ALAN SUSA

Data Engineer

✉ alansusa@email.com   📞 (123) 456-7890   📍 New York, NY

in LinkedIn

WORK EXPERIENCE

Data Engineer
Consumer Reports
📅 May 2018 - current   📍 New York, NY

- Led the migration from Oracle to Redshift using Amazon Athena and S3, resulting in an annual cost savings of $678,000 and an increase in performance of 14%
- Designed and implemented a real-time data pipeline to process semi-structured data by integrating 150 million raw records from 30+ data sources using Kafka and PySpark
- Designed the data pipeline architecture for a new product that quickly scaled from 0 to 125,000 daily active users
- Studied and revamped data dictionaries to include a more robust history for developing consistency across domain

Data Engineer
Guardian Life Insurance Company
📅 August 2016 - May 2018   📍 New York, NY

- Maintained data pipeline up-time of 99.8% while ingesting streaming and transactional data across 8 different primary data sources using Spark, Redshift, S3, and Python
- Automated ETL processes across billions of rows of data, which reduced manual workload by 29% monthly
- Ingested data from disparate data sources using a combination of SQL, Google Analytics API, and Salesforce API using Python to create data views to be used in BI tools like Tableau
- Communicated with project managers and analysts about data pipelines that drove efficiency KPIs up by 26%

Data Engineer Intern
Federal Reserve Board of Governors
📅 August 2014 - August 2016   📍 Washington, DC

- Built basic ETL that ingested transactional and event data from a web app with 12,000 daily active users that saved over $85,000 annually in external vendor costs
- Worked with client to understand business needs and translate those business needs into actionable reports in Tableau, saving 17 hours of manual work each week
- Utilized Spark in Python to distribute data processing on large streaming datasets to improve ingestion and processing speed of that data by 67%
- Supported implementation and active monitoring of controls and programs for precision and efficacy

EDUCATION

B.A.
Computer Science
University of Pittsburgh
📅 September 2010 - April 2014
📍 Pittsburgh, PA

SKILLS

Python
ETLs
SQL (Postgres, Redshift, MySQL)
NoSQL (MongoDB)
Spark, Kafka
Airflow
AWS (Athena, Lambda, S3)