# Anis Guechtouli

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

## PROFILE

Data Scientist with hands-on experience in automation (**AI, NLP**), visual analysis, and data optimization. Skilled in developing end-to-end solutions: data collection, processing, modeling, and user interfaces (**Streamlit**, **PowerBI**). Comfortable translating business needs into operational data products.

## EDUCATION

**Master's Degree in Data Science and Analytics**                                                  09/2023 – 06/2025
*University Of Algiers 1 - Benyoucef Benkhedda*  ✐

**Bachelor's Degree in Computer Science**                                                          11/2020 – 07/2023
*University Of Algiers 1 - Benyoucef Benkhedda*  ✐

## EXPERIENCE

**Business Analyst – SIG Service**  ✐                                          06/2025 – present  |  Algiers, Algeria
- Contributed to a data migration project using **PostgreSQL** and **DBeaver**.
- Analyzed existing schemas and executed **SQL** queries to ensure data integrity.
- Collaborated with the Data team to secure and streamline the migration process.

**Data Scientist Intern – Djezzy**  ✐                                          11/2024 – 06/2025  |  Algiers, Algeria
- Developed an **NLP-based** conversational AI solution using **Large Language Models** to automate repetitive information delivery tasks.
- **Processed** and **structured** Q&A data to support model training and evaluation.
- Applied **prompt engineering** techniques to improve the consistency and relevance of generated answers.

**Data Analyst Intern – CAAR**  ✐                                             02/2023 – 07/2023  |  Algiers, Algeria
- Designed an interactive dashboard with **Streamlit** to analyze automotive branch performance indicators.
- Extracted data from an **Oracle database** using **SQL** queries to generate reliable insights.
- Collaborated with the Automotive VP and IT team to support decision-making through visual analytics.
- Identified significant KPI discrepancies through targeted visualizations.

## SKILLS

- **Languages** : Arabic (Native speaker), English (Advanced proficiency), French (Intermediate proficiency)
- **Programming** : SQL, Git, Python (key libraries: PySpark, Pandas, Numpy, Matplotlib, Plotly, Scikit-learn, LangChain, Streamlit)
- **Data Managment :** Spark, SQLite, Oracle, PostgreSQL
- **Dashboarding & BI Tools :** Excel, PowerBI, Tableau
- **Communication & Presentation :** Canva, PowerPoint, Markdown

## PROJECTS

**Expected Goals (xG) Model**                                                                      01/2025 – present
- Collected free StatsBomb data, creating a dataset of **12 million rows**.
- Preprocessed and cleaned the data using **PySpark** to ensure quality and consistency.
- Engineered **7 key features**: distance, angle, preferred foot, etc.
- Built a **logistic regression** model achieving **~98% matching rate** with the official StatsBomb model and an **R²** of **0.83**.

**Exploratory Data Analysis in Retail**  ✐                                                          07/2024 – 07/2024
- Performed **exploratory analysis** on a retail dataset of **9,994 sales**, identifying key issues such as **unprofitable products** and **excessive discounts** in specific states.
- Built **visualizations**, a **dashboard**, and a detailed **report** to communicate findings and guide strategic decisions.

**Optimization Application – Knapsack Problem**  ✐                                                  03/2024 – 05/2024
- Modeled a constrained optimization problem using **mixed-integer programming** (GEKKO).
- Developed an interactive **Streamlit** interface to let users test different allocation scenarios.
- Highlighted decision logic and visualization in resource-constrained settings.

**Iris Recognition System**                                                                        01/2024 – 05/2024
- Designed a biometric system for iris recognition using **SIFT** and **OpenCV**.
- Built a complete pipeline: preprocessing, feature extraction, visual matching.
- Deployed a **Streamlit** interface for image upload and individual identification.