

## Plan de cours

DMO 3900 Informatique pour la démographie (3 cr.)

### Hiver 2023

**Plage horaire :** MERCREDI 9h00-11h50

**Local :** B-1215, pavillon 3200 Jean-Brillant

**RESPONSABLE :**

Prof. Nadine OUELLETTE

📍 C-5034, pavillon Lionel-Groulx

✉ [nadine.ouellette@umontreal.ca](mailto:nadine.ouellette@umontreal.ca)

Disponibilité : LUNDI 11h30-12h30

**AUXILIAIRE D'ENSEIGNEMENT :**

Pietro VIOLO

📍 C-5044, pavillon Lionel-Groulx

✉ [pietro.violo@umontreal.ca](mailto:pietro.violo@umontreal.ca)

Disponibilité 1 : MARDI 12h10-12h40

Disponibilité 2 : MERCREDI 17h00-17h30

## 1 Objectifs du cours

Le cours se veut une initiation aux logiciels R, SAS et Stata, qui sont largement utilisés en démographie, à la fois dans les milieux professionnel et académique. L'apprentissage des particularités du langage propre à chacun des trois logiciels constituera un sous-produit important du cours, bien que le principal objectif visé consistera à développer de solides compétences informatiques générales, incluant :

- la conception et la mise en oeuvre d'algorithmes pour résoudre des problèmes concrets ;
- l'adoption de bonnes habitudes en matière de programmation informatique, une activité se déroulant le plus souvent en groupe qu'en solo ;
- la familiarisation avec les structures conditionnelles et les boucles itératives.

Tous les exemples d'application s'inscriront dans le domaine de la démographie et mettront à profit des connaissances théoriques et méthodologiques déjà acquises par les étudiant.e.s dans d'autres cours de démographie (e.g., manipulation de données, calculs d'indicateurs, représentations graphiques pour la visualisation de données et de résultats).

Grâce aux compétences informatiques générales nouvellement développées dans ce cours, les étudiant.e.s seront ainsi plus à même d'apprendre et de maîtriser plus rapidement d'autres langages de programmation.

## 2 Méthodes pédagogiques

Les séances de cours se dérouleront en présentiel dans une salle informatique (B-1215). Les différents sujets seront d'abord introduits sous forme d'exposés magistraux, suivis de démonstrations faites par la

professeure avec le logiciel à l'étude. Afin de maximiser la consolidation des acquis, chaque séance de cours se clôturera par des exercices pratiques supervisés, à résoudre par les étudiant.e.s avec l'aide de la professeure et de l'auxiliaire d'enseignement, qui circuleront dans la salle informatique afin de répondre aux questions. À l'occasion, de courtes capsules vidéo pédagogiques préenregistrées portant sur des sujets ciblés pourraient être conçues et mises à la disposition des étudiant.e.s afin de compléter les contenus des séances de cours en présentiel.

Il convient de souligner que **l'apprentissage de nouveaux langages de programmation nécessite de la pratique sur une base régulière**. Entre deux séances de cours successives, les étudiant.e.s devront ainsi s'assurer de prendre le temps de mettre en pratique tous les sujets abordés en classe lors de la plus récente séance complétée. La maîtrise de tous les sujets précédents est incontournable pour l'acquisition de ceux de la semaine suivante.

### 3 Évaluation

**Examen intra** (à livres ouverts, sans ordinateur) – **35 %**

📅 22 février 2023, 9h00-11h00

📍 C-5143, pavillon Lionel-Groulx

Sujet : Programmation sous R

**Travaux pratiques** (individuels) – **40 %**, soit **8 %** chacun

Énoncés disponibles : TP #1 – 11 jan., TP #2 – 18 jan., TP #3 – 1 fév., TP #4 – 8 mars, TP #5 – 29 mars

Dates de remise : TP #1 – 18 jan., TP #2 – 25 jan., TP #3 – 8 fév., TP #4 – 22 mars, TP #5 – 12 avr.

Sujets : Hygiène (TP #1), programmation sous R (TP #2 et #3), sous Stata (TP #4), sous SAS (TP #5)

**Devoir** (individuel ou en groupe de deux étudiant.e.s) – **20 %**

Énoncé disponible : 25 janvier 2023

Date de remise : 20 février 2023, 23h59

Sujet : Programmation sous R

**Participation** – **5 %**

Participation aux séances de cours (en posant des questions, répondant aux questions, émettant des commentaires, etc.) et contribution au blogue du cours sur StudiUM.

### 4 Plagiat et intégrité intellectuelle

À l'Université de Montréal, le plagiat est sanctionné par le Règlement disciplinaire sur le plagiat et la fraude concernant les étudiant.e.s du premier cycle et les étudiant.e.s des cycles supérieurs. L'Université met à votre disposition d'excellentes ressources pour prévenir le plagiat et la fraude dans vos travaux et vos examens, consultez-les ici.

