

한강공원별 고객 특성 파악

송승연

7/15/2020

한강공원별 이용객 데이터 분석

나들이 가기 딱 좋은 따뜻한 봄. 4, 5월에 한강 공원 이용객들이 가장 많은 시기입니다. 하지만 사실 그 시기에는 미세먼지로 인해 우리나라 대기질이 가장 안좋은 시기입니다. 이 데이터를 보고 정말 사람들은 야외활동에 있어서 미세먼지를 신경쓰지 않는 것일까하는 의문이 들었고, 이외에도 한강공원 이용객의 특성이나 이용객 증감에 영향을 미치는 또 다른 요소들에 대해 데이터 분석을 진행하였습니다.

활용 데이터 출처

-SKT 빅데이터 허브

-서울 열린데이터 광장

-K-WEATHER

데이터 수집 범위 설정

서울 “동” 데이터로 한정 : 한강공원이라는 지역을 기점으로 분석이 필요했기 때문에 서울시 데이터를 기준으로 분석범위는 시군구가 아닌 한강공원이 소속되어 있는 동을 기준으로 한정하였습니다.

5월 데이터 분석 : 2019년 한강공원 방문객이 가장 많은 달은 4월, 5월, 그 중에 4월은 벚꽃 시즌에 여의도 공원에 너무 많은 사람이 몰려 방문객수에 한쪽으로 쏠리는 경우 발생하기 때문에 평균적으로 방문객이 가장 많은 5월을 선정하였습니다.

분석 방법

운동시설 이용객, 자전거 이용객 이렇게 세 가지 목적별로 살펴보고, 그 중에 일반이용객이 배달 주문 통화건수와 상관관계가 있을 것이라고 생각하여 진행했습니다. 그리고 한강공원 이용객들의 방문과 배달음식 주문이 미세먼지, 체감온도, 습도와 같은 다양한 날씨 요소에 어떤 변수값이 가장 영향을 많이 미치는지 알아보았습니다.

```

setwd(dir = "/Users/useonsong/Nano Degree/1차 프로젝트/외부데이터")
library(tidyverse)
library(readxl)
people_in_광나루 <- readxl::read_xlsx("광나루 이용객.xlsx")
people_in_광나루 <- people_in_광나루[, c(2,4)]
people_in_광나루 <- people_in_광나루[-c(1:4, 9:16, 18:22), ]
people_in_광나루$한강공원 <- rep('광나루', 5)
colnames(people_in_광나루)[1] <- c('이용목적')
colnames(people_in_광나루)[2] <- c('이용객')
people_in_광나루 <- people_in_광나루 %>% spread('이용목적', '이용객')
people_in_광나루[, 2:6] <- map_df(people_in_광나루[, 2:6], as.integer)
people_in_광나루 <- data.frame(people_in_광나루, rowSums(people_in_광나루[, 3:5]))
people_in_광나루 <- people_in_광나루[, -c(3:5)]
colnames(people_in_광나루)[4] <- c('일반')
people_in_광나루 <- people_in_광나루 %>% gather('이용목적', '이용객', ~한강공원)

people_in_난지 <- readxl::read_xlsx("난지 이용객.xlsx")
people_in_난지 <- people_in_난지[, c(2:3)]
people_in_난지 <- people_in_난지[-c(1:4, 9:16, 18:48), ]
people_in_난지$한강공원 <- rep('난지', 5)
colnames(people_in_난지)[1] <- c('이용목적')
colnames(people_in_난지)[2] <- c('이용객')
people_in_난지 <- people_in_난지 %>% spread('이용목적', '이용객')
people_in_난지[, 2:6] <- map_df(people_in_난지[, 2:6], as.integer)
people_in_난지 <- data.frame(people_in_난지, rowSums(people_in_난지[, 3:5]))
people_in_난지 <- people_in_난지[, -c(3:5)]
colnames(people_in_난지)[4] <- c('일반')
people_in_난지 <- people_in_난지 %>% gather('이용목적', '이용객', ~한강공원)

people_in_뚝섬 <- readxl::read_xlsx("뚝섬 이용객.xlsx")
people_in_뚝섬 <- people_in_뚝섬[, c(2:3)]
people_in_뚝섬 <- people_in_뚝섬[-c(1:3, 8:15, 17:28), ]
people_in_뚝섬$한강공원 <- rep('뚝섬', 5)
colnames(people_in_뚝섬)[1] <- c('이용목적')
colnames(people_in_뚝섬)[2] <- c('이용객')
people_in_뚝섬 <- people_in_뚝섬 %>% spread('이용목적', '이용객')
people_in_뚝섬[, 2:6] <- map_df(people_in_뚝섬[, 2:6], as.integer)
people_in_뚝섬 <- data.frame(people_in_뚝섬, rowSums(people_in_뚝섬[, 3:5]))
people_in_뚝섬 <- people_in_뚝섬[, -c(3:5)]
colnames(people_in_뚝섬)[4] <- c('일반')
people_in_뚝섬 <- people_in_뚝섬 %>% gather('이용목적', '이용객', ~한강공원)

people_in_망원 <- readxl::read_xlsx("망원 이용객1.xlsx")
people_in_망원 <- people_in_망원[, c(2,4)]
people_in_망원 <- people_in_망원[-c(1:4, 9:16, 18:48), ]
people_in_망원$한강공원 <- rep('망원', 5)
colnames(people_in_망원)[1] <- c('이용목적')
colnames(people_in_망원)[2] <- c('이용객')
people_in_망원 <- people_in_망원 %>% spread('이용목적', '이용객')
people_in_망원[, 2:6] <- map_df(people_in_망원[, 2:6], as.integer)
people_in_망원 <- data.frame(people_in_망원, rowSums(people_in_망원[, 3:5]))
people_in_망원 <- people_in_망원[, -c(3:5)]
colnames(people_in_망원)[4] <- c('일반')
people_in_망원 <- people_in_망원 %>% gather('이용목적', '이용객', ~한강공원)

people_in_반포 <- readxl::read_xlsx("반포 이용객.xlsx")

```

```

people_in_반포 <- people_in_반포[, c(2,4)]
people_in_반포 <- people_in_반포[-c(1:4, 9:16, 18:48), ]
people_in_반포$한강공원 <- rep('반포', 5)
colnames(people_in_반포)[1] <- c('이용목적')
colnames(people_in_반포)[2] <- c('이용객')
people_in_반포 <- people_in_반포 %>% spread('이용목적', '이용객')
people_in_반포[, 2:6] <- map_df(people_in_반포[, 2:6], as.integer)
people_in_반포 <- data.frame(people_in_반포, rowSums(people_in_반포[, 3:5]))
people_in_반포 <- people_in_반포[, -c(3:5)]
colnames(people_in_반포)[4] <- c('일반')
people_in_반포 <- people_in_반포 %>% gather('이용목적', '이용객', -한강공원)

people_in_여의도 <- readxl::read_xlsx("여의도 이용객.xlsx")
people_in_여의도 <- people_in_여의도[, c(2,4)]
people_in_여의도 <- people_in_여의도[-c(1:4, 9:16, 18:36), ]
people_in_여의도$한강공원 <- rep('여의도', 5)
colnames(people_in_여의도)[1] <- c('이용목적')
colnames(people_in_여의도)[2] <- c('이용객')
people_in_여의도 <- people_in_여의도 %>% spread('이용목적', '이용객')
people_in_여의도[, 2:6] <- map_df(people_in_여의도[, 2:6], as.integer)
people_in_여의도 <- data.frame(people_in_여의도, rowSums(people_in_여의도[, 3:5]))
people_in_여의도 <- people_in_여의도[, -c(3:5)]
colnames(people_in_여의도)[4] <- c('일반')
people_in_여의도 <- people_in_여의도 %>% gather('이용목적', '이용객', -한강공원)

people_in_이촌 <- readxl::read_xlsx("이촌 이용객.xlsx")
people_in_이촌 <- people_in_이촌[, c(2,4)]
people_in_이촌 <- people_in_이촌[-c(1:4, 9:16, 18:22), ]
people_in_이촌$한강공원 <- rep('이촌', 5)
colnames(people_in_이촌)[1] <- c('이용목적')
colnames(people_in_이촌)[2] <- c('이용객')
people_in_이촌 <- people_in_이촌 %>% spread('이용목적', '이용객')
people_in_이촌[, 2:6] <- map_df(people_in_이촌[, 2:6], as.integer)
people_in_이촌 <- data.frame(people_in_이촌, rowSums(people_in_이촌[, 3:5]))
people_in_이촌 <- people_in_이촌[, -c(3:5)]
colnames(people_in_이촌)[4] <- c('일반')
people_in_이촌 <- people_in_이촌 %>% gather('이용목적', '이용객', -한강공원)
people_in_이촌[2, 2] <- c('자전거')

people_in_잠실 <- readxl::read_xlsx("잠실 이용객.xlsx")
people_in_잠실 <- people_in_잠실[, c(2,4)]
people_in_잠실 <- people_in_잠실[-c(1:4, 9:16, 18:48), ]
people_in_잠실$한강공원 <- rep('잠실', 5)
colnames(people_in_잠실)[1] <- c('이용목적')
colnames(people_in_잠실)[2] <- c('이용객')
people_in_잠실 <- people_in_잠실 %>% spread('이용목적', '이용객')
people_in_잠실[, 2:6] <- map_df(people_in_잠실[, 2:6], as.integer)
people_in_잠실 <- data.frame(people_in_잠실, rowSums(people_in_잠실[, 3:5]))
people_in_잠실 <- people_in_잠실[, -c(3:5)]
colnames(people_in_잠실)[4] <- c('일반')
people_in_잠실 <- people_in_잠실 %>% gather('이용목적', '이용객', -한강공원)

people_in_잠원 <- readxl::read_xlsx("잠원 이용객.xlsx")
people_in_잠원 <- people_in_잠원[, c(2,4)]
people_in_잠원 <- people_in_잠원[-c(1:4, 9:16, 18:48), ]
people_in_잠원$한강공원 <- rep('잠원', 5)
colnames(people_in_잠원)[1] <- c('이용목적')
colnames(people_in_잠원)[2] <- c('이용객')

```

```

people_in_잠원 <- people_in_잠원 %>% spread('이용목적', '이용객')
people_in_잠원[, 2:6] <- map_df(people_in_잠원[, 2:6], as.integer)
people_in_잠원 <- data.frame(people_in_잠원, rowSums(people_in_잠원[, 3:5]))
people_in_잠원 <- people_in_잠원[, -c(3:5)]
colnames(people_in_잠원)[4] <- c('일반')
people_in_잠원 <- people_in_잠원 %>% gather('이용목적', '이용객', -한강공원)

people_in_강서 <- readxl::read_xlsx("강서 이용객.xlsx")
people_in_강서 <- people_in_강서[, c(2,4)]
people_in_강서 <- people_in_강서[-c(1:4, 9:16, 18:48), ]
people_in_강서$한강공원 <- rep('강서', 5)
colnames(people_in_강서)[1] <- c('이용목적')
colnames(people_in_강서)[2] <- c('이용객')
people_in_강서 <- people_in_강서 %>% spread('이용목적', '이용객')
people_in_강서[, 2:6] <- map_df(people_in_강서[, 2:6], as.integer)
people_in_강서 <- data.frame(people_in_강서, rowSums(people_in_강서[, 3:5]))
people_in_강서 <- people_in_강서[, -c(3:5)]
colnames(people_in_강서)[4] <- c('일반')
people_in_강서 <- people_in_강서 %>% gather('이용목적', '이용객', -한강공원)

공원별_이용객 <- rbind(people_in_강서, people_in_광나루, people_in_난지, people_in_뚝섬, people_in_망원, people_in_반포, people_in_여의도, people_in_이촌, people_in_잠실, people_in_잠원)

