Data Analytics

SESSION 2: INTRODUCTION to working with R

Assignment 2.2

1

Data Analytics

Table of Contents

1.Introduction ............................................................................................................................................... 3

2.Objective .................................................................................................................................................... 3

3.Prerequisites .............................................................................................................................................. 3

4.Associated Data Files ................................................................................................................................. 3

5.Problem Statement .................................................................................................................................... 3

6.Expected Output ........................................................................................................................................ 3

7.Approximate Time to Complete Task ........................................................................................................ 3

2

Data Analytics

**1. Introduction**

This assignment will help you understand the concepts learnt in the session.

**2. Objective**

This assignment will test your skills on the basics of R.

**3. Prerequisites**

Not applicable.

**4. Associated Data Files**

Not applicable.

**5. 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.

**Answer:**

**install.packages("rjson") # install the package**

**setwd() # set the working directory**

**library(rjson) # load the package**

**json1 <- fromJSON(file = "text1.json") # import the first file in R**

**json2 <- fromJSON(file = "text2.json") #import the second file in R**

**json3 <- fromJSON(file = "text3.json") # import the third file in R**

**print(json1) # print first file.**

**print(json2) #print second file**

**print(json3) # print third file.**

**# to convert Json into Data frame.,**

**Json\_data <- data.frame(json1, json2, json3)**

**Print(json\_data)**

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 }'

Answer:

install.package(RJSONIO) # install the package RJSONIO

require(RJSONIO) # to load RJSONIO

js<-'{ "name": null, "release\_date\_local": null, "title": "3 (2011)", "opening\_weekend\_take": 1234, "year": 2011, "release\_date\_wide": "2011-09-16", "gross": 59954 }' # call object js

js <-fromJSON(js)

print(js)

# convert it into data frame.

do.call(“cbind”, js) # use do.call method to convert the JSON

Output will be:

title "opening\_weekend\_take year release\_date\_wide gross

[1,] "3 (2011)" "1234" "2011" "2011-09-16" "59954"

3. Write a script for variable binning using R.

Answer:

x <- c(4,7,9,1,10,15,18,109,3,160,100,16,120,22,2,23,16,17)

binned.x <- as.factor(ifelse(x>50, "10+",x))

binned.x

Output:

[1] 4 7 9 1 10 15 18 10+ 3 10+ 10+ 16 10+ 22 2 23 16 17

Levels: 1 10 10+ 15 16 17 18 2 22 23 3 4 7 9

Data Analytics

**6. Expected Format**

1. R file should be submitted where applicable. 2. R file should be in PDF or in .r format 3. Proper screenshots of the outputs should be submitted as well 4. The r codes, if submitted in any other format, will be subjected to deduction in marks

Note: Your solution will not be entertained if it is any other format, e.g., .zip, .doc, .rtf etc.

**7. Approximate Time to Complete Task**

20 mins.

4