

ANALYSE DE DONNEES AVEC STATA – SESSION 1

COMMANDES STATA ET TESTS UNIVARIES

DIRECTEURS DE COURS

Monsieur Loïc CHARTIER

LYSARC - Centre Hospitalier Lyon-Sud
Secteur Sainte Eugénie – Bâtiment 2D
69495 Pierre-Bénite Cedex
Tél. : +33 (0)4 72 66 93 33
Email : loic.chartier@lysarc.org

Monsieur Arnaud FONTANET

Institut Pasteur
Unité d'Epidémiologie des Maladies Emergentes
25/28 rue du Dr Roux
Tél. : +33 (0)1 40 61 37 63
75724 Paris Cedex 15
Email : arnaud.fontanet@pasteur.fr

CHEFFES DES TRAVAUX

Madame Laura SCHAEFFER

Institut Pasteur
Unité d'Epidémiologie des Maladies Emergentes
25/28 rue du Dr Roux
Tél. : +33 (0)1 45 68 83 23
75724 Paris Cedex 15
Email : laura.schaeffer@pasteur.fr

Madame Laura TONDEUR

Institut Pasteur
Unité d'Epidémiologie des Maladies Emergentes
25/28 rue du Dr Roux
Tél. : +33 (0)1 40 61 39 85
75724 Paris Cedex 15
Email : laura.tondeur@pasteur.fr

ADRESSE & CONTACT

Institut Pasteur – Centre d'enseignement
28 rue du Docteur Roux, 75015 Paris
Salle 5 – Bâtiment 6

Pôle administratif

Aurore LAFFOREST
28 rue du Docteur Roux
Bâtiment 10A – Pavillon Emile Roux – RDC
Tél. : 06 82 08 89 82
Email : aurore.lafforest@pasteur.fr / stata@pasteur.fr

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.

* * * * *

PROGRAMME DU 12 AU 16 OCTOBRE 2026

LUNDI 12 OCTOBRE 2026 - SALLE 5, BATIMENT 6

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 13 OCTOBRE 2026 - SALLE 5, BATIMENT 6

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 14 OCTOBRE 2026 - SALLE 5, BATIMENT 6

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 15 OCTOBRE 2026 - SALLE 5, BATIMENT 6

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 16 OCTOBRE 2025 - SALLE 5, BATIMENT 6

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	