

Visualisation des données (FAS 1003)

Florence Vallée-Dubois

Automne 2021

Courriel: florence.vallee-dubois@umontreal.ca

Site: florencevdubois.github.io

Heures de bureau: Sur rendez-vous

Horaire du cours: Mercredis, 16h-19h

Bureau: C-3114, pavillon Lionel-Groulx

Salle de classe: Pavillon Lionel-Groulx, local C-4019

Version: août 2021

Étant donnée la situation sanitaire, ce plan de cours est sujet à changements sans préavis. Le cours est prévu en mode présentiel, mais il se peut qu'il bascule en mode hybride ou à distance si les conditions sanitaires devaient se dégrader.

Table des matières

1. Approche et objectifs
2. Pédagogie
3. Évaluations
4. Responsabilités
5. Présence en classe
6. Calendrier
7. Politique concernant les retards
8. Auxiliaires et délai de correction
9. Règlements pédagogiques
10. Prévention du plagiat
11. Le harcèlement, y compris à caractère sexuel
12. Santé mentale
13. Bibliothécaire

1. Approche et objectifs

La visualisation des données est une compétence de plus en plus convoitée, non seulement en recherche universitaire, mais aussi dans des secteurs d'emplois comme les communications, l'économie, les politiques publiques et le journalisme. Au croisement du design et des sciences des données, elle permet d'illustrer des phénomènes complexes, de vulgariser des questions de recherche, de rapporter des résultats statistiques abstraits et de déceler des erreurs dans une banque de données.

L'objectif du cours *FAS 1003: Visualisation des données* est d'initier les étudiant.e.s à la visualisation des données en général, et des données sociales en particulier. Pour ce faire, nous étudierons les bases théoriques de la visualisation et du design, puis approfondirons le fonctionnement du langage **R** et de sa librairie **ggplot2**, qui permet de visualiser un large éventail de données. Les étudiant.e.s seront amené.e.s à manipuler différents types de données (numériques, textuelles, catégorielles, géographiques, etc.) et à créer différents types de graphiques (univariés, multivariés, des diagrammes à points, à bandes, des cartes, etc.).

Au terme de ce cours, les étudiant.e.s devraient être en mesure (1) comprendre l'utilité et l'importance de la visualisation des données pour différents secteurs; (2) de manipuler des banques de données pour préparer la visualisation; (3) de sélectionner le type de graphique approprié pour visualiser un phénomène et (4) réaliser ces visualisations.

Disponibilités: Je suis disponible sur rendez-vous (en personne le mercredi / par visioconférence tous les autres jours ouvrables). Si vous souhaitez céder une rencontre, contactez-moi par courriel.

Questions courtes: Je peux répondre aux questions courtes par courriel (par ex.: demande de reformuler une notion, aide avec un petit bug informatique, demande d'information rapide concernant une évaluation).

Délai de réponse aux courriels: Je garantis une réponse aux courriels dans les 48h suivant leur réception. Assurez-vous de prévoir ce délai si vos questions concernent une évaluation à venir.

2. Pédagogie

Le cours repose sur l'enseignement magistral, les exercices en classe, les lectures et les devoirs à la maison.

Le cours est prévu entièrement en mode présentiel, mais il se peut qu'il bascule en mode hybride ou à distance si les conditions sanitaires devaient se dégrader.

La partie pratique du cours repose sur l'utilisation du langage **R**. Il est attendu que les étudiant.e.s se présentent en classe muni.e.s de leur ordinateur portable, sur lequel devrait être installés et fonctionnels:

- **R:** Pour l'installer, suivre les indications pour le bon système d'exploitation [ici](#).
- **R Studio:** Pour l'installer, suivre les indications pour le bon système d'exploitation [ici](#) (*R Studio Desktop Free*).

La partie théorique du cours repose sur des ouvrages ou chapitres de livres qui seront rendus disponibles sur Studium ou auxquels il sera possible d'accéder via URL (voir plus bas dans l'horaire des séances).

