**R LAB – 07**

**Task – 01 :**

**Aim :** Apply cbind , rbind , cast , melt functions on any inbuilt dataset in R soft

* Ware .

**Program :**

# package installations

require("plyr")

library("plyr")

install.packages("data.table")

require("data.table")

library("data.table")

install.packages("reshape2")

require("reshape2")

library("reshape2")

sport <- c("Hockey", "Baseball", "Football")

league <- c("NHL", "MLB", "NFL")

trophy <- c("Stanley Cup", "Commissioner’s Trophy", "Vince Lombardi Trophy")

trophies1 <- cbind(sport,league,trophy) # vectors will become as columns

trophies0 <- rbind(sport,league,trophy) # vectors become as rows

trophies0

trophies1

class(trophies1)

trophies2<- data.frame(sport=c("Basketball", "Golf"), league=c("NBA", "PGA"),

trophy=c("Larry Championship Trophy", "Wanamaker Trophy"), stringsAsFactors=FALSE)

trophies2

trophies <- rbind(trophies1, trophies2)

trophies

# Doing the JOINS

getwd()

setwd("C:/Users/itadmin/Documents/")

getwd()

df1 = data.frame(StudentId = c(101:106),

Product = c("Hindi", "English",

"Maths", "Science",

"Political Science",

"Physics"))

df1

df2 = data.frame(StudentId = c(102, 104, 106,

107, 108),

State = c("Manglore", "Mysore",

"Pune", "Dehradun", "Delhi"))

df2

df = merge(x = df1, y = df2, by = "StudentId")

df

# left outer

df3 = merge(x = df1, y = df2, by = "StudentId",

all.x = TRUE)

df3

# right outer

df4 = merge(x = df1, y = df2, by = "StudentId",

all.y = TRUE)

df4

# full join

df5 = merge(x = df1, y = df2, by = "StudentId",

all = TRUE)

df5

# cross join

df6 = merge(x = df1, y = df2, by = NULL)

df6

# semi join (doubt topic)

install.packages("dplyr")

library(dplyr)

df7 = df1 %>% semi\_join(df2, by = "StudentId")

df7

# anti join (doubt topic)

df8 = df1 %>% anti\_join(df2, by = "StudentId")

df8

# doing CAST and MELT

# 1. CAST : transforming rows into columns

# 2. MELT : transforming columns into row

# ships data set is a default dat set in the r packages.

install.packages("MASS")

install.packages("reshape")

library("MASS")

library("reshape")

library("reshape2")

ships

sd = (head(ships, n = 10)) # taking as data frame

class(sd)

# in id attribute putting the type and year column constant

molten\_ships = melt(sd, id = c("type","year"))

molten\_ships

rec <- reshape2::dcast(molten\_ships, type+year~variable,sum)

rec

**Output :**

sport <- c("Hockey", "Baseball", "Football")

> league <- c("NHL", "MLB", "NFL")

> trophy <- c("Stanley Cup", "Commissioner’s Trophy", "Vince Lombardi Trophy")

> trophies1 <- cbind(sport,league,trophy) # vectors will become as columns

> trophies0 <- rbind(sport,league,trophy) # vectors become as rows

> trophies0

[,1] [,2] [,3]

sport "Hockey" "Baseball" "Football"

league "NHL" "MLB" "NFL"

trophy "Stanley Cup" "Commissioner’s Trophy" "Vince Lombardi Trophy"

> trophies1

sport league trophy

[1,] "Hockey" "NHL" "Stanley Cup"

[2,] "Baseball" "MLB" "Commissioner’s Trophy"

[3,] "Football" "NFL" "Vince Lombardi Trophy"

> class(trophies1)

[1] "matrix" "array"

trophies2<- data.frame(sport=c("Basketball", "Golf"), league=c("NBA", "PGA"),

+ trophy=c("Larry Championship Trophy", "Wanamaker Trophy"), stringsAsFactors=FALSE)

> trophies2

sport league trophy

1 Basketball NBA Larry Championship Trophy

2 Golf PGA Wanamaker Trophy

>

> trophies <- rbind(trophies1, trophies2)

> trophies

sport league trophy

1 Hockey NHL Stanley Cup

2 Baseball MLB Commissioner’s Trophy

3 Football NFL Vince Lombardi Trophy

4 Basketball NBA Larry Championship Trophy

5 Golf PGA Wanamaker Trophy

getwd()

[1] "E:/venkat sai/rstudio\_language"

> setwd(choose.dir())

> getwd()

[1] "E:/venkat sai/rstudio\_language"

df1 = data.frame(StudentId = c(101:106),

+ Product = c("Hindi", "English",

+ "Maths", "Science",

+ "Political Science",

+ "Physics"))

> df1

StudentId Product

1 101 Hindi

2 102 English

3 103 Maths

4 104 Science

5 105 Political Science

6 106 Physics

>

> df2 = data.frame(StudentId = c(102, 104, 106,

+ 107, 108),

+ State = c("Manglore", "Mysore",

+ "Pune", "Dehradun", "Delhi"))

> df2

StudentId State

1 102 Manglore

2 104 Mysore

3 106 Pune

4 107 Dehradun

5 108 Delhi

df = merge(x = df1, y = df2, by = "StudentId")

