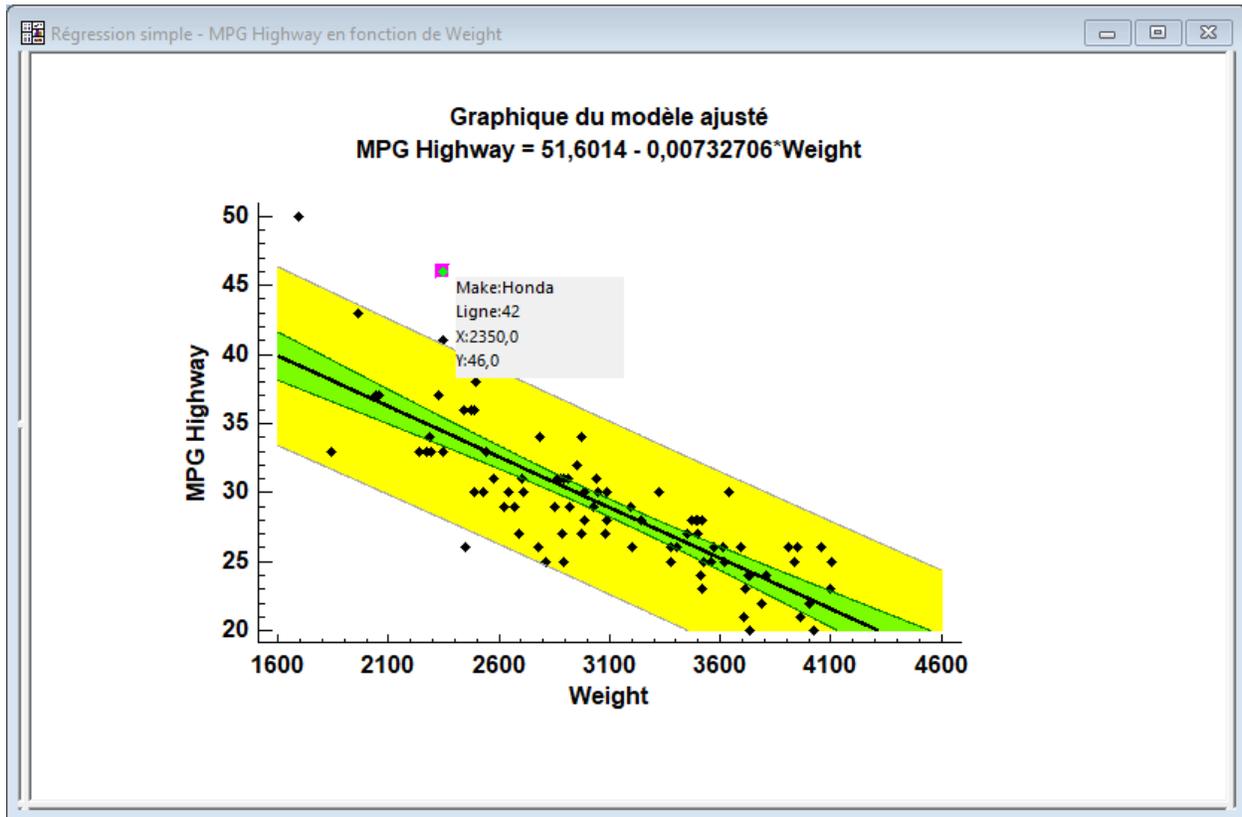
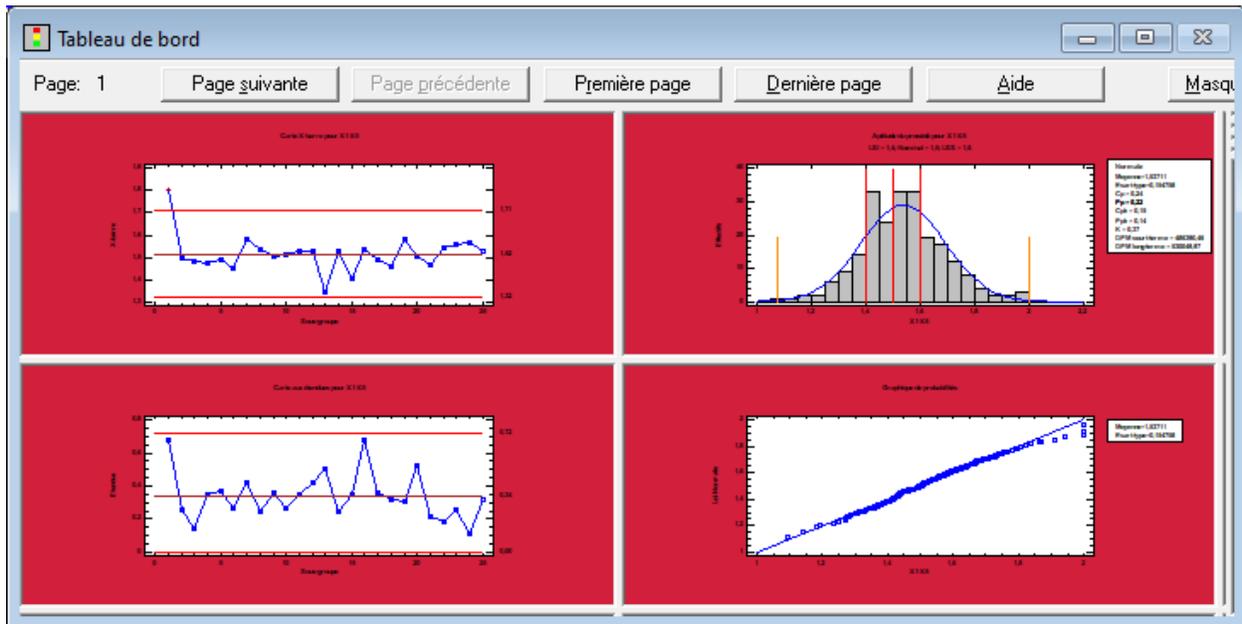


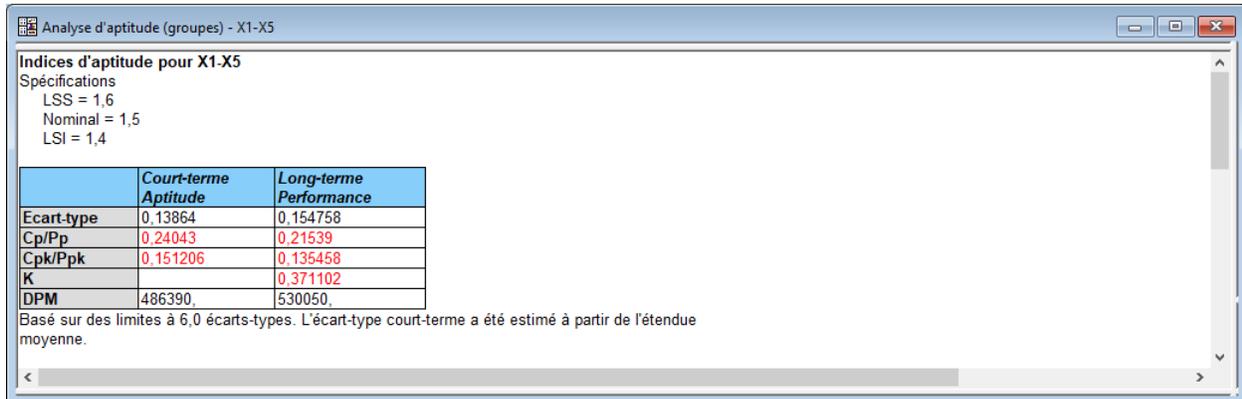
Tableau de bord

Une nouvelle fenêtre appelée *Tableau de bord* a été ajoutée aux StatFolios pour afficher des tableaux et des graphiques sélectionnés dans de multiples fenêtres d'analyse. Chaque tableau et graphique est lié à une fenêtre d'analyse et mis à jour immédiatement dès que l'analyse est modifiée. Dans des procédures comme les cartes de contrôle et les régressions, il est possible de définir la couleur de fond de chaque fenêtre dans le tableau de bord à vert, jaune ou rouge pour indiquer des alertes. Par exemple, un fond rouge dans une carte de contrôle peut indiquer qu'il y a un ou plusieurs points au-delà des limites de contrôle à 3 écarts-types.



En-têtes des lignes et des colonnes

Les utilisateurs peuvent choisir de colorer les en-têtes dignes et des colonnes des tableaux.

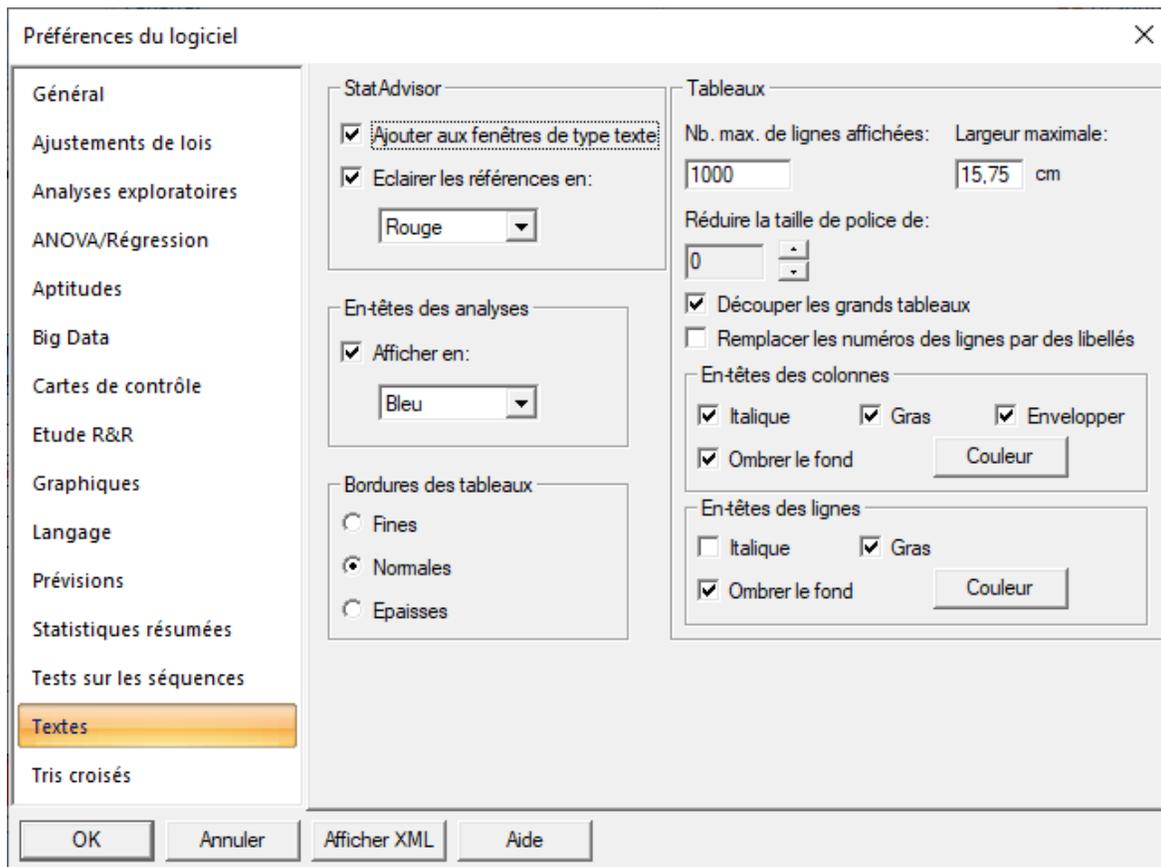


Indices d'aptitude pour X1-X5
 Spécifications
 LSS = 1,6
 Nominal = 1,5
 LSI = 1,4

	Court-terme Aptitude	Long-terme Performance
Ecart-type	0,13864	0,154758
Cp/Pp	0,24043	0,21539
Cpk/Ppk	0,151206	0,135458
K	0,371102	
DPM	486390,	530050,

Basé sur des limites à 6,0 écarts-types. L'écart-type court-terme a été estimé à partir de l'étendue moyenne.

L'épaisseur de la bordure des tableaux peut également être modifiée.



