

Crimes in boston

SaiNagaChandra VivekGarimella

September 21, 2019

```
library(dplyr)

##
## Attaching package: 'dplyr'

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

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

library(ggplot2)
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.2.1

## v tibble  1.4.2      v purrr   0.2.5
## v tidyr   0.8.2      v stringr  1.3.1
## v readr   1.3.1      vforcats  0.3.0

## -- Conflicts ----- tidyverse_conflicts()
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()   masks stats::lag()

library(maps)

##
## Attaching package: 'maps'

## The following object is masked from 'package:purrr':
##
##     map

library(lubridate)

##
## Attaching package: 'lubridate'

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

crime1<-read.csv("C:/Users/garim/OneDrive/Desktop/NEU/Introduction to Data Management and Processing/data/crime.csv")
crime1<-drop_na(crime1)

new<-crime1 %>% group_by(OFFENSE_CODE_GROUP) %>% summarise(total=n()) %>% arrange(desc(total)) %>% top_n(10)

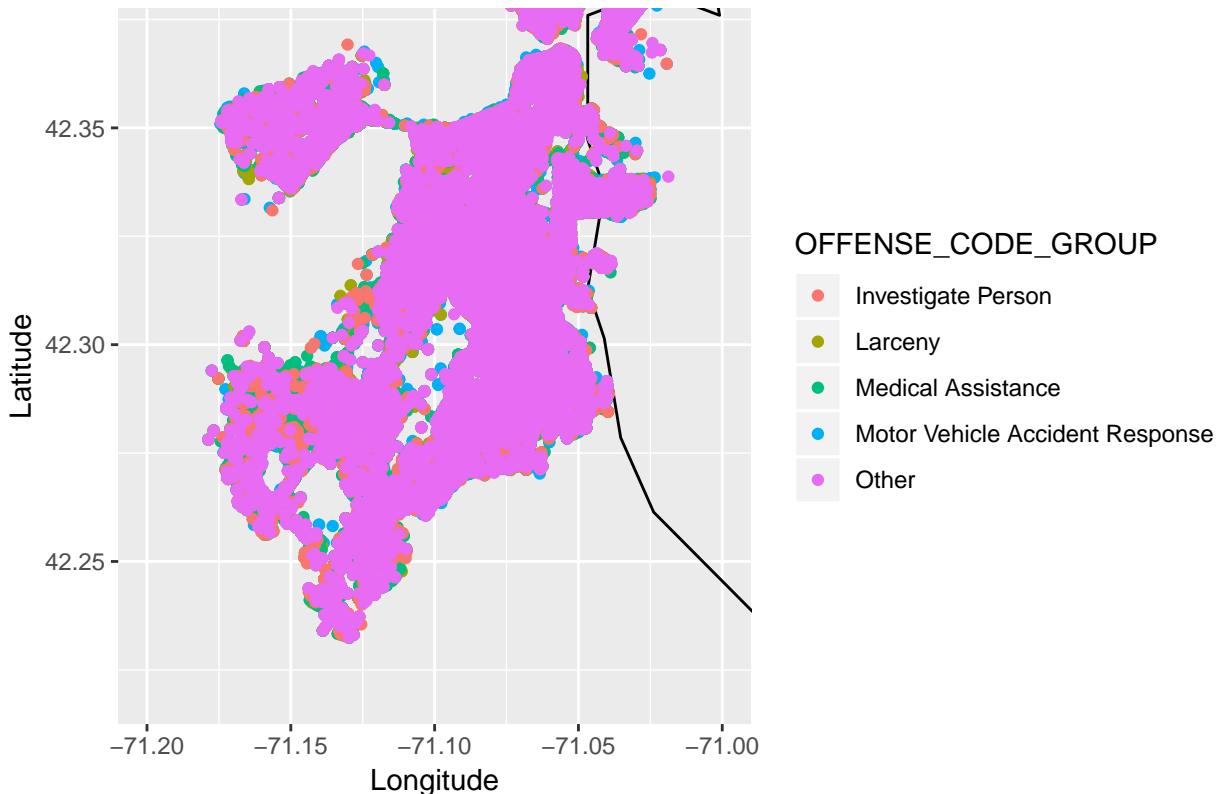
## Selecting by total
crime1<-left_join(new,crime1)

## Joining, by = "OFFENSE_CODE_GROUP"
```

```
#locations  
corners<-map_data('state',region='massachusetts')
```

