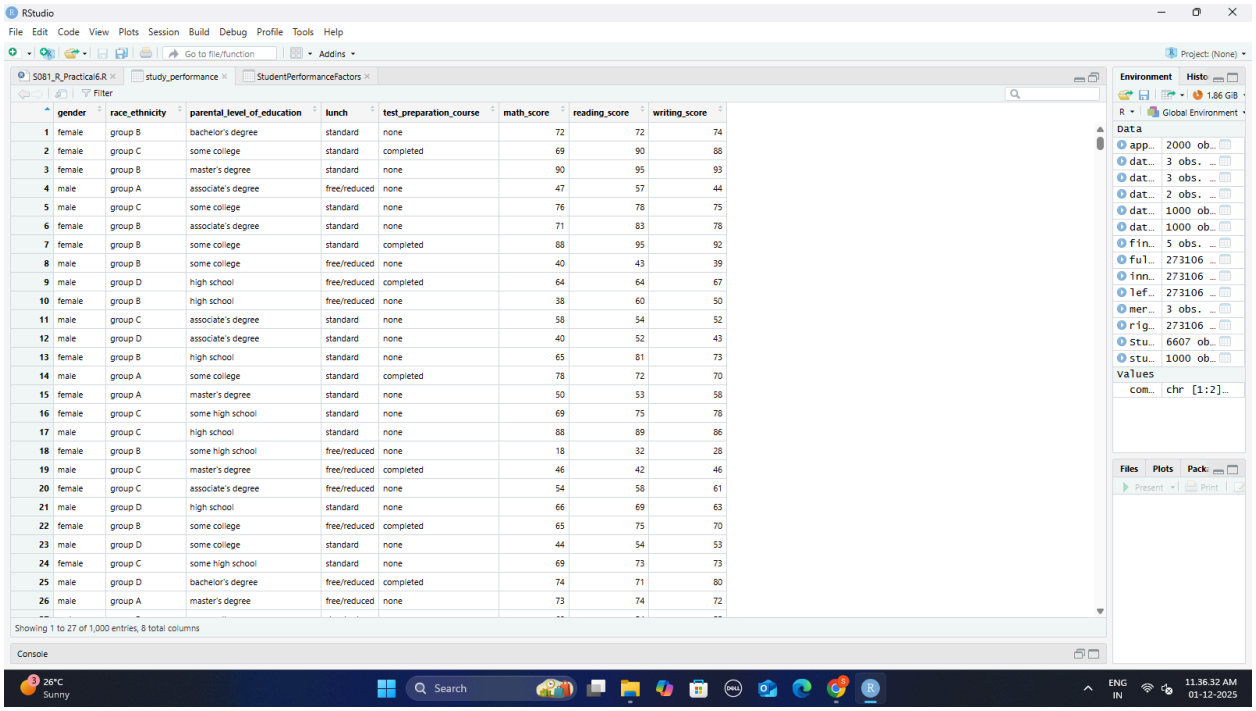
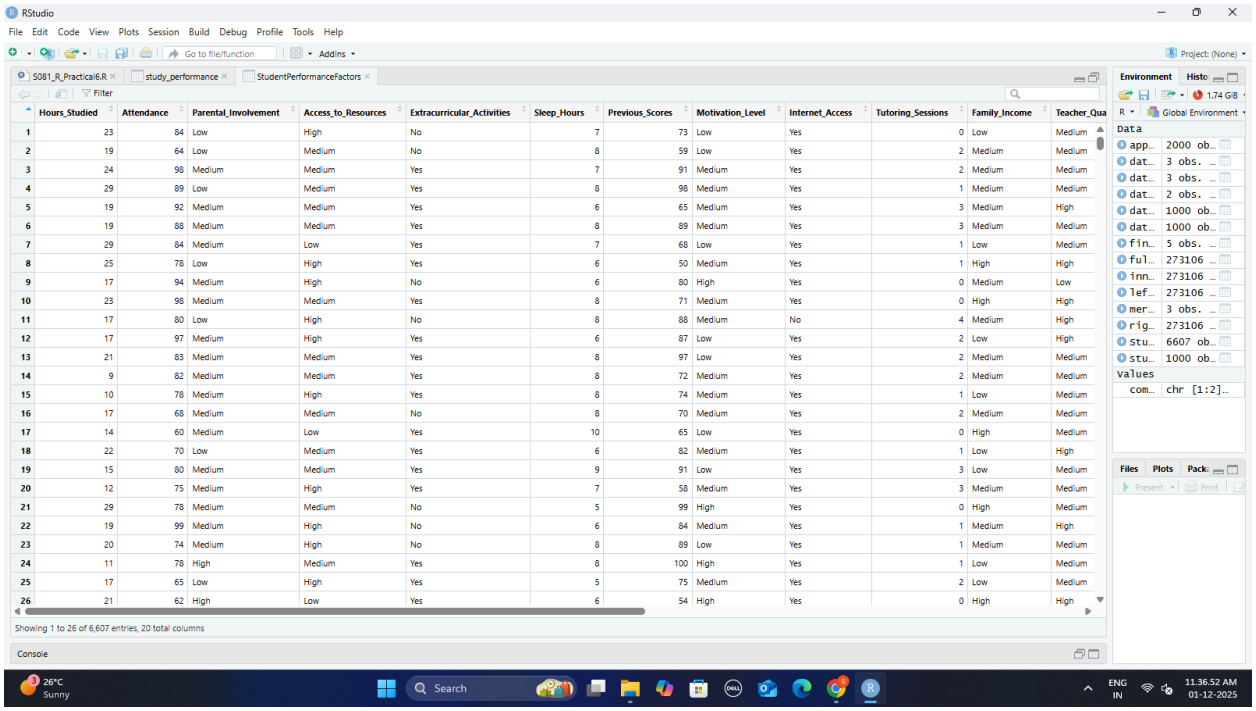
Sheth L.U.J. & Sir M.V. College

6. Combining and appending datasets using merge() or bind_rows() in R.



	gender	race_ethnicity	parental_level_of_education	lunch	test_preparation_course	math_score	reading_score	writing_score
1	female	group B	bachelor's degree	standard	none	72	72	74
2	female	group C	some college	standard	completed	69	90	88
3	female	group B	master's degree	standard	none	90	95	93
4	male	group A	associate's degree	free/reduced	none	47	57	44
5	male	group C	some college	standard	none	76	78	75
6	female	group B	associate's degree	standard	none	71	83	78
7	female	group B	some college	standard	completed	88	95	92
8	male	group B	some college	free/reduced	none	40	43	39
9	male	group D	high school	free/reduced	completed	64	64	67
10	female	group B	high school	free/reduced	none	38	60	50
11	male	group C	associate's degree	standard	none	58	54	52
12	male	group D	associate's degree	standard	none	40	52	43
13	female	group B	high school	standard	none	65	81	73
14	male	group A	some college	standard	completed	78	72	70
15	female	group A	master's degree	standard	none	50	53	58
16	female	group C	some high school	standard	none	69	75	78
17	male	group C	high school	standard	none	88	89	86
18	female	group B	some high school	free/reduced	none	18	32	28
19	male	group C	master's degree	free/reduced	completed	46	42	46
20	female	group C	associate's degree	free/reduced	none	54	58	61
21	male	group D	high school	standard	none	66	69	63
22	female	group B	some college	free/reduced	completed	65	75	70
23	male	group D	some college	standard	none	44	54	53
24	female	group C	some high school	standard	none	69	73	73
25	male	group D	bachelor's degree	free/reduced	completed	74	71	80
26	male	group A	master's degree	free/reduced	none	73	74	72