Préférences du logiciel

Général

Ajustements de lois

Analyses exploratoires

ANOVA/Régression

Aptitudes

Big Data

Cartes de contrôle

Etude R&R

Graphiques

Langage

Prévisions

Statistiques résumées

Tests sur les séquences

Textes

Tris croisés

StatAdvisor

Ajouter aux fenêtres de type texte

Eclairer les références en: Rouge

En-têtes des analyses

Afficher en: Bleu

Bordures des tableaux

Fines

Normales

Épaisses

Tableaux

Nb. max. de lignes affichées: 1000

Largeur maximale: 15,75 cm

Réduire la taille de police de: 0

Découper les grands tableaux

Remplacer les numéros des lignes par des libellés

En-têtes des colonnes

Italique Gras Envelopper

Ombrer le fond

En-têtes des lignes

Italique Gras

Ombrer le fond

OK Annuler Afficher XML Aide

Interface Python

Statgraphics 19 inclut une interface avec le langage de programmation Python. Cette interface est similaire à celle pour R introduite dans la version 18 de Statgraphics.

Installation et configuration

Statgraphics 19 contient un menu qui aide les utilisateurs à installer Python et les modules requis. Python est disponible gratuitement en téléchargement sur Internet.

Python – Installation et configuration ✕

1. Pour installer Python, cliquer sur un des liens 'Download' sur le site internet Python :

2. Après l'installation de Python, entrer le chemin vers python.exe dans le champ ci-dessous :

3. Définir le temps d'attente maximal pour l'exécution par Python d'un ensemble de commandes :
 secondes

4. Installer les modules Python pour les procédures que vous souhaitez utiliser en cliquant sur les boutons ci-dessous :

<input type="button" value="Lister les modules installés"/>	
<input type="button" value="Installer IPython, Jupyter, Numpy, Pandas, Scipy"/>	Nécessaire pour toutes les procédures.
<input type="button" value="Installer Matplotlib"/>	Nécessaire pour créer les graphiques.
<input type="button" value="Installer Scikit-learn"/>	Pour les algorithmes d'apprentissage machine.

Echange des données

Les données contenues dans le classeur de Statgraphics peuvent être aisément envoyées vers Python. Une boîte de dialogue est disponible pour préciser les variables contenant les données à envoyer et le nom du fichier CSV à utiliser pour contenir ces données. En cliquant sur *Exporter*, le fichier et les commandes Python nécessaires pour lire les données sont générés. La boîte de dialogue aide également l'utilisateur dans l'importation des résultats créés par Python en les stockant dans un data frame Python et en les chargeant dans une feuille du classeur de Statgraphics 19.

Interface avec Python – Echanger des données ✕

Charger Python

C:\Users\Christian\Documents\PYTHON\python.exe

Exporter des données dans Python

Etape 1: Préciser le data frame Python à créer :

Nom: Enregistrer les données de type texte en variables qualitatives

Etape 2: Sélectionner les colonnes à transférer (sauf si toutes).

Etape 3: Exporter les données dans un fichier CSV:

C:\Users\CHRIST~1\AppData\Local\Temp\statgraphics_data.csv

Etape 4: Entrer les commandes suivantes dans la fenêtre Python :

Etape 5: Pour afficher les données, entrer la commande suivante dans la fenêtre Python :

Importer des données de Python

Etape 1: Préciser le data frame Python à importer :

Nom:

Etape 2: Préciser le fichier temporaire à créer:

Nom du fichier:

Etape 3: Entrer la commande suivante dans la fenêtre Python :

Etape 4: Importer les données dans la feuille de données Statgraphics:

Feuille: A B C D E F G H I J K L M
 N O P Q R S T U V W X Y Z Supprimer les données présentes

Exécution de scripts Python

Comme avec R, les utilisateurs de Statgraphics 19 peuvent créer des scripts Python, les enregistrer dans un StatFolio et plus tard les exécuter. Les données peuvent être transférées à Python avant l'exécution du script et les résultats importés après l'exécution des commandes. Le script montré ci-dessous envoie les données à Python, exécute une classification par la méthodes des K-moyennes et transfère les numéros des classes obtenues à Statgraphics.

Interface avec Python – Options d'exécution d'un script ✕

Chemin vers Python:

Données exportées

Data frame Python à créer: Enregistrer les données de type texte en variables qualitatives
 Supprimer les lignes non sélectionnées

Script Python

Largeur des graphiques: pouces Hauteur des graphiques: pouces Attente maximale: secondes

Commandes Python

```

from sklearn.cluster import KMeans as km
kmeans=km(n_clusters=3).fit(data)
clusters=kmeans.predict(data)
clusters
results=pandas.DataFrame(clusters,columns=['cluster'])
results.info()
results.to_csv(r'C:\temp\Python_results.csv',index=False)

```

Données importées

Fichier CSV à importer (s'il y en a un):

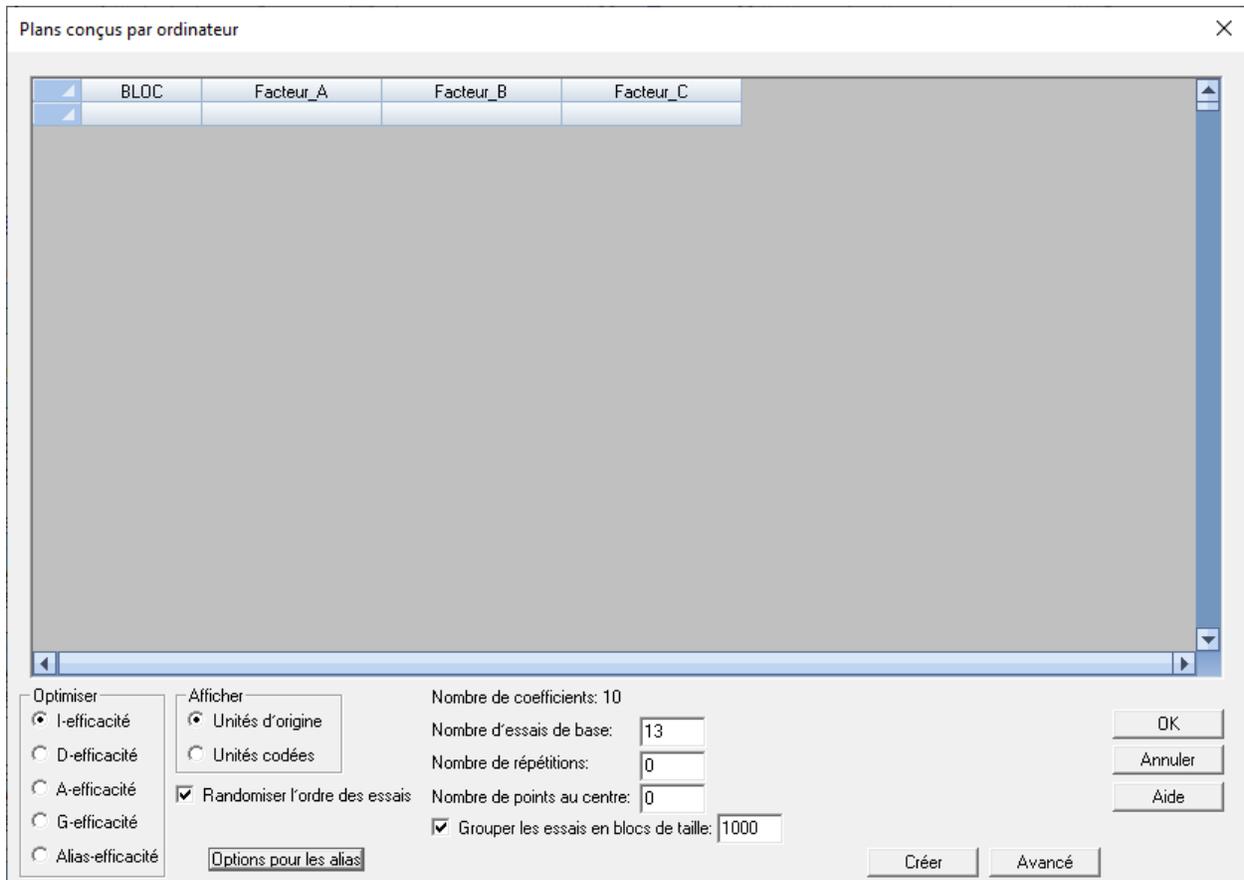
Feuille: A B C D E F G H I J K L M Supprimer les données présentes
 N O P Q R S T U V W X Y Z

Plans d'expériences

Deux fonctionnalités significatives ont été ajoutées à l'assistant pour les plans d'expériences : la possibilité de créer des 'Plans alias-optimaux' et d'améliorer des plans déjà existants en ajoutant des essais additionnels optimaux.

Plans alias-optimaux

Les plans d'expériences construits pour optimiser un critère comme la D-efficacité sont d'excellents choix lorsque les facteurs et les interactions sont connus à l'avance. Dans les cas où des termes non présents dans le modèle peuvent avoir des effets significatifs sur les réponses, il est souvent avantageux d'accepter une petite diminution de la D-efficacité pour réduire la structure des alias si ces effets sont actifs. Statgraphics 19 ajoute les 'Plans alias-optimaux' à l'assistant pour les plans d'expériences. Pour construire un tel plan, l'utilisateur doit sélectionner 'Plans conçus par ordinateur' et préciser le modèle statistique supposé. A l'étape 5, l'option pour optimiser les alias 'Options pour les alias' est alors sélectionnée :



Plans conçus par ordinateur

BLOC	Facteur_A	Facteur_B	Facteur_C

Optimiser

- I-efficacité
- D-efficacité
- A-efficacité
- G-efficacité
- Alias-efficacité

Afficher

- Unités d'origine
- Unités codées

Randomiser l'ordre des essais

Options pour les alias

Nombre de coefficients: 10

Nombre d'essais de base:

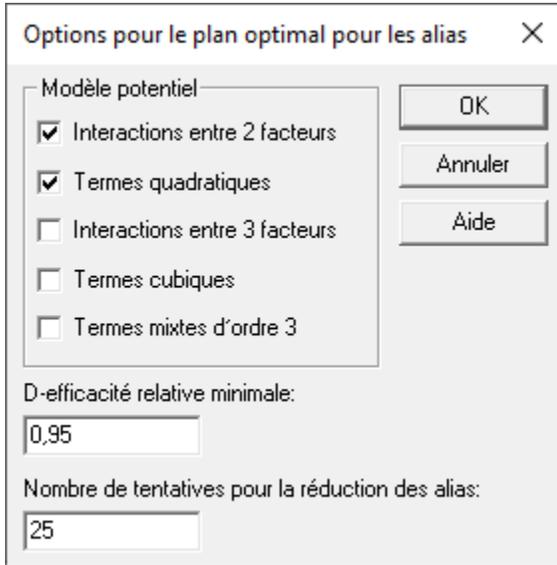
Nombre de répétitions:

Nombre de points au centre:

Grouper les essais en blocs de taille:

Créer Avancé OK Annuler Aide

En cliquant sur le bouton 'Options pour les alias', l'utilisateur peut préciser les termes non présents dans l'actuel modèle dont la structure d'alias doit être minimisée :



Options pour le plan optimal pour les alias

Modèle potentiel

- Interactions entre 2 facteurs
- Termes quadratiques
- Interactions entre 3 facteurs
- Termes cubiques
- Termes mixtes d'ordre 3

D-efficacité relative minimale:

Nombre de tentatives pour la réduction des alias:

OK
 Annuler
 Aide

L'utilisateur peut préciser la D-efficacité relative minimale acceptée pour réduire la structure d'alias.

Augmentation optimale de plans existants

L'étape 'Augmenter le plan' de l'assistant pour les plans d'expériences permet à l'utilisateur d'ajouter des essais dans un plan déjà existant pour augmenter l'efficacité du plan augmenté.

