Lerato Matamela

Johannesburg, South Africa | leratomatamela1@gmail.com | +27 78 430 4420 | GitHub: https://github.com/Lerato-leo |
LinkedIn: www.linkedin.com/in/lerato-matamela-b7697118a/ | Portfolio: https://personal-portfolio-nine-iota-32.vercel.app/

EDUCATION

Richfield Graduate Institute of Technology   **Johannesburg**
Bachelor of Science in Information Technology   *Final Year: 2026*
- o **Specialisation:** Data Science & Machine Learning
- o **Related Coursework:** Data Structures, Database Systems, Data Science, Machine Learning, Statistics, Cloud Computing, Software Engineering, Information Systems and Calculus.

EXPERIENCE

CAPACITI   **Johannesburg**
Data Engineer Associate (Contract)   *Nov 2025 – Present*
- Design and implement scalable ETL pipelines and data workflows in cloud-based environments.
- Develop and maintain data pipelines supporting enterprise analytics, reporting, and automation initiatives.
- Optimize data ingestion, transformation, and storage processes for performance and reliability.

Project Y   **Sandton, South Africa**
Engineering Cohort Member   *Nov 2025 – Present*
- Selected through a highly competitive, multi-month evaluation process for South Africa's elite tech talent program.
- Complete employer-aligned technical training in data engineering, cloud computing, and modern software development.
- Work on hands-on, production-focused projects emphasizing scalable systems, automation, and real-world problem solving.

PROJECTS

Healthcare Insurance Cost Prediction
Machine Learning & Data Science   *Oct 2025 – Dec 2025*
- Built a Gradient Boosting regression model achieving 83.8% R^2 accuracy for cost prediction.
- Developed an interactive Streamlit application with real-time predictions, scenario analysis, and user history tracking.
- Performed feature importance analysis revealing smoking status as the dominant cost driver.

GitHub Wrapped
Data Engineering | Analytics | Automation   *Nov 2025 – Dec 2025*
- Built an automated data analytics project that processes GitHub activity data to generate a personalized "GitHub Wrapped" summary.
- Designed data ingestion, transformation, and aggregation workflows to extract repository activity, commit frequency, and contribution trends into meaningful insights.

CERTIFICATIONS

- Google Data Analytics Professional Certificate
- IBM – Python Project for Data Engineering
- IBM – Introduction to Data Engineering
- AWS – Generative AI with Large Language Models
- Microsoft Power Platform

SKILLS

Programming & Data: Python, SQL, JavaScript, ETL Pipelines, Data Modelling, Data Analysis
Tools & Platforms: Git, GitHub, Node.js, Power Apps, Streamlit, Jupyter Notebook
Soft skills: Problem Solving • Analytical Thinking • Technical Communication • Team Collaboration • Time Management
Currently Learning / Developing: TensorFlow, PyTorch, Apache Airflow, Spark

REFERENCES

- Available on request