Recommandations de cours sur DataCamp

École interdisciplinaire outils & méthodes (EIOM), 6e édition

L'équipe de l'EIOM

2026-08-10

Table des matières

Pourquoi utiliser DataCamp? 1
Comment utiliser ce guide 1
Bloc 1 : d'ici le 24 août 2
Bloc 2 : pendant la semaine intensive 3
Bloc 3 : pendant la session d'automne 3
Pour aller plus loin 3
Une dernière chose 3

Pourquoi utiliser DataCamp?

DataCamp est une plateforme d'apprentissage en ligne spécialisée dans les compétences en science des données, analyse de données et programmation. Elle offre une variété de cours interactifs qui couvrent des sujets comme la programmation en R et en Python, l'analyse statistique, la visualisation de données, l'apprentissage automatique (*machine learning*) et, depuis peu, l'intelligence artificielle générative.

L'EIOM offre gratuitement DataCamp à toutes les personnes étudiantes inscrites au programme, pour toute la session d'automne. C'est une occasion précieuse de développer des compétences essentielles sans frais supplémentaires, et elle ne se limite pas à la semaine intensive : votre accès reste actif jusqu'à la fin de la session.

Nous encourageons vivement l'apprentissage et la pratique du langage R sur la plateforme, peu importe le parcours que vous choisissez. R est un outil puissant pour l'analyse statistique et la visualisation de données, et le maîtriser vous donnera un avantage certain dans vos études comme dans votre carrière.

Certains parcours ne demandent aucune connaissance préalable en programmation. C'est le cas des parcours *Méthodes mixtes*, *Le cas : méthodes et inférence* et *Mise à niveau*. Il est toutefois probable que R soit mentionné pendant la semaine et que des exercices soient proposés. Ne vous inquiétez pas : même en partant de zéro, DataCamp vous permet de progresser à votre rythme.

Dans tous les cas, apprendre un langage de programmation en 2026 est une excellente idée pour toute personne aux études supérieures. La maîtrise d'un langage comme R ou Python renforce vos compétences analytiques et de traitement des données, et vous dote d'outils essentiels pour mener vos recherches de façon efficace et autonome. La programmation favorise également la résolution de problèmes complexes, l'automatisation des tâches répétitives et l'analyse de grands volumes de données, ce qui est aujourd'hui indispensable dans un très grand nombre de disciplines.

Comment utiliser ce guide

Ce guide est organisé en trois blocs, selon le moment où vous vous y mettez. Vous n'avez pas à tout faire, et surtout pas tout de suite.

1. **D'ici le 24 août** : le minimum utile pour arriver à l'aise le lundi matin. Environ 30 minutes par jour suffisent.
2. **Pendant la semaine intensive** : quoi consolider le soir, selon votre parcours.
3. **Pendant la session d'automne** : un plan de 12 semaines pour celles et ceux qui veulent vraiment apprendre R, sans se presser.

Tous les titres sont cliquables et mènent directement au cours sur la plateforme.

Bloc 1 : d'ici le 24 août

Il reste environ deux semaines. À 30 minutes par jour, cela représente à peu près 7 heures. Voici ce que nous vous suggérons d'en faire.

Pour tout le monde, peu importe le parcours

Si vous n'avez jamais programmé, commencez ici et n'allez pas plus loin. C'est déjà très bien.

1. Introduction to R
2. Intermediate R

Selon votre parcours

Mise à niveau : les outils et les méthodes dans le processus scientifique

Ce parcours ne demande aucun préalable technique. Si vous voulez tout de même prendre de l'avance, ces cours portent sur ce qui vient *avant* l'analyse, là où l'on bloque le plus souvent en pratique.

— Introduction to Importing Data in R
— Cleaning Data in R
— Exploratory Data Analysis in R

La science des données : modélisation et prédiction

C'est le parcours le plus exigeant sur le plan technique. Une base en R est fortement recommandée, et les deux cours ci-dessous préparent directement aux séances sur la régression puis sur l'apprentissage automatique.

