Simon Mufara

simon.mufara1@gmail.com • +27620552382 • 20 station road, Cape Town, Western Cape, 7700, ZA • [simon-mufara](#)

SUMMARY

Bioinformatics & Healthcare Data Scientist with a background in medical microbiology and public health and over 3 years of experience. Experienced in variant calling, multi-omics data integration, and computational modeling for cancer and infectious disease research. Skilled at delivering reproducible, data-driven solutions.

Education

Honours Bioinformatics
Jan '25 - Dec '25

University of Cape Town

Minor: Genomic & Systems Biology, Phylogenetics, Protein Annotation, Machine Learning, Gene Regulation, Sequence Alignment Algorithms, Biological Ontologies

Relevant Coursework: Developing a transformer-based multi-omics drug response prediction model for African breast cancer patients

- Applied **transfer learning** to fine-tune models on African-specific TNBC and HER2+ cohorts, achieving improved accuracy vs. baseline models
- Integrated 5+ datasets (BCaPE, H3Africa, GEO, TCGA, GDSC) covering 1,000+ samples acros**s gene expression, mutations, and methylation**

Bachelor BHSc
Jan '20 - Dec '22

University of the Witwatersrand

Minor: public health, epidemiology, and health policy, applying data analysis to evaluate healthcare trends

Relevant Coursework: Improved projects using Excel, SPSS, and R, developing skills in quantitative research and evidence-based decision-making

GPA: 7.2

Volunteer Experience

Voluntary Facilitator
Feb '25 - Present

EH!Woza

- Facilitated 5+ interactive science workshops for township students on **HIV/TB awareness**
- Designed and delivered monthly **youth engagement** sessions, reaching 50+ students per session

Work Experience

Variant Calling Freelancer
Mar '25 - Jun '25

Drug Discovery & Drug Development (H3D) | Cape Town

- Performed end-to-end variant calling on 5+ **Mycobacterium tuberculosis clinical isolates** using bcftools, samtools, and Snakemake
- Annotated variants with SnpEff, predicting functional consequences and highlighting **drug resistance-associated mutations**
- Utilized **data visualization** tools and generative AI to present bioinformatics findings and support healthcare decision-making
- Published full pipeline and documentation is available on **GitHub with cleaned QC** which achieved a score of 98%

Global Health Equity Intern
Jan '24 - Dec '24

Johnson & Johnson Innovative Medicine | Midrand

- Presented strategic vision and progress reports directly to the **CEO and senior leadership,** securing executive sponsorship and resources for sustainability
- Supported 3 healthcare access programs across Sub-Saharan Africa, contributing to **policy recommendations** impacting ~10,000 patients
- Developed 5+ dashboards and reports to communicate findings to internal stakeholders and external partners, which leveraged **business operations**
- Collaborated with cross-functional teams of 10+ staff to integrate data insights into **strategic planning,** which improved healthcare management

Early Childhood Development - Oncology & Nephrology Intern   Jan '23 - Dec '23

Reach For A Dream | Johannesburg

- Supervised **pedesterian wards** at Chris Hani Baragwanath Hospital and Charlotte Maxeke Johannesburg Academic Hospital working with over 30+ health practitioners
- Managed 200+ **Electronic Medical Records** using Redcap, improving data accessibility
- Delivered 10+ structured **health education** sessions per month to children and families

KEY SKILLS

- **Programming & Data Analysis**: Python (pandas, scikit-learn, PyTorch, transformers), R (tidyverse, Seurat, DESeq2), SQL, Bash/Linux, Jupyter
- **Bioinformatics & Genomics Tools**: bcftools, samtools, Snakemake, SnpEff, VEP
- **Data Visualization & Analytics**: Python (matplotlib, seaborn), R (ggplot2), Excel, Power BI
- **Collaboration:** Coordinated with cross-functional teams and researchers to integrate data insights
- **Communication**: Delivered presentations and reports to academic, industry, and executive stakeholders
- **Project Management**: Led research workflows and multi-step analyses ensuring reproducibility and quality

Awards

Top 10 Most Employable Graduate   Dec '24 - Dec '24

GradStar | Johhanesburg

Out of 500 graduates who participated in the awards, it was narrowed to the top 100 who went to a 3-day workshop event. I performed well and earned myself a trophy and a spot in the top 10.

Power BI Essential Training   Aug '25 - Aug '25

Linkedin Learning | Remote

learned data modeling, DAX formulas, interactive dashboards, and visualization best practices