

Programme et cadre de la formation 2021-2022 « Parcours quantitatif »

Julien Barnier (Centre Max Weber)

Sofiane Bouzid (Archéorient)

Nathalie De Jong (Cercrid / Triangle)

Karine Pietropaoli (Centre Max Weber)

Plateforme technologique PANELS (MSH-LSE)

CERC**CRID**
UMR5137
CENTRE DE RECHERCHES CRITIQUES SUR LE DROIT


Archéorient
Environnements et sociétés de l'Orient ancien


RIS
Réseau Interdisciplinaire
autour de la Statistique


Centre
Max Weber

msh 
Lyon St-Étienne

triangle
UMR 5206

Formations 2021-2022

« Parcours quantitatif »

Contexte et objectifs

Le parcours quantitatif est une initiative de formation inter-laboratoires d'Archéorient, du Centre Max Weber, du Cercriid et de Triangle, soutenue par la MSH Lyon St-Etienne qui propose un programme de formation en traitement et analyse quantitative de données. L'objectif est d'offrir un parcours progressif qui permettra d'apprendre à organiser, traiter, analyser et présenter l'information.

Ce « parcours quantitatif » repose sur les principes suivants :

- proposer une série de formations et une dynamique d'apprentissage progressives, adaptées aux débutants
- utilisation tout au long de la formation du logiciel R (libre et gratuit) afin de se familiariser avec l'outil
- mise en application systématique afin d'acquérir une première expérience et une connaissance pratique des concepts et des méthodes abordés

Public visé

Tous les personnels des laboratoires SHS du site de Lyon et de Saint-Étienne (chercheurs, enseignants-chercheurs, ingénieurs, techniciens, doctorants) débutant en analyse quantitative et/ou souhaitant se remettre à niveau.

Organisation

Le parcours est composé de trois sessions :

- *Session 1 : Introduction au quanti.* Trois séances de formation : une séance dédiée à l'enquête par questionnaire et une séance sur les aspects réglementaires en matière de traitement de données à caractère personnel, une séance d'introduction à l'analyse quantitative.
- *Session 2 : Outils de base.* Deux séances de formations à des outils de base pour la production, la manipulation et la visualisation de données et d'analyses, une séance sur la représentation de cartes statistique.
- *Session 3 : Méthodes.* Trois séances de formations d'approfondissement sur des méthodes d'analyse statistiques présentées et mises en pratique sous R.

Les personnes totalement débutantes ou n'ayant jamais mis de méthodes quantitatives en application devront suivre la séance d'introduction à l'analyse quantitative, qui fournira des rappels et bases indispensables pour la suite. Les deux autres séances de la session 1, ainsi que les séances des sessions 2 et 3 sont indépendantes et peuvent être suivies sous condition de respecter

les prérequis. La participation aux séances de la session 3 nécessitent d'avoir suivi l'*Introduction à R et au tidyverse* de la session 2, ou d'avoir une bonne maîtrise de R (là aussi, se référer aux prérequis).

Les inscriptions à chaque session seront lancées après la fin de la session précédente, laissant la possibilité à chacun.e de juger s'il/elle souhaite approfondir la formation.

Ces formations se dérouleront de novembre 2021 à mars 2022 dans les locaux de la MSH Lyon St-Etienne à Lyon. Le nombre de participants sera limité à 15 personnes par journée de formation. Les séances n'auront lieu que si le nombre de participants est suffisant (5 personnes minimum).

Les participants devront venir avec leur ordinateur portable.

Programme

Programme complet de chaque session en ligne :

<https://www.msh-lse.fr/parcours-quantitatif-2021-2022/>

Session 1 séance 1 - Méthodologie de l'enquête par questionnaire / Initiation à l'outil Limesurvey

Nathalie De Jong, 2 jours (12h), 22 novembre et 23 novembre 2021

L'objectif de cette séance est d'initier les participants à la mise en place d'une enquête par questionnaire.

Une première demi-journée théorique abordera les différentes étapes de mise en oeuvre d'une enquête par questionnaire d'un point de vue méthodologique (de la définition de la question de départ, au rapport d'analyse).

Puis, deux demi-journées d'initiation à l'élaboration et à la passation d'un questionnaire en ligne, à l'aide de l'outil Limesurvey (application web open-source) seront proposées aux participants. Il s'agira, à partir d'un questionnaire commun, de découvrir les caractéristiques, les possibilités techniques, mais également les limites de cet outil.

Enfin, une dernière demi-journée sera proposée sous forme d'atelier pour les participants ayant déjà un projet de construction de questionnaire.

Note : les personnes maîtrisant la méthodologie de l'enquête par questionnaire qui souhaitent seulement suivre *l'initiation à l'outil Limesurvey* peuvent ne pas participer à la première ½ journée.

Session 1 séance 2 - Traitement de données à caractère personnel : aspects réglementaires

Nathalie De Jong, 1 journée (6h), 29 novembre 2021

L'objectif de cette séance est de sensibiliser les participants à la réglementation en vigueur en matière de traitement de données à caractère personnel, souvent au coeur de l'activité des sciences

sociales. Elle abordera les notions fondamentales et les grands principes du RGPD, ainsi que la mise en conformité d'un traitement (aspects pratiques).

