

Semaine [Data-SHS]

Enquêtes, données, et méthodes quantitatives
pour les Sciences Humaines et Sociales

Du 05 au 09 décembre 2022 | Strasbourg

 En présentiel | Sur inscription

C'est la quatrième année que la [Plateforme Universitaire de Données de Strasbourg](#) (PUD-S) organise 5 jours de séminaires et d'ateliers consacrés à l'utilisation de données quantitatives en sciences humaines et sociales (SHS).

Cette semaine est organisée au niveau national par la très grande infrastructure de recherche [Progedo](#) et le [Réseau National des Maisons des Sciences de l'Homme](#).

Chaque journée sera composée de deux temps :

- la matinée sera l'occasion d'assister à la présentation de sources de données quantitatives ou de travaux de recherche les mobilisant ;
- l'après-midi verra se dérouler un atelier pratique en rapport.

Ce sont au total 17 intervenant·e·s aux statuts, rattachements et disciplines différents qui se succéderont au cours de cet événement.

Objectifs

- **Acquérir** des notions théoriques sur les bases de données et enquêtes disponibles afin de les mobiliser dans le cadre d'un travail de recherche.
- **Initier** les participant·e·s au traitement, à la structuration et à l'analyse de données d'enquête ainsi qu'à l'automatisation de la collecte grâce aux ateliers.

Publics

Les communications sont ouvertes à toutes et tous quand les ateliers s'adressent aux doctorant·e·s, enseignant·e·s-chercheur·e·s, ainsi qu'aux chercheur·e·s et ingénieur·e·s.

Informations et inscriptions

La participation est libre, mais nécessite inscription.

Inscription en ligne sur [LimeSurvey](#)

Date limite d'inscription : la veille de chaque événement

 Le nombre de places disponibles par ateliers étant limité, vous serez notifié·e si le nombre maximal d'inscrit·e·s est atteint.

PROGRAMME (les ateliers sont matérialisés par une puce de couleur rouge • et entre crochets [l'outil utilisé])

Lundi 5 décembre 2022

L'INED : données, outils et projets en cours

- 9 h 30 - 12 h : Présentations du projet *BigStat* et du Datalab de l'Ined

 Salle de conférence | MISHA

[Laurent Toulemon](#), Directeur de recherche, INED

[Ariana Caporali](#), Ingénieure de recherche, INED

- 14 h - 17 h : « [R] Quelles données et outils pour cartographier les naissances multiples ? »

 Salle de formation | MISHA

[Catalina Torres](#), chercheuse postdoctorale en démographie, INED
[Plateforme Universitaire de Données Strasbourg](#), MISHA - UAR 3227

Mardi 6 décembre 2022

Données de santé et vieillissement en Europe

- 9 h 15 - 12 h 45 : Présentations des enquêtes SHARE et MFV-2 et travaux associés

 Salle de conférence et en visioconférence | MISHA

[Thomas Renaud](#), Ingénieur de recherche, LEDa - UMR 8007-260

[Mélanie Lepori](#), Maîtresse de conférences en démographie, SAGE - UMR 7363

[Guillaume Noirel](#), Doctorant en démographie, SAGE - UMR 7363

[Didier Breton](#), Professeur des Universités, SAGE - UMR 7363

[Maude Crouzet](#), Chercheuse postdoctorale en sociologie quantitative, IRIS - UMR 8156

- 14 h - 17 h : « [R] Première prise en main d'une enquête avec panel : le jeu de données pédagogique EASY-SHARE »

 Salle de formation | MISHA

[Plateforme Universitaire de Données Strasbourg](#), MISHA - UAR 3227

Mercredi 7 décembre 2022

Utilisation de données confidentielles et reproductibilité

- 9 h 00 - 10 h 30 : Le Centre d'Accès Sécurisé aux Données : présentation, fonctionnement et retours d'expérience
- 10 h 30 - 12 h 00 : CASCaD : un service pour la reproductibilité des travaux de recherche