Outre l'enseignement, le cours comportera les évaluations suivantes:

- 4 mini travaux pratiques, réalisés en partie en classe (4 x 5 points = 20 points)
- 2 exercices de visualisation, réalisés en partie en classe (2 x 15 points = 30 points)
- 1 travail de mi-session pour 20 points
- 1 travail final (incluant une présentation orale) pour 30 points

Total = 100 points

3. Évaluations

3.1 Minis travaux pratique (5 points chaque) et Exercices de visualisation (15 points chaque)

Les **minis travaux pratiques** visent à exercer vos compétences avec les outils technologiques utilisés en classe. Ces travaux me servent aussi à mesurer votre cheminement. Ils pourront en grande partie être réalisés en classe, puisque des ateliers pratiques sont prévus à presque toutes les séances de cours. Je serai donc présente pour vous accompagner et répondre à vos questions.

Les **exercices de visualisation** servent à mesurer votre acquisition des notions de visualisation des données plus spécifiquement. Tout comme pour les travaux pratiques, les exercices pourront être réalisés (ou du moins, entamés) en classe. Ils devraient cependant requérir un peu plus de travail que les travaux pratiques, c'est pourquoi ils comptent pour plus de points.

Les minis travaux pratiques et les exercices de visualisation devront être réalisés individuellement. Vous pouvez vous entraider, mais chaque étudiant.e doit remettre son propre travail. Plus d'informations sur ces évaluations vous seront fournies en classe et les documents explicatifs seront disponibles sur Studium.

Les travaux pratiques et les exercices de visualisation sont à remettre **sur Studium** au plus tard le **vendredi suivant le cours, à 23h59**. Veuillez consulter le calendrier du cours et la section sur les dates importantes (plus bas) pour connaître les jours exacts.

- Mini travail pratique 1 (5 pts): Introduction à R
- Mini travail pratique 2 (5 pts): Outils de travail
- Mini travail pratique 3 (5 pts): Les rudiments de la visualisation des données avec `ggplot`
- Mini travail pratique 4 (5 pts): Graphiques multivariés
- Exercice de visualisation 1 (15 points): Graphiques univariés
- Exercice de visualisation 2 (15 points): Graphiques multivariés, modèles et incertitude

3.2 Travail de mi-session (20 points) et Travail final (30 points)

Le travail de mi-session et le travail final visent à vous faire raconter une histoire visuelle avec des données. Ces travaux devront être réalisés individuellement. Vous devrez:

- Sélectionner une des banques de données mises à votre disposition pour vulgariser un phénomène de votre choix;
- Sélectionner une série de graphiques univariés et multivariés à concevoir;

- Réaliser ces visualisation en tenant compte des notions vues en classe;
- Communiquer votre histoire de façon claire et concise en y ajoutant du texte et des chiffres.

Le but de ces travaux est (1) de vulgariser un phénomène en images, en mots et en chiffres. Pour ce faire, les étudiant.e.s, devront démontrer qu'ils sont en mesure (2) de manipuler différents types de données dans le but de les visualiser et (3) d'exécuter différents types de visualisations.

Les deux évaluations auront le même format, mais le travail final devra intégrer un plus grand nombre d'acquis. C'est pourquoi il compte pour plus de points. Il comprend également une présentation, contrairement au travail de mi-session.

Plus d'informations sur ces deux évaluations vous seront fournies en classe et un document explicatif détaillé sera disponible sur Studium. Le travail de mi-session est à remettre au plus tard le vendredi **15 octobre 2021, à 23h59**. Le mercredi 13 octobre, il n'y aura pas de cours magistral, seulement une séance d'atelier en classe pour vous aider à avancer le travail de mi-session.

La séance de présentation des travaux finaux se tiendra le **1er décembre 2021**. La semaine précédente (le mercredi 24 novembre), il n'y aura pas de cours magistral, seulement une séance d'ateliers en classe pour vous aider à avancer le travail final. Les travaux finaux devront être remis au plus tard le **3 décembre 2021, à 23h59**. Vous bénéficierez donc de 2 jours suivant les présentations pour apporter des corrections à votre travail, si nécessaire.

