ap3650_nyc_crime_data_visualization

Anita

March 13, 2018

```
library(data.table)
library(vcdExtra)
library(extracat)
library(ggplot2)
library(dplyr)
library(tidyverse)
library(lubridate)

#fread("NYPD_Complaint_Data_Historic.csv",na.strings="",colClasses = c(PARKS_NM="c",HADEVELOPT="c"))->c
#fread("NYPD_Complaint_Data_Historic.csv",na.strings="")->crime_df
crime_df <- read.csv("NYPD_Complaint_Data_Historic.csv", header=TRUE)
```

Data Manipulations

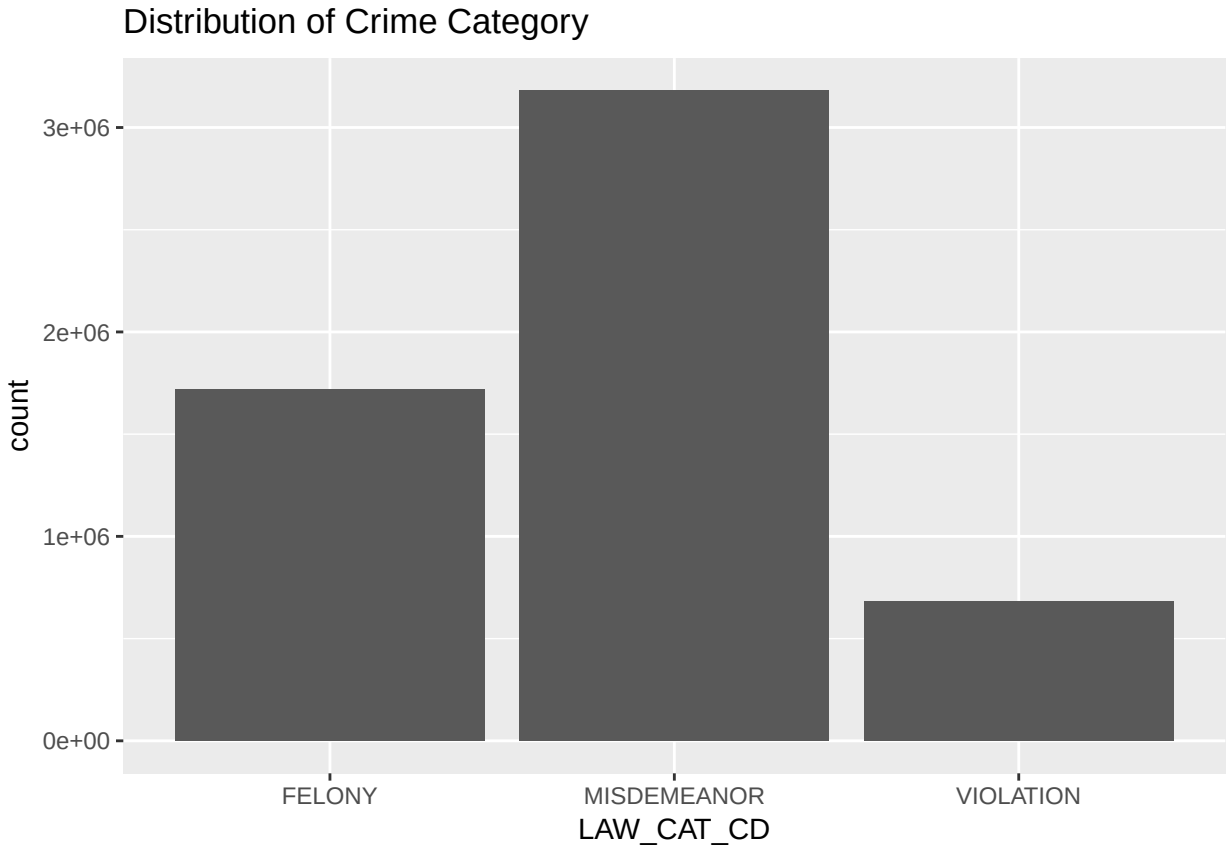
```
#Convert dates and times to correct format
crime_df$CMPLNT_FR_DT <- as.Date(crime_df$CMPLNT_FR_DT, format='%m/%d/%Y')
crime_df$CMPLNT_TO_DT <- as.Date(crime_df$CMPLNT_TO_DT, format='%m/%d/%Y')
crime_df$RPT_DT <- as.Date(crime_df$RPT_DT, format='%m/%d/%Y')

crime_df$CMPLNT_FR_TM <- as.POSIXct(crime_df$CMPLNT_FR_TM, format='%H:%M:%S')
crime_df$CMPLNT_TO_TM <- as.POSIXct(crime_df$CMPLNT_TO_TM, format='%H:%M:%S')
```