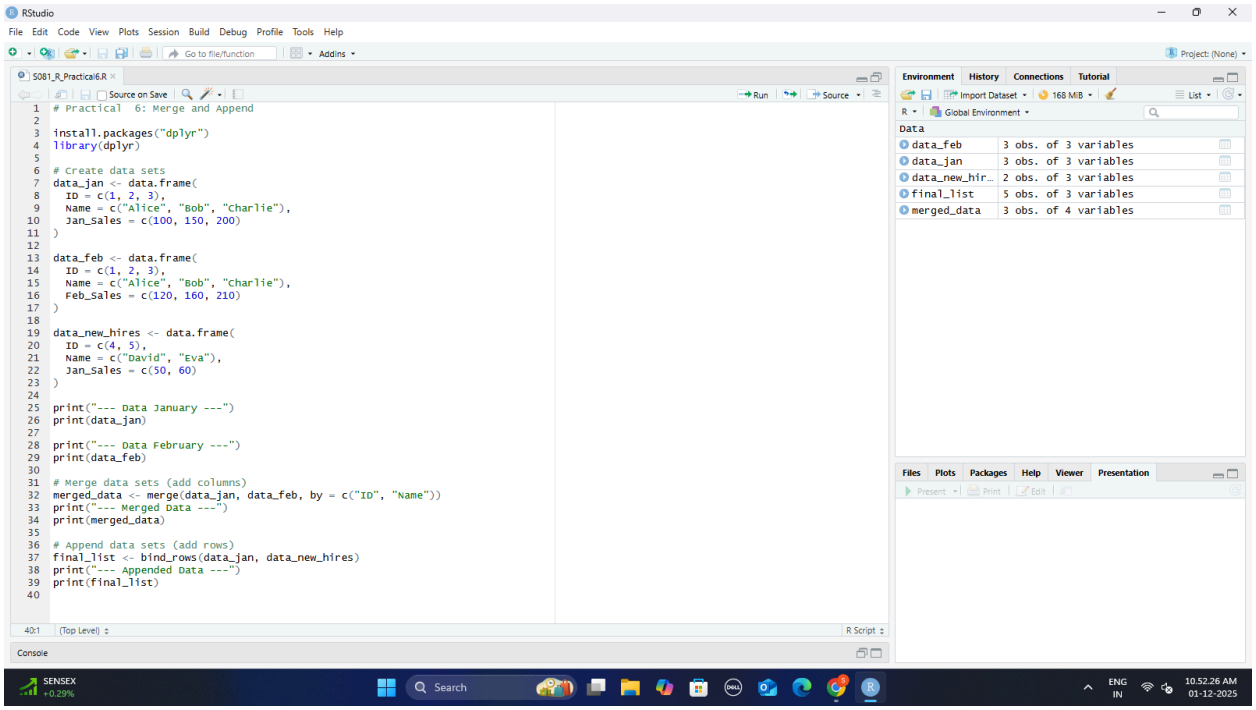


	Hours_Studied	Attendance	Parental_Involvement	Access_to_Resources	Extracurricular_Activities	Sleep_Hours	Previous_Scores	Motivation_Level	Internet_Access	Tutoring_Sessions	Family_Income	Teacher_Qua
1	23	84	Low	High	No	7	73	Low	Yes	0	Low	Medium
2	19	64	Low	Medium	No	8	59	Low	Yes	2	Medium	Medium
3	24	98	Medium	Medium	Yes	7	91	Medium	Yes	2	Medium	Medium
4	29	89	Low	Medium	Yes	8	98	Medium	Yes	1	Medium	Medium
5	19	92	Medium	Medium	Yes	6	65	Medium	Yes	3	Medium	High
6	19	88	Medium	Medium	Yes	8	89	Medium	Yes	3	Medium	Medium
7	29	84	Medium	Low	Yes	7	68	Low	Yes	1	Low	Medium
8	25	78	Low	High	Yes	6	50	Medium	Yes	1	High	High
9	17	94	Medium	High	No	6	80	High	Yes	0	Medium	Low
10	23	98	Medium	Medium	Yes	8	71	Medium	Yes	0	High	High
11	17	80	Low	High	No	8	88	Medium	No	4	Medium	High
12	17	97	Medium	High	Yes	6	87	Low	Yes	2	Low	High
13	21	83	Medium	Medium	Yes	8	97	Low	Yes	2	Medium	Medium
14	9	82	Medium	Medium	Yes	8	72	Medium	Yes	2	Medium	Medium
15	10	78	Medium	High	Yes	8	74	Medium	Yes	1	Low	Medium
16	17	68	Medium	Medium	No	8	70	Medium	Yes	2	Medium	Medium
17	14	60	Medium	Low	Yes	10	65	Low	Yes	0	High	Medium
18	22	70	Low	Medium	Yes	6	82	Medium	Yes	1	Low	High
19	15	80	Medium	Medium	Yes	9	91	Low	Yes	3	Low	Medium
20	12	75	Medium	High	Yes	7	58	Medium	Yes	3	Medium	Medium
21	29	78	Medium	Medium	No	5	99	High	Yes	0	High	Medium
22	19	99	Medium	High	No	6	84	Medium	Yes	1	Medium	High
23	20	74	Medium	High	No	8	89	Low	Yes	1	Medium	Medium
24	11	78	High	Medium	Yes	8	100	High	Yes	1	Low	Medium
25	17	65	Low	High	Yes	5	75	Medium	Yes	2	Low	Medium
26	21	62	High	Low	Yes	6	54	High	Yes	0	High	High

