



Données : Enjeux, Collecte et Analyse – 1A / Semestre 2

Conférence de méthodes DECA2 – Année 2023-2024

Horaires : mercredi de 09h45 à 11h15 (01h30)

Salle : salle A.136

Enseignant : Mignot Léo

Bureau : B.133

Disponibilité : sur rendez-vous

Courriel : l.mignot@sciencespobordeaux.fr

Merci d'indiquer DECA2 dans l'objet de tout courriel adressé à l'enseignant.

PRESENTATION

DECA est un enseignement pluri-disciplinaire autour des sources et données en sciences sociales qui se développe de façon progressive chaque semestre des 2 premières années de Sciences Po Bordeaux (4 semestres en tout).

Étant entendu que les fonctions de responsabilité amènent à prendre des décisions en se fondant sur des informations diverses et actant le fait que les données sont tout à la fois de plus en plus multiformes et de plus en plus utilisées dans le débat public, il s'agira de montrer leurs rôles et les controverses qui les accompagnent, en termes de qualité, fiabilité, compréhension, etc.

Au cours du 1er semestre (**DECA1**), un cours magistral a fait intervenir différent.es enseignant.es de sciences sociales de l'établissement (Droit, Économie, Histoire, Sciences Politique et sociologie). Il a pour objectif de vous familiariser à la notion de données, en abordant leurs diversités (au cours du temps, en termes de forme, de qualité, de modalités de collecte, de visualisation). Ce faisant, vous avez été sensibilisé.es aux enjeux autour des données et au caractère fondamental d'en maîtriser les contours.

Économie générale de DECA

PREMIERE ANNEE		DEUXIEME ANNEE	
Semestre 1	Semestre 2	Semestre 3	Semestre 4
DECA 1	DECA 2	DECA 3	DECA 4
-----	-----	-----	-----
	Administration de la preuve (données quantitatives)	Administration de la preuve (données qualitatives)	
	<i>Atelier</i>	<i>Atelier</i>	

PREMIERE ANNEE	DEUXIEME ANNEE
<p>La notion de données dans tous ses états</p> <p><i>Cours magistral</i></p>	<p>Approfondissement de l'usage d'une méthode de collecte et d'analyse de données. A la carte, en fonction de propositions de conférences quantitatives ou qualitatives.</p> <p><i>Atelier d'approfondissement</i></p>

Objectifs de ce semestre DECA2

DECA2 s'intéresse à l'administration de la preuve à partir de données quantitatives (statistiques), **en particulier dans le cadre de l'enquête par questionnaire**. Le semestre aborde notamment les questions de la représentativité des échantillons de données, de l'analyse descriptive, des relations entre deux variables et de la significativité de ces relations, et offrira (s'il reste du temps) une introduction à l'analyse multivariée.

L'objectif consiste notamment à vous doter des compétences requises en matière de lecture de tableaux et de visualisation de données.

Les exercices de validation sont variés. On privilégie la pratique et vous serez initié.es à l'usage d'un logiciel libre de traitement de donnée : le **logiciel JAMOVI**.

SOLIDARITE ! Cette conférence entend valoriser les vertus du travail en groupe et l'importance de la réflexion cumulative et collective. Il s'agit de travailler et réfléchir ensemble, dans la bienveillance, la tolérance et la solidarité. Nous avons conscience que vous n'avez pas toutes les mêmes facilités et/ou dispositions quant aux compétences enseignées dans cette conférence. Par conséquent, le travail au sein de chaque groupe, mais également le soutien aux autres sont attendus.

Pré-requis

1/ Vous traiterez un jeu de données à partir de VOTRE **ordinateur portable** personnel. Cela implique de disposer d'un ordinateur qui ne se limite pas à la consultation de mails, photos et vidéos ou au partage de fichiers (en d'autres termes, cette conférence requiert un matériel informatique qui dispose par exemple d'un pointeur – souris / trackpad – et rend possible le traitement de texte et, par extension, le traitement de données). Une tablette n'est pas adaptée.

En cas d'impossibilité matérielle, les personnes concernées devront constituer leur groupe de travail avec des membres de la conférence équipé.es d'un ordinateur adéquat. Qu'elles n'hésitent par ailleurs pas à contacter Viviane Le Hay par mail (v.le.hay@sciencespobordeaux.fr).

