

Méthodes quantitatives pour les sciences sociales

Collège universitaire de Sciences Po, Campus de Poitiers

Mattéo Lanoë

Doctorant en sociologie

26 janvier 2026



Qui suis-je ?

Position actuelle

- ▶ Doctorant en sociologie, rattaché au Centre de recherche sur les inégalités sociales (Sciences Po/CNRS), à l'Institut national d'études démographiques (Ined), et au projet de recherche européen « Genre et inégalités sociales de santé » (Gendhi).
- ▶ Contact : matteo.lanoe@sciencespo.fr

Intérêts de recherche, disciplinaires

- ▶ Sociologie, épidémiologie sociale, science politique
- ▶ Sociologie de la santé, sociologie de la socialisation, sociologie des inégalités, parcours de vie
- ▶ Sujet de thèse : inégalités dans la gestion et l'appréhension des syndromes post-infectieux, comme le « Covid long ».

Organisation du cours

Ce cours est une introduction aux méthodes quantitatives en sciences sociales, et n'a pas pour objectif de vous rendre experts de ces méthodes.

Il est divisé en trois grands modules pédagogiques : outils statistiques et conceptuels de base pour les méthodes quantitatives ; utiliser et traiter des données ; tester la robustesse des associations réalisées.

- ▶ Séance 1 : À quoi peuvent bien servir les statistiques ?
- ▶ Séance 2 : D'où viennent les données statistiques ?
- ▶ Séance 3 : Décrire une variable
- ▶ Séance 4 & 5 : Relier des variables entre elles
- ▶ Séance 6 & 7 : Construire et catégoriser des données
- ▶ Séance 8 & 9 : Tester la significativité des associations
- ▶ Séance 10 : Introduction à la régression linéaire
- ▶ Séance 11 : Visualisation de données

Évaluation du cours

Ce cours sera évalué en deux temps :

1. De courts devoirs sur table après la fin de chaque module pédagogique, destinés à vérifier la compréhension et l'apprentissage des notions abordées (**$\frac{2}{3}$ de la note finale**).
2. Un devoir sur table final, commun à tous les étudiants suivant ce cours (**$\frac{1}{3}$ de la note finale**).
3. La participation orale pendant le cours comptera pour **10%** de la note finale.

À vous de jouer, présentez-vous !

Vos intérêts disciplinaires, thématiques, ambitions professionnelles, et potentielles peurs vis-à-vis de ce cours.

Séance 1 :

À quoi peuvent bien servir les statistiques ?

Mettre en chiffres la réalité sociale

Semestre de printemps 2026

Campus de Poitiers, Sciences Po Paris

Introduction de la séance

L'objectif de la séance d'aujourd'hui sera d'introduire les étudiants à la manière dont les statistiques se sont construites depuis la fin du XIXe siècle, en examinant certaines affirmations communément faites sur les méthodes quantitatives.

- ▶ La fin de la séance sera dédiée à l'installation de R puis de RStudio sur vos ordinateurs.
- ▶ Ceux qui n'ont pas la possibilité d'installer le logiciel sur leurs ordinateurs passeront par la version web (*via* <https://posit.cloud>, ou Posit Cloud).

Maintenant, Quizz !

<https://app.wooclap.com/events/SCIENCESPOQUANTI/live-session>

Discussion des résultats du Quizz

Retours sur les questions du quizz, discussion collective, et éléments de réponse sociologique et épistémologique.

Les statistiques sont un outil de contrôle de l'État sur sa population

Les statistiques ne sont pas neutres, elles sont liées à l'avènement de l'État moderne, qui s'en sert comme outils. Alain Desrosières distingue trois traditions :

Les statistiques sont un outil de contrôle de l'État sur sa population

Les statistiques ne sont pas neutres, elles sont liées à l'avènement de l'État moderne, qui s'en sert comme outils. Alain Desrosières distingue trois traditions :

- ▶ **La "Statistik" allemande (XVII^e)** : Approche descriptive et **holiste**. L'État cherche à connaître son territoire (climat, ressources, coutumes) pour unifier un empire morcelé.