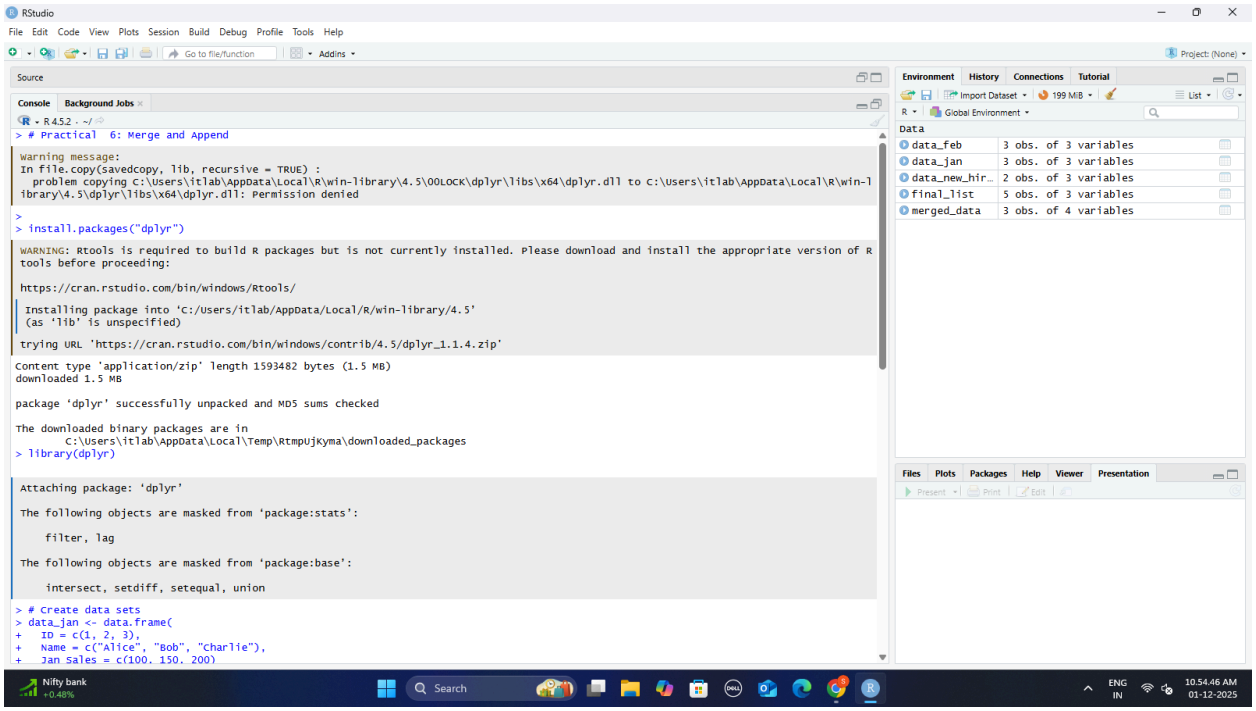
Name :- Priya Gupta

Roll no. :- S081

Sheth L.U.J. & Sir M.V. College



```
# Practical  6: Merge and Append

install.packages("dplyr")
library(dplyr)

# Create data sets
data_jan <- data.frame(
  ID = c(1, 2, 3),
  Name = c("Alice", "Bob", "Charlie"),
  Jan_Sales = c(100, 150, 200)
)

data_feb <- data.frame(
  ID = c(1, 2, 3),
  Name = c("Alice", "Bob", "Charlie"),
  Feb_Sales = c(120, 160, 210)
)

data_new_hires <- data.frame(
  ID = c(4, 5),
  Name = c("David", "Eva"),
  Jan_Sales = c(50, 60)
)

print("--- Data January ---")
print(data_jan)

print("--- Data February ---")
print(data_feb)

# Merge data sets (add columns)
merged_data <- merge(data_jan, data_feb, by = c("ID", "Name"))
print("--- Merged Data ---")
print(merged_data)

# Append data sets (add rows)
final_list <- bind_rows(data_jan, data_new_hires)
print("--- Appended Data ---")
print(final_list)
```

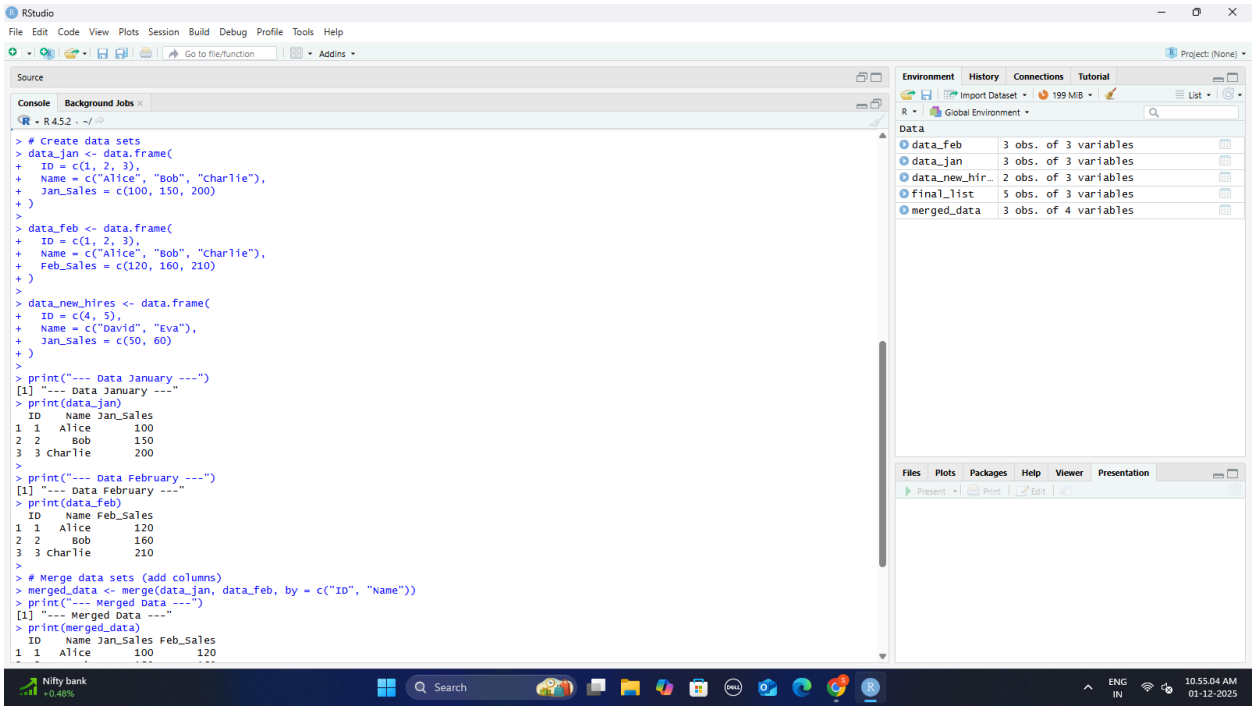


```
> # Practical  6: Merge and Append
Warning message:
In file.copy(savedcopy, lib, recursive = TRUE) :
  problem copying C:\Users\itlab\AppData\Local\R\win-library\4.5\00LOCK\dplyr\libs\x64\dplyr.dll to C:\Users\itlab\AppData\Local\R\win-library\4.5\dplyr\libs\x64\dplyr.dll: Permission denied

>
> install.packages("dplyr")
WARNING: Rtools is required to build R packages but is not currently installed. Please download and install the appropriate version of Rtools before proceeding:

https://cran.rstudio.com/bin/windows/Rtools/

Installing package into 'C:/Users/itlab/AppData/Local/R/win-library/4.5'
(as 'lib' is unspecified)

trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.5/dplyr_1.1.4.zip'
Content type 'application/zip' length 1593482 bytes (1.5 MB)
downloaded 1.5 MB

package 'dplyr' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\itlab\AppData\Local\Temp\RtmpUjKyma\downloaded_packages
> library(dplyr)

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union

> # Create data sets
> data_jan <- data.frame(
+   ID = c(1, 2, 3),
+   Name = c("Alice", "Bob", "Charlie"),
+   Jan_Sales = c(100, 150, 200)
```

