

PLAN DE COURS¹

SIGLE ET TITRE DU COURS : IAF2000 – Statistiques avancées et outils d'analyse de données

GROUPE : 05 - Formation à distance

ENSEIGNANT² : Jérôme Deschênes

PRÉSENTATION DE L'ENSEIGNANT

Jérôme Deschênes est professeur régulier en intelligence d'affaires. Il est directeur du Laboratoire sur l'acceptabilité sociale de l'exploitation des ressources naturelles (LASERN). Il est professeur à l'UQAT depuis novembre 2020. Il a complété son doctorat (Ph. D.) en comptabilité ainsi que sa maîtrise en administration des affaires (MBA) en responsabilité sociale des organisations à l'Université Laval. Ses recherches portent sur la gouvernance d'entreprise, la responsabilité sociale des entreprises, le processus d'audit et la structure du risque financier. Il enseigne notamment l'intelligence d'affaires, les méthodes quantitatives, la finance corporative et la comptabilité financière.

ACCEPTATION DU DÉPARTEMENT³

HUMAINE
CRÉATIVE
AUDACIEUSE

¹ [Procédure relative aux plans de cours](#)

² Dans ce document, le générique masculin est utilisé sans aucune discrimination et uniquement dans le but d'alléger la lecture du texte.

³ Conformément à la [Procédure relative aux plans de cours](#), un plan de cours qui n'a pas été accepté par le département ne peut être présenté aux étudiants.

OBJECTIFS DE LA FORMATION DE 1ER CYCLE

La formation de premier cycle trouve sa spécificité dans les caractéristiques suivantes :

- elle vise le développement chez l'étudiant de ses capacités d'analyse et de synthèse par l'étude relativement approfondie d'une discipline ou d'un champ d'études dans les aspects aussi bien théoriques qu'empiriques et en misant sur l'appropriation des fondements et de la méthodologie propre à cette discipline ou à ce champ d'études ;
- dans la perspective d'une éducation permanente, elle développe la capacité d'apprendre par soi-même de façon continue ;
- elle le rend capable de s'adapter facilement aux changements, de relier son champ de compétence aux autres spécialités et de collaborer avec des tiers ;
- elle le rend capable de discerner la valeur objective des affirmations qu'il fait ou qui lui sont faites, de bien comprendre, interpréter et commenter l'information ;
- elle développe chez lui une compétence professionnelle ou scientifique qui le rend apte à intervenir efficacement et à mesurer la portée sociale et éthique de ses activités ;
- elle l'amène à maîtriser le langage propre à son domaine de connaissances, à pouvoir produire un discours cohérent et pertinent, à l'articuler de façon précise, claire et concise, tant à l'écrit qu'à l'oral et, ainsi, à être capable de communiquer ses connaissances dans son milieu professionnel ou scientifique et dans l'ensemble de la société ;
- elle développe son esprit d'initiative et sa créativité, qui le rendent actif dans son milieu et l'amènent à appliquer ses connaissances à des situations et des problèmes nouveaux.

DESCRIPTION DU COURS

Introduction aux bases de données structurées. Initiation à la programmation informatique. Langages de programmation. Initiation à R. Optimisation numérique. Théorie des graphes. Régressions linéaires et logistiques. Séries chronologiques. Analyses de variance. Programmation linéaire et algorithmie. Éléments de mathématiques discrètes. Utilisation avancée des tableurs (statistiques avancées, tableaux croisés dynamiques, filtres, macros, etc.).

OBJECTIFS DU COURS

GÉNÉRAUX

- Utiliser des méthodes statistiques, mathématiques et numériques propres aux sciences des données.
- Approfondir les tableurs traditionnels.
- Tirer profit de logiciels statistiques avancés.
- Comprendre les fondements de la programmation informatique.
- Décrire les étapes clés dans le développement d'une application informatique.
- Prendre contact avec la structure et le fonctionnement des bases de données.

SPÉCIFIQUES

Au terme de ce cours, l'étudiant(e) sera en mesure de :