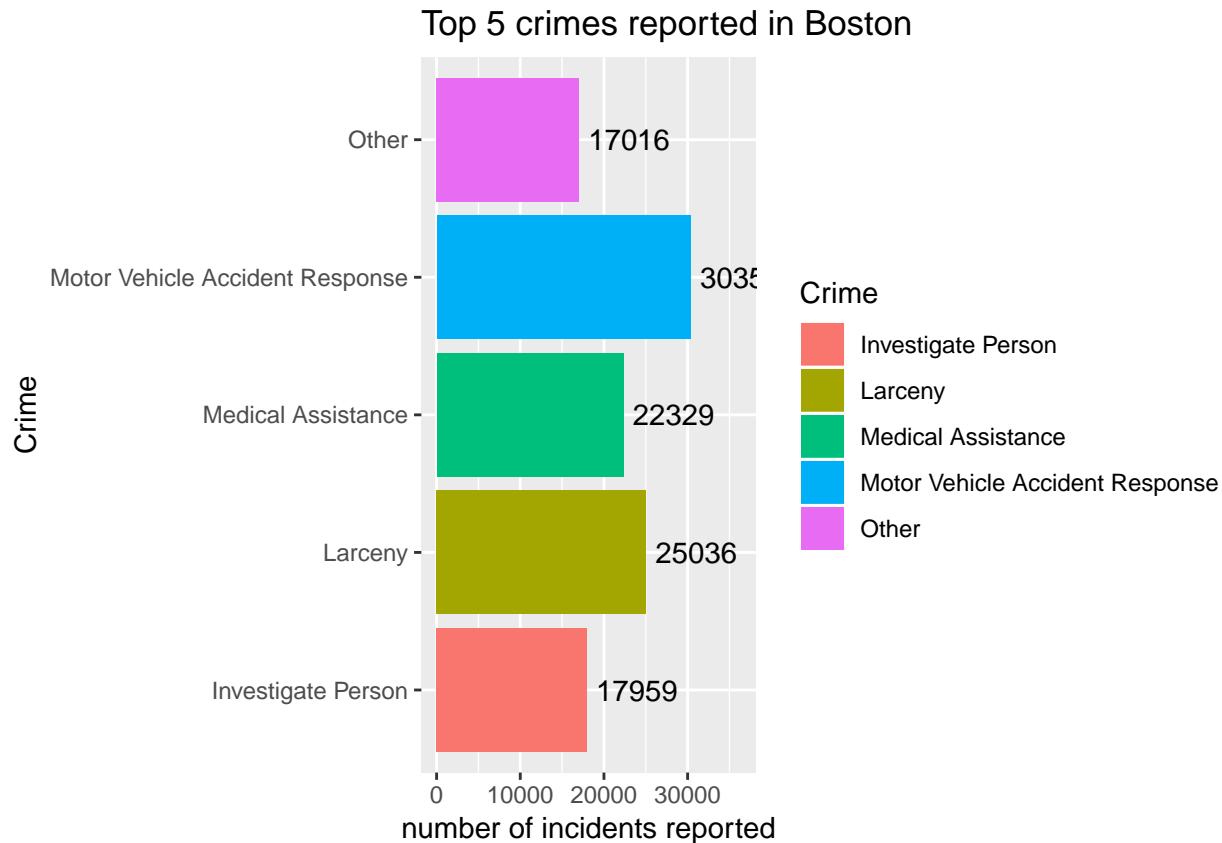
```
ggplot(corners)+geom_polygon(mapping = aes(x=long,y=lat),fill=NA,color="black")+ geom_point(data = crime)
```

Top 5 Crimes reported in Boston



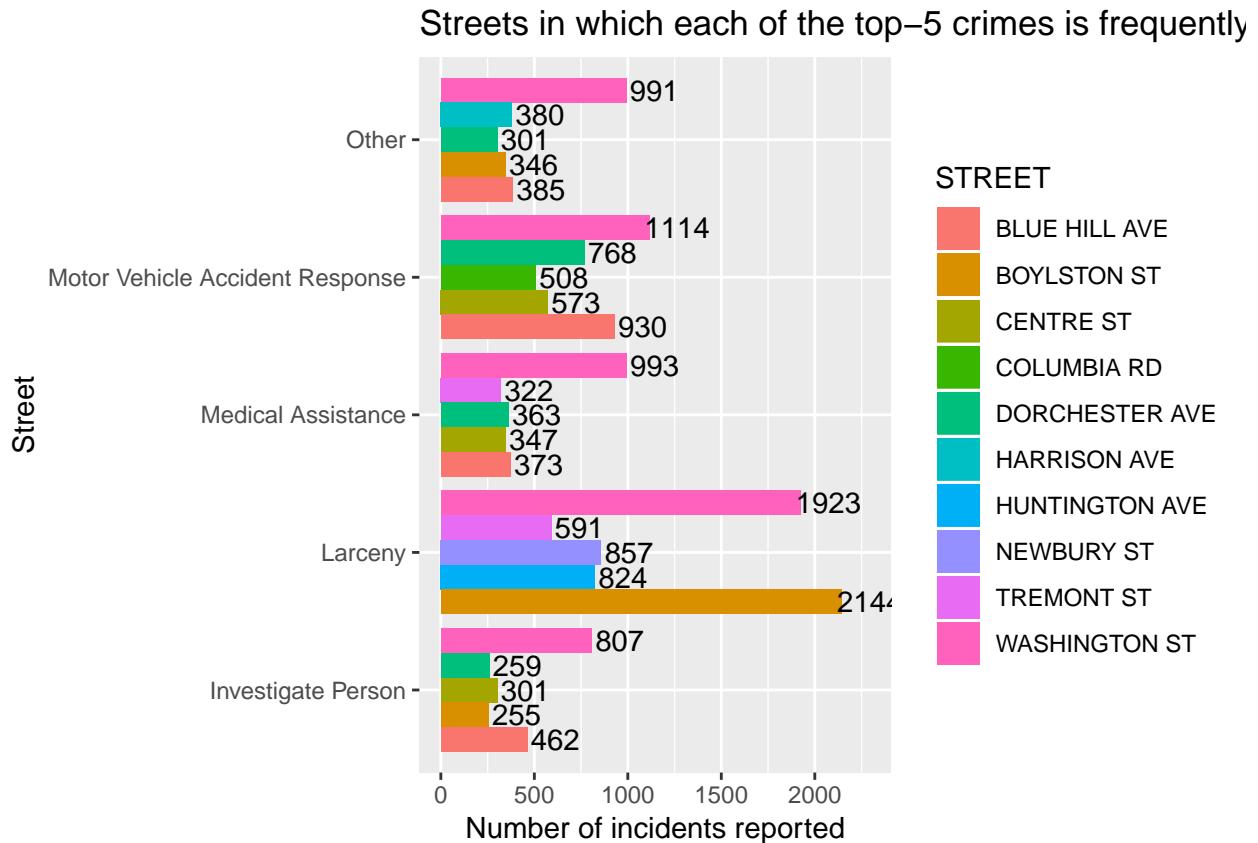
```
#OFFENSE_CODE_GROUP
```

```
new %>% ggplot(mapping = aes(x=OFFENSE_CODE_GROUP,y=total,fill=OFFENSE_CODE_GROUP))+geom_bar(stat = 'identity')
```



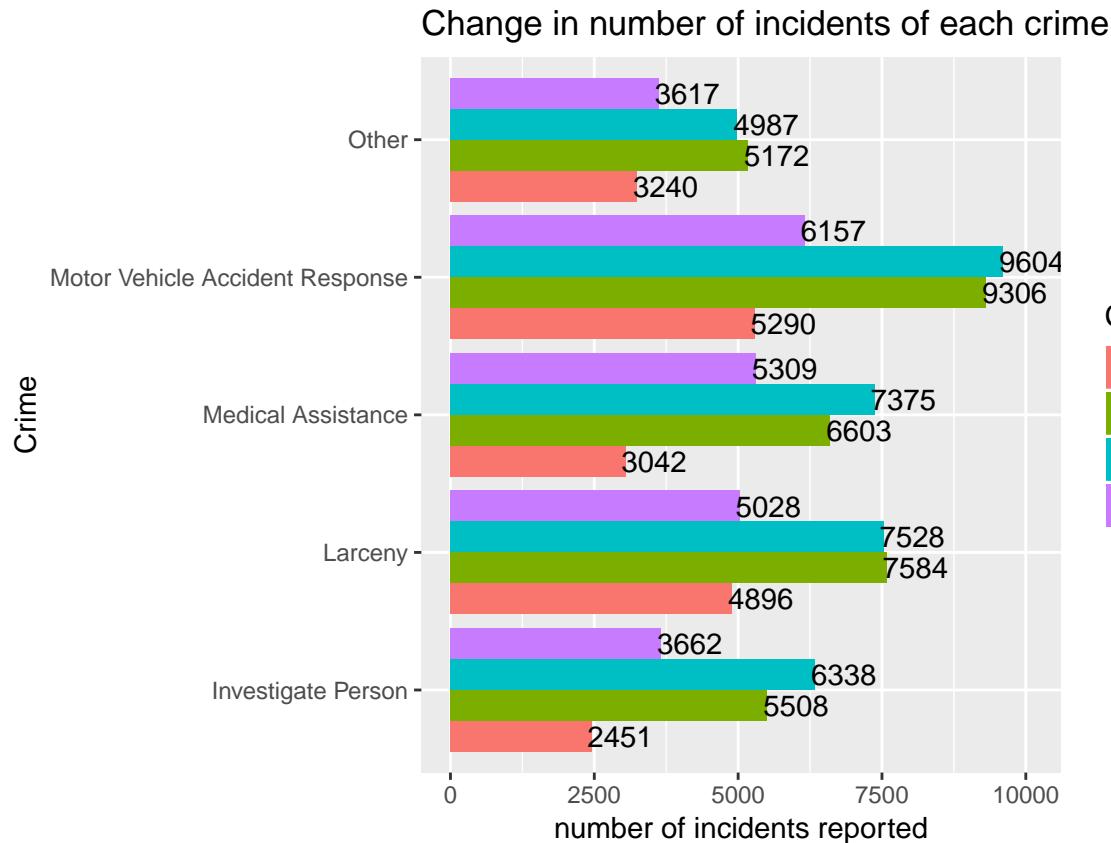
#CRIMES IN RESPECTIVE STREETS

```
first<-crime1%>% group_by(OFFENSE_CODE_GROUP,STREET) %>% summarise(crimes_in_respective_streets=n()) %>%  
