Task-4-Data-Frame-Operations-2347249.R

rojal

2023-11-10

```
#Load the dataset from "employee_data.csv" into a data frame in R.
employee_data<-read.csv("employee_data.csv")
print(employee_data)
```

```
##    Employee_ID   Name  Department Salary Joining_Date
## 1            1   Riya          IT   5000    2000-8-10
## 2            2    Sam       Sales   1000    2001-4-10
## 3            3  Jenny   Marketing   2000    2005-10-4
## 4            4   Rose          IT   3000    2009-10-8
## 5            5 Antony       Sales   8000     2014-2-1
## 6            6  Karen   Marketing  10000    2014-5-25
## 7            7  Haily Devolpement   4000   2016-12-30
## 8            8 Justin Devolpement   1000    2000-7-16
## 9            9   Mary       Sales  12000    2023-10-1
## 10          10   Alen   Designing  15000     2022-5-4
```

```
df<-data.frame(employee_data)
df
```

```
##    Employee_ID   Name  Department Salary Joining_Date
## 1            1   Riya          IT   5000    2000-8-10
## 2            2    Sam       Sales   1000    2001-4-10
## 3            3  Jenny   Marketing   2000    2005-10-4
## 4            4   Rose          IT   3000    2009-10-8
## 5            5 Antony       Sales   8000     2014-2-1
## 6            6  Karen   Marketing  10000    2014-5-25
## 7            7  Haily Devolpement   4000   2016-12-30
## 8            8 Justin Devolpement   1000    2000-7-16
## 9            9   Mary       Sales  12000    2023-10-1
## 10          10   Alen   Designing  15000     2022-5-4
```

```
#Display the structure of the data frame, including column names and data types.