- Comprendre la structure générale d'un programme informatique
- Résoudre des problèmes de programmation informatique.
- Effectuer des recherches efficaces et efficaces en ligne liées à la programmation informatique.
- Solutionner des problèmes complexes à l'aide d'un tableur informatique
- Distinguer les différents types de données
- Faire un usage avancé du logiciel d'Excel et de R pour la manipulation de données
- Calculer des statistiques (simples et avancés) dans Excel et R
- Représenter graphiquement des statistiques dans Excel et R
- Comprendre les concepts de base de l'algorithmie
- Expérimenter avec le calcul informatique parallélisé
- Analyser des relations entre plusieurs variables à l'aide de différents types de régressions linéaires
- Produire des analyses de variance dans Excel et R
- Évaluer la performance des modèles de régression utilisés
- Comprendre et utiliser l'analyse factorielle
- Appliquer des méthodes prévisionnelles dans l'analyse de données
- Résoudre des problèmes d'optimisation linéaire à deux et à plusieurs variables
- Démontrer leur compréhension d'éléments de la théorie des probabilités
- Suggérer une structure pertinente pour une base de données relationnelle
- Effectuer des requêtes SQL élémentaires dans Excel et R
- Comprendre, construire et représenter des graphes
- Effectuer des analyses de réseau

CALENDRIER

Séance (semaine)	Contenu	Lectures, travaux et directives
1	Principes fondateurs <ul style="list-style-type: none"> – Introduction au cours – Présentation du plan de cours – Introduction à la programmation – Introduction au développement de logiciels 	
2	Mises à jour <ul style="list-style-type: none"> – Notions d'algèbre linéaire – Rappels statistiques – Rappels sur Excel et installations de modules complémentaires – Introduction du concept de macro – Types de données – Fichier de données plats – Autres environnements statistiques programmables 	
3	Introduction à R <ul style="list-style-type: none"> – Présentation de R – Exploration de R – Données manquantes dans R – Types d'objets dans R – Éléments de programmation propres à R – Création d'une fonction dans R 	
4	Manipulation de données 1 <ul style="list-style-type: none"> – Importation avancée de données dans Excel – Manipulation des données dans Excel – Manipulation des données avancées dans Excel 	
5	Manipulation de données 2 <ul style="list-style-type: none"> – Importation avancée de données dans R – Exportation de données dans R – Ajout de données dans R – Manipulation des données dans R – Manipulation avancée de données dans R 	
6	Statistiques descriptives <ul style="list-style-type: none"> – Statistiques descriptives dans Excel – Représentations graphiques simples dans Excel – Statistiques descriptives dans R – Représentations graphiques simples dans R – Rappels de tests statistiques – Rappels intervalles de confiance sur moyennes – Rappels intervalles de confiance sur proportions 	

7	Statistiques et algorithmie <ul style="list-style-type: none"> – Tests statistiques dans Excel – Tests statistiques dans R – Introduction à l'algorithmie – Calculs parallélisés dans R 	
8	Examen mi-session en ligne	
9	Régression 1 - Aspects théoriques <ul style="list-style-type: none"> – Régression OLS à plusieurs variables (postulats et exemple simple) – Régressions dans Excel – Régressions avancées – Analyse de variance 	
10	Régression 2 - Aspects pratiques <ul style="list-style-type: none"> – Régressions dans R – Tests diagnostiques avancés sur les résultats d'une régression dans R – Sauvegarde et utilisation des résultats de régressions dans R – Analyse de variance dans R 	
11	Modélisation avancée <ul style="list-style-type: none"> – Analyse factorielle – Analyse factorielle dans R – Lissage – Lissage dans R – Tests hors échantillon – Tests hors échantillon dans R 	
12	Optimisation à plusieurs variables <ul style="list-style-type: none"> – Rappel d'optimisation à deux variables – Optimisation à plusieurs variables – Programmation linéaire dans R – Dénombrement et analyse combinatoire 	
13	Bases de données structurées et SQL <ul style="list-style-type: none"> – Introduction aux bases de données structurées – Introduction aux requêtes SQL – Requêtes SQL dans Excel (base de données externes) – Requêtes SQL dans R (base de données internes) – Requêtes SQL dans R (base de données externes) <p>* Évaluation des enseignements disponible</p>	
14	Analyse de réseaux <ul style="list-style-type: none"> – Théorie des graphes 1 – Introduction à l'analyse de réseaux avec R – Théorie des graphes 2 	

	<ul style="list-style-type: none">- Représentation de réseaux dans R- Théorie des graphes 3- Analyse de réseaux avec R	
15	Examen final en ligne	

ÉVALUATION DE L'ÉTUDIANT

Notation

Conformément à l'article 2.13.1 du [Règlement 3 des études de 1^{er} cycle de l'UQAT](#), la notation littérale pour ce cours représente :

A+, A, A-; B+, B, B-; C+, C, C-; D+, D; E

Une cote fixe sera utilisée pour établir la notation littérale pour ce cours selon la grille suivante :

A+	90,00 %	100,00 %	C+	72,00 %	74,99 %
A	87,00 %	89,99 %	C	69,00 %	71,99 %
A-	84,00 %	86,99 %	C-	66,00 %	68,99 %
B+	81,00 %	83,99 %	D+	63,00 %	65,99 %
B	78,00 %	80,99 %	D	60,00 %	62,99 %
B-	75,00 %	77,99 %	E	0,00 %	59,99 %

Pondération d'une activité

Conformément la résolution M10-GES-38-07 du Conseil de module des sciences de la gestion, aucune activité d'évaluation ne peut représenter plus de 50 % de la note finale d'un cours.

