ABÉCÉDAIRE DE SAS® VISUAL STATISTICS

SAS Visual Statistics est un nouvel outil, disponible depuis août 2014, qui répond aux besoins d'analyse de très gros volumes de données. Il permet de construire des modèles prédictifs de façon interactive, de les comparer et sélectionner le meilleur.

Caractéristiques :

Catégories : Visual Statistics
OS : Serveur Windows, Linux
Versions : 6.4, 7.1
Vérifié en décembre 2014

Nous allons ici décliner quelques lettres de l'alphabet pour vous permettre de vous familiariser avec SAS Visual Statistics.

Sommaire

A comme SAS Visual Analytics 1
C comme Clients 2
F comme en Francais 2
I comme Interface 2
L comme SAS LASR™ Analytic Server 2
P comme Programmation 3
S comme Serveurs 3
T comme Techniques statistiques 3
U comme URL 3
V comme Versions 4
Liens utiles 4
Conclusion 4

A COMME SAS VISUAL ANALYTICS

SAS Visual Statistics fait partie de la gamme de nos produits "SAS Visual". Il s'agit ici d'une application Web qui se déploie au sein de SAS® Visual Analytics. Il vient donc le compléter. Entendez que, techniquement et contractuellement, il est nécessaire de disposer de SAS Visual Analytics pour faire fonctionner SAS Visual Statistics.

Fonctionnellement, les deux outils sont complémentaires. SAS Visual Analytics s'utilise en premier lieu pour explorer les données. Il permet de répondre à des questions du type : « que s'est-il passé ? », « pourquoi cela est-il arrivé ? », en explorant les données et en produisant des statistiques descriptives.
Puis quand se pose la question : « comment puis-je prédire ce qu'il va arriver ? » SAS Visual Statistics entre en scène. La phase de

SAS Visual Analytics

- Qui sont aujourd'hui les clients les plus profitables ? Quelles sont leurs caractéristiques ?
- Quelles sont les composantes les plus corrélées en détection de fraude ?
- Quels sont les facteurs pertinents pour identifier le risque de départ de ses clients vers la concurrence (churn) ?

➔ Exploration et statistiques descriptives

SAS Visual Statistics

- Qui seront les clients les plus susceptibles d'être réceptifs à notre prochaine campagne ?
- Comment identifier les demandes ou réclamations frauduleuses ?
- Comment exploiter nos facteurs pour prédire les clients qui sont susceptibles de se tourner vers la concurrence, avant qu'il ne soit trop tard ?

➔ Comprendre et prévoir par la modélisation

modélisation peut commencer, avec l'utilisation de modèles prédictifs pour permettre d'expliquer des problématiques métier.

En d'autres termes, avec SAS Visual Statistics, vous disposerez des outils pour prendre de l'avance, et anticiper les faits à venir, que ce soit pour proposer le bon produit à l'achat ou encore en termes de fraude, de perte de clients, etc.

C COMME CLIENTS

SAS Visual Statistics consiste en une interface interactive, accessible depuis :

- un navigateur
  - Chrome, version 15 ou supérieure
  - Firefox, version 6 ou supérieure
  - ou Internet Explorer, version 9 ou supérieure

Plus de détails sont disponibles ici avec la liste des systèmes d'exploitation sur lesquels ils sont supportés, et les versions des navigateurs correspondantes.

- une tablette tactile
  - iPad d'Apple
    - La version d'iOS minimale est 6.0
    - Les types d'iPad supportés sont : iPad 2, 3, 4, iPad Mini et iPad Air
  - tablette Android
    - Pour Android 4.4 : Google Nexus 7 (éditions 2012 et 2013)
    - Pour Android 4.1 à 4.3 : Samsung Galaxy Tab 2 10.1, Samsung Galaxy Note 10.1 et 8, Google Nexus 7, Google Nexus 10.1
    - Pour Android 4.0.3 et 4.0.4 : Samsung Galaxy Tab 10.1, ASUS EEE Pad Transformer TF101.

L'application à télécharger

depuis l'App Store ou Google Play est SAS Mobile BI (non payante)

F COMME EN FRANCAIS

SAS Visual Statistics est disponible en français. La langue, dans laquelle l'interface est affichée, est déterminée par l'option « Langues » du navigateur en premier lieu, puis par les préférences de SAS Visual Statistics en second lieu.

I COMME INTERFACE / INTERACTIF / ITERATIF

L'interface web est conviviale et similaire à SAS Visual Analytics, pour proposer une expérience de navigation comparable et faciliter la prise en main. Les variables se sélectionnent par glisser-lâcher. Il est donc aisé :

- d'ajouter des variables, et des interactions entre variables,
- ou encore de supprimer des variables,
- de retirer des valeurs considérées comme aberrantes,
- de modifier des paramétrages au niveau du modèle,

pour voir l'impact, immédiat, sur la modélisation en cours. Le processus est ici à la fois interactif et itératif.

L COMME SAS LASR™ ANALYTIC SERVER

Les données sont montées en mémoire, dans le serveur SAS LASR. Toutes les analyses sont ainsi réalisées en mémoire, pour une restitution en quasi temps réel dans l'interface.
Un changement dans les paramètres du modèle, dans le choix des variables, etc, ne nécessite donc pas de répéter le processus intégral. L'interactivité est au rendez-vous.