Alain Desrosières

Alain Desrosières (2010). *La politique des grands nombres : Histoire de la raison statistique*. Poche / Sciences humaines et sociales. Paris : La Découverte

Les statistiques sont un outil de contrôle de l'État sur sa population

Les statistiques ne sont pas neutres, elles sont liées à l'avènement de l'État moderne, qui s'en sert comme outils. Alain Desrosières distingue trois traditions :

- ▶ **La "Statistik" allemande (XVII^e)** : Approche descriptive et **holiste**. L'État cherche à connaître son territoire (climat, ressources, coutumes) pour unifier un empire morcelé.
- ▶ **L'Arithmétique politique anglaise** : Approche **libérale**. L'État intervient peu, on utilise des registres et des calculs estimatifs.



Alain Desrosières

Alain Desrosières (2010). *La politique des grands nombres : Histoire de la raison statistique*. Poche / Sciences humaines et sociales. Paris : La Découverte

Les statistiques sont un outil de contrôle de l'État sur sa population

Les statistiques ne sont pas neutres, elles sont liées à l'avènement de l'État moderne, qui s'en sert comme outils. Alain Desrosières distingue trois traditions :

- ▶ **La "Statistik" allemande (XVII^e)** : Approche descriptive et **holiste**. L'État cherche à connaître son territoire (climat, ressources, coutumes) pour unifier un empire morcelé.
- ▶ **L'Arithmétique politique anglaise** : Approche **libérale**. L'État intervient peu, on utilise des registres et des calculs estimatifs.
- ▶ **La synthèse française (Révolution)** : Tentative d'unification ("Adunation") via les Préfets. Mélange de description littéraire et de quantification.



Alain Desrosières

Alain Desrosières (2010). *La politique des grands nombres : Histoire de la raison statistique*. Poche / Sciences humaines et sociales. Paris : La Découverte

Les statistiques sont un outil de contrôle de l'État sur sa population

Est-ce que connaître la population signifie la contrôler ?

Les statistiques sont un outil de contrôle de l'État sur sa population

Est-ce que connaître la population signifie la contrôler ?

- ▶ Michel Foucault montre le passage d'une logique de discipline (exclure les malades) à une logique de sécurité (gérer des probabilités) (**Foucault 2004**).



Michel Foucault

Michel Foucault (2004). *Sécurité, territoire, population : cours au Collège de France, 1977-1978*. Paris : Seuil/Gallimard

Les statistiques sont un outil de contrôle de l'État sur sa population

Est-ce que connaître la population signifie la contrôler ?

- ▶ Michel Foucault montre le passage d'une logique de discipline (exclure les malades) à une logique de sécurité (gérer des probabilités) (**Foucault 2004**).
- ▶ **L'exemple de la Variole** : On ne cherche plus à éradiquer totalement, mais à maintenir un taux de mortalité "socialement acceptable" via l'inoculation.



Michel Foucault

Michel Foucault (2004). *Sécurité, territoire, population : cours au Collège de France, 1977-1978*. Paris : Seuil/Gallimard

Les statistiques sont un outil de contrôle de l'État sur sa population

Est-ce que connaître la population signifie la contrôler ?

- ▶ Michel Foucault montre le passage d'une logique de discipline (exclure les malades) à une logique de sécurité (gérer des probabilités) (**Foucault 2004**).
- ▶ **L'exemple de la Variole** : On ne cherche plus à éradiquer totalement, mais à maintenir un taux de mortalité "socialement acceptable" via l'inoculation.
- ▶ La statistique permet de voir la population comme un ensemble de **processus** et de régularités.



Michel Foucault

Michel Foucault (2004). *Sécurité, territoire, population : cours au Collège de France, 1977-1978*. Paris : Seuil/Gallimard

Les statistiques sont un outil de contrôle de l'État sur sa population

Est-ce que connaître la population signifie la contrôler ?

- ▶ Michel Foucault montre le passage d'une logique de discipline (exclure les malades) à une logique de sécurité (gérer des probabilités) (**Foucault 2004**).
- ▶ **L'exemple de la Variole** : On ne cherche plus à éradiquer totalement, mais à maintenir un taux de mortalité "socialement acceptable" via l'inoculation.
- ▶ La statistique permet de voir la population comme un ensemble de **processus** et de régularités.
- ▶ **Conclusion** : Les données servent à la connaissance ET au contrôle (gestion des risques).



Michel Foucault

Michel Foucault (2004). *Sécurité, territoire, population : cours au Collège de France, 1977-1978*. Paris : Seuil/Gallimard

Les données statistiques sont les seules données empiriques qui permettent de généraliser des résultats de sciences sociales.

