

Science des données biologiques



Rendons les étudiants actifs de leurs apprentissages

Guyliann Engels & Philippe Grosjean

Université de Mons, Belgique
Laboratoire d'Écologie numérique des Milieux aquatiques



<http://biodatascience-course.sciviews.org>
sdd@sciviews.org

Crise de la reproductibilité, Open Science, Open Data, Open Knowledge

Les statistiques “classiques” sont de plus en plus insuffisantes dans le contexte actuel de la recherche, à tel point qu’on parle maintenant d’une autre discipline : **la science des données**.

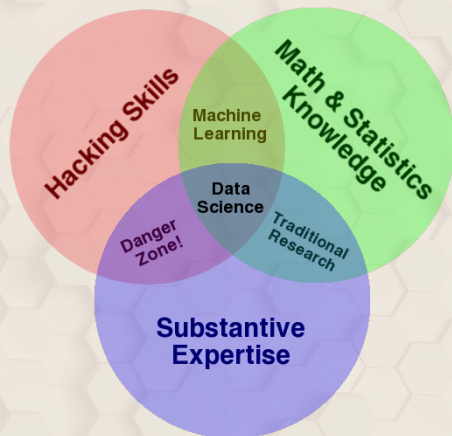


MONS

Science des données : à l'interface entre plusieurs disciplines

- La Science des Données, c'est la discipline qui s'intéresse à l'analyse de données *sous toutes ses formes*
- Très large et **interdisciplinaire** :
 - (Bio)statistiques et visualisation
 - Utilisation d'**outils informatiques**
 - Expertise dans le domaine (biologie)
- Il faut maîtriser simultanément les 3 domaines pour être un scientifique des données.

C'est notre objectif que nos étudiants montois en biologie se distinguent par une formation plus complète en "science des données biologiques" que dans les autres universités belges.



Un job recherché

1 Data Scientist



4.8 / 5

Job Score

4.2 / 5

Job Satisfaction

\$110,000

Median Base Salary

4,524

Job Openings

[View Jobs](#)

In 2017, [Glassdoor also ranked Data Scientist as the best job in America](#), with the same job score and the same median base salary, but slightly higher job satisfaction. In 2016, [Glassdoor ranked Data Scientist as no. 1 job in USA](#) for the first time, with median salary \$117K, and about 1,700 jobs listed.

Here are the top 5 jobs in 2018 and also other jobs in top 50 related to Analytics, Big Data, Data Science, according to Glassdoor:

2018 Rank	Job Title	Job Score	Job Satisfaction	Median Base Salary	Job Openings
1	Data Scientist	4.8	4.2	\$110,000	4,524

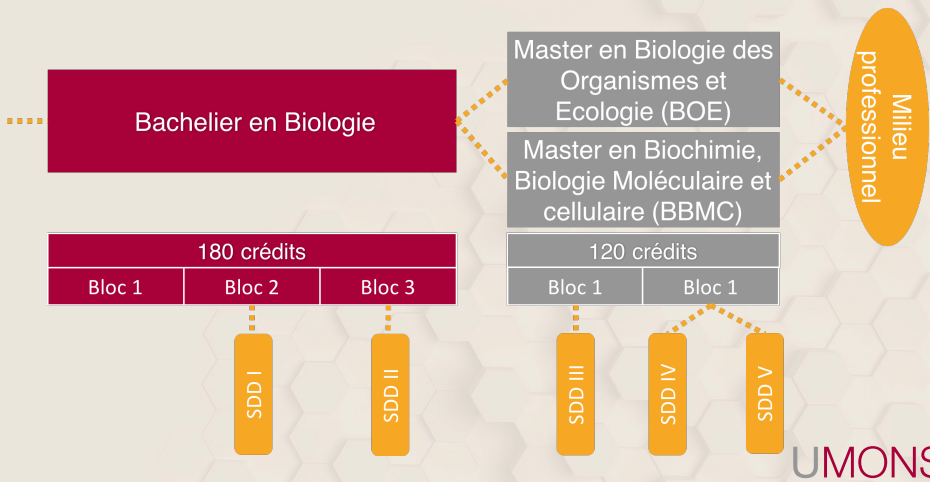
- Formation solide en science des données
 - “Vendeur” pour attirer des étudiants vers l'UMONS
 - Avantage à nos étudiants sur le marché de l'emploi

A data scientist is a
statistician who is useful.
– *Hadley Wickham*

- Classes renversées
- Cours et T.P. remplacés par des activités où les étudiants sont **(pro)actifs**
- Apprentissage en continu
- **Méthodes et matériel pédagogiques modernisés**
- Évolution vers un MOOC, version anglaise

Processus d'apprentissage continu

L'apprentissage est réparti sur 4 années pour un total de 16 crédits (200h en présentiel)