Plots

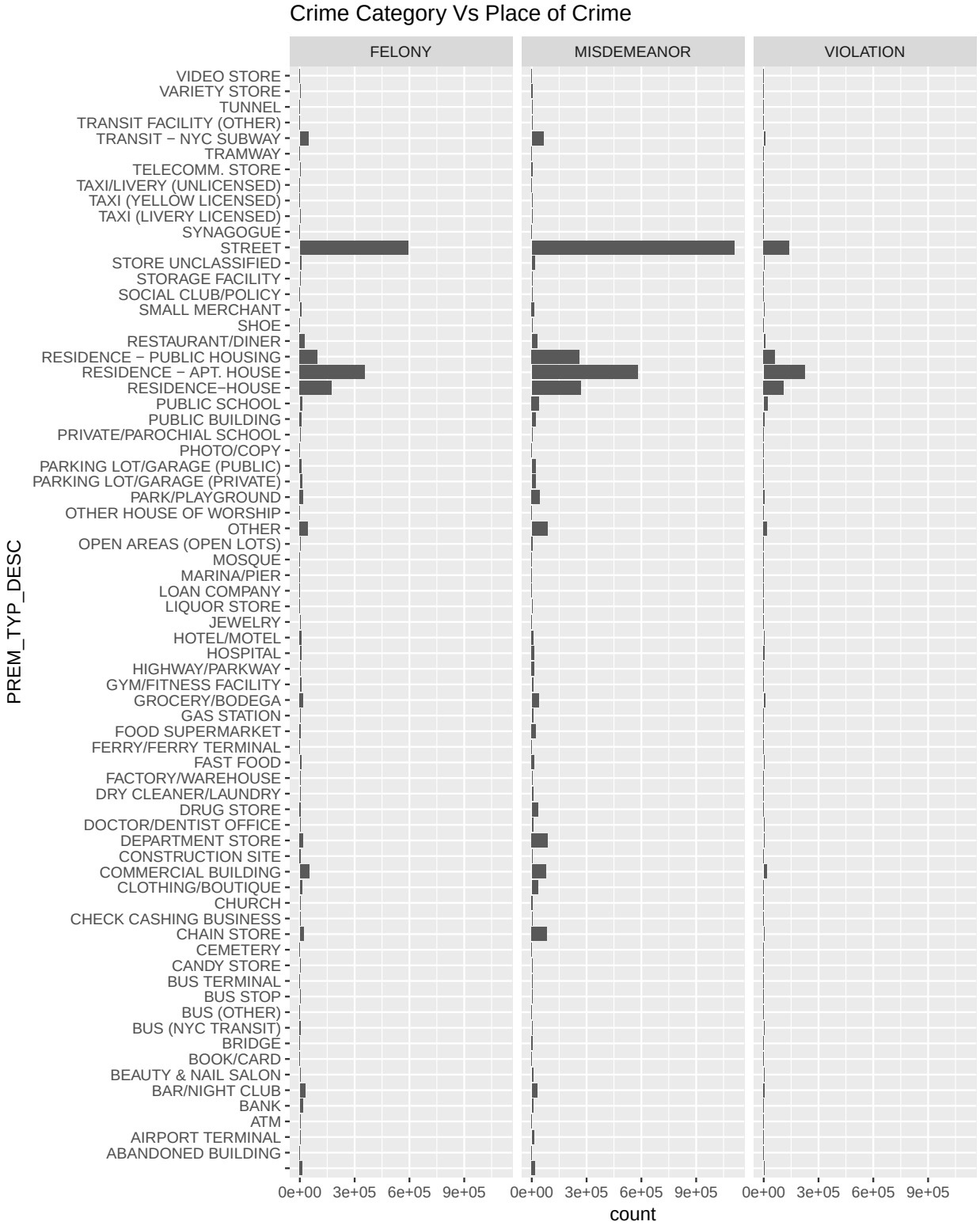
Warm-up Plot :-) Bar Chart

```
  ggplot(crime_df,aes(LAW_CAT_CD)) +
    geom_bar() +
    ggtitle("Distribution of Crime Category")
```



