Mohamed Atef Fahmy

Dubai,United Arab Emirates
+971585448785
muhamedfahmy7474@gmail.com
GitHub | LinkedIn | Medium

Professional Experience

e& (Etisalat) — *Senior Data Engineer*
Nov 2024 - present, Dubai

- Design and execute **real-time and ad hoc campaigns** using **HDFS,SQL**, **Spark**, **Hive**, and **Impala**.
- Develop **scalable data workflows** for efficient campaign management.
- Engineer and maintain a **custom ETL framework** with dynamic Scala compilation for business logic, enabling rapid development of data pipelines and real-time customer segmentation.
- Optimize workflows with advanced SQL for accurate and timely execution

LigaData — *Data Engineer.*
Oct 2022 - Nov 2024

- Spearheaded the ingestion and transformation of large datasets from various sources into **HDFS**, enhancing the analytics capabilities of the team.
- Developed and optimized ETL pipelines to facilitate efficient data flow, enhance reporting capabilities, and create reports on Power BI.
- Reduced report generation time by 40% by automating data workflows using **Python**, **Hive**, and **Power BI**.
- Contributed to the development of a data fabric product on **Kubernetes**, enabling secure and scalable data integration across cloud environments.

Upwork -Freelance Data Engineer & Analyst
Jun 2022 - Dec 2023

- Built **ETL processes** to gather and analyze large-scale job market data using web scraping techniques, contributing to clients' tasks.
- Developed and deployed machine learning models to predict trends with an 85% accuracy, streamlining clients' forecasting processes.
- Earned consistent 5-star feedback due to timely deliverables and high-quality data insights.

Cube— *Data Analyst and Operations.*
Aug 2019 - Sep 2021.

- Automated data processes with Python and SQL, cutting manual work by 40%, and built Power BI dashboards to improve decision-making.

Projects

Covid-19 Data Pipeline And Analysis Using Azure

- Utilized Azure Data Factory, Data Lake Gen2, Blob Storage, Azure SQL, HDInsight, and Databricks.
- Implemented Power BI for data visualization and reporting.
- Developed CI/CD pipelines via Azure DevOps for automated releases.
- Used Azure Data Studio for data warehouse queries and Azure Storage Explorer for storage management.

ArsenalFC End to End Data Pipeline and Analysis

- Led the development of a robust data engineering pipeline to analyze data concerning Arsenal FC's players, goalkeepers, and match performances.
- Processed and analyzed data to derive insights on player performance, match outcomes, and goalkeeper stats.
- Tools Used: Python, PySpark, Docker, PostgreSQL, Power BI, Airflow,Jupyter.

Olympic Data Analysis on Azure.

- Utilized Azure cloud services to perform historical data analysis and visualization of Olympic Games data.
- Employed Azure Databricks, Azure SQL Database, Azure Data Factory, Azure Synapse Analytics, and Azure Storage to provide a scalable data processing solution.

Objective

Experienced Data Engineer and Analyst aiming to apply expertise in data ingestion, ETL, and Reporting to enhance data-driven decision-making. Proficient in Python, SQL, and cloud technologies, I seek to contribute to projects that utilize cutting-edge technologies and foster business efficiency in a dynamic team environment.

Education

ITI—*Machine Learning Diploma.*

Mansoura University—*BSc. In Computer Engineering.*

Vocational Skills

Operating Systems and Programming: Python, Pandas, Numpy, Scikit-Learn, SQL, Presto, Bash Script, Linux, Scala

Data Visualization and Reporting: Power BI, Matplotlib, Seaborn

Big Data and Distributed Systems: Apache Hive, Apache Hadoop, Apache Spark, Kafka, Impala, Flink.

ETL and Data Warehousing: Airflow, Talend, Dbt, IGC.

Cloud and Containerization: Docker, Kubernetes, Terraform

Version Control and Collaboration: Git, GitHub, Jira, Confluence

Web Scraping: BeautifulSoup

Soft Skills

Communication & Interpersonal Skills.

Problem-solving

Critical thinking

Resilience

Leadership

Attention to detail.

Problem Solving.

Courses & Certifications

Azure Data Engineering DP-203

AWS Cloud Practitioner

DataEngineering ZoomCamp

Udacity AWS DataEngineering

Udacity Data Streaming Nanodegree