ERICK SANG
Data Engineer

Email: ericksang@gmail.com
Mobile: +57 313 481 6342
LinkedIn: linkedin.com/in/esangc
GitHub: github.com/leonsang
Portfolio: ericksang.dev
Location: Madrid, Colombia

PROFESSIONAL SUMMARY

Data Engineer with 5+ years of experience designing, developing, and maintaining scalable data pipelines, data warehouses, and data lakes. Expert in Python, SQL, SAS, ETL/ELT, and Big Data technologies including Hadoop, Spark, and Kafka. Proven ability to collaborate with cross-functional teams and DevOps to deliver high-impact data solutions. Specialized in on-premises to cloud data migration (AWS, Azure, GCP), process automation, optimization of relational and NoSQL databases, and data visualization with Power BI, SAS VIYA, LookerML, and BigQuery.

WORK EXPERIENCE

SohoHumantech   **Bogotá, Colombia**
Data Engineer   *June 2025 - July 2025*
- Developed and implemented data pipelines for human resources information processing
- Designed data warehouse architecture to centralize employee data and HR metrics
- Implemented ETL/ELT solutions using Python and SQL to automate data processes
- Created interactive dashboards with LookerML and BigQuery for talent management KPIs visualization
- Developed frontend with React and Tailwind CSS for API integration and dashboard component generation
- Optimized BigQuery queries achieving 40% improvement in response times
- Collaborated with HR teams to define data requirements and business metrics
- Implemented data quality validations and automated process monitoring

Corporación Unificada Nacional (CUN)   **Bogotá, Colombia**
Data Engineer   *April 2025 - June 2025*
- Designed and implemented Single Source of Truth (SSOT) architecture in AWS for the university
- Developed data pipelines using Python full-stack with AWS Glue, Lambda, and S3 for academic data consolidation
- Implemented data lake in AWS S3 with optimized partitioning for efficient queries
- Utilized Amazon Redshift for data warehouse with query and index optimization
- Implemented Apache Airflow in AWS for data workflow orchestration
- Developed REST APIs with AWS API Gateway and Lambda for consolidated data access
- Implemented monitoring and alerting with CloudWatch for data pipelines
- Achieved 60% reduction in consolidated data access time

FACTECH   **Bogotá, Colombia**
Data Engineer   *September 2024 - March 2025*
- Designed and implemented efficient data pipelines using Apache Airflow, achieving 90% improvement in processing
- Developed ETL/ELT solutions for large data volumes and real-time flows with Python and SQL

- Migrated legacy systems to modern cloud architectures in AWS/Azure, ensuring integrity and optimization
- Implemented automated data quality validations reducing errors by 35% and guaranteeing data quality
- Optimized relational databases (SQL Server, PostgreSQL, Oracle) and NoSQL (MongoDB, Cassandra), reducing query times by 50%
- Built infrastructures for Machine Learning solution integration and analytical processing
- Collaborated with DevOps teams in implementing Infrastructure as Code (IaC) solutions and CI/CD processes
- Implemented data modeling and data warehousing using Snowflake and Redshift

Teleperformance   **Bogotá, Colombia**
Data Specialist   *December 2021 - March 2025*
- Designed and developed data pipelines for ingesting data from various sources (REST APIs, web scraping, RPA, log files)
- Transformed and cleaned raw data using Python and SQL, ensuring data quality and consistency
- Maintained monitoring and uptime of all production databases and processes (SQL jobs, .NET framework executables)
- Diagnosed and fixed database and process errors using debugging tools
- Optimized system performance and functionality, recommending optimizations to existing queries
- Built data models for data warehousing and analysis using data modeling techniques
- Implemented data visualization and reporting using Power BI and Excel VBA

EDUCATION

Polítécnico Grancolombiano, Bogotá | *Software Engineering*   **2023 - Present**
- Status: Eighth semester approved
- Relevant Courses: Software Architecture, Database Systems, Cloud Computing, Machine Learning

ECCI, Bogotá | *Automation and Robotics Technology*   **2017 - 2021**
- Status: Degree option
- Relevant Courses: Control Systems, Industrial Automation, Robotics Programming

Universidad Distrital Francisco José de Caldas, Bogotá | *Electronic Engineering*
2012 - 2016
- Status: Deferred-retired
- Relevant Courses: Digital Systems, Microprocessors, Signal Processing

KEY PROJECTS

Single Source of Truth (SSOT) in AWS | *Corporación Unificada Nacional (CUN)*
2025
- Design and implementation of SSOT architecture in AWS for academic data consolidation
- Development of data lake in S3 with optimized partitioning
- Implementation of data warehouse in Redshift with query optimization
- Development of REST APIs for consolidated data access
- Full Python stack with AWS Glue, Lambda, S3, Redshift
- 60% reduction in consolidated data access time

HR System with Data Warehouse | *SohoHumantech* 2025

- Development of data warehouse for employee data centralization
- Implementation of ETL/ELT pipelines for process automation
- Creation of interactive dashboards with LookerML and BigQuery
- Frontend development with React and Tailwind CSS
- BigQuery query optimization with 40% improvement in response times

On-Premises to Cloud Data Migration | *FACTECH* 2024

- Successfully migrated legacy systems to modern cloud architectures in AWS/Azure
- Implemented data lakes and data warehouses using Snowflake and Redshift
- Performance optimization with 50% reduction in query times
- Implementation of data governance and compliance frameworks

TECHNICAL SKILLS

Data Engineering ETL/ELT, Apache Airflow, Data Pipeline Development, Data Warehouse Design, Data Lake Implementation

Big Data Hadoop, Apache Spark, Apache Kafka, Data Processing, Real-time Data Flows

Cloud Platforms AWS (Glue, Lambda, S3, Redshift, CloudWatch), Azure, GCP, Snowflake

Databases SQL Server, PostgreSQL, Oracle, MongoDB, Cassandra, BigQuery

Programming Python, SQL, Java, JavaScript, R, Excel VBA

Data Visualization Power BI, SAS VIYA, Looker, LookerML, Tableau, Dashboard Development

Frontend Development React, Tailwind CSS, API Integration, Dashboard Components

DevOps CI/CD, Jenkins, Git, GitHub, Infrastructure as Code (Terraform, CloudFormation)

Data Quality Data Validation, Data Cleansing, Data Governance, Automated Monitoring

CERTIFICATIONS

Python Essential, LinkedIn Learning	2024
Python Certify, ProgrammingHub	2024
SQL Basics, SoloLearn	2024
SQL Intermediate, SoloLearn	2024
SQL Advanced, SoloLearn	2024
Power BI, LinkedIn Learning	2024
Associate Data Engineer in SQL, DataCamp (In Progress)	2024
Data Engineer in Python, DataCamp (In Progress)	2024

LANGUAGES

Spanish Native

English Intermediate (B2) - TOEFL/IELTS certification in progress

ACHIEVEMENTS & METRICS

Business Impact 90% improvement in data processing speed, 60% reduction in consolidated data access time, 40% optimization in HR queries, 35% reduction in data errors, 50% optimization in database query times

Technologies Implemented 6+ cloud technologies (AWS, Azure, GCP, Snowflake, Redshift, S3), 3+ programming languages (Python, SQL, Java), 4+ databases (SQL Server, PostgreSQL, MongoDB, Cassandra), 2+ visualization tools (Power BI, SAS VIYA)

Availability 24/7 availability of critical systems, 100% compliance with regulatory requirements