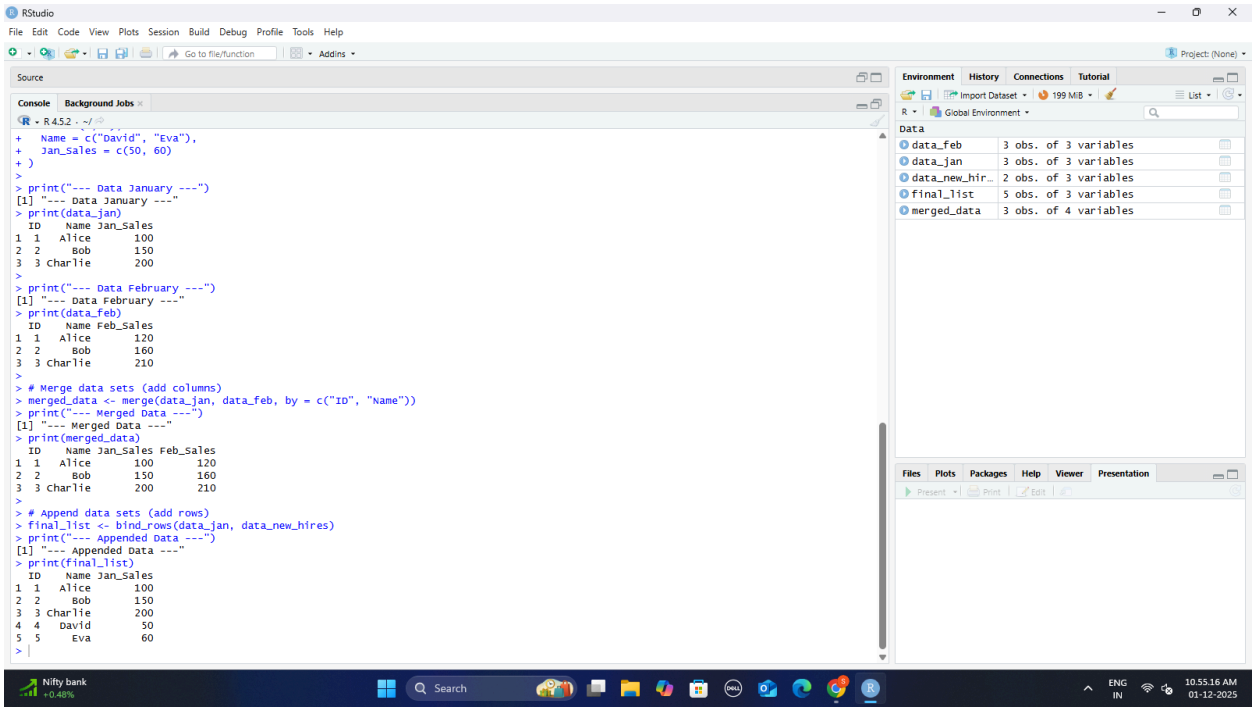
Name :- Priya Gupta

Roll no. :- S081

Sheth L.U.J. & Sir M.V. College



```
> # Create data sets
> data_jan <- data.frame(
+   ID = c(1, 2, 3),
+   Name = c("Alice", "Bob", "Charlie"),
+   Jan_Sales = c(100, 150, 200)
+ )
>
> data_feb <- data.frame(
+   ID = c(1, 2, 3),
+   Name = c("Alice", "Bob", "Charlie"),
+   Feb_Sales = c(120, 160, 210)
+ )
>
> data_new_hires <- data.frame(
+   ID = c(4, 5),
+   Name = c("David", "Eva"),
+   Jan_Sales = c(50, 60)
+ )
>
> print("--- Data January ---")
[1] "--- Data January ---"
> print(data_jan)
  ID    Name Jan_Sales
1  1   Alice       100
2  2     Bob       150
3  3 Charlie       200
>
> print("--- Data February ---")
[1] "--- Data February ---"
> print(data_feb)
  ID    Name Feb_Sales
1  1   Alice       120
2  2     Bob       160
3  3 Charlie       210
>
> # Merge data sets (add columns)
> merged_data <- merge(data_jan, data_feb, by = c("ID", "Name"))
> print("--- Merged Data ---")
[1] "--- Merged Data ---"
> print(merged_data)
  ID    Name Jan_Sales Feb_Sales
1  1   Alice       100       120
```



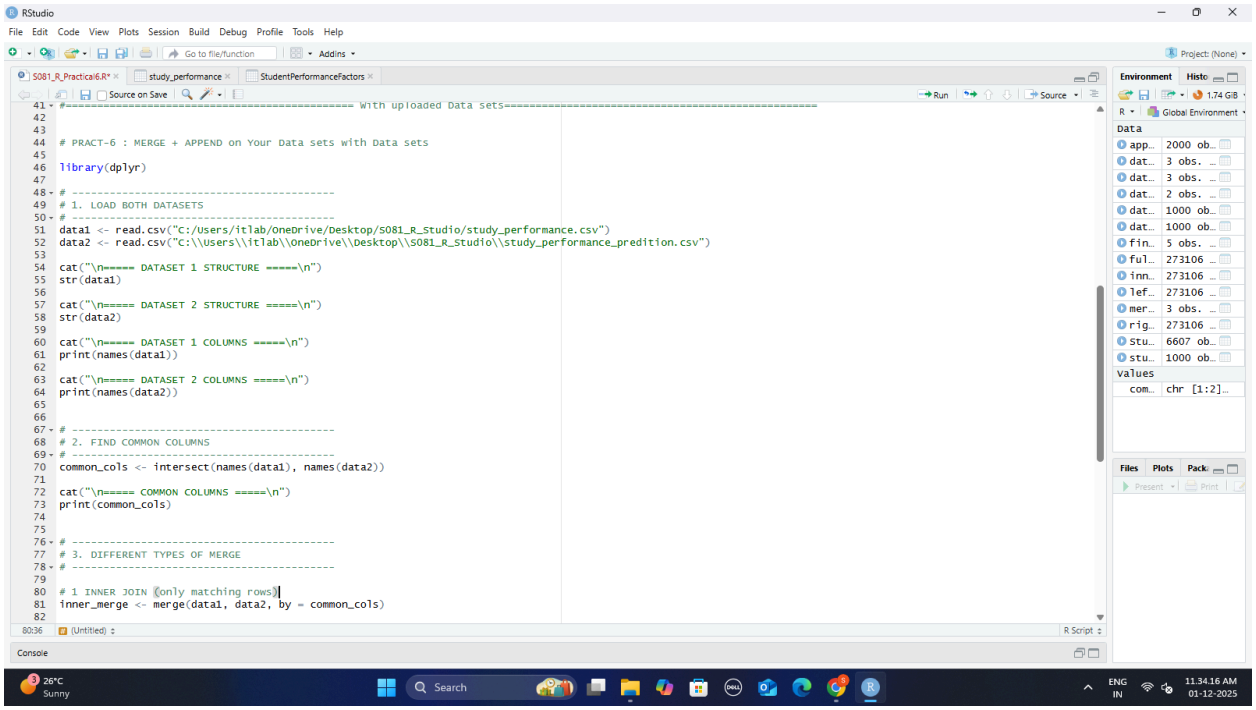
```
+   Name = c("David", "Eva"),
+   Jan_Sales = c(50, 60)
+ )
>
> print("--- Data January ---")
[1] "--- Data January ---"
> print(data_jan)
  ID    Name Jan_Sales
1  1   Alice       100
2  2     Bob       150
3  3 Charlie       200
>
> print("--- Data February ---")
[1] "--- Data February ---"
> print(data_feb)
  ID    Name Feb_Sales
1  1   Alice       120
2  2     Bob       160
3  3 Charlie       210
>
> # Merge data sets (add columns)
> merged_data <- merge(data_jan, data_feb, by = c("ID", "Name"))
> print("--- Merged Data ---")
[1] "--- Merged Data ---"
> print(merged_data)
  ID    Name Jan_Sales Feb_Sales
1  1   Alice       100       120
2  2     Bob       150       160
3  3 Charlie       200       210
>
> # Append data sets (add rows)
> final_list <- bind_rows(data_jan, data_new_hires)
> print("--- Appended Data ---")
[1] "--- Appended Data ---"
> print(final_list)
  ID    Name Jan_Sales
1  1   Alice       100
2  2     Bob       150
3  3 Charlie       200
4  4   David        50
5  5     Eva        60
>
```

