

## DURÉE / LIEU

### 21 jours

1 cycle de 5 modules de 2 à 4 jours

1 cycle de 3 modules de 2 jours

Présentiel : 126 heures

Rennes

## TARIF

5.000 € net

## MODALITÉS DE CANDIDATURE

- **Date limite de candidature : le 24/02/2023**
- **Un CV** est à joindre à votre demande d'inscription. En fonction de votre parcours professionnel, la responsable pédagogique pourra vous contacter pour un entretien.

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.

## PARTENARIAT



AGENCE TECHNIQUE  
DE L'INFORMATION  
SUR L'HOSPITALISATION

### Une question concernant votre candidature ?

Direction de la scolarité

[candidaturefc@ehesp.fr](mailto:candidaturefc@ehesp.fr)

Tél. : 02.99.02.25.00

### Qui contacter si vous êtes en situation de handicap ?

Référent handicap EHESP

[referent-handicap@ehesp.fr](mailto:referent-handicap@ehesp.fr)

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](http://formation-continue.ehesp.fr)



## Inscription et modalités sur notre site

PAGE [SCIENCE DES DONNEES DE SANTE](#)

OU EN RECHERCHANT LA FORMATION PAR THÈME

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

[formation-continue.ehesp.fr](http://formation-continue.ehesp.fr)



E H E S P

Formation  
continue

DIPLÔME D'ÉTABLISSEMENT (DE)

Traitement des données de santé : PMSI et SNDS

# SCIENCE DES DONNÉES DE SANTÉ

Nouveauté

## Les +

- Une formation en collaboration avec l'ATIH
- Un diplôme avec une double expertise données de santé et Data science
- La maîtrise de trois langages leader de la data science (Python, R et SQL)
- L'appropriation des algorithmes de Machine Learning



1 cycle de 5 modules  
de 2 à 4 jours

+

1 cycle de 3 modules  
de 2 jours

Modules ouverts à la carte



Rennes



21 jours  
AVRIL à  
NOVEMBRE 2023



## OBJECTIFS PROFESSIONNELS

- Connaître les 3 langages leader de data science et de data analyse : Python, R et SAS SQL
- Maîtriser l'architecture de la base PMSI MCO-SSR, ainsi que la plateforme de données hospitalières ATIH
- Connaître les fonctionnalités des outils Anaconda, Rstudio et SAS Enterprise Guide
- Elaborer une démarche d'étude du plan d'extraction au plan d'analyse
- Savoir implémenter des algorithmes classiques de machine learning
- Appliquer et paramétrer ces méthodes sur les données de santé



## PUBLIC VISÉ

Toute personne souhaitant exploiter les données issues du PMSI ou valoriser les données de santé en général via des langages et outils de la data science tels que Python, R, SAS SQL, SAS Enterprise Guide et la machine learning (médecin DIM, médecin de santé publique, statisticien, data-manager, ingénieur hospitalier, chercheur, personnels ARS, Assurance maladie, bureau d'étude)



## PRÉREQUIS

Connaissance du secteur de la santé



## PARCOURS PÉDAGOGIQUE

Ce diplôme est structuré autour de 8 modules permettant d'aborder les principales thématiques de données en santé et la méthodologie de data-science.

La formation peut être suivie sur 2 ans, à raison d'un cycle par an.

## PROGRAMME ET MODULES

### CYCLE TRAITEMENT DES DONNÉES PMSI

#### MODULE 1 - PLATEFORME DES DONNÉES HOSPITALIÈRES ET TECHNIQUES DE BASE DE REQUÊTAGE

##### Contenu de formation

- Présentation de la plateforme de l'ATIH contenant des bases PMSI
- Notions de base de données
- Architecture de la base PMSI MCO mise en place par l'ATIH
- Prise en main du logiciel SAS Enterprise Guide
- Exercices d'analyse de données à partir d'une base PMSI MCO

4 jours • Session 1 : 3 au 6 avril 2023  
Session 2 : 22 au 25 mai 2023

#### MODULE 3 - SAS PROGRAMMATION SQL ET TECHNIQUES DE REQUÊTAGE AVANCÉES

##### Contenu de formation

- Architecture des bases PMSI MCO-SSR mise en place par l'ATIH
- Programmation SQL SAS (join, Subqueries, SQL procedure, SET Operators, Advanced PROC SQL Features)
- Exercices et études de cas sur le chaînage des séjours à partir des données PMSI MCO-SSR via programmation SQL

