

N° DU CONCOURS : **2022-03**

GRADE : **Ingénieur d'études de classe normale (catégorie A)**

BAP : **E - Informatique, Statistiques et Calcul scientifique**

FAP : **Statistiques**

EMPLOI-TYPE : **E2D46 - Ingénieur-e statisticien-ne**

NATURE : **Ingénieur-e data**

AFFECTATION : **Service des méthodes statistiques de l'Ined**

NOMBRE DE POSTE A POURVOIR : **1**

NIVEAU DE DIPLÔME **MINIMUM** REQUIS : **Titre ou diplôme classé au moins niveau 6 (Bac+3)**

Descriptif de la structure

L'Institut national d'études démographiques (Ined) est un établissement public à caractère scientifique et technologique (EPST). Il a pour mission de mener des recherches sur les différentes problématiques démographiques (fécondité et famille, santé et mortalité, migrations...) et d'informer la communauté scientifique et le public des évolutions en cours. Fort de ses 11 unités de recherche, l'Ined est au cœur d'un réseau de partenariats nationaux et internationaux de recherche.

Le service Méthodes Statistiques est composé de 9 personnes, et a pour missions :

- Le conseil auprès des responsables de projets de l'Ined pour les guider dans le choix des méthodes et outils statistiques adaptés à leurs problématiques ainsi que la valorisation des résultats
- La formation des chercheurs de l'institut aux outils et méthodes statistiques utiles à leurs projets, en particulier dans l'exploitation des données
- L'accompagnement des équipes de projets de l'Ined dans la structuration de leurs données d'analyse
- L'animation de séminaires scientifiques axés autour de la pratique statistique
- Participation, en lien avec les autres services, à la politique de gestion et de traitement des données de recherche à l'Ined
- La veille sur les méthodes statistiques et leurs applications

Missions principales du poste

L'ingénieur-e data aura pour mission principale de conseiller et d'organiser la structuration, l'exploration et la valorisation de la Donnée, en lien avec les équipes métiers et unités de recherche.

Son activité sera découpée en 3 pôles complémentaires :

1. Préparation de la donnée (intégration, structuration, analyse de la qualité, ...),
2. Exploration de la donnée (production automatisée de codebooks, visualisations exploratoires),
3. Valorisation de la donnée (publication, production de tableaux synthétiques, graphiques

interactifs).

L'ingénieur·e data accompagnera les équipes de recherche Ined, dans la mise en œuvre de développements et de conseils dans ces trois pôles d'activité précisés ci-dessus.

La méthodologie de prise en charge sera basée sur un processus en quatre étapes :

1. Une étape initiale de proposition de solutions adaptées, basée sur la compréhension de la problématique de recherche, sa traduction en termes analytiques.
2. Une veille sur les outils disponibles, la mise en œuvre de tests, et la proposition de solutions et procédures adaptés à cette ingénierie de la donnée.
3. Le développement de tout ou partie de cette procédure.
4. Le transfert de connaissances des solutions proposées.

Cette activité de support et de développement sur projet constituera la majeure partie de son activité et sera complétée par une veille, une valorisation et un transfert de connaissances à l'ensemble du personnel des techniques mises en œuvre par le biais de formations, rédaction de notes, alimentation de pages Web d'information, avec un objectif de diffusion d'une culture de la donnée.

Dans un contexte de science ouverte, il veillera à produire une traçabilité dans le cadre des chaînes de traitement dont il assurera le développement. Il travaillera en parallèle sur plusieurs projets, ce qui nécessitera rigueur et organisation.

Son activité sera tournée à la fois non seulement vers les unités de recherche mais aussi vers l'ensemble des services d'appui dans la préparation, mise en forme, et valorisation des données, internes et externes. Il pourra être amené à collaborer plus particulièrement avec le service des éditions de l'Ined, la direction de la communication et le DataLab de l'Ined.

Activités d'un·e ingénieur·e data

Activité de structuration de la donnée :

- Garantir et assurer une veille sur la qualité des données (détection des valeurs aberrantes, méthodologies de nettoyage, imputations, gestion des anomalies, ...). Assurer le transfert de compétences pour la mise en œuvre du processus de validation des données.
- Structurer et organiser l'automatisation de règles de gestion et de validité des données (typages autorisés, normalisation des valeurs selon le type de données, traitement des valeurs manquantes) en regard des différents types de traitements statistiques à effectuer.
- Assurer un support pour la création de variables (règles de développement et de programmation structurée, aide à la mise en œuvre) nécessaires pour les procédures d'analyses exploratoires et de modélisation statistique.
- Maîtriser les techniques d'extraction de données à partir de divers systèmes de gestion de bases.
- Maîtriser les techniques d'appariement de données, et l'optimisation des espaces de stockage
- Assurer une veille sur les sources d'appariement pour améliorer la qualité des données
- Assurer une veille sur les techniques d'extraction et de structuration en fonction du type de données (structurées, réseaux, textuelles, géographiques, données issues du Web, ...)

Activité d'exploration de la donnée

- Maîtriser les techniques d'exploration systématique de données (techniques de production de codebooks automatisés, processus de regroupements de valeurs, techniques de corrélations, ...)
- Maîtriser les techniques de datavisualization en fonction du type de données. Connaître et valoriser l'utilisation des principaux packages R adaptés à cette thématique (activité de tests, veille, support)

Activité de valorisation de la donnée

- Proposer les outils adaptés (graphiques, tableaux de bord) en fonction de la demande, du public, et du degré d'interactivité souhaité. Connaître au moins un outil de développement d'application. Activité de développement d'applicatifs.
- Organiser le développement d'outils d'aide à la décision analysant des données complexes.

Il participera aux activités collectives du service en termes de valorisation, formations, rédaction de notes techniques, alimentation de sites Web dédiés.

Profil du.de la candidat.e

Formation initiale :

- Statistiques, ingénierie informatique/mathématique ou data science

Compétences :

- Connaissances solides en R
- Bonne maîtrise du logiciel SAS® Foundation (Programmation structurée, procédures de statistique exploratoire)
- La connaissance du langage Python serait un plus
- Connaissance d'un outil de développement d'applications (de type Shiny)
- Maîtrise des techniques de publication sur le Web (markdown)
- Maîtrise des techniques de traçabilité des traitements (Connaissance des systèmes GiT)
- Une culture des principales analyses statistiques utilisées en sciences sociales est souhaitée.

Qualités recherchées :

Savoir-faire :

- Etre force de proposition pour la mise en œuvre de solutions techniques sur un projet, savoir élaborer un cahier des charges technique
- Assurer une veille, tant au niveau des méthodologies que des technologies à utiliser, dans la valorisation de la donnée.
- Savoir mettre au point ou adapter des techniques nouvelles
- Accompagner et conseiller. Savoir communiquer et faire preuve de pédagogie
- Rédiger des rapports ou des documents
- Développer un sens critique
- Développer des capacités de conceptualisation
- Gérer la sécurité de l'information

Savoir être :

- Capacité à interagir avec différents personnels
- Bon relationnel et goût pour le travail en équipe
- Disponibilité et réactivité
- Autonomie et esprit d'initiative
- Maîtrise de l'anglais technique et scientifique du domaine, notamment à l'écrit.
- Capacités relationnelles et rédactionnelles.
- Organisation et rigueur : une attention particulière sera donnée à la capacité de la personne recrutée à travailler en équipe, et à son aptitude à travailler en parallèle sur plusieurs projets de recherche.

Contexte du Poste:

- Localisation du poste : Siège de l'Ined, 9 cours des humanités, 93300 Aubervilliers (Site du Campus Condorcet)