# Strategy

The strategy is to read the JSON file using the fromJSON function of the jsonlite package. The output will be presented as a list of lists. Read individual lists, and with the help of rapply and unique functions, extract the value of the labels. Repeat this for all the data that is required to form a data frame.

The value section of the JSON file returns the elements in the form of a numeric vector. Read the vectors by adding three into their indexes and assign them to a new variable. Remember to start from the first, second, and third place to read the right element. Repeat this logic three times to create three variables. Use the same logic and create two more variables, one for the year and another for statistics.

# Code

Here is the working copy of the code for your scrutiny. Please comment if you have a better and more optimized way of handling this data.

*############################################################################# ###*

*## ## Reading a JSON file and preparing data for analysis ############################################################################# ###*

*#Using jsonlite to read .json file*

**library** jsonlite

*#Using function fromJSON from jsonlite package to read the file*

djson <- fromJSON( https://statbank.cso.ie/StatbankServices/StatbankServices.svc/jsonservice/resp onseinstance/CIS78

*#Preparing the data frame from the list of lists djson created above #Reading individual lists and preparing columns*

df <- data.frame

*#Reading dimension Type of Cooperation Partner*

unique(rapply(djson$dataset$dimension$`Type of Cooperation Pa rtner`$category$label, **function** lst) head(lst, ))),

*#Reading first and every other third value from there for eac*

*h observation*

)]

*h observation*

)]

*h observation*

)]

V2 = djson$dataset$value[seq( , length(djson$dataset$value), *#Reading second and every other third value from there for ea* V3 = djson$dataset$value[seq( , length(djson$dataset$value), *#Reading third and every other third value from there for eac* V4 = djson$dataset$value[seq( , length(djson$dataset$value), *#Reading first and every other third value from there for eac*

*h observation but for dimension called year*

V5 = djson$dataset$value[seq( , length(djson$dataset$value),

)],

*#Reading first and every other third value from there for eac*

*h observation but for dimension called Statistic*

V6 = djson$dataset$value[seq( , length(djson$dataset$value),

)]

## 'data.frame':

12 obs. of 6 variables

## $ Type of Cooperation Partner

chr "Any type of cooperation" "Cooperation from clients and or customers" "Cooperation from competitors" "Cooperation other enterprises within own ente rprise group" ..

## $ Industries (05 to 39)

num 54.7 34.9 21.5 29 30 45.8 43.8 27 15.9 44.6 ..

## $ Industries and selected services (05 to 39,46,49 to 53,58 to 63,64 to 6

6,71 to 73): num 50.8 32.9 20.2 27.4 25.8 40.1 38.7 23.3 17.3 41.9 ..

## $ Selected Services (46, 49-53, 58-63, 64-66, 71-73)

num 47.8 31.5 19.3 26.3 22.7 35.9 34.8 20.5 18.5 39.9 ..

## $ 2018

num 54.7 34.9 21.5 29 30 45.8 43.8 27 15.9 44.6 ..

## $ Co-operation by Technological Innovative Enterprises (%) num 50.8 32.9 20.2 27.4 25.8 40.1 38.7 23.3 17.3 41.9 ..

*#Printing data frame*

df

##

Type of Cooperation Partner ## 1

Any type of cooperation ## 2

ation from clients and or customer ## 3

Cooperation from competitors ## 4

rprises within own enterprise group ## 5

ies and or third level institution ## 6

t, materials, components or software

Cooper

Cooperation other ente Cooperation from Universit

Cooperation from suppliers of equipmen

## 7 Cooperation from consultants and or commercial laboratories or private

research and development institutes

## 8

nment or public research institutes ## 9

public sector clients or customers

Cooperation from Gover

Cooperation from

*#Assigning column names from vectors to match the data presented on the site given below*

*# https://data.gov.ie/dataset/7b6c5d4c-955c-4eeb-a9d0-e35fb58bf200/resource/5a856b72-f470-4c71-ab 1f-fbb0ef3b1e22#&r=Type%20of%20Cooperation%20Partner&c=NACE%20Rev%202%20Sector* colnames(df) = c(djson$dataset$dimension$`Type of Cooperation Partner`$label,

unique(rapply(djson$dataset$dimension$`NACE Rev 2 Sector`$ca tegory$label, **function** lst) head(lst, )))

unique(rapply(djson$dataset$dimension$Year$category$label,

**unction** lst) head(lst, )))

unique(rapply(djson$dataset$dimension$Statistic$category$ ab el, **function** lst) head(lst, )))

*#Structure of the data frame*

str(df

## 10 Cooperation from private business enterp rises outside your enterprise group

## 11

Cooperation from other enterprises

## 12 Cooper

ation from non-profit organisations ## Industries (05 to 39)

## 1 54.7

## 2 34.9

## 3 21.5

## 4 29.0

## 5 30.0

## 6 45.8

## 7 43.8

## 8 27.0

## 9 15.9

## 10 44.6

## 11 22.7

## 12 11.3

## Industries and selected services (05 to 39,46,49 to 53,58 to 63,64 to 6

6,71 to 73)

## 1

50.8

## 2

32.9

## 3

20.2

## 4

27.4

## 5

25.8

## 6

40.1

## 7

38.7

## 8

23.3

## 9

17.3

## 10

41.9

## 11

21.5

## 12

12.5

## Selected Services (46, 49-53, 58-63, 64-66, 71-73) 2018

## 1 47.8 54.7

## 2 31.5 34.9

## 3 19.3 21.5

## 4 26.3 29.0

## 5 22.7 30.0

## 6 35.9 45.8

## 7 34.8 43.8

## 8 20.5 27.0

## 9 18.5 15.9

## 10 39.9 44.6

## 11 20.6 22.7

## 12 13.4 11.3

## Co-operation by Technological Innovative Enterprises (% ## 1 50.8

## 2 32.9

## 3 20.2

## 4 27.4

## 5 25.8

## 6 40.1

## 7 38.7

## 8 23.3

## 9 17.3

## 10 41.9

## 11 21.5

## 12 12.5

I hope you would like this short article. …