Day 04

Topic Covered: Data Structures in R

Summary:

Today's session focused on data structures in R, which are essential for organizing and storing data efficiently. Data structures help manage memory usage and improve computation speed, which is important when working with large datasets in data analytics. We explored the four main data structures in R: vectors, lists, matrices, and arrays. Each structure has its own properties and use cases. Understanding these structures allows us to store, manage, and manipulate data effectively, whether we deal with simple numeric sequences or multi-dimensional datasets.

New Concepts Learned:

- Vectors: One-dimensional homogeneous collections of data created using c(). Functions include length(), sort(), rep(), and seq().

- Lists: One-dimensional heterogeneous collections created using list(). We learned accessing elements using indexing, checking existence using %in%, combining lists using c(), and appending values.

- Matrices: Two-dimensional homogeneous structures created using matrix(). Rows run horizontally and columns vertically. Learned to access elements, delete rows or columns, combine matrices using cbind() and rbind(), and check dimensions using dim().

- Arrays: Multi-dimensional homogeneous structures created using array(). Learned indexing in multiple dimensions, using length() to check data size, and dim() to check shape.

Activity:

- Created vectors and used functions like length(), sort(), rep(), and seq()

- Built lists with different types of elements and accessed elements using indexing

- Created matrices, performed row and column operations, and combined matrices using cbind() and rbind()

- Worked with arrays and practiced multi-dimensional indexing and dimension checking

Challenges Faced:

Remembering the syntax and functions for each data structure, especially arrays with multiple dimensions, was challenging. Understanding how to use each structure effectively for real-world data analysis required careful practice.

Key Takeaway:

Mastering data structures is essential for data storage, manipulation, and analysis in R. Each structure serves a specific purpose, and knowing when to use vectors, lists, matrices, or arrays is an important skill in data analytics.