4. Responsabilités

En tant que votre enseignante pour ce cours, je m'engage à assumer certaines responsabilités:

- Faire preuve de motivation et d'engagement dans mon enseignement.
- Me rendre disponible pour répondre à vos questions concernant le contenu du cours et les évaluations.
- Être à votre écoute ou vous diriger vers les ressources disponibles pour répondre à tout enjeu extra-curriculaire qui pourrait affecter votre réussite du cours.
- Interagir avec vous de façon respectueuse.

En échange, je m'attends à ce que vous assumiez aussi certaines responsabilités. La réussite d'un cours dépend de l'engagement de l'enseignant.e, mais aussi de celle des étudiant.e.s.

- Être présent.e et démontrer un intérêt pour la matière enseignée.
- Faire le travail demandé au moment demandé. Votre progression dépend en grande partie de votre assiduité au travail.
- Poser des questions lorsqu'une notion ou une évaluation n'est pas claire.
- Interagir avec moi et avec vos collègues de façon respectueuse.

5. Présence en classe

La présence en classe n'est pas obligatoire à proprement dit. Je ne prendrai pas les présences lors des cours (sauf exceptions, voir plus bas). Or, votre présence pourrait être garante de votre réussite dans ce cours. La visualisation des données comprend plusieurs aspects techniques. Certaines de ces notions ne peuvent pas facilement être acquises à la simple lecture des notes de cours. Voir des démonstrations et s'entraîner à réaliser les visualisations est extrêmement important. **C'est pourquoi l'horaire prévoit des moments de travail pratique à presque toutes les séances.** Vous pourrez profiter de ces moments pour avancer vos travaux et poser des questions sur les notions vues en classe.

Votre présence en classe est donc grandement encouragée. Il est bien sûr possible que des imprévus vous empêchent d'assister au cours. Dans cette éventualité, il est fortement recommandé de consulter les notes de cours et de faire les exercices pratiques pour vous tenir à jour et être en mesure de suivre lors de la séance qui suivra. Vous pouvez bien sûr me poser des questions sur la matière que vous avez manquée, mais il est essentiel que vous ayez d'abord fait l'effort de consulter les notes et complété les exercices (s'il y a lieu). Je ne peux pas « ré-enseigner » un cours ou une section entière de cours à un.e étudiant.e qui se serait absenté.e.

Exceptions: La présence est requise à seulement deux reprises durant la session.

1. Lors de la séance du 13 octobre, qui est une séance d'atelier pour le travail de mi-session. Vous aurez l'opportunité de poser des questions et d'obtenir mon aide lors de cette séance. Des points de présence seront alloués dans la note du travail de mi-session.
2. Lors de la séance de présentation des travaux finaux, le 1er décembre. La présence est bien sûr requise puisque vous présenterez vos travaux. Des points de participation seront aussi alloués aux personnes qui poseront des questions sur les travaux de leurs collègues. Tous ces points seront inclus dans la note du travail final.

Pour ces deux séances, les seules absences permises sont celles qui auront été approuvées via le formulaire d'absence à une évaluation (disponible sur [Synchro](#)).

6. Calendrier

6.1 Calendrier détaillé

1 1er septembre 2021: Introduction

Enseignement	Présentation détaillée du plan de cours; Explication des évaluations; Introduction à R et R Studio: espace de travail, ouverture de données et logique générale
Travail pratique	Commencer le mini travail pratique 1
Objectifs	Comprendre le fonctionnement et les attentes du cours; se familiariser avec R
Lectures	Plan de cours; Arel-Bundock (2021), chapitre 21 (voir plus bas pour le lien vers l'ouvrage)

2 8 septembre 2021: Introduction à R et à la visualisation (suite); Outils de travail

