

# R pour la statistique et la science des données

Le logiciel R est un outil incontournable de statistique, de visualisation de données et de science des données tant dans le monde universitaire que dans celui de l'entreprise. Ceci s'explique par ses trois principales qualités : il est gratuit, très complet et en essor permanent. Récemment, il a su s'adapter pour entrer dans l'ère du big-data et permettre de recueillir et traiter des données hétérogènes et de très grandes dimensions (issues du Web, données textuelles, etc.).

Ce livre s'articule en deux grandes parties : la première est centrée sur le fonctionnement du logiciel R tandis que la seconde met en œuvre une trentaine de méthodes statistiques au travers de fiches. Ces fiches sont chacune basées sur un exemple concret et balayent un large spectre de techniques pour traiter des données.

Ce livre s'adresse aux débutants comme aux utilisateurs réguliers de R. Il leur permettra de réaliser rapidement des graphiques et des traitements statistiques simples ou élaborés.

## Les auteurs

**Pierre-André Cornillon,**  
Université Rennes 2

**Arnaud Guyader,**  
Sorbonne Université

**François Husson,**  
Agrocampus-Ouest

**Nicolas Jégou,**  
Université Rennes 2

**Julie Josse,**  
École polytechnique

**Nicolas Klutchnikoff,**  
Université Rennes 2

**Erwan Le Pennec,**  
École polytechnique

**Eric Matzner-Løber,**  
Cepe ENSAE

**Laurent Rouvière,**  
Université Rennes 2

**Benoît Thieurmél,**  
Datastorm