

# Dooraedata

coop711

2025-10-08

## Data

- Reading data with `xlsx`
  - Note the arguments `stringsAsFactors` and `encoding`
  - knitr 에서 default로 R 폴더를 작업 디렉토리로 설정함에 유의.

```
# library(xlsx)
library(knitr)
library(magrittr)
library(readxl)
dooraedata_rest <- read_excel("../data/Data0924.xlsx",
                              sheet = 1,
                              range = "B1:BB98")
dooraedata_name <- read_excel("../data/Data0924.xlsx",
                              sheet = 1,
                              range = "A1:A98")
dooraedata <- data.frame(dooraedata_name, dooraedata_rest)
# dooraedata_rest <- read.xlsx("../data/Data0924.xlsx",
#                               sheetIndex = 1,
#                               colIndex = 2:54, endRow = 98,
#                               stringsAsFactors = TRUE,
#                               encoding = "UTF-8")
# dooraedata_name <- read.xlsx("../data/Data0924.xlsx",
#                               sheetIndex = 1,
#                               colIndex = 1, endRow = 98,
#                               stringsAsFactors = FALSE,
#                               encoding = "UTF-8")
# dooraedata <- dooraedata_name %>%
#   data.frame(dooraedata_rest)
# dooraedata <- dooraedata_rest %>%
#   data.frame(dooraedata_name, .)
# dooraedata <- data.frame(dooraedata_name, dooraedata_rest)
```

## Data for Years in Business

- 업력(years in business) 계산

```
dooraedata %>%
  str
```

```

## 'data.frame':    97 obs. of  54 variables:
## $              : chr  " " " " " " " " " " ...
## $              : chr  " " " " " " " " " " ...
## $              : chr  " " " " " " " " " " ...
## $ PD.1          : chr  " " " " " " " " " " ...
## $ PD.2          : chr  " " " " " " " " " " ...
## $              : chr  " " " " " " " " " " ...
## $              : chr  " " " " " " " " " " ...
## $              : num  4 4 4 5 5 5 40 67 5 5 ...
## $              : chr  " " " " " " " " " " ...
## $              : POSIXct, format: "2014-05-01" "2015-03-01" ...
## $              : chr  "NA" "513-80-29465" "NA" "124-87-43816" ...
## $              : chr  "NA" "NA" "NA" " " " " ...
## $              : chr  "NA" "NA" "NA" " " " " ...
## $              : chr  " " " " " " " " " " ...
## $              : chr  " " " " " " " " " " ...
## $              : chr  " " " " " " " " " " ...
## $              : chr  " " " " " " " " " " ...
## $              : num  1 1 1 1 0 0 0 0 0 0 ...
## $              : chr  " " " " " " " " " " ...
## $ . .          : num  12000 12000 10000 5000 0 0 0 0 0 0 ...
## $              : chr  "SW+HW" "SW+HW" "HW" "SW" ...
## $              : chr  " " " " " " " " " " ...
## $              : chr  " " " " " " " " " " ...
## $              : chr  " " " " " " " " " " ...
## $ _2012         : num  0 0 0 0 0 ...
## $ _2013         : num  0 0 0 5000 0 0 0 0 0 0 ...
## $ _2014         : num  0 0 0 4000 0 ...
## $ _2012         : num  0 0 0 0 0 0 653000 0 0 0 ...
## $ _2013         : num  0 0 0 80000 0 0 0 0 0 0 ...
## $ _2014         : num  0 0 0 70000 0 ...
## $ _            : chr  "No" "No" "No" "Yes" ...
## $ _            : chr  "No" "No" "Yes" "Yes" ...
## $ _            : chr  "No" "Yes" "No" "Yes" ...
## $ _            : chr  "No" "No" "No" "Yes" ...
## $ _            : chr  "No" "No" "No" "Yes" ...
## $ _            : chr  "No" "No" "No" "No" ...
## $ _            : chr  "No" "No" "No" "No" ...
## $ _            : chr  "No" "No" "No" "Yes" ...
## $ _            : chr  "No" "No" "No" "No" ...
## $ _            : chr  "Yes" "Yes" "No" "No" ...
## $ _SNS          : chr  "Yes" "Yes" "No" "No" ...
## $ _2014         : num  0 0 0 0 0 ...
## $ _2015_3       : num  0 0 20800 7080 0 0 0 0 0 0 ...
## $ _2014         : num  0 0 0 0 0 ...
## $ _2015_3       : num  0 0 5200 760 0 0 0 0 0 0 ...
## $ _2014         : num  0 0 0 0 0 0 0 0 0 0 ...
## $ _2015         : num  0 0 0 0 0 0 0 0 0 0 ...
## $ _2014         : num  0 0 0 0 0 0 0 0 0 0 ...
## $ _2015         : num  0 0 0 3 0 0 0 0 0 0 ...
## $ _            : num  1 1 0 1 0 2 2 2 2 0 ...
## $ _            : num  0 0 0 0 0 1 1 0 0 0 ...
## $ _            : num  1 0 0 0 0 2 3 2 6 0 ...
## $ _            : num  0 0 0 0 0 3 3 1 0 0 ...