Assistant pour les plans d'expériences - Augmenter le plan

BLOCK	feed rate liters/min	catalyst %	agitation rpm	temperature degrees	concentration %
1	12,5	1,5	110,0	160,0	4,5
2	10,0	1,0	100,0	140,0	6,0
3	15,0	1,0	100,0	140,0	3,0
4	10,0	2,0	100,0	140,0	3,0
5	15,0	2,0	100,0	140,0	6,0
6	10,0	1,0	120,0	140,0	3,0
7	15,0	1,0	120,0	140,0	6,0
8	10,0	2,0	120,0	140,0	6,0
9	15,0	2,0	120,0	140,0	3,0
10	12,5	1,5	110,0	160,0	4,5
11	10,0	1,0	100,0	180,0	3,0
12	15,0	1,0	100,0	180,0	6,0
13	10,0	2,0	100,0	180,0	6,0

Action

Répliquer le plan

Ajouter une fraction

Clarifier les effets directs

Clarifier un facteur

Ajout de points en étoile

Optimiser Ajouter essais Options de recherche Nouveau modèle

Nombre total d'essais: 19
 Nombre total de blocs: 1
 Coefficients dans le modèle: 16

OK Annuler Défaire Aide

L'utilisateur précise le nombre d'essais à ajouter puis clique sur 'Nouveau modèle' pour préciser le modèle à ajuster dans le plan augmenté :

Options pour le modèle du plan

Facteurs procédé

- Moyenne
- Linéaire (effets directs)
- Interactions d'ordre 2
- Quadratique
- Cubique

Composants du mélange

- Moyenne
- Linéaire
- Quadratique
- Spécial-cubique
- Cubique

Inclure:

- A: feed rate
- B: catalyst
- C: agitation
- D: temperature
- E: concentration
- AA
- AB
- AC
- AD
- AE
- BB
- BC
- BD
- BE

Exclure:

OK

Annuler

Aide

Le bouton 'Options de recherche' permet de préciser le critère utilisé pour sélectionner l'ensemble optimal d'essais additionnels :

Options pour le plan augmenté par ordinateur

Optimiser

- I-efficacité
- D-efficacité
- A-efficacité
- G-efficacité

Nombre de niveaux à considérer pour les facteurs continus :

5

Incréments entre les niveaux des composants de mélange :

Créer un nouveau bloc

Nombre de départs aléatoires :

100

Nombre maximum d'itérations par départ :

100

OK

Annuler

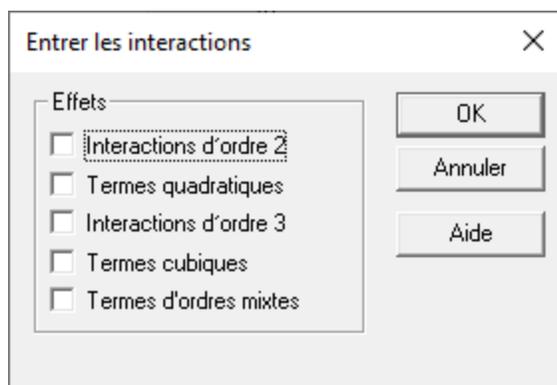
Aide

Régression et analyse de variance

Plusieurs nouvelles procédures et fonctionnalités ont été ajoutées pour les ajustements de modèles de régression et pour les calculs des analyses de variance.

Modèle linéaire général – Spécification du modèle et sélection pas à pas

La procédure ‘Modèle linéaire général’ (GLM) a été enrichie : définition plus aisée des modèles contenant des interactions et des termes polynomiaux, sélection pas à pas des variables pour les modèles contenant des facteurs quantitatifs et qualitatifs. Lors de la définition des modèles, une nouvelle boîte de dialogue a été conçue pour préciser les termes à ajouter dans la boîte de dialogue de spécification du modèle :



Cela diminue grandement le temps nécessaire pour définir de tels modèles.

Également, la boîte de dialogue des options d’analyse permet à l’utilisateur de mettre en œuvre une sélection pas à pas des facteurs :

Options pour le modèle linéaire général

Sommes des carrés

Type I

Type III

Affichage: MPG Highway

Modèle avec constante

Inclure MANOVA

Facteur: Domestic

Niveaux du facteur (glisser pour modifier l'ordre):

0

1

Transformation Box-Cox

Puissance (Lambda1): 1,0

Décalage (Lambda2): 0,0

Optimiser

Sélection pas à pas des variables

Aucune

Ascendante

Descendante

P en entrée: 0,05

P en sortie: 0,05

Nb étapes max.: 100

Afficher chaque étape

Conserver les effets d'ordres inférieurs

Facteur:

A

B

C

D

Terme d'erreur:

Automatique

Aucun

Résidu

A

B

C

D

Sélections:

A - Automatique

B - Automatique

C - Automatique

D - Automatique

OK

Annuler

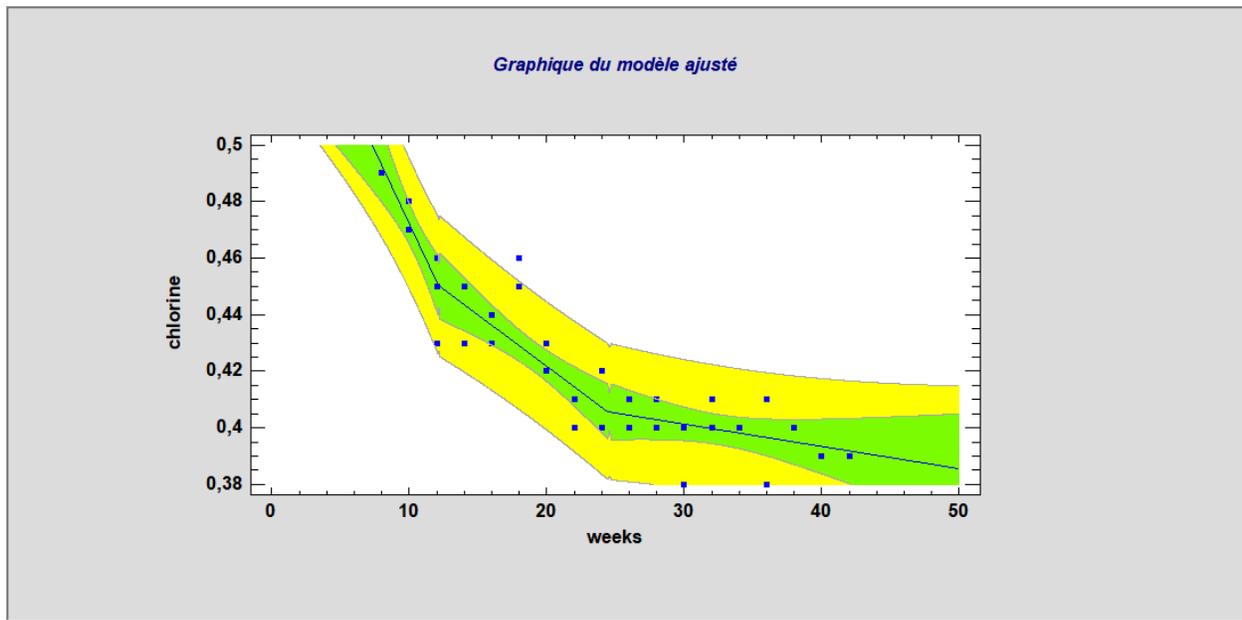
Aide

Deux fonctionnalités sont particulièrement importantes :

1. Lorsqu'une variable qualitative est ajoutée ou retirée du modèle, toutes les indicatrices pour ce facteur sont traitées comme un groupe.
2. La case à cocher 'Conserver les effets d'ordres inférieurs' permet de respecter la hiérarchie des termes dans le modèle. Par exemple, il ne sera pas possible de définir un modèle contenant l'interaction AB interaction si les effets principaux A et B ne sont pas inclus.

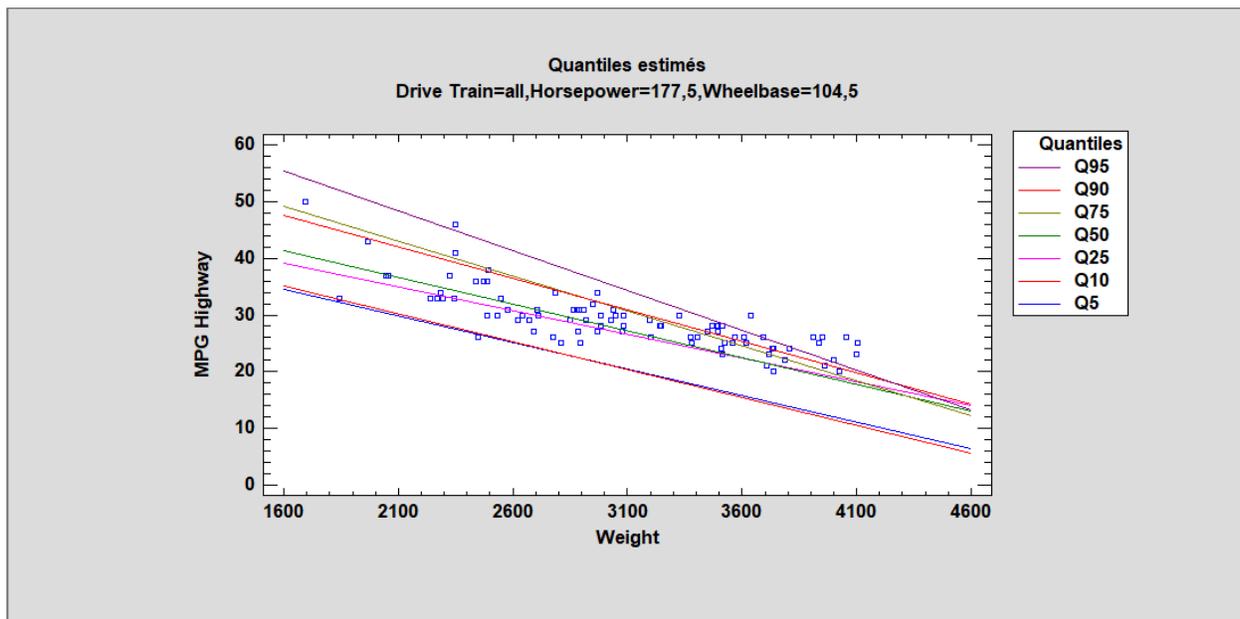
Régression linéaire par morceaux

La procédure 'Régression linéaire par morceaux' est conçue pour ajuster un modèle de régression dans lequel la relation entre la variable à expliquer Y et la variable explicative X est une fonction continue constituée de 2 ou plus morceaux linéaires. La fonction est estimée en utilisant les moindres carrés non linéaires. L'utilisateur précise le nombre de morceaux et les positions initiales estimées où les morceaux se rejoignent.



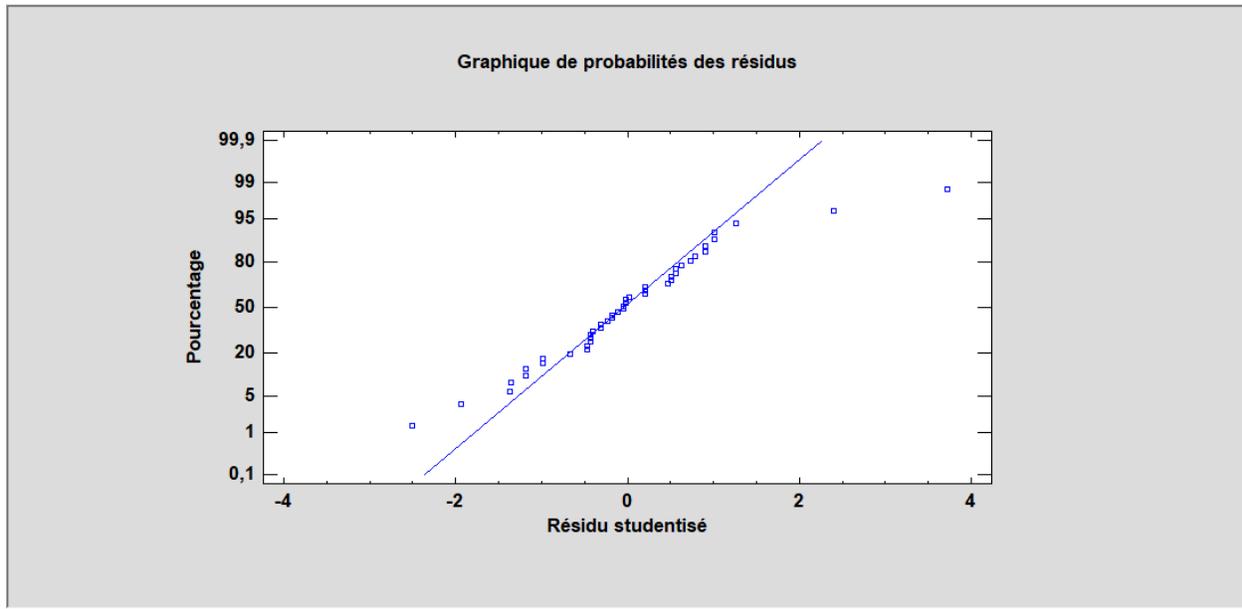
Régression quantile

La procédure Régression quantile ajuste des modèles linéaires pour décrire la relation entre des quantiles sélectionnés de la variable à expliquer Y et une ou plusieurs variables explicatives. Les variables explicatives peuvent être quantitatives ou qualitatives. A la différence des procédures classiques de régression multiple dans lesquelles le modèle est utilisé pour prévoir la réponse moyenne, les modèles de régression quantile peuvent être utilisés pour prévoir tout centile. La régression médiane est un cas spécial où le quantile à prévoir est le 50ème centile. La régression quantile est fréquemment utilisée en écologie et en économétrie.