Peut-on généraliser sans statistiques ?

Idée reçue : "Seuls les chiffres permettent de généraliser."

Réponse : Faux. Il existe deux régimes de preuve distincts.

Les données statistiques sont les seules données empiriques qui permettent de généraliser des résultats de sciences sociales.

Peut-on généraliser sans statistiques ?

Idée reçue : "Seuls les chiffres permettent de généraliser."

Réponse : Faux. Il existe deux régimes de preuve distincts.

Approche Quantitative

- Principe : **Représentativité**.

Approche Qualitative

Mario Luis Small (2009). *How many cases do I need ? On science and the logic of case selection in field-based research.* In : *Ethnography* 10.1, p. 5-38

Les données statistiques sont les seules données empiriques qui permettent de généraliser des résultats de sciences sociales.

Peut-on généraliser sans statistiques ?

Idée reçue : "Seuls les chiffres permettent de généraliser."

Réponse : Faux. Il existe deux régimes de preuve distincts.

Approche Quantitative

- ▶ Principe : **Représentativité**.
- ▶ Outil : Échantillon aléatoire ou par quotas.

Approche Qualitative

Mario Luis Small (2009). *How many cases do I need ? On science and the logic of case selection in field-based research.* In : *Ethnography* 10.1, p. 5-38

Les données statistiques sont les seules données empiriques qui permettent de généraliser des résultats de sciences sociales.

Peut-on généraliser sans statistiques ?

Idée reçue : "Seuls les chiffres permettent de généraliser."

Réponse : Faux. Il existe deux régimes de preuve distincts.

Approche Quantitative

- ▶ Principe : **Représentativité**.
- ▶ Outil : Échantillon aléatoire ou par quotas.
- ▶ Objectif : Dresser un "portrait" fidèle à grande échelle (Loi des grands nombres).

Approche Qualitative

Mario Luis Small (2009). *How many cases do I need ? On science and the logic of case selection in field-based research.* In : *Ethnography* 10.1, p. 5-38

Les données statistiques sont les seules données empiriques qui permettent de généraliser des résultats de sciences sociales.

Peut-on généraliser sans statistiques ?

Idée reçue : "Seuls les chiffres permettent de généraliser."

Réponse : Faux. Il existe deux régimes de preuve distincts.

Approche Quantitative

- ▶ Principe : **Représentativité**.
- ▶ Outil : Échantillon aléatoire ou par quotas.
- ▶ Objectif : Dresser un "portrait" fidèle à grande échelle (Loi des grands nombres).

Approche Qualitative

- ▶ Principe : **Saturation**.

Mario Luis Small (2009). *How many cases do I need ? On science and the logic of case selection in field-based research.* In :

Ethnography 10.1, p. 5-38

Les données statistiques sont les seules données empiriques qui permettent de généraliser des résultats de sciences sociales.

Peut-on généraliser sans statistiques ?

Idée reçue : "Seuls les chiffres permettent de généraliser."

Réponse : Faux. Il existe deux régimes de preuve distincts.

Approche Quantitative

- ▶ Principe : **Représentativité**.
- ▶ Outil : Échantillon aléatoire ou par quotas.
- ▶ Objectif : Dresser un "portrait" fidèle à grande échelle (Loi des grands nombres).

Approche Qualitative

- ▶ Principe : **Saturation**.
- ▶ Concept de Mario L. Small (**Small 2009**).

Mario Luis Small (2009). *How many cases do I need ? On science and the logic of case selection in field-based research.* In :

Ethnography 10.1, p. 5-38

Les données statistiques sont les seules données empiriques qui permettent de généraliser des résultats de sciences sociales.

Peut-on généraliser sans statistiques ?

Idée reçue : "Seuls les chiffres permettent de généraliser."

Réponse : Faux. Il existe deux régimes de preuve distincts.

Approche Quantitative

- ▶ Principe : **Représentativité**.
- ▶ Outil : Échantillon aléatoire ou par quotas.
- ▶ Objectif : Dresser un "portrait" fidèle à grande échelle (Loi des grands nombres).

Approche Qualitative

- ▶ Principe : **Saturation**.
- ▶ Concept de Mario L. Small (**Small 2009**).
- ▶ Logique : On arrête quand les nouveaux entretiens n'apportent plus d'information nouvelle (répétition).

