



## MASTER 2 SCIENCES ECONOMIQUES ET SOCIALES

### PARCOURS-TYPE INGENIERIE DE LA TRANSITION DES TERRITOIRES

- **Titre du cours** : Traitement statistique des données
- **Programme** : M2 SES mention ITT
- **Crédits ECTS** : 2
- **Langue d'enseignement** : français
- **Modalités pédagogiques** : Présentiel
- **Total heures** : 20h
- **Heures CM** : aucune
- **Heures TD** : 20h
- **Heures de travail personnel par étudiant** : 20h
- **Responsable de la formation** : Pascal ROGGERO
- **Responsable du cours** : Céline VIGNES
- **Chargé de TD** : Céline VIGNES

## DESCRIPTION

### ● **Présentation** :

Les différentes méthodes d'analyse statistique offrent un cadre de traitement des données quantitatives. L'objectif de ce cours est d'introduire certaines de ces méthodes. L'accent sera mis sur la pratique et l'interprétation des résultats.

Les méthodes statistiques étudiées sont : la statistique descriptive uni- et bi-variée, les tests d'hypothèse (essentiellement les tests de liaison : test du coefficient de corrélation linéaire à 0, test du Khi-deux d'indépendance, test de comparaison de plusieurs moyennes) et la régression linéaire dont le but est d'expliquer les variations d'une variable quantitative continue (à beaucoup de valeurs différentes) à l'aide d'une ou plusieurs variables explicatives qui peuvent être quantitatives ou qualitatives.

Toutes les méthodes étudiées sont appliquées sur un exemple réel traité avec le logiciel R et l'interface Rstudio. R est un logiciel libre, gratuit et multiplateforme très répandu que les étudiants pourront installer sur leur ordinateur personnel.

### ● **Objectifs d'apprentissage** :

A l'issue de ce cours, l'étudiant est capable de :

- identifier le type des variables.
- produire des graphiques adaptés pour la visualisation des variables selon leur type.
- interpréter en termes concrets les indicateurs numériques et les valeurs fournies dans les tableaux de distribution.
- établir précisément les hypothèses de liaison à tester selon le type des variables, mettre en œuvre avec RStudio le test statistique approprié à la liaison étudiée, conclure sur le test en utilisant la p-valeur et faire une conclusion plus générale en revenant aux données.
- écrire l'équation d'un modèle de régression linéaire et mettre en œuvre cette régression à l'aide de Rstudio.
- interpréter en termes concrets les effets des variables explicatives quantitatives ou qualitatives sur la variable à expliquer après avoir réalisé les tests de significativité.

### ● **Compétences transverses** :

Ce cours prépare à certaines compétences transverses « Usages avancés et spécialisés des outils numériques (RNCP34846BC01) » :

- identifier les usages numériques et les impacts de leur évolution sur le ou les domaines concernés par la mention
- se servir de façon autonome des outils numériques avancés pour un ou plusieurs métiers ou secteurs de recherche du domaine

- **International**  
Il n'y a pas de dimension internationale dans le cours.
- **Responsabilité sociale et environnementale**  
Il n'y a pas de dimension RSE dans le cours.

- **Prérequis :**

Savoir utiliser un logiciel de traitement de texte.

Notions de base de statistique descriptive (moyenne, écart-type) et notions de probabilités (loi normale).

- **Contenu, organisation :**

Durant les deux premières séances, d'une durée d'1h30, sont présentées les méthodes d'analyses statistiques descriptives univariées et bivariées avec les tests statistiques de liaison, sous la forme de cours magistraux.

Les autres séances se déroulent en salle informatique afin de réaliser les analyses statistiques grâce au logiciel R avec l'interface R Studio. Une attention soutenue est accordée aux interprétations en termes concrets. Ces séances de TD sont organisées en 4 parties :

1. Prise en main du logiciel R avec l'interface Rstudio et opérations de base sur un jeu de données (1 séance de 2h)
2. Analyses univariées (1 séance de 3h)
3. Analyses bivariées, avec tests statistiques de liaison (2 séances de 3h)
4. Régression linéaire (2 séances de 3h)

L'ensemble des documents nécessaires au cours sont disponibles sur Moodle : support de cours, sujets de TP, programme corrigé et commenté en détails de chaque TP, annales.

- **Modalités d'évaluation :**

Contrôle terminal :

Le contrôle a lieu sur ordinateur, en binôme. Les étudiants devront répondre à des questions sur un jeu de données en mettant en œuvre sous RStudio les méthodes étudiées. Les documents sont autorisés.

- **Bibliographie :**

- Lafaye de Micheaux P., Drouilhet R., Liquet B., *Le logiciel R - Maîtriser le langage - Effectuer des analyses statistiques*. Springer, 2010.
- Cornillon P.-A., Matzner-Løber E., *Régression avec R*. Springer, 2011.
- Guay J.-H., *Statistiques en sciences sociales avec R*. Les Presses de l'Université de Laval, 2013.
- Millot G., *Comprendre et réaliser les tests statistiques à l'aide de R*. De Boeck, 2014.

*Clause de non-responsabilité*

*Les informations fournies dans ce syllabus sont provisoires et peuvent être modifiées. Les modifications du plan de cours seront annoncées en classe.*