

Anis Guechtouli

guechtoulianiss7@gmail.com | +213 555648000 | Mahelma, Alger, Algérie | in/anis-guechtouli/ | /GuechtouliAnis

PROFIL

Data Scientist avec une expérience concrète en **automatisation (IA, NLP)**, **analyse visuelle** et **optimisation de données**. Capable de développer des solutions de bout en bout : collecte, traitement, modélisation et interface utilisateur (**Streamlit**, **PowerBI**). À l'aise pour transformer des besoins métier en produits data opérationnels.

ÉDUCATION

Master Analyse et Sciences de Données 09/2023 – 06/2025
Université d'Alger 1 - Benyoucef Benkhedda

Licence Systèmes Informatiques 11/2020 – 07/2023
Université d'Alger 1 - Benyoucef Benkhedda

EXPÉRIENCE

Business Analyst - SIG Service 06/2025 – Actuel | Alger, Algérie

- Contribution à un projet de migration de données via **PostgreSQL** et **DBeaver**.
- Analyse de schémas existants et exécution de requêtes **SQL** pour assurer l'intégrité des données.
- Collaboration avec l'équipe Data pour fiabiliser le processus de migration.

Stagiaire Data Scientist – Djezzy 11/2024 – 06/2025 | Alger, Algérie

- Développé une solution d'IA conversationnelle basée sur le **NLP** et les **LLM** pour automatiser des tâches répétitives de transmission d'informations.
- Traitement et structuration** des données au format Q/R pour l'entraînement et l'évaluation du modèle.
- Appliqué des techniques de **prompt engineering** pour améliorer la cohérence et la pertinence des réponses générées.

Stagiaire Analyste de Données – CAAR 02/2023 – 07/2023 | Alger, Algérie

- Conçu un **tableau de bord interactif avec Streamlit** pour analyser les indicateurs de performance de l'agence automobile.
- Exploité des données issues d'une **base Oracle** via requêtes **SQL** pour générer des insights fiables.
- Collaboré avec le VP de la branche automobile et l'équipe IT pour orienter les décisions à partir d'analyses visuelles.
- Identifié des écarts significatifs sur des **KPI** grâce à la visualisation ciblée.

COMPÉTENCES

- Langues** : Arabe (langue maternelle), Anglais (avancé), Français (professionnel)
- Programmation** : SQL, Git, Python (bibliothèques principales : PySpark, Pandas, Numpy, Matplotlib, Plotly, Scikit-learn, LangChain, Streamlit)
- Gestion des données** : Spark, SQLite, Oracle, PostgreSQL
- Visualisation et reporting** : Excel, PowerBI, Tableau
- Communication et présentation** : Canva, PowerPoint, Markdown

PROJETS

Modèle Expected Goals (xG) 01/2025 – Actuel

- Collecte de données gratuites StatsBomb, création d'un dataset de **12 millions de lignes**.
- Prétraitement et nettoyage avec **PySpark** pour garantir qualité et cohérence.
- Ingénierie de **7 caractéristiques** clés : distance, angle, pied préféré...etc
- Modèle de **régression logistique** atteignant **~98 % matching rate** avec le modèle de StatsBomb et **R² de 0,83**.

Analyse Exploratoire des Données dans le Commerce de Détail 07/2024 – 07/2024

- Réalisé une **analyse exploratoire** sur un dataset retail de **9 994 ventes**, identifiant des problèmes clés comme des **produits non rentables** et des **remises excessives** dans certains états.
- Créé des **visualisations**, un **tableau de bord** et un **rapport détaillé** pour communiquer clairement les résultats et orienter les **décisions stratégiques**.

Application d'Optimisation – Sac à Dos (Knapsack) 03/2024 – 05/2024

- Modélisation d'un problème d'optimisation sous contrainte via **programmation mixte entière** (GEKKO).
- Développement d'une **interface Streamlit interactive** permettant aux utilisateurs de tester différents scénarios d'allocation.
- Mise en avant de la logique décisionnelle et de la visualisation dans des situations à ressources limitées.

Système de Reconnaissance de l'Iris 01/2024 – 05/2024

- Conception d'un système biométrique basé sur **SIFT + OpenCV** pour la reconnaissance par iris.
- Pipeline complet : prétraitement, extraction de caractéristiques, matching visuel.
- Déploiement d'une interface Streamlit** permettant de charger une image et identifier l'individu.