

42 Avenue des genottes , 95800 Cergy | 0767034955 | <https://chihab-mezrigui.onrender.com/> | <https://www.linkedin.com>

## Profil

Ingénieur Data passionné, spécialisé dans les pipelines Big Data (Kafka, Spark, Hadoop). Expérience sur GCP, Data Warehouse et Data Lakehouse. Maîtrise du traitement et de la transformation de données à grande échelle, avec une forte appétence pour la visualisation (Power BI, ReactJS). Curieux et rigoureux, prêt à relever de nouveaux défis.

## Formation

MASTER BIG DATA & IA | 2025 | IPSSI PARIS

LICENCE GÉNIE LOGICIEL | 2023 | UNIVERSITE DE CARTHAGE - TUNIS

## Expérience Professionnelle

### INGENIEUR DATA - ALTERNANCE | RDIUP | DEPUIS SEPTEMBRE 2023

**Tech Stack:** Python, JavaScript, Kafka, PySpark, Hadoop, Pandas, CassandraDB, MongoDB, Postgresql , Docker, GitLab CI/CD , Swagger

#### Missions principales :

- Traitement temps réel avec Kafka
- Création de pipelines ETL pour données CSV
- Visualisation via dashboards et interfaces ReactJS
- Analyse exploratoire et extraction d'insights
- Conteneurisation Docker et automatisation CI/CD

### DEVELOPPEUR FULLSTACK – STAGE PFE | RDIUP | APR 2023 – AOUT 2023

**Tech Stack :** React Native, Python (Django), Docker , MongoDB, GitLab, Swagger

#### Missions principales :

- Développement d'une appli mobile en React Native.
- Intégration d'API externes, notamment Mapbox pour la géolocalisation.
- Création et maintenance d'API REST avec Django.
- Gestion du versioning et déploiement automatisé via GitLab.

### DEVELOPPEUR FULLSTACK - STAGE | CNI | MAR 2022 – MAI 2022

**Stack:** Typescript (Angular), Java (Spring boot),Postgresql,Github,Postman

#### Missions principales :

- Développement et maintenance d'API RESTful avec Spring Boot pour le back-end
- Utilisation de Swagger pour la documentation et les tests des API
- Réalisation de tests unitaires et d'intégration afin de garantir une haute qualité du code
- Gestion du contrôle de version et de l'intégration continue avec Git et GitHub

## Projets Personnels :

### Prédiction aux JO 2024 à Paris.

**Tech Stack :** Node.js, Express, React, Ant Design, PostgreSQL, Scikit-learn, Pandas, Docker .

- Développement backend (Node.js/Express) & frontend (React/Ant Design)
- Création d'APIs REST pour les prédictions
- Entraînement de modèles ML via Jupyter Notebooks
- Analyse de données (athlètes, pays, médailles)
- Déploiement d'une interface web interactive
- Conteneurisation avec Docker

[🔗 Démo](#) | [Code GitHub](#)

### Prédiction d'Assurance Médicale

**Tech Stack :** Python, Pandas, Scikit-learn, Matplotlib, Streamlit

- Création d'un modèle de régression linéaire pour estimer les frais médicaux.
- Analyse des données (EDA) à partir d'un dataset Kaggle.
- Déploiement d'une application web interactive avec Streamlit.
- Score obtenu :  $R^2 = 0.78$ ,  $MSE = 0.23$

[🔗 Démo](#) | [Code GitHub](#)

### Intégration et Analyse Google Sheets via API Google Cloud

**Tech Stack :** Python, Google Sheets API, Pandas, NLTK, Matplotlib, Wordcloud, Google Cloud Platform

- Projet d'analyse de contenu Google Sheets via l'API Google Cloud.
- Extraction et traitement des données Sheets
- Analyse de texte avec n-grammes (fréquence des mots)
- Visualisation : graphiques et nuages de mots

[🔗 Démo](#) | [Code GitHub](#)

## Compétences et capacités

### Exploration & Valorisation des données :

Pandas, NumPy, Matplotlib, Seaborn, Sickit-Learn ,Jupyter Notebook, Power BI

### Big Data & Traitement de données :

Apache Kafka, Apache Spark, Pyspark, Hadoop (HDFS, MapReduce, Pig), ETL, Apache Parquet , Google Cloud Platform (BigQuery, Cloud Storage), Data Warehouse & Data Lakehouse .

### Développement & Programmation

Python(Django,Flask) , JavaScript (ReactJS, React Native), Django, SQL(PostgreSQL), NOSQL (Cassandra, MongoDB)

### Cloud & DevOps

Docker, Google Cloud platform, GitLab CI/CD, Swagger,Postman, Mapbox, API REST

### Méthodologies & Gestion de Projet :

Scrum, Kanban

## Langues

Français , Anglais, Allemand