— Introduction to the Tidyverse
— Introduction to Regression in R

Analyse causale et expérimentation

Il n'existe pas de cours d'inférence causale sur DataCamp qui soit à la hauteur de ce parcours. La meilleure préparation reste donc la régression, qui est le socle de toutes les méthodes qui seront abordées.

— Introduction to Regression in R
— Intermediate Regression in R

L'intelligence artificielle en recherche

Ces trois cours sont conceptuels et ne demandent aucune programmation. Ils totalisent moins de 5 heures et donnent le vocabulaire commun du parcours.

— Large Language Models (LLMs) Concepts (débutant, 2 h 15)
— Understanding Prompt Engineering (débutant, 1 h 11)
— Introduction to AI Agents (débutant, 1 h 30)

Méthodes mixtes : le meilleur des mondes et Le cas : méthodes et inférence

Aucun préalable technique n'est exigé. Si le volet outillé de l'analyse textuelle vous intéresse, ce cours en donne un bon aperçu.

— Introduction to Natural Language Processing in R

Bloc 2 : pendant la semaine intensive

Les journées sont chargées et personne ne s'attend à ce que vous fassiez des exercices tous les soirs. Ces cours sont là si vous voulez consolider une notion vue dans la journée.

- **Après une séance sur la visualisation** : Introduction to Data Visualization with ggplot2
- **Après une séance sur la manipulation de données** : Data Manipulation with dplyr
- **Après une séance de statistiques** : Introduction to Statistics in R
- **Après le midi-métho du mardi sur LaTeX, Markdown et Quarto** : Reporting with R Markdown. À noter que DataCamp n'a pas encore de cours dédié à Quarto; ce cours porte sur R Markdown, dont Quarto est le successeur direct, et les principes s'appliquent presque intégralement.
- **Après l'atelier du mardi sur les données textuelles** : Text Mining with Bag-of-Words in R

Bloc 3 : pendant la session d'automne

Votre accès à DataCamp reste actif jusqu'à la fin de la session. Pour celles et ceux qui préfèrent y aller progressivement, voici un plan de 12 semaines, à raison d'un chapitre ou deux par semaine.

- Semaine 1 : Introduction to R, chapitres *Intro to basics* et *Vectors*
- Semaine 2 : Introduction to R, chapitres *Matrices* et *Factors*
- Semaine 3 : Introduction to R, chapitres *Data frames* et *Lists*
- Semaine 4 : Intermediate R, chapitre *Conditionals and Control Flow*
- Semaine 5 : Introduction to the Tidyverse, chapitres *Data Wrangling* et *Grouping and Summarizing*
- Semaine 6 : Introduction to Data Visualization with ggplot2, chapitres *Introduction* et *Aesthetics*
- Semaine 7 : Introduction to Data Visualization with ggplot2, chapitres *Geometries* et *Themes*
- Semaine 8 : Introduction to Regression in R, chapitre *Simple Linear Regression*
- Semaine 9 : Introduction to Regression in R, chapitre *Predictions and model objects*
- Semaine 10 : Introduction to Regression in R, chapitre *Assessing model fit*
- Semaine 11 : Introduction to Regression in R, chapitre *Simple logistic regression*
- Semaine 12 : Intermediate Regression in R, chapitre *Multiple linear regression*

Pour aller plus loin

Apprentissage automatique. Si le parcours *Science des données* vous a donné le goût d'aller plus loin que ce qui a été couvert pendant la semaine :

- Machine Learning with caret in R
- Supervised Learning in R: Regression

Analyse de textes. Pour traiter et analyser des corpus textuels avec R :

- Text Mining with Bag-of-Words in R
- Introduction to Natural Language Processing in R

Intelligence artificielle et éthique. Pour prolonger la réflexion amorcée dans le parcours IA et dans les ateliers de l'après-midi :

- AI Ethics
- Generative AI Concepts

Une dernière chose

N'oubliez pas que vous faites désormais partie de la communauté Slack de l'EIOM. Si vous bloquez sur un exercice, posez la question : d'autres participantes et participants sont probablement passés par là, et les auxiliaires d'enseignement sont là pour vous accompagner jusqu'à la remise du travail final.

Bonne école!

L'équipe de l'EIOM