

Ces formations pourraient également vous intéresser

DIPLÔME D'ÉTABLISSEMENT

SCIENCE DES DONNÉES DE SANTÉ

CYCLE

TRAITEMENT DES DONNÉES PMSI

PARTENAIRE

atih

AGENCE TECHNIQUE
DE L'INFORMATION
SUR L'HOSPITALISATION

Ecole des hautes études en santé publique
15, avenue du Professeur-Léon-Bernard
CS 74312 - 35043 Rennes Cedex
Tél. + 33(0)2 99 02 22 00

formation-continue.ehesp.fr

DURÉE

6 jours le cycle complet
3 modules de 2 jours

DATES

Module 1 : 14 au 15 octobre 2024
Module 2 : 16 au 17 octobre 2024
Module 3 : 25 au 26 novembre 2024

TARIFS

2050 € net le cycle complet
900 € le module de 2 jours

MODALITÉS DE CANDIDATURE

Votre candidature sera soumise au responsable pédagogique. Une confirmation d'inscription vous sera adressée au plus tard 3 semaines avant le début de la formation.

Conditions générales de vente disponibles sur notre site.

Qui contacter si vous êtes en situation de handicap ?

Référent handicap
referent-handicap@ehesp.fr

Une question concernant votre inscription ?

Direction de la scolarité
candidaturefc@ehesp.fr
Tél. : 02.99.02.25.00



Inscription et modalités sur notre site

PAGE [PYTHON POUR L'ANALYSE DES DONNÉES DE SANTÉ](#)
OU EN RECHERCHANT LA FORMATION PAR THEME

[Traitement des données de santé : PMSI et SNDS](#)

formation-continue.ehesp.fr

L'EHEP est organisme de formation enregistré sous le n° 53 35 09162 35
Cet enregistrement ne vaut pas agrément de l'Etat

Organisme enregistré par l'Agence nationale du DPC N° 2371

Les personnes en situation de handicap souhaitant bénéficier d'un accueil spécifique peuvent s'adresser par courriel à : referent-handicap@ehesp.fr

Conformément à la loi «informatique et libertés» du 6 janvier 1978 et au RGPD 2016/679 du 27 avril 2016 vous bénéficiez d'un droit d'accès et de rectification aux informations qui vous concernent. Vous pouvez également, pour des motifs légitimes, vous opposer au traitement des données vous concernant. Si vous souhaitez exercer ce droit et obtenir communication des informations vous concernant, veuillez-vous adresser à DDFC – École des hautes études en santé publique, 15 avenue du Professeur-Léon-Bernard 35043 Rennes et/ou par email à fc-communication@ehesp.fr



CYCLE SPÉCIALISÉ

Traitement des données de santé

PYTHON POUR L'ANALYSE DES DONNÉES DE SANTÉ ET LA DATA SCIENCE

Les +

- Formation mise en place en **collaboration avec l'ATIH**
- Formateur externe **CNAMTS SNDS et ENSAE-ENSAI FORMATION CONTINUE**
- Nombre de participants restreint pour **favoriser les échanges** et permettre une **mise en pratique optimale**
- Les aspects techniques et informatiques seront abordés de **façon progressive** au cours de la formation
- **Conseils sur les bonnes pratiques et les pièges à éviter**

6 jours
(3 modules de 2 jours)
Modules ouverts à la
carte

Rennes



Octobre à
novembre 2024

CYCLE PYTHON POUR L'ANALYSE DES DONNÉES DE SANTÉ ET LA DATA SCIENCE

OBJECTIFS PROFESSIONNELS

- Connaître les fonctionnalités du langage Python et l'outil Anaconda
- Savoir effectuer des traitements sur les données de santé plus rapidement et plus efficacement via Python
- Acquérir les compétences sur la thématique spécifique : la data science et le machine learning avec Python
- Appliquer et paramétrer ces méthodes sur les données de santé



