SUAMI EVELIN ROCHA DE MEDEIROS

Junior Data Engineer | Data Analyst | Business Intelligence
Araguaína, Tocantins, Brazil
+55 63 99275-4343 | suamirochati@gmail.com
LinkedIn | GitHub | Portfolio

PROFESSIONAL SUMMARY

Data Engineer with strong foundation in Systems Analysis and Development and specialization in Data Analytics. Proven experience building scalable ETL/ELT pipelines, distributed processing of large data volumes, and developing business-oriented analytical solutions. Proficient in cloud-native technologies, complex workflow orchestration, and modern data architecture implementation. Committed to delivering high-performance solutions that drive data-driven strategic decisions. Passionate problem-solver with hands-on approach to tackling complex operational challenges.

TECHNICAL SKILLS

Data Engineering: ETL/ELT Pipelines, Data Warehousing, Data Lake Architecture, Data Modeling, Data Quality, Data Governance, Batch Processing, Stream Processing, Workflow Orchestration
Programming Languages: Python (Advanced), SQL (Intermediate), Bash Scripting.
Frameworks & Libraries: Pandas, NumPy, PySpark, PyTorch, SQLAlchemy, Apache Spark.
Big Data & Streaming: Apache Spark, Apache Kafka, Apache Airflow, Hadoop Ecosystem
Cloud Platforms: AWS (S3, Glue, Redshift, Lambda), Azure (Data Factory, Synapse), Google Cloud (BigQuery), Databricks. **DevOps & Infrastructure**: Docker, Git/GitHub, CI/CD Concepts, Infrastructure as Code, Container Orchestration, Business Intelligence: Power BI, Grafana (learning), Metabase (familiar), Data Visualization, Dashboard Development, KPI Tracking, Excel. **Advanced Databases:** PostgreSQL, MySQL, Amazon Redshift, Google BigQuery, Data Warehouse Design. A**I Tools:** Proficient in leveraging AI assistants for code development, pattern recognition, and accelerating insight generation

PROFESSIONAL EXPERIENCE

Mobile Development Intern | Branvier
Belém, Pará, Brazil | February 2024 – October 2024
Developed and implemented analytical pipeline for mobile application metrics monitoring, processing data from over 50,000 monthly active users
Structured dimensional model for user behavior analysis, reducing report generation time by 35%
Automated operational data collection and transformation processes, eliminating 15 hours of weekly manual work. Implemented interactive Power BI dashboards for critical KPI monitoring, supporting product decisions.
Collaborated with cross-functional team in agile environment, participating in bi-weekly sprints and Scrum ceremonies. Key Results: Delivered 10+ analytical features with direct product impact, Reduced recurring error rate by approximately 20% through automated data validation
Supported 8 successful releases based on data-driven insights.

DATA ENGINEERING PROJECTS

ETL Pipeline – Epidemiological Analysis of Dengue in Brazil
Ongoing Development | 2025
Architected and implemented robust ETL pipeline for public health epidemiological data processing, handling datasets with 100,000+ historical dengue case records.
Responsibilities & Achievements:
Developed Python modules for automated data ingestion from government public sources
Implemented data cleaning and normalization routines, handling inconsistencies and missing values in complex datasets
Designed optimized analytical dimensional model for temporal and geographical aggregation queries
Executed complex SQL transformations to generate epidemiological metrics and trend indicators
Created analysis-ready datasets for consumption by visualization and statistical analysis tools

Tech Stack: Python, Pandas, SQL, NumPy, Data Transformation
Repository: github.com/surocham/dengue_no_brasil
Machine Learning Pipeline – Handwritten Digit Classification
Practical Project | 2025
Built end-to-end data pipeline for deep learning model training and deployment, processing MNIST dataset with 70,000 images.
Responsibilities & Achievements:

Developed automated data preparation pipeline with quality validation and transformations
Implemented convolutional neural network model using PyTorch with accuracy exceeding 95%
Optimized training process, reducing execution time by 40% through efficient batch processing
Automated inference workflow for batch classification of new samples

Tech Stack: Python, PyTorch, NumPy, Deep Learning, Computer Vision
Orchestrated Pipeline with Apache Airflow
Practical Project | 2025
Developed data pipeline orchestration architecture using Apache Airflow in containerized environment.
Responsibilities & Achievements:

Created complex DAGs for multi-stage ETL workflow orchestration with dependency management
Implemented Docker containers for environment isolation and pipeline portability
Configured schedulers for automated job execution at defined intervals
Developed retry strategies and error handling to ensure pipeline resilience
Simulated real-world scenarios for analytical data ingestion, transformation, and loading

Tech Stack: Apache Airflow, Docker, Python, Workflow Orchestration, DAG Design

EDUCATION
Postgraduate Specialization in Data Analytics
FIAP | 2025
Advanced specialization in enterprise analytical solutions, applied machine learning, big data analytics, and modern data engineering. Focus on scalable architectures and data-driven decision making.
Key Subjects: Big Data Analytics • Machine Learning for Business • Analytical Data Engineering • Advanced Analytics Solutions • Cloud Data Platforms
Associate Degree in Systems Analysis and Development
UNINTER | 2022 – 2025
Complete training in software development, system architecture, data modeling and persistence, software engineering, and advanced programming.
Key Subjects: Data Modeling • Software Engineering • Advanced Programming • Database Systems • System Architecture

CERTIFICATIONS & COMPLEMENTARY COURSES
Data Engineering Fundamentals – Data Science Academy
Databricks Lakehouse Fundamentals
IBM Data Engineering Fundamentals
Power BI for Business Intelligence and Data Analytics
Data Science with Python – Santander Open Academy

BEHAVIORAL SKILLS
Effective Communication • Teamwork • Complex Problem Solving • Analytical Thinking • Continuous Learning • Adaptability • Attention to Detail • Time Management • Results-Oriented • Cross-Functional Collaboration • Quick Learner • Hands-on Approach • No-excuses Mindset

LANGUAGES
Portuguese: Native
English: Intermediate (Advanced technical reading, conversational proficiency)

ADDITIONAL INFORMATION
Availability: Flexible for night shift positions
Work Style: Proactive problem-solver with firefighter mentality – focused on preventing issues before they escalate
Learning Approach: Hands-on learner who thrives on understanding complex systems through practical experience
AI Proficiency: Experienced in using AI tools to accelerate development, automate analysis, and extract insights from data