str(df)
```

```
## 'data.frame':    10 obs. of  5 variables:
##  $ Employee_ID : int  1 2 3 4 5 6 7 8 9 10
##  $ Name        : chr  "Riya" "Sam" "Jenny" "Rose" ...
##  $ Department  : chr  "IT" "Sales" "Marketing" "IT" ...
##  $ Salary      : int  5000 1000 2000 3000 8000 10000 4000 1000 12000 15000
##  $ Joining_Date: chr  "2000-8-10" "2001-4-10" "2005-10-4" "2009-10-8" ...
```

```
#Calculate and add a new column named "Years of Service" to the data frame, representing the
number of years each employee has worked in the company.
Years_of_Service<-c(10,5,2,4,13,5,6,5,8,10)
df<-cbind(df,Years_of_Service)

#Create a new data frame named "Senior Employees" containing records of employees who have wo
rked for the company for 5 or more years.
s=subset(df,Years_of_Service>5)
Senior_Employees<-data.frame(s)
Senior_Employees
```

```
##    Employee_ID   Name  Department Salary Joining_Date Years_of_Service
## 1            1   Riya          IT   5000    2000-8-10               10
## 5            5 Antony       Sales   8000     2014-2-1               13
## 7            7  Haily Devolpement   4000   2016-12-30                6
## 9            9   Mary       Sales  12000    2023-10-1                8
## 10          10   Alen   Designing  15000     2022-5-4               10
```

```
#Calculate and print the average salary of employees in each department
mean(df$Salary)
```

```
## [1] 6100
```

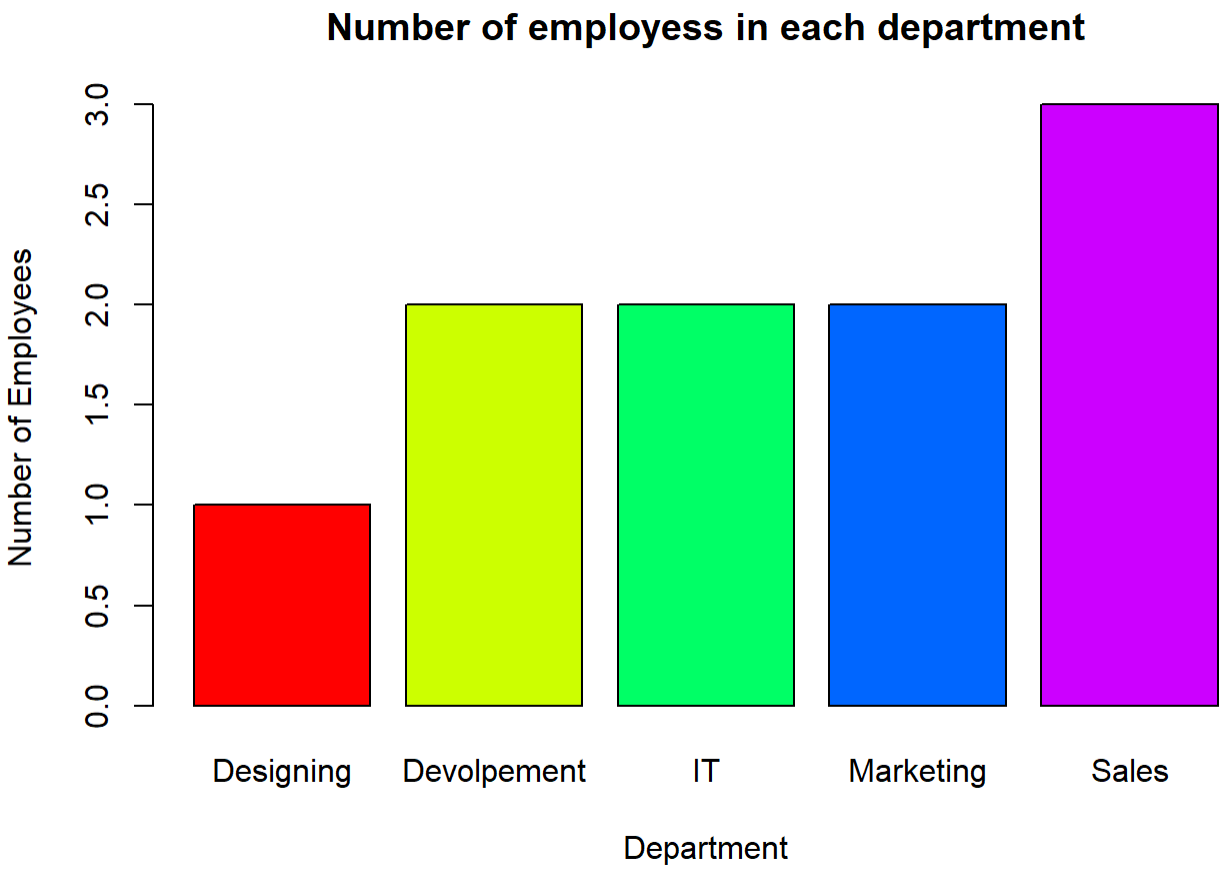
```
#Determine the highest and lowest salaries in the entire dataset and identify the employees w
ith these salaries.