Mario Luis Small (2009). *How many cases do I need ? On science and the logic of case selection in field-based research*. In : *Ethnography* 10.1, p. 5-38

Les statistiques essentialisent et figent les individus dans des catégories alors que le monde social est beaucoup plus complexe

Les statistiques figent la réalité pour pouvoir la traiter, mais elles ne doivent pas l'essentialiser.

Les statistiques essentialisent et figent les individus dans des catégories alors que le monde social est beaucoup plus complexe

Les statistiques figent la réalité pour pouvoir la traiter, mais elles ne doivent pas l'essentialiser.

- ▶ **La méthode** : En quanti, les catégories sont construites *a priori* (avant l'enquête), alors qu'en quali elles sont souvent construites *a posteriori*.

Alain Desrosières (2008). *Pour une sociologie historique de la quantification : L'argument statistique I*. Paris : Presses des Mines

Les statistiques essentialisent et figent les individus dans des catégories alors que le monde social est beaucoup plus complexe

Les statistiques figent la réalité pour pouvoir la traiter, mais elles ne doivent pas l'essentialiser.

- ▶ **La méthode** : En quanti, les catégories sont construites *a priori* (avant l'enquête), alors qu'en quali elles sont souvent construites *a posteriori*.
- ▶ **Le danger** : Oublier que ces catégories sont des construits sociaux.

Alain Desrosières (2008). *Pour une sociologie historique de la quantification : L'argument statistique I*. Paris : Presses des Mines

Les statistiques essentialisent et figent les individus dans des catégories alors que le monde social est beaucoup plus complexe

Les statistiques figent la réalité pour pouvoir la traiter, mais elles ne doivent pas l'essentialiser.

- ▶ **La méthode** : En quanti, les catégories sont construites *a priori* (avant l'enquête), alors qu'en quali elles sont souvent construites *a posteriori*.
- ▶ **Le danger** : Oublier que ces catégories sont des construits sociaux.
- ▶ Alain Desrosières rappelle que **Quantifier = Convenir + Mesurer**.

Alain Desrosières (2008). *Pour une sociologie historique de la quantification : L'argument statistique I*. Paris : Presses des Mines

Les statistiques essentialisent et figent les individus dans des catégories alors que le monde social est beaucoup plus complexe

Les statistiques figent la réalité pour pouvoir la traiter, mais elles ne doivent pas l'essentialiser.

- ▶ **La méthode** : En quanti, les catégories sont construites *a priori* (avant l'enquête), alors qu'en quali elles sont souvent construites *a posteriori*.
- ▶ **Le danger** : Oublier que ces catégories sont des construits sociaux.
- ▶ Alain Desrosières rappelle que **Quantifier = Convenir + Mesurer**.
 1. **Convenir** : Se mettre d'accord sur la définition (Ex : Qu'est-ce qu'un immigré ? Né à l'étranger ET de nationalité étrangère à la naissance).

Alain Desrosières (2008). *Pour une sociologie historique de la quantification : L'argument statistique I*. Paris : Presses des Mines

Les statistiques essentialisent et figent les individus dans des catégories alors que le monde social est beaucoup plus complexe

Les statistiques figent la réalité pour pouvoir la traiter, mais elles ne doivent pas l'essentialiser.

- ▶ **La méthode** : En quanti, les catégories sont construites *a priori* (avant l'enquête), alors qu'en quali elles sont souvent construites *a posteriori*.
- ▶ **Le danger** : Oublier que ces catégories sont des construits sociaux.
- ▶ Alain Desrosières rappelle que **Quantifier = Convenir + Mesurer**.
 1. **Convenir** : Se mettre d'accord sur la définition (Ex : Qu'est-ce qu'un immigré ? Né à l'étranger ET de nationalité étrangère à la naissance).
 2. **Mesurer** : Mettre en chiffre cette convention.

Alain Desrosières (2008). *Pour une sociologie historique de la quantification : L'argument statistique I*. Paris : Presses des Mines

Qui produit la recherche quantitative ?

Idée reçue : "Les travaux quantitatifs sont l'affaire des statisticiens, pas des sociologues."

Qui produit la recherche quantitative ?

Idée reçue : "Les travaux quantitatifs sont l'affaire des statisticiens, pas des sociologues."