- | | |
|------------------|---|
| Enseignement | Retour sur le travail pratique 1; Suite de l'introduction à R; Outils de travail pour la session: R Markdown et autres |
| Travail pratique | Commencer le mini travail pratique 2 |
| Objectifs | Savoir naviguer R; Connaître les outils de travail et les modes de remises des travaux |
| Lectures | Healy (2019), chapitre 2; Si ce n'est pas fait: Arel-Bundock (2021), chapitre 21 (voir plus bas pour le lien vers les ouvrages) |
- 3 15 septembre 2021: Le graphisme**
- | | |
|--------------|---|
| Enseignement | Retour sur le travail pratique 2; Principes de la visualisation et du graphisme; Conférence d'une graphiste invitée |
| Objectifs | Comprendre les principes de la cognition qui influencent la visualisation; Connaître les principes du "beau" |
| Lectures | À déterminer |
- 4 22 septembre 2021: Les rudiments de la visualisation des données avec ggplot**
- | | |
|------------------|---|
| Enseignement | La librairie ggplot2 |
| Travail pratique | Commencer le mini travail pratique 3 |
| Objectifs | Connaître la "logique interne" de ggplot2 |
| Lectures | Arel-Bundock (2021), p.19 à 26.; Healy (2019), chapitre 3 |
- 5 29 septembre 2021: Les graphiques univariés et les distributions**
- | | |
|------------------|--|
| Enseignement | Retour sur le travail pratique 3; Les graphiques à bandes, à lignes, à points pour visualiser une seule variable; les graphiques pour représenter la densité |
| Travail pratique | Commencer l'exercice de visualisation 1 |
| Objectifs | Savoir naviguer différents types de graphiques univariés |
| Lectures | Arel-Bundock (2021), p.26 à 29; Healy (2019), chapitre 4 |
- 6 6 octobre 2021: Aspects visuels**
- | | |
|------------------|---|
| Enseignement | Retour sur l'exercice de visualisation 1; Thèmes dans ggplot, modification des axes, des couleurs, etc. |
| Travail pratique | Exercices en classe |
| Objectifs | Pouvoir modifier les aspects visuels qui se trouvent autour des données |
| Lectures | Healy (2019), chapitres 5 et 8; Document explicatif sur le travail de mi-session |
- 7 13 octobre 2021: Atelier sur le travail de mi-session**
- | | |
|-----------|---|
| Objectifs | Avancer le travail; Poser des questions |
|-----------|---|
- 18-22 octobre : Semaine de relâche**
- 8 27 octobre 2021: Graphiques multivariés**
- | | |
|------------------|--|
| Enseignement | Retour sur le travail de mi-session; Graphiques à plus d'une variable (deux ou plusieurs variables continues, deux ou plusieurs variables catégorielles, mélange de types, etc.) |
| Travail pratique | Commencer le mini travail pratique 4 |
| Objectifs | Ajouter des types de graphiques à son coffre à outil; Savoir quand utiliser les graphiques multivariés (vs univariés) |

Lectures Arel-Bundock (2021), p.29 à 36

9 3 novembre 2021: Modèles statistiques et incertitude

Enseignement Retour sur le mini travail pratique 5; Modèle de régression linéaire; Prédic-
tions; Visualisation des marges d'erreurs
Travail pratique Commencer l'exercice de visualisation 2
Objectifs Savoir visualiser les modèles statistiques et intégrer l'incertitude aux
graphiques
Lectures Arel-Bundock (2021), p.29 à 36, Healy (2019), chapitre 6

10 10 novembre 2021: Cartes

Enseignement Retour sur l'exercice de visualisation 2; Principes de la géolocalisation et
de la cartographie; Cartes avec ggplot2
Travail pratique Exercices en classe
Objectifs Se familiariser avec les données géographiques et la création de cartes
Lectures Healy (2019), chapitre 7; Document explicatif du travail final

11 17 novembre 2021: Conclusion – Utilisation appliquée de la visualisation des données

Enseignement Semaine de "buffer"; Retour sur certaines notions moins bien comprises;
Conférence d'un journaliste en visualisation des données

12 24 novembre 2021: Atelier sur le travail final

Objectifs Avancer le travail; Poser des questions

13 1er décembre 2021: Présentations

Objectifs Présenter votre travail; Obtenir des commentaires pour améliorer votre tra-
vail en vue de la remise 2 jours plus tard

6.2 Ouvrages

Les extraits à lire dans Arel-Bundock (2021) se trouvent [ici](#) (gratuitement).

