

Paul LO

Alternance - Data Analyst

✉ hsiaowenpaul@gmail.com
🏠 PARIS
🔗 drakoriz.github.io
📍 75013
île-de-France
☎ +33 767694454

Mes réalisations

Application web de comparaison de villes développée sous Streamlit dans un délai court – <https://projetville.streamlit.app>

Langues

Anglais
B2 Anglais professionnel

Certificat

Cambridge English Certificate (CEC)
LEVEL B2, Sept 2022, Paris

Centres d'intérêt

Art Digital sur Procreate

Réseaux sociaux

in Paul LO

Je suis étudiant en premier cycle d'ingénieur à l'EFREI, spécialisé en Big Data et Machine Learning, et je recherche un contrat d'alternance sur un rythme de deux jours en formation et trois jours en entreprise.

Diplômes et Formations

Cycle Ingénieur - Big Data & Machine Learning

De septembre 2025 à août 2028 EFREI - Villejuif

Formation d'ingénieur spécialisée dans l'exploitation des données massives et le déploiement de solutions d'IA. Compétences en analyse de données, pipelines Big Data, optimisation des traitements, cloud computing, et intégration de solutions aux systèmes existants.

BUT - Science des Données

De septembre 2022 à juin 2025 IUT de Paris - Rives de Seine - Paris

Visualisation, analyse, traitement des données et modélisation statistique

Baccalauréat Général

2022 Lycée Pierre-Gilles de Gennes - ENCPB - Paris

Enseignements de spécialités : Mathématiques et Physique-Chimie

Options : Mathématiques expertes et Section européenne en anglais.

Expérience professionnelle

Alternant Data Analyst - Pôle Data, Direction Digital & Data

De septembre 2024 à septembre 2025 RATP Dev Paris

Participation à des projets variés (DataLake, maintenance digitale/prédictive, exploitation des données de transport). Analyse et valorisation des données pour optimiser les performances opérationnelles, les flux passagers et la segmentation des voyageurs.

Compétences

Data - Base de données

Requêtage, gestion et structuration des données
Python, SQL, SAS, R, MongoDB

Analyse & Apprentissage Automatique

Classification, Séries temporelles, Régression linéaire, Analyse factorielle
Estimation, Statistiques inférentielles, Arbres de régression CART

Visualisation & Outils Data

Shiny, Power BI, Tableau, Excel, Dataiku

Développement & Cloud

Django, Streamlit, Web Scraping, HTML, CSS, Outils Cloud (AWS S3)

Autres librairies

Pandas, NumPy, Scikit-learn, TensorFlow, Keras, Plotly, Seaborn, Matplotlib

Projet professionnels et universitaires

Alternance - Dashboard d'une filiale basée à Al'Ula, Arabie saoudite (2024-2025)

Développement d'un **dashboard interactif** Streamlit pour la visualisation des **KPI** du **réseau de transports** d'Al'Ula à partir d'API, avec stockage sur **AWS S3**. **Méthode** : Agile
Résultat : le dashboard a été adopté par la filiale, suscitant une demande d'évolutions fonctionnelles supplémentaires et son extension à d'autres cas d'usage.

Alternance - Etablissement des principes de l'éco-conduite (2024-2025)

Analyse exploratoire et **prétraitement** des données IoT de bus
Nettoyage des données, gestion des valeurs manquantes et calcul de la consommation moyenne sur fenêtres glissantes de 20 km pour lisser le bruit.
Exploration des corrélations entre vitesse et consommation, **visualisation** via notebooks Python et premiers graphiques sur Superset à partir de **requêtes SQL/Snowflake**.
Résultat : Identification de **tendances générales** et **préparation des données** pour analyses plus fines, avec transmission des insights pour poursuite par un **expert dédié**.

Université – Ciblage rachat sur données (2024-2025)

Projet de prédiction des clients susceptibles de réaliser un rachat : exploration des **méthodes de scoring** (aléatoires, ciblées, profilées), mise en œuvre de modèles avancés, et création d'un dashboard de synthèse sous **Power BI**.
Résultat : meilleur score de la promotion obtenu grâce à l'utilisation d'un modèle Random Forest associé à un feature engineering approfondi.