

LA STATISTIQUE AVEC UN TABLEUR

Bernard Burtschy

ENST

46 rue Barrault

75013 PARIS

Le tableur a révolutionné l'informatique et l'informatique a révolutionné la statistique
La relation est-elle transitive ?

Première application historique du micro-ordinateur et moteur essentiel sinon exclusif dans la conquête des utilisateurs professionnels, le tableur a été le cheval de Troie qui a permis de détrôner l'informatique centralisée, ses lourdeurs et ses pompes. Il n'a pas si longtemps que cela, il fallait une armée de programmeurs Cobol pour établir et maintenir un simple tableau de bord consolidé. Cette opération, comme tant d'autres, s'effectue maintenant en quelques heures, sur un coin de bureau. De fait, le tableur est devenu l'outil à tout faire, et même le symbole, du cadre moderne.

Que l'informatique ait révolutionné l'analyse des données, composante importante (mais non exclusive) de la statistique, il n'y a pas besoin de faire un dessin au lecteur assidu de la revue de MODULAD. Il en est intimement persuadé. La vague de fond commence à toucher la statistique en général.

Quelles sont les relations entre le tableur et la statistique ? Les deux travaillent sur des tableaux de données qui sont leurs matières de base. L'esprit est différent. Le principe de l'analyse des données est de prendre un tableau de données et de « l'essorer » à fond. Les logiciels spécialisés proposent moult méthodes de traitements, mais font peu de cas de l'entrée et de la gestion des données. A quelques exceptions près, les manipulations des lignes ou des colonnes relèvent du parcours du combattant. Les logiciels spécialisés de statistique sont des pachydermes en matière de gestion de données. Dans bien des cas, il vaut mieux savoir programmer pour se créer un logiciel jetable *ad hoc* qui ne servira qu'une seule fois.

Le tableur quant à lui, est spécialisé dans la gestion des données sous toutes ses formes. Saisie des données, importations, créations de formules et de liaisons, outils de mise en forme, tris, filtrages, création de tableaux croisés, n'ont pas de secrets et s'effectuent aisément par des simples clicks de la souris ou à l'aide d'assistants. Le traitement statistique n'est pas son fort. Mais, fait nouveau, ces logiciels ne se cantonnent plus à la gestion des données. Version après version, les outils graphiques et statistiques s'amoncellent, sans bruit et sans que les statisticiens s'en aperçoivent.

Le marché des outils logiciels pour statisticiens (Tab 1) se partage maintenant en deux. D'un côté, il existe, depuis quelque temps déjà, des logiciels riches en méthodes statistiques. Réservés à une élite qui est et se sent dépositaire d'un savoir-faire, ils leur permettent de soumettre les données à un intense pilonnage de méthodes afin qu'elles expriment jusqu'à la dernière goutte, toute la substantifique information qu'elles recèlent. Ces logiciels sont à confier aux mains de statisticiens professionnels aguerris.

Par ailleurs, il existe tout un ensemble de données qui ne nécessitent qu'une analyse statistique rapide et légère. Les logiciels professionnels de statistique sont des marteaux-pilons pour ces données. Ils sont aussi beaucoup trop lourds pour une utilisation pédagogique. Les élèves passent plus de temps à apprendre le jargon des logiciels statistiques qu'à assimiler les méthodes. Faute de logiciels adaptés, beaucoup de données ne sont pas analysées du tout et restent en jachère.