With Dataset

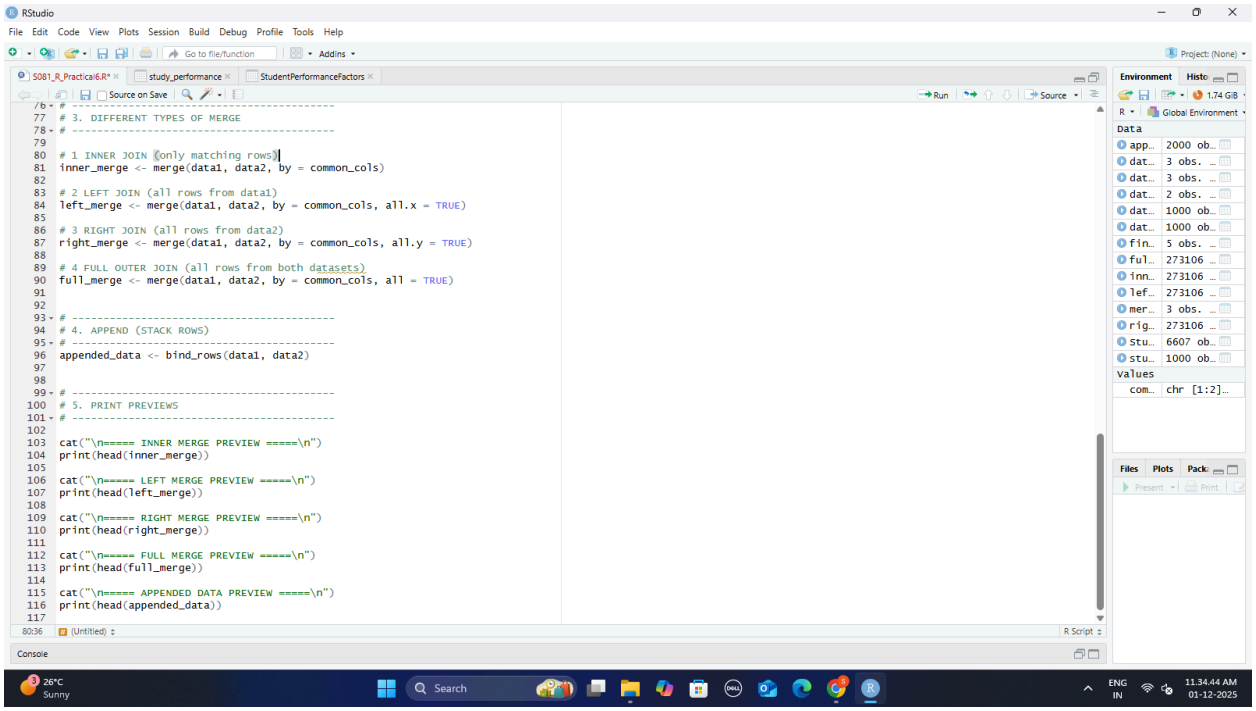
Name :- Priya Gupta

Roll no. :- S081

Sheth L.U.J. & Sir M.V. College



```
#============================================ With Uploaded Data sets=============================================


# PRACT-6 : MERGE + APPEND on Your Data sets with Data sets

library(dplyr)

# ----------------------------------------
# 1. LOAD BOTH DATASETS
# ----------------------------------------
data1 <- read.csv("C:/Users/itlab/OneDrive/Desktop/S081_R_Studio/study_performance.csv")
data2 <- read.csv("C:\\Users\\itlab\\OneDrive\\Desktop\\S081_R_Studio\\study_performance_predition.csv")

cat("\n===== DATASET 1 STRUCTURE =====\n")
str(data1)

cat("\n===== DATASET 2 STRUCTURE =====\n")
str(data2)

cat("\n===== DATASET 1 COLUMNS =====\n")
print(names(data1))

cat("\n===== DATASET 2 COLUMNS =====\n")
print(names(data2))


# ----------------------------------------
# 2. FIND COMMON COLUMNS
# ----------------------------------------
common_cols <- intersect(names(data1), names(data2))

cat("\n===== COMMON COLUMNS =====\n")
print(common_cols)


# ----------------------------------------
# 3. DIFFERENT TYPES OF MERGE
# ----------------------------------------

# 1 INNER JOIN (only matching rows)
inner_merge <- merge(data1, data2, by = common_cols)
```

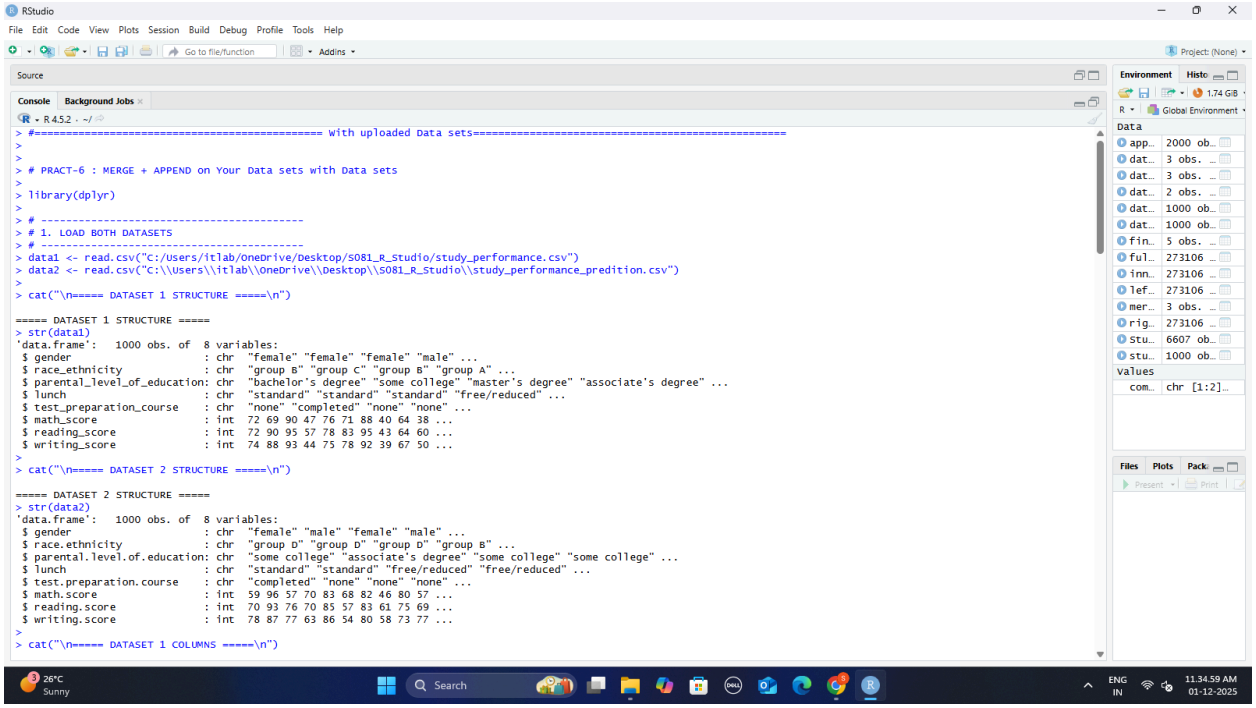


```
# ----------------------------------------
# 3. DIFFERENT TYPES OF MERGE
# ----------------------------------------

# 1 INNER JOIN (only matching rows)
inner_merge <- merge(data1, data2, by = common_cols)

# 2 LEFT JOIN (all rows from data1)
left_merge <- merge(data1, data2, by = common_cols, all.x = TRUE)

# 3 RIGHT JOIN (all rows from data2)
right_merge <- merge(data1, data2, by = common_cols, all.y = TRUE)

# 4 FULL OUTER JOIN (all rows from both datasets)
full_merge <- merge(data1, data2, by = common_cols, all = TRUE)


# ----------------------------------------
# 4. APPEND (STACK ROWS)
# ----------------------------------------
appended_data <- bind_rows(data1, data2)


# ----------------------------------------
# 5. PRINT PREVIEWS
# ----------------------------------------

cat("\n===== INNER MERGE PREVIEW =====\n")
print(head(inner_merge))

cat("\n===== LEFT MERGE PREVIEW =====\n")
print(head(left_merge))

cat("\n===== RIGHT MERGE PREVIEW =====\n")
print(head(right_merge))

cat("\n===== FULL MERGE PREVIEW =====\n")
print(head(full_merge))

cat("\n===== APPENDED DATA PREVIEW =====\n")
print(head(appended_data))
```

