

ALFRED AMACHREE

Kansas City, MO | 559-753-6543 | amachreealfred@yahoo.com

alfredamachree.com

linkedin.com/in/alfredamachree/

github.com/alfredama

EDUCATION

University of Kansas, Missouri July 2020

Certificate, Data Science (Bootcamp)

Covenant University, Nigeria July 2017

Bachelor of Engineering, Information and Communication Engineering

WORK EXPERIENCE

Relias, Morrisville, NC (Remote) June 2023 – Present

Data Engineer II

- Developed Azure Synapse notebooks using python/pyspark and SQL to ingest click stream data from multiple API sources like (Iterable, Pendo) and landed them in azure serverless SQL pools for use by data scientist and analysts to develop actionable insights for business.
- Orchestrated the data ingestion of Advertisement/User Acquisition data from multiple social media sources/channels through funnel IO and google analytics using multiple Data factory pipelines/pyspark notebooks and Data modelling skills, enabling the marketing team to analyze the relationship between ad spend and user acquisition. This process also helped the business in evaluating the success of paid ad campaigns /organic campaigns across multiple benchmarks like channels and user interactions, cost per click, impressions, conversions, ad Spend etc.
- Developed ETL pipeline in Azure data factory and pyspark notebook to Ingest data tracking user activity/engagement with pendo user features/ web pages on Relias learning website. Initial historical load backlogged to an estimated 600 million rows of data.
- Ingested data for Customer engagement dashboard from multiple sources via API and static files and made use of power Bi DAX, SQL and Python to create informative dashboard/visualizations relating to customer engagement.
- Streamlined and optimized Azure Synapse instance by consolidating redundant pipelines into a template pipeline, enabling seamless data consumption from APIs and efficient storage in data lake. This helped to unify our teams API ingestion making teamwork more consolidated.
- Created documentation on data consumption processes for the team.
- Developed relationship with Business stakeholders and continued to advocate for Data driven decision making.

H&R Block, Kansas City MO

October 2020 – May 2023

Data Engineer II

- Developed multiple ETL pipelines to ingest data from various on prem and cloud sources leveraging Microsoft Azure data factory and IBM DataStage to form the backbone data sources for enterprise reporting in Looker.
- Developed databricks notebook using pyspark and sparksql to incrementally load delta live tables to Azure SQL Database for consumption by data scientists.
- Performed data analysis and data quality checks on tables loaded to ensure data loaded was properly cleansed and accurately normalized for downstream data users.
- Worked with Senior Leadership team to scope and define KPIs for new banking product. Modelled Data, Designed and maintained multiple ETL Pipelines for this product leveraging Databricks, Pyspark, SparkSql and Azure data factory.

- Developed a python library that helped to speed up how data was migrated from OLTP DB into Enterprise Data Platform saving engineers time spent re-writing redundant sql queries.
- Developed an automated date dimension table script to generate date dimension table used companywide for reporting.
- Created technical documentation on migrating Jupyter notebooks and developing standard JDBC connection strings to enterprise cloud platform.
- Maintained and installed= new and custom python and R packages/libraries for data scientists.
- Worked with Data Analysts/Information Governance teams to understand and scope data needs for accurate reporting/visualizations.

Oracle Cerner Corporation, Kansas City MO

October 2018 – October2020

Interface Engineer (Data Integration/Ingestion)

- Designed and implemented robust interface pipelines, using ETL concepts in the migration of big data from multiple healthcare organizations, transformed the data from variety of sources like csv's, HI7 etc., to meet Cerner specifications using CCL and successfully loaded the refined data into Oracle Databases.
- Made use of multiple commands to navigate Linux to manipulate and transfer large data files through multiple directories.
- Designed Training guide and trained new engineers on interface design and types of data manipulation scripts using Cerner Millennium tools alongside dB technologies.
- Travelled to client site to provide support for go-lives.
- Created a model to predict new employee billability using knowledge of Linear regression also tools like pandas dataframes and matplotlib based on data Ingested from Survey monkey.

TECHNICAL PROJECTS

- **Weather RESTAPI App:** Created a RESTAPI using Python Flask library that ingest data retrieved via get request, transforms retrieved data and stores in my SQLAlchemy DB and delivers JSON responses to a user over the web.
- **Web scraping App:** Used Python Beautiful soup library to Perform Web scraping of NASA webpage and stored retrieved information in MongoDB and created an API endpoint that can interact with the DB and return results in a HTML webpage.
- **Predictive Modelling:** Used Python/Matplotlib and knowledge of Machine learning and Linear regression to model possible temperatures in each location based on their latitude in a dataset gotten from Kaggle.

SKILLS

-
- **Programming Languages:** Python, HTML/CSS, JavaScript, SQL, VBA, CCL, Pyspark, SparkSQL
 - **Libraries:** - Flask, Pandas, Sklearn, NLP, Matplotlib, NumPy, Beautiful Soup, SQLAlchemy, Plotly, D3, Leaflet.
 - **Cloud:** Azure, IBM cloudpak, Databricks
 - **Databases:** MongoDB, PostgreSQL, Oracle SQL, AzureSQL, Cosmos **Tools:** Git, Tableau, Looker.
 - **Exploring:** Docker, GraphQL, AWS, CI/CD
 - **Soft Skills:** Lean Process Improvement, Problem Solving, Collaboration, Project Management, Agile dev.

CERTIFICATIONS

-
- Microsoft certified: Azure Data Engineer Associate
 - Microsoft certified: Azure Fundamentals