

# 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 a data engineer with three years of experience building scalable data pipelines, managing data lakes, and integrating cloud solutions. Skilled in PySpark, AWS, Docker, and Linux, I focus on data quality, automation, and efficiency. I hold the AWS Developer Associate and KCNA certifications, and am currently preparing for the CKA and AWS DevOps Engineer Professional exams.

### Professional Experience

---

#### Data Engineer (Full-Time)

Jul 2022 - Present

*Ti.Saúde*

*Recife, Brazil*

- Create data pipelines
- Maintain and build data storage systems in the cloud to be used by other areas of the company
- Analyze 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

*Instituto de Educação Superior de Brasília (IESB)*

#### Bachelor's degree in Architecture and Urbanism

Graduated on September 16, 2020

*Centro Universitário de Brasília (UniCEUB)*

### Technical Skills

---

- Programming: Python, SQL, Bash, Java
- Cloud & Infrastructure: AWS (Boto3, Lambda, Glue, S3, EC2, AWS CLI, CloudFormation, VPC, RDS, IAM, SSM, QuickSight, Lake Formation, Amazon Athena, EventBridge, SNS, RedShift, DynamoDB, Kinesis, Amazon Q, Bedrock, CloudWatch, KMS, CodeBuild, EKS, ECR, API Gateway, Cognito, Route53, ACM)
- DevSecOps: Docker, Git, CI/CD, Chaos Engineering, Kubernetes, Security, Least Privilege Access, Cryptography, Terraform, Snyk
- Data Engineering: PySpark, Airflow, Kafka, Data Quality
- AI & Machine Learning: Scikit-learn, MLlib (Spark), spaCy, LLMops, AI Agents
- Databases: PostgreSQL, MySQL, Redis, MongoDB, Hadoop, SQLite
- Data Analysis: Pandas, EDA, Inference, Scipy, NumPy
- Monitoring & Testing: Prometheus, Grafana, Pytest
- Web & Visualization: Flask, Django, PowerBI, Plotly, MetaBase, Streamlit
- Systems: Linux Servers

**Certificates:** View my certifications

### Languages

---

**English:** TOEFL (Oct 2020) - View certificate: (Link)

**English:** Casa Thomas Jefferson (Sep 2017) - View certificate: (Link)