Kgathola Puka

0715763427 ● Johannesburg, GP 2001 ● kgatholapuka@gmail.com

● [Linkedin](#)

● [GitHub](#)

Python | SQL | PowerBI

Data Scientist familiar with gathering, cleaning and organizing data for use by technical and non-technical personnel. Advanced understanding of statistical, algebraic and other analytical techniques. Highly organized, motivated and diligent. Well-qualified Data Scientist experienced working with vast data sets to break down information, gather relevant points and solve advanced business problems. Skilled in predictive modeling, data mining and hypothetical testing.

CORE COMPETENCIES
Communication Skills | Organizational Skills | Teamwork
Machine Learning | Statistics | Data Science | Problem Solving

WORK EXPERIENCE

OpenServe South Africa - *Johannesburg, GP*   Sep 2021 – Dec 2021

Data Analytics

- **Utilized advanced querying, visualization and analytics tools to analyze and process complex data sets.**
- **Worked with stakeholders to develop quarterly roadmaps based on impact, effort and test coordination's.**
- **Set up SQL database on cloud servers to store client data for query analysis.**
- **Created and implemented new forecasting models to increase company productivity.**
- **Assessed accuracy and effectiveness of new and existing data sources and data analysis techniques.**
- **Created a New Automated Dashboard using PowerBi with updated weekly.**

Explore Data Science Academy - *Johannesburg, GP*   Jan 2021 – Dec 2021

Data Science Intern

- Cleaned data and monitored where errors are coming from using SQL and Python.
- Used SQL to aggregate, sort data, and do calculations.

- Wrote SQL JOIN statements, Sub-Queries, organized data, defining table relationships using primary and foreign keys, and implemented table normalization up to 3rd normal form.
- Worked with SQL Views, Optimized SQL queries, data imputation, string manipulation, updating and cleaning tables, and data transformations.
- Validated data accuracy identified and removed duplicates and standardized all data.
- Descriptive data analysis using measures of central tendency and measures of spread.
- Inferential data analysis using Correlation, Regression, and Analysis of Variance.
- Text data analysis using Natural language processing on climate change twitter data.
- Data interpretation and report building using PowerBI.
- Understanding of structured, unstructured data, clob, and blob data types.
- Wrote Python scripts for error checking, data imputation, and data sorting.
- Worked on a project to predict the estimated price of apples using advanced regression techniques..
- Used Natural Language Processing on Twitter data to predict whether a person believes in climate change or not.
- Built a movie recommender system inspired by Netflix, and used it to predict what a user will rate a movie they haven't seen.

EDUCATION

Data Science – Information System – Explore Data Science Academy, *Johannesburg*

Data Engineering – information System – Explore Data Science Academy, Johannesburg

PROFESSIONAL SKILLS

- Mastery of Microsoft Office programs (Word, Excel, and PowerPoint).
- Comfortable working with video conferencing software like Zoom, Google Meet, Slack, and google docs.
- Excellent communication skills with a focus on team building and customer relations.
- Outstanding organizational, multitasking, and problem-solving abilities.

LANGUAGE COMPETENCIES

- English: fluent (speaking, reading, writing).
- Afrikaans: Basic (speaking, reading, writing).
- Sepedi: fluent (speaking, reading, writing).

REFERENCES

- Available on request.