Session 1 séance 3 - Introduction à l'analyse quantitative

Sofiane Bouzid, Julien Barnier et Karine Pietropaoli, 2 jours, 13 et 14 décembre 2021

L'objectif de ces séances est de donner à la fois un aperçu de l'analyse quantitative de données, son intérêt, ses outils, ses concepts mais aussi ses limites et ses difficultés. Après une présentation générale sur ce que signifie concrètement "faire du quanti", les concepts de base en statistique seront présentés via des manipulations sur tableur, puis aura lieu une présentation d'un outil de traitement et d'analyse de données, le langage R.

- Introduction générale à l'analyse quantitative de données (½ journée)
- Introduction aux concepts de base de la statistique sur tableur (1 journée)
- Découverte du logiciel R (½ journée)

Session 2 séance 1 - Introduction à R et au *tidyverse*

Julien Barnier, 4 jours (6h par jour), 17, 18, 24 et 25 janvier 2022

Cette séance est une introduction au langage R (logiciel libre et gratuit de traitement et analyse de données) et aux extensions dites du *tidyverse*, qui facilitent les opérations les plus courantes, notamment la manipulation et le recodage des données, la production de graphiques et l'exportation des résultats.

Note : les personnes connaissant déjà R et souhaitant seulement une introduction au *tidyverse* peuvent ne pas participer à la première journée.

- *Premier jour* : introduction de base au langage R et à l'interface RStudio, sans prérequis. Éléments de base du langage, analyses univariées et bivariées.
- *Deuxième jour* : production de graphiques avec ggplot2, manipulation des données avec dplyr
- *Troisième jour* : recodage des données avec *questionr* et *forcats*, export des résultats avec R Markdown.
- *Quatrième jour* : mise en application de ce qui a été vu précédemment sur un jeu de données d'exemple.

Session 2 séance 2 - Réaliser des cartes avec R

Karine Pietropaoli, 1 journée (6h), 01 février 2022

Cette séance est dédiée à la représentation de cartes statistiques avec R. La formation abordera les étapes préparatoires à la mise en carte, notamment l'acquisition et le traitement des données cartographiques, puis la réalisation de carte statiques en figuré proportionnel et cartes choroplèthes et enfin la réalisation de cartes dynamiques avec Leaflet.

Note : Il est recommandé d'avoir quelques notions de géomatique et d'avoir suivi les séances *introduction à R et au tidyverse* (ou bien de connaître dplyr) pour suivre cette formation.

Session 3 séance 1 - Introduction aux tests statistiques

Sofiane Bouzid et Julien Barnier, 1 jour et demi (6h par jour), 28 février et 1er mars 2022

L'objectif de cette séance est d'initier les participants à deux grandes notions statistiques utilisées lorsque l'on travaille sur des échantillons : l'inférence (essayer de déterminer une valeur ou une caractéristique d'une population à partir d'un échantillon) et le test (essayer d'éprouver une hypothèse sur une population à partir d'un échantillon).

- Présentation du principe des tests statistiques, leur utilisation, leurs limites et leur interprétation en se basant sur le test du χ^2 d'indépendance sur un tableau croisé
- Mise en application sous R de quelques tests parmi les plus utilisés en SHS

Session 3 séance 2 - Analyse de données multidimensionnelles

Sofiane Bouzid et Karine Pietropaoli, 3 journées (6h par jour), 21, 22 et 24 mars 2022

La formation a pour but de comprendre et appliquer les méthodes de l'analyse des données pour rendre les participants autonomes dans la mise en œuvre et l'interprétation d'analyses multidimensionnelles.

- Analyse en composantes principales.
- Analyse factorielle des correspondances.
- Analyse des correspondances multiples
- Classification de données.

Session 3 séance 3 - Introduction à l'analyse de données textuelles

Sofiane Bouzid, 1 journée (6h), 28 mars 2022

Mettre en œuvre des méthodes de la statistique, de l'analyse de données sur des corpus textuels (questions ouvertes, entretiens, articles, Web, etc.) pour présenter et extraire les connaissances du texte.

La formation abordera les différentes étapes de traitement, d'analyse et de visualisation d'un corpus de texte. Une application sur des données issues de twitter est prévue en fin de séance

Informations pratiques

Lieu : MSH Lyon St-Etienne, 14 avenue Berthelot 69363 Lyon cedex 07 (salle Elise Rivet)

Inscription : L'inscription aux formations se fait via un formulaire en ligne, **à raison d'un formulaire par session**. Chaque formulaire sera mis en ligne à la fin de la session précédente.

Contenus des sessions :

Julien Barnier : julien.barnier@cnrs.fr - Sofiane Bouzid : sofiane.bouzid@mom.fr

Nathalie De Jong : nathalie.dejong@univ-st-etienne.fr - Karine Pietropaoli : k.pietropaoli@univ-lyon2.fr