MLW / KUHeS R & statistics short course

When: 19/10, 21/10, 26/10, 28/10, 02/11 and 04/11/2021, 9am-1pm on each day (6 sessions total)

Where: Nyika 1, Learning & Training Centre, Malawi – Liverpool – Wellcome Trust Clinical Research Programme.

Audience: Anybody who wants to learn or refresh their basic statistics & R skills: MSc & PhD students, post-docs, …

Equipment: Please bring your own laptop, with R installed (installation instructions will be shared before Session 1).

Cost / registration: $10 per participant for the full course; register by emailing mhenrion@mlw.mw before 12 October 2021.

Session 1: Introduction to R
Tuesday 19 October 2021, 9am-1pm

- Help, documentation, tutorials
- Interfaces: R console, R GUI, R Studio, R projects and script editors (Emacs, Vim, Notepad++, …)
- Reading data, writing data
- R objects / variables
- R workspace: saving it, recovering it
- Manipulating data in R: editing, subsetting, recoding, creating variables, R objects
- R scripts, useful editors, version control

Session 2: Data visualisation in R
Thursday 21 October 2021, 9am-1pm

- Different types of data: nominal, ordinal, interval, ratio
- graphics & ggplot2 libraries
- Data distributions
  - Bar plots, pie charts
  - Histograms
- Covariation
  - Box & whisker, violin plots
  - Scatter plots
  - Line & time plots
  - Heat maps, contour plots
- Multi-panel figures
- A note on creating maps for visualizing spatial data

Session 3: Statistical principles & study design
Tuesday 26 October 2021, 9am-1pm

- Probability theory
- Random variables
- Discrete & continuous distributions
  - Bernoulli, binomial, Poisson, geometric
  - Normal, beta, gamma, exponential
- Simulating data (using R)
- Observational designs: cross-sectional, longitudinal, case-control
- Experimental designs: clinical trials and other intervention studies, lab & field experiments
- General design considerations
- Common & adaptive designs for clinical trials

Session 4: Parameter estimation & statistical hypothesis testing (using R)
Thursday 28 October 2021, 9am-1pm

- What is a statistical hypothesis test & why do it?
- Type I and II errors: significance level & power
- Estimation & confidence intervals
- Statistical hypothesis testing (parametric and non-parametric tests)
- Tests for parameter estimates
- Comparing 2 or more groups
- Comparing 2 proportions
- Power & sample size
- A note on p-values, null hypothesis significance testing and the reproducibility crisis.

Session 5: Basic regression modelling (using R)
Tuesday 02 November 2021, 9am-1pm

- Linear model, correlation & ANOVA
- Generalised linear model
  - Special cases: logistic, Poisson & negative binomial regression
- Estimation, model fitting
  - Maximum likelihood, least squares, optimisation techniques (brief overview only)
  - Model diagnostics
  - Note on Bayesian modelling
- Statistical tests:
  - Wald, Likelihood ratio & score tests
  - Omnibus vs. single coefficient tests
- Interaction terms and interpretation
- Variable & model selection
- Regression and smoothing splines, Generalised Additive Models
- Correlated data (introduction only)

Session 6: Reproducible Research
Thursday 04 November 2021, 9am-1pm

- R markdown
- GitHub
- R packages
- Licenses for (i) software, (ii) research outputs