

Outils pour chercher de l'information sur R et se former

N. Villa-Vialaneix^a and C. Bontemps^b and S. Déjean^c

^aMIAT, Université de Toulouse, INRA
31326 Castanet-Tolosan cedex, France
nathalie.villa@toulouse.inra.fr

^bToulouse School of Economics, Université de Toulouse, GREMAQ-INRA
31000 Toulouse, France
christophe.bontemps@toulouse.inra.fr

^cInstitut de Mathématiques de Toulouse, Université de Toulouse
31062 Toulouse cedex 9, France
sebastien.dejean@math.univ-toulouse.fr

Mots clefs : R, formation, documentation.

Dans cette proposition de communications, nous nous proposons de faire le tour de ressources disponibles en ligne pour rechercher des informations sur R, son installation, son utilisation ainsi que de celles qui permettent de se former. Notre ambition n'est pas de fournir une liste exhaustive de ces ressources mais, devant le foisonnement et le développement de sites web, blogs et ressources diverses concernant le logiciel, de faire un descriptif organisé de celles que nous avons utilisées ou appréciées.

R : où commencer ?

Si vous êtes novice sous R, beaucoup de sites web vous aideront à faire les premiers pas et nombre d'enseignants mettent en ligne leurs polycopiés de formation : il serait trop long et fastidieux d'en faire un inventaire, même superficiel. On peut toutefois citer quelques initiatives pédagogiques innovantes :

- des *MOOCs*, sur diverses plateformes et en diverses langues. La plupart de ses initiatives sont des cours de statistique illustrés avec R mais on peut aussi trouver des MOOCs spécifiquement conçus pour apprendre à utiliser R, comme par exemple le MOOC « R programming » qui est un MOOC assez dense et complet, enseignant les bases du langage de programmation en 4 semaines <https://www.coursera.org/learn/r-programming> ou bien le MOOC « Introduction to R for data science » <https://www.edx.org/course/introduction-r-data-science-microsoft-dat204x> de la plateforme edX¹ ;
- en français, le tutoriel « R pour débutants », disponible en ligne https://cran.r-project.org/doc/contrib/Paradis-rdebuts_fr.pdf est une bonne entrée dans le logiciel en 80 pages environ. On trouve également l'initiative originale **Begin'R** de l'université de Bordeaux qui proposent des tutoriels en ligne de l'installation de R à son utilisation en statistique ;
- une manière originale d'apprendre à utiliser R depuis R est **swirl**, <http://swirlstats.com> : les enseignants y trouveront également un support pour créer leurs propres formations interactives. Dans un esprit similaire (en ligne), le site <http://tryr.codeschool.com> propose des tutoriels progressifs pour se familiariser avec la syntaxe de R.

1. <https://www.edx.org/>

R : les incontournables

Une fois lancés, certaines pages deviennent des incontournables : on peut citer, par exemple, les manuels R du CRAN <https://cran.r-project.org/manuals.htm> qui sont la documentation officielle sur la structure du langage, l'installation du logiciel ou bien la création de packages. Sur le site du CRAN, on trouve aussi les « Task Views » <https://cran.r-project.org/web/views> qui sont une liste de plusieurs pages thématiques (« Bayesian », « Genetics », « Optimization »...) gérées chacune par un « mainteneur » identifié qui recense les principaux packages de la thématique et décrivent leurs fonctionnalités.

En dehors du CRAN, le dépôt Bioconductor <http://bioconductor.org> a une place croissante et propose des packages orientés vers l'analyse bioinformatique et biostatistique. On y trouve un moteur de recherche spécifique, des documentations, organisées par type de données ou type d'analyses, des formations (également sous forme de vidéos) mais aussi des « workflows ». Chaque package Bioconductor est accompagné d'au moins une vignette qui complète l'aide en ligne classée par une mise en œuvre pratique des fonctionnalités du package.

Enfin, le site web R-bloggers <http://www.r-bloggers.com>, agrégeant les articles de plus de 573 blogs sur R, est devenu un incontournable pour se tenir informé-e des dernières avancées du logiciel. En français, <http://r.developpez.com> proposent également de nombreux tutoriels pour approfondir ses connaissances, notamment sur les aspects programmation et administration.

R : aides mémoire

Pour tous ceux dont la mémoire est défaillante, certains sites sont une collection de ressources organisées sur R : le « R programming wikibooks » https://en.wikibooks.org/wiki/R_Programming, site web collaboratif rappelant les bases de R, les feuille de triche de RStudio <https://www.rstudio.com/resources/cheatsheets>, organisées en thématique pour vous aider à utiliser RMarkdown, RStudio, **shiny**, ... En français, le moteur de recherche En français, le site recherche <http://www.duclert.org/Aide-memoire-R> fournit également un aide mémoire organisé des principales fonctions du langage.

R : à l'aide!!!

Si vous cherchez de l'aide, vous pouvez vous adresser aux divers forums spécialisés sur le langage : <http://stackoverflow.com/questions/tagged/r>, <http://r.789695.n4.nabble.com/R-help-f789696.html> ou bien le groupe des utilisateurs francophones du logiciels R <http://forums.cirad.fr/logiciel-R>.

Si R présente un inconvénient majeur, c'est son nom : taper « R » suivi de quelques mots clés dans son moteur de recherche préféré ne mène pas toujours vers une réponse pertinente. Pourquoi donc ne pas utiliser un moteur de recherche spécifiquement conçu pour faire des requêtes sur R : Sasha Goodman a créé <http://rseek.org>, un moteur de recherche décomposé en onglets thématiques (« Packages », « For Beginners », « Support »...) qui donne souvent des résultats pertinents. Le site web <http://www.rdocumentation.org> rassemble, quant à lui, la documentation de plus de 9 000 packages et permet une recherche rapide par mot clé ou bien avancée par nom de package, fonction, auteur, etc.

En conclusion, la présentation actuelle n'est pas exhaustive mais nous aurons 6 minutes pour vous présenter ces sites et d'autres, en espérant que vous découvrirez certains d'entre eux et qu'ils vous seront utiles.