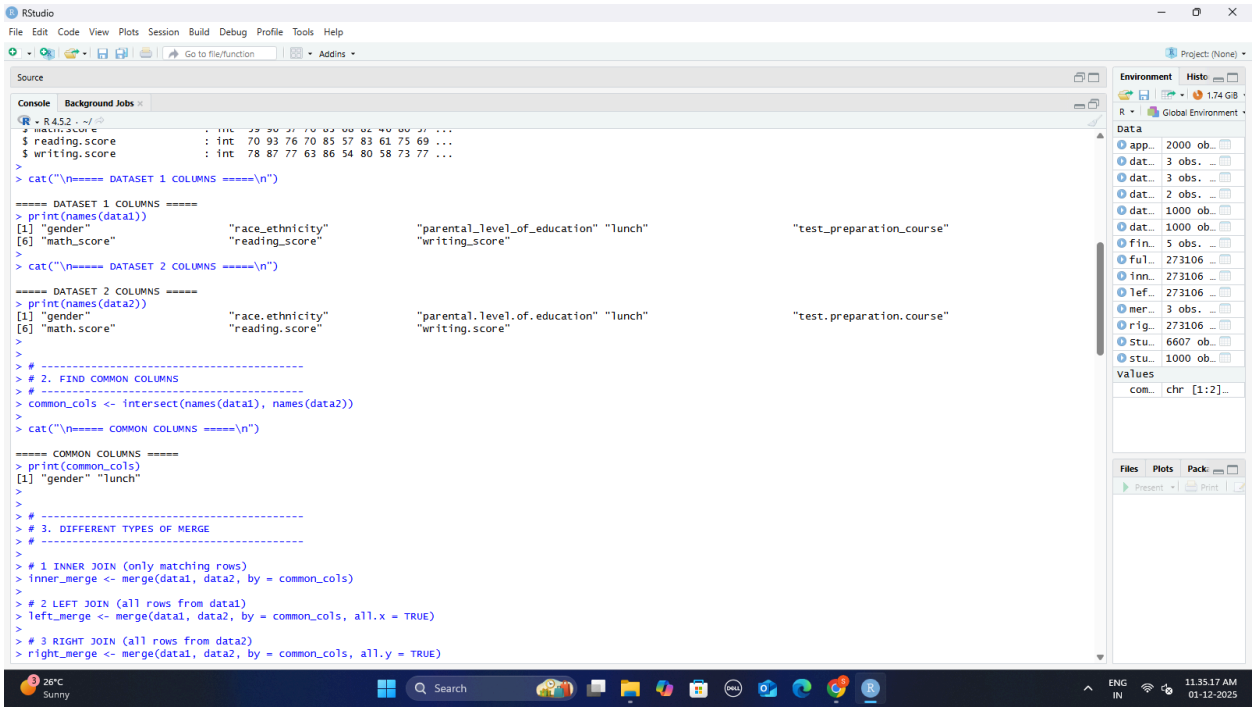
Name :- Priya Gupta

Roll no. :- S081

Sheth L.U.J. & Sir M.V. College



```
> #============================================== With uploaded Data sets=================================================
>
>
> # PRACT-6 : MERGE + APPEND on Your Data sets with Data sets
>
> library(dplyr)
>
> # ------------------------------------------
> # 1. LOAD BOTH DATASETS
> # ------------------------------------------
> data1 <- read.csv("C:/Users/itlab/OneDrive/Desktop/S081_R_Studio/study_performance.csv")
> data2 <- read.csv("C:\\Users\\itlab\\OneDrive\\Desktop\\S081_R_Studio\\study_performance_predition.csv")
>
> cat("\n===== DATASET 1 STRUCTURE =====\n")

===== DATASET 1 STRUCTURE =====
> str(data1)
'data.frame':   1000 obs. of  8 variables:
 $ gender                    : chr  "female" "female" "female" "male" ...
 $ race_ethnicity            : chr  "group B" "group C" "group B" "group A" ...
 $ parental_level_of_education: chr  "bachelor's degree" "some college" "master's degree" "associate's degree" ...
 $ lunch                     : chr  "standard" "standard" "standard" "free/reduced" ...
 $ test_preparation_course   : chr  "none" "completed" "none" "none" ...
 $ math_score                : int  72 69 90 47 76 71 88 40 64 38 ...
 $ reading_score             : int  72 90 95 57 78 83 95 43 64 60 ...
 $ writing_score             : int  74 88 93 44 75 78 92 39 67 50 ...
>
> cat("\n===== DATASET 2 STRUCTURE =====\n")

===== DATASET 2 STRUCTURE =====
> str(data2)
'data.frame':   1000 obs. of  8 variables:
 $ gender                    : chr  "female" "male" "female" "male" ...
 $ race.ethnicity            : chr  "group D" "group D" "group D" "group B" ...
 $ parental.level.of.education: chr  "some college" "associate's degree" "some college" "some college" ...
 $ lunch                     : chr  "standard" "standard" "free/reduced" "free/reduced" ...
 $ test.preparation.course   : chr  "completed" "none" "none" "none" ...
 $ math.score                : int  59 96 57 70 83 68 82 46 80 57 ...
 $ reading.score             : int  70 93 76 70 85 57 83 61 75 69 ...
 $ writing.score             : int  78 87 77 63 86 54 80 58 73 77 ...
>
> cat("\n===== DATASET 1 COLUMNS =====\n")
```

