

# HOSSAM METWALLY

## Data Engineer and Governance Intern

+20(106) 849 1535

Hossammetwally1@outlook.com

Egypt - Giza, 6th of October

---

### SUMMARY

A results-driven Data Engineer with solid expertise in data ingestion, transformation, and modeling. Skilled in designing and optimizing ETL pipelines, managing databases, and leveraging machine learning techniques. Proficient in Python, SQL, and big data tools, with a strong passion for creating innovative data solutions that drive business impact.

---

### EDUCATION

- **BSc in Computer Science, Ahram Canadian University – Giza, 6th of October**
- Major: Computer Science
- 09/2018 - 07/2022
- GPA: 3.12

---

### EXPERIENCE

#### Data Engineer and Governance Intern

##### Giza Systems – Headway Program (09/2024 – 02/2025)

- Designed and implemented data ingestion and extraction pipelines using ETL tools (SSIS, Informatica PowerCenter).
- Developed data models and schemas for OLAP and OLTP systems, including dimensional modeling techniques.
- Ensured data quality by implementing validation procedures and lifecycle management for data governance.
- Created detailed documentation and workflows for data pipelines and governance processes.

#### Data Engineer

##### Corelia – Maadi (06/2024 – 09/2024)

- Built automated data extraction pipelines leveraging vector databases for machine learning workflows.
- Designed schemas for machine learning applications, optimizing data storage and retrieval from APIs and on-premises servers.
- Focused on scalability and reliability for large-scale data processing.

#### Data Analyst Intern (Remote)

##### Technocolabs Software – India (08/2022 – 10/2022)

- Preprocessed and analyzed datasets for a genre classification project using machine learning algorithms.
- Collaborated on predicting startups' financial status using real-time Crunchbase data, applying feature engineering and visualization techniques.

## PROJECTS

- **Fake News Detection**
  - Built a machine learning pipeline for English and Arabic fake news detection using TF-IDF, Word2Vec, and pre-trained AraBERT models.
  - Conducted data preprocessing and tokenization to improve accuracy.
  - **Road Pothole Detection**
  - Designed and trained a deep learning model using MobileNet to classify road images for pothole detection.
  - Implemented preprocessing techniques for image datasets.
  - **Wikipedia Data ETL Pipeline**
  - Developed an ETL project to scrape, clean, and transform Wikipedia data into structured formats for analysis.
  - Used Python to automate loading data into SQLite databases and CSV files, ensuring error tracking with logging.
- 

## CERTIFICATIONS AND COURSES

- **Certifications (Coursera)**
  - Machine Learning with Apache Spark (IBM) – 11/2024
  - Introduction to Big Data with Spark and Hadoop (IBM) – 11/2024
  - Introduction to NoSQL Databases (IBM) – 10/2024
  - Data Warehouse Fundamentals (IBM) – 06/2024
  - Hands-on Introduction to Linux Commands and Shell Scripting (IBM) – 03/2024
  - Databases and SQL for Data Science with Python (IBM) – 02/2024
  - Python for Data Science, AI & Development (IBM) – 12/2023
  - Machine Learning by Andrew Ng (Stanford University) – 07/2023
  - **Courses (Udemy)**
  - Data Governance with Axon and Enterprise Data Catalog (EDC) – 01/2025
  - Data Governance: The Complete Course for Beginners – 01/2025
  - Informatica Tutorial: Beginner to Expert Level – 12/2024
  - **ITI Tech-Leaps (Online)**
  - Data Analytics & Visualization Track – 10/2022 to 02/2023
  - Developed data visualization skills using Python and Power BI.
  - Learned dimensional modeling for data warehouses and business intelligence systems.
  - **Udacity**
  - Data Analysis Challenger & Professional Track – 06/2021 to 11/2021
  - Gained proficiency in Python programming, data acquisition, cleaning, and preprocessing.
- 

## SKILLS

- **Technical Skills:**
- Programming Languages: Python, SQL, Java
- ETL Tools: Informatica PowerCenter, SSIS, SSAS
- Big Data Processing: Apache Spark, Kafka, Airflow
- Database Management: SQL Server, PostgreSQL, MongoDB
- Visualization Tools: Power BI, Tableau, Matplotlib
- **Soft Skills:**
- Leadership, Teamwork, Problem-Solving, Analytical Thinking, Communication
- **Languages**
- Arabic (Native), English (Fluent)