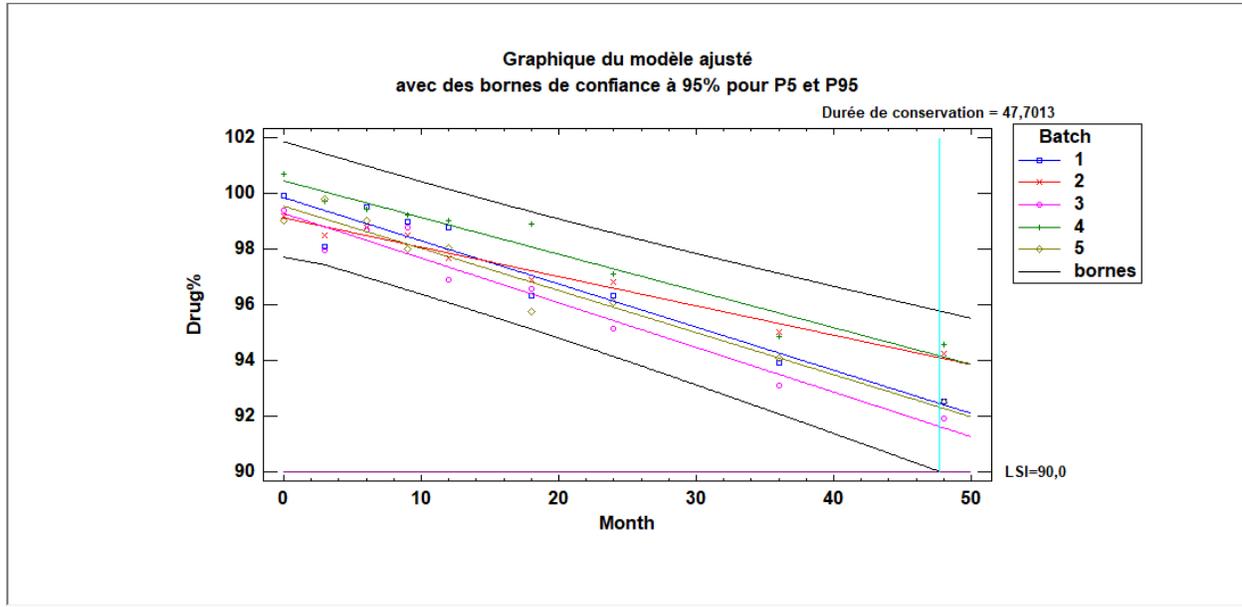
Graphique de probabilités des résidus

Ce graphique a été ajouté comme une des options graphiques dans de nombreuses procédures de modélisation.



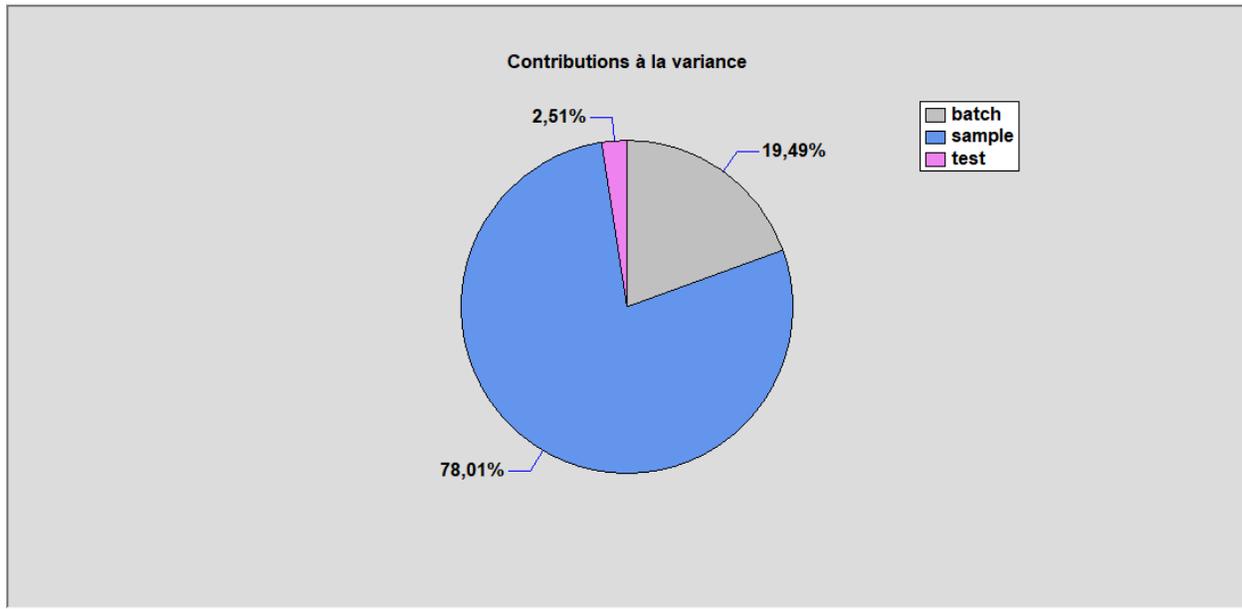
Etudes de stabilité

Les études de stabilité sont couramment utilisées par les industries pharmaceutique et alimentaire pour estimer le taux de dégradation d'un produit périssable et pour établir une durée de conservation. Des mesures sont faites sur des échantillons provenant de plusieurs lots à différents instants. L'intérêt premier est d'estimer l'instant auquel la limite inférieure de prévision du modèle de dégradation intercepte la limite inférieure de spécification pour le produit.



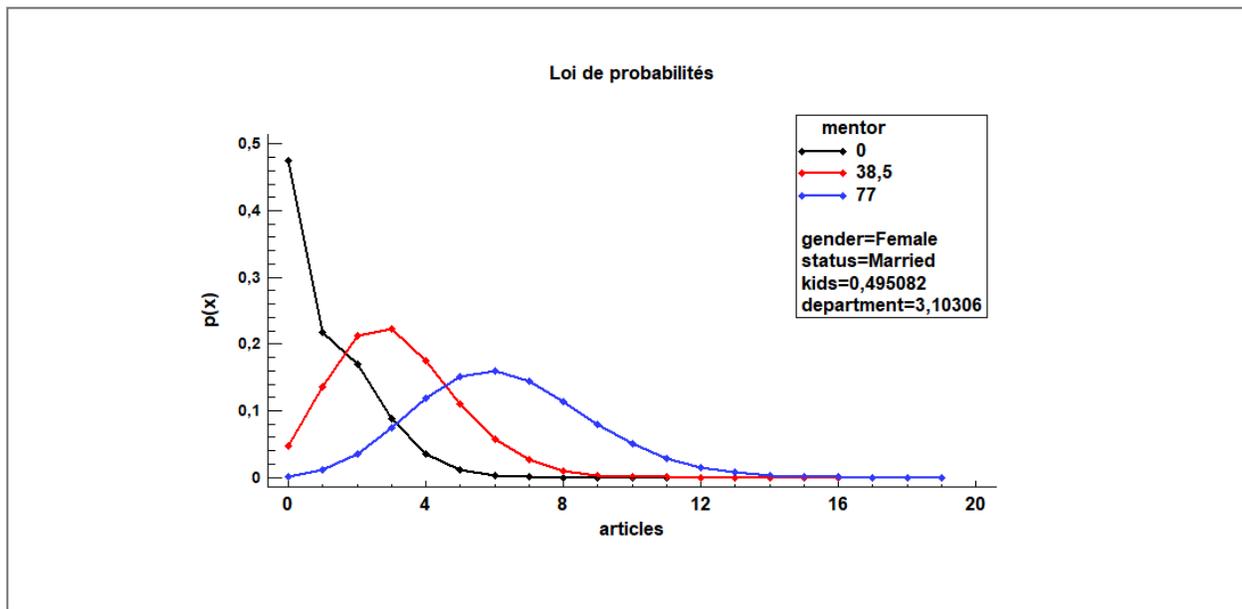
Graphique des contributions en analyse de variance à facteurs imbriqués

Ce graphique affiche les contributions en pourcentages de chacun des composants de la variance à la variance totale.



Régressions de Poisson et binomiale négative à inflation de 0

Dans la vie réelle, la loi d'une variable discrète contient souvent plus de zéros que les lois de Poisson ou binomiale négative ne prévoient. Dans de tels cas, un terme additionnel prenant en compte cette structure des zéros supplémentaires est ajouté à la loi. Cela est important non seulement lors de l'ajustement de la loi à un jeu de données distribuées de façon identique mais également lorsqu'une telle variable est utilisée comme variable à expliquer dans un modèle de régression. Dans le graphique ci-dessous, le terme mentor = 0 prend en compte ces zéros supplémentaires.



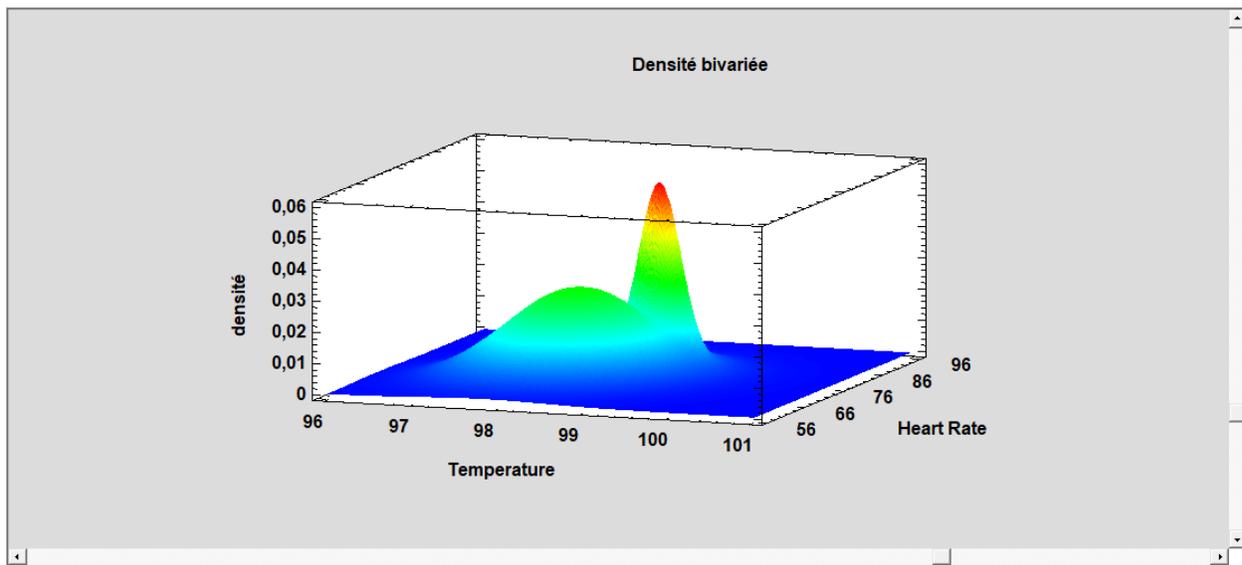
Ajustement de lois

Des améliorations ont été apportées aux procédures d'ajustement de lois, dont la possibilité d'ajuster :

1. Des mélanges gaussiens bivariés
2. Des lois de Johnson
3. Des mélanges gaussiens univariés
4. Les lois de Poisson et binomiale négative à inflation de 0

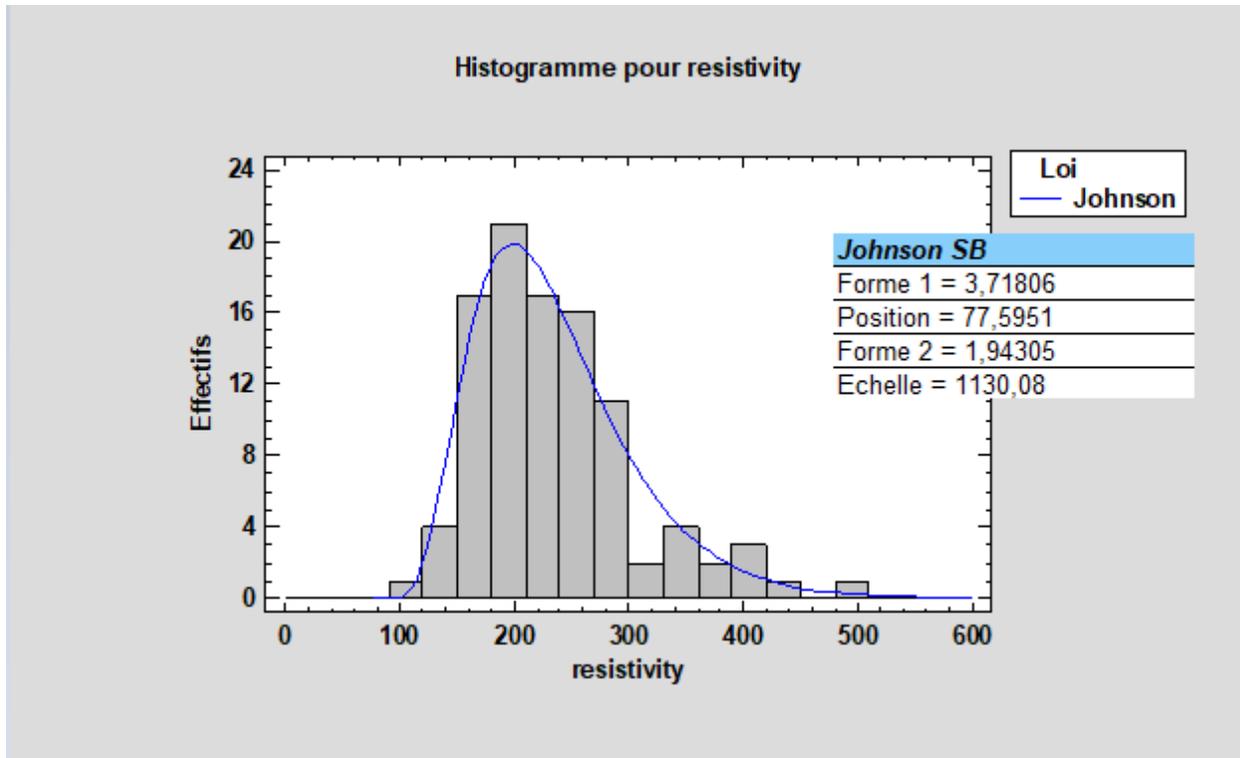
Mélanges gaussiens bivariés

Les données composées de deux variables aléatoires peuvent parfois être bien représentées par un mélange de deux lois normales bivariées.