- ▶ **C'est faux.** Le producteur est avant tout un chercheur en sciences sociales.

Qui produit la recherche quantitative ?

Idée reçue : "Les travaux quantitatifs sont l'affaire des statisticiens, pas des sociologues."

- ▶ **C'est faux.** Le producteur est avant tout un chercheur en sciences sociales.
- ▶ **La compétence requise :** Une "culture générale" statistique (comprendre le raisonnement) est nécessaire, mais pas une expertise mathématique pure. Même sans pratiquer, il faut pouvoir lire ses collègues.

Qui produit la recherche quantitative ?

Idée reçue : "Les travaux quantitatifs sont l'affaire des statisticiens, pas des sociologues."

- ▶ **C'est faux.** Le producteur est avant tout un chercheur en sciences sociales.
- ▶ **La compétence requise :** Une "culture générale" statistique (comprendre le raisonnement) est nécessaire, mais pas une expertise mathématique pure. Même sans pratiquer, il faut pouvoir lire ses collègues.
- ▶ **Le rôle du politiste/sociologue :**

Qui produit la recherche quantitative ?

Idée reçue : "Les travaux quantitatifs sont l'affaire des statisticiens, pas des sociologues."

- ▶ **C'est faux.** Le producteur est avant tout un chercheur en sciences sociales.
- ▶ **La compétence requise :** Une "culture générale" statistique (comprendre le raisonnement) est nécessaire, mais pas une expertise mathématique pure. Même sans pratiquer, il faut pouvoir lire ses collègues.
- ▶ **Le rôle du politiste/sociologue :**
 - ▶ Donner du **sens social** aux chiffres.

Qui produit la recherche quantitative ?

Idée reçue : "Les travaux quantitatifs sont l'affaire des statisticiens, pas des sociologues."

- ▶ **C'est faux.** Le producteur est avant tout un chercheur en sciences sociales.
- ▶ **La compétence requise :** Une "culture générale" statistique (comprendre le raisonnement) est nécessaire, mais pas une expertise mathématique pure. Même sans pratiquer, il faut pouvoir lire ses collègues.
- ▶ **Le rôle du politiste/sociologue :**
 - ▶ Donner du **sens social** aux chiffres.
 - ▶ Nuancer et relativiser les résultats bruts.

Qui produit la recherche quantitative ?

Idée reçue : "Les travaux quantitatifs sont l'affaire des statisticiens, pas des sociologues."

- ▶ **C'est faux.** Le producteur est avant tout un chercheur en sciences sociales.
- ▶ **La compétence requise :** Une "culture générale" statistique (comprendre le raisonnement) est nécessaire, mais pas une expertise mathématique pure. Même sans pratiquer, il faut pouvoir lire ses collègues.
- ▶ **Le rôle du politiste/sociologue :**
 - ▶ Donner du **sens social** aux chiffres.
 - ▶ Nuancer et relativiser les résultats bruts.
 - ▶ Faire dialoguer les données avec la littérature existante.

Faut-il être fort en maths ?

"Peut-on manipuler des données sans avoir étudié ou aimé les maths?"

Faut-il être fort en maths ?

"Peut-on manipuler des données sans avoir étudié ou aimé les maths?"

- ▶ **Réponse : OUI.** C'est précisément l'objectif de ce module.

Faut-il être fort en maths ?

"Peut-on manipuler des données sans avoir étudié ou aimé les maths?"

- ▶ **Réponse : OUI.** C'est précisément l'objectif de ce module.
- ▶ Les statistiques sont ici des **outils**, des objets relativement simples à manipuler une fois démystifiés.

Faut-il être fort en maths ?

"Peut-on manipuler des données sans avoir étudié ou aimé les maths ?"

- ▶ **Réponse : OUI.** C'est précisément l'objectif de ce module.
- ▶ Les statistiques sont ici des **outils**, des objets relativement simples à manipuler une fois démystifiés.
- ▶ **Le levier d'apprentissage** : L'ancrage dans la réalité.

Faut-il être fort en maths ?

"Peut-on manipuler des données sans avoir étudié ou aimé les maths ?"