```

우선 서울에 있는 10개의 한강공원 중 이용객이 가장 많은 공원을 살펴보았는데, 예상대로 뚝섬과 여의도가 가장 많은 것을 볼 수 있습니다.

```

setwd(dir = "/Users/useonsong/Nano Degree/1차 프로젝트/외부데이터")

library(tidyverse)
library(readxl)
library(stringr)

people_in_광나루 <- readxl::read_xlsx("광나루 이용객.xlsx")
people_in_광나루 <- people_in_광나루[, c(2,4)]
people_in_광나루 <- people_in_광나루[-c(1:4, 9:16, 18:22), ]
people_in_광나루$한강공원 <- rep('광나루', 5)
colnames(people_in_광나루)[1] <- c('이용목적')
colnames(people_in_광나루)[2] <- c('이용객')
people_in_광나루 <- people_in_광나루 %>% spread('이용목적', '이용객')
people_in_광나루[, 2:6] <- map_df(people_in_광나루[, 2:6], as.integer)
people_in_광나루 <- data.frame(people_in_광나루, rowSums(people_in_광나루[, 3:5]))
people_in_광나루 <- people_in_광나루[, -c(3:5)]
colnames(people_in_광나루)[4] <- c('일반')
people_in_광나루 <- people_in_광나루 %>% gather('이용목적', '이용객', -한강공원)

people_in_난지 <- readxl::read_xlsx("난지 이용객.xlsx")
people_in_난지 <- people_in_난지[, c(2:3)]
people_in_난지 <- people_in_난지[-c(1:4, 9:16, 18:48), ]
people_in_난지$한강공원 <- rep('난지', 5)
colnames(people_in_난지)[1] <- c('이용목적')
colnames(people_in_난지)[2] <- c('이용객')
people_in_난지 <- people_in_난지 %>% spread('이용목적', '이용객')
people_in_난지[, 2:6] <- map_df(people_in_난지[, 2:6], as.integer)
people_in_난지 <- data.frame(people_in_난지, rowSums(people_in_난지[, 3:5]))
people_in_난지 <- people_in_난지[, -c(3:5)]
colnames(people_in_난지)[4] <- c('일반')
people_in_난지 <- people_in_난지 %>% gather('이용목적', '이용객', -한강공원)

people_in_뚝섬 <- readxl::read_xlsx("뚝섬 이용객.xlsx")
people_in_뚝섬 <- people_in_뚝섬[, c(2:3)]
people_in_뚝섬 <- people_in_뚝섬[-c(1:3, 8:15, 17:28), ]
people_in_뚝섬$한강공원 <- rep('뚝섬', 5)
colnames(people_in_뚝섬)[1] <- c('이용목적')
colnames(people_in_뚝섬)[2] <- c('이용객')
people_in_뚝섬 <- people_in_뚝섬 %>% spread('이용목적', '이용객')
people_in_뚝섬[, 2:6] <- map_df(people_in_뚝섬[, 2:6], as.integer)
people_in_뚝섬 <- data.frame(people_in_뚝섬, rowSums(people_in_뚝섬[, 3:5]))
people_in_뚝섬 <- people_in_뚝섬[, -c(3:5)]
colnames(people_in_뚝섬)[4] <- c('일반')
people_in_뚝섬 <- people_in_뚝섬 %>% gather('이용목적', '이용객', -한강공원)

people_in_망원 <- readxl::read_xlsx("망원 이용객1.xlsx")
people_in_망원 <- people_in_망원[, c(2,4)]
people_in_망원 <- people_in_망원[-c(1:4, 9:16, 18:48), ]
people_in_망원$한강공원 <- rep('망원', 5)
colnames(people_in_망원)[1] <- c('이용목적')
colnames(people_in_망원)[2] <- c('이용객')
people_in_망원 <- people_in_망원 %>% spread('이용목적', '이용객')
people_in_망원[, 2:6] <- map_df(people_in_망원[, 2:6], as.integer)
people_in_망원 <- data.frame(people_in_망원, rowSums(people_in_망원[, 3:5]))
people_in_망원 <- people_in_망원[, -c(3:5)]
colnames(people_in_망원)[4] <- c('일반')

```

```

people_in_망원 <- people_in_망원 %>% gather('이용목적', '이용객', -한강공원)

people_in_반포 <- readxl::read_xlsx("반포 이용객.xlsx")
people_in_반포 <- people_in_반포[, c(2,4)]
people_in_반포 <- people_in_반포[-c(1:4, 9:16, 18:48), ]
people_in_반포$한강공원 <- rep('반포', 5)
colnames(people_in_반포)[1] <- c('이용목적')
colnames(people_in_반포)[2] <- c('이용객')
people_in_반포 <- people_in_반포 %>% spread('이용목적', '이용객')
people_in_반포[, 2:6] <- map_df(people_in_반포[, 2:6], as.integer)
people_in_반포 <- data.frame(people_in_반포, rowSums(people_in_반포[, 3:5]))
people_in_반포 <- people_in_반포[, -c(3:5)]
colnames(people_in_반포)[4] <- c('일반')
people_in_반포 <- people_in_반포 %>% gather('이용목적', '이용객', -한강공원)

people_in_여의도 <- readxl::read_xlsx("여의도 이용객.xlsx")
people_in_여의도 <- people_in_여의도[, c(2,4)]
people_in_여의도 <- people_in_여의도[-c(1:4, 9:16, 18:36), ]
people_in_여의도$한강공원 <- rep('여의도', 5)
colnames(people_in_여의도)[1] <- c('이용목적')
colnames(people_in_여의도)[2] <- c('이용객')
people_in_여의도 <- people_in_여의도 %>% spread('이용목적', '이용객')
people_in_여의도[, 2:6] <- map_df(people_in_여의도[, 2:6], as.integer)
people_in_여의도 <- data.frame(people_in_여의도, rowSums(people_in_여의도[, 3:5]))
people_in_여의도 <- people_in_여의도[, -c(3:5)]
colnames(people_in_여의도)[4] <- c('일반')
people_in_여의도 <- people_in_여의도 %>% gather('이용목적', '이용객', -한강공원)

people_in_이촌 <- readxl::read_xlsx("이촌 이용객.xlsx")
people_in_이촌 <- people_in_이촌[, c(2,4)]
people_in_이촌 <- people_in_이촌[-c(1:4, 9:16, 18:22), ]
people_in_이촌$한강공원 <- rep('이촌', 5)
colnames(people_in_이촌)[1] <- c('이용목적')
colnames(people_in_이촌)[2] <- c('이용객')
people_in_이촌 <- people_in_이촌 %>% spread('이용목적', '이용객')
people_in_이촌[, 2:6] <- map_df(people_in_이촌[, 2:6], as.integer)
people_in_이촌 <- data.frame(people_in_이촌, rowSums(people_in_이촌[, 3:5]))
people_in_이촌 <- people_in_이촌[, -c(3:5)]
colnames(people_in_이촌)[4] <- c('일반')
people_in_이촌 <- people_in_이촌 %>% gather('이용목적', '이용객', -한강공원)
people_in_이촌[2, 2] <- c('자전거')

people_in_잠실 <- readxl::read_xlsx("잠실 이용객.xlsx")
people_in_잠실 <- people_in_잠실[, c(2,4)]
people_in_잠실 <- people_in_잠실[-c(1:4, 9:16, 18:48), ]
people_in_잠실$한강공원 <- rep('잠실', 5)
colnames(people_in_잠실)[1] <- c('이용목적')
colnames(people_in_잠실)[2] <- c('이용객')
people_in_잠실 <- people_in_잠실 %>% spread('이용목적', '이용객')
people_in_잠실[, 2:6] <- map_df(people_in_잠실[, 2:6], as.integer)
people_in_잠실 <- data.frame(people_in_잠실, rowSums(people_in_잠실[, 3:5]))
people_in_잠실 <- people_in_잠실[, -c(3:5)]
colnames(people_in_잠실)[4] <- c('일반')
people_in_잠실 <- people_in_잠실 %>% gather('이용목적', '이용객', -한강공원)

people_in_잠원 <- readxl::read_xlsx("잠원 이용객.xlsx")
people_in_잠원 <- people_in_잠원[, c(2,4)]
people_in_잠원 <- people_in_잠원[-c(1:4, 9:16, 18:48), ]

```

```

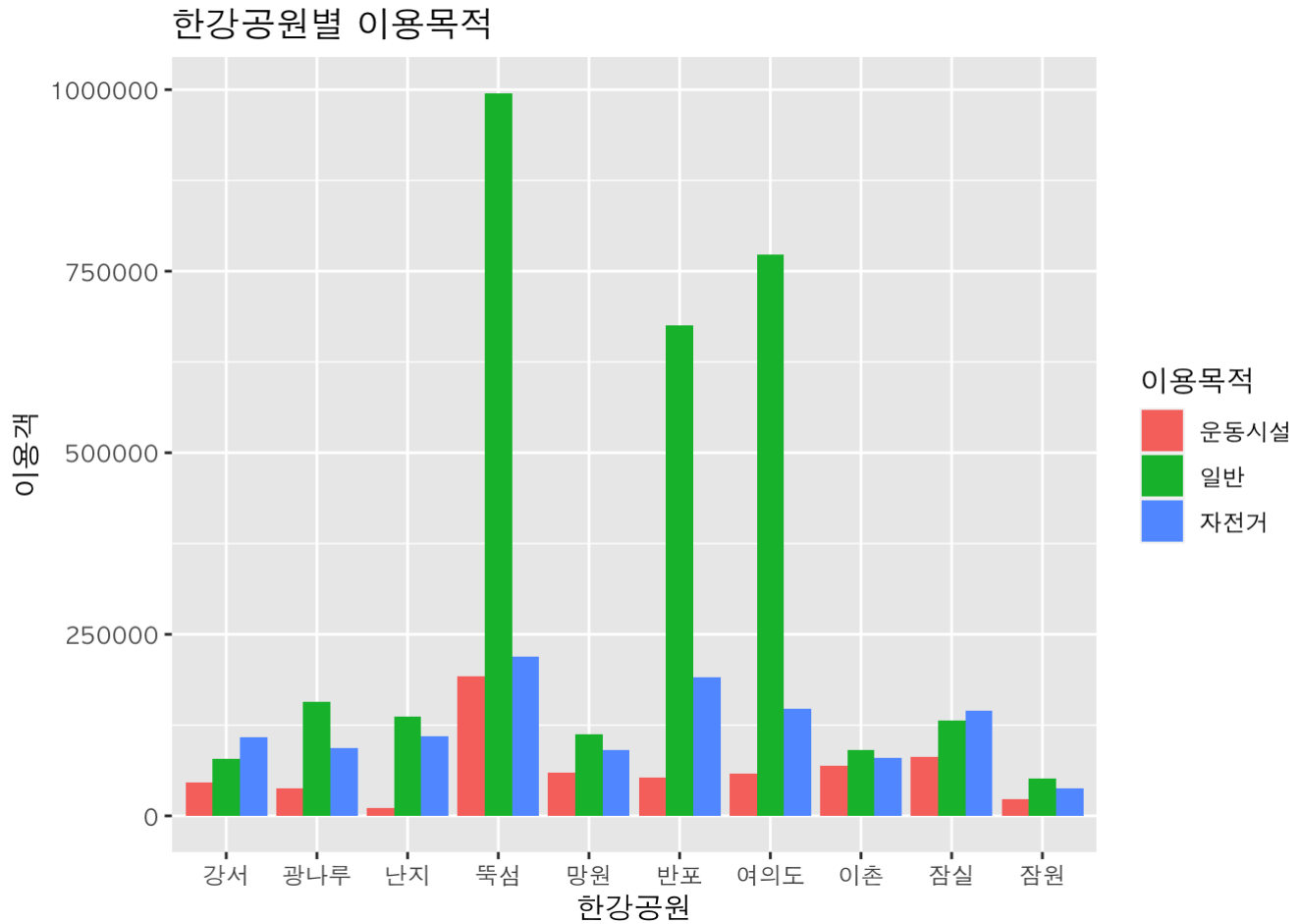
people_in_잠원$한강공원 <- rep('잠원', 5)
colnames(people_in_잠원)[1] <- c('이용목적')
colnames(people_in_잠원)[2] <- c('이용객')
people_in_잠원 <- people_in_잠원 %>% spread('이용목적', '이용객')
people_in_잠원[, 2:6] <- map_df(people_in_잠원[, 2:6], as.integer)
people_in_잠원 <- data.frame(people_in_잠원, rowSums(people_in_잠원[, 3:5]))
people_in_잠원 <- people_in_잠원[, -c(3:5)]
colnames(people_in_잠원)[4] <- c('일반')
people_in_잠원 <- people_in_잠원 %>% gather('이용목적', '이용객', -한강공원)

people_in_강서 <- readxl::read_xlsx("강서 이용객.xlsx")
people_in_강서 <- people_in_강서[, c(2,4)]
people_in_강서 <- people_in_강서[-c(1:4, 9:16, 18:48), ]
people_in_강서$한강공원 <- rep('강서', 5)
colnames(people_in_강서)[1] <- c('이용목적')
colnames(people_in_강서)[2] <- c('이용객')
people_in_강서 <- people_in_강서 %>% spread('이용목적', '이용객')
people_in_강서[, 2:6] <- map_df(people_in_강서[, 2:6], as.integer)
people_in_강서 <- data.frame(people_in_강서, rowSums(people_in_강서[, 3:5]))
people_in_강서 <- people_in_강서[, -c(3:5)]
colnames(people_in_강서)[4] <- c('일반')
people_in_강서 <- people_in_강서 %>% gather('이용목적', '이용객', -한강공원)

공원별_이용객 <- rbind(people_in_강서, people_in_광나루, people_in_난지, people_in_뚝섬, people_in_망원, people_in_반포, people_in_여의도, people_in_이촌, people_in_잠실, people_in_잠원)

library(ggplot2)
theme_set(theme_grey(base_family = 'AppleGothic'))
par(family = 'AppleGothic')
ggplot(공원별_이용객, aes(x = 한강공원, y = 이용객, fill = 이용목적)) +
  geom_bar(stat = 'identity', position = 'dodge') +
  ggtitle('한강공원별 이용목적') +
  ylab('이용객')

