```
In [2]:   # Experiment No: 6 - Install the necessary R packages and apply data manipul
```

```
In [4]:   # suppress messages when loading dplyr
          suppressPackageStartupMessages(library(dplyr))

          # create a data frame
          stats <- data.frame(
            player = c('A', 'B', 'C', 'D'),
            runs = c(100, 200, 408, 19),
            wickets = c(17, 20, NA, 5)
          )

          # fetch players who scored more than 100 runs
          result <- filter(stats, runs > 100)

          # Output the result
          print(result)
```

```
  player runs wickets
1      B  200      20
2      C  408      NA
```

```
In [5]:   # create a data frame
          stats <- data.frame(
            player = c('A', 'B', 'C', 'D', 'A', 'A'),
            runs = c(100, 200, 408, 19, 56, 100),
            wickets = c(17, 20, NA, 5, 2, 17)
          )

          # remove duplicate rows
          result_distinct <- distinct(stats)

          # remove duplicates based on a column (player) while keeping all columns
          result_distinct_player <- distinct(stats, player, .keep_all = TRUE)

          # Output the results
          print(result_distinct)
          print(result_distinct_player)
```

```
  player runs wickets
1      A  100      17
2      B  200      20
3      C  408      NA
4      D   19       5
5      A   56       2
  player runs wickets
1      A  100      17
2      B  200      20
3      C  408      NA
4      D   19       5
```