Rana Badawy

Irving, TX 75063 • +1-442-224-3461 • rana_badawy@outlook.com
https://www.linkedin.com/in/ranabadawy/

Data Engineer

Data Engineer with 3 years of professional experience in designing and implementing scalable ETL data pipelines, building data warehouses, and delivering actionable insights. Proficient in Python, SQL, Azure, and Power BI, with a strong foundation in data modeling, data integration, and automation.

ETL • Data Warehouse • Data Lake • Data Modeling • Data Integration • Data Analysis • Reporting • Web Scraping
Algorithms • Data Structures • OOP • APIs • RPA • AI

Languages: Python, Java
Cloud: Azure (Databricks, Data Factory, Event Hubs, Stream Analytics, Synapse Analytics, Data Lake, SQL, Logic Apps, DevOps), AWS (S3, Lambda, RDS, DynamoDB)
Big Data: Spark, Pandas, Hadoop, Hive, Sqoop, Flume, Power BI
DevOps & CI/CD: Terraform, Azure DevOps, Docker
AI/ML: LangChain, scikit-learn, TensorFlow, Keras, PyTorch
Databases: SQL, NoSQL, T-SQL, MySQL, MongoDB, HBase
Web Development: JavaScript, HTML, CSS, REST, Node.js, React.js
SDLC: Agile, Scrum
Tools: KNIME, Power Apps, Power Automate, Jira, Git, PyCharm, VS Code, Postman, UiPath, Blue Prism, Excel

---

Professional Experience

Orion360 Digital Transformation (01/2021 – 08/2023)
Data Engineer (07/2021 – 08/2023)

- Implemented scalable ELT Data Pipelines using Azure Databricks, Synapse Analytics, ADF, and ADLS, accommodating adding 10+ new data sources without disrupting existing workflows.
- Restructured data warehouse design by applying data modeling techniques to design multidimensional structures, including star and snowflake schemas, resulting in reduced storage usage, and faster processing times.
- Developed data analysis scripts using Pandas and SQL to extract actionable insights on stock, purchase requisitions, and planned orders, enhancing operational decision-making efficiency by 20%.
- Built custom data quality validation scripts ensuring accuracy and consistency in business-critical datasets, reducing errors by 15%.
- Designed Power BI dashboards and DAX measures to visualize RPA achievements, directly contributing to new business opportunities and increasing revenue by 5%.
- Created efficient Python scripts to orchestrate shipment organization and optimize orders distribution, reducing processing time from 2 hours to 7 minutes.

Technologies Used: Python, PySpark, Pandas, MS SQL, Stored Procedures, Azure Databricks, Azure Data Factory, Azure Synapse Analytics, Azure SQL, Azure Data Lake, Azure Key Vault, Azure Logic Apps, Azure DevOps, Microsoft Power BI, DAX, KNIME, Jira

Robotic Process Automation Engineer (01/2021 – 07/2021)

- Automated data extraction and transformation from SAP using Blue Prism, applying Excel formulas and Macros for detailed analysis and generation of reports on stock and order workflows.
- Developed Python web scraping scripts to extract and compare product data from websites with client datasets, enhancing market analysis.
- Consolidated and transformed data from multiple systems into centralized repositories, using Python APIs, Blue Prism, and Excel reducing manual work by 20%.
- Designed the MySQL database for an internal Leave Management System, integrating it with the back-end to facilitate efficient data storage and retrieval.
- Leveraged Power Apps to develop interfaces for managing internal datasets, reducing manual data entry and increasing data integrity by 15%.

Technologies Used: Blue Prism, Microsoft Power Apps, Microsoft Power Automate, Excel Macros (VBA Macros), Python, MySQL, Node.js, JavaScript, React, HTML, CSS, Bootstrap

Projects

Natural Language to SQL: This AI and Machine Learning project provides an interface to convert natural language input into SQL queries, execute them against a MySQL database, and returns the answer in natural language. Link
Technologies Used: Python, LangChain, Docker, Google PaLM, Chroma vector database, Streamlit, ML, LLM, RAG, Few-shots learning

Azure Streaming: This project demonstrates a real-time data streaming pipeline using Azure services. It simulates data generation, ingestion, and processing with a focus on real-time capabilities. Link
Technologies Used: Python, Azure Event Hubs, Azure Stream Analytics, Azure Databricks, Azure Data Lake

---

Certifications

OCI Generative AI Certified Professional
By Oracle
Obtained 05/2024

Education

Master of Science in Computer Science
Maharishi International University, Fairfield, Iowa
In progress via distance education

Bachelor's Degree in Computer Engineering
Zagazig University, Zagazig, Egypt