Rappel : les étudiant.es à SCPOBX peuvent installer gratuitement la suite microsoft office sur leur ordinateur.

Aide numérique de l'établissement

Afin de lutter contre la précarité numérique, Sciences Po Bordeaux met en place un dispositif d'aide sous certaines conditions (présentation d'une facture d'achat ; avis d'imposition N-1 du foyer fiscal de référence ; notification de bourse CROUS ; certificat de scolarité de l'année en cours).

Aide numérique de l'établissement

<https://www.sciencespobordeaux.fr/attachment/cadre-reglementaire-accordeon/4.%20R%C3%A8glement%20aide%20num%C3%A9rique%20du%2030-09-22%20VF.pdf?download=true>

lien vers le questionnaire de demande :

<https://survey.sciencespobordeaux.fr/index.php/876469?lang=fr>

Par ailleurs, Sciences Po Bordeaux a également mis en place un service de prêt de machines (voir auprès de la DSI).

En cas de difficulté matérielle, vous pouvez également vous rapprocher de la scolarité de 1^{er} cycle, mais aussi de la responsable du service de la Vie étudiante et de l'égalité des chances (aidesciencespobordeaux.fr), bureau A 101.

2/ les principales **manipulations de bureautique doivent être maîtrisées** en amont et en dehors de cette conférence !! (Sur son ordinateur, SAVOIR télécharger un fichier, l'ouvrir dans un logiciel / une application, l'enregistrer dans un répertoire dédié). Assurez-vous de savoir faire cela à partir d'Excel par exemple (ou autre tableur de cette nature).

3/ L'équipe pédagogique a des ambitions pour vous ! **L'assiduité et la ponctualité** sont non négociables (une absence justifiée nécessitera le **rattrapage** préalable du cours pour être capable de suivre la séance suivante sans ralentir le groupe).

Calendrier et contenu des séances

NB : Le calendrier est susceptible de bouger à la marge en fonction du rythme d'avancée de chacun des groupes.

La plus grande partie des séances s'articulera autour de 3 grands volets :

- Volet 1/ Retour sur la(les) lecture(s) obligatoire(s) du jour (voir ci-dessous) avec interrogation express sur le texte (de type : indiquer ce que montre un graphique tiré d'un texte à lire pour la séance).
- Volet 2/ Apprentissage des notions statistiques et application sur le jeu de données ;
- Volet 3/ Application/travail en groupe pour avancer sur le '*4 pages*'.

La mise en œuvre des apprentissages au traitement de données se fera à partir du jeu de données tiré **de l'enquête web menée en décembre 2023 auprès des 1A et 3A scpobx**.

En plus des rendus intermédiaires dans le cadre du '*4 pages*', une à deux lectures obligatoires COURTES seront demandées pour la bonne compréhension des séances.

Il est attendu que les manipulations informatiques dans JAMOVI abordées pendant une séance ainsi que les notions statistiques en lien soient maîtrisées lors de la séance suivante (révision des manip' requise avant chaque séance). Si vous ne comprenez pas des notions, n'hésitez pas à réviser à plusieurs et à solliciter ceux d'entre vous les plus à l'aise !

Validation / évaluation → Contrôle continu

A. Réalisation d'un '4 pages' par groupe de 3 étudiant.es

Date de rendu du '4 pages' : **mercredi 15 avril 2023**

Attention : Les groupes doivent s'assurer qu'ils disposent d'au moins 1 ordinateur capable d'utiliser JAMOVI (ce qui n'est pas le cas des tablettes par exemple).

Les G4P (groupes de travail pour le '4 pages') peuvent tout à fait, à leur convenance, se servir du jeu de données utilisé pendant les séances pour leur travail en groupe, ou mobiliser les jeux proposés dans le moodle.

Aucun autre jeu de données (issus d'autres sources) ne sera en revanche accepté (raisons : risques RGPD ; risque de choisir un jeu de données de mauvaise qualité ou peu fiable ; ou risque de nécessiter un trop gros travail de préparation des données).