- ▶ **Réponse : OUI.** C'est précisément l'objectif de ce module.
- ▶ Les statistiques sont ici des **outils**, des objets relativement simples à manipuler une fois démystifiés.
- ▶ **Le levier d'apprentissage :** L'ancrage dans la réalité.
- ▶ Appliquer ces méthodes sur des sujets politiques ou sociaux qui vous intéressent rend la technique beaucoup plus accessible et concrète.

L'environnement de travail : R et RStudio

Avant d'installer, il faut comprendre la distinction. On utilise souvent l'analogie de la voiture.

1. Le langage R *(Le Moteur)*

- ▶ C'est le cœur du système.

2. Le logiciel RStudio *(Le Tableau de bord)*

L'environnement de travail : R et RStudio

Avant d'installer, il faut comprendre la distinction. On utilise souvent l'analogie de la voiture.

1. Le langage R *(Le Moteur)*

- ▶ C'est le cœur du système.
- ▶ Il effectue les calculs et stocke les données.

2. Le logiciel RStudio *(Le Tableau de bord)*

L'environnement de travail : R et RStudio

Avant d'installer, il faut comprendre la distinction. On utilise souvent l'analogie de la voiture.

1. Le langage R *(Le Moteur)*

- ▶ C'est le cœur du système.
- ▶ Il effectue les calculs et stocke les données.
- ▶ Austère et difficile à utiliser seul.

2. Le logiciel RStudio *(Le Tableau de bord)*

L'environnement de travail : R et RStudio

Avant d'installer, il faut comprendre la distinction. On utilise souvent l'analogie de la voiture.

1. Le langage R *(Le Moteur)*

- ▶ C'est le cœur du système.
- ▶ Il effectue les calculs et stocke les données.
- ▶ Austère et difficile à utiliser seul.

2. Le logiciel RStudio *(Le Tableau de bord)*

- ▶ C'est l'interface que vous ouvrirez.

L'environnement de travail : R et RStudio

Avant d'installer, il faut comprendre la distinction. On utilise souvent l'analogie de la voiture.

1. Le langage R *(Le Moteur)*

- ▶ C'est le cœur du système.
- ▶ Il effectue les calculs et stocke les données.
- ▶ Austère et difficile à utiliser seul.

2. Le logiciel RStudio *(Le Tableau de bord)*

- ▶ C'est l'interface que vous ouvrirez.
- ▶ Il rend R utilisable (fenêtres, graphiques, aide).

L'environnement de travail : R et RStudio

Avant d'installer, il faut comprendre la distinction. On utilise souvent l'analogie de la voiture.

1. Le langage R *(Le Moteur)*

- ▶ C'est le cœur du système.
- ▶ Il effectue les calculs et stocke les données.
- ▶ Austère et difficile à utiliser seul.

2. Le logiciel RStudio *(Le Tableau de bord)*

- ▶ C'est l'interface que vous ouvrirez.
- ▶ Il rend R utilisable (fenêtres, graphiques, aide).
- ▶ **Il ne fonctionne pas sans R !**

Installation : La marche à suivre

Attention : L'ordre d'installation est important.

Installation : La marche à suivre

Attention : L'ordre d'installation est important.

1. Étape 1 : Installer R (Le moteur)

Installation : La marche à suivre

Attention : L'ordre d'installation est important.

1. Étape 1 : Installer R (Le moteur)

- ▶ Allez sur le CRAN : <https://cloud.r-project.org/>

Installation : La marche à suivre

Attention : L'ordre d'installation est important.

1. Étape 1 : Installer R (Le moteur)

- ▶ Allez sur le CRAN : <https://cloud.r-project.org/>
- ▶ Téléchargez la version pour votre système (Mac/Windows).

Installation : La marche à suivre

Attention : L'ordre d'installation est important.

1. Étape 1 : Installer R (Le moteur)

- ▶ Allez sur le CRAN : <https://cloud.r-project.org/>
- ▶ Téléchargez la version pour votre système (Mac/Windows).
- ▶ Installez-le comme un logiciel classique.

Installation : La marche à suivre

Attention : L'ordre d'installation est important.

1. Étape 1 : Installer R (Le moteur)

- ▶ Allez sur le CRAN : <https://cloud.r-project.org/>
- ▶ Téléchargez la version pour votre système (Mac/Windows).
- ▶ Installez-le comme un logiciel classique.

2. Étape 2 : Installer RStudio (L'interface)

Installation : La marche à suivre

Attention : L'ordre d'installation est important.