L'enseignant peut exiger, s'il le précise dans son plan de cours, que les étudiants obtiennent une note minimale pour un ou plusieurs éléments d'évaluation pour la réussite du cours.

Participation

Conformément la résolution M10-GES-38-08 du module des sciences de la gestion, aucune pondération ne doit être attribuée à la participation des étudiants dans tous les cours du module.

Présentation des travaux écrits

Conformément à la résolution M10-GES-38-09 du module des sciences de la gestion, tout travail remis à l'enseignant doit respecter les règles de présentation de la dernière version du [Guide de présentation des travaux écrits de l'UQAT](#).

Délai de correction

Sauf dans les cours à formule particulière ou dérogeant du calendrier universitaire (intensif, de moins de trois crédits, s'échelonnant sur plus d'un trimestre, stage, projet d'application, etc.), un élément d'évaluation comptant pour au moins 20 % de la note finale doit être administré et corrigé pour que les résultats soient connus des étudiants avant la date limite d'abandon sans mention d'échec au dossier universitaire et sans remboursement prévue au calendrier universitaire.

Évaluation de la qualité du français

Conformément à la [Politique linguistique de l'UQAT](#) et la résolution M00-GES-10-06 du module des sciences de la gestion, jusqu'à 10 % des points des travaux individuels et des travaux en équipe peuvent être déduits de la note obtenue en regard de la qualité du français écrit.

Omission ou retard dans remise d'un travail

Aucun retard dans la remise d'un travail ne sera toléré sans la présentation d'une justification appropriée. Un courriel doit être envoyé directement à l'enseignant détaillant les motifs du retard appréhendé dès que l'étudiant sait qu'il ne sera pas en mesure de remettre le travail au moment prescrit.

Absence à un examen

Conformément à la [Politique d'absence et de reprise à un examen du module des sciences de la gestion](#), seul un motif sérieux permet à l'étudiant de reporter, de s'absenter et de reprendre un examen. Le motif invoqué doit relever de circonstances importantes et indépendantes de la volonté de l'étudiant.

L'étudiant qui doit reporter ou s'absenter à un examen doit compléter, signer et envoyer le formulaire de demande de reprise d'examen dans les délais prescrits avec les pièces justificatives par courriel à l'adresse gestion.distance@uqat.ca.

Plagiat

Conformément au [Règlement 12 sur le plagiat ou la fraude pour les étudiants de l'UQAT](#), tout acte (incluant la tentative et la participation) de plagiat ou de fraude relatif autant aux travaux académiques qu'aux documents officiels à caractère scolaire peut entraîner une ou plusieurs sanctions.

Échec à un cours

Conformément à l'article 2.12.3 du [Règlement 3 des études de 1^{er} cycle de l'UQAT](#), l'évaluation est continue en ce qu'elle tient compte, pendant toute la durée du cours, de tous les éléments susceptibles de révéler le niveau d'apprentissage et de formation atteint. C'est pourquoi, en cas d'échec d'un cours, il n'y a pas d'examen de reprise tenant lieu d'évaluation globale.

Travaux évalués	% de la note finale	Date limite de remise
Six quiz	15 %	Semaines 1, 2, 6, 9, 12, et 14
Devoir 1 – Manipulation de données	15 %	Le dimanche de la semaine 7 à 23h59
Devoir 2 – Régressions	20 %	Le dimanche de la semaine 12 à 23h59
Examen mi-session (en ligne)	25 %	Voir sur Moodle – Séance 8
Examen fin de session (en ligne)	25 %	Voir sur Moodle – Séance 15

Détail des travaux évalués

(Les consignes et les grilles détaillées des évaluations sont disponibles sur le portail du cours)

Devoir 1 : importation, nettoyage, fusion et exportation de bases de données en Excel et en R.

Devoir 2 : analyse d'une base de données à l'aide de régressions afin de répondre à certaines questions.

Six quiz

Six courts quiz à compléter individuellement sur Moodle suivront les semaines portant sur des notions plus théoriques.