## Selecting by crimes_in_respective_streets  
left_join(first,new) %>% ggplot(mapping = aes(x=OFFENSE_CODE_GROUP,y=crimes_in_respective_streets,fill=  
## Joining, by = "OFFENSE_CODE_GROUP"
```



```
#CRIMES IN RESPECTIVE YEAR
second<-crime1%>% group_by(OFFENSE_CODE_GROUP, YEAR) %>% summarise(crimes_in_respective_year=n()) %>% arrange(YEAR, OFFENSE_CODE_GROUP)

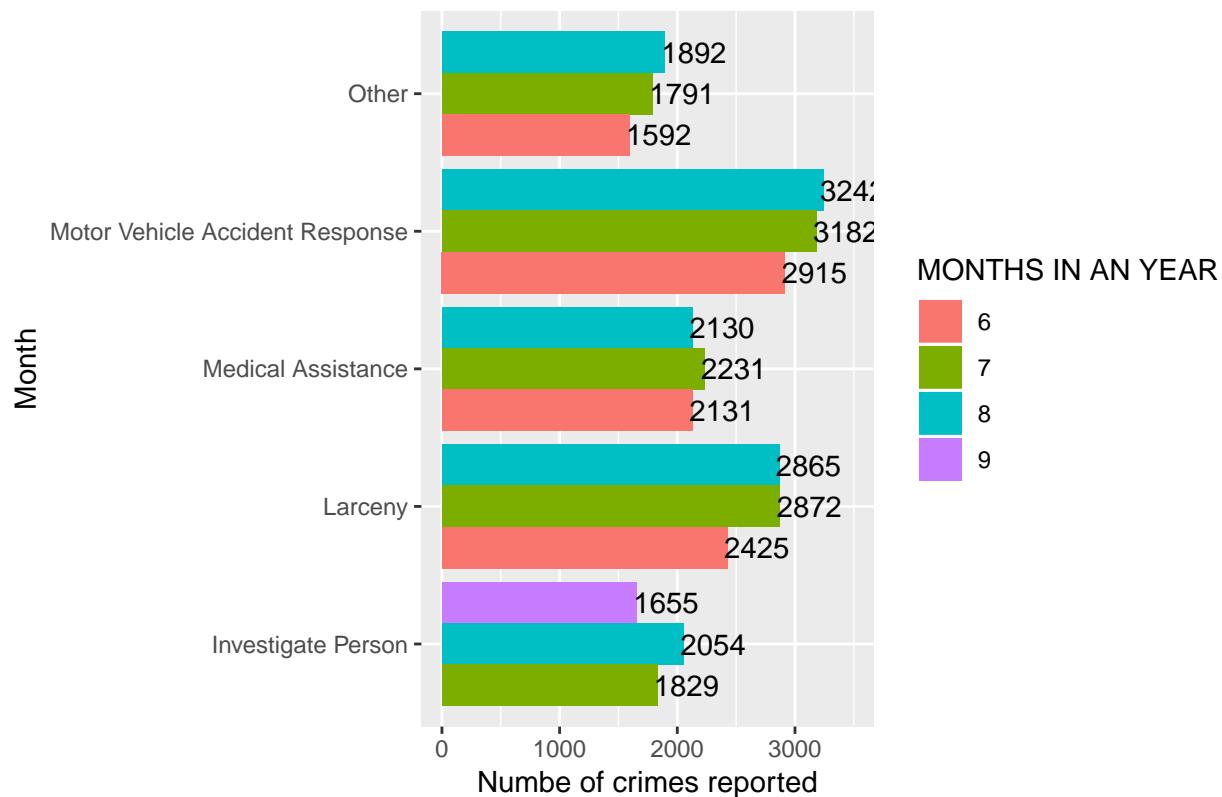
## Selecting by crimes_in_respective_year
left_join(second,new) %>% ggplot(mapping = aes(x=OFFENSE_CODE_GROUP,y=crimes_in_respective_year,fill=as.factor(OFFENSE_CODE_GROUP)))
## Joining, by = "OFFENSE_CODE_GROUP"
```



#CRIMES IN RESPECTIVE MONTH

```