### 5 Calendrier du cours

**Semaine 1** (11 janvier)

- Présentation et discussion du plan de cours, des méthodes pédagogiques et de l'évaluation
- Hygiène informatique : les principes de base
- Qu'est-ce que la programmation informatique et qu'entend-on par ses « bonnes pratiques » ?

- Le logiciel R et le convivial RStudio : description sommaire du logiciel R, anatomie de la session de travail, rubriques d'aide, etc.

\* **Travail pratique #1** : remise de l'énoncé

### Semaine 2 (18 janvier)

- Scalaires, vecteurs et matrices : définitions et opérations de base
- Programmation sous R, partie 1 : premiers pas (affectation, objets, fonctions internes, sauvegarde, ...)
- Bonnes pratiques (plan, terminologies, présentation, commentaires)

\* **Travail pratique #2** : remise de l'énoncé

\*\*\* **Travail pratique #1** : remise des copies

### Semaine 3 (25 janvier)

- Programmation sous R, partie 2 : structures de données sous R, importation et exportation de données
- Bonnes pratiques (séquence d'exécution)

\* **Devoir** : remise de l'énoncé

\*\*\* **Travail pratique #2** : remise des copies

### Semaine 4 (1<sup>er</sup> février)

- Programmation sous R, partie 3 : installation et utilisations de *packages*, conditions, boucles itératives (et à pré-condition), *arrays*
- Bonnes pratiques (présentation, commentaires)

\* **Travail pratique #3** : remise de l'énoncé

### Semaine 5 (8 février)

- Programmation sous R, partie 4 : outils graphiques, quelques mots sur *tidyverse*
- Bonnes pratiques (visualisation)

\*\*\* **Travail pratique #3** : remise des copies

### Semaine 6 (15 février)

- Programmation sous R, partie 5 : créer ses propres fonctions, révision en vue de l'examen intra
- Bonnes pratiques (organisation)

### Semaine 7 (22 février)

- **Examen intra** : 9h00 à 11h00

\*\* **Devoir** : remise des copies (**20 février 2023**)

### Semaine 8 (1<sup>er</sup> mars) \*\*\* Pas de cours : période d'activités libres \*\*\*

### Semaine 9 (8 mars) \* avec la Prof. Sophia CHAE \*

- Programmation sous Stata, partie 1

\* **Travail pratique #4** : remise de l'énoncé

### Semaine 10 (15 mars) \* avec la Prof. Sophia CHAE \*

- Programmation sous Stata, partie 2

### Semaine 11 (22 mars)

- Programmation sous SAS, partie 1 : description sommaire du logiciel, anatomie de la session de travail, rubriques d'aide, premiers pas

\*\*\* **Travail pratique #4** : remise des copies

### Semaine 12 (29 mars)

- Programmation sous SAS, partie 2 : étape DATA et importation de fichiers de données (procédure IMPORT)

\* **Travail pratique #5** : remise de l'énoncé

### Semaine 13 (5 avril)

- Programmation sous SAS, partie 3 : fusion de bases de données (procédure SQL)

### Semaine 14 (12 avril)

- Activité spéciale en programmation informatique située à l'intersection des *bots*, de Twitter, de R et de l'horloge démographique

\*\*\* **Travail pratique #5** : remise des copies

## 6 Ressources utiles

CODY, Ron (2018). **Learning SAS by example : A programmer's guide**, Second Edition, Cary (NC), SAS Institute, 536 pages.

DELWICHE, Laura D. et SLAUGHTER, Susan J. (2012). **The little SAS book : A primer**, Fifth Edition, Cary (NC), SAS Institute, 380 pages.

KABACOFF, Robert I. (2015). **R in action**, Second Edition, Shelter Island (NY), Manning Publications Co., 579 pages.

LONG, J.D et TEETOR, Paul (2019). **R Cookbook**, 2nd Edition, O'Reilly Media, 413 pages. En accès libre [ici](#).

MUENCHEN, Robert A. (2011). **R for SAS and SPSS Users**, Second Edition, New York (NY), Springer Series in Statistics and Computing, Springer, 686 pages. Une ancienne version intégrale est disponible [ici](#).

PRESSAT, Roland (1983). **L'analyse démographique**, Paris, Presses universitaires de France, 4<sup>e</sup> édition, 295 pages.

PRESTON, Samuel H., HEUVELINE, Patrick et GUILLOT, Michel (2000). **Demography. Measuring and modeling population processes**, Oxford, Blackwell Publishers, 291 pages.

R Core Team (2016). **R Data Import/Export**, Vienne, R Foundation for Statistical Computing. Une version électronique intégrale est disponible [ici](#).

VENABLES, W. N., SMITH, D. M. et le R Core Team (2018). **An introduction to R**, 99 pages. En accès libre [ici](#).

WACHTER, Kenneth W. (2014). **Essential demographic methods**, Cambridge, Harvard University Press, 288 pages.

WICKHAM, Hadley et GROLEMUND, Garrett (2016). **R for Data Science**, O'Reilly Media, 474 pages. En accès libre [ici](#).