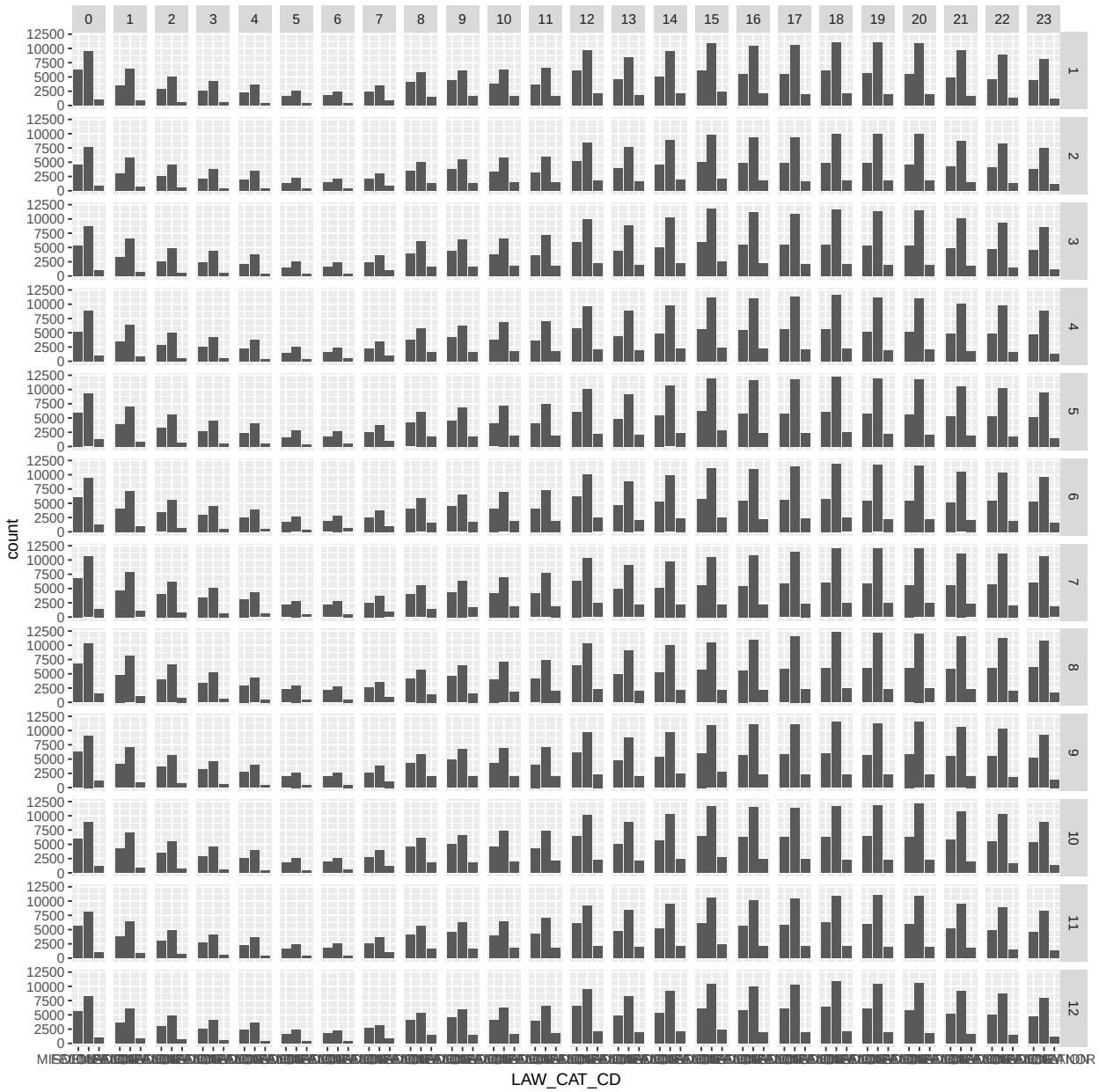
Type of Offense

```
ggplot(crime_df,aes(PREM_TYP_DESC)) +
  geom_bar() +
  facet_wrap(~LAW_CAT_CD) +
  coord_flip() +
  ggtitle("Crime Category Vs Place of Crime")
```