Université	Cours (h)	Exercices	Total (h)
UCLouvain	36	39	75
ULiège	65	30	95
UNamur	41	74	115
ULB	54	78	132
UMons (*)	75	125	200

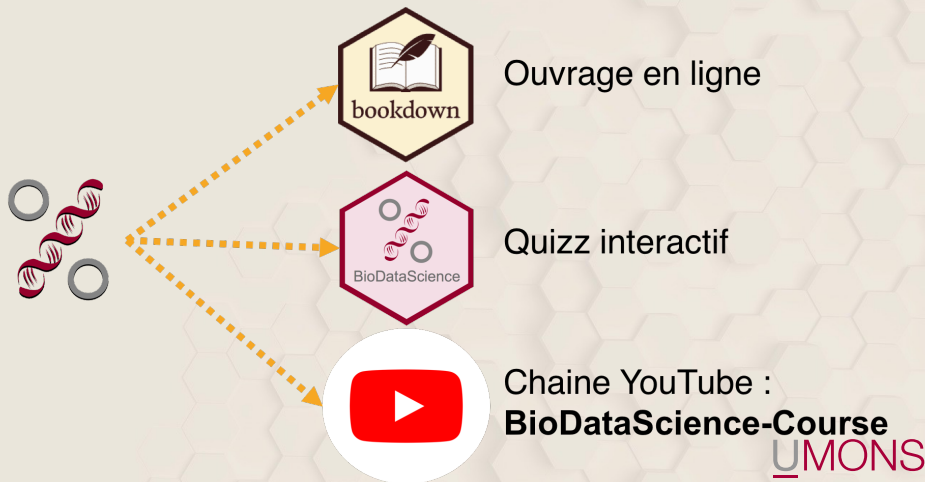
(*) Y compris SDD4 & SDD5 à option en Master

Et chez nos voisins flamands ?

Université	Cours (h)	Exercices	Total (h)
VUB	42	42	84
KUL	72	71	143
UAntwerpen	??	??	165
UGent	90	121	211

Processus d'apprentissage actif

Il n'y a pas que le nombre d'heures qui compte. *Nous voulons rendre cet enseignement le plus qualitatif possible !* Contenu des cours centralisé, varié, en ligne et en constante amélioration : <http://biodatascience-course.sciviews.org>



Amélioration du point de vue qualitatif

- Contenu numérique (vidéos, quizz, exercices en ligne)
- Enregistrement de **toutes les actions** des étudiants dans une base de données
- Feedback analysable pour déterminer ce qui fonctionne bien et moins bien
- Progression des étudiants suivie de près

Besoins

Cette année-ci, nous déployons la partie du cours 1 (75h en présenciel, 12 modules). Pour progresser, nous avons besoin d'outils techniques (e.g., annotation dans les vidéos), et d'une aide institutionnelle (SAP, informatique, pédagogues, main d'oeuvre pour l'aide à l'élaboration du matériel pédagogique)

Prenons le module 2 du cours de science des données I : visualisation et inférence (<http://biodatascience-course.sciviews.org/sdd-umons/visu1.html>)

Etudiant

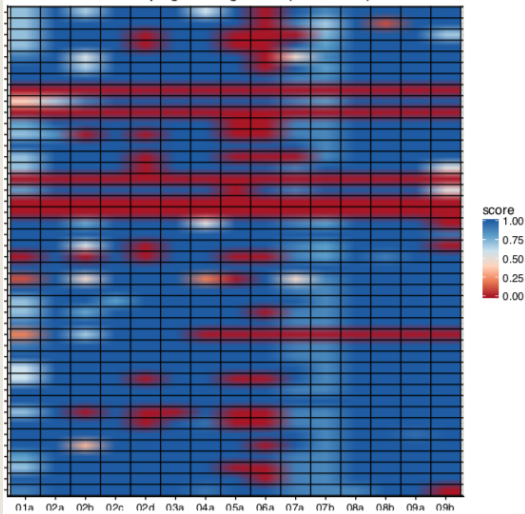
- Théorie sous diverse forme
- Exercice individuelle
- Projet par équipe

Professeur

- **leanrR**
- **GitHub Classroom**
- **GitHub**

Progression des étudiants via les quizz et exercices en ligne.

SDD I – progression générale (2018–11–25)



Progression des étudiants

Lors de la réalisation des exercices, toutes les données sont récupérées et stockées sur une base de données.

- Progression générale
- Progression individuelle
- Notez le taux de participation élevé !

Pour aller plus loin

Prenons le module 2 du cours de science des données I : visualisation et inférence (<http://phgrosjean.sciviews.org/bookdown-test/visu1.html>)

- shiny app
- H5p

Présentation d'un poster à la conférence : Rencontres R à Rennes 2018.

Introduction de nouveaux outils (learnr, Github classroom,. . .) dans un cours de Science des Données Biologiques

Grand intérêt marqué par les professeurs présents