

EDA_Executive_Summary

Elizabeth Hora

12/6/2020

Loading Packages

```
library(tidyverse)
library(ggplot2)
library(readxl)
library(lubridate)
```

Coronavirus Pandemic

Loading Code from My Final Project

Each step matches what is found in my Final Project

```

# Step 1
paths_to_open <- c("../data/mass_pike/framingham",
                  "../data/mass_pike/cochituate",
                  "../data/mass_pike/newton",
                  "../data/mass_pike/pre_allston",
                  "../data/mass_pike/post_allston"
                  )

files <- dir(paths_to_open, pattern = "\\\\.xlsx$", full.names = TRUE)

pike_file_list <- list()

# Step 2
times <- as.character(0:23)

for (i in files){
  pike_file_list[[i]] <- read_excel(i, skip = 10, col_names = c("day", times, "Total", "QC Status")) %>%
    pivot_longer(c(`0`, `1`, `2`, `3`, `4`, `5`, `6`, `7`, `8`, `9`, `10`, `11`, `12`, `13`, `14`,
                  `15`, `16`, `17`, `18`, `19`, `20`, `21`, `22`, `23`),
                 names_to = "hour", values_to = "cars_per_hour") %>%
    select("day", "hour", "cars_per_hour", "Total")
}

# Step 3
months <- str_extract(names(pike_file_list), "[A-Z]B\\d+") %>%
  str_sub(start = 5) %>%
  as.double()

years <- str_extract(names(pike_file_list), "\\d{4}") %>%
  as.double()

travel_direction <- str_extract(names(pike_file_list), "[A-Z]B_") %>%
  str_sub(start = 2L, end = -2L)

location <- str_extract(names(pike_file_list), "/mass_pike/..") %>%
  str_sub(start = 12L) %>%
  as.character()

# Step 4
for (i in 1:length(files)){
  pike_file_list[[i]] <- pike_file_list[[i]] %>%
    mutate(hour = as.numeric(hour)) %>%
    mutate(full_date = make_datetime(year = years[i], month = months[i], day = day, hour = hour)) %>%
    mutate(day_of_week = wday(full_date, label = TRUE, abbr = FALSE)) %>%
    mutate(direction = travel_direction[i]) %>%
    mutate(location = location[i])
}

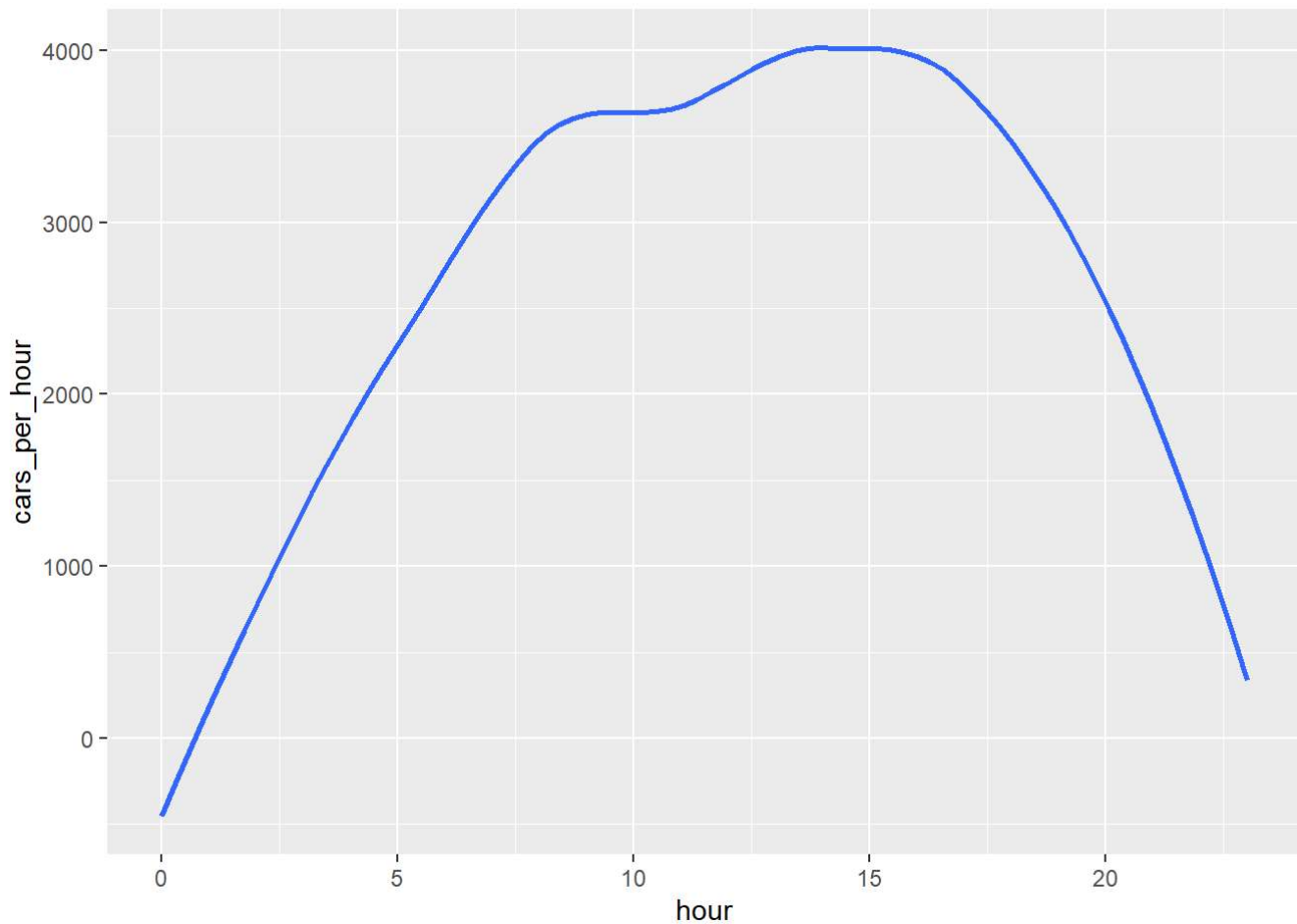
# Step 5

