Bharati Vidyapeeth
Deemed to be University

BVUDET

Department of Engineering and Technology
Plot no. KC-1, Sector 3, Kharghar, Navi Mumbai-410210

Subject: Computing Lab - III | Experiment No - 07 (3rd YEAR CSE-AIML 2023-2024)

Roll No: 11	Name: Kamran Khan
Class: CSE-AIML	Batch: B1
PRN: 2143110133	Date of Experiment: __ / __ / 2024
Marks (Out of 25):	Date of Submission: __ / __ / 2024

Aim: To implement group manipulation in R programming language.

Theory:

Group manipulation in R involves performing operations on subsets of data based on specific grouping variables. Several functions facilitate group manipulation:

1. **apply:** The `apply()` function applies a function to the rows or columns of a matrix, array, or data frame.

```
matrix_data <- matrix(1:12, nrow = 4, ncol = 3)
apply(matrix_data, 1, mean)  # Apply mean function to rows
```

```
## [1] 5 6 7 8
```

2. **lapply:** It applies a function to each element of a list and returns a list.

```
list_data <- list(a = 1:5, b = 6:10, c = 11:15)
lapply(list_data, mean)  # Compute mean for each element in the list
```

```
## $a
## [1] 3
##
## $b
## [1] 8
##
## $c
## [1] 13
```

3. **sapply:** Similar to `lapply()` , but it simplifies the result to a vector or matrix if possible.

```
sapply(list_data, mean)  # Simplify the result to a vector
```

```
##  a  b  c
##  3  8 13
```

4. **mapply:** It applies a function to multiple arguments in parallel.

```
mapply(rep, 1:4, 4:1)  # Repeat values from 1 to 4, 4 to 1 times respectively
```

```
## [[1]]
## [1] 1 1 1 1
##
## [[2]]
## [1] 2 2 2
##
## [[3]]
## [1] 3 3
##
## [[4]]
## [1] 4
```

5. **tapply:** This function applies a function over subsets of a vector or data frame, split by a factor.

```
tapply(iris$Sepal.Length, iris$Species, mean)  # Compute mean Sepal.Length for each Species
```

```
##     setosa versicolor  virginica 
##      5.006      5.936      6.588
```

6. **Aggregation:** Aggregation functions like `aggregate()` are used to compute summary statistics for subsets of data.

```
aggregate(Sepal.Length ~ Species, data = iris, FUN = mean)  # Compute mean Sepal.Length for each Species using aggregate
```

```
##      Species Sepal.Length
## 1     setosa        5.006
## 2 versicolor        5.936
## 3  virginica        6.588
```

Conclusion:

In this experiment, we explored various functions for group manipulation in R programming. These functions are essential for analyzing data at different levels of granularity and summarizing information within groups.

Signature of Lab Incharge

(Prof. Supriya Khaitan)