

# LINDA KARANI

## Geospatial Data Scientist

I am a dynamic data scientist with 5 years of experience, specializing in geospatial analysis, spatial modeling, and data automation. My expertise lies in transforming complex spatial datasets into actionable insights, leveraging advanced python and R libraries like GDAL, GeoPandas, Shapely and raster. I have a deep understanding of Earth Observation data (including satellite and aerial imagery), spatial regression techniques, and the application of statistical models to solve real-world challenges.

A key strength of mine is developing innovative methodologies for spatial data analysis, uncovering patterns and correlations that inform strategic decision-making. I am proficient in automating data workflows in Python and R, allowing for efficient processing of large geographic datasets. My approach combines problem-solving with technical rigor, ensuring that my work not only meets but exceeds client expectations.

I am dedicated to making a lasting impact by applying advanced spatial models and innovative geospatial solutions to tackle real-world challenges, with a particular focus on humanitarian and development contexts.



## EDUCATION

2021  
|  
2024


### Strathmore University

Masters of Science in Data Science and Analytics  Nairobi, Kenya

Relevant units: Advanced machine learning concepts, Advanced statistical analysis, Geospatial analysis, Data visualization and communication, Database management


2022

### ETH Zurich

Dissertation: Importance of reproducibility in advancing Sustainable Development Goal(SDG) 6 in R : A comparison of statistical models  Zurich, Switzerland

2014  
|  
2018

### Strathmore University


Bachelor of Business Science in Financial Economics  Nairobi, Kenya



## CERTIFICATIONS

2024

### IBM

Data Engineering & Data science  Remote


Relevant units: Data Management, ETL, Apache, Databases, Data Warehousing, Distributed computing systems, SQL. Generative AI

## CONTACT INFO

 [karanilinda@gmail.com](mailto:karanilinda@gmail.com)

 [github.com/KaraniLinda](https://github.com/KaraniLinda)

 [in LinkedIn](#)

 +254 708054713

## DATA TECH STACK

**Database/Server** : MySQL, PostgreSQL

**Programming languages** : Python, R programming

**Frameworks/libraries** : NumPy, Pandas, matplotlib, Scikit-learn, ggplot, tidyverse, dplyr, leaflet, caret

**Data visualization/Reporting** : R Markdown, Quarto, Power BI, Tableau , Google Data Studio, Excel

**Geospatial packages**: Uber-H3(Hexagon), ARCGIS, R(raster,tmap), Python (geopandas, rasterio, xarray, shapely)

**Cloud computing** : Google cloud platform

Git and Version Control

## LANGUAGES

English : Native

French : Limited working proficiency

2024

## Esri

Spatial data science

📍 Remote

Relevant units: spatiotemporal clustering, suitability analysis, hotspot analysis, object detection using deep learning, spatial regression and interpolation



## PROFESSIONAL EXPERIENCE

November,  
2022

|  
Date

### Data scientist

World data lab

📍 Nairobi, Kenya

- Led the comprehensive data collection, analysis, and forecasting for multiple high-impact projects including the in-depth analysis of youth labor force survey data in collaboration with the Mastercard Foundation. Executed the project for Kenya, Rwanda, and Ghana, leading to the establishment of a youth unemployment clock providing real-time insights on jobs. Utilized advanced statistical modelling techniques such as linear regression and time series models to forecast future labour market trends
- Developed and deployed predictive models utilizing high-resolution satellite imagery, drone data, and thermal imagery, alongside demographic data, to create detailed maps of aggregate spending for European countries. These insights helped clients pinpoint optimal retail expansion sites at the grid and hexagon level (Uber H3 - Res 9).
- Leveraged machine learning algorithms, such as Random Forest and regularized regression, combined with geospatial covariates to disaggregate spending data at the grid level, achieving a prediction accuracy of 86%.
- Designed and automated spatial data analysis pipelines using Python and R, streamlining workflows and ensuring efficient handling of large geospatial datasets.
- Organized and executed student hackathons with participation from over 100 students and in collaboration with the top 3 universities in Kenya, resulting in the development of innovative solutions for real-world problems and fostering collaboration within the academic community.
- Conducting rigorous data quality assurance checks, ensuring the continuous enhancement of models with the latest and highest-quality data.
- Demonstrated expertise in the in-depth analysis of household and census survey data.
- Delivered high-impact, actionable insights to clients, focusing on providing data-driven solutions that directly informed strategic decision-making and drove meaningful results.

April,  
2021  
|  
November,  
2022

### Senior associate- Data analysis and business modelling

Sanergy

📍 Nairobi, Kenya

- Spearheaded the development of spatial models, leveraging GIS tools and satellite imagery to identify optimal locations for non-sewered sanitation services. Collaborated on data management and analysis to pinpoint high-potential urban areas for future scale-up
- Applied spatial regression and clustering techniques to assess sanitation coverage, providing critical insights that guided future infrastructure investments.
- Developed several robust quantitative and qualitative data collection tools for citywide inclusive sanitation planning to determine market share and predict an appropriate mix of sanitation products based on key indicators and metrics.
- Analyzed data collected by the Cost and Climate in Urban Sanitation (CACTUS) project led by the University of Leeds and supported by the Bill and Melinda Gates Foundation and used the findings to create a set of improved data collection tools and manuals to facilitate widespread contribution to a database of benchmark cost estimates to inform sanitation investment decisions.
- Conducted data analysis for a UNICEF implemented project to inform progress towards safely managed sanitation
- Communicated complex data-driven insights effectively to senior leadership and external stakeholders, supporting informed decision-making processes.
- Produced detailed reports and technical documentation, including manuals, ensuring sustainable handover of the tools and insights developed during the projects.
- Proactively managed stakeholder relationships, providing advisory services and capacity-building training to enhance project outcomes.

September,  
2019  
|  
March,  
2021

### Finance/Data analyst

Sanergy.

📍 Nairobi, Kenya

- Identified key organizational metrics, cleaned and analyzed data from all departments and presented results on Tableau improving accuracy of results by 35%.
- Aggregated data across all departments that was previously in multiple different sources to create an interactive organization wide dashboard based on key KPIs that led to increased efficiency in operations and improved decision making and that resulted in a 30% reduction in redundancy.
- Using data analytics to identify business needs and requirements and deliver data-driven recommendations and reports to executives and stakeholders.
- Supported teams to understand, monitor and control key business metrics to achieve success and base decisions on data.
- Prepared financial statements and other financial reports and reported performance to senior management and the finance board



## REFERENCES

References available upon request

This resume was made with the R package [pagedown](#).

Last updated on 2024-12-12.