2 jours • 1er au 2 juin 2023

#### MODULE 2 - SAS ENTERPRISE GUIDE ET TECHNIQUES DE REQUÊTAGE INTERMÉDIAIRES

##### Contenu de formation

- Architecture des bases PMSI MCO-SSR mise en place par l'ATIH
- Logiciel SAS Enterprise Guide
- Requête via SAS Guide
- Etude de cas 1 : Analyse de l'activité d'un GHT
- Etude de cas 2 : Parcours patient MCO-SSR
- Règles de bon usage

2 jours • 30 au 31 mai 2023

#### MODULE 4 - PLATEFORME DES DONNÉES ATIH VIA R

##### Contenu de formation

- De SAS à R depuis le portail ATIH
- Requêtes PMSI MCO-SSR de type SQL avec R : Tidiverse
- Requêtes simples de prise en main
- Requêtes de synthèse : les statistiques médicales par groupe (GHM, DA, GP, CAS...), diagnostic, acte, case-mix par établissement, taux de recours, part de marché, taux d'attractivité et fuite, statistiques des GHT
- Utilisation de nomenclatures (évolution GHM)
- Graphes d'indicateurs d'activités PMSI (ggplot)
- Etudes de cas : prise en charge d'infarctus du myocarde non compliqué ; Analyse de l'activité d'un GHT ; Flux patient MCO-SSR
- Règles de bon usage

3 jours • 11 au 13 septembre 2023

### MODULE 5 - UTILISATION DE LA DATASCIENCE POUR EXPLOITER LES DONNÉES PMSI

#### Contenu de formation

- Manipulation et préparation des données PMSI MCO-SSR avec R : dataFrame, data.table et tidyverse (niveau intermédiaire)
- Machine Learning : méthodologie générale
- L'apprentissage supervisé via workflow tidymodel R
  - ◊ Préparation des données
  - ◊ Implémentation et comparaison des modèles supervisés de classification ou de régression (Régression, Gradient Boosting, Random Forest, svm, etc)
  - ◊ Paramétrage et entraînement du modèle
  - ◊ Evaluation du modèle
  - ◊ Diagnostic, gestion du sur-apprentissage ou du sous-apprentissage, cross-validation
- Cas d'utilisation sur les données médico-administratives (PMSI, ScanSanté...)

Ce module vise à implémenter informatiquement les algorithmes d'apprentissage supervisé avec langage R, et comprendre le principe d'un projet de Machine Learning. Les théories fondamentales mathématiques des algorithmes ne sont pas abordées. Des références d'ouvrages mathématiques seront recommandées par la responsable pédagogique

3 jours • 22 au 24 novembre 2023



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

#### MODULE 1 - PYTHON : INITIATION

##### Contenu de formation

- Introduction à Python (distributions, installations, interfaces shell, notebook, IDE)
- Présentation d'Anaconda et Jupyter lab
- Base du langage : instruction, variable, scripts
- Les objets de collection de python : dict, list, set, tuple
- Instruction des contrôles : conditions, boucles, fonctions
- Présentation 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)
- Cas d'utilisation sur les données médico-administratives

2 jours • 2 au 3 octobre 2023

#### MODULE 2 - PYTHON : INTERMÉDIAIRE

##### Contenu de formation

- Pandas avancé : agrégation, recodage de variables, jointure, crosstab, apply, optimisation, datetime...
- Gestion des formats de données JSON et accès à des données en ligne via une API web
- Data visualisation avec matplotlib et seaborn
- Cas d'utilisation sur les données médico-administratives

2 jours • 4 au 5 octobre 2023

#### MODULE 3 - PYTHON : AVANCÉ - LA DATA SCIENCE ET LE MACHINE LEARNING

##### Contenu de formation

- Machine Learning : méthodologie générale
- Le pipeline de l'apprentissage supervisé via la librairie Scikit-Learn
  - ◊ Construction du jeu de données et préparation des données
  - ◊ Comparaison des modèles supervisés de classification ou de régression (Gradient Boosting, Random Forest, Regressions...)
  - ◊ Paramétrage et entraînement du modèle
  - ◊ Evaluation 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)

2 jours • 20 au 21 novembre 2023



#### Responsable pédagogique

**Fei GAO**

Professeure Département Méthodes quantitatives en santé publique, EHESP



#### MODALITÉS PÉDAGOGIQUES

Présentiel  
Apports théoriques  
Exercices pratiques  
Études de cas  
Mise en situation réelle à la plateforme sécurisée des données via des comptes de formation...