> df

StudentId Product State

1 102 English Manglore

2 104 Science Mysore

3 106 Physics Pune

>

> # left outer

> df3 = merge(x = df1, y = df2, by = "StudentId",

+ all.x = TRUE)

> df3

StudentId Product State

1 101 Hindi <NA>

2 102 English Manglore

3 103 Maths <NA>

4 104 Science Mysore

5 105 Political Science <NA>

6 106 Physics Pune

>

> # right outer

> df4 = merge(x = df1, y = df2, by = "StudentId",

+ all.y = TRUE)

> df4

StudentId Product State

1 102 English Manglore

2 104 Science Mysore

3 106 Physics Pune

4 107 <NA> Dehradun

5 108 <NA> Delhi

>

> # full join

> df5 = merge(x = df1, y = df2, by = "StudentId",

+ all = TRUE)

> df5

StudentId Product State

1 101 Hindi <NA>

2 102 English Manglore

3 103 Maths <NA>

4 104 Science Mysore

5 105 Political Science <NA>

6 106 Physics Pune

7 107 <NA> Dehradun

8 108 <NA> Delhi

>

> # cross join

> df6 = merge(x = df1, y = df2, by = NULL)

> df6

StudentId.x Product StudentId.y State

1 101 Hindi 102 Manglore

2 102 English 102 Manglore

3 103 Maths 102 Manglore

4 104 Science 102 Manglore

5 105 Political Science 102 Manglore

6 106 Physics 102 Manglore

7 101 Hindi 104 Mysore

8 102 English 104 Mysore

9 103 Maths 104 Mysore

10 104 Science 104 Mysore

11 105 Political Science 104 Mysore

12 106 Physics 104 Mysore

13 101 Hindi 106 Pune

14 102 English 106 Pune

15 103 Maths 106 Pune

16 104 Science 106 Pune

17 105 Political Science 106 Pune

18 106 Physics 106 Pune

19 101 Hindi 107 Dehradun

20 102 English 107 Dehradun

21 103 Maths 107 Dehradun

22 104 Science 107 Dehradun

23 105 Political Science 107 Dehradun

24 106 Physics 107 Dehradun

25 101 Hindi 108 Delhi

26 102 English 108 Delhi

27 103 Maths 108 Delhi

28 104 Science 108 Delhi

29 105 Political Science 108 Delhi

30 106 Physics 108 Delhi

df7 = df1 %>% semi\_join(df2, by = "StudentId")

> df7

StudentId Product

1 102 English

2 104 Science

3 106 Physics

>

> # anti join (doubt topic)

> df8 = df1 %>% anti\_join(df2, by = "StudentId")

> df8

StudentId Product

1 101 Hindi

2 103 Maths

3 105 Political Science

library("reshape2")

> ships

type year period service incidents

1 A 60 60 127 0

2 A 60 75 63 0

3 A 65 60 1095 3

4 A 65 75 1095 4

5 A 70 60 1512 6

6 A 70 75 3353 18

7 A 75 60 0 0

8 A 75 75 2244 11

9 B 60 60 44882 39

10 B 60 75 17176 29

11 B 65 60 28609 58

12 B 65 75 20370 53

13 B 70 60 7064 12

14 B 70 75 13099 44

15 B 75 60 0 0

16 B 75 75 7117 18

17 C 60 60 1179 1

18 C 60 75 552 1

19 C 65 60 781 0

20 C 65 75 676 1

21 C 70 60 783 6

22 C 70 75 1948 2

23 C 75 60 0 0

24 C 75 75 274 1

25 D 60 60 251 0

26 D 60 75 105 0

27 D 65 60 288 0

28 D 65 75 192 0

29 D 70 60 349 2

30 D 70 75 1208 11

31 D 75 60 0 0

32 D 75 75 2051 4

33 E 60 60 45 0

34 E 60 75 0 0

35 E 65 60 789 7

36 E 65 75 437 7

37 E 70 60 1157 5

38 E 70 75 2161 12

39 E 75 60 0 0

40 E 75 75 542 1

> sd = (head(ships, n = 10)) # taking as data frame

> class(sd)

[1] "data.frame"

molten\_ships = melt(sd, id = c("type","year"))

> molten\_ships

type year variable value

1 A 60 period 60

2 A 60 period 75

3 A 65 period 60

4 A 65 period 75

5 A 70 period 60

6 A 70 period 75

7 A 75 period 60

8 A 75 period 75

9 B 60 period 60

10 B 60 period 75

11 A 60 service 127

12 A 60 service 63

13 A 65 service 1095

14 A 65 service 1095

15 A 70 service 1512

16 A 70 service 3353

17 A 75 service 0

18 A 75 service 2244

19 B 60 service 44882

20 B 60 service 17176

21 A 60 incidents 0

22 A 60 incidents 0

23 A 65 incidents 3

24 A 65 incidents 4

25 A 70 incidents 6

26 A 70 incidents 18

27 A 75 incidents 0

28 A 75 incidents 11

29 B 60 incidents 39

30 B 60 incidents 29

rec <- dcast(molten\_ships, type+year~variable,sum)

> rec

type year period service incidents

1 A 60 135 190 0

2 A 65 135 2190 7

3 A 70 135 4865 24

4 A 75 135 2244 11

5 B 60 135 62058 68

**Result :** Successfully executed all lines in the program…