

KAVYA BASAPPA

Data Analyst | Data Engineer | MSc Data Engineering & Cloud Computing

Disponible immédiatement

kavyamona777@gmail.com

linkedin.com/in/kavya-basappa-670b08215 | github.com/Kavyamona7 | kavyamona7.github.io

PROFIL

Profil data (MSc Data Engineering & Cloud Computing en cours) avec une expérience pratique de la conception de pipelines de données, du contrôle de la qualité des données et de la création de tableaux de bord BI. Compétences en Python, SQL, Power BI, Spark, Kafka et Azure : transformation de données brutes multi-sources en jeux de données prêts à analyser et en tableaux de bord d'aide à la décision. Disponible immédiatement pour des postes de Data Analyst, BI ou Data Engineer junior.

COMPÉTENCES

Analytique & BI: SQL, Power BI, Excel avancé, tableaux de bord KPI, Streamlit, visualisation de données

Data Engineering: Python, ETL/ELT, dbt, Apache Spark, Kafka, validation et contrôles qualité des données

Cloud & Outils: Microsoft Azure, Docker, Git, BeautifulSoup, Tesseract OCR, Flask

Langues: Anglais ; Français (A1)

EXPÉRIENCE PROFESSIONNELLE

AI Solutions Developer

2026

Aivancity AI Clinic, partenaire industriel Craft AI | Paris

- Conception d'un pipeline de transformation de données en 7 étapes, convertissant des données brutes multi-formats (PDF, images, pages web) en jeux de données propres et prêts à analyser.
- Mise en place de contrôles de validation et de qualité garantissant l'exhaustivité, l'exactitude et la cohérence de l'ensemble des résultats traités.
- Livraison d'un tableau de bord Streamlit permettant aux parties prenantes d'interroger, d'explorer et d'extraire des insights des données traitées en langage naturel.
- Mise en œuvre d'un stockage horodaté des résultats avec piste d'audit complète, assurant la reproductibilité des analyses et la traçabilité des exécutions.

Stack : Python, SQL, Streamlit, Google Gemini, BeautifulSoup, Tesseract OCR, Docker

Stagiaire AI Product Developer

2025

Aivancity AI Clinic | Paris

- Développement d'un système automatisé de notation et de reporting basé sur un LLM pour l'évaluation de code Python, réduisant de 85 % le temps de relecture manuelle.
- Conception de l'architecture des flux de données, garantissant 100 % d'intégrité des données sur l'ensemble des enregistrements des défis pédagogiques.
- Intégration d'un microservice Deep Evaluation produisant des métriques de qualité quantifiées sur 60 % de dimensions d'évaluation en plus.

Stack : Python, SQL, Flask, LLM, DeepEval, Streamlit, React

Data Quality Analyst

2024

Evolve BPM Pvt Ltd | Bengaluru, Inde

PROJETS

Tableau de bord d'analyse crypto en temps réel

2025

- Pipeline complet de la source à l'insight : données de marché BTC/ETH/SOL en direct diffusées via Kafka, agrégées par fenêtres d'une minute avec Spark et visualisées dans un tableau de bord KPI Streamlit interactif.

Tableau de bord d'évaluation GenAI

2025

- Tableau de bord multi-métriques comparant la qualité de modèles d'IA (pertinence, cohérence, exactitude factuelle) pour étayer le choix du modèle.

FORMATION

MSc Data Engineering & Cloud Computing

En cours

Aivancity School for Technology, Business & Society | Paris

Bachelor of Computer Applications (BCA)

2021

St. Joseph's College (Autonomous) | Bengaluru, Inde