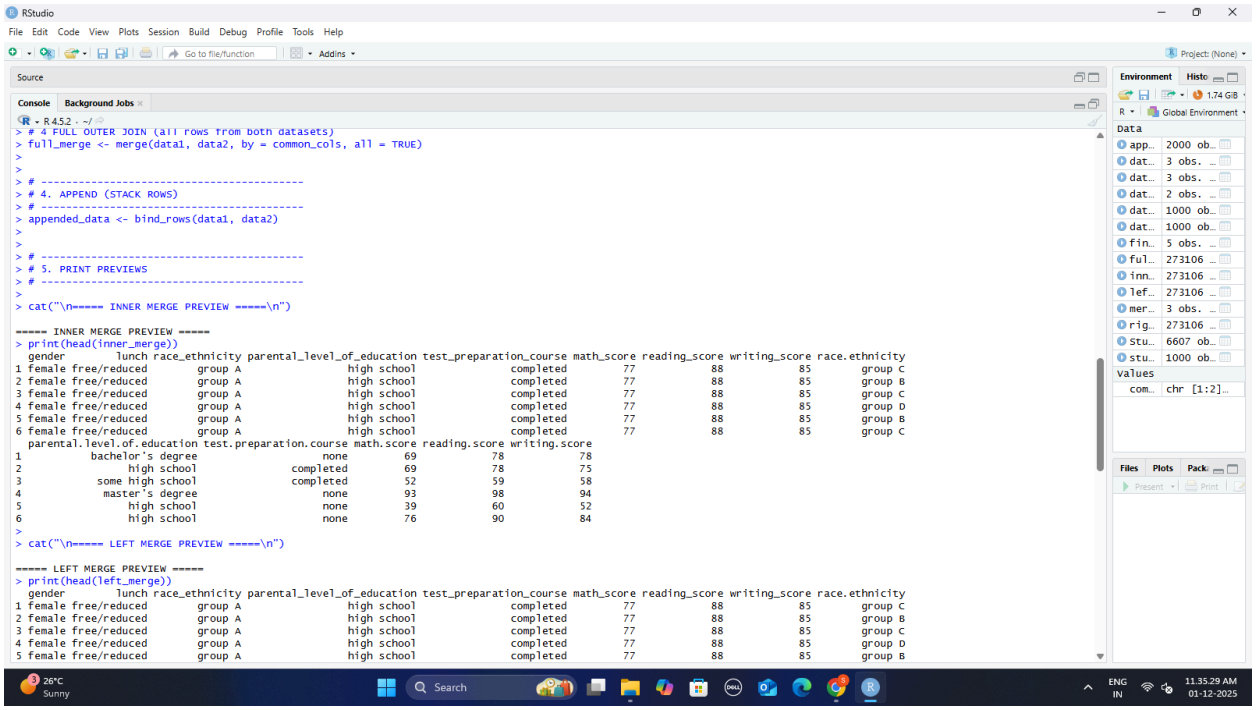


```
 $ reading.score             : int  70 93 76 70 85 57 83 61 75 69 ...
 $ writing.score             : int  78 87 77 63 86 54 80 58 73 77 ...
>
> cat("\n===== DATASET 1 COLUMNS =====\n")

===== DATASET 1 COLUMNS =====
> print(names(data1))
[1] "gender"                 "race_ethnicity"         "parental_level_of_education" "lunch"                      "test_preparation_course"
[6] "math_score"             "reading_score"          "writing_score"
>
> cat("\n===== DATASET 2 COLUMNS =====\n")

===== DATASET 2 COLUMNS =====
> print(names(data2))
[1] "gender"                 "race.ethnicity"         "parental.level.of.education" "lunch"                      "test.preparation.course"
[6] "math.score"             "reading.score"          "writing.score"
>
>
> # ------------------------------------------
> # 2. FIND COMMON COLUMNS
> # ------------------------------------------
> common_cols <- intersect(names(data1), names(data2))
>
> cat("\n===== COMMON COLUMNS =====\n")

===== COMMON COLUMNS =====
> print(common_cols)
[1] "gender" "lunch"
>
>
> # ------------------------------------------
> # 3. DIFFERENT TYPES OF MERGE
> # ------------------------------------------
>
> # 1 INNER JOIN (only matching rows)
> inner_merge <- merge(data1, data2, by = common_cols)
>
> # 2 LEFT JOIN (all rows from data1)
> left_merge <- merge(data1, data2, by = common_cols, all.x = TRUE)
>
> # 3 RIGHT JOIN (all rows from data2)
> right_merge <- merge(data1, data2, by = common_cols, all.y = TRUE)
```

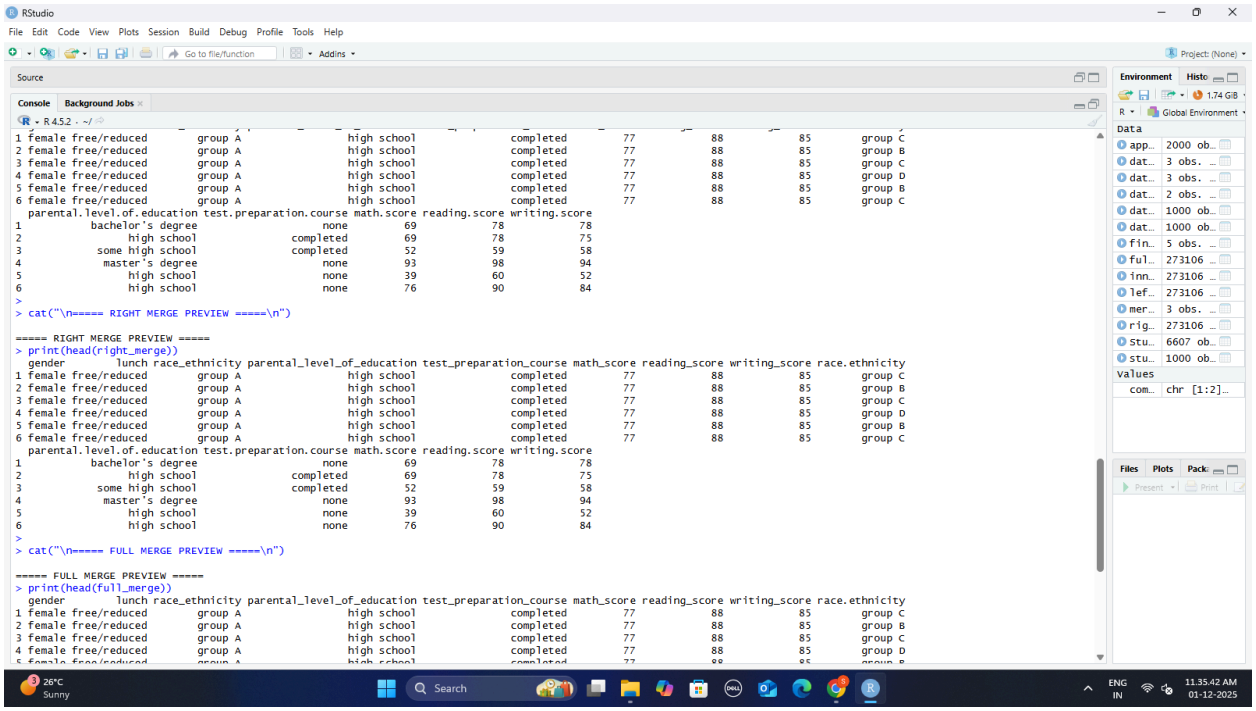
Name :- Priya Gupta

Roll no. :- S081

Sheth L.U.J. & Sir M.V. College



```
> # 4 FULL OUTER JOIN (all rows from both datasets)
> full_merge <- merge(data1, data2, by = common_cols, all = TRUE)
>
>
> # ----------------------------------------
> # 4. APPEND (STACK ROWS)
> # ----------------------------------------
> appended_data <- bind_rows(data1, data2)
>
>
> # ----------------------------------------
> # 5. PRINT PREVIEWS
> # ----------------------------------------
>
> cat("\n===== INNER MERGE PREVIEW =====\n")

===== INNER MERGE PREVIEW =====
> print(head(inner_merge))
  gender        lunch race_ethnicity parental_level_of_education test_preparation_course math_score reading_score writing_score race.ethnicity
1 female free/reduced        group A                high school                completed         77            88            85        group C
2 female free/reduced        group A                high school                completed         77            88            85        group B
3 female free/reduced        group A                high school                completed         77            88            85        group C
4 female free/reduced        group A                high school                completed         77            88            85        group D
5 female free/reduced        group A                high school                completed         77            88            85        group B
6 female free/reduced        group A                high school                completed         77            88            85        group C
  parental.level.of.education test.preparation.course math.score reading.score writing.score
1           bachelor's degree                    none         69            78            78
2                 high school               completed         69            78            75
3            some high school               completed         52            59            58
4             master's degree                    none         93            98            94
5                 high school                    none         39            60            52
6                 high school                    none         76            90            84
>
> cat("\n===== LEFT MERGE PREVIEW =====\n")

===== LEFT MERGE PREVIEW =====
> print(head(left_merge))
  gender        lunch race_ethnicity parental_level_of_education test_preparation_course math_score reading_score writing_score race.ethnicity
1 female free/reduced        group A                high school                completed         77            88            85        group C
2 female free/reduced        group A                high school                completed         77            88            85        group B
3 female free/reduced        group A                high school                completed         77            88            85        group C
4 female free/reduced        group A                high school                completed         77            88            85        group D
5 female free/reduced        group A                high school                completed         77            88            85        group B
```

