Fredrick Lambuth

+17138395385 | flambuth@gmail.com | linkedin.com/in/fredricklambuth/ | fredlambuth.com/

Summary

I build end-to-end data solutions, not mere ETL jobs between nodes. With classic software engineering techniques, I create modular and easy to use data infrastructure. My professional history is a mix of liberal arts education, IT operations, and data analytics. That eclectic mix has given me a wide set of faculties for solving problems by blending analytical thinking to solve everyday problems and creativity for the difficult ones. Identifying problems comes from collaborating, and double-checking, with the business side to ensure my work solves real needs, and coordination with peers to make sure our outputs sync.

Professional Experience

Logic2020   **Remote**
Data Engineer   *October 2023 - Present*
- Back-end software development with Palantir Foundry
- Developing data pipelines for triggering customer notifications of electrical power outages.

Aspirent   **Remote**
Data Consultant   *September 2022 - September 2023*
- Developed a PySpark module for data quality checks in pipeline steps, reducing the turn around time to report errors to data providers.
- Created Pandas scripts to enhance options for analysts conducting anomaly detection on Medicare claims.
- Converted stored SQL procedures into PySpark notebooks held in Azure Synapse before contract deadline despite lack of promised documentation from client.
- Built pipeline to write API calls Shopify to extract sales data, transform to fit model and extract to client's Redshift database. Developed without client support caused by departure of tech lead.
- Authored the GIT playbook for developing cloud data warehouse in adherence to the client's company policy and trained client's for hand-off of version control management.

KYM Advisors   **Remote**
Senior Data Analyst   *October 2021 - September 2022*
- Refactored pipelines made from a handful of legacy SQL script into loosely coupled Pyspark modules. Modules allowed for data quality checks, parallelization within Spark jobs, unit testing, quicker development, and much faster debugging.
- Designed, implemented, and administered hierarchical Hive tables that reduced query time, complexity of analysts' queries and efficiency of scheduled jobs.
- Redesigned single table archives of previous reports into directories partitioned by date, rapidly increasing the query time on historic data.

Wells Fargo   **Remote**
Analytics Consultant   *January 2020 - September 2021*
- Replaced disparate manual SAS queries and Excel macros into a single, shareable Pandas module allowing all analysts on the team to execute all data reports from a shared drive.
- Managed SharePoint collecting user text entries among regional managers. Created SQL job to extract from API, transform data and load as a data source for Tableau Server, usable for all production dashboards.
- Developed data model and ETL pipeline to extract customer data into HR data warehouse, allowing for daily analysis between customer and contact center operations data.
- Pioneered the migration of local SSIS jobs into an Azure Data Factory runtime with parameterized inputs, easing development of further migrations.

CACI   **Norfolk, VA, USA**
Unix System Administrator   *February 2018 - December 2019*
- Collected admin tasks into testable Bash scripts with a standardized parameter and naming schema into shared directory for use by other system administrators.
- Developed Python script to detect system log anomalies, summarize into HTML document, and emailed to concerned parties when detected, or on scheduled basis.
- Automated the creation of containerized deployment packages for software updates ready for testing environment, to include notifications to parties involved.

United States Navy   **Norfolk, VA, USA**
Electronic Technician   *February 2012 - February 2018*
- Served as a maintenance operator for wireless network equipment aboard the USS Ford. Collateral duties included IT help-desk support, Windows endpoint setups, office management, and cryptographic security.

Certifications

Databricks - Associate Developer for Apache Spark
Microsoft - Azure Data Engineer

Education

Texas A&M University - College Station   **September 2004 - July 2009**
Bachelor's, Political Science

Skills

Skills: Python, SQL, Apache Spark, Pandas, Microsoft Azure, SQL Server, Databricks, Linux/Unix, HTML/CSS, REST APIs, PowerShell, Tableau, Python NLTK, Natural Language Processing (NLP), UI/UX Design, Data Analysis, Flask, Django, Bash