language

lim

2022 5 12

library(readxl)

## Warning: 패키지 'readxl'는 R 버전 4.1.3에서 작성되었습니다

library(dplyr)

## Warning: 패키지 'dplyr'는 R 버전 4.1.3에서 작성되었습니다

##   
## 다음의 패키지를 부착합니다: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(ggplot2)

## Warning: 패키지 'ggplot2'는 R 버전 4.1.3에서 작성되었습니다

lan <- read\_excel("C:/Bdata/mini/table/1-3-1-8. 개발언어별 SW 전문인력.xlsx")

소프트웨어 정책 연구소의 통계자료를 통하여 개발 언어별 전문 인력 비율을 시각화 하였다.

# 컬럼명 변경  
names(lan) <- c("language", "2018", "2019", "2020")   
  
str(lan$`2018`)

## chr [1:10] "103.85077731936165" "-" "122.40955573868297" ...

# 숫자형으로 변경  
lan$`2018` <- ifelse(lan$`2018` == "-", 0, lan$`2018`)  
lan$`2018` <-as.numeric(lan$`2018`)  
  
lan$`2020`<-gsub("\\\*","", lan$`2020`)  
lan$`2020`<-as.numeric(lan$`2020`)  
  
# 소수점 한자리 값으로 변경  
lan$`2018`<-round(lan$`2018`, 1)  
lan$`2019`<-round(lan$`2019`, 1)  
lan$`2020`<-round(lan$`2020`, 1)  
  
# 행 열 교체  
library(reshape2)

## Warning: 패키지 'reshape2'는 R 버전 4.1.3에서 작성되었습니다

lan\_melt<-melt(lan)

## Using language as id variables

names(lan\_melt) <- c("language", "year", "value")   
View(lan\_melt)  
  
  
# 시각화  
ggplot(data = lan\_melt, aes(x=lan\_melt$language, y=lan\_melt$value, group = lan\_melt$year,  
 fill=year))+  
 geom\_line(aes(color=year))+  
 geom\_point(aes(color=year), size=2)+  
 xlab("개발언어")+  
 ylab("연도")+  
 ggtitle("개발언어별 SW 전문인력")

## Warning: Use of `lan\_melt$language` is discouraged. Use `language` instead.

## Warning: Use of `lan\_melt$value` is discouraged. Use `value` instead.

## Warning: Use of `lan\_melt$year` is discouraged. Use `year` instead.

## Warning: Use of `lan\_melt$language` is discouraged. Use `language` instead.

## Warning: Use of `lan\_melt$value` is discouraged. Use `value` instead.

## Warning: Use of `lan\_melt$year` is discouraged. Use `year` instead.