B. Travaux intermédiaires en vue de la réalisation du '4 pages'

Pour chaque rendu intermédiaire, il est demandé de préciser qui a fait quoi dans le trio (y compris tâches d'organisation, etc.).

1. Projet de '4 pages' (question posée - jeu de données mobilisé) et retroplanning (avec calage de créneaux de réunion de travail collectif + répartition du travail-qui a fait quoi).
2. Identification de 3 travaux de référence sur la thématique choisie pour le '4 pages' (articles, ouvrages scientifiques) accompagnés pour chacune d'une fiche synthétique de lecture d'une page (problématique - jeu de données et variables utilisées - principaux résultats **en lien avec le projet du '4 pages'**) + répartition du travail (qui a fait quoi).

NB : il est attendu que les groupes continuent d'enrichir ces premières lectures au cours du semestre, pour améliorer le travail final. La bibliographie finale se doit donc d'être davantage étoffée.

3. Schéma causal (opérationnalisation) : formulation de 2 à 3 hypothèses de travail avec leur mode d'opérationnalisation (quelles variables ? quels recodages ? quelle(s) méthode(s) statistique(s) ? + Mise en œuvre pour une des hypothèses). + répartition du travail (qui a fait quoi).
4. Exposé collectif en séance 11 sur le 4P. En 6 min - quel objet / quelles données / 1 résultat marquant / + 1 commentaire libre sur leur travail.

➔ Ces 4 travaux intermédiaires feront l'objet d'explicitations sur les attendus et les critères d'évaluation au fur et à mesure du semestre.

C. Lectures préparatoires aux séances

- Participation / interrogations à propos des lectures préparatoires (lectures courtes).

D. Validations individuelles des manipulations statistiques et informatiques abordées pendant les conférences

- Vous remettrez individuellement des exercices sur les manipulations apprises pendant les séances (tableaux de statistiques descriptives / statistiques bivariées et représentations graphiques / recodages / tests statistiques). Vous pouvez tout à fait dans ce cadre choisir une application en lien avec le thème du 4 pages pour avancer le projet collectif.

E. Participation (+ assiduité, ponctualité et solidarité) / Plagiat

Assiduité :

L'appel est systématiquement fait. Toute absence doit être justifiée d'abord auprès des enseignant·e·s, puis de la scolarité de 1A (Mme Amandine Forest et M. Pascal Davoury). Vous devez fournir un justificatif auprès de l'enseignant.e (certificat médical, convocation administrative, etc.).

En cas de rattrapage d'une séance en dehors de son horaire habituel, la règle de l'assiduité est assouplie.

Au-delà de **4 absences injustifiées sur l'année**, l'étudiant·e est renvoyé·e en 2^e session. Les motifs d'absence justifiée sont les suivants :

- Absence maladie avec certificat : une absence sans certificat de maladie n'est plus justifiée ;
- Absence avec justificatif : toute absence autre que pour maladie (RDV administratif incontournable, raison familiale impérieuse, examen dans le cadre d'un double cursus, etc.) ;
- Élu.e.s étudiant.e.s : participation de ces étudiants aux Commissions et Conseils de Sciences Po Bordeaux ;
- Motif non encore connu : en attendant l'éventuel certificat ou justificatif.

Toutes les autres absences seront considérées comme des absences injustifiées.

Ponctualité :

Les étudiant·e·s sont également tenu·e·s d'arriver à l'heure au cours. En cas de problème qui nécessite des aménagements particuliers, vous êtes invité·e·s à vous rapprocher de la scolarité.

Rendez-vous individuel :

Vous pouvez demander un rendez-vous individuel avec l'enseignant en cas de difficultés pédagogiques ou personnelles (prise de RV après la séance ou par courriel).

En dehors de questions spécifiques et d'un suivi personnalisé, toutes les questions pédagogiques relatives aux enseignements doivent être posées durant la conférence. En revanche, les questions administratives sont à adresser à la scolarité.

Il va de soi que le plagiat n'est pas toléré et sera durement sanctionné (citez les auteurs et les documents, pas de paraphrase).

De même, les deadlines doivent être respectées.

