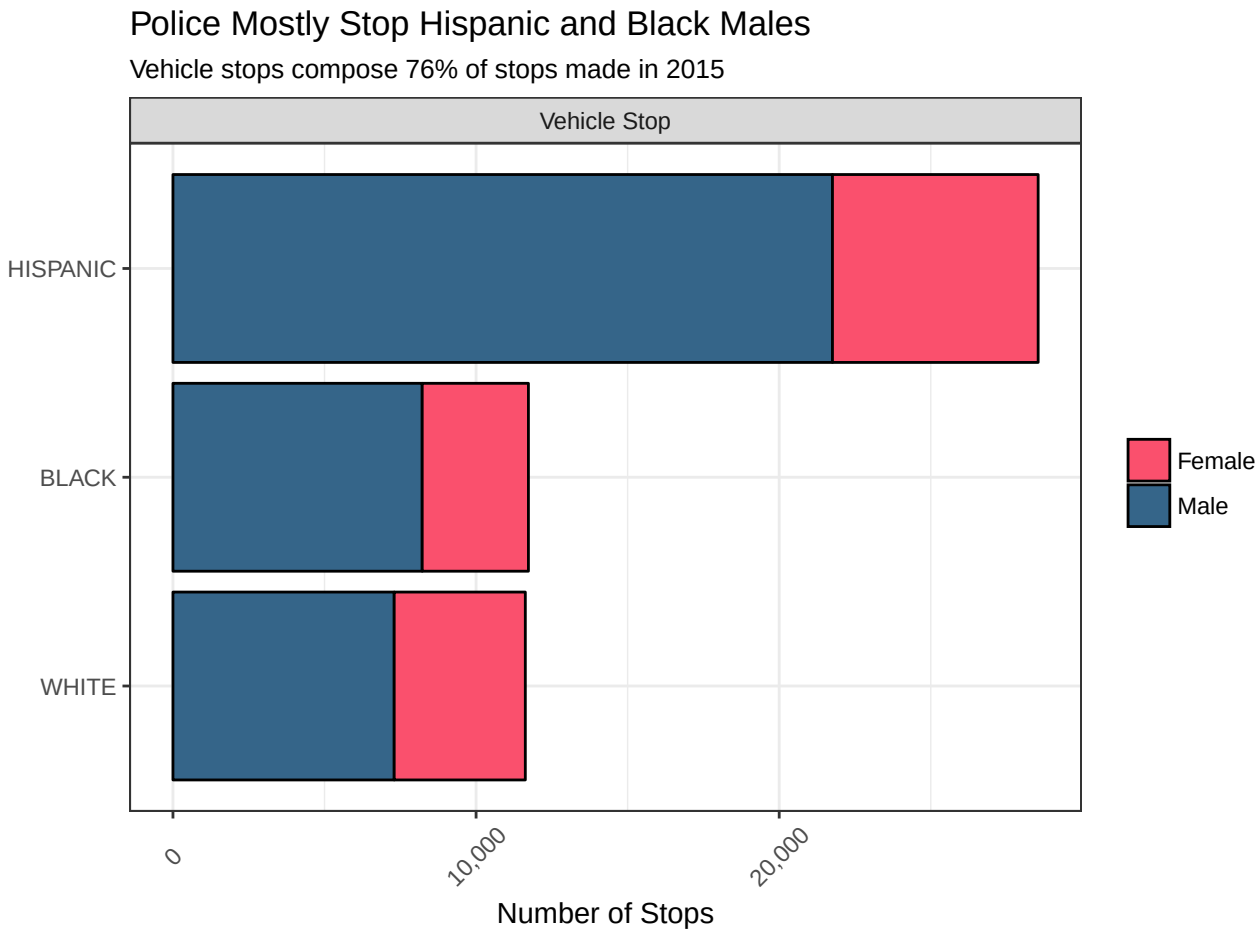
Stop Data Analysis 1

```
library(tidyverse)
library(scales)
library(DT)
library(forcats)
library(ggthemes)
library(lubridate)

source('../../R/helpers.R')

# thanks to notebooks default working directory >.<
stops <- read_rds('../../processed_data/prepared_stops.rds')
```

