SESSION 2: INTRODUCTION to working with R

Assignment 2

1. Problem Statement

1. Read multiple json files into a working directory for further converting into a dataset.

I have files text1, text2, text3 in the directory json.

Solution :

```
library(jsonlite)
library(dplyr)

y1<-read_json("C:\\Users\\vmeiyazhaga2\\OneDrive - DXC Production\\Vadivu
Personal\\ACAD_Gild\\JSON\\text1.json")
View(y1)
result1<- as.data.frame(do.call("rbind",y1))
result1

y2<-read_json("C:\\Users\\vmeiyazhaga2\\OneDrive - DXC Production\\Vadivu
Personal\\ACAD_Gild\\JSON\\text2.json")
View(y2)

result2<- as.data.frame(do.call("rbind",y2))
result2

y3<-read_json("C:\\Users\\vmeiyazhaga2\\OneDrive - DXC Production\\Vadivu
Personal\\ACAD_Gild\\JSON\\text3.json")
View(y3)
result3<- as.data.frame(do.call("rbind",y3))
result3
```

Out Put

```
y1<-read_json("C:\\Users\\vmeiyazhaga2\\OneDrive - DXC Production\\Vadivu Personal\\ACAD_
> View(y1)
> result1<- as.data.frame(do.call("rbind",y1))
> result1
  name species                    foods
1  abx     cat tuna, catnip, ham, zucchini
2  Bcd     dog        bones, carrots, tuna
3  egg     cat          mice, nn, cookies
```

```
> y2<-read_json("C:\\Users\\vmeiyazhaga2\\OneDrive - DXC Production\\Vadivu Personal\\ACA
> View(y2)
> result2<- as.data.frame(do.call("rbind",y2))
> result2
   name species                      foods
1   Meo     cat  tuna, catnip, ham, zucchini
2 Barky     dog         bones, carrots, tuna
3 Purws     cat                    a, nn, s


>
> y3<-read_json("C:\\Users\\vmeiyazhaga2\\OneDrive - DXC Production\\Vadivu Personal\\ACA
> View(y3)
> result3<- as.data.frame(do.call("rbind",y3))
> result3
      name species                      foods
1   Meowsy     cat  tuna, catnip, ham, zucchini
2    Barky     dog         bones, carrots, tuna
3 Purrpaws     cat          mice, nn, cookies
```

2. Parse the following JSON into a data frame

> js<-'{
> "name": null, "release_date_local": null, "title": "3 (2011)",
>
> "opening_weekend_take": 1234, "year": 2011,
> "release_date_wide": "2011-09-16", "gross":
> 59954 }'

js<-'{
"name": null, "release_date_local": null, "title": "3 (2011)",
"opening_weekend_take": 1234, "year": 2011, "release_date_wide": "2011-09-16", "gross":
59954 }'

Solution :

Question2<- fromJSON(js)
Question2

OutPut
> Question2
$`name`
NULL

$release_date_local
NULL

$title
[1] "3 (2011)"

$opening_weekend_take
[1] 1234

$year
[1] 2011

$release_date_wide
[1] "2011-09-16"

$gross
[1] 59954

3. Write a script for variable binning using R.

Solution:

```
#Writing binning() function for dividing the variable named age into 4 bins named as "group1-(1 to 25)","group2-(26 to 50)","group3-(51 to 75)","group4-(76 to 90)"
age <- c(1:90)
age

binning <- function(x)
{
  for(i in c(1:90))
  {
    ifelse(i <= 25, paste(i,"group1"),
        ifelse(i <= 50, paste(i,"group2"),
            ifelse(i <= 75, paste(i,"group3"),
                paste(i,"group4"))))
    break
  }
}

binning(age)

# Example 1:Let us consider a vector consisting of values from 1 to 90 and we need to create 4 bins named "group1","group2","group3","group4".
# VARIABLE BINNING USING cut() function

x<-c(1:90)
cut(x,4,labels=c("group1","group2","group3","group4"))


# Example 2: Import a mtcars.csv file into R-Studio and divide the variable named mpg into 5 bins named "FIRST","SECOND","THIRD","FOURTH" and "FIFTH"

library(readr)
mtcars <- read_csv("C:/Users/vmeiyazhaga2/OneDrive - DXC Production/Vadivu Personal/ACAD_Gild/R_Script/mtcars.csv")
mpg<- mtcars$mpg
mpg
cut(mpg,5)
cut(mpg,5,labels=c("FIRST","SECOND","THIRD","FOURTH","FIFTH"))
```