```



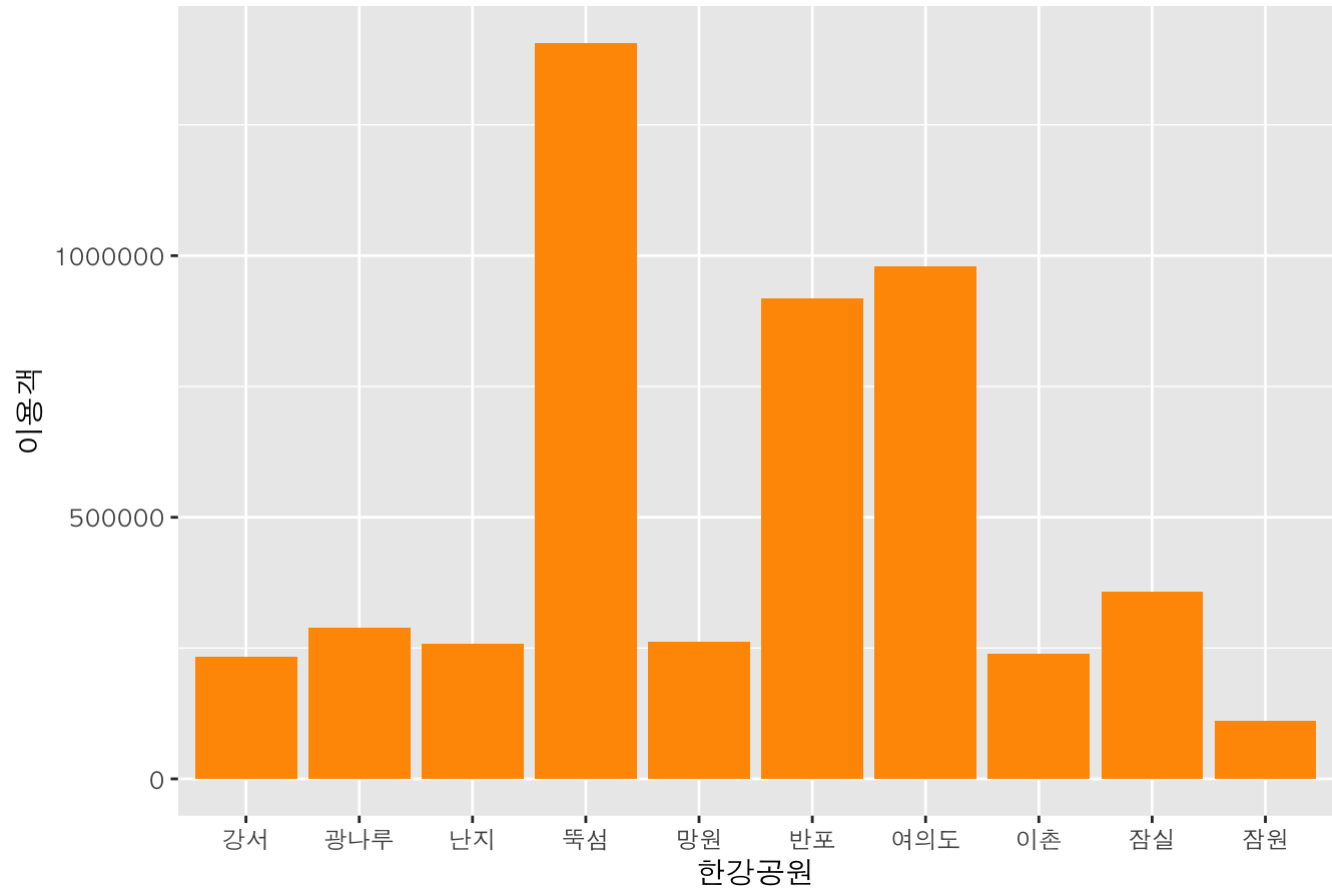
```

공원별_이용객 <- 공원별_이용객 %>% spread(한강공원, 이용객)
공원별_이용객 <- 공원별_이용객[, -1]
공원별_이용객 <- t(as.matrix(공원별_이용객)) %>% as.data.frame()
공원별_이용객$합계 <- data.frame(rowSums(공원별_이용객[, 1:3]))
공원별_이용객 <- 공원별_이용객[, -c(1:3)]
colnames(공원별_이용객)[1] <- c('합계')
공원별_이용객$한강공원 <- rownames(공원별_이용객)

library(ggplot2)
theme_set(theme_grey(base_family = 'AppleGothic'))
par(family = 'AppleGothic')
ggplot(공원별_이용객, aes(x = 한강공원, y = 합계)) +
  geom_bar(stat = 'identity', position = 'dodge', fill = "#FF9900") +
  ggtitle('한강공원별 이용객') +
  ylab('이용객') +
  scale_y_continuous(breaks = seq(0, 1500000, 500000),
    labels = as.integer(seq(0, 1500000, 500000)))

```


한강공원별 이용객



이렇게 많은 이용객 중에는 운동시설을 이용하러 온 이용객도 있고, 자전거를 타러온 이용객들도 있었는데, 목적별로 살펴봐도 뚝섬, 여의도가 가장 이용객이 많았고, 이중에서도 일반 이용객들이 월등하게 높은 비율을 차지하는 것을 알 수 있었습니다.

```

setwd(dir = "/Users/useonsong/Nano Degree/1차 프로젝트/외부데이터")

people_in_뚝섬 <- read_xlsx('뚝섬 이용객.xlsx')
people_in_뚝섬 <- people_in_뚝섬[-c(3, 8:15, 17:28), -3 ]

colnames(people_in_뚝섬) <- people_in_뚝섬[1, ]
people_in_뚝섬 <- people_in_뚝섬[-1, -1]
colnames(people_in_뚝섬)[1] <- c('이용목적')

#일자별 뚝섬 데이터 long으로 변환
people_in_뚝섬 <- people_in_뚝섬 %>%
  gather('일자', '이용객', -이용목적, na.rm = T)
#결측값 삭제
people_in_뚝섬 <- people_in_뚝섬 %>%
  filter(is.na(people_in_뚝섬$이용목적) == FALSE)

library(tidyverse)
people_in_뚝섬[, 2:3] <- map_df(people_in_뚝섬[, 2:3], as.integer)

#일반이용자 3구분을 더해서 하나로 묶기
library(tidyr)
people_in_뚝섬 <- people_in_뚝섬 %>% spread(key = 이용목적, value = 이용객, fill = 0)

people_in_뚝섬 <- data.frame(people_in_뚝섬, rowSums(people_in_뚝섬[, 3:5]))
people_in_뚝섬 <- people_in_뚝섬[, -c(3:5)]
colnames(people_in_뚝섬)[4] <- c('일반')

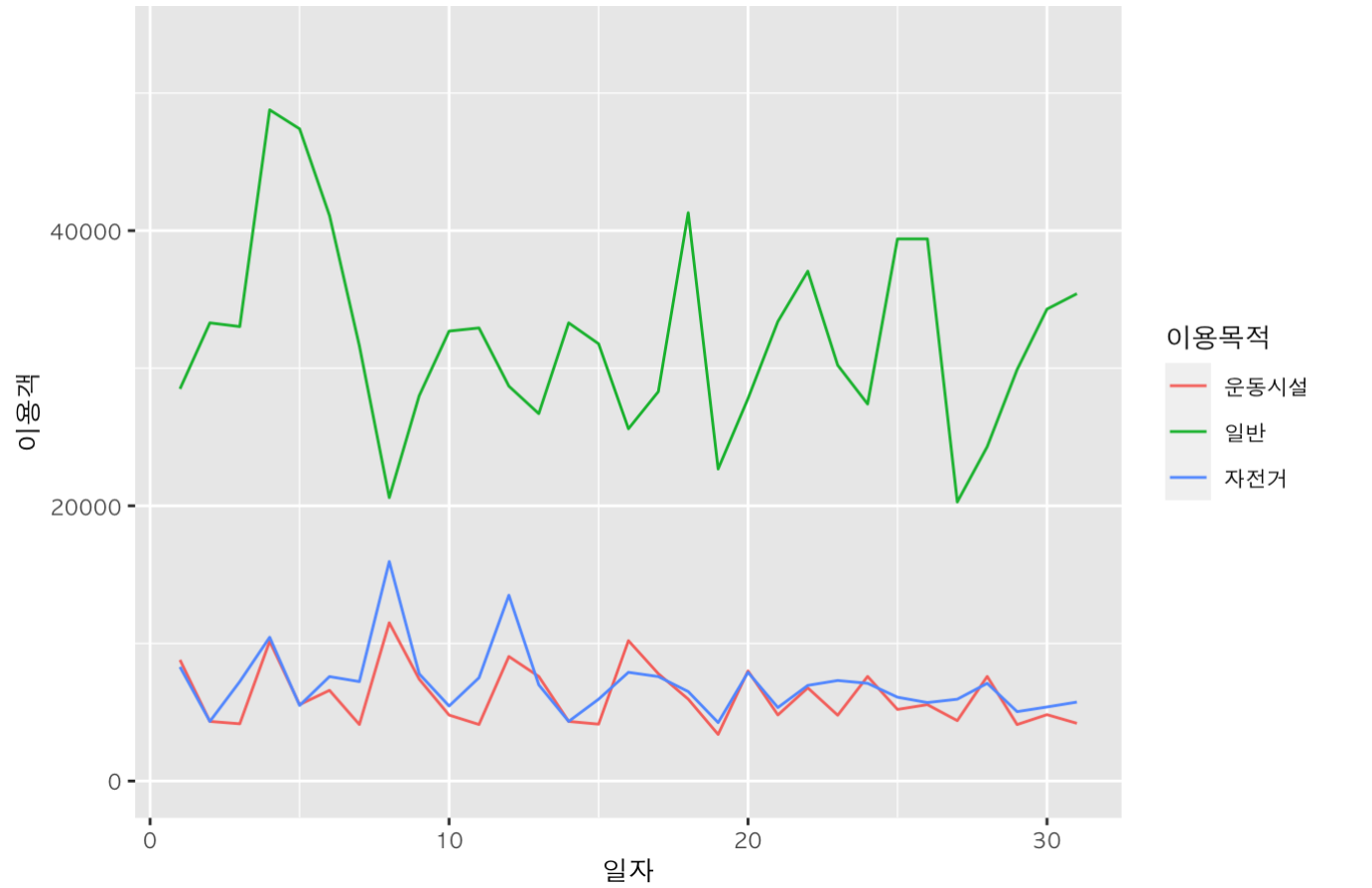
people_in_뚝섬 <- people_in_뚝섬 %>%
  gather(이용목적, 이용객, -일자)

#이방법은 왜 안돼는지?
#people_in_뚝섬[, 1] <- str_replace_all(people_in_뚝섬[, 1], '일반이용자\\(아침\\)', '일반')
#*괄호는 문자로 인식하지 않기 때문에 \\를 꼭 붙여줘야한다.
#일반이용자(아침)을 일반으로 바꿔줌
#people_in_뚝섬$이용목적[seq(1, 93, 3)] <- '일반'

#뚝섬 일자별 이용객 선그래프
library(ggplot2)
ggplot(people_in_뚝섬, aes(x = 일자, y = 이용객, fill = 이용목적, colour = 이용목적)) +
  geom_line() +
  ylim(0, max(people_in_뚝섬$이용객*1.1)) +
  ggtitle('뚝섬 한강공원 일자별 이용객') +
  theme(text = element_text(size = 10, family = "AppleGothic"))

```

독섬 한강공원 일자별 이용객



#뚝섬 요일별 그래프 그리기

```
people_in_뚝섬 <- read_xlsx('뚝섬 이용객.xlsx')
people_in_뚝섬 <- people_in_뚝섬[-c(3, 8:15, 17:28), -3 ]
```

```
colnames(people_in_뚝섬) <- str_c(people_in_뚝섬[1, ], people_in_뚝섬[2, ])
people_in_뚝섬 <- people_in_뚝섬[-c(1, 2), -1]
colnames(people_in_뚝섬)[1] <- c('이용목적')
```