Les extraits à lire dans Healy (2019) se trouvent [ici](#) (gratuitement – cliquez sur “Menu” pour voir les chapitres).

6.3 Dates importantes

Date	Élément important
3 sept., 23h59	Remise du mini travail pratique 1
10 sept., 23h59	Remise du mini travail pratique 2
21 sept.	Date limite pour la modification d'un choix de cours
24 sept., 23h59	Remise du mini travail pratique 3
1er oct., 23h59	Remise de l'exercice de visualisation 1
13 oct.	Temps pour travailler sur votre travail de mi-session
15 oct., 23h59	Remise du travail de mi-session
Semaine du 18 oct.	Se reposer
29 oct., 23h59	Remise du mini travail pratique 4
5 nov.	Date limite pour l'abandon d'un cours
5 nov., 23h59	Remise de l'exercice de visualisation 2
24 nov.	Temps pour travailler sur votre projet final
1er déc.	Présentation du travail final en classe
3 déc., 23h59	Remise du travail final

7. Politique concernant les retards

Les travaux en retard seront acceptés jusqu'à 4 jours après la date de remise, sauf le travail final, qui pourra être accepté avec un maximum de 7 jours de retard. Les dates de remise sont les vendredis à 23h59, ce qui signifie que vous avez jusqu'au **mardi suivant** (23h59) pour rendre vos travaux (vendredi suivant, pour le travail final). La pénalité est de 10 points de pourcentage pour chaque jour de retard (samedi, dimanche, lundi, mardi).

- Un mini travail pratique rendu le samedi perd 0.5/5; le dimanche, 1/5, et ainsi de suite.
- Un exercice de visualisation rendu le samedi perd 1/10; le dimanche, 2/10, et ainsi de suite.
- Un travail de mi-session rendu le samedi perd 2/20; le dimanche, 4/20, et ainsi de suite.

À chaque cours, nous effectuerons un retour en classe sur l'évaluation de la semaine précédente (s'il y a lieu).

La pénalité pour les travaux finaux remis en retard est la même: 10 points de pourcentage (3/30) par jour de retard. La note de 0 sera attribuée aux étudiant.e.s qui dépassent ce délai.

8. Auxiliaires et délai de correction

Selon le nombre d'étudiants inscrits, il est possible qu'un.e auxiliaire d'enseignement travaille avec nous. Elle/Il serait alors responsable de la correction de vos travaux, et serait parfois présent.e en classe pour vous aider lors des ateliers pratiques.

Étant donné la fréquence des évaluations, je ne garantis pas de délai spécifique pour la correction. Je (ou l'auxiliaire) ferai mon possible pour rendre les notes dans un délai raisonnable. Ceci dit,

nous reviendrons sur chaque évaluation lors du cours suivant. La correction ne devrait donc pas être une surprise.

9. Règlements pédagogiques

Veillez prendre note que le trimestre se termine le 22 décembre 2021 (incluant la période des examens). Aucune demande d'examen différé (dans notre cas, de travail final différé) ne sera acceptée sans motif valable. Nous entendons par motif valable, un motif indépendant de la volonté de l'étudiant, tel que la force majeure, le cas fortuit ou une maladie attestée par un certificat de médecin.

L'étudiant.e doit motiver, en remplissant le formulaire disponible dans le Centre étudiant, toute absence à une évaluation ou à un cours faisant l'objet d'une évaluation continue dès qu'il est en mesure de constater qu'il ne pourra être présent.e à une évaluation et fournir les pièces justificatives dans les sept jours suivant l'absence.

L'étudiant.e doit motiver, en remplissant le formulaire disponible dans le Centre étudiant, toute demande de délai pour la remise d'un travail dès qu'il est en mesure de constater qu'il ne pourra remettre à temps le travail et fournir les pièces justificatives.