1. Étape 1 : Installer R (Le moteur)

- ▶ Allez sur le CRAN : <https://cloud.r-project.org/>
- ▶ Téléchargez la version pour votre système (Mac/Windows).
- ▶ Installez-le comme un logiciel classique.

2. Étape 2 : Installer RStudio (L'interface)

- ▶ Allez sur le site de Posit : <https://posit.co/download/rstudio-desktop/>

Installation : La marche à suivre

Attention : L'ordre d'installation est important.

1. Étape 1 : Installer R (Le moteur)

- ▶ Allez sur le CRAN : <https://cloud.r-project.org/>
- ▶ Téléchargez la version pour votre système (Mac/Windows).
- ▶ Installez-le comme un logiciel classique.

2. Étape 2 : Installer RStudio (L'interface)

- ▶ Allez sur le site de Posit : <https://posit.co/download/rstudio-desktop/>
- ▶ Choisissez "RStudio Desktop".

Installation : La marche à suivre

Attention : L'ordre d'installation est important.

1. Étape 1 : Installer R (Le moteur)

- ▶ Allez sur le CRAN : <https://cloud.r-project.org/>
- ▶ Téléchargez la version pour votre système (Mac/Windows).
- ▶ Installez-le comme un logiciel classique.

2. Étape 2 : Installer RStudio (L'interface)

- ▶ Allez sur le site de Posit : <https://posit.co/download/rstudio-desktop/>
- ▶ Choisissez "RStudio Desktop".
- ▶ Installez-le.

Installation : La marche à suivre

Attention : L'ordre d'installation est important.

1. Étape 1 : Installer R (Le moteur)

- ▶ Allez sur le CRAN : <https://cloud.r-project.org/>
- ▶ Téléchargez la version pour votre système (Mac/Windows).
- ▶ Installez-le comme un logiciel classique.

2. Étape 2 : Installer RStudio (L'interface)

- ▶ Allez sur le site de Posit : <https://posit.co/download/rstudio-desktop/>
- ▶ Choisissez "RStudio Desktop".
- ▶ Installez-le.

3. Pour travailler : Vous n'ouvrirez que RStudio (l'icône bleue ronde), jamais R directement.

Étape 1 : Préparer le terrain

Pour importer des données, R doit savoir où elles sont rangées.

Étape 1 : Préparer le terrain

Pour importer des données, R doit savoir où elles sont rangées.

1. Téléchargez le fichier titanic.csv (sur le site : <https://matteolanoe.github.io/teaching.html>).

Étape 1 : Préparer le terrain

Pour importer des données, R doit savoir où elles sont rangées.

- 1. Téléchargez** le fichier `titanic.csv` (sur le site :
<https://matteolanoe.github.io/teaching.html>).
- 2. Déplacez** ce fichier dans le dossier de votre projet sur votre ordinateur.

Étape 1 : Préparer le terrain

Pour importer des données, R doit savoir où elles sont rangées.

1. **Téléchargez** le fichier `titanic.csv` (sur le site :
<https://matteolanoe.github.io/teaching.html>).
2. **Déplacez** ce fichier dans le dossier de votre projet sur votre ordinateur.
3. **Important** : Le fichier de données (`.csv`) et votre script R (`.R`) doivent idéalement être dans le **même dossier**.

Étape 1 : Préparer le terrain

Pour importer des données, R doit savoir où elles sont rangées.

1. **Téléchargez** le fichier `titanic.csv` (sur le site :
<https://matteolanoe.github.io/teaching.html>).
2. **Déplacez** ce fichier dans le dossier de votre projet sur votre ordinateur.
3. **Important** : Le fichier de données (`.csv`) et votre script R (`.R`) doivent idéalement être dans le **même dossier**.
4. Dans RStudio, vérifiez que vous voyez le fichier apparaître dans l'onglet **Files** (en bas à droite).

Étape 2 : La commande d'importation

Nous allons utiliser la fonction `read.csv()` et l'opérateur d'assignation `<-`.

Le Code à taper

```
titanic <- read.csv("titanic.csv")
```

Étape 2 : La commande d'importation

Nous allons utiliser la fonction `read.csv()` et l'opérateur d'assignation `<-`.

Le Code à taper

```
titanic <- read.csv("titanic.csv")
```

Décortiquons cette ligne :

- ▶ `titanic` : Le nom que je donne à mon objet dans R (le conteneur).

