

Roberto Moreira Diniz

Data Engineer

City: Brasilia, Brazil • Phone: +55(61) 98234-0088 • LinkedIn: @s33ding • Email: robertomdiniz@protonmail.com • Github: @s33ding

About Me

I am an accomplished data professional with nearly three years of experience in data engineering, including two years in a full data engineer role. My expertise includes crafting robust data pipelines, implementing efficient data lakes, and orchestrating seamless data integration within cloud environments. I have honed my skills in PySpark, Exploratory Data Analysis, serverless architecture, AWS Cloud services, Docker, and Linux. I actively manage data quality routines and implement automated testing to ensure data integrity. Currently, I am pursuing AWS Developer certification and studying for the KCNA exam while continuing my education in Machine Learning.

Professional Experience

Data Engineer (Full-Time)

Jul, 2022 - Present

Ti.Saúde

Recife, Brazil

- Create data pipelines
- Maintain and build data storage systems on the cloud to be used by other areas of the company
- Analyzed and extract insights from the data

Data Engineer (Internship)

Nov, 2021 - Apr, 2022

Ministry of Communications (Government)

Brasilia, Federal District, Brazil

- Support the analysis of problems and technical solutions of data integration on projects related to internet connectivity in the national territory

Education

Bachelor's degree in Data Science and Artificial Intelligence

Feb, 2021 - Present

The Higher Education Institute of Brasilia (IESB)

Technical Skills

Programming: Python, SQL, Bash

Cloud: AWS (Boto3, Lambda, Glue, S3, EC2, AWS CLI, Terraform, VPC, RDS, Secret Manager, IAM, SSM, QuickSight, AWS Lake Formation, Amazon Athena, EventBridge, SNS, Rekognitio, RedShift, DynamoDB)

Databases: PostgreSQL, MySQL, Redis, MongoDB, Hadoop

Data Analysis: Pandas, EDA, Probability, Statistics, Sympy, Scipy, Itertools, NumPy

Data Processing: Docker, Git, PySpark, Shell scripting, Regex, Vim, Airflow, Crontab

Machine Learning: Scikit-learn, MLlib (Spark), spaCy

Others Flask, Django, PowerBI, Plotly, MetaBase, Linux Servers

View certificates: (Link)

Idioms

English: TOEFL (Oct, 2020) - View certificates: (Link)

English: Casa Thomas Jefferson (Sep, 2017) - View certificates: (Link)