Zylloo Consult
Etudes & Conseils

2nd Floor, Abraham's Arcade,
Links Road, Nyali,
Mombasa, Kenya

info.cash@zylloo.com
https://zylloo.com/events

ABOUT R FOR ANALYTICS

This course provides a comprehensive introduction to programming in R, equipping learners with the skills to perform effective data analysis. Participants will gain hands-on experience in setting up and configuring the tools required for a robust statistical programming environment. The course also covers essential programming concepts, as implemented in R, a high-level statistical language widely used in data science.

Course Objectives

- Develop proficiency in statistical computing, including:
  - Programming in R.
  - Importing and managing data within R.
  - Utilizing and accessing R packages.
  - Writing custom R functions to streamline analysis.
  - Debugging and profiling R code for optimization.
  - Organizing, documenting, and commenting on R scripts for clarity and collaboration.

Who Should enrol?

- Aspiring data scientists looking to build foundational R programming skills.
- Current business data analysts or scientists seeking to enhance their expertise in statistical computing and data analysis.

Mode of Delivery: Online

Estimated Effort: 58 hours

Instructors:

- GIDEON CHIKAMAI INGUTIAH, Msc, MBA (Executive Director – ZYLLOO CONSULT LTD, Deputy Director - NCTCCA)
- DR. LEACKY MUCHENE (Janssen Pharmaceutical Companies of J&J, Netherlands)

Organizers: ZYLLOO CONSULT LTD

Technical assistance: CHESA KWEYU (Statistician)

Course Timetable for 2025

Course	Delivery Mode	Start Date	End Date	Duration	Frequency
Artificial Intelligence in R	Online/Physical	Jan 15, 2026	Feb 26, 2026	6 weeks	Twice a year
R for Analytics intermediate	Online/Physical	Mar 15, 2026	Apr 26, 2026	6 weeks	Twice a year
Data Science in R	Online/Physical	May 15, 2026	Jun 26, 2026	6 weeks	Quarterly
Data Science in Python	Online/Physical	Jul 15, 2026	Aug 26, 2026	6 weeks	Quarterly
Artificial Intelligence in R	Online/Physical	Sep 15, 2026	Oct 26, 2026	6 weeks	Twice a year
R for Analytics Advanced	Online/Physical	Nov 15, 2026	Dec 26, 2026	6 weeks	Twice a year

R FOR ANALYTICS COURSE CONTENTS & PROGRAMS

Course 1: Artificial Intelligence in R

Online: Jan 15 – Feb 26th 2025

Course Description:

This course focuses on leveraging R for Artificial Intelligence (AI), including machine learning, deep learning, and natural language processing. It equips learners with hands-on experience in implementing AI models in R.

Time	Session	Content	Facilitator(s)
	Opening remarks	Speakers and organizers introduction	
	Introduction to the R environment	○ Set up a project ○ R data structures ○ Load internal and external data files ○ Base R plotting	
	Data Wrangling	○ Modify objects ○ Reshape: melt, cast, gather ○ Tidyverse approach ○ Summary, mutate, filter, distict, ungroup	
	Visualization with ggplot	○ Single layer plots ○ Multi-layer plots ○ Faceting ○ Saving in .eps, .pdf	
	Reproducible research (Rmarkdown + GIT)	○ Basic Rmarkdown structure ○ Transfer day's code to RMarkdown ○ Compile html, pdf	
	Introduction to AI Concepts in R:	○ Overview of AI, ML, and DL. ○ Importance of R in AI. ○ Installation and configuration of relevant R libraries (e.g., caret, tensorflow, keras, nnet).	
	Supervised Learning Models:	○ Linear and Logistic Regression. ○ Decision Trees, Random Forests, Gradient Boosting (e.g., XGBoost, LightGBM). ○ Performance metrics (accuracy, precision, recall, F1 score).	
	Unsupervised Learning Models:	○ Clustering algorithms (K-means, Hierarchical Clustering). ○ Dimensionality reduction (PCA, t-SNE).	
	Deep Learning:	○ Introduction to Neural Networks. ○ Building feedforward, convolutional, and recurrent networks in R using keras and tensorflow. ○ Hyperparameter tuning.	

