

## Training in Data Science for Health: 2025 WASHA Takwimu Short Course Series

### Ignite Your Data Science Journey with our upcoming Short Courses...

The University of KwaZulu-Natal (UKZN), in partnership with Harvard University School of Public Health (HSPH) and Heidelberg Institute of Global Health (HIGH), is pleased to announce the 2025 series of short courses under the WASHA Takwimu initiative, a specialized training grant funded by the National Institutes of Health (U2RTW012140).

**WASHA (Working on Applications for Data Science and Health in Africa) Takwimu**, meaning “Ignite Data” in Swahili, aims to build capacity in data science to address pressing global health and climate challenges across Africa. This programme seeks to develop a cadre of health professionals and scientists equipped with the skills needed to leverage data science in innovative and impactful ways.

The 2025 course offerings provide an immersive learning experience, focusing on both foundational and advanced data science techniques with a focus on health. Courses are designed to bridge the gap between theoretical knowledge and practical applications, with an emphasis on addressing real-world issues in health systems, epidemiology, and food systems, climate change and planetary health.

### Eligibility Criteria

- This programme is designed for individuals with a background in health sciences, biological sciences, statistics, mathematics, computer science, or related fields, with a demonstrated interest in applying data science methods to health, climate resilience, and planetary health challenges.
- Students may choose to take **one or all** courses depending on their interest and prior training/experience.

**We look forward to you joining us on this journey to advance your expertise in health data science!**

### Programme Features and Benefits

1. **Expert Faculty:** Engage with leading experts from UKZN, Harvard, and Heidelberg, bringing extensive experience in data science and global health.
2. **Hands-On Training:** Participate in practical sessions designed to build competence in data manipulation, analysis, and advanced methodologies with relevance to the local context.
3. **Scholarships Available:** Travel scholarships to participate in the short course training programme will be awarded on a competitive basis to **two** applicants from each of the following institutions: **UKZN, Ghana, Uganda, Tanzania and Nigeria**. The scholarship will cover the course facilitation, airport transfers, materials, flights (if applicable), bed and breakfast only for the course duration. The scholarship excludes: Visa costs, vaccine requirements, travel insurance, and daily dinner for the duration of the course.
4. **Networking Opportunities:** Collaborate with a diverse group of health professionals, researchers, and scientists across Africa, fostering a collaborative network for future projects.
5. **Additional candidates from our Spokes, other African institutions and South Africa are welcome to apply.** However, spaces are very limited and these applicants will have to fund their own travel and accommodation. We can co-fund the workshop teaching and conferencing costs.

### Application and Contact Information

Spaces are limited, and interested candidates are encouraged to apply early. Please submit a completed application form by the closing date using the link in the table below.

For more information, please visit our website: <https://www.hsph.harvard.edu/dsi-africa/>. If you have any questions or other inquiries, please contact Mr. Themba Manqele: [Manqelet@ukzn.ac.za](mailto:Manqelet@ukzn.ac.za). Please use the subject heading: Application for Short Courses in WASHA Takwimu Health Data Science: [Name of short course course]

## Course Series for 2025 - 2026

2025	2026
<ul style="list-style-type: none"> <li>• Introduction to R and Python for Data Science</li> <li>• Introduction to Data Science and Epidemiology for Public Health</li> <li>• Data mining for Public Health</li> <li>• Supervised Machine Learning</li> <li>• Unsupervised Machine Learning</li> <li>• Causal Inference</li> <li>• Manuscript writing</li> </ul>	<ul style="list-style-type: none"> <li>• Deep Learning</li> <li>• Health Systems in Data Science</li> <li>• Proposal Development and Grant Writing</li> <li>• Climate change and food systems in Data Science</li> </ul>

