**PROGRAMS :**

**1) Write an R program to create a new dataframe with the following columns:**

**a) Rollno**

**b) Sname**

**c) dept**

**d) Sem**

**e) Subject1 IT marks**

**f) Subject2 IT marks**

**g) Subject3 IT marks**

**h) Subject1 Semester Theory marks**

**i) Subject2 Semester Theory marks**

**j) Subject3 Semester Theory marks**

**k) Attendance**

**l) Grade**

**data <- data.frame(**

**Rollno = c(1, 2, 3, 4, 5),**

**Sname = c("John", "Alice", "Bob", "Emma", "Michael"),**

**dept = c("Computer Engineering", "IT", "Computer Engineering", "IT", "Computer Engineering"),**

**Sem = c(3, 3, 5, 5, 3),**

**Subject1\_IT\_marks = c(85, 78, 90, 82, 88),**

**Subject2\_IT\_marks = c(75, 80, 85, 77, 79),**

**Subject3\_IT\_marks = c(88, 82, 91, 79, 83),**

**Subject1\_Semester\_Theory\_marks = c(75, 80, 85, 77, 79),**

**Subject2\_Semester\_Theory\_marks = c(70, 75, 80, 73, 76),**

**Subject3\_Semester\_Theory\_marks = c(80, 85, 90, 83, 86),**

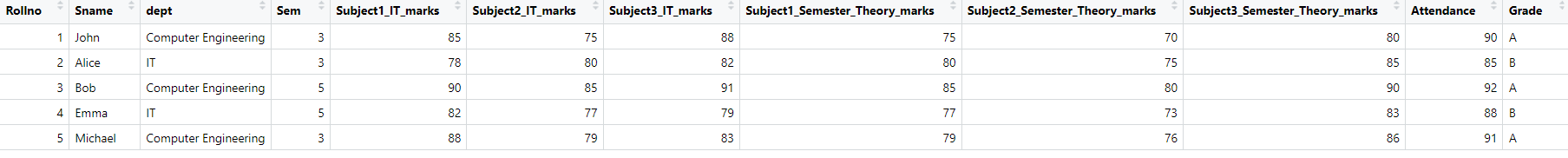
**Attendance = c(90, 85, 92, 88, 91),**

**Grade = c("A", "B", "A", "B", "A")**

**)**

**View(data)**

**Output:**

****

**2) For the above mentioned dataset display the data of computer engineering students for**

**semester 3**

**data <- data.frame(**

**Rollno = c(1, 2, 3, 4, 5),**

**Sname = c("John", "Alice", "Bob", "Emma", "Michael"),**

**dept = c("Computer Engineering", "IT", "Computer Engineering", "IT", "Computer Engineering"),**

**Sem = c(3, 3, 5, 5, 3),**

**Subject1\_IT\_marks = c(85, 78, 90, 82, 88),**

**Subject2\_IT\_marks = c(75, 80, 85, 77, 79),**

**Subject3\_IT\_marks = c(88, 82, 91, 79, 83),**

**Subject1\_Semester\_Theory\_marks = c(75, 80, 85, 77, 79),**

**Subject2\_Semester\_Theory\_marks = c(70, 75, 80, 73, 76),**

**Subject3\_Semester\_Theory\_marks = c(80, 85, 90, 83, 86),**

**Attendance = c(90, 85, 92, 88, 91),**

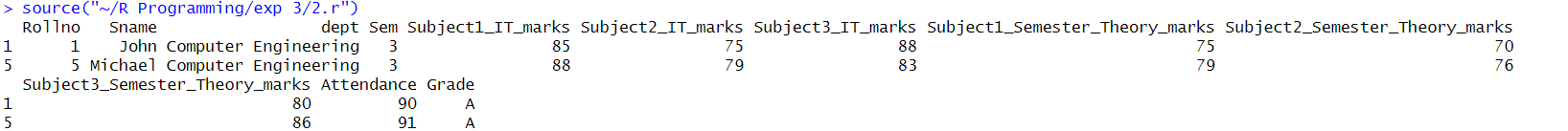
**Grade = c("A", "B", "A", "B", "A")**

**)**

**computer\_eng\_sem\_3 <- subset(data, dept == "Computer Engineering" & Sem == 3)**

**print(computer\_eng\_sem\_3)**

**Output:**

****

**3) For the above mentioned data frame ,display the average marks of students of sem 5**

**data <- data.frame(**

**Rollno = c(1, 2, 3, 4, 5),**

**Sname = c("John", "Alice", "Bob", "Emma", "Michael"),**

**dept = c("Computer Engineering", "IT", "Computer Engineering", "IT", "Computer Engineering"),**

**Sem = c(3, 3, 5, 5, 3),**

**Subject1\_IT\_marks = c(85, 78, 90, 82, 88),**

**Subject2\_IT\_marks = c(75, 80, 85, 77, 79),**

**Subject3\_IT\_marks = c(88, 82, 91, 79, 83),**

**Subject1\_Semester\_Theory\_marks = c(75, 80, 85, 77, 79),**

**Subject2\_Semester\_Theory\_marks = c(70, 75, 80, 73, 76),**

**Subject3\_Semester\_Theory\_marks = c(80, 85, 90, 83, 86),**

**Attendance = c(90, 85, 92, 88, 91),**

**Grade = c("A", "B", "A", "B", "A")**

**)**

**sem\_5\_avg\_marks <- mean(data$Subject1\_IT\_marks[data$Sem == 5] + data$Subject2\_IT\_marks[data$Sem == 5] + data$Subject3\_IT\_marks[data$Sem == 5])**