La pénalité imposée pour les retards dans la remise des travaux est de 10 points de pourcentage par jour. Cette pénalité est calculée en déduisant 10 points de pourcentage à la note obtenue pour le travail en question. À noter, il s'agit de la politique « par défaut »; les enseignant.e.s sont libres d'imposer une pénalité plus élevée s'ils le désirent. Voir la section « Politique concernant les retards » pour plus d'informations sur la politique dans le cadre de ce cours.

10. Prévention du plagiat

Le Département porte une attention toute particulière à la lutte contre le plagiat, le copiage ou la fraude lors des examens. Le plagiat consiste à utiliser de façon totale ou partielle, littérale ou déguisée le texte d'autrui en le faisant passer pour sien ou sans indication de référence à l'occasion d'un travail, d'un examen ou d'une activité faisant l'objet d'une évaluation. Cette fraude est lourdement sanctionnée.

Tous les étudiant.e.s sont invité.e.s à consulter le site web [suivant](#) et à prendre connaissance du Règlement disciplinaire sur le plagiat ou la fraude concernant les étudiants. Plagier peut entraîner un échec, la suspension ou le renvoi de l'Université.

11. Le harcèlement, y compris à caractère sexuel

Il incombe à chaque membre de la communauté universitaire de se conduire avec respect en tout temps envers tout le monde. En particulier, l'Université de Montréal s'engage à créer un milieu accueillant et sécuritaire pour toutes et tous, quelle que soit leur identité.

Les documents suivants ont des démarches pratiques à suivre :

- Si vous pensez que vous vivez du harcèlement, consultez cette [ressource](#).
- Si on s'est confié à vous ou si vous êtes témoin de harcèlement, consultez cette [ressource](#).
- Pour toute autre question, cliquez [ici](#).

12. Santé mentale

Une [étude de la FAÉCUM en 2016](#) a révélé que les étudiant.e.s de l'Université de Montréal, surtout les étudiants de premier cycle, souffraient davantage de détresse psychologique que le reste de la population. Si vous vous sentez isolé.e, déprimé.e, épuisé.e ou avez des pensées suicidaires, sachez qu'il existe des ressources. Voici une liste de celles-ci sur le campus et dans les environs:

- Les étudiants ont accès à un service [24 heures/7 jours](#) offert par l'Alliance pour la santé étudiante au Québec. Le numéro est le suivant : 1-833-851-1363.
- Les [sentinelles](#) de l'UdeM: « La sentinelle est une présence attentive, ouverte et empathique pour accueillir les demandes d'aide spontanées. Elle sait repérer les signes de détresse chez un.e étudiant.e qui fait un pas vers elle et est proactive à offrir son soutien. » Pour voir la liste complète des sentinelles de la Faculté des Arts et des Sciences, cliquez [ici](#). Par bâtiment, cliquez [ici](#).
- Le [Centre de santé et de consultation psychologique](#) offre des services de consultation psychologique à prix modique pour la communauté universitaire et des [trousses d'information](#) sur différents thèmes liés à la santé psychologique.
- La [Clinique universitaire de psychologie](#) offre des services de consultation psychologique pour les citoyens de la région métropolitaine.
- Vous pouvez aussi avoir accès aux services du [Centre de services psychologiques de l'UQÀM](#).
- Le [Centre de psychothérapie et de formation Desjardins](#) offre des services de psychothérapie par des jeunes psychologues.
- En cas d'urgence, vous pouvez consulter le [bottin de ressources](#).
- Le [CÉSAR](#) offre des évaluations neuropsychologiques individuelles pour déceler des difficultés d'apprentissage et développer des stratégies d'intervention.
- Tous les liens sont rassemblés sur cette page [Ça va aller](#), sous l'onglet *Services de soutien*.

13. Bibliothèque

N'hésitez pas à profiter des services de Caroline Patenaude, bibliothécaire disciplinaire spécialisée en données statistiques. Vous pouvez lui [écrire](#) ou consulter son [guide internet](#).