```

```
## $ : num 0 1 0 0 0 4 4 2 0 0 ...
```

```
doorae0924$ " " <- doorae0924$ " " %>%  
# `[(`, " ") %>%  
  difftime("2015-09-24", ., units = "days") %>%  
  `/(365) %>%  
  as.numeric %>%  
  round(digits = 1)  
doorae0924$ " " %>%  
  str
```

```
## num [1:97] 1.4 0.6 0.4 2.7 0.4 1.6 4.1 2.1 1.4 0.1 ...
```

```
doorae0924$ " " %>%  
  head
```

```
## [1] 1.4 0.6 0.4 2.7 0.4 1.6
```

```
doorae0924 %>%  
  str
```

```
## 'data.frame': 97 obs. of 55 variables:  
## $ : chr " " " _ " " " " ...  
## $ : chr " " " " " " " ...  
## $ : chr " " " " " " " ...  
## $ PD.1 : chr " " " " " " " ...  
## $ PD.2 : chr " " " " " " " ...  
## $ : chr " " " " " " " ...  
## $ : chr " " " " " " " ...  
## $ : num 4 4 4 5 5 5 40 67 5 5 ...  
## $ : chr " " " " " " " ...  
## $ : POSIXct, format: "2014-05-01" "2015-03-01" ...  
## $ : chr "NA" "513-80-29465" "NA" "124-87-43816" ...  
## $ : chr "NA" "NA" "NA" " " " ...  
## $ : chr "NA" "NA" "NA" " " " ...  
## $ : chr " " " " " " " ...  
## $ : chr " " " " " " " ...  
## $ : chr " " " " " " " ...  
## $ : chr " " " " " " " ...  
## $ : num 1 1 1 1 0 0 0 0 0 0 ...  
## $ : chr " " " " " " " ...  
## $ . . : num 12000 12000 10000 5000 0 0 0 0 0 0 ...  
## $ : chr "SW+HW" "SW+HW" "HW" "SW" ...  
## $ : chr " " " " " " " ...  
## $ : chr " " " " " " " ...  
## $ : chr " " " " " " " ...  
## $ _2012 : num 0 0 0 0 0 ...  
## $ _2013 : num 0 0 0 5000 0 0 0 0 0 0 ...  
## $ _2014 : num 0 0 0 4000 0 ...  
## $ _2012 : num 0 0 0 0 0 0 653000 0 0 0 ...  
## $ _2013 : num 0 0 0 80000 0 0 0 0 0 0 ...  
## $ _2014 : num 0 0 0 70000 0 ...
```

```
## $ _ : chr "No" "No" "No" "Yes" ...
## $ _ : chr "No" "No" "Yes" "Yes" ...
## $ _ : chr "No" "Yes" "No" "Yes" ...
## $ _ : chr "No" "No" "No" "Yes" ...
## $ _ : chr "No" "No" "No" "Yes" ...
## $ _ : chr "No" "No" "No" "No" ...
## $ _ : chr "No" "No" "No" "No" ...
## $ _ : chr "No" "No" "No" "Yes" ...
## $ _ : chr "No" "No" "No" "No" ...
## $ _ : chr "Yes" "Yes" "No" "No" ...
## $ _SNS : chr "Yes" "Yes" "No" "No" ...
## $ _2014 : num 0 0 0 0 0 ...
## $ _2015_3 : num 0 0 20800 7080 0 0 0 0 0 ...
## $ _2014 : num 0 0 0 0 0 ...
## $ _2015_3 : num 0 0 5200 760 0 0 0 0 0 ...
## $ _2014 : num 0 0 0 0 0 0 0 0 0 ...
## $ _2015 : num 0 0 0 0 0 0 0 0 0 ...
## $ _2014 : num 0 0 0 0 0 0 0 0 0 ...
## $ _2015 : num 0 0 0 3 0 0 0 0 0 ...
## $ _ : num 1 1 0 1 0 2 2 2 0 ...
## $ _ : num 0 0 0 0 0 1 1 0 0 ...
## $ _ : num 1 0 0 0 0 2 3 2 6 ...
## $ _ : num 0 0 0 0 0 3 3 1 0 ...
## $ : num 0 1 0 0 0 4 4 2 0 ...
## $ : num 1.4 0.6 0.4 2.7 0.4 1.6 4.1 2.1 1.4 0.1 ...
```

- 사업자등록여부와 업력을 함께 새로운 data frame 으로

```
doorae0924$ " " <- doorae0924$ " " %>%
  `==`("NA") %>%
  ifelse("No", "Yes") %>%
  factor(levels = c("Yes", "No"))
doorae0924_ys <- doorae0924$ " " %>%
  factor(levels = c(" ", " ")) %>%
  data.frame(status = ., years = doorae0924$ " ")
```

Save data

```
save.image("./doorae_new_data.RData")
```