



# Aymane MAGHOUTI

## DATA ENGINEERING STUDENT

✉ aymanemaghouti16@gmail.com  
🏠 Aknoul, Taza, Morocco  
📅 Date of birth 05/07/2002  
📞 +212 656 155867  
🔗 <https://aymane-maghouti.github.io/>  
💻 Open to remote work

### PORTFOLIO

🌐 @aymane-maghouti  
🐙 @aymane-maghouti

### LANGUAGES

**Arabic**  
Native  
  
**French**  
intermediate  
  
**English**  
intermediate

### TECHNICAL SKILLS

**Programming languages**  
Java  
Python  
HTML/CSS , JS (Web programming)  
SQL - PL/SQL  
shell (Basic notion)  
  
**Database**  
MYSQL  
Oracle (sql developer)  
Microsoft SQL Server  
  
**Operating systems**  
Linux (Ubuntu)  
Windows  
  
**Data Warehouse**  
Google BigQuery  
Snowflake

### INTERESTED

Learning new tech  
Sport (football , basketball ...)

I am currently pursuing my studies in data engineering at the national school of applied sciences al Hoceima and I am actively looking for an end-of-year internship opportunity in the field of computer science / data engineering.

### EDUCATION

- **Data Engineering**  
Since September 2022  
National School of Applied Sciences of Al Hoceima Al Hoceima
- **Preparatory cycle**  
From September 2020 to June 2022  
National School of Applied Sciences of Al Hoceima Al Hoceima
- **baccalaureate of Science in Physics and Chemistry**  
From September 2018 to June 2019 high school 2 octobre 1955 Aknoul

### LATEST PROJECTS

#### Real-time data pipeline with kafka

- **Python** : the *psutil* library for collecting the data
- **Apache Kafka** : For data streaming.
- **SQL server Database** : the data destination.
- **Power BI** : For creating the dashboard.

#### Mobile Data Analysis

- **MySQL** : data source.
- **Apache Sqoop** : transferring data between MySQL and Hive .
- **Apache Hive** : the datawarehouse solution.
- **Power BI** : For creating the dashboard.

#### Human Resources data pipeline

- **Azure data Factory** : for building data pipeline
- **Azure Blob Storage** : the data storage solution.
- **Azure Databricks (PySpark)** : for data transformation.
- **Power BI** : for data visualization.

#### Sales data pipeline

- **Python** : For data extraction and transformation.
- **Apache Airflow** : To orchestrate the ETL pipeline.
- **Docker** : To run the Apache Airflow instance in a container.
- **Google Cloud Platform (GCP)** : For BigQuery data warehousing.
- **Looker Studio** : For creating the dashboard.

#### Human Resources data pipeline

- **Oracle Database & PL/SQL** : as a data source and for data cleaning
- **Informatica** : for data integration
- **Snowflake Data Warehouse** : for store the transformed data
- **Power BI** : for creating the dashboard.

#### YouTube data pipeline

- **Python & Pandas** : Used for data extraction & transformation.
- **Snowflake Data Warehouse** : Used as the central repository for storing.
- **Power BI** : used for creating interactive dashboard.