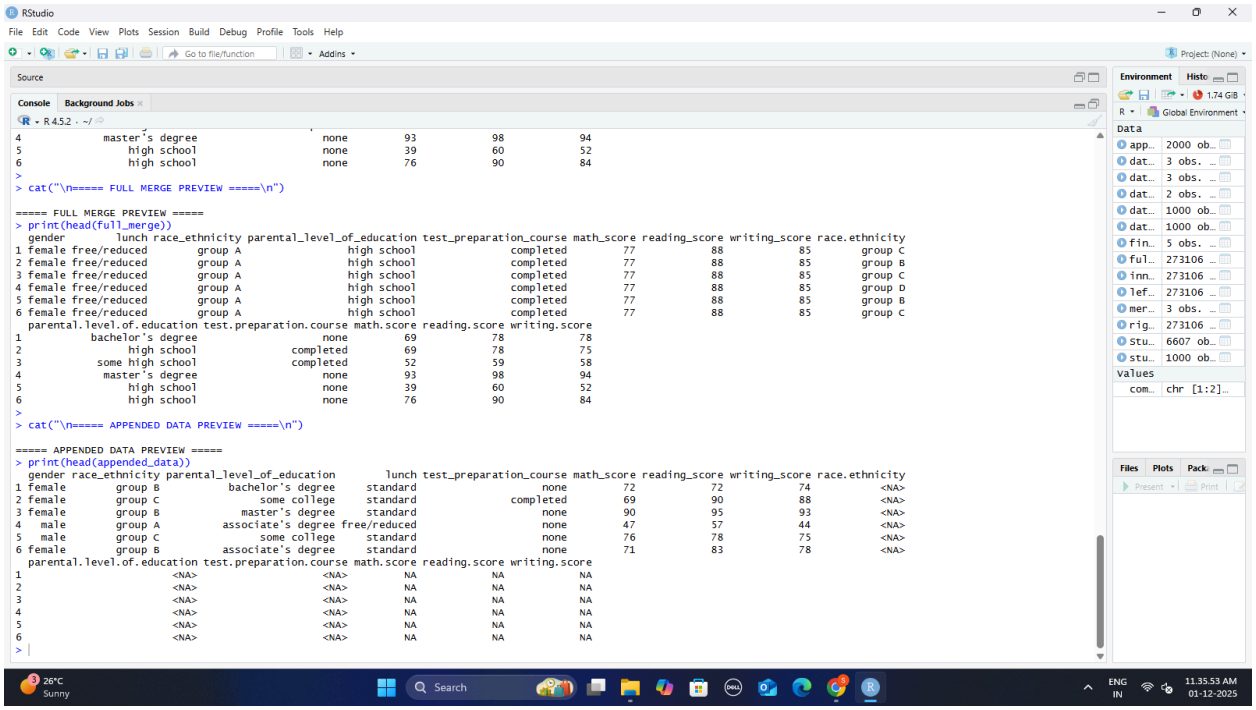


```
1 female free/reduced        group A                high school                completed         77            88            85        group C
2 female free/reduced        group A                high school                completed         77            88            85        group B
3 female free/reduced        group A                high school                completed         77            88            85        group C
4 female free/reduced        group A                high school                completed         77            88            85        group D
5 female free/reduced        group A                high school                completed         77            88            85        group B
6 female free/reduced        group A                high school                completed         77            88            85        group C
  parental.level.of.education test.preparation.course math.score reading.score writing.score
1           bachelor's degree                    none         69            78            78
2                 high school               completed         69            78            75
3            some high school               completed         52            59            58
4             master's degree                    none         93            98            94
5                 high school                    none         39            60            52
6                 high school                    none         76            90            84
>
> cat("\n===== RIGHT MERGE PREVIEW =====\n")

===== RIGHT MERGE PREVIEW =====
> print(head(right_merge))
  gender        lunch race_ethnicity parental_level_of_education test_preparation_course math_score reading_score writing_score race.ethnicity
1 female free/reduced        group A                high school                completed         77            88            85        group C
2 female free/reduced        group A                high school                completed         77            88            85        group B
3 female free/reduced        group A                high school                completed         77            88            85        group C
4 female free/reduced        group A                high school                completed         77            88            85        group D
5 female free/reduced        group A                high school                completed         77            88            85        group B
6 female free/reduced        group A                high school                completed         77            88            85        group C
  parental.level.of.education test.preparation.course math.score reading.score writing.score
1           bachelor's degree                    none         69            78            78
2                 high school               completed         69            78            75
3            some high school               completed         52            59            58
4             master's degree                    none         93            98            94
5                 high school                    none         39            60            52
6                 high school                    none         76            90            84
>
> cat("\n===== FULL MERGE PREVIEW =====\n")

===== FULL MERGE PREVIEW =====
> print(head(full_merge))
  gender        lunch race_ethnicity parental_level_of_education test_preparation_course math_score reading_score writing_score race.ethnicity
1 female free/reduced        group A                high school                completed         77            88            85        group C
2 female free/reduced        group A                high school                completed         77            88            85        group B
3 female free/reduced        group A                high school                completed         77            88            85        group C
4 female free/reduced        group A                high school                completed         77            88            85        group D
```

Name :- Priya Gupta

Roll no. :- S081

Sheth L.U.J. & Sir M.V. College



```
4              master's degree                 none         93            98            94
5                  high school                 none         39            60            52
6                  high school                 none         76            90            84
>
> cat("\n===== FULL MERGE PREVIEW =====\n")

===== FULL MERGE PREVIEW =====
> print(head(full_merge))
  gender        lunch race_ethnicity parental_level_of_education test_preparation_course math_score reading_score writing_score race.ethnicity
1 female free/reduced        group A                high school                completed         77            88            85        group C
2 female free/reduced        group A                high school                completed         77            88            85        group B
3 female free/reduced        group A                high school                completed         77            88            85        group C
4 female free/reduced        group A                high school                completed         77            88            85        group D
5 female free/reduced        group A                high school                completed         77            88            85        group B
6 female free/reduced        group A                high school                completed         77            88            85        group C
  parental.level.of.education test.preparation.course math.score reading.score writing.score
1           bachelor's degree                    none         69            78            78
2                some college               completed         69            78            75
3                some college               completed         52            59            58
4           master's degree                    none         93            98            94
5                high school                    none         39            60            52
6                high school                    none         76            90            84
>
> cat("\n===== APPENDED DATA PREVIEW =====\n")

===== APPENDED DATA PREVIEW =====
> print(head(appended_data))
  gender race_ethnicity parental_level_of_education        lunch test_preparation_course math_score reading_score writing_score race.ethnicity
1 female        group B           bachelor's degree     standard                    none         72            72            74           <NA>
2 female        group C                some college     standard                completed         69            90            88           <NA>
3 female        group B             master's degree     standard                    none         90            95            93           <NA>
4   male        group A          associate's degree free/reduced                    none         47            57            44           <NA>
5   male        group C                some college     standard                    none         76            78            75           <NA>
6 female        group B          associate's degree     standard                    none         71            83            78           <NA>
  parental.level.of.education test.preparation.course math.score reading.score writing.score
1                        <NA>                    <NA>         NA            NA            NA
2                        <NA>                    <NA>         NA            NA            NA
3                        <NA>                    <NA>         NA            NA            NA
4                        <NA>                    <NA>         NA            NA            NA
5                        <NA>                    <NA>         NA            NA            NA
6                        <NA>                    <NA>         NA            NA            NA
>
```

Name :- Priya Gupta

Roll no. :- S081