Muhammad Hafez

Linkedin | +20 1117447466 | Mohammedhafez1902@gmail.com | Github

Objective

Data modeler with hands-on experience in optimizing data processes and improving model performance. Worked on migrating and streamlining data pipelines, achieving major reductions in runtime while processing billions of records. Comfortable working with cross-functional teams to build strategic business models that support real-world decision-making.

Experience

Sr. Big Data Modeler, Ejada – Riyadh May 2025 – present

- Designed a metadata tracking model for Denodo to support data model monitoring and automate TDD validation processes.
- Co-designed dimensional models for a Lakehouse PoC using Apache Hive, Iceberg, Spark SQL, and Airflow, enabling daily budget reporting instead of yearly snapshot-based reporting and reducing reliance on on-the-fly Denodo queries.
- Migrated seven end-to-end systems in the NCGR MASAR platform from Qlik and Oracle to Impala, Hive, and Kudu using Denodo, after analyzing existing scripts, clarifying business requirements with business teams, and reviewing source system logic to implement accurate data model and reporting changes.
- Contributed to migrating two main operational versions of the Etimad platform with different workflows, aligning them with the legacy workflow to preserve consistent reporting logic and ensure NCGR reports remained unaffected.
- Designed and delivered analytical data models aligned with internal and external stakeholder requirements, partnering closely with Business Analysts and Product Owners.

Big Data Modeler, Orange – Cairo Dec 2023 – Apr 2025

- Migrated over five end-to-end data streams from ClickHouse to Trino using dbt, applying best practices to improve model maintainability, scalability, and reporting reliability.
- Stabilized, reviewed, and optimized over 20 production data models, improving performance, reducing memory usage, and ensuring reliable execution.
- Refactored the monthly customer base usage aggregation model, improving scalability and reducing runtime from 12+ hours to under 1 hour while processing billions of records.
- Improved data quality and reconciliation accuracy in **Voice** and **SMS** switch data models, reducing variance from 50% to 17%.
- Developed analytical data models to support the finance department's strategic objectives.
- Contributed to the development of an **Orange dbt** package that streamlined the development cycle.
- Set up data pipeline monitoring using Metabase dashboards to track pipeline health and execution status.

SAP HCM & MM Consultant, Emisal – Cairo Jan 2022 – Nov 2023

- Contributed to reducing monthly external support costs by supporting the in-house design and implementation of internal applications.
- Helped restructure the SAP HCM payroll pipeline, improving processing speed despite limited technical documentation.
- Supported the transition of wage type configuration from a static setup to a more dynamic structure.
- Contributed to digitizing the purchasing process, helping reduce annual operational costs.
- Supported the migration of warehouse reports to improve KPI tracking and reporting visibility.

Skills

tools: SQL, DBT, Airflow, Trino, Iceberg, HDFS, Hive, Nifi, Metabase, Denodo and Impala.

Software: VS Code, DBeaver, mRemoteNG, Cloudera Hue

Education

University of Helwan, BS in Computer Science Sept 2015 – May 2019

- GPA: 3.02/4.0

Projects

UIPath Smart-Shopper-Robot Github Link

- The RPA process utilizes the UiPath tool to perform a series of tasks: first, it collects product details from users via an input dialog. Next, it launches the Amazon website and conducts a search based on user input. It then extracts comprehensive product details, URLs, and prices. Following this, it filters and cleans the dataset, sorting it by lowest price. The filtered results are then written into an Excel sheet, with the lowest priced product highlighted in color. Finally, the process repeats these steps on the Jumia website to ensure comprehensive data retrieval and analysis across multiple platforms.
- Tools Used: UIPath, DataTable, Data Scrapping, Regix and Excel.

Shopping Cart Front-End – Back-End

- The MERN Stack E-Commerce application caters to two distinct user roles: Users can register, log in to view all products on the homepage, add items to their cart or favorites, manage their shopping cart, and complete purchases securely using Stripe API for credit card payments. Meanwhile, administrators access an admin dashboard to effortlessly Add, Edit, or Delete products from the store, ensuring seamless management and maintenance of the E-commerce platform's inventory and offerings.
- Tools Used: React Js, Mongoose, Node Js, Express Js, Redux, Rest-API.

The Tracker Github

- This Ionic-based mobile application enables parents to monitor their children's locations using the Google Maps API. It detects real-time locations and notifies parents if a child remains in an unfamiliar place for over 15 minutes, ensuring timely alerts and proactive parental intervention.
- Tools Used: Typescript, Angular 6, SQL, Rest-API.

TRAINING'S and COURSES

IBM: Big Data Engineer - Mastery Award for Students (2019) Certificate Link

Coach Academy: Data structure and Algorithms (2017) Conducted a bootcamp designed to teach participants how to measure the memory usage and performance of our software applications.