low=min(df$Salary)
subset(df,df$Salary==low)
```

```
##   Employee_ID   Name  Department Salary Joining_Date Years_of_Service
## 2           2    Sam       Sales   1000    2001-4-10                5
## 8           8 Justin Devolpement   1000    2000-7-16                5
```

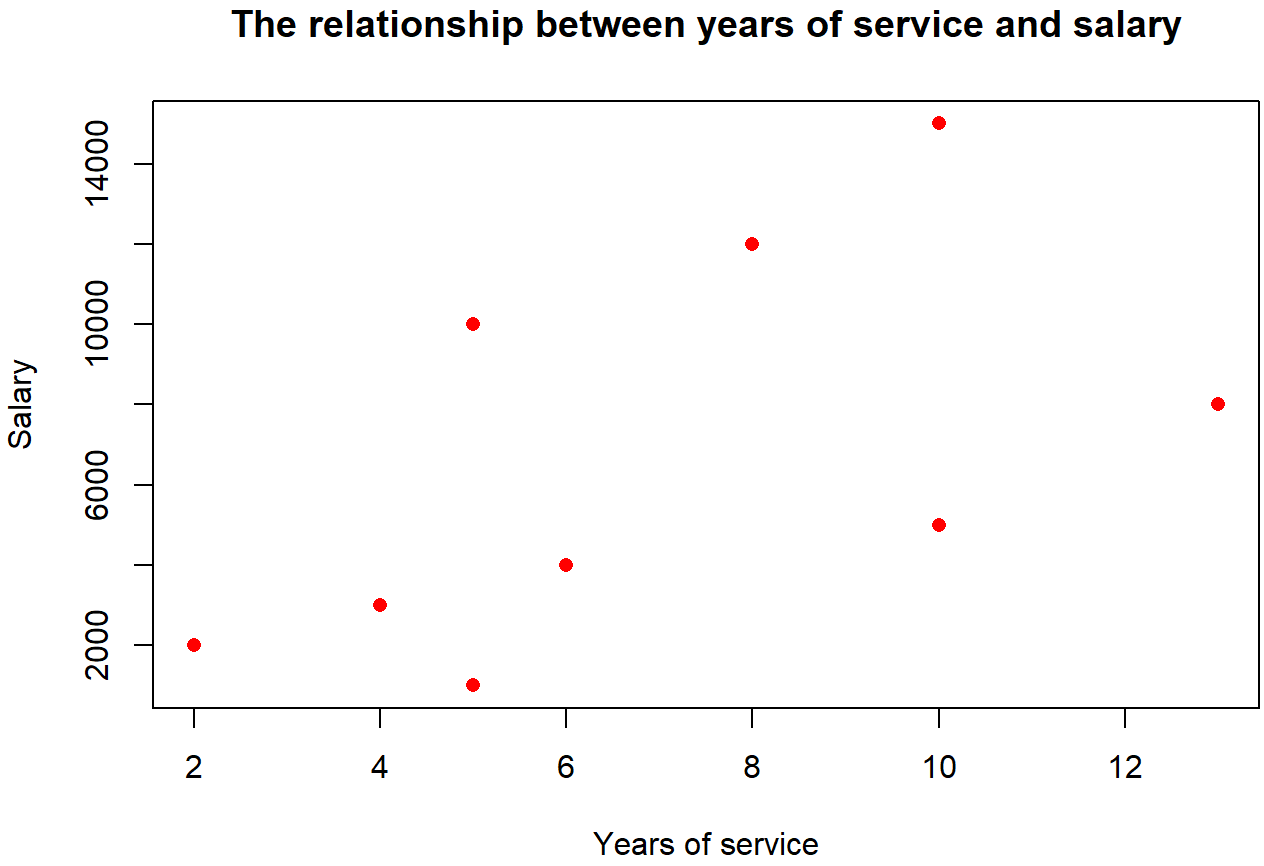
```
high=max(df$Salary)
subset(df,df$Salary==high)
```

```
##    Employee_ID Name Department Salary Joining_Date Years_of_Service
## 10          10 Alen  Designing  15000     2022-5-4               10
```

```
#Create a bar plot to visualize the number of employees in each department.
barplot(table(df$Department),
        main="Number of employess in each department",
        xlab = "Department",
        ylab = "Number of Employees",
        col=rainbow(length(unique(df$Department))))
```




```
#Generate a scatter plot to explore the relationship between years of service and salary.
plot(df$Years_of_Service,df$Salary,
     main="The relationship between years of service and salary",
     xlab="Years of service",
     ylab="Salary",
     pch=16,
     col="red"
     )
```




```
#Save the "Senior Employees" data frame as a CSV file named "senior_employees.csv."
write.csv(Senior_Employees,"senior_employees.csv.",row.names = FALSE)
```