#요일별 뚝섬 데이터 long으로 변환

```
people_in_뚝섬 <- people_in_뚝섬 %>%
  gather('요일', '이용객', -이용목적, na.rm = T)
```

```
#str(people_in_뚝섬) #각 컬럼별 class확인, 이용객 정수형으로 변환
people_in_뚝섬[, 3] <- as.integer(people_in_뚝섬$이용객)
```

#요일끼리 더해줘야함

#우선 요일이름을 맞춰주기 위해 앞에 날짜를 다 지워준다.

```
people_in_뚝섬$요일 <- str_remove_all(people_in_뚝섬$요일, '\\d')
```

#그리고 같은 요일끼리 이용목적에 따라 이용객 다 더해준다.

```
people_in_뚝섬 <- people_in_뚝섬 %>% group_by(이용목적, 요일) %>% summarise(이용객 = sum(이용객))
```

#일반이용자 3가지 구분 더해주기 위해

```
people_in_뚝섬 <- people_in_뚝섬 %>% spread(이용목적, 이용객, fill = 0)
people_in_뚝섬 <- data.frame(people_in_뚝섬, rowSums(people_in_뚝섬[, 3:5]))
people_in_뚝섬 <- people_in_뚝섬[, -c(3:5)]
colnames(people_in_뚝섬)[4] <- c('일반')
```

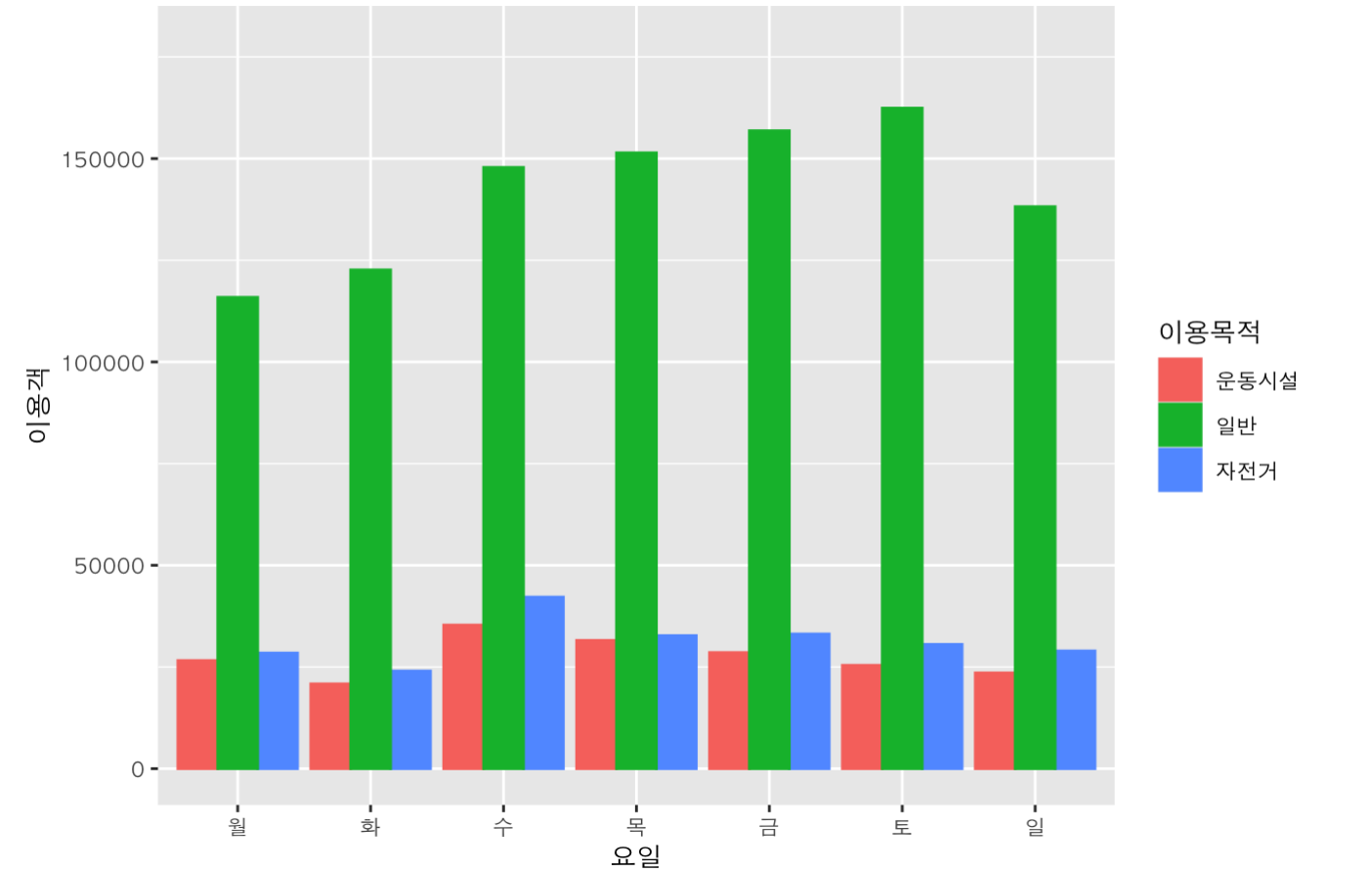
```
people_in_뚝섬 <- people_in_뚝섬 %>% gather(이용목적, 이용객, -요일)
```

#그래프 그리기

```
library(ggplot2)
```

```
ggplot(people_in_뚝섬, aes(x = 요일, y = 이용객, fill = 이용목적, colour = 이용목적)) +
  geom_col(stat = 'identity', position = 'dodge') +
  scale_x_discrete(limit = c('월', '화', '수', '목', '금', '토', '일')) +
  ylim(0, max(people_in_뚝섬$이용객*1.1)) +
  ggtitle('뚝섬 한강공원 요일별 이용객') +
  theme(text = element_text(size = 10, family = "AppleGothic"))
```

독섬 한강공원 요일별 이용객



가장 사람들이 많이 가는 독섬 한강공원의 이용객들을 목적에 따라 일자별로, 요일별로 분석해본 결과입니다. 독섬은 평일, 주말 관계없이 운동시설이나 자전거 이용객보다는 일반이용객들이 훨씬 많다는 걸 알 수 있었습니다. 하지만 다른 공원들에 비하면 운동시설이나 자전거를 이용하는 사람들도 많은 편입니다.

```

setwd(dir = "/Users/useonsong/Nano Degree/1차 프로젝트/외부데이터")

people_in_여의도 <- read_xlsx('여의도 이용객.xlsx')
people_in_여의도 <- people_in_여의도[-c(1, 4, 9:16, 18:36), -c(3, 4)]

colnames(people_in_여의도) <- people_in_여의도[1, ]
people_in_여의도 <- people_in_여의도[-1, -1]
colnames(people_in_여의도)[1] <- c('이용목적')

#일자별 여의도 데이터 long으로 변환
people_in_여의도 <- people_in_여의도 %>%
  gather('일자', '이용객', -이용목적, na.rm = T)
#결측값 삭제
people_in_여의도 <- people_in_여의도 %>%
  filter(is.na(people_in_여의도$이용목적) == FALSE)

library(tidyverse)
people_in_여의도[, 2:3] <- map_df(people_in_여의도[, 2:3], as.integer)

#일반이용자 3구분을 더해서 하나로 묶기
library(tidyr)
people_in_여의도 <- people_in_여의도 %>% spread(key = 이용목적, value = 이용객, fill = 0)

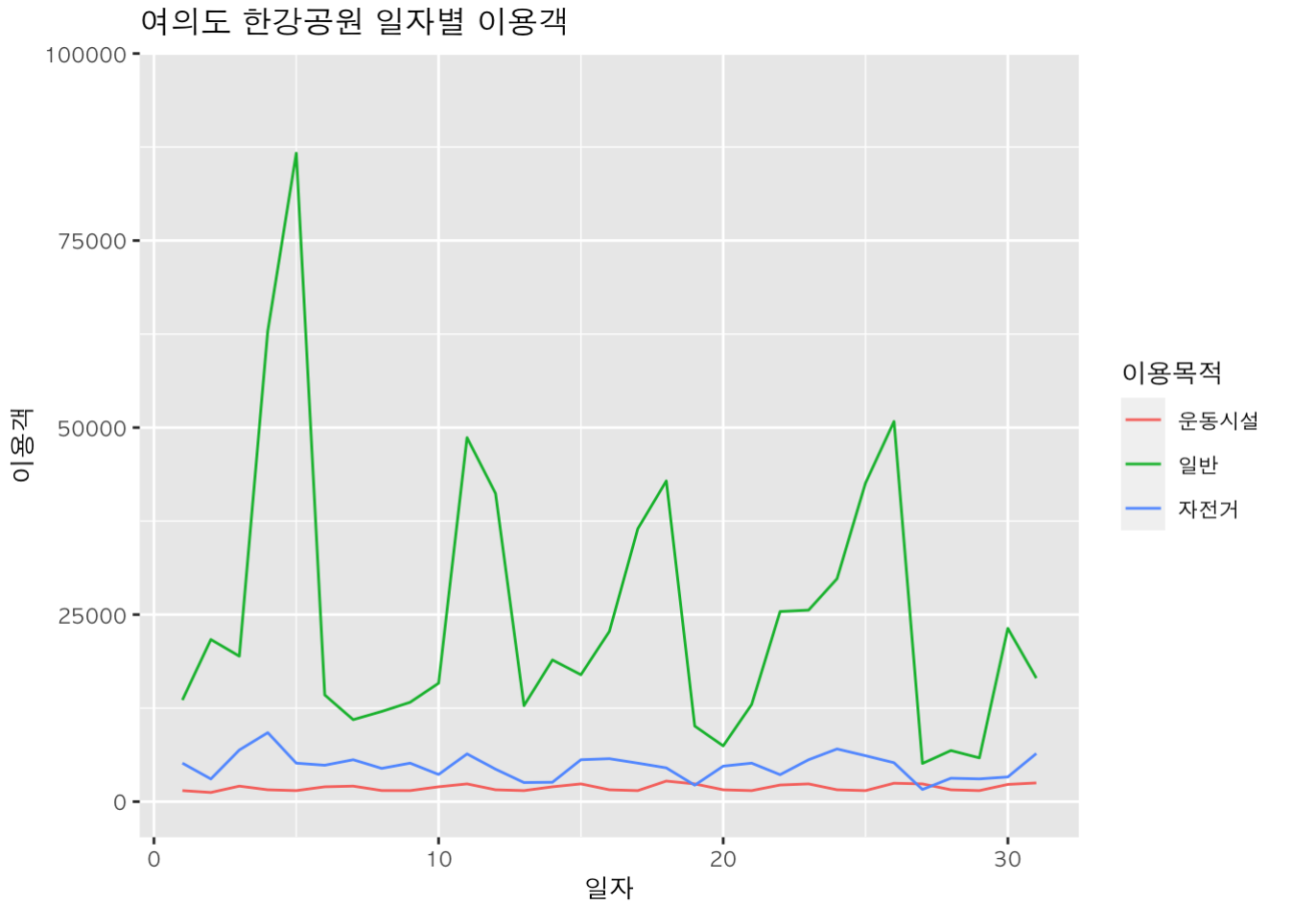
people_in_여의도 <- data.frame(people_in_여의도, rowSums(people_in_여의도[, 3:5]))
people_in_여의도 <- people_in_여의도[, -c(3:5)]
colnames(people_in_여의도)[4] <- c('일반')

people_in_여의도 <- people_in_여의도 %>%
  gather(이용목적, 이용객, -일자)

#이방법은 왜 안되는지?
#people_in_여의도[, 1] <- str_replace_all(people_in_여의도[, 1], '일반이용자\\(아침\\)', '일반')
#*괄호는 문자로 인식하지 않기 때문에 \\를 꼭 붙여줘야한다.
#일반이용자(아침)을 일반으로 바꿔줌
#people_in_여의도$이용목적[seq(1, 93, 3)] <- '일반'

#여의도 일자별 이용객 선그래프
library(ggplot2)
ggplot(people_in_여의도, aes(x = 일자, y = 이용객, fill = 이용목적, colour = 이용목적)) +
  geom_line() +
  ylim(0, max(people_in_여의도$이용객*1.1)) +
  ggtitle('여의도 한강공원 일자별 이용객') +
  theme(text = element_text(size = 10, family = "AppleGothic"))

