HW04_XU ZHANG_DSO545

Xu Zhang

March 5, 2018

```
setwd("C:/Users/Xu Zhang/Desktop/DSO545")
load("hubwaytrip2012.rda")
library(lubridate)

##
## Attaching package: 'lubridate'

## The following object is masked from 'package:base':
##
##     date

library(dplyr)

##
## Attaching package: 'dplyr'

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

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

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

library(ggplot2)
```

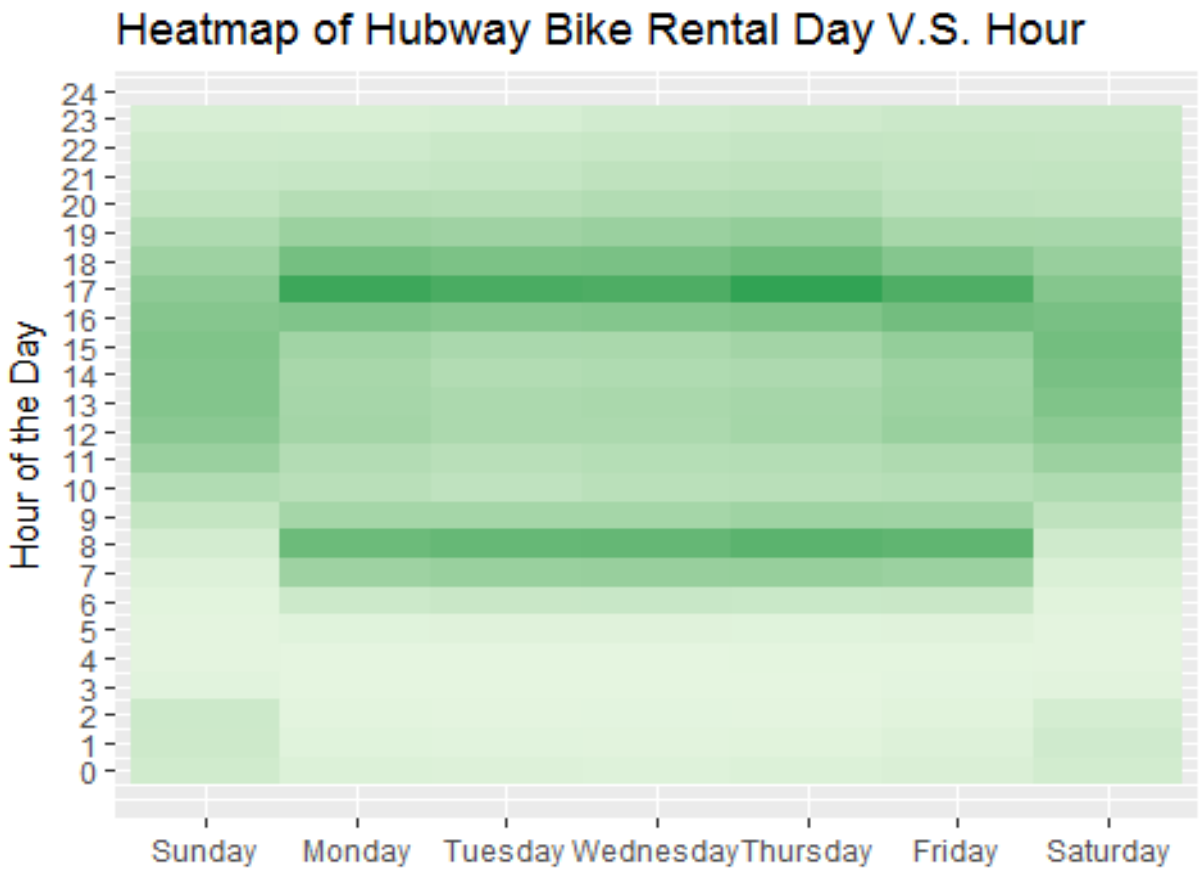
1. Use functions from dplyr and lubridate to summarise by Day and Start Time (Hour of the day) of bike rental usage. Save your new dataframe in a variable called trips2012Time, and output head(trips2012Time) to show the first few observations of the new dataframe. (Make sure to set the timezone to "EST")

```
trips2012$start_date=mdy_hm(trips2012$start_date)
trips2012$start_date=force_tz(trips2012$start_date,"EST")
trips2012Time=trips2012 %>%
  mutate(Day=wday(start_date,label=TRUE,abbr=FALSE),
           Start_time=hour(start_date)) %>%
  group_by(Day,Start_time) %>%
  summarise(count=n())
head(trips2012Time)
```

```
## # A tibble: 6 x 3
## # Groups: Day [1]
##   Day    Start_time count
##   <ord>       <int> <int>
## 1 Sunday          0  1384
## 2 Sunday          1  1607
## 3 Sunday          2  1660
## 4 Sunday          3   284
## 5 Sunday          4   128
## 6 Sunday          5   148
```

2. Create the following heatmap , use color #e5f5e0 and #31a354. Summarise what you see.

```
ggplot(trips2012Time,aes(x=Day,y=Start_time,fill=count))+
  geom_tile()+
  scale_fill_gradient(low = "#e5f5e0",high ="#31a354",guide = FALSE)+
  labs(title="Heatmap of Hubway Bike Rental Day V.S. Hour",x="",y="Hour of th
e Day")+
  scale_y_continuous(breaks = seq(0,24,1))
```



People rent bike more on weekdays than on weekends. People rent bike more on 7:00 AM to 9:00 AM and 4:00 PM to 6:00 in a day than other time.