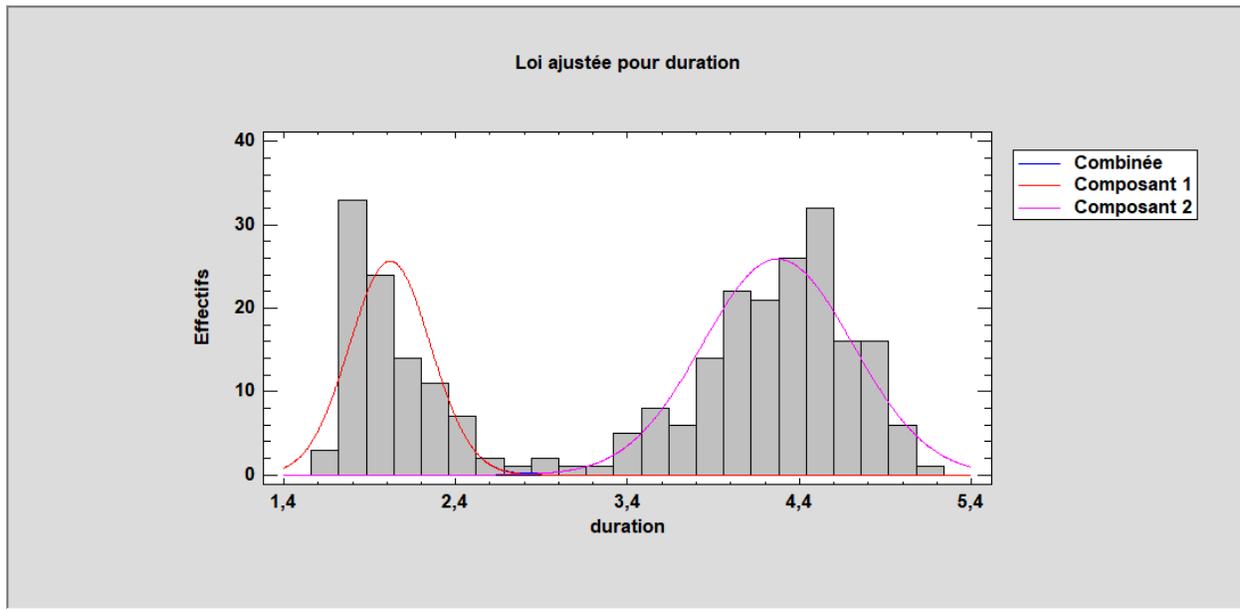
Lois de Johnson

La famille des lois de Johnson peut ajuster les quatre premiers moments de tout ensemble de mesures observées. Elle est utile pour générer un ajustement empirique à des données qui ne correspondent pas à une quelconque loi standard. Lors de l'ajustement des données, Statgraphics 19 sélectionne automatiquement la forme de la loi de Johnson à partir des types SB, SL et SU.



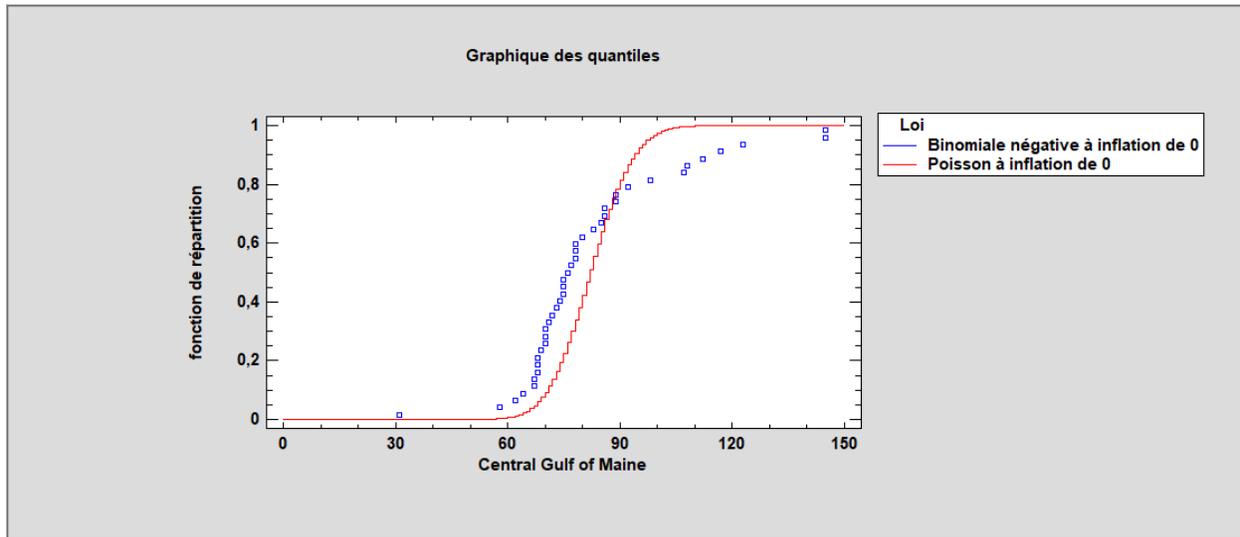
Mélanges gaussiens univariés

Les données réelles sont fréquemment des mélanges de plus d'une loi de probabilités. Souvent, ces données peuvent être représentées par un mélange de deux lois normales ou plus. Statgraphics 19 peut ajuster de telles lois et estime les proportions des observations provenant de chacune des lois.



Lois de Poisson et binomiale négative à inflation de 0

Les données de comptages ont fréquemment plus de zéros qu'attendu pour des lois de Poisson ou binomiale négative. En ajoutant un terme additionnel à la loi pour représenter ces zéros additionnels, une meilleure estimation peut souvent être obtenue.



Tests statistiques

Analyse de conformité

L'analyse de conformité a été ajoutée à l'analyse d'aptitude pour des attributs. Elle est utilisée pour estimer la probabilité pour un lot contenant un nombre fixe d'items de ne pas avoir plus d'un nombre donné de non conformes. Un indice appelé Cpc est également calculé pour comparer la probabilité estimée de non conformes à une valeur cible.

Analyse de conformité

Distribution: Binomiale

pourcentage moyen de défectueux = 23,1333

Taille du lot: 500

Nombre maximum autorisé de non-conformités: 100

Pourcentage de conformité cible: 99,73%

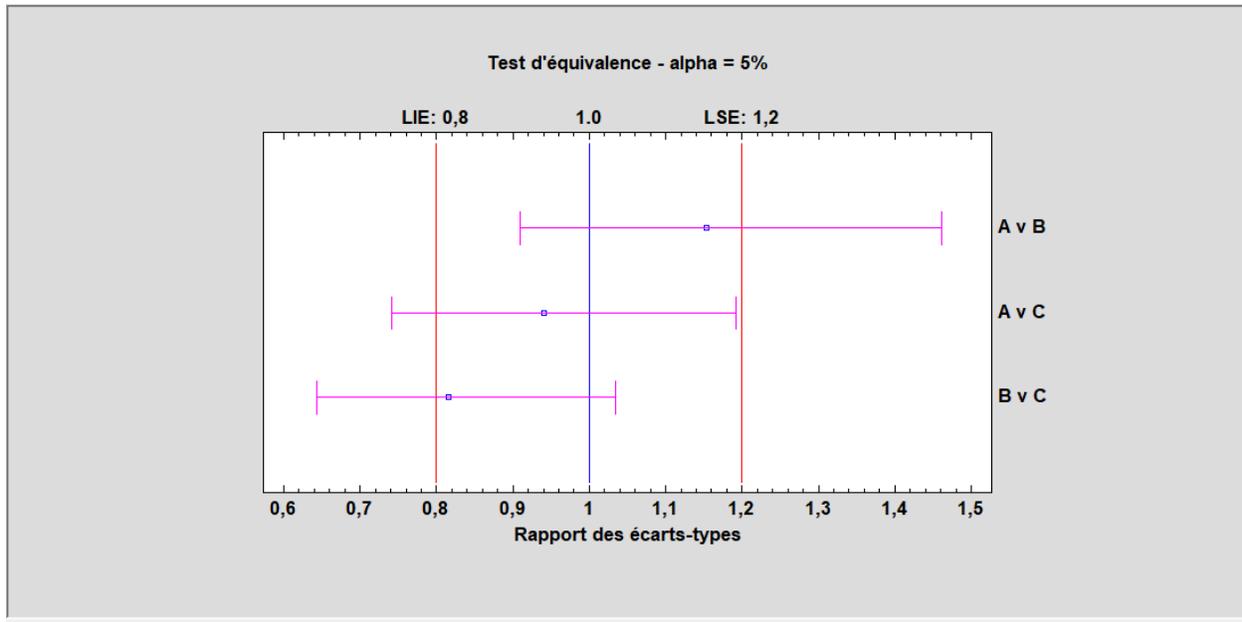
	<i>Estimation</i>	<i>Limite inf. à 95%</i>	<i>Limite sup. à 95%</i>
Proba. de conformité	0,0520954	0,0028932	0,309578
Cpc	0,00284839	0,00270783	0,00391065

Le StatAdvisor

La probabilité de ne pas avoir plus de 100 items non conformes dans un lot de 500 est estimée égale à 0,0520954. Avec un niveau de confiance de 95,0%, cette probabilité est comprise entre 0,0028932 et 0,309578. L'indice d'aptitude Cpc indique que la probabilité d'obtenir un lot non conforme est de 351,076 fois celle autorisée (l'inverse de Cpc).

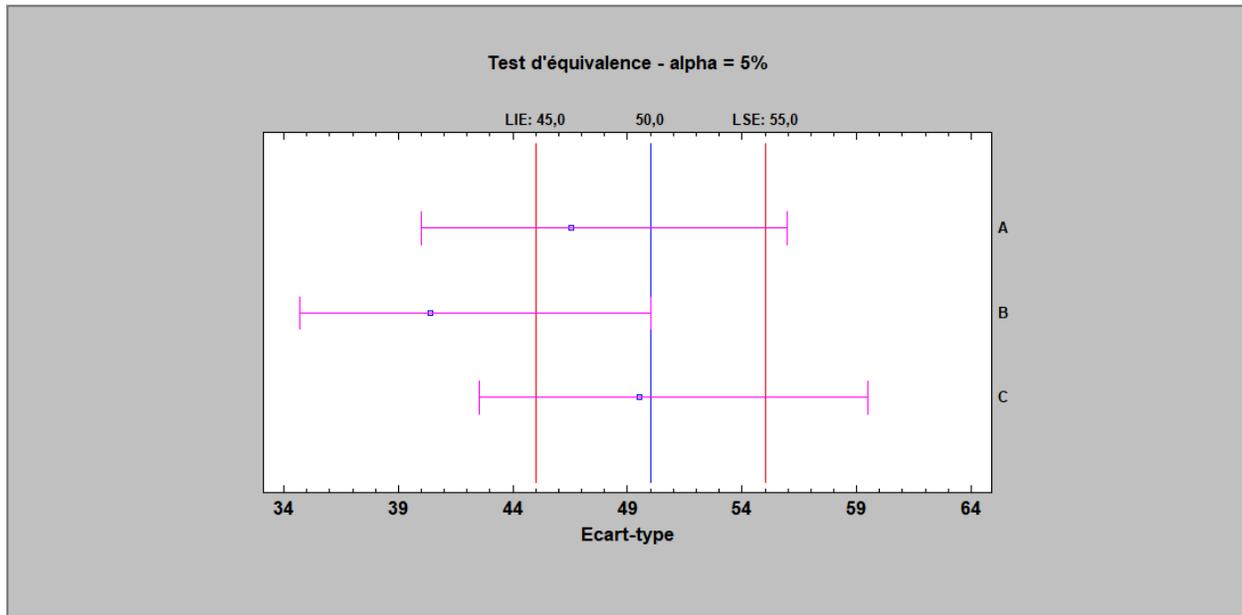
Tests d'équivalence et de non-infériorité – Comparer deux variances

Des tests d'équivalence et de non-infériorité ont été ajoutés pour démontrer que le rapport de deux écarts-types est suffisamment proche de 1.