Le nombre de tables chargées dans le serveur LASR n'est pas limité, ni le nombre d'utilisateurs concurrentiels. La limite sera uniquement fonction de la mémoire disponible et du « hardware ».

P COMME PROGRAMMATION

Aucune connaissance en programmation n'est nécessaire, puisque les manipulations se font « à la souris » dans SAS Visual Statistics, que ce soit le choix des données, du modèle et de ses paramètres, ou la sélection du modèle champion.
Une fois le modèle champion sélectionné, la finalité est de déployer ce modèle, en l'appliquant à un nouveau jeu de données, pour prédire la variable réponse. Cette étape passe par la génération, automatique, de code SAS, pour appliquer le score à un nouveau jeu de données. Ce programme SAS, une étape data incluant toutes les variables utilisées par le modèle, sera ensuite exécuté dans une session SAS classique.

S COMME SERVEURS

SAS Visual Statistics doit être installé sur un serveur 64-bit. Il doit s'agir de :

- Linux,
- ou Windows (en mode non-distribué uniquement)

Attention, une installation sur un PC Windows 64-bit n'est pas supportée. Il convient de disposer d'une licence serveur de Windows 64-bit.
Les informations détaillées, techniques, sur les composants impliqués, sont disponibles à la fois en mode non-distribué (configuration de SAS LASR Analytic Server sur une seule machine) et distribué.

T COMME TECHNIQUES STATISTIQUES

Voyons les techniques statistiques disponibles dans SAS Visual Statistics. Les familles de méthodes étant :

➢ La classification, avec :

La régression logistique

Le modèle linéaire généralisé

L'arbre de décision

➢ La régression, avec :

La régression linéaire

Le modèle linéaire généralisé

L'arbre de décision

➢ Le « clustering », avec :

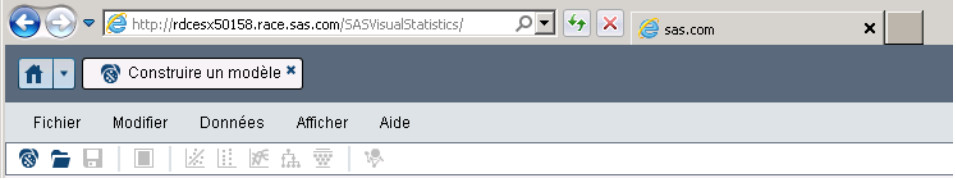
Cluster

➢ Ajoutons la notion de « Group-by »
La modélisation est en effet possible pour chaque modalité d'une variable.

U COMME URL

Dans un navigateur, l'adresse à utiliser se présente sous cette forme :
http://nom_de_votre_serveur/SASVisualStatistics

Avant toute chose, il est alors nécessaire de s'authentifier, avec un nom d'utilisateur / mot de passe, fourni par l'administrateur SAS.



Il est également possible d'accéder à SAS Visual Statistics depuis la page d'accueil de SAS Visual Analytics (le « hub ») : http://nom_de_votre_serveur/SASVisualAnalyticsHub/, en cliquant soit sur l'icône « Construire un modèle » ou sur le bouton « Créer un modèle » :

V COMME VERSIONS

La première version de SAS Visual Statistics est la 6.4. Elle est sortie début août 2014 et est livrée avec un dépôt SAS 9.4 Maintenance 2.
Pourquoi 6.4 ? Pour être en adéquation avec la numérotation de version de SAS Visual Analytics.

La seconde version est la 7.1 et est disponible depuis début octobre 2014.

LIENS UTILES

- La page de référence sur support.sas.com, pour accéder entre autres :
  - aux informations générales,
  - aux documentations (l'identifiant et le mot de passe à utiliser seront indiqués dans le fichier instructions.html généré lors de l'installation du produit).
  - à des vidéos de présentation
- Le site en français, pour découvrir des exemples concrets d'utilisation, et des témoignages clients
- Le site de la « communauté » des utilisateurs, pour poser des questions, et proposer et/ou voter pour les idées d'amélioration du produit.

CONCLUSION

Si vous recherchez l'interactivité, lorsque vous établissez vos modèles statistiques, et que vous travaillez sur de très gros volumes de données, pensez à SAS Visual Statistics !

Nous vous encourageons à vous connecter sur notre site pour un essai totalement gratuit :
http://www.sas.com/en_us/offers/14q3/visual-stats-try-before-you-buy.html
Vous pourrez vous authentifier avec votre profil SAS, ou remplir le formulaire, pour avoir ensuite accès :
- à SAS Visual Analytics pour vous familiariser avec son interface
- au guide de démarrage pour SAS Visual Statistics
- et enfin à SAS Visual Statistics

Les seuls pré-requis sont de disposer de l'un de ces navigateurs : Internet Explorer 9, Firefox 6 ou Chrome 15 (ou leurs versions supérieures), et d'Adobe Flash Palyer 11.1 minimum.

Karine Chrillesen
Consultante Support Clients SAS France