homework3

```
library(tidyverse)
```

```
── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
✔ dplyr     1.1.4     ✔ readr     2.1.5
✔ forcats   1.0.0     ✔ stringr   1.5.1
✔ ggplot2   3.4.4     ✔ tibble    3.2.1
✔ lubridate 1.9.3     ✔ tidyr     1.3.1
✔ purrr     1.0.2
── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
✖ dplyr::filter() masks stats::filter()
✖ dplyr::lag()    masks stats::lag()
ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to
```

```
library(ggplot2)
library(lubridate)
library(dplyr)
```

Problem 1:

• Load the wmata_ridership data frame into R from https://dcgerard.github.io/stat_412_612/data/wmata_ridership.csv. For each month, calculate the proportion of rides made on a given day of the month. Then make box plots of the proportions of ridership vs day of the weak. But exclude any days from 2004 and 2005.

```
wmata_ridership <- read.csv("https://dcgerard.github.io/stat_412_612/data/wmata_ridersh
```

```
head(wmata_ridership)
```

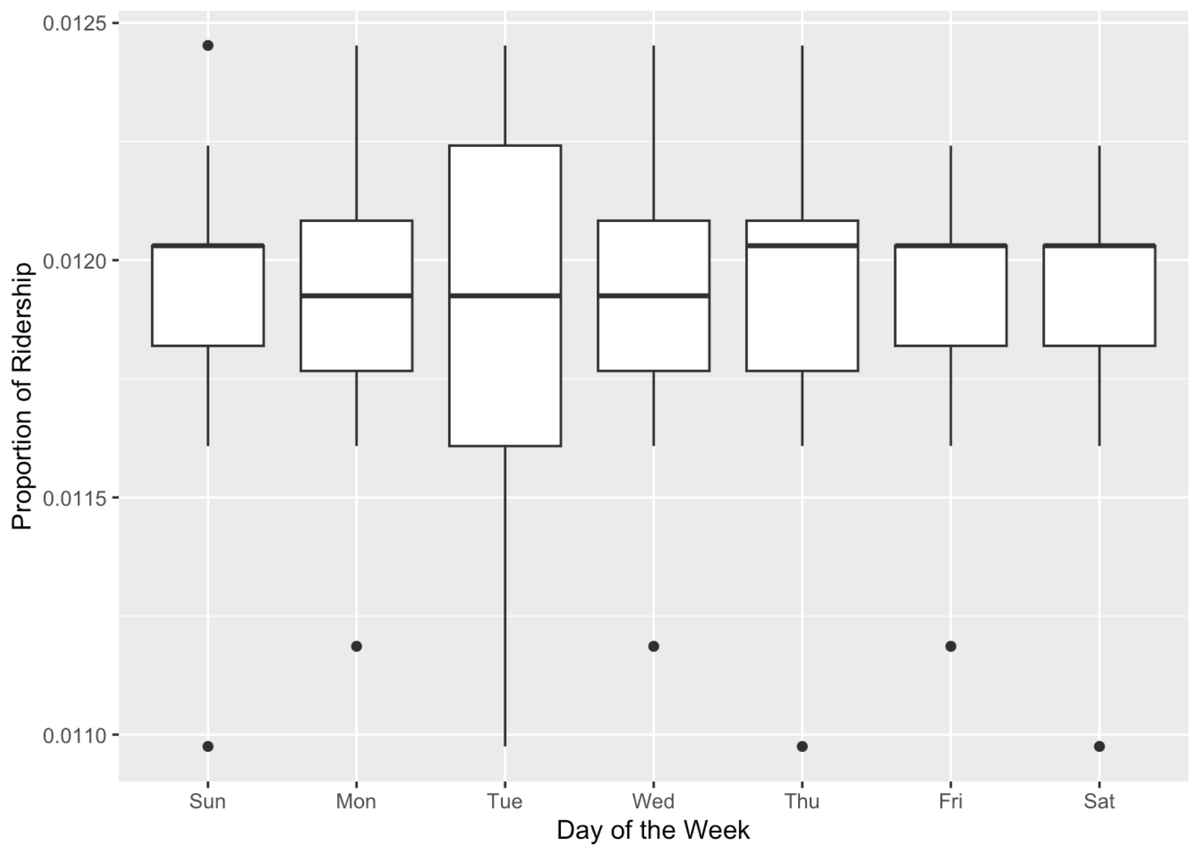
```
        Date  Total
1 2004-01-01 129000
2 2004-01-02 419000
3 2004-01-03 222000
4 2004-01-04 140000
5 2004-01-05 564000
6 2004-01-06 609000
```

```
wmata_ridership$Date <- as.Date(wmata_ridership$Date)
```

```
wmata_ridership <- wmata_ridership %>% filter(year(Date) != 2004 & year(Date) != 2005)
```

```
wmata_ridership <- wmata_ridership %>%
  mutate(Day_of_month = day(Date),
         Day_of_week = wday(Date, label = TRUE, abbr = TRUE)) %>%
  group_by(month(Date), Day_of_week) %>%
  summarise(prop_rides = n() / nrow(wmata_ridership), .groups = 'drop')
```

```
ggplot(wmata_ridership, aes(x = Day_of_week, y = prop_rides)) +
  geom_boxplot() +
  labs(x = "Day of the Week", y = "Proportion of Ridership")
```



Problem 2:

• Write only one line of the code to return only the leap years fo years 2024, 3021, 2000 or 1800, 800, 5050, 600 and 7156.

```
leap_years <- c(2024, 3021, 2000, 1800, 800, 5050, 600, 7156)
leap_years[leap_years %% 4 == 0 & (leap_years %% 100 != 0 | leap_years %% 400 == 0)]
```

```
[1] 2024 2000  800 7156
```