F. Synthèse des modalités d'évaluation

Exercices	Coefficient
Assiduité/participation/solidarité	1
'4 pages' rendu final	3
Rendus intermédiaires du '4 pages'	1
Interrogation sur les lectures préparatoires aux séances + validation individuelle des manipulations apprises en séances	1
Exposé	2
TOTAL	8

Références

Pierre Bréchon (dir), *Enquêtes qualitatives, enquêtes quantitatives*, Grenoble, PUG, 2011 (chapitres 6 à 10).

Flora Chanvril-Ligneel et Viviane Le Hay, *Méthodes statistiques pour les sciences sociales*, Paris, Ellipses, 2014, (chapitres 1, 2, 5, 6, 8).

William Fox, *Statistiques sociales*, Paris, De Boeck, 1999.

Olivier Martin, *L'enquête et ses méthodes : l'analyse de données quantitatives*, Paris, Armand Colin, 2005.

Serge Paugam (dir), *L'enquête sociologique*, Paris, PUF, 2010 (chapitres 4, 5 et 12).

Aide en ligne pour Jamovi

Guide utilisateur (pour les étudiant-es en psychologie) :

<https://jmeunierp8.github.io/ManuelJamovi/index.html>

Blog : <https://blog.jamovi.org/>

Plan de la conférence et lectures obligatoires :

BLOC 1 (3 SEANCES) : Utiliser un nouvel outil logiciel : JAMOVI Problématiser en quanti

Séance 1 – Séance introductive (17 janvier 2024)

- ° Présentation du semestre (programme, thèmes, exercices & validation, le moodle) ;
- ° Lecture d'un '4 pages' pour comprendre l'exercice attendu (identifier dans ce '4 pages' la question et la mise en œuvre empirique → passage des concepts à l'expérience empirique) ;
- ° Découverte du logiciel JAMOVI (Installation, Ouvrir un jeu de données) ;
- ° "Jeux" de prise en main avec JAMOVI (idéalement à partir de l'enquête maison : SCPOBX-ETU) ;
- ° Échanges pour former les groupes autour du '4 pages'.

A FAIRE AVANT LA SEANCE 2 :

Les groupes '4 pages' sont constitués et ont envoyé par mail avant la séance 2 (à l'enseignant.e) les membres du groupe, leur idée de thème et le jeu de données a priori utilisé (tiré de moodle).

Séance 2 – L'enquête par questionnaire 1/2 (24 janvier 2024)

- Lecture obligatoire : voir Moodle & indication en cours
- Notions abordées :
Bref panorama de la survey research (1/2) : les grands principes de [l'échantillonnage](#), [incertitude](#) et [marge d'erreur](#) ;
[Problématiser/conceptualiser](#) (spécificités) puis [Opérationnaliser](#) ;
➔ Application au jeu de données SCPOBX-ETU24 (quels questionnements possibles ?) et à la lecture obligatoire ;
- Travail en groupe pendant la séance (projet 4 pages) :
Exploration des jeux de données mis à disposition dans moodle avec leurs questionnaires pour ancrer les choix de thématique ;
Constitution d'une 1ère bibliographie + répartition des 1ères lectures.
- Bibliographie complémentaire (le cas échéant) : voir Moodle.

A RENDRE POUR LA SEANCE 3 :

Rendu intermédiaire 1 (projet avec jeu de données choisi et rétroplanning)

Séance 3 – L'enquête par questionnaire 2/2 (31 janvier 2024)

- Lecture obligatoire : voir Moodle & indication en cours
- Notions abordées :

Bref panorama de la survey research (2/2) : histoire, effet des [formulations](#) de questions et des [modes d'administration](#) ; Sensibilisation aux [biais de sondage](#) et à la [désirabilité sociale](#).

Jeu de données : Les différents [types de variables](#).

- Travail en groupe pendant la séance (projet 4 pages) :
Travail sur la question (problématique) ;
Identification des variables utiles (d'intérêt) dans le jeu de données & détermination de leur type.
- Bibliographie complémentaire (le cas échéant) : voir Moodle.

