AI and Omics Research Internship

Module I: Getting Started With R Programming

This module provides a comprehensive foundation in **R programming**, essential for computational biology, data preprocessing, machine learning & bioinformatics workflows.

Participants will begin with environment setup and progress through core programming concepts required for high-throughput genomics data analysis.

Learning Objectives

- Set up the R programming environment (R, RStudio, RTools)
- Understand R syntax, data types, and control structures
- Create and manage projects in R
- Write custom functions and use key R packages (dplyr, ggplot2, DataExplorer)
- Perform basic data exploration and visualization

Module Content

1. Getting Started with R

- Installing **R**, **RStudio**, and **RTools**
- Exploring the RStudio interface
- Setting the working directory
- Creating a project in RStudio
- Understanding how R executes code
- Saving and loading your work: .R, .RData, save(), load()

2. Basic Syntax in R

- Declaring variables
- Writing comments for clarity
- Understanding reserved keywords

3. Operators in R

- Arithmetic, assignment, relational, and logical operators

4. Conditional Statements

- if, else if, else control structures

5. Loops in R

- for loop
- while loop

6. User-Defined Functions

- Writing and using custom functions in R

7. Data Types in R

- Numeric, character, logical, and factor types

8. Data Structures & Data Assessment

- Vectors, lists, matrices, data frames
- Basic data validation and inspection

9. Introduction to DataExplorer package

- Automated data exploration and summary tools

10. Introduction to dplyr package

- Data manipulation: filter(), select(), mutate(), summarise(), group_by()

11. Introduction to ggplot2 package

- Data visualization using ggplot2

Hands-on Practice

Hands-on practice materials will be released with each session to reinforce learning in a real time

For every topic, you will receive:

- Practical scripts
- Short guided exercises
- All content will be available in this repository, organized by module and session.

GitHub Repository: https://github.com/AI-Biotechnology-Bioinformatics/AI_and_Omics_Research_Internship_2025

Recommendation: Create your own GitHub repository to document your learning journey. This is not a submission requirement, but it is highly recommended. Maintaining a personal repository helps you organize your work, track your progress, and showcase your skills to peers, mentors, and potential recruiters.