### Course 1 in the Series: Introduction to R and Python for Data Science

<b>Description</b>	This intensive course is designed to introduce participants to the fundamentals of R and Python, the two leading programming languages used in data science. With a focus on hands-on learning, this course will cover essential skills for data manipulation, analysis, and visualization in R and Python. By the end of the week, participants will be able to apply these tools to real-world data, especially in the context of public health and other scientific disciplines.
<b>Proposed dates</b>	10 – 14 February 2025
<b>Application Closing Date</b>	10 <sup>th</sup> January 2025
<b>Application Form Link</b>	<a href="#">2025 Introduction to R and Python for Data Science Application Form</a>

### Upcoming courses: Closing Dates and Application form links to be shared in 2025

Course Name	Description	Proposed dates
<b>Introduction to Data Science and Epidemiology for Public Health</b>	This course is suitable for individuals new to both data science and epidemiology. The course covers essential concepts and practical activities to introduce participants to how these fields intersect and are applied in public health.	10 – 14 March 2025
<b>Data mining for Public Health</b>	This course provides an overview of advanced analytical techniques to large and complex datasets to discover patterns, hypothesis formulation, trends, and insights that can improve public health outcomes, resource allocation, and decision-making. It involves using algorithms from fields like machine learning, statistics, and database systems to extract useful information from data. In public health, data mining helps analyse vast amounts of information collected from various sources such as epidemiological surveys, hospital records, social media, and environmental data, enabling public health officials to make evidence-based decisions.	7 – 11 April 2025
<b>Supervised Machine Learning</b>	This course offers a comprehensive exploration of supervised learning techniques for health data science. Participants will learn how to build predictive models, tune algorithms, and evaluate their performance to address real-world challenges in healthcare and public health. The course will cover essential supervised learning techniques such as linear regression, decision trees, random forests, support vector	9 – 13 June 2025

	machines, and neural networks. Through hands-on exercises using health data, participants will apply these methods to tasks like disease prediction based on different data types, risk stratification, and resource optimisation.	
<b>Unsupervised Machine Learning</b>	This course will provide an in-depth understanding of unsupervised learning methods for analysing health data. Participants will explore how to uncover hidden structures, patterns, and relationships in large, complex health datasets. The course covers key techniques such as clustering, dimensionality reduction, anomaly detection, and association rule mining. Through hands-on exercises using real-world health data, participants will apply these methods to practical health data science challenges.	21 – 25 July 2025
<b>Causal Inference</b>	This course will start with the counterfactual model to introduce the concepts of causal reasoning in statistics and epidemiology. It will use both statistical and graphical approaches to provide rigorous introductions to experiments and key quasi-experimental methods. Experimental approaches important in the health sciences will be covered, including parallel-arm, stepped-wedge, and adaptive trials with both individual and cluster units of randomization. Quasi-experimental approaches will include instrumental variable, regression discontinuity, and difference-in-differences analyses. Finally, this course will provide theoretical and practical bases for integrating prediction, and causal understanding and explanation. After completion of this course, students will be able to design and execute experiment in the health sciences and identify and execute quasi-experimental studies in routine healthcare data.	18 – 22 August 2025
<b>Manuscript writing</b>	Principles of scientific writing will be taught. The course will cover organization of scientific papers, presentation of data in graphical and tabular forms, and style. The course is designed for advanced trainees who are beginning to work on a paper for publication. Each section of a paper will be discussed extensively. Trainees will advance and ideally complete a manuscript on a topic related to health systems strengthening or food systems, climate change, and planetary health in SSA for submission to an international or regional scientific journal for publication. Harvard, UKZN and Heidelberg faculty will be available for final external review of the manuscript prior to submission to scientific journals. The instructors will guide the discussion and use the paper to make additional points of constructive criticism, which will serve to illustrate the principles shared. Guidelines for journals and co-author criteria will be discussed. This course will be especially useful for students who are beginning to write up results from their research.	14 – 18 October 2025

Acknowledgement: This training programme is supported by the Fogarty International Center of the National Institutes of Health under Award Number U2RTW012140. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health