 Salle de conférence et en visioconférence | MISHA

[Rémy Marquier](#), Délégué à la protection des données, CASD

[Christophe Perignon](#), Professeur de finance, CASCaD - UAR 2007

[Anaïs Hamelin](#), Professeure des Universités, LaRGE - EA 2364, Université de Strasbourg

[Wilda Jean-Baptiste](#), Ingénieur d'études en traitement et analyse de données, LISST - UMR 5193, Université de Toulouse

- 14 h 00 - 17 h 00 : « [R] Principes RGPD et traitement de données personnelles d'enquêtes »

 Salle de formation en visioconférence | MISHA

[Alice Gralhien](#), Déléguée à la protection des données, Université de Strasbourg

[Quentin Gallis](#), Ingénieur d'études en traitement et analyse de données, CDSP



Jeudi 8 décembre 2022

Construction et modes de collecte des enquêtes en sociologie

- 9 h 00 – 10 h 30 : Modes de collecte d'une enquête longitudinale : l'exemple de l'enquête ELFE (Étude Longitudinale Française depuis l'Enfance)
- 10 h 30 – 12 h 00 : Dans les coulisses de la collecte, projection et discussion autour du documentaire « En quête de chiffres » (2019)

Salle de la table ronde | MISHA

[Xavier Thierry](#), Chargé de recherche, INED

[Elsa Martin](#), Maîtresse de conférences en sociologie, 2L2S – EA 3478, Université de Lorraine

[Simon Paye](#), Maître de conférences en sociologie, 2L2S – EA 3478, Université de Lorraine

- 14 h – 17 h : « [R] *Webscraping* et réseaux sociaux : de nouvelles voies pour la recherche en sociologie ? »

Salle de formation | MISHA

[Plateforme Universitaire de Données Strasbourg](#), MISHA – UAR 3227



Vendredi 9 décembre 2022

Le plan de gestion de données : fondements théoriques et pratiques

- 9 h 30 – 12 h 00 : Comment se conçoit, se structure et se rédige un plan de gestion des données ?(PGD)

Salle de la table ronde et de formation | MISHA

[Stéphanie Chevion](#), *Data librarian*, BU, Université de Strasbourg

[Gaëlle Gauvrit](#), *Data librarian*, BU, Université de Strasbourg



Vendredi 9 décembre 2022

Quiz interactif : données et méthodes quantitatives en SHS

- 14 h 00 – 16 h 00 : « Culture-PUD – 2^e édition » (sous réserve d'un nombre d'inscrit·e·s suffisant)

En visioconférence

[Plateforme Universitaire de Données Strasbourg](#), MISHA – UAR 3227



Maison Interuniversitaire des
Sciences de l'Homme
- Alsace -



RÉSUMÉS DES COMMUNICATIONS ET ATELIERS

« L'INED : données, outils et projets en cours »

Établissement de recherche français connu et reconnu, l'Institut National des Études Démographiques (INED) est également l'un des principaux producteurs et diffuseurs de données d'enquêtes en Sciences Humaines et Sociales (SHS). Cette matinée sera l'occasion de présenter ces deux versants au cours tout d'abord d'une présentation du projet BigStat, dédié à l'exploitation des principales sources de données en démographie, puis du Datalab de l'Ined ; plateforme des données de l'Institut, rassemblant catalogue d'enquêtes et bases de données contextuelles.

« Quels données et outils pour cartographier les naissances multiples ? »

L'atelier débutera par une présentation de la base de données contextuelle sur les naissances multiples : *Human Multiple Births Database* (HMBD). Cette dernière, constituée à partir d'une multitude de sources, fournit les nombres et les taux annuels de naissances gémellaires pour différents pays d'Europe et du monde depuis 1815. L'objectif de l'atelier sera de concevoir à partir de ce matériel des cartes pour analyser l'évolution de ce phénomène à la fois entre les territoires, mais également entre les époques.