Exploratory Plots

All Stops Breakdown by Race, Sex, and Stop Type

```
rst_break <- stops %>%
  group_by(DESCENT_DESC, PERSN_GENDER_CD, STOP_TYPE) %>%
  summarize(num_stops = n()) %>%
  ungroup() %>%
  mutate(
    DESCENT_DESC = fct_reorder(DESCENT_DESC, num_stops)
  )

rst_break %>%
  ggplot(aes(DESCENT_DESC, num_stops, fill = PERSN_GENDER_CD)) +
  geom_bar(stat = 'identity', color = 'black') +
  facet_grid(~ STOP_TYPE) +
  theme_bw() +
  coord_flip() +
  xlab('') + ylab('Number of Stops') +
  scale_y_continuous(labels = comma) +
  theme(legend.title = element_blank()) +
  theme(axis.text.x = element_text(angle = 45, vjust = .5)) +
  scale_fill_manual(values = c('#FA506D', '#356589')) +
  ggtitle('Police Mostly Stop Hispanic and Black Males',
          subtitle = 'Vehicle stops compose 76% of stops made in 2015')
```

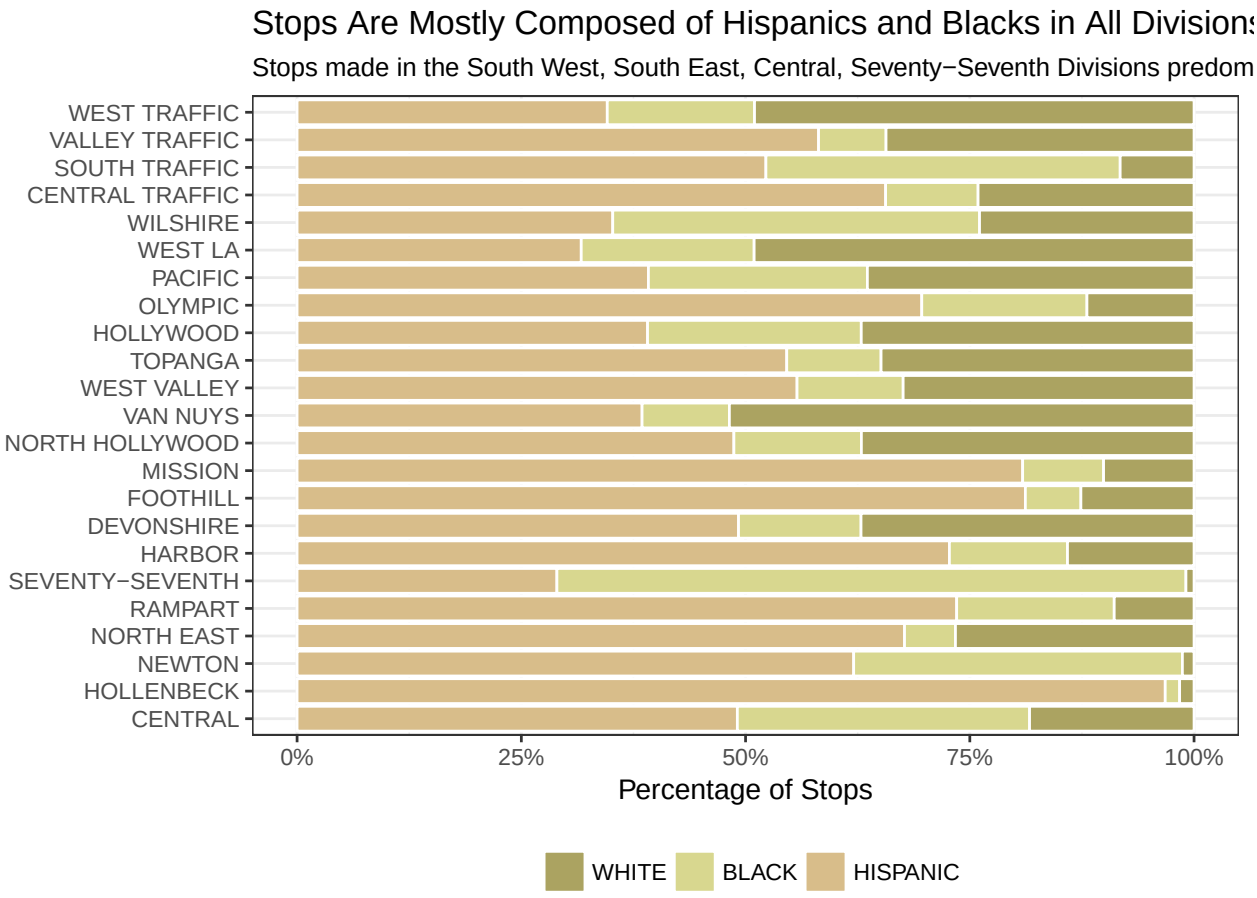


All Stop Breakdown by Selected Division

There are 52 divisions present in the data, but we'll look at the XXXX interesting ones.

```
# filter to divisions we will test later
area_divs <- c(
  'CENTRAL', 'HOLLENBECK', 'NEWTON', 'NORTH EAST', 'RAMPART',
  'SEVENTY-SEVENTH', 'HARBOR', 'SOUTHEAST', 'SOUTHWEST',
