

**JSFS**

**Avant-propos**

*Journal de la société française de statistique*, tome 142, n° 1 (2001), p. 3

[http://www.numdam.org/item?id=JSFS\\_2001\\_\\_142\\_1\\_3\\_0](http://www.numdam.org/item?id=JSFS_2001__142_1_3_0)

© Société française de statistique, 2001, tous droits réservés.

L'accès aux archives de la revue « Journal de la société française de statistique » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

## AVANT-PROPOS

Pour un journal dont l'ambition est de développer un outil d'information et de communication entre statisticiens, il était opportun et important d'aborder le thème développé dans ce volume.

Je ne chercherai pas à faire en quelques lignes une synthèse du débat présenté : ce serait illusoire et difficile, à moins de reprendre la « réponse courte » de J. Friedman (voir contribution de R. D. De Veaux) qui, je l'avoue, me satisfait bien. Je suis en effet de ceux qui donnent une définition très large au mot statistique, ayant toujours pensé que c'était simplement la science (ou l'art ?) d'exploiter des données, d'où ma surprise à l'avènement du *data mining*, comme d'ailleurs naguère à celui de *l'analyse des données*. Ne s'agit-il pas d'avatars de la même démarche ? Il est clair que la recherche, en statistique comme ailleurs, doit évoluer en suivant les besoins et en s'adaptant aux possibilités de les satisfaire ; il est clair aussi que le développement et la diffusion des méthodes demandent des efforts de communication. Peut-être un vocabulaire adapté est-il utile dans ce contexte. Mais on peut aussi se poser quelques questions : pourquoi, apparemment, le mot statistique fait-il peur ? ou, au moins, pourquoi est-il souvent compris dans un sens étroit ? Ce sont des questions très légitimes dans le Journal d'une Société de Statistique puisqu'elles portent sur notre identité fondamentale. Et cela nous ramène à l'utilité actuelle et l'importance du débat présenté ici. Merci donc à tous ceux qui ont bien voulu y contribuer en élaborant la « réponse longue » à la comparaison entre *data mining* et *statistique*, ouvrant ainsi des perspectives variées, même si, pour l'essentiel, un large accord semble se dessiner et l'émergence du *data mining* n'engendre pas de stériles querelles de chapelles, ce dont on ne peut que se réjouir (vue de l'utilisateur, l'efficacité primerait-elle sur les luttes d'influence et le marquage de territoire ?). Mais le débat n'est certainement pas clos et le Journal reste ouvert à tous pour le poursuivre.

Henri Caussinus