Crime Category Vs Place of Crime
FELONY
MISDEMEANOR
VIOLATION
PREM_TYP_DESC
VIDEO STORE
VARIETY STORE
TUNNEL
TRANSIT FACILITY (OTHER)
TRANSIT − NYC SUBWAY
TRAMWAY
TELECOMM. STORE
TAXI/LIVERY (UNLICENSED)
TAXI (YELLOW LICENSED)
TAXI (LIVERY LICENSED)
SYNAGOGUE
STREET
STORE UNCLASSIFIED
STORAGE FACILITY
SOCIAL CLUB/POLICY
SMALL MERCHANT
SHOE
RESTAURANT/DINER
RESIDENCE − PUBLIC HOUSING
RESIDENCE − APT. HOUSE
RESIDENCE−HOUSE
PUBLIC SCHOOL
PUBLIC BUILDING
PRIVATE/PAROCHIAL SCHOOL
PHOTO/COPY
PARKING LOT/GARAGE (PUBLIC)
PARKING LOT/GARAGE (PRIVATE)
PARK/PLAYGROUND
OTHER HOUSE OF WORSHIP
OTHER
OPEN AREAS (OPEN LOTS)
MOSQUE
MARINA/PIER
LOAN COMPANY
LIQUOR STORE
JEWELRY
HOTEL/MOTEL
HOSPITAL
HIGHWAY/PARKWAY
GYM/FITNESS FACILITY
GROCERY/BODEGA
GAS STATION
FOOD SUPERMARKET
FERRY/FERRY TERMINAL
FAST FOOD
FACTORY/WAREHOUSE
DRY CLEANER/LAUNDRY
DRUG STORE
DOCTOR/DENTIST OFFICE
DEPARTMENT STORE
CONSTRUCTION SITE
COMMERCIAL BUILDING
CLOTHING/BOUTIQUE
CHURCH
CHECK CASHING BUSINESS
CHAIN STORE
CEMETERY
CANDY STORE
BUS TERMINAL
BUS STOP
BUS (OTHER)
BUS (NYC TRANSIT)
BRIDGE
BOOK/CARD
BEAUTY & NAIL SALON
BAR/NIGHT CLUB
BANK
ATM
AIRPORT TERMINAL
ABANDONED BUILDING
0e+00
3e+05
6e+05
9e+05
count