```



#여의도 요일별 그래프 그리기

```
people_in_여의도 <- read_xlsx('여의도 이용객.xlsx')
people_in_여의도 <- people_in_여의도[-c(1, 4, 9:16, 18:48), -c(3, 4)]
```

```
colnames(people_in_여의도) <- str_c(people_in_여의도[1, ], people_in_여의도[2, ])
people_in_여의도 <- people_in_여의도[-c(1, 2), -1]
colnames(people_in_여의도)[1] <- c('이용목적')
```

#요일별 여의도 데이터 long으로 변환

```
people_in_여의도 <- people_in_여의도 %>%
  gather('요일', '이용객', -이용목적, na.rm = T)
```

```
#str(people_in_여의도) #각 컬럼별 class확인, 이용객 정수형으로 변환
people_in_여의도[, 3] <- as.integer(people_in_여의도$이용객)
```

#요일끼리 더해줘야함

#우선 요일이름을 맞춰주기 위해 앞에 날짜를 다 지워준다.

```
people_in_여의도$요일 <- str_remove_all(people_in_여의도$요일, '\\d')
```

#그리고 같은 요일끼리 이용목적에 따라 이용객 다 더해준다.

```
people_in_여의도 <- people_in_여의도 %>% group_by(이용목적, 요일) %>% summarise(이용객 = sum(이용객))
```

#일반이용자 3가지 구분 더해주기 위해

```
people_in_여의도 <- people_in_여의도 %>% spread(이용목적, 이용객, fill = 0)
people_in_여의도 <- data.frame(people_in_여의도, rowSums(people_in_여의도[, 3:5]))
people_in_여의도 <- people_in_여의도[, -c(3:5)]
colnames(people_in_여의도)[4] <- c('일반')
```

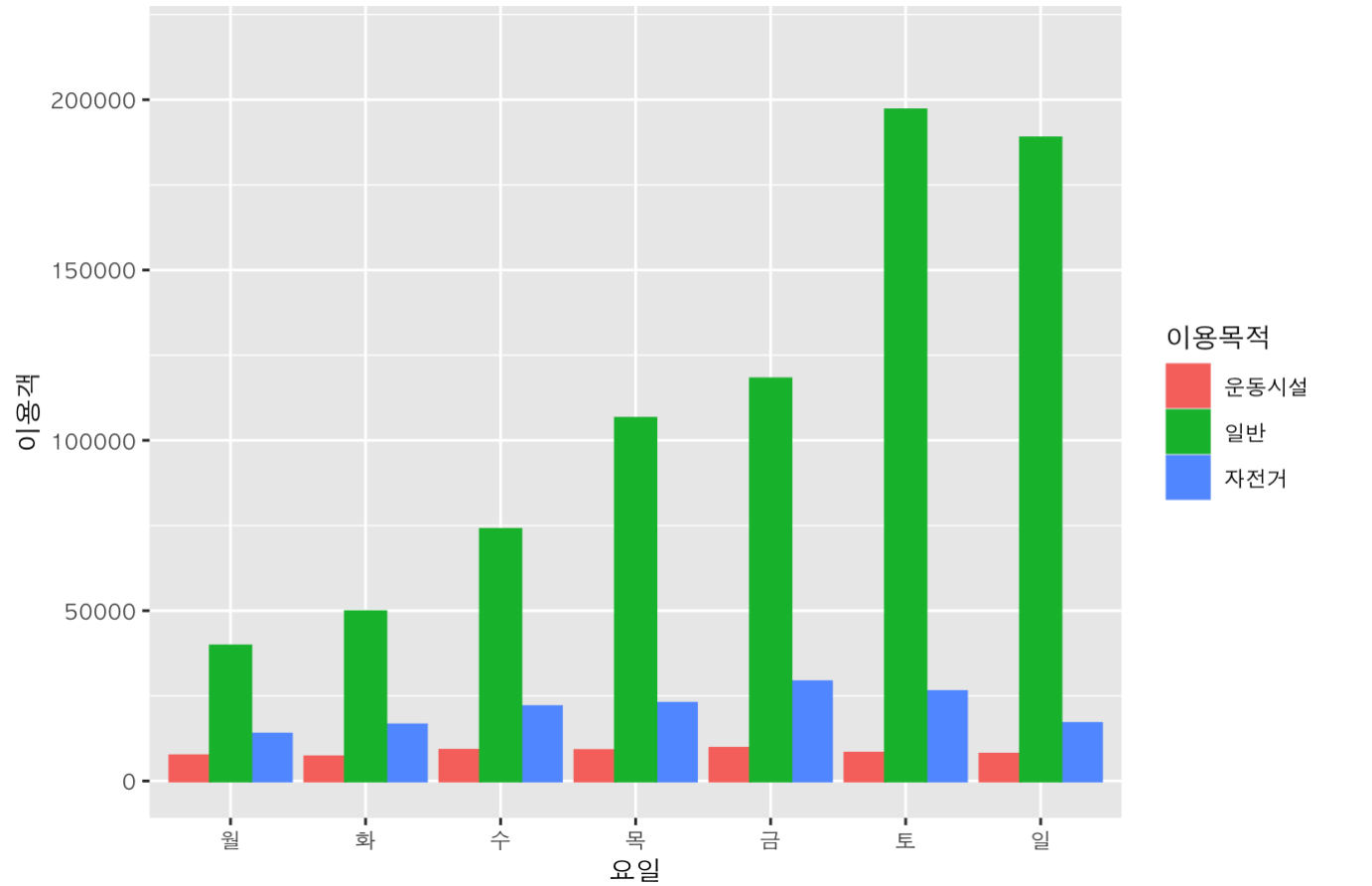
```
people_in_여의도 <- people_in_여의도 %>% gather(이용목적, 이용객, -요일)
```

#그래프 그리기

```
library(ggplot2)
```

```
ggplot(people_in_여의도, aes(x = 요일, y = 이용객, fill = 이용목적, colour = 이용목적)) +
  geom_col(stat = 'identity', position = 'dodge') +
  scale_x_discrete(limit = c('월', '화', '수', '목', '금', '토', '일')) +
  ylim(0, max(people_in_여의도$이용객*1.1)) +
  ggtitle('여의도 한강공원 요일별 이용객') +
  theme(text = element_text(size = 10, family = "AppleGothic"))
```


여의도 한강공원 요일별 이용객



그리고 독섬만큼 유명한 한강공원인 여의도도 마찬가지로 일반 이용객이 훨씬 높은 것을 볼 수 있습니다. 하지만 여의도는 독섬한강공원만큼 평일에 이용객이 많지는 않고 주말에 몰리는 결과입니다.

```

setwd(dir = "/Users/useonsong/Nano Degree/1차 프로젝트/외부데이터")
people_in_잠실 <- read_xlsx('잠실 이용객.xlsx')
people_in_잠실 <- people_in_잠실[-c(1, 4, 9:16, 18:48), -c(3, 4)]

colnames(people_in_잠실) <- people_in_잠실[1, ]
people_in_잠실 <- people_in_잠실[-1, -1]
colnames(people_in_잠실)[1] <- c('이용목적')

#일자별 잠실 데이터 long으로 변환
people_in_잠실 <- people_in_잠실 %>%
  gather('일자', '이용객', -이용목적, na.rm = T)
#결측값 삭제
people_in_잠실 <- people_in_잠실 %>%
  filter(is.na(people_in_잠실$이용목적) == FALSE)

library(stringr)
library(tidyverse)
people_in_잠실[, 2:3] <- map_df(people_in_잠실[, 2:3], as.integer)

#일반이용자 3구분을 더해서 하나로 묶기
library(tidyr)
people_in_잠실 <- people_in_잠실 %>% spread(key = 이용목적, value = 이용객, fill = 0)

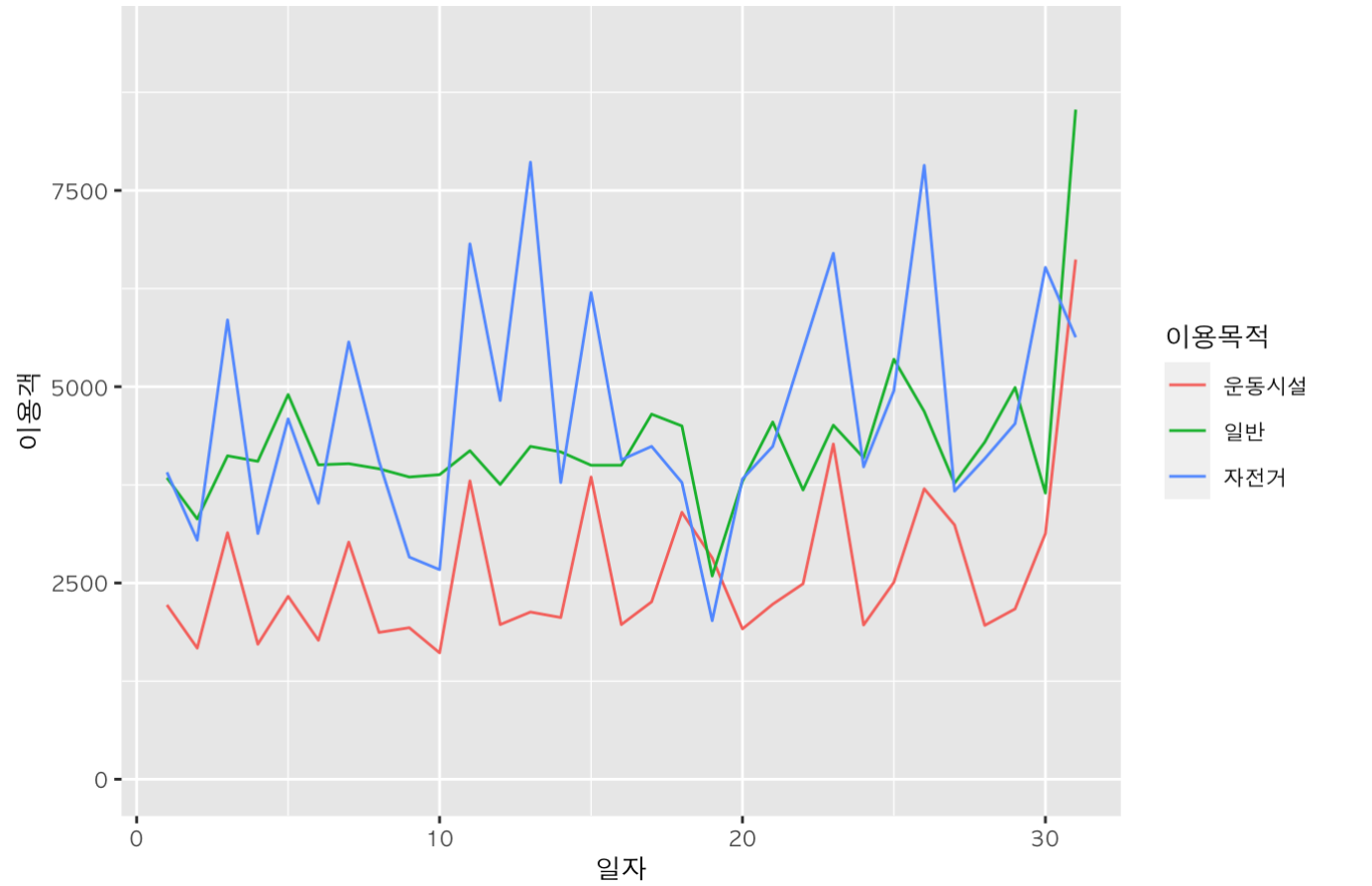
people_in_잠실 <- data.frame(people_in_잠실, rowSums(people_in_잠실[, 3:5]))
people_in_잠실 <- people_in_잠실[, -c(3:5)]
colnames(people_in_잠실)[4] <- c('일반')

people_in_잠실 <- people_in_잠실 %>%
  gather(이용목적, 이용객, -일자)

library(ggplot2)
ggplot(people_in_잠실, aes(x = 일자, y = 이용객, fill = 이용목적, colour = 이용목적)) +
  geom_line() +
  ylim(0, max(people_in_잠실$이용객*1.1)) +
  ggtitle('잠실 한강공원 일자별 이용객') +
  theme(text = element_text(size = 10, family = "AppleGothic"))