PUBLIC VISÉ

Toute personne souhaitant s'initier à Python/Anaconda ET/OU désirant connaître et exploiter les données PMSI via Python: médecin DIM, de santé publique, statisticien, data-manager, ingénieur en établissement de santé, ARS, centre d'investigation clinique, bureau d'étude, équipe de recherche



MODALITÉS PÉDAGOGIQUES

- Présentiel
- Apports théoriques
- Exercices pratiques

Nombre maximum de participants : 10



PRÉREQUIS

Connaissance du secteur de la santé



ÉVALUATION/ VALIDATION

Questionnaire de positionnement et de vérification des acquis



Responsable pédagogique

Dr Sahar BAYAT
Professeure
Département Méthodes quantitatives en santé publique, EHESP

Ce cycle peut également être suivi dans le cadre du **Diplôme d'établissement Science des données de santé** à retrouver sur :

formation-continue.ehesp.fr

Recherche par thème : Traitement des données de santé

PROGRAMME PRÉVISIONNEL

MODULE 1 : PYTHON POUR L'ANALYSE DES DONNÉES EN SANTÉ : INITIATION

Compétences visées

- Distinguer les objets Python
- Apprendre les fonctionnalités de base du langage Python
- Mettre en œuvre des fonctions de Python
- Manipuler un jeu de données en santé via pandas

Contenu de la formation

- Introduction Python (distributions, installations, interfaces shell, notebook, IDE)
- Présentation d'anaconda et jupyter lab
- Base du langage : instruction, variable, scripts
- Les objets de collection en python : dict, list, set, tuples

2 jours • 14 au 15 octobre 2024

MODULE 2 : PYTHON POUR L'ANALYSE DES DONNÉES EN SANTÉ : INTERMÉDIAIRE

Compétences visées

- Effectuer des traitements sur les données de santé, plus rapidement et plus efficacement
- Connaître les principaux packages d'analyse statistique de python

Contenu de la formation

- Data management avec pandas : agrégation, jointure, crosstab...
- Data visualisation avec matplotlib et seaborn

2 jours • 16 au 17 octobre 2024

- Instruction des contrôles : conditions, boucles, fonctions
- Présentation rapide de Programmation Orienté Object : Classe, attribut, méthode, module, package. Importation des modules/packages
- Calcul scientifique avec numpy : vecteur et matrice
- Data management avec pandas (DataFrame et Serie, import/Export, indexation, filtering, traitement des manquants)
- Data visualisation avec matplotlib
- Cas d'utilisation sur les données médico-administratives (PMSI, ScanSanté...)

- Utilisation des méthodes d'analyses statistiques classiques : T-tests + Anova, Chi2, PCA, régression logistique avec p-values
- Librairie scikit-learn pour l'apprentissage statistique (les modèles usuels de régression, régression logistique et arbres de décision...)
- Cas d'utilisation sur les données médico-administratives (PMSI, ScanSanté...)



MODULE 3 : PYTHON POUR L'ANALYSE DES DONNÉES EN SANTÉ – AVANCÉ : LA DATA SCIENCE ET LE MACHINE LEARNING

Compétences visées

- Connaître la librairie scikit-learn et les implémentations des algorithmes classiques de machine learning
- Être capable d'appliquer et de paramétrer ces méthodes sur les données de santé

Contenu de la formation

- Machine Learning : méthodologie générale
- Le pipeline de l'apprentissage supervisé
 - Construction du jeu de données et préparation des données
 - Comparaison des modèles supervisés de classification ou de régression (Gradient Boosting, Random Forest, Régressions, etc)

2 jours • 25 au 26 novembre 2024

- Paramétrage et entraînement du modèle
- Évaluation du modèle
- Diagnostic, gestion du sur-apprentissage ou du sous-apprentissage, Cross-Validation
- Introduction à l'apprentissage non-supervisé
- Cas d'utilisation sur les données médico-administratives (PMSI, ScanSanté...)

Programmes complets sur :
formation-continue.ehesp.fr