Time	Session	Content	Facilitator(s)
	Natural Language Processing:	○ Text preprocessing and tokenization. ○ Sentiment analysis and topic modeling. ○ Text generation with RNNs.	
	AI in Practice:	○ Case studies on real-world applications. ○ Ethical considerations in AI.	

Course 2: R for Analytics (Intermediate)

Online: Mar 5 – Apr 26th 2025

Course Description:

This course builds on foundational R knowledge to develop intermediate analytics skills, including advanced data manipulation, visualization, and exploratory data analysis.

Time	Session	Content	Facilitator(s)
	Opening remarks	Speakers and organizers introduction	
	Data Manipulation:	○ Working with dplyr and data.table. ○ Advanced data cleaning techniques. ○ Handling missing and outlier data.	
	Data Visualization:	○ Customizing plots with ggplot2. ○ Interactive visualizations with plotly and shiny. ○ Dashboard creation.	
	Exploratory Data Analysis (EDA):	○ Techniques to summarize data. ○ Advanced statistical measures (e.g., correlation, skewness, kurtosis). ○ Data distribution visualization	
	Introduction to Time Series Analysis:	○ Basic concepts and applications. ○ Visualization and decomposition.	
	Hands-on Projects:	○ Analyzing business/real-time datasets. ○ Data-driven decision-making use cases.	

Course 3: Data Science in R

Online: May 15 – Jun 26th 2025

Course Description:

This course covers core data science techniques, focusing on machine learning and statistical modelling in R.

Time	Session	Content	Facilitator(s)
	Opening remarks	Speakers and organizers introduction	
	Introduction to Data Science	○ Data science workflow. ○ Role of R in data science.	
	Statistical Foundations	○ Descriptive and inferential statistics. ○ Hypothesis testing ○ Confidence intervals.	
	Predictive Modeling	○ Regression and classification models. ○ Cross-validation and model evaluation.	
	Feature Engineering	○ Variable selection techniques. ○ Encoding categorical variables.	
	Model deployment	○ Exporting models for production. ○ Integrating R with web applications.	
	Capstone Project	○ Full-cycle data science project.	

Course 4: Data Science in Python

Online: Jul 15 – Aug 26th 2025

Course Description:

This course introduces Python as a powerful tool for data science, covering essential libraries like pandas, numpy, and scikit-learn.

Time	Session	Content	Facilitator(s)
	Opening remarks	Speakers and organizers introduction	
	Introduction to Python for Data Science	○ Setting up Python for data science. ○ Overview of pandas, numpy, and matplotlib.	
	Data processing	○ Data wrangling with pandas. ○ Handling large datasets with dask.	
	Machine Learning with scikit-learn	○ Supervised and unsupervised learning. ○ Hyperparameter tuning.	
	Feature Engineering	○ Variable selection techniques. ○ Encoding categorical variables.	
	Advanced topics in Python for Data science	○ Deep learning with tensorflow and keras. ○ Natural Language Processing with nltk and spaCy.	
	End-to-End Data Science Project	○ From Exploratory analysis to deployment.	

Course 5: Artificial Intelligence in R (Advanced Version)

Online: Sept 15 – Oct 26th 2025

Course Description:

This advanced course delves into complex AI models and their implementation in R.

Time	Session	Content	Facilitator(s)
	Opening remarks	Speakers and organizers introduction	
	Advanced machine learning	○ Ensemble methods and stacking models. ○ Anomaly detection.	
	Deep Learning Optimization	○ Transfer learning. ○ Optimizing neural networks for efficiency.	
	AI for Big Data	○ Working with large datasets using sparklyr. ○ Distributed machine learning	
	AI Applications	○ Recommender systems. ○ Image recognition.	

Course 6: R for Analytics (Advanced)

Online: Nov 15 – Dec 26th 2025

Course Description:

This advanced course equips learners with skills for sophisticated analytics, focusing on predictive modeling and advanced visualization.

Time	Session	Content	Facilitator(s)
	Opening remarks	Speakers and organizers introduction	
	Advanced Statistical Analysis	○ Ensemble Bayesian inference. ○ Advanced time series forecasting.	
	Predictive Analytics	○ Building predictive pipelines. ○ Advanced regression techniques.	
	Data Visualization Techniques	○ Advanced themes and interactivity. ○ Geographic visualizations.	
	Real-world Applications	Analytics in marketing, healthcare, and finance	