```

잠실 한강공원 일자별 이용객



```

people_in_잠실 <- read_xlsx('잠실 이용객.xlsx')
people_in_잠실 <- people_in_잠실[-c(1, 4, 9:16, 18:48), -c(3, 4)]

colnames(people_in_잠실) <- str_c(people_in_잠실[1, ], people_in_잠실[2, ])
people_in_잠실 <- people_in_잠실[-c(1, 2), -1]
colnames(people_in_잠실)[1] <- c('이용목적')

#요일별 잠실 데이터 long으로 변환

people_in_잠실 <- people_in_잠실 %>%
  gather('요일', '이용객', -이용목적, na.rm = T)

#str(people_in_잠실) #각 컬럼별 class확인, 이용객 정수형으로 변환
people_in_잠실[, 3] <- as.integer(people_in_잠실$이용객)

#요일끼리 더해줘야함
#우선 요일이름을 맞춰주기 위해 앞에 날짜를 다 지워준다.
people_in_잠실$요일 <- str_remove_all(people_in_잠실$요일, '\\d')

#그리고 같은 요일끼리 이용목적에 따라 이용객 다 더해준다.
people_in_잠실 <- people_in_잠실 %>% group_by(이용목적, 요일) %>% summarise(이용객 = sum(이용
객))

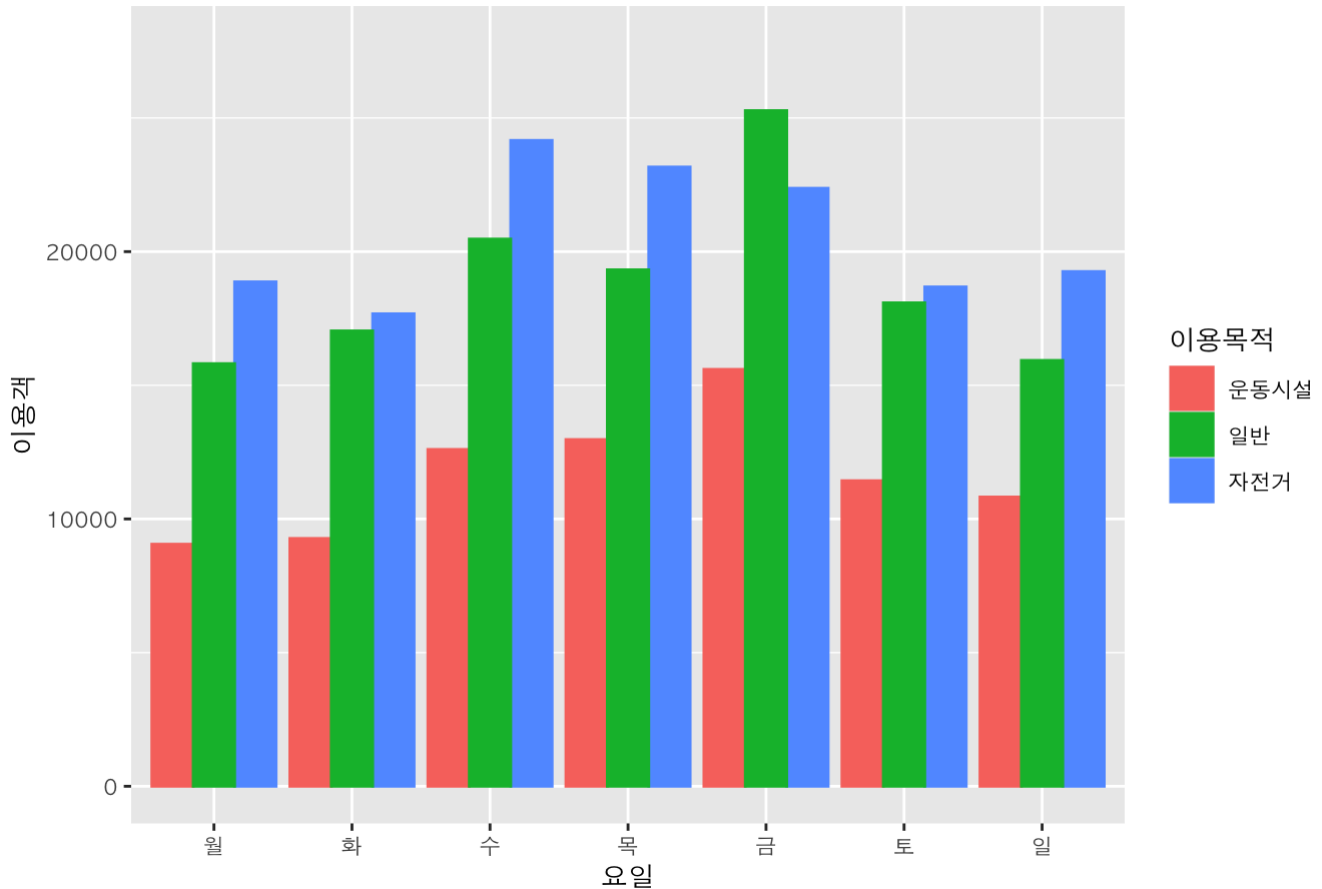
#일반이용자 3가지 구분 더해주기 위해
people_in_잠실 <- people_in_잠실 %>% spread(이용목적, 이용객, fill = 0)
people_in_잠실 <- data.frame(people_in_잠실, rowSums(people_in_잠실[, 3:5]))
people_in_잠실 <- people_in_잠실[, -c(3:5)]
colnames(people_in_잠실)[4] <- c('일반')

people_in_잠실 <- people_in_잠실 %>% gather(이용목적, 이용객, -요일)

#그래프 그리기
library(ggplot2)
ggplot(people_in_잠실, aes(x = 요일, y = 이용객, fill = 이용목적, colour = 이용목적)) +
  geom_col(stat = 'identity', position = 'dodge') +
  scale_x_discrete(limit = c('월', '화', '수', '목', '금', '토', '일')) +
  ylim(0, max(people_in_잠실$이용객*1.1)) +
  ggtitle('잠실 한강공원 요일별 이용객') +
  theme(text = element_text(size = 10, family = "AppleGothic"))

```

잠실 한강공원 요일별 이용객



그리고 잠실은 잠원에서 이어지는 자전거도로가 잘되어 있는만큼 오히려 일반이용객들보다 자전거 이용객이 평일 주말 관계 없이 높습니다.

```
setwd(dir = "/Users/useonsong/Nano Degree/1차 프로젝트/외부데이터")
```

```
library(readxl)
```

```
library(dplyr)
```

```
list.files(pattern = 'xlsx')
```

```
## [1] "강서 이용객.xlsx" "광나루 이용객.xlsx" "난지 이용객.xlsx"
## [4] "뚝섬 이용객.xlsx" "망원 이용객1.xlsx" "반포 이용객.xlsx"
## [7] "여의도 이용객.xlsx" "이촌 이용객.xlsx" "잠실 이용객.xlsx"
## [10] "잠원 이용객.xlsx"
```

```

people_in_광나루 <- readxl::read_xlsx("광나루 이용객.xlsx")

people_in_광나루 <- people_in_광나루[-c(1, 4, 9:16, 18:22), -c(3)]

colnames(people_in_광나루) <- people_in_광나루[1, ]
people_in_광나루 <- people_in_광나루[-c(1, 2), -c(1,4:34)]

people_in_광나루[, 2] <- as.integer(unlist(people_in_광나루[, 2]))

일반 <- c('일반', sum(people_in_광나루[1:3, 2]))
people_in_광나루 <- rbind(일반, people_in_광나루)

people_in_광나루 <- people_in_광나루[-c(2:4), ]

colnames(people_in_광나루)[1:2] <- c('이용목적', '이용객')

people_in_광나루 <- people_in_광나루 %>% spread('이용목적', '이용객')

people_in_광나루 <- as.data.frame(people_in_광나루)
rownames(people_in_광나루)[1] <- c('광나루')

#난지

people_in_난지 <- readxl::read_xlsx("난지 이용객.xlsx")

people_in_난지 <- people_in_난지[-c(1:4, 9:16, 18:48), -c(1, 4:34)]

people_in_난지[, 2] <- as.integer(unlist(people_in_난지[, 2]))

일반 <- c('일반', sum(people_in_난지[1:3, 2]))
people_in_난지 <- rbind(일반, people_in_난지)

people_in_난지 <- people_in_난지[-c(2:4), ]

colnames(people_in_난지)[1:2] <- c('이용목적', '이용객')

people_in_난지 <- people_in_난지 %>% spread('이용목적', '이용객')

people_in_난지 <- as.data.frame(people_in_난지)
rownames(people_in_난지)[1] <- c('난지')

#뚝섬

people_in_뚝섬 <- readxl::read_xlsx("뚝섬 이용객.xlsx")

```

```

people_in_뚝섬 <- people_in_뚝섬[-c(1:3, 8:15, 17:28), -c(1, 4:34)]
people_in_뚝섬[, 2] <- as.integer(unlist(people_in_뚝섬[, 2]))

일반 <- c('일반', sum(people_in_뚝섬[1:3, 2]))
people_in_뚝섬 <- rbind(일반, people_in_뚝섬)

people_in_뚝섬 <- people_in_뚝섬[-c(2:4), ]

colnames(people_in_뚝섬)[1:2] <- c('이용목적', '이용객')

people_in_뚝섬 <- people_in_뚝섬 %>% spread('이용목적', '이용객')

people_in_뚝섬 <- as.data.frame(people_in_뚝섬)
rownames(people_in_뚝섬)[1] <- c('뚝섬')

#망원
people_in_망원 <- readxl::read_xlsx("망원 이용객1.xlsx")

people_in_망원 <- people_in_망원[-c(1:4, 9:16, 18:48), -c(1, 3, 5:35)]

people_in_망원[, 2] <- as.integer(unlist(people_in_망원[, 2]))

일반 <- c('일반', sum(people_in_망원[1:3, 2]))
people_in_망원 <- rbind(일반, people_in_망원)

people_in_망원 <- people_in_망원[-c(2:4), ]

colnames(people_in_망원)[1:2] <- c('이용목적', '이용객')

people_in_망원 <- people_in_망원 %>% spread('이용목적', '이용객')

people_in_망원 <- as.data.frame(people_in_망원)
rownames(people_in_망원)[1] <- c('망원')

#반포
people_in_반포 <- readxl::read_xlsx("반포 이용객.xlsx")

people_in_반포 <- people_in_반포[-c(1:4, 9:16, 18:48), -c(1, 3, 5:36)]

people_in_반포[, 2] <- as.integer(unlist(people_in_반포[, 2]))

```

```

일반 <- c('일반', sum(people_in_반포[1:3, 2]))
people_in_반포 <- rbind(일반, people_in_반포)

people_in_반포 <- people_in_반포[-c(2:4), ]

colnames(people_in_반포)[1:2] <- c('이용목적', '이용객')

people_in_반포 <- people_in_반포 %>% spread('이용목적', '이용객')

people_in_반포 <- as.data.frame(people_in_반포)
rownames(people_in_반포)[1] <- c('반포')


#여의도
people_in_여의도 <- readxl::read_xlsx("여의도 이용객.xlsx")

people_in_여의도 <- people_in_여의도[-c(1:4, 9:16, 18:36), -c(1, 3, 5:35)]

people_in_여의도[, 2] <- as.integer(unlist(people_in_여의도[, 2]))

일반 <- c('일반', sum(people_in_여의도[1:3, 2]))
people_in_여의도 <- rbind(일반, people_in_여의도)

people_in_여의도 <- people_in_여의도[-c(2:4), ]

colnames(people_in_여의도)[1:2] <- c('이용목적', '이용객')

people_in_여의도 <- people_in_여의도 %>% spread('이용목적', '이용객')

people_in_여의도 <- as.data.frame(people_in_여의도)
rownames(people_in_여의도)[1] <- c('여의도')


#이촌
people_in_이촌 <- readxl::read_xlsx("이촌 이용객.xlsx")

people_in_이촌 <- people_in_이촌[-c(1:4, 9:16, 18:22), -c(1, 3, 5:35)]

people_in_이촌[, 2] <- as.integer(unlist(people_in_이촌[, 2]))

