Data Science for Public Health workshop
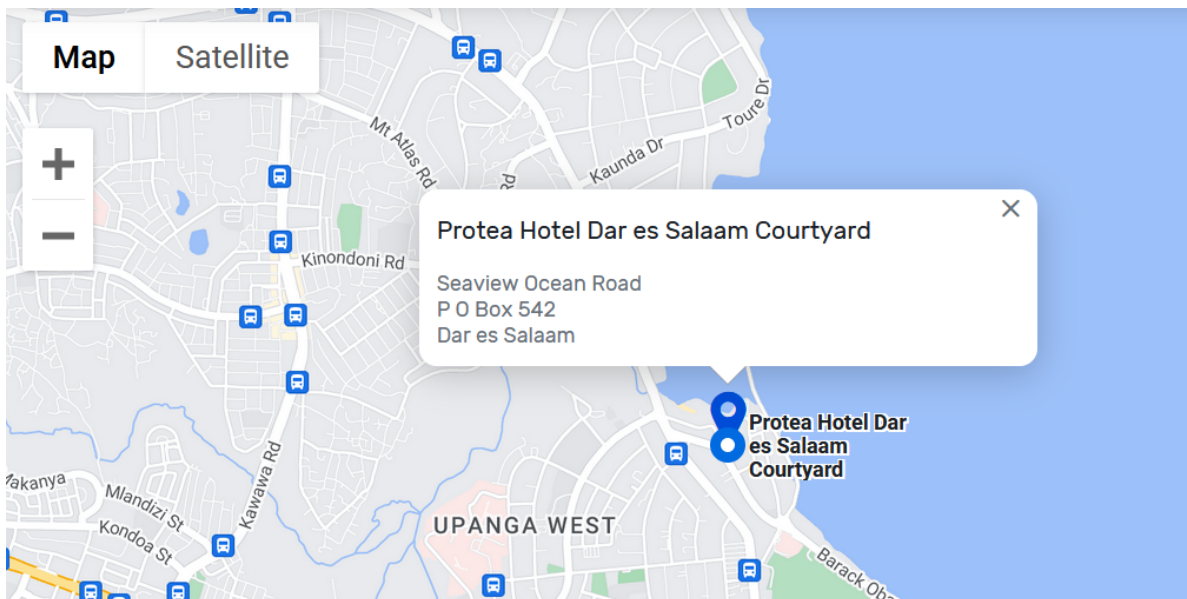
Information

September 26-28, 2022
09:00 - 17:00
Dar-es-Salaam, Tanzania

This workshop targets public health researchers and aims to accompany them in their journey towards Data Science. Looking forward to meeting you all on Monday!

1 Venue

The workshop takes place at **Protea Hotel Dar es Salaam Courtyard** on Seaview Ocean Road. The conference room is **Nelson Hall**.

2 Before the workshop

1. Fill out the online pre-workshop questionnaire

2. If you have not done it yet, please go through the following five steps (about 30 minutes).

   a. Install R

   b. Install and setup RStudio Desktop

   c. Install Quarto

   d. Install the rmarkdown R package

   e. Create a GitHub account

   f. Install GitHub Desktop

> **ℹ Note**
>
> - All software used in this workshop are free.
> - If you have any difficulties with the installation, support can be provided on the first day of the workshop before the first session or during breaks.

> **❗ Important**
>
> For those who have already been introduced to Git, R/RStudio and RMarkdown, note that you need to additionally install Quarto for this workshop.

3 Schedule

3.1 Day 1: Monday September 26th

Table 2: Overview Day 1

Time	Session
08.30 - 09.00	Welcome
	Support for software installation
09.00 - 09.15	Introduction to data science tools
	Overview of objectives for Day 1
09.15 - 10.15	Version control with Git
10.15 - 10.45	☕ Break
10.45 -11.45	Introduction to dynamic documents and Quarto
11.45 - 12.15	Use Quarto with Stata
12.15 - 13.15	☕ Lunch break
13.15 - 14.00	Import external data
14.00 - 15.00	Manipulate data
15.00 - 15.30	☕ Break
15.30 - 16.15	Visualise data
16.15 - 17.00	Share code and collaborate with Git

3.2 Day 2: Tuesday September 27th

Table 3: Overview Day 2

Time	Session
08.30 - 09.00	Welcome
09.00 - 09.20	Introduction to Data Science for Public Health Overview of objectives for Day 2
09.20 - 10.15	Discussion on concepts related to to public health data for decision-making
10.15 - 10.45	☕ Break
10:45-12:15	Malaria practical #1
12.15- 13.15	☕ Lunch break
13.15- 13.30	Prepare presentation of data findings
13.30 - 14.00	Present and discuss data findings
14.00 - 14.30	Interdisciplinary discussion on malaria, dataset, uncertainties, challenges
14.00 - 14.30	Malaria practical #2
15.00 - 15.30	☕ Break
15.30 - 16.15	Malaria practical #2
16.15 - 16.30	Prepare presentation of data findings
16.30 - 17.00	Present and discuss data findings

3.3 Day 3: Wednesday September 26th

Table 4: Overview Day 3

Time	Session
08.30 - 09.00	Welcome
09.00 - 09.20	Interdisciplinary introduction to big data and machine Learning
	Overview of objectives for Day 3
09.20 - 10.15	Discussion on secondary data sources
	(Public Datasets, e.g. DHS, Facebook, facilities, etc)
	Benefits and drawbacks between primary and secondary data
	sources
10.15 - 10.45	☕ Break
10.45 - 12.15	[Analysis stream] Introduction to machine learning
	[Interpretation stream] Critically discuss data surveys/reports
12:15-13:15	☕ Lunch break
13:15-13:45	Speed talks - research presentations
13:45-15.00	[Analysis stream] Data practical
	[Interpretation stream] Critically discuss data surveys/reports
15.00 - 15.30	☕ Break
15:30-16:00	Feedback on findings from practicals
16:00-16:30	Feedback on workshop - Wrap-up

4 After the workshop

1. Fill out the online post-workshop questionnaire (link to be provided after the workshop)