Month and Time and Type of Crime

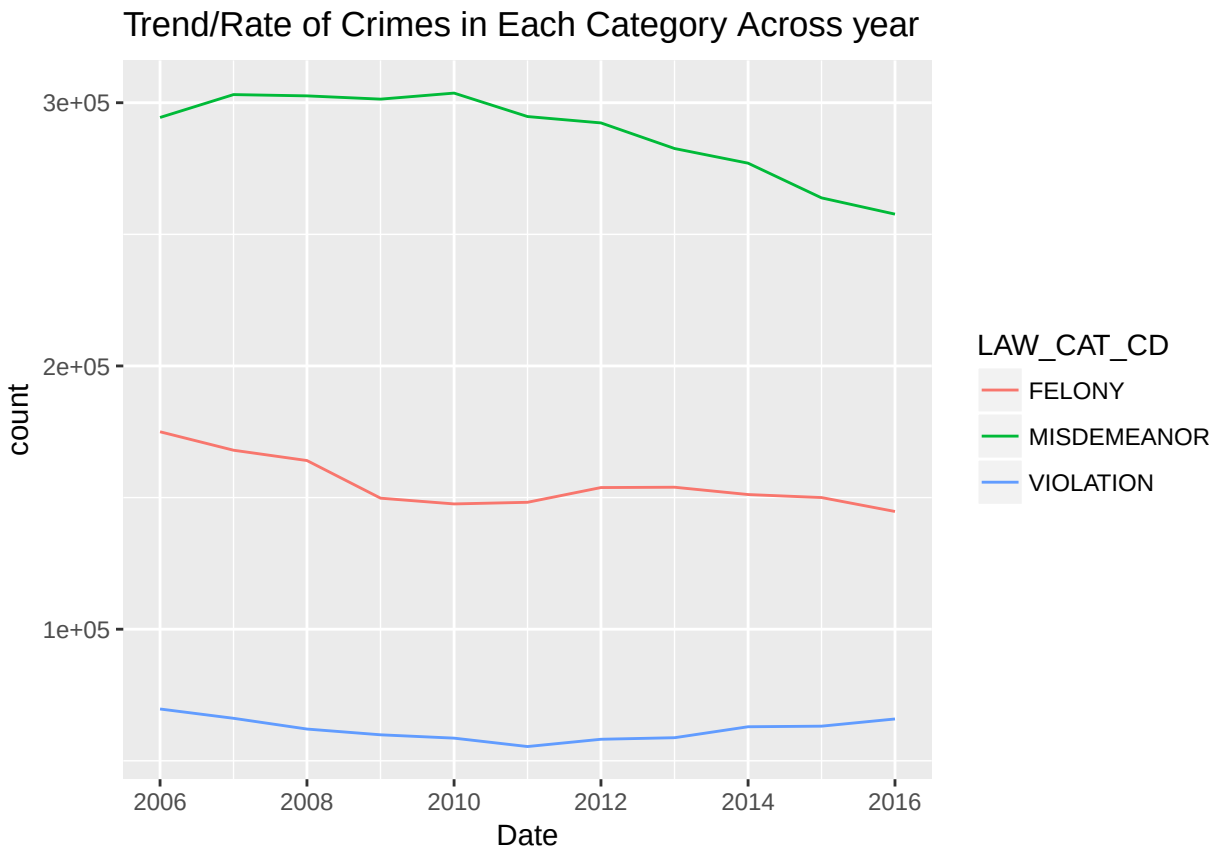
```
crime_df <- crime_df %>% drop_na()
ggplot(crime_df, aes(LAW_CAT_CD)) +
geom_bar() +
#facet_wrap(~month(CMPLNT_FR_DT))
#facet_wrap(~hour(CMPLNT_FR_TM))
facet_grid(month(CMPLNT_FR_DT)~hour(CMPLNT_FR_TM))
```