**print(sem\_5\_avg\_marks)**

**Output:**

****

**4) For the above mentioned dataframe display the student name,dept,sem and attendance**

**data <- data.frame(**

**Rollno = c(1, 2, 3, 4, 5),**

**Sname = c("John", "Alice", "Bob", "Emma", "Michael"),**

**dept = c("Computer Engineering", "IT", "Computer Engineering", "IT", "Computer Engineering"),**

**Sem = c(3, 3, 5, 5, 3),**

**Subject1\_IT\_marks = c(85, 78, 90, 82, 88),**

**Subject2\_IT\_marks = c(75, 80, 85, 77, 79),**

**Subject3\_IT\_marks = c(88, 82, 91, 79, 83),**

**Subject1\_Semester\_Theory\_marks = c(75, 80, 85, 77, 79),**

**Subject2\_Semester\_Theory\_marks = c(70, 75, 80, 73, 76),**

**Subject3\_Semester\_Theory\_marks = c(80, 85, 90, 83, 86),**

**Attendance = c(90, 85, 92, 88, 91),**

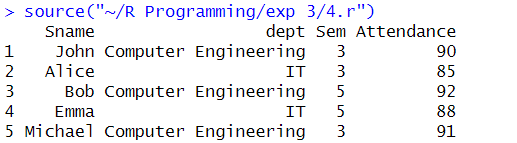
**Grade = c("A", "B", "A", "B", "A")**

**)**

**student\_info <- data[, c("Sname", "dept", "Sem", "Attendance")]**

**print(student\_info)**

**Output:**

****

**5) Display any three records for the above mentioned data frame**

**data <- data.frame(**

**Rollno = c(1, 2, 3, 4, 5),**

**Sname = c("John", "Alice", "Bob", "Emma", "Michael"),**

**dept = c("Computer Engineering", "IT", "Computer Engineering", "IT", "Computer Engineering"),**

**Sem = c(3, 3, 5, 5, 3),**

**Subject1\_IT\_marks = c(85, 78, 90, 82, 88),**

**Subject2\_IT\_marks = c(75, 80, 85, 77, 79),**

**Subject3\_IT\_marks = c(88, 82, 91, 79, 83),**

**Subject1\_Semester\_Theory\_marks = c(75, 80, 85, 77, 79),**

**Subject2\_Semester\_Theory\_marks = c(70, 75, 80, 73, 76),**

**Subject3\_Semester\_Theory\_marks = c(80, 85, 90, 83, 86),**

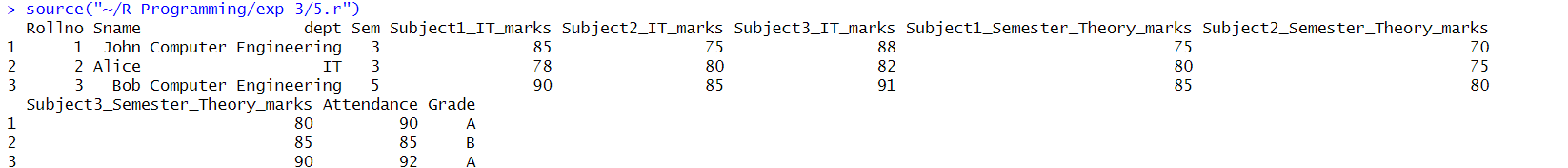