BLOC 2 (3 SEANCES) :
Premières opérationnalisations - Explorer, préparer, nettoyer :
bref, CONNAITRE ses données

Séance 4 – Statistique descriptive univariée (7 février 2024)

- Lecture obligatoire : voir Moodle & indication en cours
- Notions abordées :
Tendance centrale et dispersion ([moyenne](#), [mode](#), [médiane](#) / [écart-type](#), [quantiles](#)) ;
Tableau des [effectifs](#) et des [pourcentages](#) ;
Repérer les [données manquantes](#) & les [modalités rares](#).
- Travail en groupe pendant la séance (projet 4 pages) :
Application des notions abordées à son jeu de données.
- Bibliographie complémentaire (le cas échéant) : voir Moodle.

A RENDRE POUR LA SEANCE 5 :

Rendu intermédiaire 2 (travail bibliographique)

Séance 5 – Statistique bivariée : relation entre 2 variables quantitatives (14 février 2024)

- Lecture obligatoire : voir Moodle & indication en cours
- Notions abordées :
[Corrélation](#) et [nuage de points](#) (et [covariation](#)) ;
Données manquantes et [outliers](#) ;
Corrélation vs [causalité](#).
- Travail en groupe pendant la séance (projet 4 pages) :
Application des notions abordées à son jeu de données.
- Bibliographie complémentaire (le cas échéant) : voir Moodle.

Séance 6 – Statistique bivariable : relation entre 2 variables qualitatives (21 février 2024)

- Lecture obligatoire : voir Moodle & indication en cours
- Notions abordées :
 - Tableau croisé ([Pourcentages lignes](#) et [colonnes](#)) ;
 - Précautions statistiques descriptives (contrôle des effectifs) ;
 - Test statistiques et [Test du Khi2](#) (application à 2 variables qualitatives).
- Travail en groupe pendant la séance (projet 4 pages) :
 - Application des notions abordées à son jeu de données ;
 - Identifier les données manquantes et modalités rares (réfléchir aux options possibles).
- Bibliographie complémentaire (le cas échéant) : voir Moodle.

BLOC 3 (2 SEANCES + 1 SEANCE DE TRAVAIL) Opérationnalisations avancées : Adapter ses variables à son questionnaire théorique

A RENDRE POUR LA SEANCE 7 :

Rendu intermédiaire 3 (schéma causal)

Séance 7 – Recodages et construction d'indicateurs (6 mars 2024)

- Lecture obligatoire : voir Moodle & indication en cours
- Notions abordées :
 - [Recodages](#) et [construction d'indicateurs](#).
- Travail en groupe pendant la séance (projet 4 pages) :
 - Application des notions abordées à son jeu de données.
- Bibliographie complémentaire (le cas échéant) : voir Moodle.

VACANCES

Séance 8 – Travail en groupe sur le 4 pages (13 mars 2024)

- Lecture obligatoire (le cas échéant) : voir Moodle & indication en cours

!! Les plus avancé.es aident les groupes en difficulté !!

Séance 9 – Visualisation graphique de résultats d'enquête (20 mars 2024)

- Lecture obligatoire : voir Moodle & indication en cours
- Notions abordées :
Introduction aux principes de [visualisation graphique](#)
- Travail en groupe pendant la séance (projet 4 pages) :
Application des notions abordées à son jeu de données.
- Bibliographie complémentaire (le cas échéant) : voir Moodle.

BLOC 4 (1 SEANCE + 2 SEANCES DE TRAVAIL)
Inférer à partir de ses résultats (= conclure sur sa population d'intérêt)

Séance 10 – De l'échantillon à l'inférence – Retour sur la notion de test statistique (27 mars 2024)

- Lecture obligatoire : voir Moodle & indication en cours
- Notions abordées :
[Échantillonnage](#) - Retour sur les notions de représentativité (marge d'erreur) et sur la notion de test statistique - explication par la [loi normale](#).
- Travail en groupe pendant la séance (projet 4 pages) :
Application des notions abordées à son jeu de données.
- Bibliographie complémentaire (le cas échéant) : voir Moodle.

Séance 11 – [EXPOSE](#) en groupe sur le 4 pages (3 avril 2024)

- Lecture obligatoire (le cas échéant) : voir Moodle & indication en cours

Séance 12– [Travail en groupe sur le 4 pages](#) (10 mars 2024)

- Lecture obligatoire (le cas échéant) : voir Moodle & indication en cours

!! Les plus avancé.es aident les groupes en difficulté !!