Nombre maximum de participants : 10



#### ÉVALUATION/ VALIDATION

Questionnaire de positionnement/ évaluation à chaque module

L'obtention du DE repose sur la réalisation d'un projet data-informatique complet



#### PERSPECTIVES

Health data expert, data manager / data analyst en santé, statistiques en santé

# Formulaire de candidature

## DIPLÔME D'ÉTABLISSEMENT

### SCIENCE DES DONNÉES DE SANTÉ (2023)



#### ○ Je souhaite suivre le diplôme en 2023 :

Je choisis la date du module 1 du cycle Traitement des données PMSI :

du 3 au 6 avril 2023 /  du 22 au 25 mai 2023

#### ○ Je souhaite suivre le diplôme sur 2 ans (2023/2024)

○ Cycle Traitement PMSI en 2023 + Cycle Python en 2024

Je choisis la date du module 1 du cycle Traitement des données PMSI :

du 3 au 6 avril 2023 /  du 22 au 25 mai 2023

○ Cycle Python en 2023 + Cycle Traitement PMSI en 2024

#### JE COMPLETE LE FORMULAIRE ET JE JOINS MON CV

Nom : ..... Nom de naissance : .....

Prénom : ..... Date de naissance : .....

Lieu de naissance : .....

E-mail (obligatoire) : .....

Fonction : .....

Corps de rattachement et profession : .....

Direction/Service : .....

Nom de l'établissement : .....

Statut juridique de l'établissement :  Public  Privé

N° de SIRET : ..... Code APE : .....

Adresse professionnelle : .....

Code postal : ..... Ville : .....

Téléphone : .....

Nom et prénom de votre responsable formation : .....

Adresse de votre responsable formation (si différente de l'employeur)

.....

Code postal : ..... Ville : .....

E-mail (obligatoire) : .....

Adresse de facturation (si différente de l'établissement) : .....

.....

Code postal : ..... Ville : .....

Financier (si différent de l'employeur) : .....

Adresse : .....

Code postal : ..... Ville : .....

J'accepte d'être informé(e) par e-mail des formations EHESP

Fait le ..... le ..... le .....

Signature du candidat ..... Visa de l'employeur ..... Visa du responsable Formation

#### Tarif :

5 000 € net

#### Pour vous inscrire,

scannez ou photocopiez ce bulletin puis retournez-le dûment complété par e-mail ou par courrier à :

[candidatureFC@ehesp.fr](mailto:candidatureFC@ehesp.fr) / Tél 02.99.02.25.00

École des hautes études en santé publique

Direction de la scolarité - Formation continue  
15 Av. du Professeur-Léon-Bernard  
CS 74312 - 35043 RENNES Cedex

#### 1 fiche par stage et par personne

L'EHESP 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](mailto:referent-handicap@ehesp.fr)

#### Modalités financières :

Dans le cas d'un règlement par chèque, merci de libeller et d'adresser indépendamment votre chèque à l'ordre de l'agent comptable de l'EHESP, accompagné de la copie de la convention de formation.

Si le client peut bénéficier d'un financement par un tiers, il lui appartient de faire la demande de prise en charge avant le début de la formation et l'accord de financement doit être fourni avant la date de la formation. Le client doit s'assurer du suivi de cette demande et de la réalisation du paiement par le financeur.

La signature de ce bulletin vaut acceptation sans réserve des conditions générales de vente accessibles sur notre site internet ou notre catalogue en vigueur.

La signature de ce bulletin vaut acceptation sans réserve des conditions générales de vente accessibles sur notre site internet ou notre catalogue en vigueur.

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](mailto:fc-communication@ehesp.fr)

Retrouvez toute l'offre de formation continue EHESP sur notre site [formation-continue.ehesp.fr](http://formation-continue.ehesp.fr)