

EXPLORATION DE DONNÉES ET INTRODUCTION À L'INFÉRENCE STATISTIQUE

LMAFY1101

Anouar El Gouch

LSBA, UCLouvain

PLAN

À PROPOS DU COURS

INFO PRATIQUE

À PROPOS DU COURS

MÉTHODE

- Ce cours possède l'avantage d'offrir simultanément une **formation progressive à la statistique et au langage (logiciel) R/Rstudio** en réduisant au strict nécessaire les notions théoriques abordées.
- Le but est de **comprendre** les notions de base de la statistique et d'**appliquer ces connaissances sur des données réelles**.
- **R est seulement un outil de calculs et non pas un objectif en soi.**
- Exposés de synthèse: les concepts et outils vous seront présentés sur base d'**exemples issus de différents domaines**.
- Les **transparents** utilisés pour la présentation du cours en constituent le support.
- Séances d'exercices en **salle informatique** en utilisant R.
- Les séances d'exercices mettront systématiquement en pratique les différentes notions structurées durant le cours sur des cas bien ciblés.

ÉVALUATION: PENDANT LES TPs

- 3 **devoirs** sous forme de mini-tests (30 – 45 min) à résoudre à l'aide de R/RStudio. Ces devoirs sont à **livre ouvert** et sont à réaliser **individuellement ou en binôme**. Il sera interdit de communiquer entre vous, sauf avec votre binôme (si le travail est réalisé en groupe).
- Chaque devoir compte pour 1 point. Au total, vous pouvez obtenir jusqu'à 3 points bonus. Aucun devoir n'est obligatoire. Vous êtes libres de participer ou pas à un ou plusieurs devoirs.
- Les bonus sont accordés individuellement: pas de devoir rendu \Rightarrow 0 bonus pour le devoir en question pour l'étudiant concerné. Le total des bonus accumulés sera ajouté à la note (non arrondie) de l'examen final selon la règle suivante:

Note de l'examen final	Bonus
< 8	33%
$[8, 9[$	66%
≥ 9	100%

ÉVALUATION: PENDANT LA SESSION DES EXAMENS

- Examen à livre ouvert **sur 20** à réaliser en **salle informatique** sur ordinateur à l'aide du logiciel R/RStudio.
- On vous fournira tous les notes du cours (les slides et le e-book R) ainsi que toute documentation nécessaire (en version PDF et/ou en HTML). Les énoncés et solutions des TP's ne seront pas accessibles. Aussi, il n'aura pas d'accès à internet pendant toute la durée de l'examen.
- L'examen sera composé par des questions ouvertes et/ou QCM. Il comportera une **partie théorique** (par exemple, calculer une probabilité à la main) et **une partie sur ordinateur**. Cette dernière consiste essentiellement à écrire et exécuter du code R et à transcrire les sorties et leurs interprétations sur les feuilles de réponses.
- Vous devez être capable d'utiliser R/Rstudio pour effectuer des analyses statistiques comme celles vues au cours et aux TP's. En principe, **vous ne devez apprendre aucune formule ou commande R par cœur**. L'accent sera mis sur la **compréhension des concepts** et non pas sur la reproduction de la matière vue au cours.

INFO PRATIQUE

- Livre de référence en français (non obligatoire): “Initiation à la statistique avec R” (2014), Dunod, 2e édition. Livre de Frédéric Bertrand et Myriam Maumy-Bertrand.
- Livre de référence en anglais (non obligatoire) : “Probability and statistics with R” (2015), Chapman and Hall/CRC, 2e édition. Livre de Maria Dolores Ugarte, Ana F. Militino et Alan T. Arnholt.
- **Inscrivez-vous sur le site du cours sur Moodle.**
 - Transparents présentés lors des exposés.
 - Énoncés et solutions des TP's.
 - Documents supplémentaires (jeux de données, code R, liens utiles ...).
 - Plus d'info sur le cours.

- Pendant le cours : posez des questions !
- Avant les TP's : lisez les notes de cours et préparez les exercices.
- Toute remarque est la bienvenue.