```



```

일반 <- c('일반', sum(people_in_이촌[1:3, 2]))
people_in_이촌 <- rbind(일반, people_in_이촌)

people_in_이촌 <- people_in_이촌[-c(2:4), ]

colnames(people_in_이촌)[1:2] <- c('이용목적', '이용객')

people_in_이촌 <- people_in_이촌 %>% spread('이용목적', '이용객')

people_in_이촌 <- as.data.frame(people_in_이촌)
rownames(people_in_이촌)[1] <- c('이촌')

#자전거가 띄어쓰기로 입력되어있어서,,
colnames(people_in_이촌)[3] <- c('자전거')


#잠실

people_in_잠실 <- readxl::read_xlsx("잠실 이용객.xlsx")

people_in_잠실 <- people_in_잠실[-c(1:4, 9:16, 18:48), -c(1, 3, 5:35)]

people_in_잠실[, 2] <- as.integer(unlist(people_in_잠실[, 2]))

일반 <- c('일반', sum(people_in_잠실[1:3, 2]))
people_in_잠실 <- rbind(일반, people_in_잠실)

people_in_잠실 <- people_in_잠실[-c(2:4), ]

colnames(people_in_잠실)[1:2] <- c('이용목적', '이용객')

people_in_잠실 <- people_in_잠실 %>% spread('이용목적', '이용객')

people_in_잠실 <- as.data.frame(people_in_잠실)
rownames(people_in_잠실)[1] <- c('잠실')


#잠원

people_in_잠원 <- readxl::read_xlsx("잠원 이용객.xlsx")

people_in_잠원 <- people_in_잠원[-c(1:4, 9:16, 18:48), -c(1, 3, 5:35)]

```

```

people_in_잠원[, 2] <- as.integer(unlist(people_in_잠원[, 2]))

일반 <- c('일반', sum(people_in_잠원[1:3, 2]))
people_in_잠원 <- rbind(일반, people_in_잠원)

people_in_잠원 <- people_in_잠원[-c(2:4), ]

colnames(people_in_잠원)[1:2] <- c('이용목적', '이용객')

people_in_잠원 <- people_in_잠원 %>% spread('이용목적', '이용객')

people_in_잠원 <- as.data.frame(people_in_잠원)
rownames(people_in_잠원)[1] <- c('잠원')


#강서

people_in_강서 <- readxl::read_xlsx("강서 이용객.xlsx")

people_in_강서 <- people_in_강서[-c(1:4, 9:16, 18:48), -c(1, 3, 5:35)]

people_in_강서[, 2] <- as.integer(unlist(people_in_강서[, 2]))

일반 <- c('일반', sum(people_in_강서[1:3, 2]))
people_in_강서 <- rbind(일반, people_in_강서)

people_in_강서 <- people_in_강서[-c(2:4), ]

colnames(people_in_강서)[1:2] <- c('이용목적', '이용객')

people_in_강서 <- people_in_강서 %>% spread('이용목적', '이용객')

people_in_강서 <- as.data.frame(people_in_강서)

rownames(people_in_강서)[1] <- c('강서')


#people_in_ rbind로 다 합치기

레이더 <- rbind(people_in_강서, people_in_광나루, people_in_난지, people_in_뚝섬, people_in_
망원, people_in_반포, people_in_여의도, people_in_이촌, people_in_잠실, people_in_잠원)

레이더[, 1:3] <- map_df(레이더[, 1:3], as.integer)

레이더 <- t(as.matrix(레이더))
레이더 <- as.data.frame(레이더)

레이더 <- rbind(rep(995200, 10), rep(0, 10), 레이더)


#방사형 차트 그리기


library(fmsb)
par(family = 'AppleGothic', mar = c(0, 0, 0, 0))

```

```

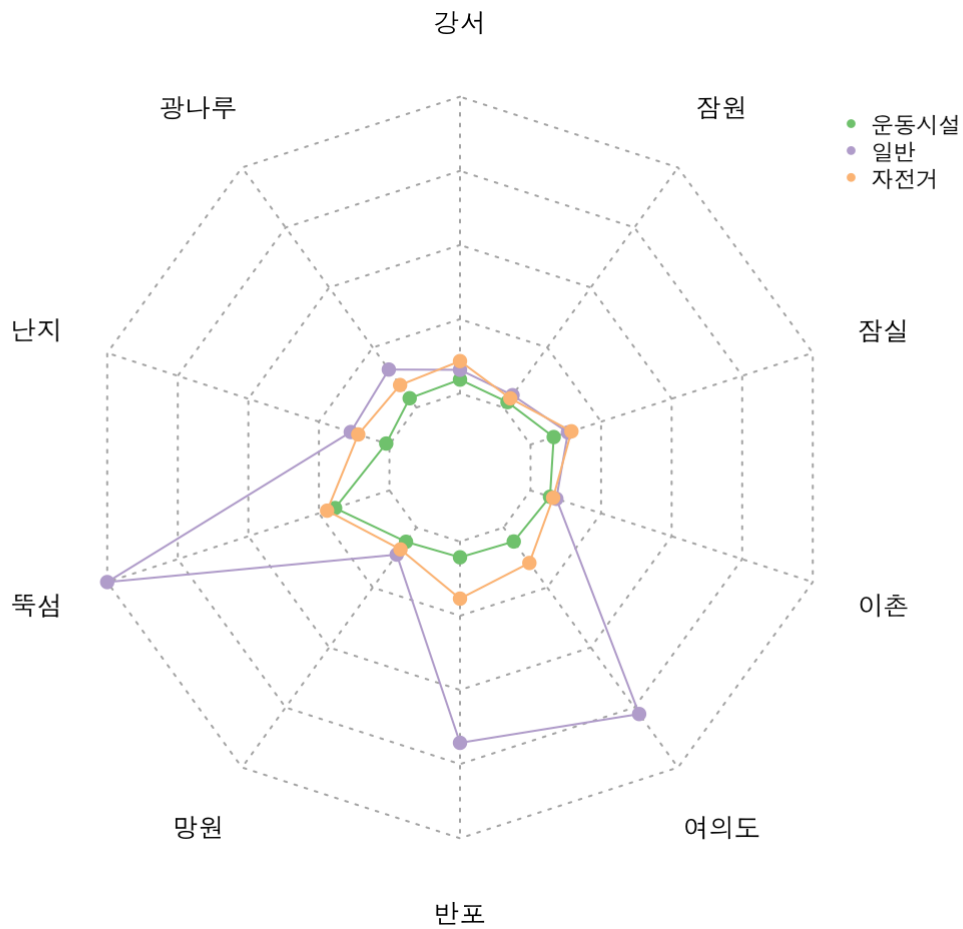
library(RColorBrewer)

colors_border = brewer.pal(n = 3, name = 'Accent')

radarchart(레이더,
           cglty = 3, cglcol = 'gray70', cglwd = 1,
           pcol = colors_border, plwd = 1, plty = 1,
           vlce = 0.8)

legend(x = 1, y = 1, legend = rownames(레이더[-c(1, 2), ]), bty = 'n', pch = 20, col =
       colors_border, cex = 0.7)

```



이처럼 한강공원 이용객 분석 결과 독섬, 여의도, 반포 한강공원이 일반 이용객이 눈에 띄게 높은 것을 알 수 있었고, 이런 높은 일반 이용객 비율이 배달 음식 주문으로 이어지는지 살펴보았습니다.

```

setwd(dir = "/Users/useonsong/Nano Degree/1차 프로젝트/외부데이터")

library(readxl)
par(family = 'AppleGothic')
뚝섬_주문량 <- readRDS("MAY_Ttukseom.RDS")
반포_주문량 <- readRDS("MAY_BANPO.RDS")
여의도_주문량 <- readRDS("MAY_Yeouido.RDS")
난지_주문량 <- readRDS("MAY_NANJI.RDS")

#뚝섬

people_in_뚝섬 <- read_xlsx('뚝섬 이용객.xlsx')
people_in_뚝섬 <- people_in_뚝섬[-c(3, 8:15, 17:28), -3 ]
colnames(people_in_뚝섬) <- people_in_뚝섬[1, ]
people_in_뚝섬 <- people_in_뚝섬[-1, -1]
colnames(people_in_뚝섬)[1] <- c('이용목적')

#일자별 뚝섬 데이터 long으로 변환
people_in_뚝섬 <- people_in_뚝섬 %>%
  gather('일자', '이용객', -이용목적, na.rm = T)
#결측값 삭제
people_in_뚝섬 <- people_in_뚝섬 %>%
  filter(is.na(people_in_뚝섬$이용목적) == FALSE)

#str(people_in_뚝섬) #각 컬럼별 class확인, 일자, 이용객 정수형으로 변환
library(stringr)
library(tidyverse)
people_in_뚝섬[, 2:3] <- map_df(people_in_뚝섬[, 2:3], as.integer)

#일반이용자 3구분을 더해서 하나로 묶기
library(tidyr)
people_in_뚝섬 <- people_in_뚝섬 %>% spread(key = 이용목적, value = 이용객, fill = 0)
people_in_뚝섬 <- data.frame(people_in_뚝섬, rowSums(people_in_뚝섬[, 3:5]))
people_in_뚝섬 <- people_in_뚝섬[, -c(3:5)]
colnames(people_in_뚝섬)[4] <- c('일반')
people_in_뚝섬 <- people_in_뚝섬 %>%
  gather(이용목적, 이용객, -일자)

#일반이용객만 비교하면 되니까 운동시설, 자전거 이용객 삭제
people_in_뚝섬 <- people_in_뚝섬[-c(1:62), ]

#뚝섬 조인
뚝섬_주문량[, 1] <- people_in_뚝섬[, 1]
뚝섬비교 <- left_join(people_in_뚝섬, 뚝섬_주문량, '일자')
뚝섬비교 <- as.data.frame(뚝섬비교)

par(mar = c(5, 4, 4, 6) + 0.1)

plot(뚝섬비교$이용객 ~ 뚝섬비교$일자, type = 'l', col = 'black', axes = FALSE, xlab = '', ylab = '', main = '뚝섬 한강공원 이용객과 통화건수 비교')
axis(side = 2, )
mtext("이용객", side = 2, line = 2.5)
box()

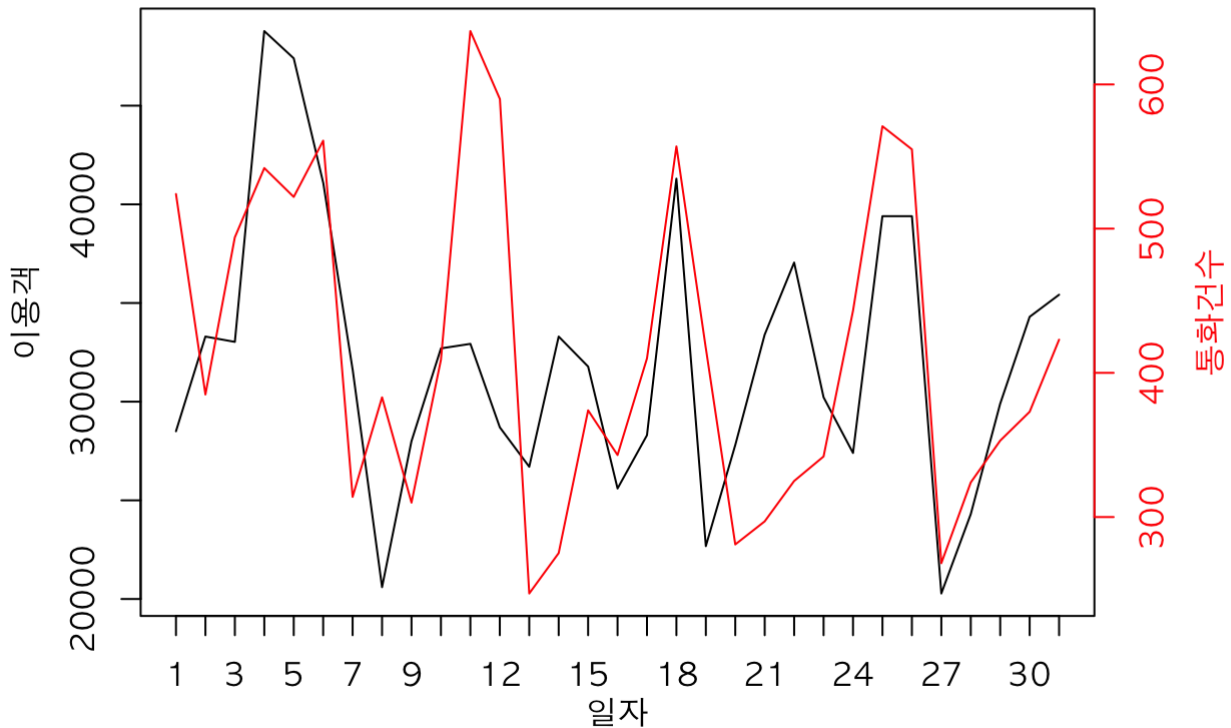
```

```

par(new = TRUE)
plot(독섬비교$통화건수 ~ 독섬비교$일자, type = "l", col = 'red', axes = FALSE, xlab = "", ylab = "")
axis(side = 4, col = "red", col.axis = "red")
mtext("통화건수", side = 4, col = "red", line = 2.5)
axis(side = 1, at = 1:31, labels = 독섬비교$일자)
mtext("일자", side = 1, line = 2)

