



# Knight Peterson JEAN-BAPTISTE

À la recherche d'un poste de Data Scientist / Data Analyst Junior à partir d'octobre 2026, dans les domaines de la statistique, de l'IA et de la data science

## CONTACT

94 avenue de la République 63000 Clermont-Ferrand  
+33 06 28 67 93 66  
knightpetersonjeanbaptiste682@gmail.com

[www.linkedin.com/in/knight-peterson-jean-baptiste-78b96a1ba](https://www.linkedin.com/in/knight-peterson-jean-baptiste-78b96a1ba)

## PERMIS B

26 ans

## COMPÉTENCES EN INFORMATIQUE

- Python** : pandas, Numpy, scikit-learn, Pytorch, keras, tensorflow
- R** : survival, tidyverse, caret, ggplot2, kableExtra
- SAS** : statistiques, data management
- SQL** : requêtes, extraction et structuration de données
- Excel , VBA** : analyse, reporting, automatisations
- Word, Powerpoint** : documentation, présentations
- Power Apps** : développement d'applications métiers
- Power BI** : tableaux de bord, visualisation

## QUALITÉS

- Rigueur analytique (validation statistique, AIC/BIC)
- Créativité (prototypage de modèles, visualisations)
- Ouverture d'esprit (travail interdisciplinaire, recherche appliquée)

## LANGUES

- Français** : seconde langue
- Anglais** : professionnel, bonne aisance à l'écrit et à l'oral.
- Espagnol** : courant
- Créole** : langue maternelle

## CENTRES D'INTÉRÊT

- Sports collectifs : football, volleyball
- IA générative et data science
- Veille technologique

## FORMATION

- 2024/26 **Master 2 mathématiques appliquées, statistiques**  
{parcours statistique et traitement de données}  
Université Clermont Auvergne
- 2020/24 **Licence mathématiques et informatique appliquées aux sciences humaines et sociales**  
Université de Strasbourg

## PROJETS

- 2026 **Modélisation de la survie (R)**
- 3 mois **Projet : Prédiction de la survie de patients atteints de carcinome corticosurrénalien (données TCGA)**
- Modèles de survie (Kaplan-Meier, cox)
  - Approches ML/DL : Arbres de décision, Random Forest, Deepsurv
  - Gestion de la censure, validation croisée, évaluation des performance
  - Tests de conformité des modèles et critères d'information (AIC, BIC)
  - Évaluation des performances et sélection du meilleur modèle
- 2025 **Régression et modélisation (Python)**
- 3 mois **Projet : Prédiction des émissions de CO2 d'un véhicule**
- Prétraitement des données, Encodage des variables catégorielles
  - Régression : Ridge, Lasso, VIF, OLS
  - ACP, AFC, ACM, tests statistiques, Diagnostic de modèles

## EXPÉRIENCE

- 2026 **Stagiaire Data Scientist - Modèles de diffusion (IA générative)**
- 5 mois **Université Clermont Auvergne (UCA) / MICHELIN**
- Analyse théorique des modèles de diffusion et du score-based modeling
  - Prétraitement et exploration de données industrielles Michelin.
  - Implémentation de modèles génératifs en Python (PyTorch, keras, tensorflow)
  - Étude de la convergence, évaluation des performances et validation.
  - Documentation et restitution des résultats (rapports, présentations).
- 2025 **Stagiaire aux méthodes industrielles avec Power BI et Power Apps**
- 3 mois **MICHELIN -Site Ladoux, Cébazat**
- Projet 1: Création d'un outil digital constitué de deux applications Power Apps pour enregistrer les tests de débits avec des visuels **Power Bi** intégrés pour suivre l'évolution des performances en temps réel .
- Projet 2 : Mise en place d'un processus automatisé en Power Query (M) pour l'importation de fichiers Excel stockés dans un dossier SharePoint sécurisé. Création de requêtes et colonnes personnalisées pour le traitement des données suivi de leur intégration dans un rapport Power Bi interactif avec des visuels dynamiques facilitant l'analyse.