

Experience

Employee

Senior Data Engineer

2019 - 2020

Data Engineering:

- Built a self-service Airflow DAG in Python and AWS to enable table sync from Redshift to Snowflake, aiding in faster pipeline migration to Snowflake. Paired with engineers on the team for development
- Developed Python tooling to help other data engineers and scientists to work efficiently and consistently
- Provided technical best practices on various data architecture components via documentation and knowledge sharing and drove adoption to ensure all teams are working consistently
- Worked closely with engineers to plan enterprise-wide RBAC (Role Based Access Control) for Snowflake and successfully implemented it across teams
- Assigned to be a primary engineer for team service desk and on-call rotation within 2 months of joining

Analytical Data Modeling:

- Set up data connectors and collaborated with BI engineers to productionize the core Finance Data
- Model used for reporting WeWork's financial metrics . Technologies used: Airflow, Fivetran, dbt, Python and Snowflake
- The dbt output included the data lineage and data dictionary for the Finance Data Model
- Company-wide Training and Development (Bootcamp):
Team lead for the Snowflake bootcamp for data engineers/scientists across business and data units in New York, Tel Aviv and Shanghai to successfully enable teams to learn and use Snowflake for their business goals
- Prepared lab sessions, presentation, logistics etc; conducted follow-up office hours to assist engineers/scientists

Senior DataOps Engineer

2015 - 2019

Engineering and Modeling:

- Built and maintained a robust custom orchestration framework and connectors in Python to extract, load and transform (ELT) data from 50+ relational databases, 10+ non-relational collections, various external sources into an analytics datawarehouse to provide timely and accurate data to multiple business functions
- Led to successfully complete a 50TB Enterprise-wide Cloud Data Migration Project to migrate automated pipelines, dozens of reports, a couple hundred users and thousands of tables without any disruption to business
- Partnered with data scientists to set up ingestion, monitoring and alerting for mission-critical data processes like the Loss Prediction, Loan Grading and Credit Risk Score models and software engineering teams to design data models for new applications
- Set up an automated Financial Reconciliations system and developed a custom parser in Python for BAI2 transactions reducing manual effort by 60 man-hours a month

Data Infrastructure:

- Responsible for developing tools for data transfer and validation to ensure data parity between multiple data and communication with stakeholders

Primary Engineer

2015 - 2015

Reporting and Business Intelligence:

- Developed ETLs to set up the analytical truth layer containing standardized metrics for Finance and Marketing reporting including investor loan sales and performance reporting, company-wide reporting for loan products, NPS and marketing campaigns
- Developed an Email and FTP reporting delivery framework in Python used for all external stakeholders such as investors, agencies, partners and affiliates and internal stakeholders including executives
- Mentored data analysts on query optimization and tuning techniques and best practices

Data Operations and Support:

- Developed an alerting mechanism for the Data Engineering/Ops team on incidents for critical data pipelines/reports
- Performed Level 1-2 activities including query optimization / tuning, capacity planning and reporting, custom scripting /automation, developing runbooks, project-driven engineering. Level 3 activities including root-cause analysis (RCA) for incidents, regular database maintenance, preparing Incident and Problem reports, mentoring Level 1-2 engineers

Senior Associate

2013 - 2015

Data Analytics and Visualization; AML Code Remediation and New Scenario Design and Development:

- Team lead for assessment of a Fortune 500 American multinational bank's anti-money laundering (AML) system by identifying gaps between the master business glossary and the Analytics environment
- Developed and implemented the AML detection scenarios to remediate the issues and gaps identified, reducing false-positive alerts by 18%
- Developed Strategy and Data Visualizations to assess the current state, evaluate ideas and make strategic decisions to present to senior management including the Chief Compliance Officer
- Data Quality Assessment, Validation and Reporting:
- Led the evaluation of a Mexico-based global bank's Data Quality controls to conduct an end-to-end evaluation of the Transaction Monitoring environment and performing heavy analysis for data reconciliation of 500 million transactions
- Analyzed the bank's AML system for data quality assurance to ensure proper monitoring and alert generation, providing detailed reports on data inconsistencies, data reconciliation as well as data validation across various systems