```

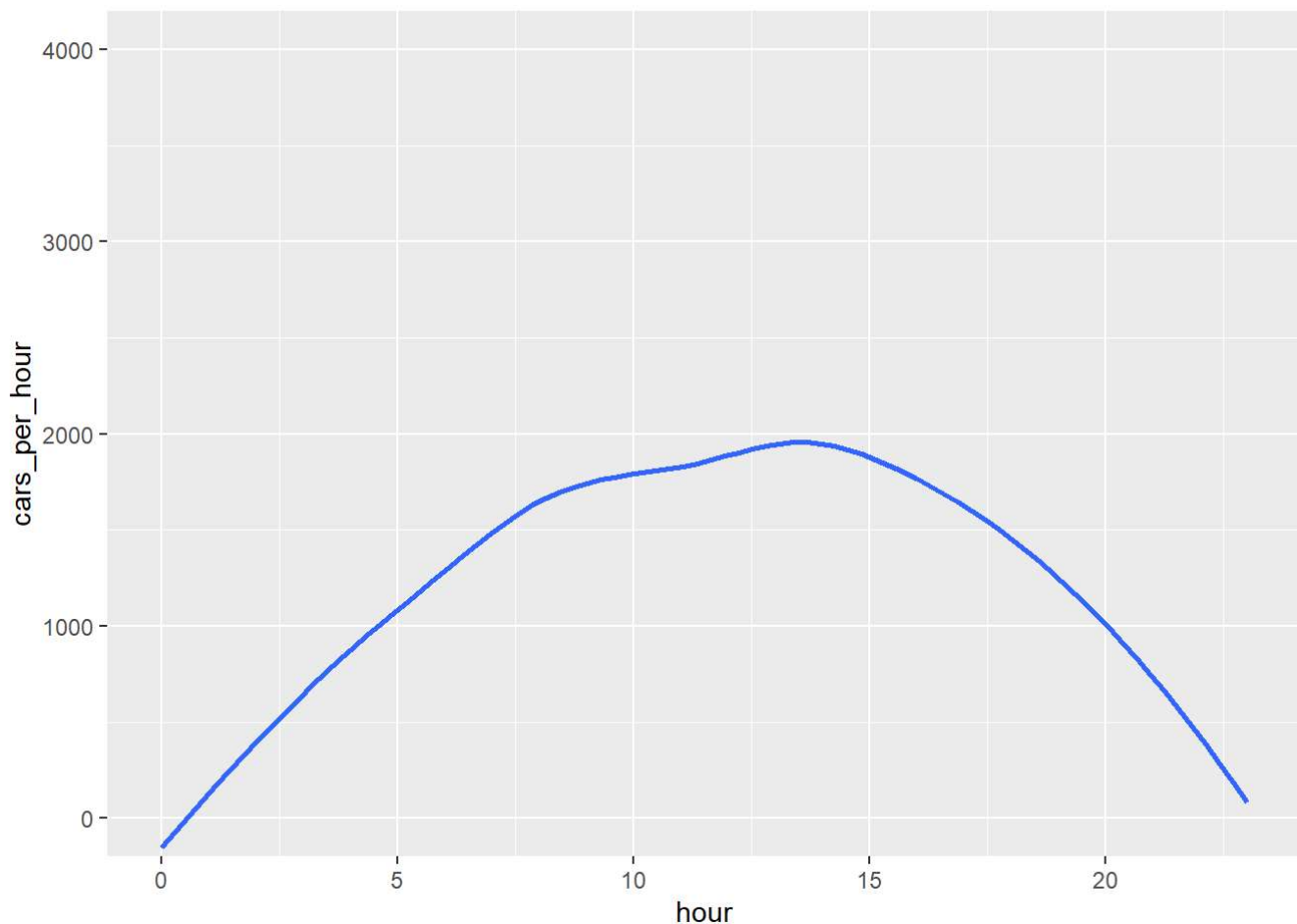
```
pike_master <- bind_rows(pike_file_list) %>%  
  mutate(daily_percent = cars_per_hour / Total * 100)
```

Before and After State of Emergency Issued in March at the Cochituate Location

```
# Before  
pike_master %>%  
  filter(day_of_week != "Saturday" & day_of_week != "Sunday",  
         location == "co",  
         full_date > ymd(20200301) & full_date < ymd(20200318)) %>%  
  ggplot(aes(x = hour, y = cars_per_hour)) +  
  geom_smooth(se = FALSE)
```



```
# After
pike_master %>%
  filter(day_of_week != "Saturday" & day_of_week != "Sunday",
         location == "co",
         full_date >= ymd(20200318) & full_date <= ymd(20200331)
  ) %>%
  ggplot(aes(x = hour, y = cars_per_hour)) +
  geom_smooth(se = FALSE) +
  coord_cartesian(ylim = c(0, 4000))
```



Conclusion:

The most interesting (and positive) finding of the project was seeing how people reacted to the coronavirus pandemic. Although Massachusetts was one of the top five states in terms of COVID cases early on, the state did a good job to combat that. Massachusetts was one of the states to issue a State of Emergency early to reduce in person contact when possible. It is good to see that more people have been working from home or coming into work on a reduced basis. This is illustrated by the average number of `cars_per_hour` changing dramatically within the month of March. Looking at other months, the second half of March and pretty much all of April had the lowest amount of traffic. This corresponds to when people took the pandemic seriously, but as people started to get complacent by traveling more frequently (months starting at May start to have more `cars_per_hour`), cases started to go back up again. COVID cases both in Massachusetts and throughout the country have started to go back up again, so it will be interesting to see if we go back to a shutdown as severe as March and April.