« Données de santé et vieillissement en Europe »

Deux enquêtes partageant la thématique commune du vieillissement seront ici présentées. La première à dimension longitudinale et européenne : l'enquête SHARE. La seconde, transversale et se concentrant sur les outre-mers français : MFV2. Cadres institutionnels, financements, enjeux et résultats seront évoqués. Des chercheurs locaux mobilisant ou ayant mobilisé ces matériels d'enquêtes viendront compléter ces interventions.

« Première prise en main d'une enquête avec panel : le jeu de données pédagogique EASY-SHARE »

L'enquête SHARE existe sous la forme d'un jeu de données pédagogiques permettant son utilisation en dehors du cadre de la recherche : EASY-SHARE. Cette base de données simplifiée, qui permet une prise en main plus rapide de l'enquête ; encourage et facilite à la fois la réutilisation des données collectées et l'enseignement du « quanti » en Sciences Humaines et Sociales.

« Utilisation de données confidentielles et reproductibilité »

Le Centre d'Accès Sécurisé aux Données (CASD) permet d'accéder à distance à une infrastructure sécurisée où sont sanctuarisées des microdonnées confidentielles (de santé, fiscales, des entreprises, etc.). Outre une présentation générale de l'entité, il sera question d'évoquer le processus de prise en charge de sortie des données. Un retour d'expérience de chercheurs mobilisant ce type de données complètera la présentation. Cette matinée se conclura par la présentation de l'unité d'appui à la recherche CASCad, en charge de délivrer des certificats de reproductibilité pour des travaux utilisant notamment les données mises à disposition par le CASD.

« Principes RGPD et traitement de données personnelles d'enquêtes »

Cet atelier présentera tout d'abord les principes de bases à observer au regard du règlement général sur la protection des données (RGPD) lors de la conception, la passation et le traitement d'un matériel d'enquête. Il s'agira ensuite d'explicitier les processus d'anonymisation et de pseudonymisation mis en œuvre par les principaux diffuseurs de données d'enquêtes. Une partie pratique permettra aux participant·e·s d'expérimenter ces méthodes à partir d'un jeu de données fourni.

« Construction et modes de collecte des enquêtes en sociologie »

L'enquête ELFE (Étude Longitudinale Française depuis l'Enfance) est le premier dispositif d'envergure nationale consacré au suivi des enfants et aux multiples aspects de leur vie par le prisme des sciences sociales, de la santé et de l'environnement. Pouvoir enquêter de façon continue auprès d'enfants et de parents, dans l'intimité de la famille, nécessite un dispositif de collecte adapté qui sera ici présenté. En regard de cette expérience, la projection du documentaire « En quête de chiffres » mettra en lumière le travail des enquêtrices de l'INSEE et les coulisses de la passation.

« Webscraping et réseaux sociaux : de nouvelles voies pour la recherche en sociologie ? »

À l'heure où les données sont notre quotidien et que nos quotidiens deviennent des données, notamment sur les différentes plateformes sociales, nous verrons comment se les approprier dans le cadre d'un travail de recherche. Posts *Twitter*, commentaires *YouTube*, et publications *Instagram* feront partie des exemples présentés. Le cadre légal, la fiabilité et la représentativité des données récupérées seront également interrogés.

« Le plan de gestion de données : fondements théoriques et pratiques »

Un plan de gestion de données (PGD) ou *data management plan* (DMP) est un outil de gestion des données. Il précise la manière dont les données seront produites, traitées, décrites, partagées, protégées et conservées au cours et à l'issue d'un projet de recherche. Pièce désormais obligatoire d'une demande de financement à l'agence nationale de la recherche (ANR), cette communication viendra rappeler les fondements théoriques du plan de gestion des données : qui, comment, et pourquoi concevoir un tel document ? Une partie pratique viendra compléter la présentation.

« Culture-PUD – 2^e édition »

Fort de son précédent succès, le quiz interactif autour des données et méthodes quantitatives en SHS revient pour une deuxième édition. À partir d'une plateforme en ligne dédiée et un ensemble de questions revu et révisé, les participant·e·s pourront mesurer leurs connaissances théoriques et pratiques du « quanti ». Qui saura tirer son épingle du jeu ?