



Contexte

L'analyse statistique est en train de devenir un des pivots incontournables de la stratégie d'entreprise. Une bonne connaissance du marché, de la concurrence, des clients ou même des mécanismes internes à l'entreprise permet d'améliorer l'offre, de se positionner correctement sur les nouvelles tendances, d'anticiper l'évolution des marchés, augmenter les retours sur investissement...



L'analyse statistique se trouve classiquement confrontée à trois écueils :

- la qualité des données ainsi que la variété des sources peut entraîner des prétraitements (data management ou consolidation des données) coûteux en termes de temps humain ;
- les données sont de plus en plus volumineuses (big data), de ce fait sont difficilement traitables ;
- les temps de traitement sont parfois prohibitifs.

La solution



R++, *the Next Step* est un logiciel d'analyse statistique **hautes performances**. C'est avant tout un logiciel métier facilement accessible à destination d'utilisateurs non spécialistes, intégré dans une **Interface Homme Machine (IHM)** moderne et conviviale. Mais c'est aussi un logiciel expert qui s'appuie sur les dernières innovations en termes de gestion du **parallélisme** et de **Big Data**.

Au niveau IHM, les recherches que nous avons menées en collaboration avec l'équipe LII (ENAC) ont conduit à la conception d'une nouvelle interface permettant de diviser par 3 ou 4 le temps de préparation des données. Cette interface révolutionnaire (brevetée) économise du temps humain sur la partie du travail généralement considérée comme la moins intéressante.

Du point de vue de la performance pure, les Preuves de Concept (PoC) ont montré que l'utilisation de la carte graphique (GPU) permet d'aller entre 100 et 800 fois plus vite que les logiciels classiques R et SAS (breveté également). 800 fois plus vite signifie que *R++ the Next Step* fait en deux minutes un calcul qui aurait pris 26 h sur un autre logiciel. Enfin, une variante de la lecture de disque « *out of core* » nous a permis de traiter sur un ordinateur portable le même volume de données que le supercalculateur du CERFACS (centre de calcul de Météo France) utilisant des méthodes classiques.

Avantages, applications et maturité

1. Avantages concurrentiels

- Plus ergonomique (économie de temps humain, moins de risque d'erreur)
- Plus performant
- Plus rapide

2. Applications & Marchés

- Industrie pharmaceutique
- Aéronautique
- Grande Distribution
- Banques et Assurances
- Défense, Cybersécurité

3. Maturité du projet

- 1 directeur technique
- 4 ingénieurs de recherche
- 1 business developer
- Brevet sur le parallélisme
- Brevet sur l'IHM

Contact

Christophe Genolini
Président

@ cg@rplusplus.com

☎ 06 21 48 47 84

📍 5 place Jean Deschamps
31000 Toulouse, France



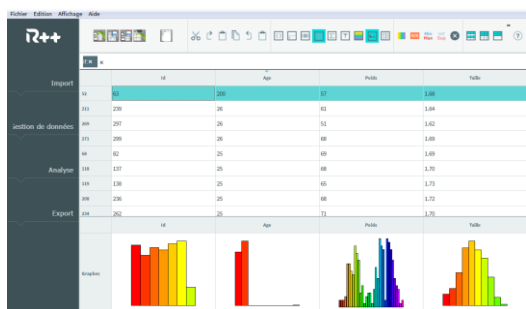
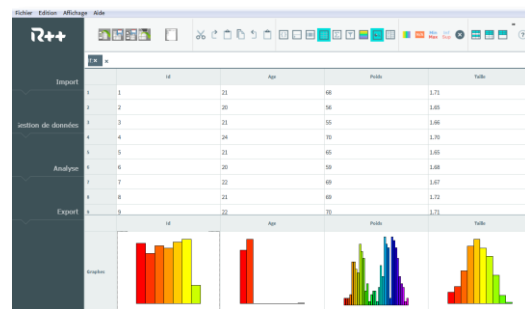
Avec le soutien de :



Quelques fonctionnalités graphiques

Graphes « en un clic »

La première étape de toute analyse statistique est de se familiariser avec les données et de détecter les valeurs aberrantes. Cela passe par une représentation graphique. R++ offre la fonctionnalité « Graphes en un clic » : pour chaque colonne, le graphe correspondant à la nature de la variable est tracé. Ainsi, les valeurs aberrantes de l'ensemble du fichier sont toutes immédiatement détectables.

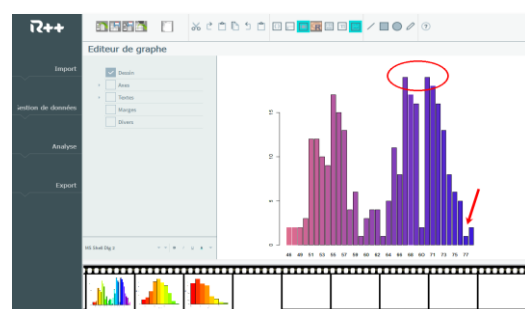


Graphes cliquables

Une fois une valeur aberrante détectée, il faut la corriger. Il n'est pas toujours simple, dans une base de données gigantesque, de trouver la ligne incriminée. Dans R++, les graphes sont cliquables : en cliquant sur la valeur aberrante, la ligne correspondante est mise en évidence (en bleu sur le dessin de gauche) et l'ordre des lignes est modifié afin que l'individu sélectionné se retrouve en haut de la base. Il est alors très simple de corriger la valeur aberrante.

Graphes éditables & exportables

Les logiciels d'analyse statistiques produisent des graphes qu'il est ensuite très compliqué de modifier. R++ intègre nativement un éditeur graphique complet. Tout est modifiable, de la couleur aux polices en passant par les axes, les titres, les marges... Il est également possible d'ajouter des formes comme des ellipses, des flèches ou même du texte, afin de mettre en évidence des éléments particuliers. Le graphe obtenu est ensuite directement exportable vers d'autres logiciels (suite Microsoft, OpenOffice, mail, ...) par simple drag & drop.



R++, c'est aussi...

- **Un logiciel simple...** son interface, ses Drag & Drop, sa simplicité, tout est fait pour le rendre intuitif
- **...mais pro !** toutes les actions sont codées, le code est accessible en permanence.
- **Un ami de vos amis :** R++ permet la prise en charge de 17 différents formats de fichiers différents, localement ou à distance (Oracle, SQL, no SQL)
- **...de tous vos amis :** R++ exporte les résultats simplement vers Microsoft Office mais aussi LaTeX, Open Office, HTML... Par fichier ou par simple Drag & Drop
- **Un logiciel évolutif...** Quand vous corrigez une erreur, R++ s'en souvient et vous proposera, ultérieurement, d'appliquer cette correction à un nouveau jeu de données.
- **...et collaboratif !** Vous pourrez aussi appliquer les correctifs de vos collègues, votre équipe... ou du monde entier, grâce à un wiki collaboratif « erreur les plus courantes » !
- **Un rassembleur :** Une interface simplifiée est spécifiquement dédiée à la fusion de bases de données.

**R++, c'est la puissance de l'informatique moderne
au service de l'analyse statistique**