Time Series - Trend of Crime Rate

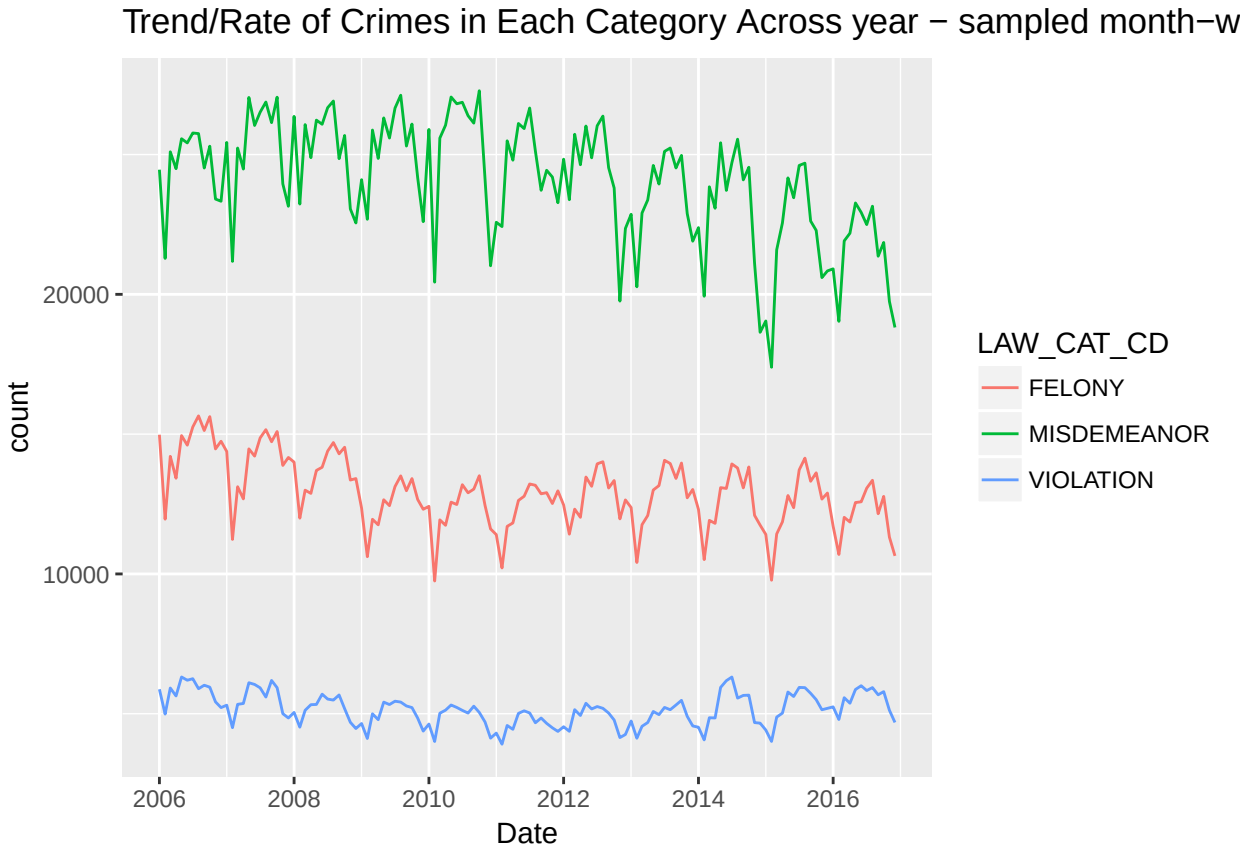
```
crime_time <- crime_df %>%
                filter(year(CMPLNT_FR_DT)>2005) %>%
                group_by(Date=floor_date(CMPLNT_FR_DT, "year"),LAW_CAT_CD) %>%
                summarize(count=n())

ggplot(crime_time, aes(Date,count, color=LAW_CAT_CD))+
      geom_line() +
      ggtitle("Trend/Rate of Crimes in Each Category Across year")
```




```
crime_time <- crime_df %>%
                filter(year(CMPLNT_FR_DT)>2005) %>%
                group_by(Date=floor_date(CMPLNT_FR_DT, "month"),LAW_CAT_CD) %>%
                summarize(count=n())

ggplot(crime_time, aes(Date,count, color=LAW_CAT_CD))+
      geom_line() +
      ggtitle("Trend/Rate of Crimes in Each Category Across year - sampled month-wise")
```



* Shows monthly pattern similar to Jingbo's * Year pattern fluctuates "` ## length of Crime Vs Type of Crime

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.