PHEMELO SEBOPELO

063 313 6167 | phemelosebopelo14@gmail.com | LinkedIn | GitHub
Pretoria, Gauteng

SUMMARY

BSc Information Technology student with strong academic performance and hands-on experience in data science, software development. Skilled in Python, SQL, JavaScript, C++, HTML, CSS, and Excel data analysis. Practical exposure through academic, professional, and volunteer projects, including telemedicine systems, PropTech solutions, and website development. Recognized for problem-solving, adaptability, and teamwork, with ongoing training in IBM and AWS Data Science certifications. Seeking an internship as a Data Scientist, Data Analyst, or Machine Learning Engineer to apply technical skills, grow professionally, and deliver impactful results.

PROJECTS

Company Bankruptcy Prediction – Python/Scikit-learn Machine Learning Pipeline
- Used Python, Pandas, and Scikit-learn for data cleaning, feature engineering, and building predictive models on bankruptcy financial data.
- Applied Logistic Regression and Random Forest to classify bankruptcy risk, using pipeline, imputation, and model evaluation techniques for robust results.
- Generated visualizations and evaluation metrics, showing model accuracy and important features; delivered clear insights for financial decision support.

South African Crime Data Analysis – Python/Pandas Data Science Project
- Python and Pandas: Used for loading, cleaning, and analyzing South African crime data, enabling efficient statistics calculation and CSV export for easy sharing.
- Matplotlib: Used for creating visual bar charts of total crimes, rates per 100k, and top crime categories, making insights simple and clear for recruiters.
- CSV files and scripting: Used to automate raw data processing, population normalization, and output generation, ensuring reproducible, readable analysis results for decision-makers.

Netflix Data Analysis – Python Data Science Project
- Used Python for data cleanup, manipulation, and visualizations; streamlined raw Netflix data into clear, actionable insights for business trends and content analysis.
- Leveraged pandas and matplotlib libraries to process, filter, and plot data; enabled easy exploration of user preferences and viewing patterns for report generation.
- Employed Jupyter Notebook for interactive coding, documentation, and sharing; made findings reproducible and accessible for team collaboration and stakeholder presentations.

EDUCATION

BSc Information Technology   February 2023 - December 2025
Richfield Graduates Institute of Technology
- Assisted with campus tours during Open Days.

SKILLS

- **Languages & Skills:** Python, HTML, CSS, JavaScript, C++, SQL
- **Tools & Frameworks:** Git, GitHub, Visual Studio/Code, Jupyter Notebook, Matplotlib, Pandas, Microsoft Office
- **Soft Skills:** Problem-solving, Teamwork, Leadership, Communication, Adaptability. Growth-oriented

CERTIFICATIONS

- IBM Data Science Badge - IBM   (In Progress)
- AWS Data Science Badge - Amazon   (In Progress)
- Microsoft Learn Copilot & M365 - Microsoft   September 2024