Étape 2 : La commande d'importation

Nous allons utiliser la fonction `read.csv()` et l'opérateur d'assignation `<-`.

Le Code à taper

```
titanic <- read.csv("titanic.csv")
```

Décortiquons cette ligne :

- ▶ `titanic` : Le nom que je donne à mon objet dans R (le conteneur).
- ▶ `<-` : La flèche d'assignation ("mets le contenu de droite dans l'objet de gauche").

Étape 2 : La commande d'importation

Nous allons utiliser la fonction `read.csv()` et l'opérateur d'assignation `<-`.

Le Code à taper

```
titanic <- read.csv("titanic.csv")
```

Décortiquons cette ligne :

- ▶ `titanic` : Le nom que je donne à mon objet dans R (le conteneur).
- ▶ `<-` : La flèche d'assignation ("mets le contenu de droite dans l'objet de gauche").
- ▶ `read.csv(...)` : La fonction (le verbe) qui lit le fichier.

Étape 2 : La commande d'importation

Nous allons utiliser la fonction `read.csv()` et l'opérateur d'assignation `<-`.

Le Code à taper

```
titanic <- read.csv("titanic.csv")
```

Décortiquons cette ligne :

- ▶ `titanic` : Le nom que je donne à mon objet dans R (le conteneur).
- ▶ `<-` : La flèche d'assignation ("mets le contenu de droite dans l'objet de gauche").
- ▶ `read.csv(...)` : La fonction (le verbe) qui lit le fichier.
- ▶ `"titanic.csv"` : Le nom exact du fichier (entre guillemets !).

Étape 3 : Vérifier que ça a marché

Si aucune erreur rouge n'apparaît, c'est bon signe. Mais il faut toujours vérifier.

Étape 3 : Vérifier que ça a marché

Si aucune erreur rouge n'apparaît, c'est bon signe. Mais il faut toujours vérifier.

- ▶ Regardez l'onglet "Environment" (haut droite) :

Étape 3 : Vérifier que ça a marché

Si aucune erreur rouge n'apparaît, c'est bon signe. Mais il faut toujours vérifier.

► Regardez l'onglet "Environment" (haut droite) :

- Vous devriez voir une ligne titanic avec le nombre d'observations (lignes) et de variables (colonnes).

Étape 3 : Vérifier que ça a marché

Si aucune erreur rouge n'apparaît, c'est bon signe. Mais il faut toujours vérifier.

► **Regardez l'onglet "Environment" (haut droite) :**

- Vous devriez voir une ligne titanic avec le nombre d'observations (lignes) et de variables (colonnes).

► **Explorez les premières lignes :**

```
head(titanic)
```

Étape 3 : Vérifier que ça a marché

Si aucune erreur rouge n'apparaît, c'est bon signe. Mais il faut toujours vérifier.

► **Regardez l'onglet "Environment" (haut droite) :**

- Vous devriez voir une ligne titanic avec le nombre d'observations (lignes) et de variables (colonnes).

► **Explorez les premières lignes :**

```
head(titanic)
```

► **Ouvrir le tableau complet** (comme Excel) :

Étape 3 : Vérifier que ça a marché

Si aucune erreur rouge n'apparaît, c'est bon signe. Mais il faut toujours vérifier.

► **Regardez l'onglet "Environment" (haut droite) :**

- Vous devriez voir une ligne titanic avec le nombre d'observations (lignes) et de variables (colonnes).

► **Explorez les premières lignes :**

```
head(titanic)
```

► **Ouvrir le tableau complet** (comme Excel) :

- Cliquez sur le nom titanic dans l'environnement, OU tapez :

```
View(titanic)
```

Étape 3 : Vérifier que ça a marché

Si aucune erreur rouge n'apparaît, c'est bon signe. Mais il faut toujours vérifier.

► **Regardez l'onglet "Environment" (haut droite) :**

- Vous devriez voir une ligne titanic avec le nombre d'observations (lignes) et de variables (colonnes).

► **Explorez les premières lignes :**

```
head(titanic)
```

► **Ouvrir le tableau complet** (comme Excel) :

- Cliquez sur le nom titanic dans l'environnement, OU tapez :

```
View(titanic)
```

- *Attention : V majuscule à View !*