third<-crime1%>% group_by(OFFENSE_CODE_GROUP,MONTH) %>% summarise(crimes_in_respective_month=n()) %>% a
## Selecting by crimes_in_respective_month
left_join(third,new) %>% ggplot(mapping = aes(x=OFFENSE_CODE_GROUP,y=crimes_in_respective_month,fill=as
## Joining, by = "OFFENSE_CODE_GROUP"
```

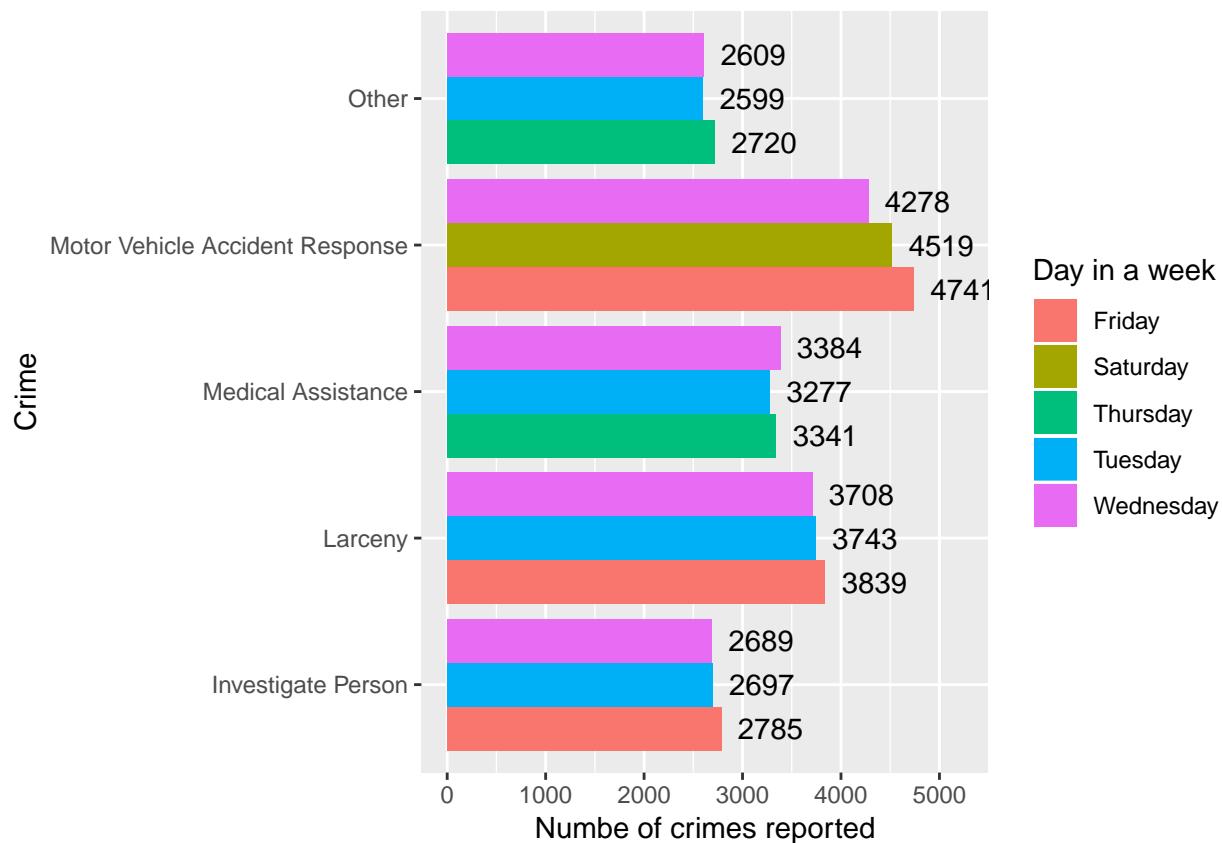
Months in which each of Top–5 crimes is frequently reported



```
#DAY OF WEEK
fourth<-crime1%>% group_by(OFFENSE_CODE_GROUP, DAY_OF_WEEK) %>% summarise(crimes_in_respective_day_of_week = sum(crimes))

## Selecting by crimes_in_respective_day_of_week
left_join(fourth,new) %>% ggplot(mapping = aes(x=OFFENSE_CODE_GROUP,y=crimes_in_respective_day_of_week,fill=DAY_OF_WEEK)) + geom_bar(stat="identity") + theme_minimal() + xlab("Offense Category") + ylab("Number of crimes reported") + scale_fill_discrete(name="MONTHS IN AN YEAR",labels=c("6","7","8","9"))

## Joining, by = "OFFENSE_CODE_GROUP"
```



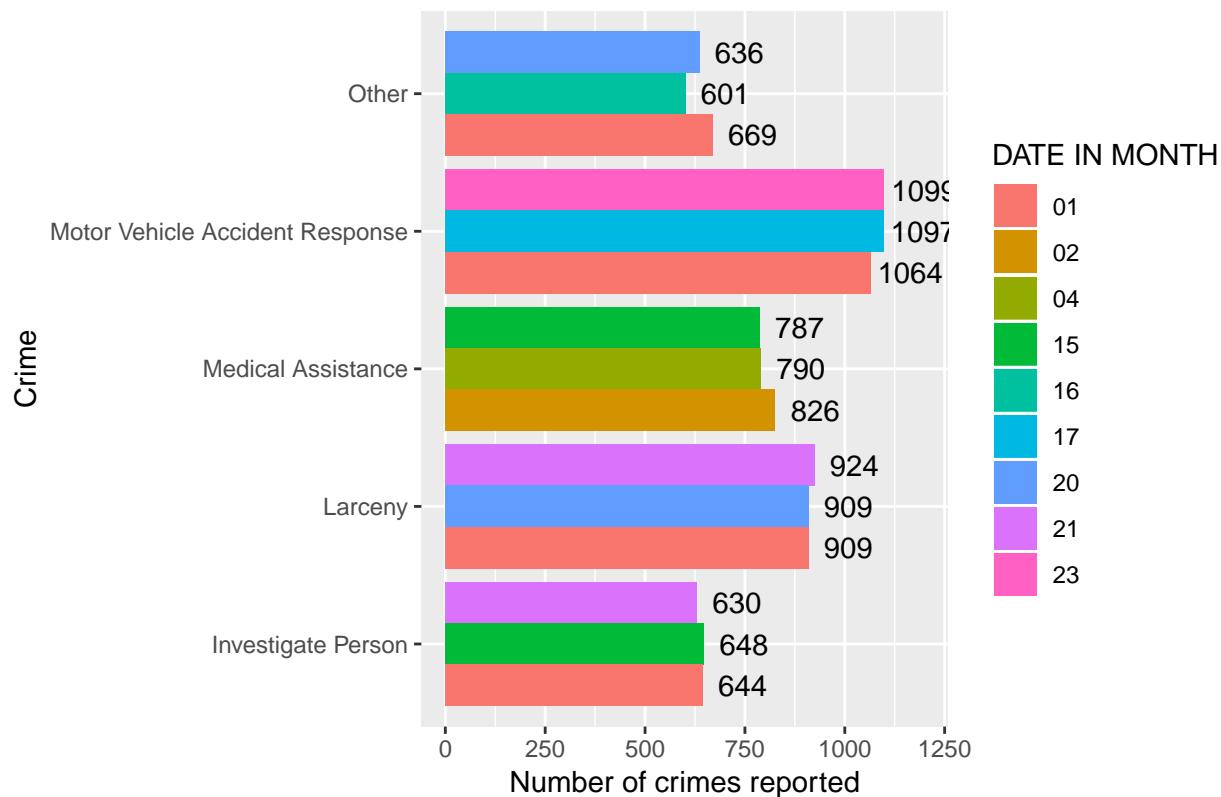
```
crime1<-crime1 %>% mutate(date=substring(crime1[['OCCURRED_ON_DATE']],9,10))

#DATE IN MONTH
fifth<-crime1%>% group_by(OFFENSE_CODE_GROUP,date) %>% summarise(crimes_in_respective_date_in_month=n())

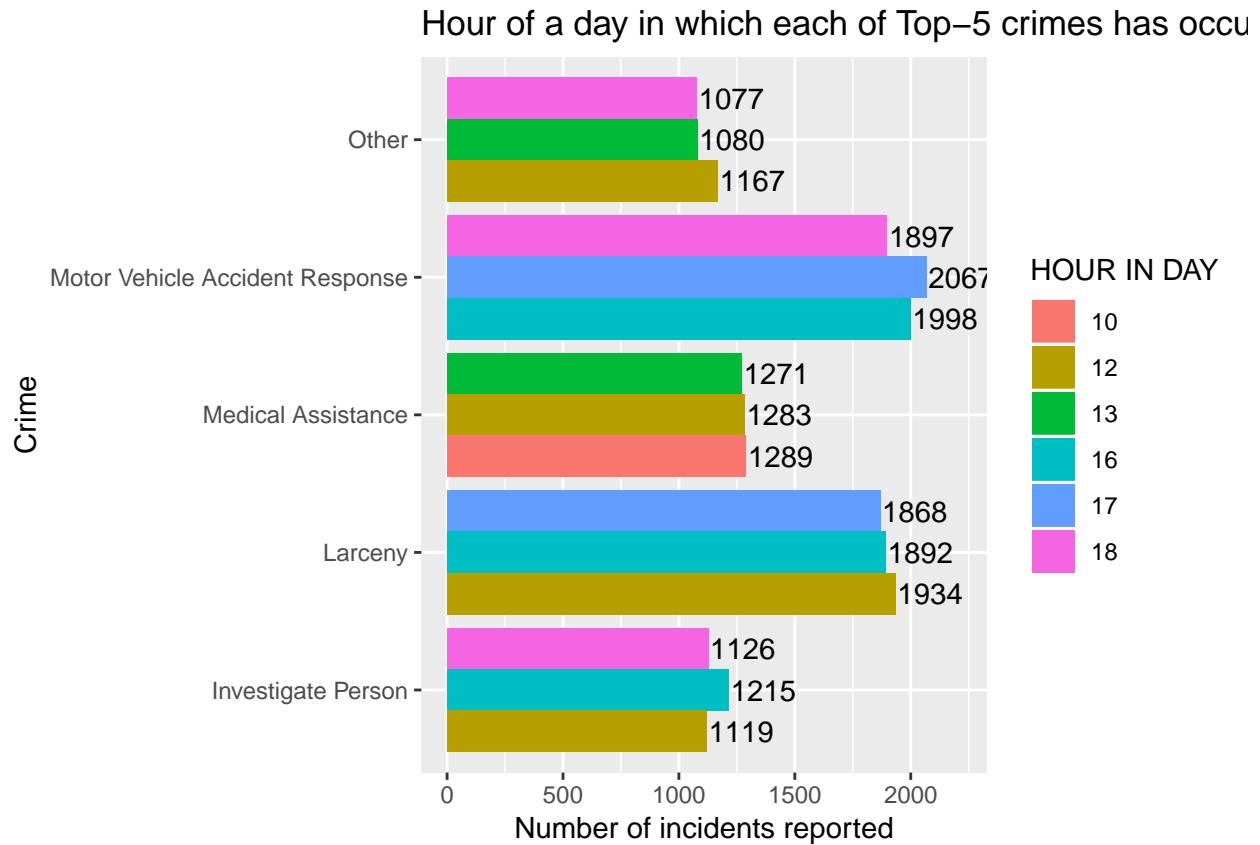
## Selecting by crimes_in_respective_date_in_month
left_join(fifth,new) %>% ggplot(mapping = aes(x=OFFENSE_CODE_GROUP,y=crimes_in_respective_date_in_month)

