Erika Jacobs

ErikaAshley3@gmail.com • (480) 529-1864 • www.erika-jacobs.com

Summary

Data professional with over 5 years of experience in data management and cleaning, as well as development and implementation of ETL pipelines. Passion for learning new data-related technologies and solving data production issues.

Skills

- **Languages:** Python, Scala, SQL, R, HTML, CSS
- **Databases:** SQL Server, MySQL, PostgreSQL, Redshift, Athena, RDS, DynamoDB, MongoDB
- **Platforms:** AWS (S3, EC2, IAM, CLI, EMR, Glue), Linux, Bash, GCP, Oracle, Microsoft Azure
- **Frameworks**: DataBricks, Spark, Kafka, Airflow, Hadoop (HDFS, MapReduce, Hive)
- **Tools:** GitHub, R Studio, Access, Excel, Tableau, Visual Studio (SSIS)
- **Application:** data modeling, schemas, cloud architecture, data mining, statistics, machine learning, agile/scrum

Project Experience

NASA Data Warehouse

Python (Boto3, Psycopg2, Requests, Pandas), AWS (Redshift, S3)

- Read unstructured weather data from the NASA space weather API. Cleaned and structured data into multiple data frames. Created security role and opened port to connect to Redshift. Staged data in S3 to load to Redshift.

Work Experience

Nationwide Insurance Company, September 2020-Present   **Scottsdale, AZ**

Data Engineer

- Design, develop, and maintain scalable ETL pipeline for telematics data quality with connected car signals using Spark SQL and S3.
- Automate and maintain manual scoring process for driven trips as part of Nationwide's Smartride program using Spark SQL.
- Conduct exploratory data quality analysis for big data from multiple IoT vendors, and identify data quality issues in existing architecture. Work with internal customers to investigate data quality issues and provide solutions.

Nationwide Insurance Company, January 2017-September 2020   **Scottsdale, AZ**

Business Specialist

- Developed ETL scripts to extract data and remove macros from hundreds of Excel documents. Cleaned data and changed variable data types to match SQL Server table, and built loop to insert dataframe into SQL Server.
- Manage monthly ETL pipelines for data production, and manage prioritization amongst responsibilities.
- Automate consolidation and merging of datasets, and verify accuracy of processed data sources for reporting.
- Conduct ad-hoc quantitative analyses, build exhibits, and summarize multiple metrics from large and complex datasets using SQL queries.

Nationwide Insurance Company, December 2015-January 2017   **Scottsdale, AZ**

Senior Analyst

- Cleaned and manipulated datasets for analysis and verified accuracy of all data.
- Built SQL tables and views for analysis. Wrote queries to consolidate and combine complex datasets.
- Generated dashboards and visualizations, with output including tables, metrics, charts and graphs.
- Maintained positive working relationships with leadership clients across different departments.

Education

Penn State University   **University Park, PA**

Masters in Applied Statistics - 3.76 GPA

Arizona State University   **Tempe, AZ**

Bachelor of Arts in Psychology, Theatre

Activities and Awards

- Mathematics Professor through the Maricopa Community Colleges 2019
- Toastmasters Competent Communicator and Competent Leader Awards