# DATA SCIENTIST

## FRIDAH GECHEMBA MACHANI

Potsdamerstrasse 70, 14974 Ludwigsfelde, Germany
**Phone:** +4915216878666

**Email**: machafrida@gmail.com
**GitHub**: https://github.com/machaniG

**LinkedIn**: www.linkedin.com/in/fmachani
**Website**: https://machanig.github.io/

---

## CORE COMPETENCIES

Data Analysis, Data Visualizations, Statistics, Machine Learning, Statistical Modeling, Customer Segmentation.

Cross-Functional Collaboration, Project Management, Analytical Problem-Solving, Adaptability, Communication.

---

## TECHNICAL SKILLS

Python, Scikit-Learn, XGBoost, TensorFlow, SQL, Power BI, Tableau, Excel, R.

---

## PROFESSIONAL EXPERIENCE

### Parental Break with Continuous Upskilling
*03/2024 – 06/2024*

- Completed advanced training in data science, machine learning, such as **IBM Data Science** and **Machine Learning Specialization** by Stanford University & DeepLearning.AI.

- Developed strong time management, multitasking, and prioritization skills as I had to balance learning with full-time childcare and household responsibilities.

- Enhanced adaptability, resilience, and problem-solving through handling sudden changes and navigating everyday parenting challenges.

### Doctoral Researcher | Max Planck Institute of Molecular Plant Physiology
*08/2020 – 09/2023, Potsdam Germany*

- Automated raw data cleaning, transformation, and statistical analyses using R scripts, reducing analytics time by 40% and enhancing productivity.

- Built easy-to-use models to estimate values like nutrient levels based on observable data, reducing the need for time-consuming and costly tests; enabled faster decision-making across 1,000+ cases.

- Optimized fertilizer use, reducing required amounts by 30% without compromising yield, ultimately cutting costs and supporting cost-effective farming.

- Analyzed how factor combinations influence KPIs like height and yield, delivering actionable insights on the exact combinations that improve yield by 10%.

- Segmented 30,000+ time-based metabolic data points, uncovered nutrient-specific

patterns and early stress signals, enabling timely interventions to prevent yield loss.

**Doctoral Researcher |** Universität Münster IEB – Münster, Germany
*10/2019 – 05/2020, Nairobi Kenya*

- Boosted seed germination rates by up to 80% by identifying optimal exudate collection time through time-series analysis, enabling better insights into host–parasite interactions.

**Population Geneticist |** World Agroforestry Centre
*01/2017 – 04/2018, Nairobi Kenya*

- Identified genetically identical trees, reducing unnecessary conservation efforts by 25% and saving up to 30% in maintenance costs.

- Improved breeding efficiency by 40% by clustering 200+ trees based on 50,000+ genetic features, enabling precise selection of diverse, high-performing cultivars.

---

**—**

**EDUCATION**

**Data Analytics with python** | ReDi School of Digital Integration
*03/2024 – 06/2024, Berlin, Germany*

Data analysis with Python: Data manipulation (Pandas), filtering, grouping, exploration, and storytelling.

Data analysis with SQL: Querying, data analysis, visualization, and dashboard creation.

Basic Statistics such as, descriptive statistics, distributions, correlation, and sampling.

**Doktor der Naturwissenschaften (PhD) |** University of Potsdam
*08/2020 – 27/09/2023, Potsdam, Germany*

**Master of Science Biotechnologie |** Kenyatta University
*May 2015 - 08/2018, Nairobi, Kenya*

**Relevant Coursework**: Statistics (basic statistics, regression, probability distributions), scientific data analysis, research methodology, inferential statistics.

---

**—**

**PROFESSIONAL DEVELOPMENT**

**SQL, Business Intelligence, Machine Learning & AI** | DataCamp and Coursera
*05/2024 – present, Online*

---

**—**

**LANGUAGES**

English (Muttersprache), Deutsch (B1)