  'DEVONSHIRE', 'FOOTHILL', 'MISSION', 'NORTH HOLLYWOOD', 'VAN NUYS',
    'WEST VALLEY', 'TOPANGA',
  'HOLLYWOOD', 'OLYMPIC', 'PACIFIC', 'WEST LA', 'WILSHIRE'
)

traffic_divs <- c(
  'CENTRAL TRAFFIC', 'SOUTH TRAFFIC', 'VALLEY TRAFFIC', 'WEST TRAFFIC'
)

selected_divs <- as.factor(c(area_divs, traffic_divs))

stops <- stops %>%
  filter(DIV1_DESC %in% selected_divs) %>%
  mutate(DIV1_DESC = factor(DIV1_DESC, levels = selected_divs))
```

```r
div_break <- stops %>%
  group_by(DIV1_DESC, DESCENT_DESC) %>%
  summarize(n = n()) %>%
  mutate(percent = n / sum(n),
         DESCENT_DESC = fct_reorder(DESCENT_DESC, percent))

race_colors <- rev(c('#4F3130', '#753742', '#AA5042', '#D8BD8A', '#D8D78F', '#ABA361'))

div_break %>%
  ggplot(aes(DIV1_DESC, percent, fill = DESCENT_DESC)) +
  geom_bar(stat = 'identity', color = 'white') +
  theme_bw() +
  coord_flip() +
  scale_y_continuous(labels = percent) +
  xlab('') + ylab('Percentage of Stops') +
  theme(
    legend.title = element_blank(),
    legend.position = 'bottom'
  ) +
  ggtitle("Stops Are Mostly Composed of Hispanics and Blacks in All Divisions", subtitle = "Stops made i
  scale_fill_manual(
    values = race_colors
  )
```

