Program

Key:

			Names [name(s)]
Mainly theory	Guided R code + theory	Mainly independent or group tasks	Lead facilitator [assisting facilitator(s)]

Day 1: 01-August-2022. Introduction to R

Time	Session	Learning outcomes	Facilitator(s)
9:00 - 9:30	Opening remarks	• Speakers and organizers introduction	Innocent Mboya and Jim Todd
9:30 - 10:30	Introduction to the R environment	• Set up a project • R data structures • Load internal and external data files • Base R plotting	Leacky Muchene [Sophia Kagoye]
10:30 - 11:00	BREAK		
11:00 - 12:00	Data wrangling	• Modify objects • Reshape: Melt, cast, gather	Philip Ayeko [Jacqueline Materu]
12:00 - 13:00	LUNCH		
13:00 – 14:00		• Tidyverse approach • Summary, mutate, filter, distinct, ungroup	Leacky Muchene [Innocent Mboya]
14:00 – 15:00	Visualization with ggplot	• Single layer plots • Multi-layer plots • Faceting • Saving in .eps, .pdf …	Leacky Muchene [Neema Mosha]
15:00 - 15:30	BREAK		
15:30 – 17:00	Reproducible research (Rmarkdown+GIT)	• Basic Rmarkdown structure • Transfer day's code to RMarkdown	Leacky Muchene [Wende Safari]

		• Compile html, pdf	

Day 2: 02-August-2022. Gaussian outcomes (Normal distribution)

Time	Session	Learning outcomes	Facilitator(s)
9:00 - 9:30	Refresher on linear regression	• Theory of univariate and multiple linear regression • Interpretation of parameter estimates • Hypothesis testing	Jacqueline Materu [Philip Ayeko]
9:30 - 10:30	Practical session • We can set this up as a mini project with questions to answer- PLFD style • Work individually or in small groups	• Model specification for both continuous and categorical predictors • Specifying interaction terms • Extracting model summary, coefficients as well as tidying this using broom::tidy() • Extracting model diagnostic plots, residuals and properly interpreting them • Model building and hypothesis testing • Specifying contrasts • Compare different models	Philip Ayeko [Jacqueline Materu]
10:30 - 11:00	BREAK		
11:00 - 12:00	Practical session continues		Philip Ayeko [Jacqueline Materu]
12:00 - 13:00	LUNCH		
13:00 – 14:00		• Basic statistical analysis report in RMarkdown	Leacky Muchene [Neema Mosha]
14:00 – 15:00	R functions and loops	• Write simple functions • Basic loops, repeat • apply functions	Neema Mosha [Leacky Muchene]
15:00 - 15:30	BREAK		
15:30 – 17:00	Introduction to simulation (Guided practical session)	• Simulate data to estimate simple parameters such as	Neema Mosha [Sophia Kagoye]

		mean, bias and variance of different sample sizes	

Day 3: 03-August-2022. Binary outcomes (binomial distribution)

Time	Session	Learning outcomes	Facilitator(s)
9:00 - 9:30	Refresher on logistic regression	- Theory of univariate and multiple logistic regression - Interpretation of parameter estimates - Hypothesis testing	Sophia Kagoye [Jim Todd]
9:30 - 10:30	Practical session (We can set this up as a mini project with questions to answer-PDA style)	- Binary versus binomial outcome specification in R - Model specification for both continuous and categorical predictors - Specifying interaction terms - Extracting model summary, coefficients as well as tidying this using broom::tidy() - classical model diagnostics and properly interpreting them - Model building and hypothesis testing - Compare different models	Neema Mosha [Sophia Kagoye]
10:30 - 11:00	BREAK		
11:00 - 12:00	Introduction to machine learning	- Basic theory of classification methods - Supervised versus unsupervised - Bias-variance trade-off - Training versus test set and overfitting - Classification algorithms	Jacqueliene Materu [Innocent Mboya]
12:00 - 13:00	LUNCH		
13:00 – 14:00	Guided session on advanced topics	- Implement logistic, LDA, kNN, - ROC curves	Jacqueliene Materu [Innocent Mboya]
14:00 – 15:00	Cross-validation approaches	- K-fold CV - Leave one out CV	Jacqueliene Materu [Innocent Mboya]
15:00 – 15:30	Refresher on Generalized Linear Mixed Models	Interpret spaghetti plots	Leacky Muchene [Sophia Kagoye]

		● Identify need for random effects	
15:30 – 17:00	Practical session	● Reshape data if needed ● Generate meaningful visualizations ● Fit GLMM in R	Leacky Muchene [Sophia Kagoye]

Day 4: 04-August-2022. Poisson regression and Survival analysis

Time	Session	Learning outcomes	Facilitator(s)
9:00 - 9:30	Refresher on Poisson regression	- Theory of univariate and multiple Poisson regression - Interpretation of parameter estimates - Overdispersion - Zero-inflated Poisson - Models with offset	Jim Todd [Wende Safari]
9:30 - 10:30	Practical session	- Model specification for both continuous and categorical predictors and interaction terms - Extracting model summary, coefficients as well as tidying this using broom::tidy() - Overdispersion due to zeros	Jim Todd [Wende Safari]
10:30 - 11:00	BREAK		
11:00 - 12:00	Introduction to survival analysis	- Introduction to survival analysis - Kaplan-Meier, Log rank test, and Cox regression	Wende Safari [Jim Todd]
12:00 - 13:00	LUNCH		
13:00 - 15:00	Practical session	Implement survival models in R	Wende Safari [Jacqueliene Materu]
15:00 - 15:30	BREAK		
15:30 – 17:00	Non-linear models (Mini-project on GAM or 4/5PL)	- Fit generalized additive models - User defined objective function optimization	Leacky Muchene [Neema Mosha]

Day 5: 05-August-2022. Recap on learning objectives and introduction to future topics

Time	Session	Learning outcomes	Facilitator(s)	
9:00 - 11:00	Introduction to R Shiny	● Understand basic concepts ● Generate a simple UI in Shiny ● Rmarkdown, interactivity	Leacky Muchene	
11:00 - 11:30	BREAK			
11:30 - 13:30	Group presentations	● participants to present their workshop tasks?	All	
13:30 - 14:00	Closing remarks			
14:00 - 16:00	General Q&A + closing lunch	● Address user-specific questions	All	