Tests d'équivalence et de non-infériorité – Comparer une variance à une valeur cible

Des tests d'équivalence et de non-infériorité ont été ajoutés pour démontrer qu'un écart-type est suffisamment proche d'une valeur cible.



Test de Mann-Kendall

Le test de Mann-Kendall a été ajouté aux procédures 'Graphique de données séquentielles' pour tester la présence d'une tendance monotone dans une série temporelle.

Test de Mann-Kendall

Hypothèse nulle: pas de tendance

Hypothèse alternative: tendance monotone ascendante ou descendante

Test	Somme	Erreur-type	Statistique	Proba.
Mann-Kendall	515	335,42	1,53241	0,125421

Le StatAdvisor

Le test bilatéral de Mann-Kendall est utilisé pour déterminer si une tendance monotone ascendante ou descendante significative existe dans les données. Puisque la probabilité calculée est supérieure ou égale à 0,05, la présence d'une tendance n'a pas été détectée au niveau de signification de 5,0%.

Test de Levene modifié

Le test de Levene modifié a été ajouté pour déterminer s'il y a des différences significatives entre les variances de groupes dans une analyse de la variance à un facteur. Il effectue une analyse de variance à un facteur sur les valeurs absolues des écarts aux données aux médianes respectives des groupes plutôt qu'aux moyennes des groupes comme dans le test de Levene non modifié.

Tests des variances

	Test	Probabilité
Test de Levene modifié	0,316515	0,8133

Comparaison	Sigma1	Sigma2	F	Probabilité
A / B	9,18753	11,9695	0,589181	0,5018
A / C	9,18753	8,81456	1,08642	0,9157
A / D	9,18753	8,86405	1,07432	0,9271
B / C	11,9695	8,81456	1,84394	0,4381
B / D	11,9695	8,86405	1,82341	0,4464
C / D	8,81456	8,86405	0,988864	0,9886

Le StatAdvisor

La statistique affichée dans ce tableau teste l'hypothèse nulle que les écarts-types de strength pour chacun des 4 niveaux de material sont les mêmes. La valeur de la probabilité est particulièrement intéressante. Comme la valeur de la probabilité est supérieure ou égale à 0,05, il n'y a pas de différence statistiquement significative entre les écarts-types au niveau de confiance de 95,0%. Ce tableau affiche également une comparaison des écarts-types pour chaque paire d'échantillons. Les probabilités au-dessous de 0,05, il y en a 0, indiquent une différence statistiquement significative entre les deux écarts-types au niveau de signification de 5%.

Etudes R&R non équilibrées avec facteurs additionnels optionnels

Une nouvelle procédure a été ajoutée pour mettre en œuvre une étude R&R utilisant le modèle linéaire général (GLM). Elle a deux avantages par rapport aux précédentes procédures :

1. Elle traite le cas de données non équilibrées.
2. Des facteurs additionnels peuvent être inclus en plus des opérateurs, pièces et essais.

Tableau de l'ANOVA					
Source	Somme des carrés	Ddl	Moyenne quadratique	Rapport F	Proba.
Opérateurs	0,0928467	2	0,0464233	72,0525	0,0000
Pièces	2,20249	9	0,244721	379,825	0,0000
Coating	0,0381633	1	0,0381633	59,2323	0,0000
Répétabilité	0,06894	107	0,000644299		
Total (corrigé)	2,40244	119			

Le StatAdvisor
Le tableau de l'ANOVA divise la variabilité totale de Coating Thickness en plusieurs composants. Le dernier composant représente l'erreur résiduelle qui correspond à la répétabilité. Des probabilités inférieures à 0,05 indiquent des sources significatives de variabilité au niveau de signification de 5,0%.

Test de Wald-Wolfowitz

Le test non paramétrique de Wald-Wolfowitz a été ajouté dans la procédure de comparaison de deux échantillons. Ce test est calculé en triant les données combinées de la plus petite à la plus grande, en remplaçant chaque observation par le numéro de l'échantillon dont elle provient puis en comptant le nombre de suites dans la séquence des numéros des échantillons.

Test de Wald-Wolfowitz						
Nombre d'ex-aequo entre les échantillons = 2						
	Statistique minimum	Proba. minimum	Statistique maximum	Proba. maximum	Statistique moyenne	Proba. moyenne
Test exact (nombre de suites)	6	0,0007	6	0,0007	6,0	0,0007
z grand échantillon	-3,12406	0,0009	-3,12406	0,0009	-3,12406	0,0009
z grand échantillon (avec correction de continuité)	-2,93662	0,0017	-2,93662	0,0017	-2,93662	0,0017

Le StatAdvisor
Cette option met en œuvre le test de Wald-Wolfowitz pour comparer les distributions des deux échantillons. Ce test est calculé en triant les données combinées de la plus petite à la plus grande, en remplaçant chaque observation par le numéro de l'échantillon dont elle provient puis en comptant le nombre de suites dans la séquence des numéros des échantillons. Dans ce cas, le nombre de suites est égal à 6. La valeur de la probabilité est particulièrement intéressante pour ce test. En se basant sur le test exact, $P = 0,0007$. Puisque la valeur de la probabilité est inférieure à 0,05, il y a une différence statistiquement significative entre les deux distributions au niveau de confiance de 95,0%.

Apprentissage machine

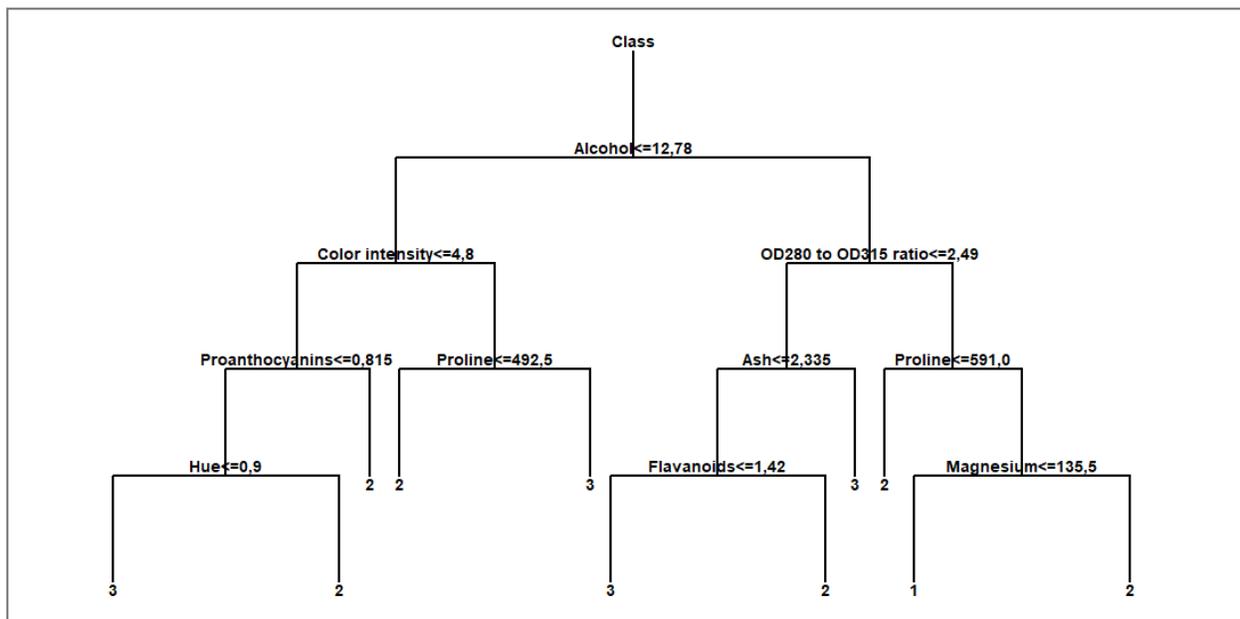
Deux nouvelles procédures ont été ajoutées pour implémenter des algorithmes d'apprentissage machine.

Forêts aléatoires décisionnelles

La procédure 'Forêts aléatoires décisionnelles' implémente une méthode d'apprentissage machine pour prévoir des observations à partir de données. Elle crée des modèles de 2 formes :

1. *Modèles décisionnels* qui découpent les observations en groupes basés sur les caractéristiques observées.
2. *Modèles de régression* qui prévoient la valeur d'une variable à expliquer.

Les modèles sont élaborés en construisant un grand nombre d'arbres de décision et en faisant la moyenne des prévisions obtenues à partir de ces arbres. De nombreux arbres sont construits en utilisant une procédure similaire à celle des Arbres de décision et de régression, avec optimisation aléatoire des nœuds et agrégation de bootstrap (bagging)

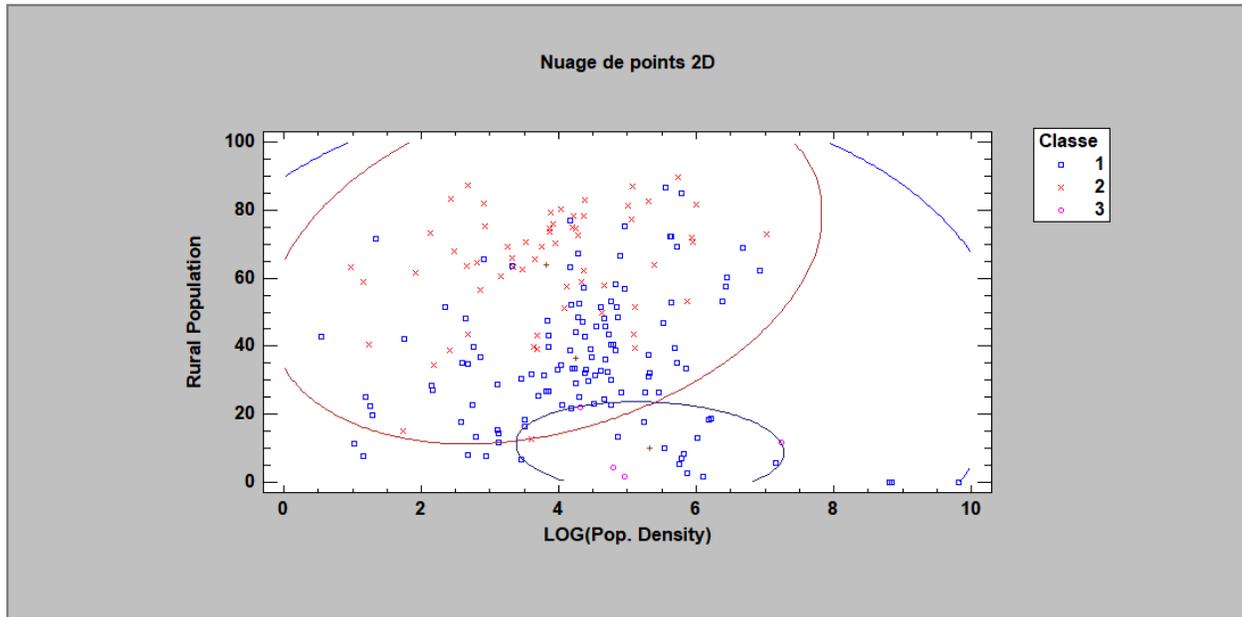


Classification par les K-moyennes

La procédure 'Classification par les K-moyennes' implémente un technique d'apprentissage machine permettant de créer des groupes ou classes d'observations caractérisées par des données quantitatives multivariées. Les classes sont créées en regroupant les observations qui sont proches dans l'espace des variables de données. A la différence de la procédure présente dans les précédentes versions, les noyaux initiaux n'ont pas à être précisés par l'utilisateur.

