

Cours Pasteur

Analyse de données avec Stata

Année universitaire 2025-2026

SESSION 1 :

Commandes Stata et tests univariés

13 - 17 octobre 2025

Co-directeurs

Arnaud FONTANET

Unité d'Epidémiologie des Maladies Emergentes
Institut Pasteur

Loïc CHARTIER

LYSARC
Centre Hospitalier Lyon Sud

Cheffes de travaux

Laura TONDEUR

Unité d'Epidémiologie des Maladies Emergentes
Institut Pasteur

Laura SCHAEFFER

Unité d'Epidémiologie des Maladies Emergentes
Institut Pasteur

DESCRIPTIF DU COURS

Cette première session Stata alternant cours théoriques et travaux dirigés débutera par la manipulation de fichiers (ouverture, importation de données, fusion de fichiers) et l'exploration d'une base de données (structure, recherche de doublons, données manquantes, valeurs aberrantes, recodage des variables). Puis seront étudiés les tests statistiques (paramétriques et non paramétriques) comparant des variables discrètes ou continues entre deux groupes indépendants (Test t de Student, test de Mann-Whitney, test du Chi² ou test exact de Fisher), plus de deux groupes indépendants (Analyse de variance, test de Kruskal-Wallis, comparaisons multiples) ou deux groupes appariés (Test de Student sur séries appariées, test des rangs signés de Wilcoxon, test de Mc Nemar). Seront également abordés les tests de corrélation entre deux variables d'un même groupe et très brièvement la régression linéaire. Le cours se terminera par le calcul de taille d'échantillon sur Stata.

Analyse de données avec Stata

- Session 1 -

Lundi 13 octobre 2025

9h00 - 9h30	Accueil des étudiants	Centre d'enseignement
9h30 - 13h00	Introduction à l'analyse de données sur Stata	L. Chartier L. Schaeffer L. Tondeur
14h20 - 17h20	Exploration d'un fichier de données	L. Chartier L. Schaeffer L. Tondeur

Mardi 14 octobre 2025

9h00 - 13h00	Création de variables	L. Chartier L. Schaeffer L. Tondeur
14h00 - 17h00	Organisation d'une base de données	L. Chartier L. Schaeffer L. Tondeur

Mercredi 15 octobre 2025

9h00 - 13h00	Contrôle de qualité des données / nettoyage	L. Chartier L. Schaeffer L. Tondeur
14h00 - 17h00	Commande pour faire des graphiques	L. Chartier L. Schaeffer L. Tondeur

Jeudi 16 octobre 2025

9h00 - 12h30	Comparaison de variables continues	A. Fontanet L. Schaeffer L. Tondeur
13h30 - 17h00	Comparaison de variables discrètes et corrélations	A. Fontanet L. Schaeffer L. Tondeur

Vendredi 17 octobre 2025

9h00 - 12h30	Taille d'échantillon / Révisions	A. Fontanet L. Schaeffer L. Tondeur
13h30 - 16h30	Examen: analyse d'une base de données à l'aide du logiciel Stata (faite de façon individuelle et guidée par des questions) Durée : 3h00, noté sur 20	