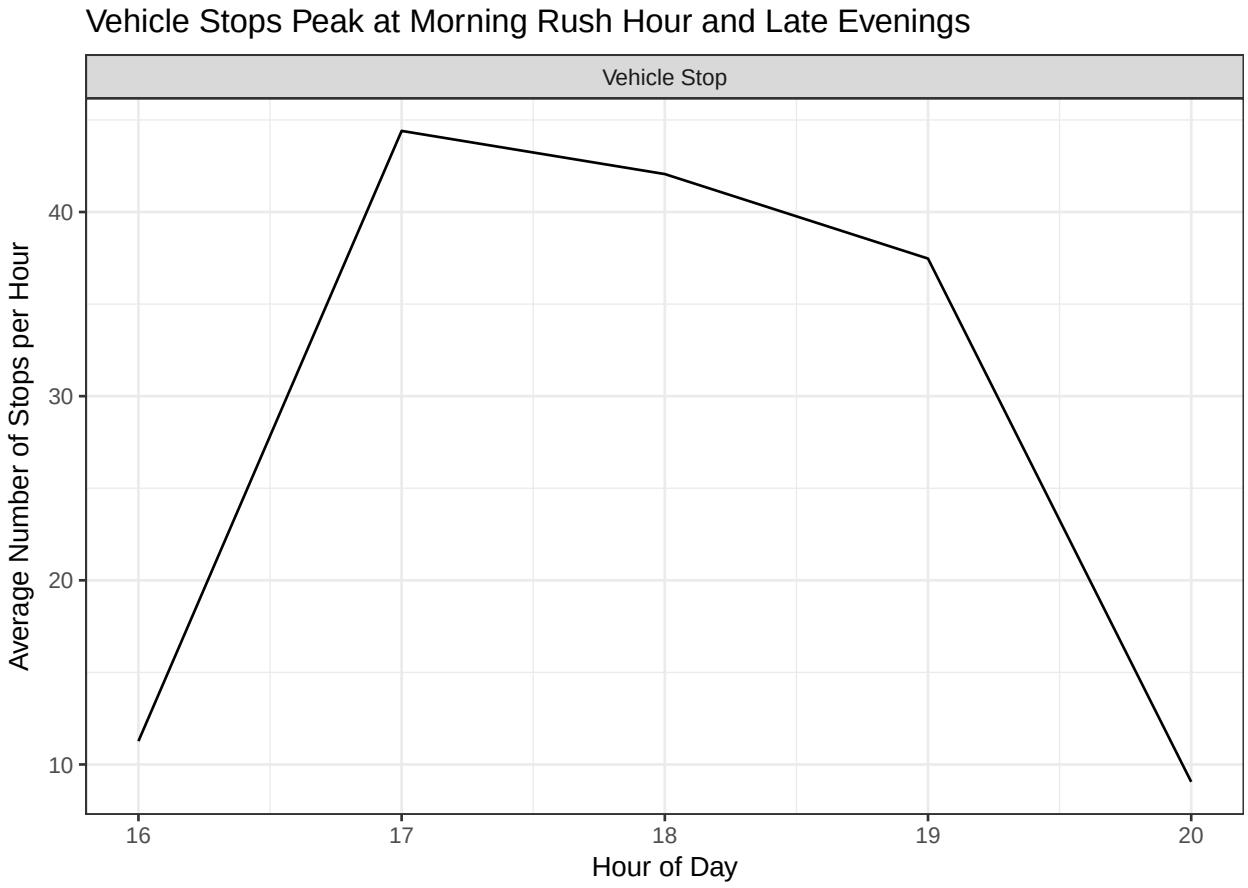


All Stops Over Time by Race and Division

```
# div_time_break <- stops %>%
#   group_by(DIV1_DESC, DESCENT_DESC, stop_month = month(STOP_DT, T)) %>%
#   summarize(num_stops = n())
#
# div_time_break %>%
#   ggplot(aes(stop_month, num_stops, color = DESCENT_DESC, group = 1)) +
#   geom_line() +
#   theme_bw() +
#   facet_grid(DIV1_DESC~ DESCENT_DESC, scales = 'free') +
#   theme(legend.position = 'bottom') +
#   scale_color_manual(values = race_colors) +
#   theme(axis.text.x = element_text(angle = 30)) +
#   labs(
#     y = "Number of Stops"
#   )
```

Noon Crime Peaks?

```
stop_hours <- stops %>%
  mutate(
    stop_hour = floor(as.numeric(STOP_TM) / 60 / 60)
  ) %>%
  group_by(stop_hour, STOP_TYPE) %>%
  summarize(num_stops_per_hour = n() / length(unique(STOP_DT)))

stop_hours %>%
  ggplot(aes(stop_hour, num_stops_per_hour)) +
  geom_line() +
  facet_grid(~ STOP_TYPE) +
  theme_bw() +
  labs(x = 'Hour of Day', y = 'Average Number of Stops per Hour',
       title = 'Vehicle Stops Peak at Morning Rush Hour and Late Evenings')
```


Vehicle Stops Peak at Morning Rush Hour and Late Evenings
Vehicle Stop
Average Number of Stops per Hour
40
30
20
10
16
17
18
19
20
Hour of Day