Résumé de la classification

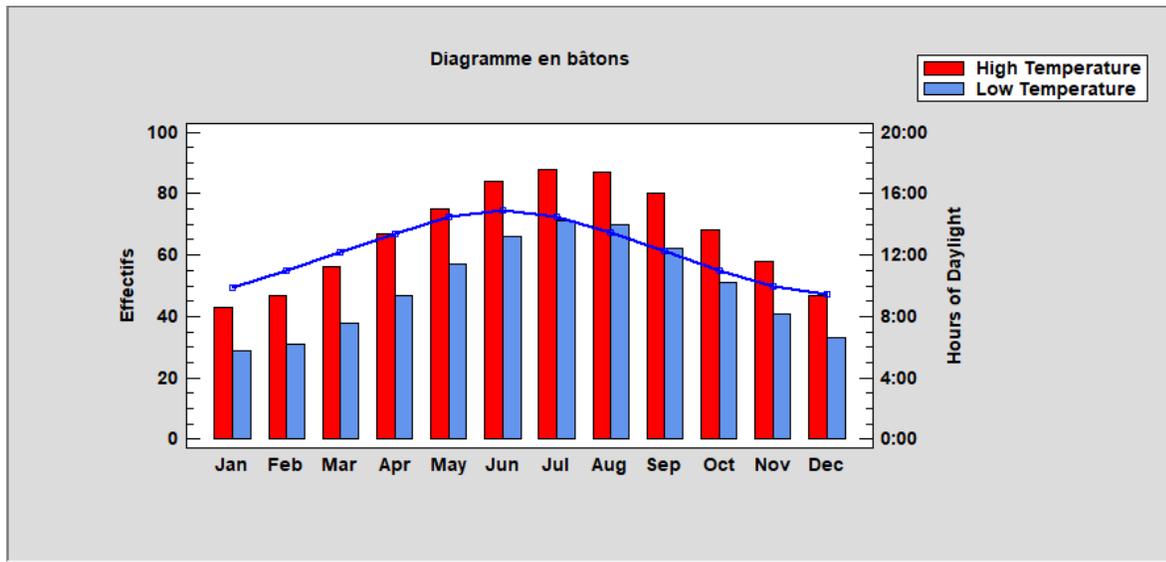
Classe	Membres	Pourcentage
1	123	69,10
2	61	34,27
3	4	2,25



Visualisation de données

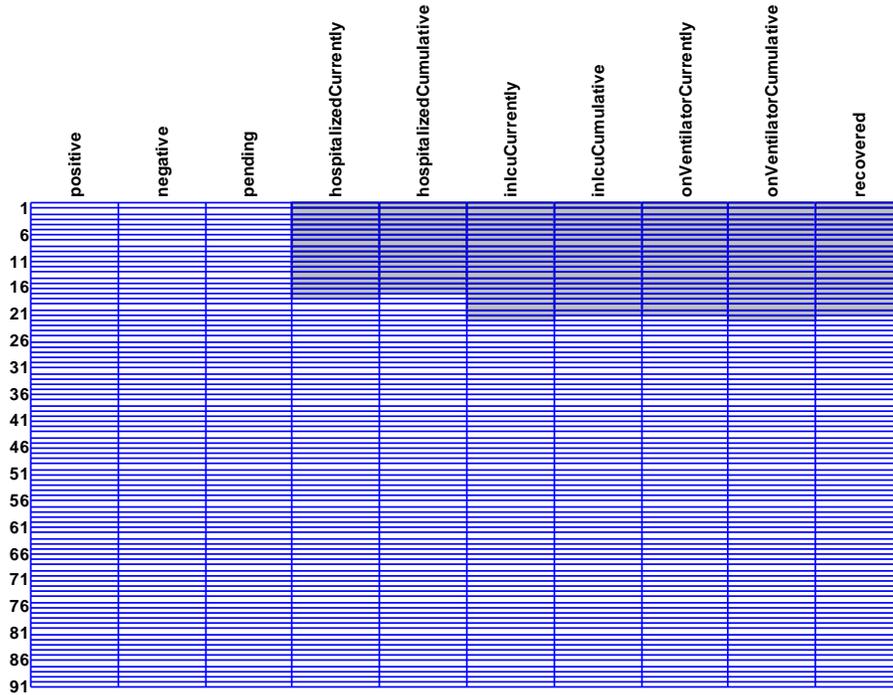
Diagramme en bâtons avec ligne avec ligne ajoutée

Une ligne peut être ajoutée dans les diagrammes en bâtons verticaux juxtaposés. Cette fonctionnalité est disponible pour les diagrammes simples et multiples en bâtons.



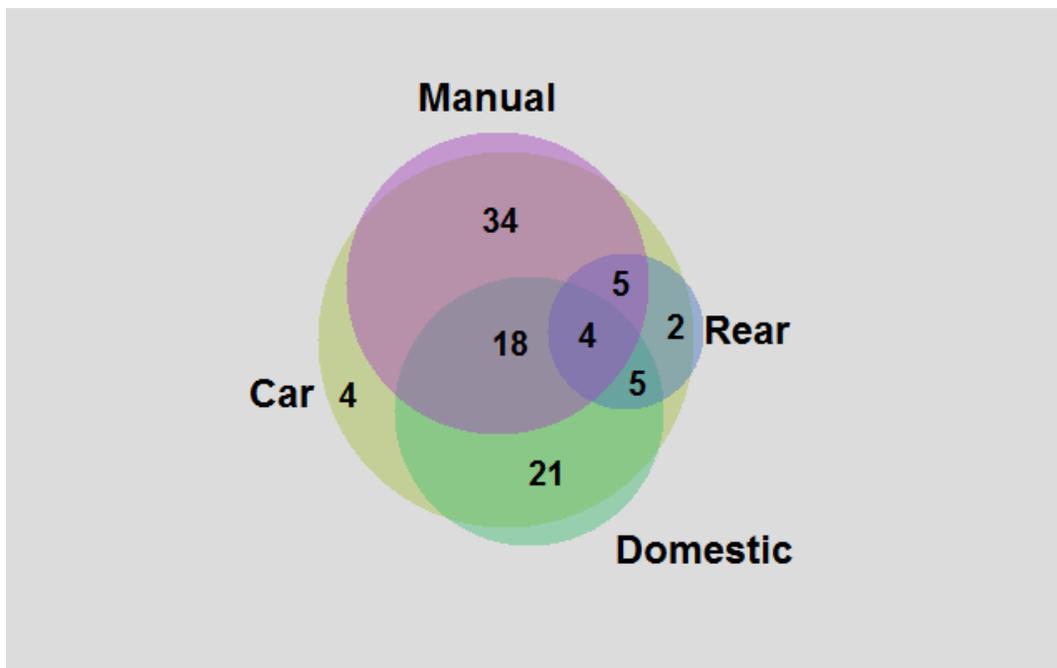
Graphique des données manquantes

Un nouveau graphique a été ajouté dans la procédure 'Visualiseur de données' pour aider à visualiser les positions des données manquantes dans un fichier de données. Les cellules ombrées indiquent les positions des données manquantes.



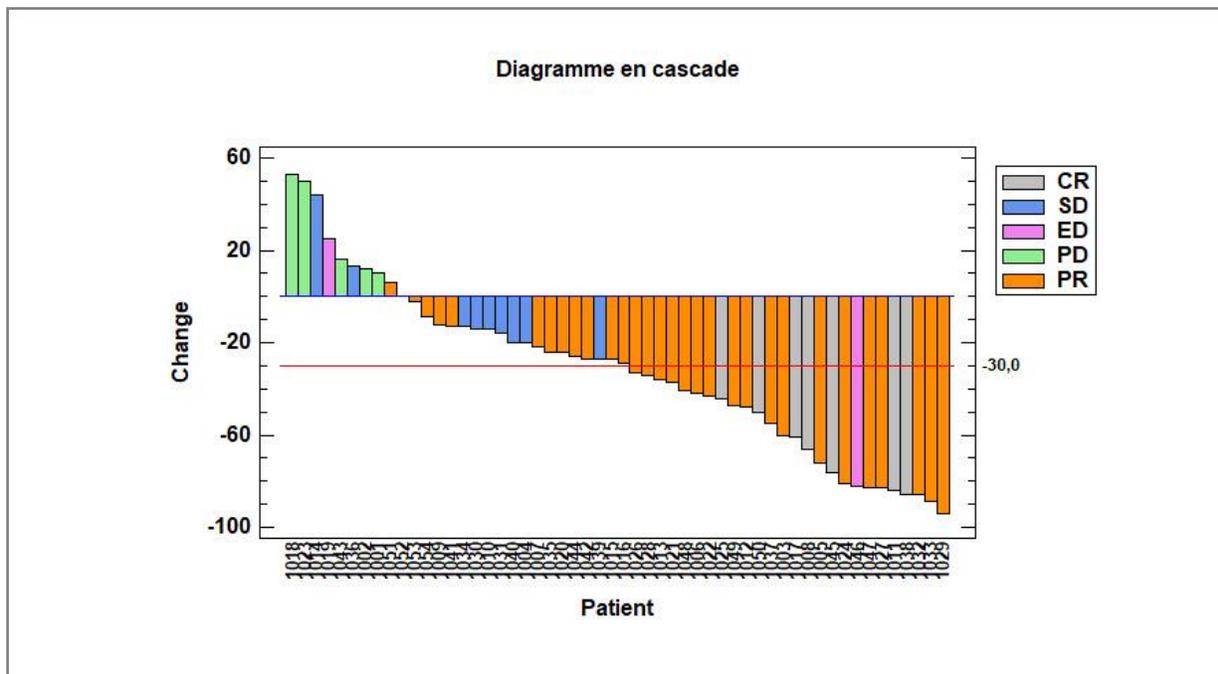
Diagrammes de Venn et Euler

La procédure 'Diagrammes de Venn et Euler' crée des diagrammes affichant des fréquences d'occurrences d'événements. Elle affiche des régions circulaires qui représentent les fréquences d'événements spécifiques dans lesquelles les chevauchements des cercles indiquent des occurrences simultanées de plus d'un événement.