**Attendance = c(90, 85, 92, 88, 91),**

**Grade = c("A", "B", "A", "B", "A")**

**)**

**print(data[1:3, ])**

**Output:**

****

**6) Create two data frames and demonstrate left outer join,right outer join,outer join ,cross join,set difference and intersect operation**

**df1 <- data.frame(**

**ID = c(1, 2, 3, 4),**

**Name = c("John", "Alice", "Bob", "Emma"),**

**Age = c(25, 30, 28, 22)**

**)**

**df2 <- data.frame(**

**ID = c(1, 2, 5),**

**Gender = c("M", "F", "F"),**

**Salary = c(50000, 60000, 55000)**

**)**

**# Left outer join**

**left\_join\_df <- merge(df1, df2, by = "ID", all.x = TRUE)**

**print(left\_join\_df)**

**cat("\n")**

**# Right outer join**

**right\_join\_df <- merge(df1, df2, by = "ID", all.y = TRUE)**

**print(right\_join\_df)**

**cat("\n")**

**# Outer join**

**outer\_join\_df <- merge(df1, df2, by = "ID", all = TRUE)**

**print(outer\_join\_df)**

**cat("\n")**

**# Cross join**

**cross\_join\_df <- merge(df1, df2, by = NULL)**

**print(cross\_join\_df)**

**cat("\n")**

**# Set difference**

**set\_difference\_df <- df1[!df1$ID %in% df2$ID, ]**

**print(set\_difference\_df)**

**cat("\n")**

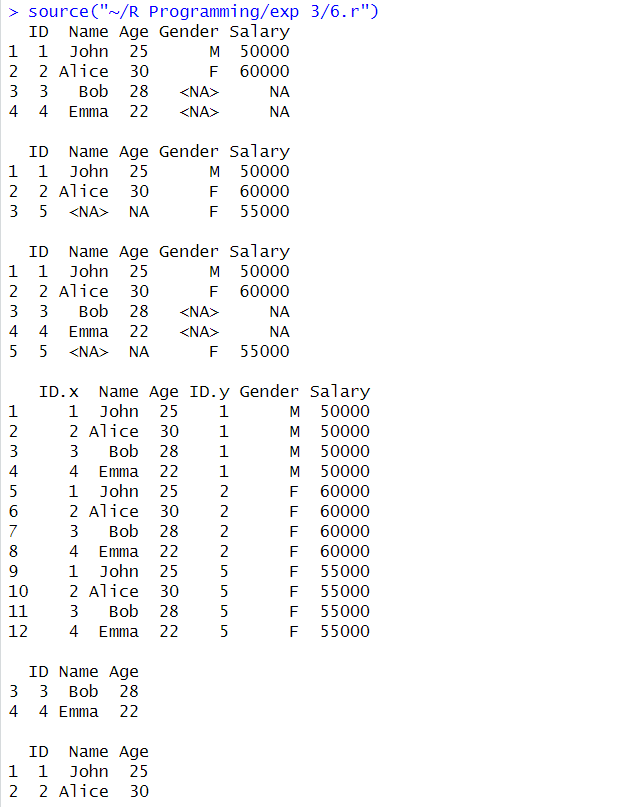
**# Intersect operation**

**intersect\_df <- df1[df1$ID %in% df2$ID, ]**

**print(intersect\_df)**

**cat("\n")**

**Output:**

****