OutPut:

Example 1:Let us consider a vector consisting of values from 1 to 90 and we need to create 4 bins named "group1","group2","group3","group4".

```r
> # VARIABLE BINNING USING cut() function
>
> x<-c(1:90)

> cut(x,4,labels=c("group1","group2","group3","group4"))

 [1] group1 group1 group1 group1 group1 group1 group1 group1 group1 group1 group1 group1
group1 group1 group1 group1
[17] group1 group1 group1 group1 group1 group1 group1 group2 group2 group2 group2
group2 group2 group2 group2 group2
[33] group2 group2 group2 group2 group2 group2 group2 group2 group2 group2 group2
group2 group2 group3 group3 group3
[49] group3 group3 group3 group3 group3 group3 group3 group3 group3 group3 group3
group3 group3 group3 group3 group3
[65] group3 group3 group3 group4 group4 group4 group4 group4 group4 group4 group4
group4 group4 group4 group4 group4
[81] group4 group4 group4 group4 group4 group4 group4 group4 group4 group4
Levels: group1 group2 group3 group4


> # Example 2: Import a mtcars.csv file into R-Studio and divide the variable named mpg into 5 bins named "FIRST","SECOND","THIRD","FOURTH" and "FIFTH"

> library(readr)
> mtcars <- read_csv("C:/Users/vmeiyazhaga2/OneDrive - DXC Production/Vadivu
Personal/ACAD_Gild/R_Script/mtcars.csv")
Parsed with column specification:
cols(
  model = col_character(),
  mpg = col_double(),
  cyl = col_integer(),
  disp = col_double(),
  hp = col_integer(),
  drat = col_double(),
  wt = col_double(),
  qsec = col_double(),
  vs = col_integer(),
  am = col_integer(),
  gear = col_integer(),
  carb = col_integer()
)


> mpg<- mtcars$mpg
> mpg
 [1] 21.0 21.0 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 17.8 16.4 17.3 15.2 10.4 10.4 14.7 32.4
30.4 33.9 21.5 15.5 15.2
[24] 13.3 19.2 27.3 26.0 30.4 15.8 19.7 15.0 21.4

> cut(mpg,5)
 [1] (19.8,24.5] (19.8,24.5] (19.8,24.5] (19.8,24.5] (15.1,19.8] (15.1,19.8] (10.4,15.1] (19.8,24.5]
(19.8,24.5]
[10] (15.1,19.8] (15.1,19.8] (15.1,19.8] (15.1,19.8] (15.1,19.8] (10.4,15.1] (10.4,15.1] (10.4,15.1]
(29.2,33.9]
[19] (29.2,33.9] (29.2,33.9] (19.8,24.5] (15.1,19.8] (15.1,19.8] (10.4,15.1] (15.1,19.8] (24.5,29.2]
(24.5,29.2]
[28] (29.2,33.9] (15.1,19.8] (15.1,19.8] (10.4,15.1] (19.8,24.5]
Levels: (10.4,15.1] (15.1,19.8] (19.8,24.5] (24.5,29.2] (29.2,33.9]

> cut(mpg,5,labels=c("FIRST","SECOND","THIRD","FOURTH","FIFTH"))
 [1] THIRD  THIRD  THIRD  THIRD  SECOND SECOND FIRST  THIRD  THIRD  SECOND SECOND
SECOND SECOND SECOND FIRST  FIRST
[17] FIRST  FIFTH  FIFTH  FIFTH  THIRD  SECOND SECOND FIRST  SECOND FOURTH FOURTH FIFTH
SECOND SECOND FIRST  THIRD
Levels: FIRST SECOND THIRD FOURTH FIFTH
```