Devoirs 1 et 2

Deux travaux pratiques à faire individuellement à l'aide des logiciels. Le premier portera sur la manipulation de bases de données dans R pour les rendre utilisables à des fins d'analyses avancées. Le deuxième travail consistera à faire ressortir la relation entre différentes variables contenues dans une base de données, et ce, toujours à l'aide de R. L'évaluation se fera à la fois sur les résultats, mais également sur le code R nécessaire pour obtenir ces résultats.

Examen de mi-session

D'une durée maximale de trois (3) heures, l'examen de mi-session se fera en ligne de façon individuelle sur Moodle. L'utilisation des logiciels Excel et R seront nécessaires pour compléter en partie ou en totalité l'examen. La matière abordée aux séances 1 à 7 est sujette à examen.

Examen final

D'une durée maximale de trois (3) heures, l'examen final se fera en ligne de façon individuelle sur Moodle. L'utilisation des logiciels Excel et R seront nécessaires pour compléter en partie ou en totalité l'examen. La matière abordée aux séances 9 à 14 est sujette à examen. Cela dit, bien que l'examen ne soit pas récapitulatif, certaines notions vues au courant des séances 1 à 7 pourraient être nécessaires pour répondre adéquatement à toutes les questions.

MÉTHODOLOGIE DE L'ENSEIGNEMENT

Rôles du professeur :

Le professeur collabore à la réussite de l'étudiant par le biais d'activités de réflexion qui contribuent à accroître les différents savoirs nécessaires à cette réussite. En d'autres mots, il transmet les informations pertinentes, oriente et guide l'étudiant dans sa démarche et répond aux demandes individuelles d'assistance en regard de l'atteinte des objectifs du cours et en évalue le degré d'atteinte. De plus, il indique les principes de fonctionnement du cours (par exemple les règles quant à la participation, aux arrivées tardives, aux consignes de sécurité s'il y a lieu, etc.).

Délai de réponse :

La résolution *DEP-2019-SG-287-14* du département des sciences de la gestion stipule que :

« Le professeur ou le chargé de cours répondra à l'étudiant dans un délai de deux (2) jours ouvrables. Pour les questions nécessitant des réponses plus élaborées, après avis transmis à l'intérieur de ce délai, une réponse complète sera fournie à l'intérieur de cinq jours ouvrables. Passé ces délais, et après avoir tenté une relance, l'étudiant peut aviser le directeur du département, pour les cours de 1^{er} cycle, ou le directeur du programme, pour les cours de 2^e cycle, afin qu'il communique par téléphone avec le professeur ou le chargé de cours concerné. Dans des circonstances particulières, le professeur ou chargé de cours qui ne serait pas en mesure de respecter momentanément ces délais doit en aviser ses étudiants. »

Rôles de l'étudiant :

De son côté, l'étudiant assume ses responsabilités quant à la prise en charge du développement de sa formation. Il s'en approprie les objectifs, se fait et réalise un plan de travail pour les atteindre et satisfait aux exigences (lectures, travaux, coopération) requises dans ce cours. Il doit évaluer l'enseignement et le cours avec objectivité.

De plus, il prend l'initiative de rencontrer individuellement le professeur au besoin. Bien sûr, ces rencontres professeur-étudiant ne peuvent avoir pour objectif de compenser l'absence au cours. L'étudiant ne peut utiliser son cellulaire ou autres équipements électroniques à l'intérieur d'un cours, à l'exception du matériel prévu dans un cadre pédagogique.

Aucun enregistrement audio ou vidéo n'est permis sans l'autorisation écrite de l'enseignant. Le non-respect de cette directive pourra entraîner des sanctions telles que prévues par les règlements et politiques de l'UQAT en vigueur.

SOURCES DOCUMENTAIRES

Matériel obligatoire :

Cornillon P.-A., Husson F., Guyader, A., et al. (2018). *R pour la statistique et la science des données*. Rennes, France, Presses universitaires de Rennes.

Veillez-vous référer au site Web <https://www.uqat.ca/etudes/distance/gestion/cours/> pour l'achat du matériel obligatoire.

Matériel complémentaire :

Cornillon P.-A. et Matzner-Løber, É. (2011). *Régression avec R, collection Pratique R*. Paris, France, Springer.

Isoz, V., & Kolani, D. (2018). *Langage R : Prise en main des statistiques*, Paris, France, Editions ENI.

Boucher, P. (2017). *Microsoft Excel 2013 et 2016 et le traitement statistique des données*. Montréal, Québec, Pearson.

Paroissin, C. (2015). *Programmation et analyse statistique avec R*. Paris, France, Ellipses.

Rigollet, P. (2016). *Tableaux de synthèse et tableaux de bord : traitez et analysez de gros volumes de données avec Excel 2016*. Paris, France, Éditions ENI.