## Joining, by = "OFFENSE_CODE_GROUP"
```

Dates on which most number of Top–5 crimes reported



```
#HOUR IN DAY
sixth<-crime1%>% group_by(OFFENSE_CODE_GROUP,HOUR) %>% summarise(crimes_in_respective_hour_in_day=n())
## Selecting by crimes_in_respective_hour_in_day
left_join(sixth,new) %>% ggplot(mapping = aes(x=OFFENSE_CODE_GROUP,y=crimes_in_respective_hour_in_day,freq))+
## Joining, by = "OFFENSE_CODE_GROUP"
```

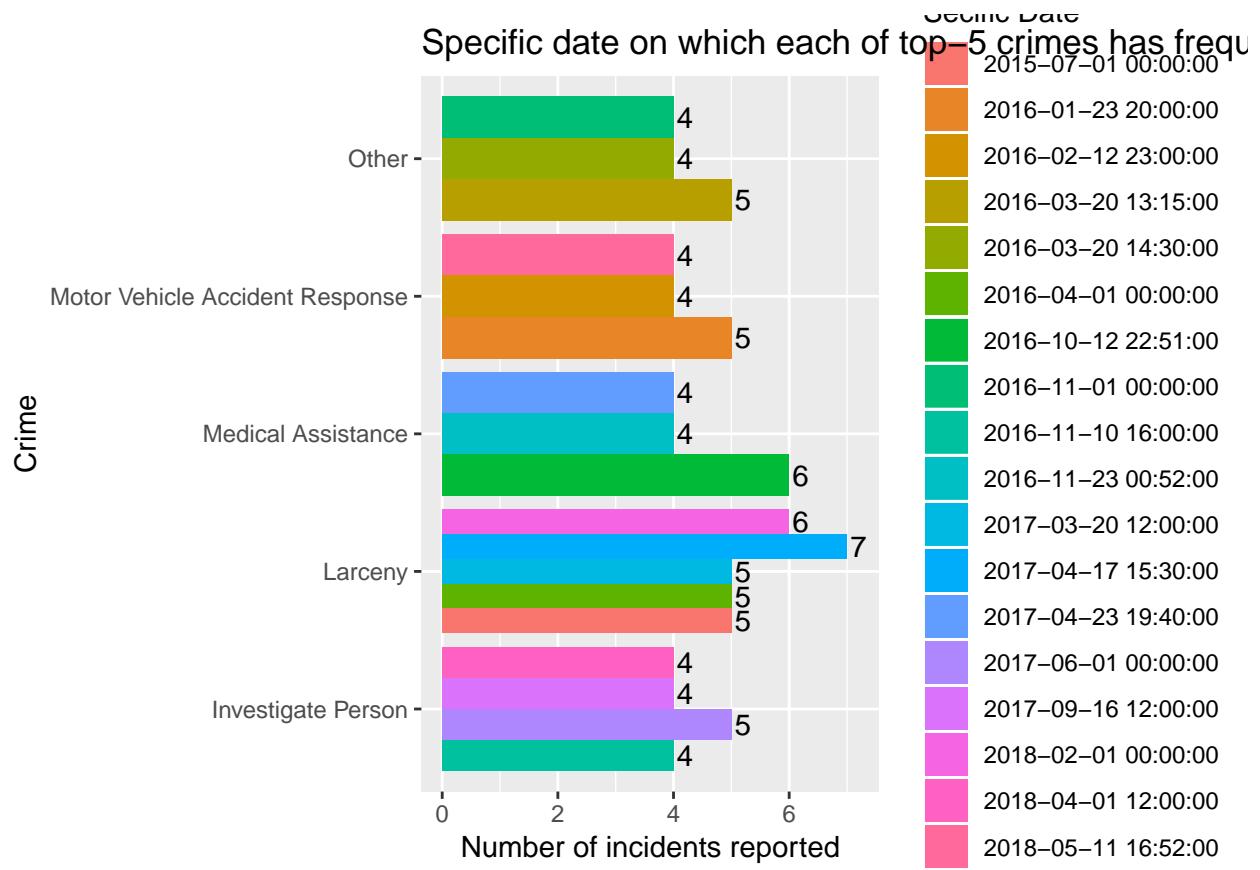


```
#ON PARTICULAR DATE

seventh<-crime1%>% group_by(OFFENSE_CODE_GROUP,OCCURRED_ON_DATE) %>% summarise(crimes_on_particular_day)

## Selecting by crimes_on_particular_day

left_join(seventh,new) %>% ggplot(mapping = aes(x=OFFENSE_CODE_GROUP,y=crimes_on_particular_day,fill=as
```



```
#REPORTING_AREA
```

```
eighth<-crime1%>% group_by(OFFENSE_CODE_GROUP,REPORTING_AREA) %>% summarise(crimes_reporting_area=n()) %>% left_join(eighth,new) %>% ggplot(mapping = aes(x=OFFENSE_CODE_GROUP,y=crimes_reporting_area,fill=as.factor(OFFENSE_CODE_GROUP))) + geom_bar(stat="identity") + theme_minimal() + theme(legend.title="Reporting Area", legend.position="top", legend.key.size=10)
```

Different UCR categories of crimes reported