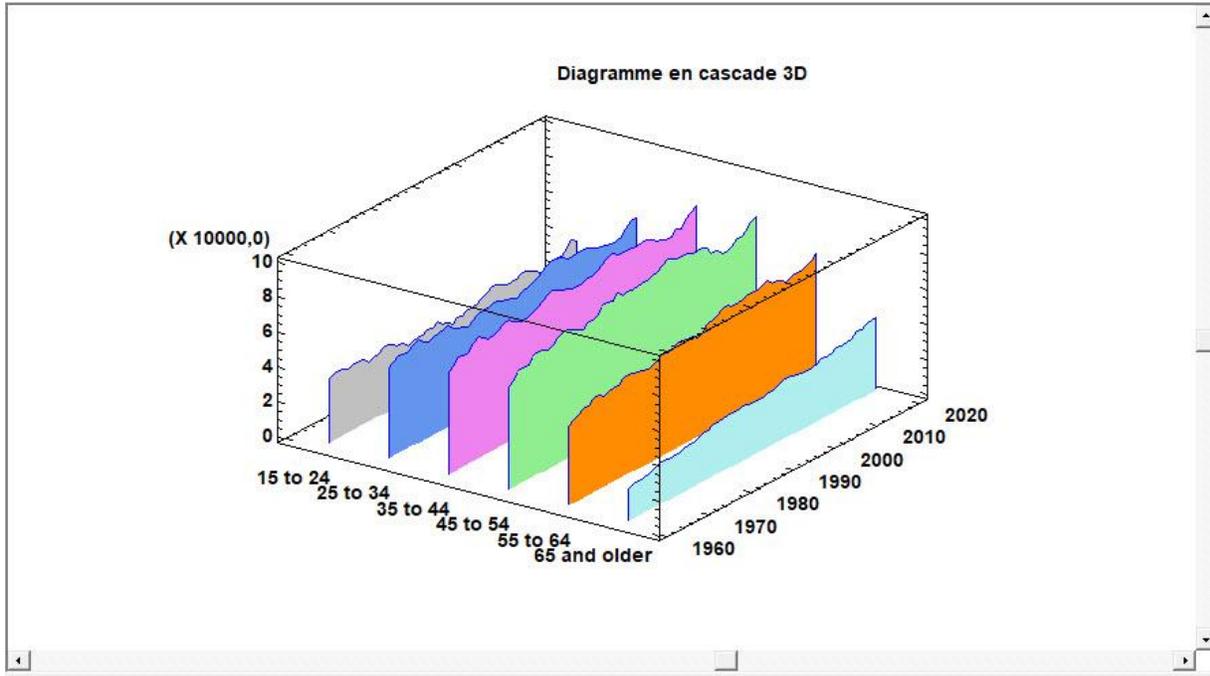
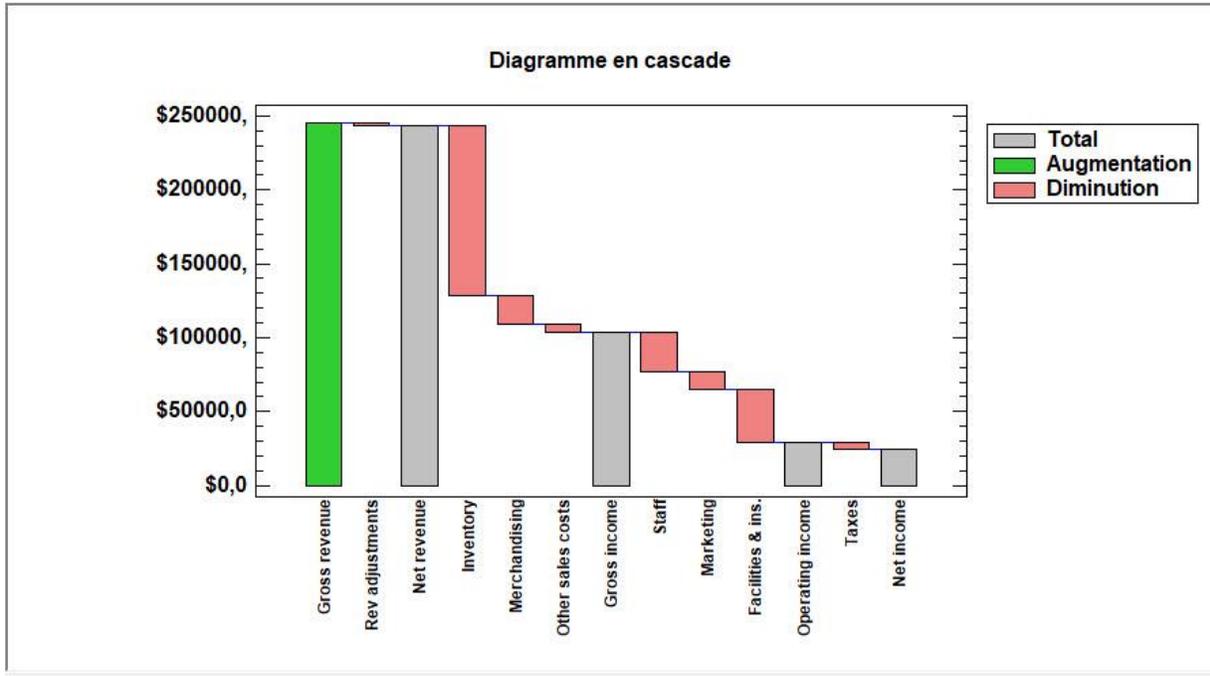
Diagrammes en cascade

Trois types de diagrammes en cascade ont été ajoutés dans la version 19 : un diagramme ordonné, un diagramme séquentiel et un diagramme 3D. Le diagramme en cascade ordonné est fréquemment utilisé pour montrer comment une variable d'intérêt augmente ou diminue dans un échantillon d'individus. Les données sont triées et affichées dans un diagramme en bâtons ayant habituellement une ligne de base à 0. Une ligne de référence peut être ajoutée dans le diagramme pour afficher une valeur cible.



Le diagramme en cascade séquentiel est fréquemment utilisé pour illustrer l'effet cumulé de contributions positives et négatives dans la décroissance d'une valeur totale. Les bâtons représentent chacune des contributions ainsi que les totaux et sous-totaux. Parmi les usages possibles de ce diagramme : finance, inventaire, analyse de performance, recrutement, analyse démographique.

Le diagramme en cascade 3D est fréquemment utilisé pour afficher plusieurs colonnes de données par rapport à une variable commune. Un exemple courant d'utilisation est le tracé de décroissance spectrale cumulative dans lequel un spectre est affiché à plusieurs instants pour visualiser les modifications de son amplitude en fonction de la fréquence et du temps. En général, ces graphiques sont utilisés pour montrer les modifications dans une variable quantitative en fonction du temps et d'un autre facteur.



Autres modifications

Modèles d'étalonnage

Une option a été ajoutée pour afficher des limites de prévision unilatérales :

Options pour le graphique du modèle ajusté

Inclure

- Limites de prévision
- Limites de confiance

Niveau de confiance: 95,0

Type de limites

- Bilatérales
- Bornes inférieures
- Bornes supérieures

Ombre les limites bilatérales

Prévoir

- Y
- X

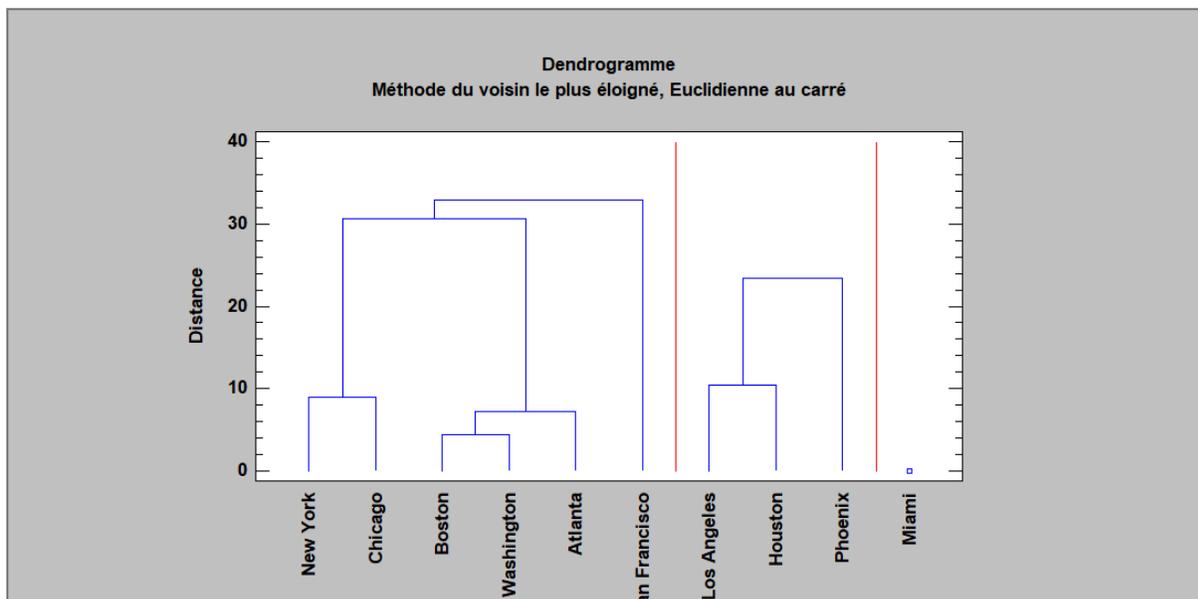
A:

Taille moyenne ou poids: 1,0

OK
Annuler
Aide

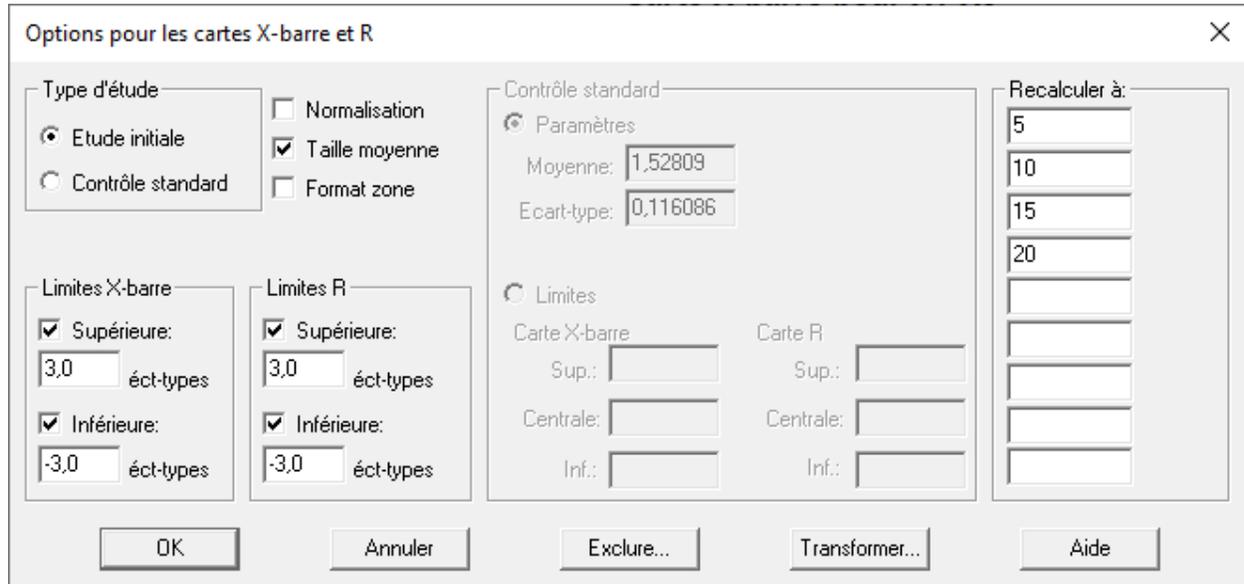
Classification

Une option a été ajoutée par séparer les classes dans un dendrogramme.



Cartes de contrôle

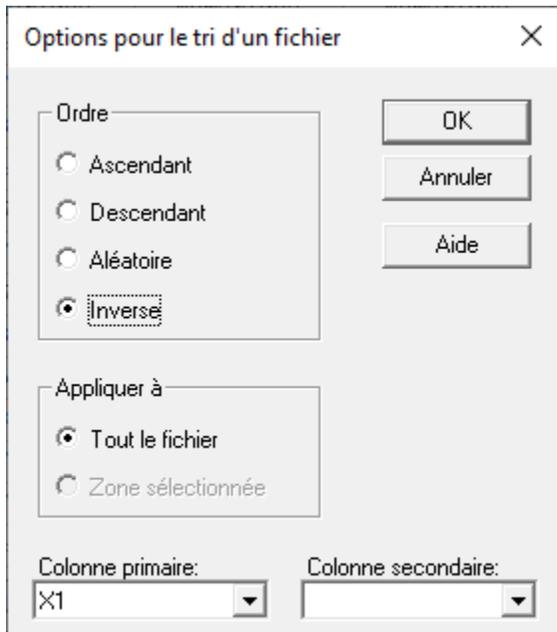
Le nombre de points auxquels les limites de contrôle peuvent être recalculées a été porté de 4 à 9.



Les entrées dans le champ 'Recalculer à' peuvent être des dates ou heures si les données sont dans ces formats.

Classeur de données – Tri en ordre inversé

Un option de tri inversé a été ajoutée dans le menu 'Editer – Trier' :



Si cette option est sélectionnée, l'ordre des lignes est inversé dans la feuille de données.

Classeur de données - Défaire

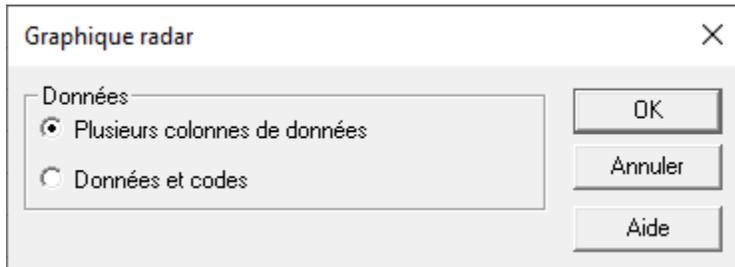
L'option 'Défaire' permet de défaire jusqu'à 30 opérations effectuées dans le classeur de données.

Importation des données

L'import direct de données aux formats projets Minitab, SAS Transport et SPSS Portable a été ajouté. Cela en plus des imports des fichiers de données classiques de Minitab, SAS et SPSS.

Graphique radar

En plus de l'entrée des données sous la forme de plusieurs colonnes, une nouvelle option sous la forme d'une colonne de données et d'une colonne de codes est proposée :



Enregistrement des graphiques

Une option a été ajoutée lorsque les graphiques sont enregistrés dans des fichiers images pour augmenter ou diminuer la tailles des symboles des points :



StatGallery

Les graphiques contenus dans la StatGallery peuvent être enregistrés dans des formats images de la même façon que les graphiques des fenêtres d'analyses.

Fond transparent dans les graphiques

La gestion de l'opacité des graphiques est maintenant conservée lorsque les graphiques sont copiés dans le presse-papiers de Windows pour les insérer dans d'autres applications.