

Introduction de nouveaux outils (learnr, Github classroom, ...) dans les cours de Science des Données en Biologie

G. Engels^a & Ph. Grosjean^a

^a Service d'Écologie Numérique des Milieux Aquatiques, Institut Complexys et Biosciences, Université de Mons (UMONS),
23 Place du Parc, 7000 Mons, Belgium
guyliann.engels@umons.ac.be



Introduction

Avec les nouvelles technologies, les biologistes sont confrontés à une quantité de données croissante qu'ils doivent réussir à interpréter de manière reproductible. Les outils nécessaires à cette compréhension dépassent de loin les statistiques classiques et mettent en avant l'importance de l'introduction de cours de sciences des données [1] au sein d'un cursus universitaire en biologie.

La crise de la reproductibilité [2] ainsi que l'ouverture des sciences (Open Science, Open Data) ont induit la recherche de nouveaux profils de chercheurs dans le milieu professionnel. Ils doivent être capable d'employer des outils tels que des programmes d'analyse et de traitement des données performants (R), d'un gestionnaire de version (Git) et d'une interface (Rstudio [3]) permettant l'analyse et la rédaction de rapport scientifique reproductible (Rmarkdown/RNotebook).

Outils et compétences à acquérir



Programme d'analyse des données



Packages pour traiter les données



Interface à R



Gestionnaire de versions



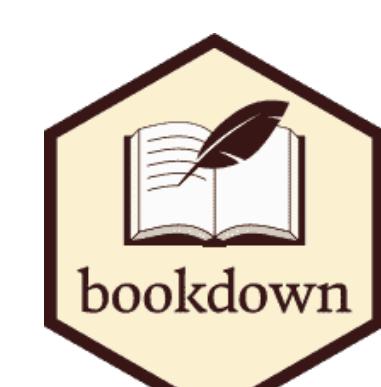
Partage des connaissances



Rédaction de rapports scientifiques

Contenu des cours centralisé, varié, en ligne et en constante amélioration

Ouvrage en ligne



Chaine YouTube : BioDataScience-Course

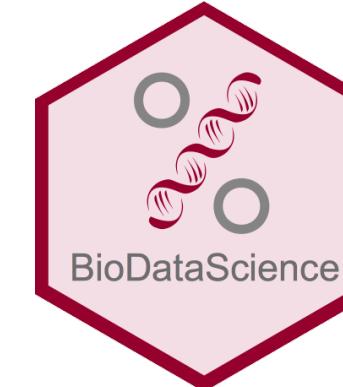


Mise à disposition d'outils performants et professionnels

Méthodes d'enseignement revisitées : classes renversées, travaux collaboratif (git)



Quizz interactif (learnr)



Outils proposés

Site en ligne

<http://biostatistics-course.sciviews.org>



Apprendre à analyser des données à l'UMONS

La science des données se développe dans un monde submergé d'informations en tous genres. En biologie, le flux grandissant des données nécessite la

Machine virtuelle complètement préconfigurée, auto-installable



SciViews Box



Système de collaboration performant

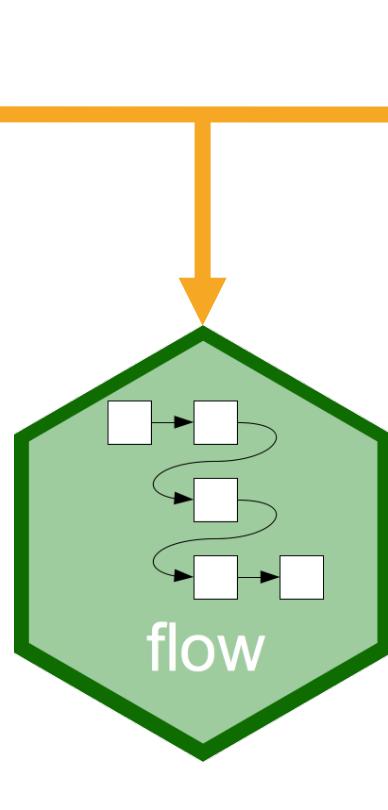


Identité SciViews

L'univers SciViews a pour objectif de fournir à des non-initiés des outils performants, utilisables dans le milieu professionnel et Open Source pour la recherche reproductible avec R. Outre la SciViews Box, d'autres outils sont développés comme les packages R chart, flow ou encore data.

Recherche reproductible facilitée

Chargement facilité des packages principaux



Opérateur « pipe » intuitif

Représentation graphique unifiée

Importation unifiée des données

L'ensemble des outils proposés est en cours d'élaboration et en continue évolution. Ils sont distribués sur CRAN & sur GitHub : <https://github.com/SciViews>

Processus d'apprentissage continu

L'Université de Mons (Belgique) fait évoluer ses cours de biostatistiques vers des cours de sciences des données avec un apprentissage réparti sur **4 années** pour un total de **16 crédits**.

Bachelier en Biologie

180 crédits

Bloc 1

Bloc 2

Bloc 3

SDD I

SDD II

Master en Biologie des Organismes et Ecologie (BOE)
Master en Biochimie, Biologie Moléculaire et cellulaire (BBMC)

120 crédits

Bloc 1

III SDD

IV SDD

V SDD

milieu professionnel

Cours	SDD I	SDD II	SDD III	SDD IV	SDD V	Total
Cours	25	15	15	10	10	75
Exercice	50	15	15	10	10	100
Crédits	6	3	3	2	2	16

- Science Des Données I : visualisation et inférence
- Science Des Données II : analyse et modélisation
- Science Des Données III : exploration et prédiction
- Science Des Données IV : pratique
- Science Des Données V : recherche reproductible

[1] Cleveland, W.S. 2001. "Data Science: An Action Plan for Expanding the Technical Areas of the Field of Statistics." *ISI Review* 69: 21–26. doi:[10.1111/j.1751-5823.2001.tb00477.x](https://doi.org/10.1111/j.1751-5823.2001.tb00477.x).

[2] Baker, M. 2016. "1,500 Scientists Lift the Lid on Reproducibility." *Nature* 533 (7604): 452–54. doi:[10.1038/533452a](https://doi.org/10.1038/533452a).

[3] RStudio and Shiny are trademarks of RStudio, Inc.