(813) 215-7748

nswatson07@gmail.com

Atlanta, GA

www.linkedin.com/in/nicoleswatson

www.github.com/nsciccarello

www.nickylovesdata.com

Nicole Watson Ciccarello

Data Engineer

SUMMARY

I am an Atlanta-based Data Engineer skilled in creating and optimizing data pipelines with Python, Scala, Spark, and SQL. I am also an expert in cloud technologies like Azure and AWS, specializing in performance optimization and ETL automation. With my unique background in journalism, my enhanced analytical and storytelling abilities enable me to deliver impactful data solutions for major clients like The Coca-Cola Company.

EXPERIENCE

Data Engineer
August 2022 - Present

The Coca-Cola Company
Atlanta, GA

- Deployed 2 new ETL pipelines using Scala, Java, Spark, and SQL to help automate data processing from legacy systems.
- Improved major bottlenecks in algorithm and metric processing by decreasing run time by 25%.
- Migrated 7PB data warehouses including stored procedures, data dictionaries, and tables from SQL Server to Synapse Analytics.

HIGHLIGHTED PROJECTS:

Global Supply Chain Project

- Developed a global data pipeline for 250+ Coca-Cola bottlers worldwide using Scala, Spark, SQL, and Microsoft Azure.
- Achieved a 63% reduction in monthly workload by automating the ETL process, efficiently handling billions of rows of data.
- Implemented a Spark-based analytics solution for processing large-scale datasets, resulting in a 50% improvement in data analysis time.

Marketing Media Project

- Created an ETL pipeline handling data from more than 50 different databases, each representing a marketing client like Google.
- Enhanced performance by 40% by converting SQL Server stored procedures to Spark SQL for Synapse Analytics integration.
- Successfully migrated 3PB of data to Azure Data Lake Storage (ADLS), optimizing storage costs and improving data accessibility.

Data Engineer
March 2022 - August 2022

Correlation One - Data Engineering Fellowship
Remote

- Selected for part-time fellowship (4% acceptance rate) with learning facilitated through real-world business cases.
- Gained expertise in AWS tools including S3, Redshift, and Lambda to enhance storage, processing, and automation of ETL pipelines.
- Mentored junior data engineers during project, promoting knowledge sharing and professional growth.

HIGHLIGHTED PROJECT:

Movie Recommendations for Domestic and International Flights

- Directed team of 6 to construct an ETL pipeline to help airliners select in-flight entertainment for passengers.
- Architected airliner ETL pipeline with Python, SQL, Hadoop, PySpark, Hive, AWS ECRs, and APIs with 1,000,000+ flight data.
- Recipient of the 'Most Distinguished Project' award, recognizing exceptional innovation and technical proficiency.

Software QA Engineer - Data Warehouse
March 2019 - July 2022

Warner Bros. Discovery
Atlanta, GA

- Identified and reported 98% of software bugs in testing phases, significantly enhancing product quality and reliability.
- Managed team collaboration as an Agile Scrum Master and increased team deliverable points by 20%.
- Implemented and validated data storage and warehousing solutions using AWS S3 and Redshift, ensuring data integrity and scalability.

HIGHLIGHTED PROJECT:

CNN Library Migration Project

- Spearheaded tests to migrate petabyte-scale CNN digital content to AWS S3 for 2000+ employees using Java and Selenium.
- Automated daily regression tests with Java and IntelliJ, reducing test duration by 60-75%.
- Ensured data integrity and accuracy during cloud migration, employing rigorous testing protocols and tools.
- Utilized AWS CloudWatch for monitoring and logging test environments, enabling proactive issue detection and resolution.

OTHER TECHNICAL CONTRIBUTIONS

Founder and Data Engineer   July 2021 - Present

Nicky Loves Data

- Founded a data engineering and analytics blog to showcase projects and skills such as data transformation and data visualization.
- Analyzed 1000+ lines of Super Bowl game data from CSV files to study correlations between outcomes, viewership, and halftime shows.
- Discovered with 95% confidence that Super Bowl viewership will decrease due to large team score differences.

TECHNICAL SKILLS

Programming Languages: Java, Python, Scala, SQL, R, JavaScript, Powershell, Bash
Cloud Technologies: AWS, Microsoft Azure
- **AWS Tools:** S3, Glue, Redshift, EMR, RDS, Lambda, Athena, DynamoDB
- **Microsoft Azure Tools:** Synapse Analytics, Data Lake Storage, Blob Storage, SQL Database, Data Factory, Databricks

Data Management & Big Data Technologies: Microsoft SQL Server, MySQL, Apache Spark, Hadoop/HDFS, Hive, Airflow, Data Modeling, Data Normalization, Data Warehousing, Data Processing, ETL
DevOps & Version Control: Git, GitHub, GitLab, Jenkins
Other Technologies: REST APIs, Selenium, .NET, Plotly Dash, Power BI, Excel

EDUCATION & CERTIFICATIONS

Correlation One Data Skills for All

Data Engineering Fellowship Certificate - Honors

Harvard University, Cambridge, MA

Courses in Computer Software Engineering

University of South Florida, Tampa, FL

B.A. Mass Communications