Luyanda Mncube

luyandamncube.github.io
luyanda.mncube@outlook.com | 061 116 7195

EDUCATION

UNISA - ONGOING
Bsc CompSci & Mathematics
Expected grad. 2025 | JHB

WETHINKCODE
2018 | JHB

CERTIFICATION

MCSA: Database Development
DP-203: Azure Data Engineer Associate
Certified SAFe ®Agilist

PROJECTS

AI-DRIVEN NOTIFICATIONS
Thames Water and Sand Technologies
A DS/DE feature that identifies bursts within a digital twin of a water network.

DIGITAL WATER BALANCE
A set of DE features to account for all water movement through a defined volume in various areas including S/N London, S/N Oxford, Guildford & SWA.

ACHIEVEMENTS

Water Industry Awards UK - Nomination
Smart Water Networks award
• AI-Driven Notifications

Architected Data Warehouse Solution
• Successfully implemented Azure-based warehousing solution, for daily PostgreSQL data backup.

Architected Bronze/Silver data ingestion
• Successfully implemented Azure-based data ingestion of various complex data sources, including telemetry data, & Geospatial data. Data used to estimate daily water leakage.

SKILLS

Azure Data Factory:
Azure SQL • MSSQL • PostgreSQL
• ADLS • Databricks

Big Data/Hadoop ecosystem:
Apache NiFi • Apache Spark
Apache Kafka • Hive • HDFS

PERSONAL SUMMARY

Eager to leverage my 5 years in Data Engineering, specializing in Azure's Data ecosystem and Python development. I have built data-driven applications within the water utilities, consulting and competitive sports analytics industries.

Excited about crafting data-driven solutions, my contributions have significantly improved business-critical processes and serve as an essential daily data source in use by hundreds of data scientists, stakeholders and the business at large.

EXPERIENCE

SAND TECHNOLOGIES (PREV. EXPLORE AI) | Data Engineer
Oct. 2021 - Present | JHB (Remote)
- I am currently in the Clean Water Operations team. We develop & maintain 8 products which service hundreds of users (engineers, data scientists stakeholders) across Thames Water Utilities.
- Data features - functionality to view daily reports, validate figures, configure assets & various other tools serving 150+ water operations specialists.
- Databricks - Cluster monitoring and troubleshooting of the Spark.
- ADF - Created ETL Pipelines to move data from different sources ensuring timely data delivery for business intelligence.
- CI CD - Created Build and Release for multiple projects (modules) in Production environment using Azure Pipelines

Tech Stack: Azure Data Factory, Databricks, Spark, MSSQL, PostgreSQL, Python

YEBOFRESH | Engineer
Mar. 2021 - Oct. 2021 | Airport Industria, CPT (Remote)

- I worked as part of the Engineering team to build features in a fast-paced environment to improve core data sets & develop data-driven features of mobile e-commerce apps for tracking customer purchases.

Tech Stack: MSSQL, Flutter, Firebase

NTT (PREV. DIMENSION DATA) | Data Engineer
Feb. 2019 - Mar. 2021 | Bryanston, JHB

- I worked as part of the Race Center Team. We were responsible for delivering real-time analytics, data warehousing, monitoring and support to the UCI World Tour (Tour De France & Paris-Roubaix) and UCI esports (Zwift).
- Maintained data pipeline up-time of 99.9% ingesting streaming and transactional data across 4 different primary data sources
- Developed maintainable, efficient batch/stream analytics features.

Tech Stack: Azure VM, Bash, Cron, Docker, Ansible, Terraform, MSSQL, Apache NiFi, Python.

KEY SKILLS & COMPENTENCIES

- Experience in Developing Spark applications using Spark
- Good understanding of Big Data Hadoop with specialization in streaming data
- Experience in Database Design and development using SQL Server & pgAdmin
- Excellent communication skills, including the ability to convey analytics insights effectively to both IT and business audiences.
- Expertise in various phases of project life cycles (Design, Analysis, Implementation and testing).