```

독섬 한강공원 이용객과 통화건수 비교



```

#반포
people_in_반포 <- read_xlsx('반포 이용객.xlsx')
people_in_반포 <- people_in_반포[-c(1, 4:5, 9:16, 18:48), -c(3, 4, 35)]
colnames(people_in_반포) <- people_in_반포[1, ]
people_in_반포 <- people_in_반포[-1, -1]
colnames(people_in_반포)[1] <- c('이용목적')

#일자별 반포 데이터 long으로 변환
people_in_반포 <- people_in_반포 %>%
  gather('일자', '이용객', -이용목적, na.rm = T)

#결측값 삭제
people_in_반포 <- people_in_반포 %>%
  filter(is.na(people_in_반포$이용목적) == FALSE)
str(people_in_반포) #각 컬럼별 class확인, 일자, 이용객 정수형으로 변환

```

```
## tibble [120 × 3] (S3: tbl_df/tbl/data.frame)
## $ 이용목적: chr [1:120] "일반이용자(낮)" "일반이용자(저녁)" "운동시설" "자전거" ...
## $ 일자      : chr [1:120] "1" "1" "1" "1" ...
## $ 이용객    : chr [1:120] "780" "29655" "4598" "18051" ...
```

```
library(stringr)
library(tidyverse)
people_in_반포[, 2:3] <- map_df(people_in_반포[, 2:3], as.integer)

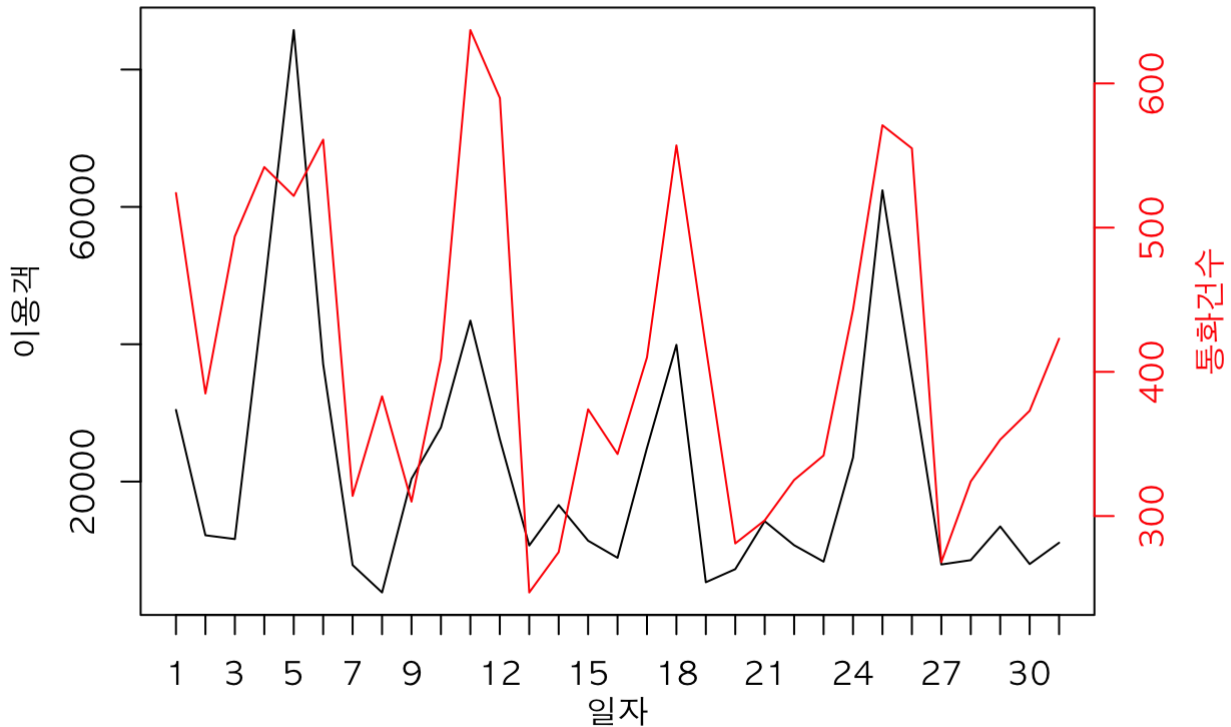
#일반이용자 3구분을 더해서 하나로 묶기
library(tidyr)
people_in_반포 <- people_in_반포 %>% spread(key = 이용목적, value = 이용객, fill = 0)
people_in_반포 <- data.frame(people_in_반포, rowSums(people_in_반포[, 3:4]))
people_in_반포 <- people_in_반포[, -c(3:4)]
colnames(people_in_반포)[4] <- c('일반')
people_in_반포 <- people_in_반포 %>%
  gather(이용목적, 이용객, -일자)

#일반이용객만 비교하면 되니까 운동시설, 자전거 이용객 삭제
people_in_반포 <- people_in_반포[-c(1:62), ]

#반포 조인
반포_주문량[, 1] <- people_in_반포[, 1]
반포비교 <- left_join(people_in_반포, 반포_주문량, '일자')

plot(반포비교$이용객 ~ 반포비교$일자, type = 'l', col = 'black', axes = FALSE, xlab = '', ylab = '', main = '반포 한강공원 이용객과 통화건수 비교')
axis(side = 2, )
mtext("이용객", side = 2, line = 2.5)
box()
par(new = TRUE)
plot(반포비교$통화건수 ~ 반포비교$일자, type = "l", col = 'red', axes = FALSE, xlab = "", ylab = "")
axis(side = 4, col = "red", col.axis = "red")
mtext("통화건수", side = 4, col = "red", line = 2.5)
axis(side = 1, at = 1:31, labels = 반포비교$일자)
mtext("일자", side = 1, line = 2)
```

반포 한강공원 이용객과 통화건수 비교



```
#여의도
```

```
people_in_여의도 <- read_xlsx('여의도 이용객.xlsx')
people_in_여의도 <- people_in_여의도[-c(1, 4, 9:16, 18:36), -c(3, 4)]
colnames(people_in_여의도) <- people_in_여의도[1, ]
people_in_여의도 <- people_in_여의도[-1, -1]
colnames(people_in_여의도)[1] <- c('이용목적')
```

```
#일자별 여의도 데이터 long으로 변환
```

```
people_in_여의도 <- people_in_여의도 %>%
  gather('일자', '이용객', -이용목적, na.rm = T)
```

```
#결측값 삭제
```

```
people_in_여의도 <- people_in_여의도 %>%
  filter(is.na(people_in_여의도$이용목적) == FALSE)
str(people_in_여의도) #각 컬럼별 class확인, 일자, 이용객 정수형으로 변환
```

```
## tibble [155 × 3] (S3: tbl_df/tbl/data.frame)
```

```
## $ 이용목적: chr [1:155] "일반이용자(아침)\r\n05:00-09:00" "일반이용자(낮)\r\n09:00-17:00"
"일반이용자(저녁)\r\n17:00-05:00" "운동시설" ...
```

```
## $ 일자 : chr [1:155] "1" "1" "1" "1" ...
```

```
## $ 이용객 : chr [1:155] "550" "7050" "5980" "1470" ...
```

```

library(stringr)
library(tidyverse)
people_in_여의도[, 2:3] <- map_df(people_in_여의도[, 2:3], as.integer)

#일반이용자 3구분을 더해서 하나로 묶기
library(tidyr)
people_in_여의도 <- people_in_여의도 %>% spread(key = 이용목적, value = 이용객, fill = 0)
people_in_여의도 <- data.frame(people_in_여의도, rowSums(people_in_여의도[, 3:5]))
people_in_여의도 <- people_in_여의도[, -c(3:5)]
colnames(people_in_여의도)[4] <- c('일반')
people_in_여의도 <- people_in_여의도 %>%
  gather(이용목적, 이용객, -일자)

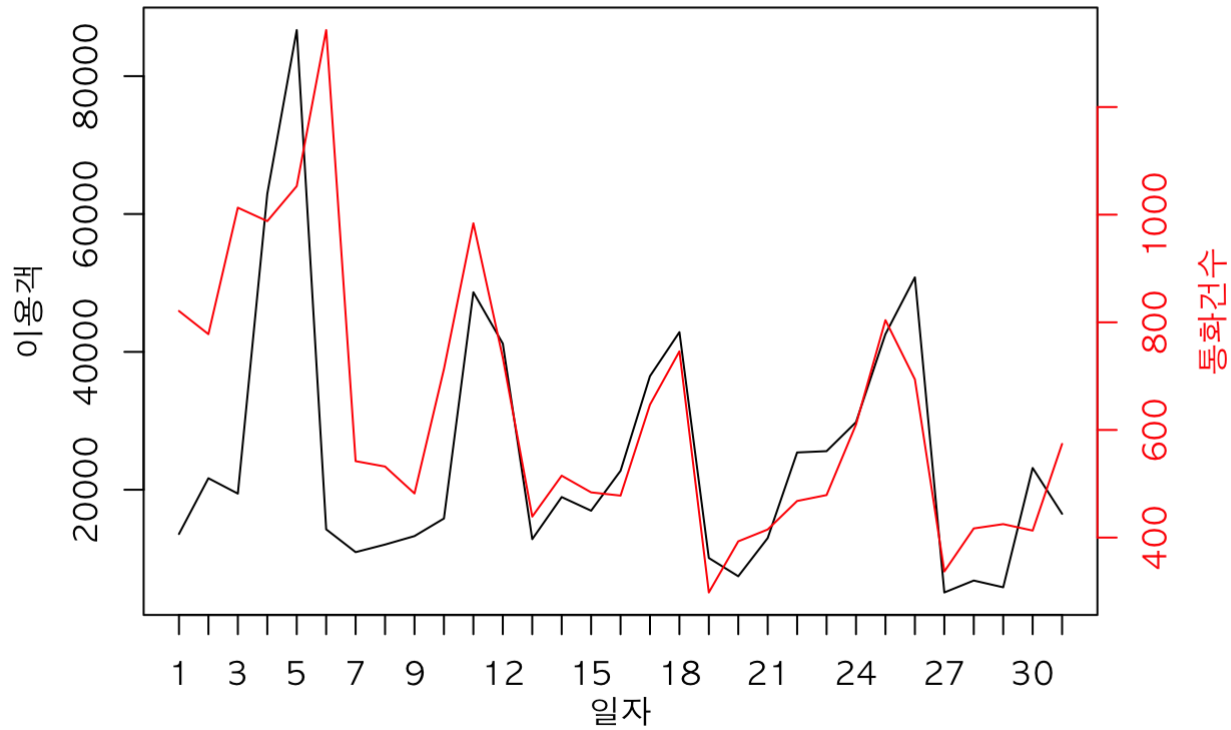
#일반이용객만 비교하면 되니까 운동시설, 자전거 이용객 삭제
people_in_여의도 <- people_in_여의도[-c(1:62), ]

#여의도 조인
여의도_주문량[, 1] <- people_in_여의도[, 1]
여의도비교 <- left_join(people_in_여의도, 여의도_주문량, '일자')

plot(여의도비교$이용객 ~ 여의도비교$일자, type = 'l', col = 'black', axes = FALSE, xlab = '',
     ylab = '', main = '여의도 한강공원 이용객과 통화건수 비교')
axis(side = 2, )
mtext("이용객", side = 2, line = 2.5)
box()
par(new = TRUE)
plot(여의도비교$통화건수 ~ 여의도비교$일자, type = "l", col = 'red', axes = FALSE, xlab = "",
     ylab = "")
axis(side = 4, col = "red", col.axis = "red")
mtext("통화건수", side = 4, col = "red", line = 2.5)
axis(side = 1, at = 1:31, labels = 여의도비교$일자)
mtext("일자", side = 1, line = 2)

```


여의도 한강공원 이용객과 통화건수 비교



일반 이용객 수와 배달 주문건수의 상관 관계

일반 이용객 비율이 높은 뚝섬, 반포, 여의도 한강공원 이용객과 이 지역의 배달음식 주문 통화건수를 이중 y축 그래프를 이용해서 살펴보았습니다. 대체로 비슷한 패턴을 보이지만 작년 5월 석가탄신일 대체공휴일에 통화건수가 튀는 것으로 보아 이 기간에는 그 지역 한강공원에서만 주문한 것이 아니라 같은 동네에 포함되어 있는 주민들의 주문건까지 포함된 것으로 생각합니다. 이 정도의 오차말고는 패턴을 분석했을 때 한강공원의 일반 이용객과 배달음식 주문 건수의 상관관계가 있다고 가정하고 다음 분석을 진행하였습니다.