# SESSION 12 – ASSIGNMENT 12.1
# GENERALIZED LINEAR MODELS

**Date:** 11th Feb 2019

**Use the given link below:**
https://archive.ics.uci.edu/ml/machine-learning-databases/communities

**Perform the below operations:**

a. **Visualize the correlation between all variable in a meaningful way, clear representation of correlations. Find out top 3 reasons for having more crime in a city.**

```
library(readr)
COBRA_YTD2017 <- read_csv("G:/DATA ANALYTICS/DATA/crime-in-atlanta-2017/COBRA-YTD2017.csv")
library(data.table)
library(Amelia)
library(Rcpp)
data<-COBRA_YTD2017
data[4:10,3] <- rep(NA,7)
data[1:5,4] <- NA
data <- data[-c(5,6)]
summary(data)
```

```
> library(readr)
> COBRA_YTD2017 <- read_csv("G:/DATA ANALYTICS/DATA/crime-in-atlanta-2017/COBRA-YTD2017.csv")
Parsed with column specification:
cols(
  .default = col_character(),
  MI_PRINX = col_double(),
  offense_id = col_double(),
  occur_time = col_time(format = ""),
  poss_time = col_time(format = ""),
  beat = col_double(),
  dispo_code = col_double(),
  MaxOfnum_victims = col_double(),
  loc_type = col_double(),
  x = col_double(),
  y = col_double()
)
See spec(...) for full column specifications.
Warning: 9 parsing failures.
  row        col expected actual                                                        file
 3239 dispo_code a double    COS 'G:/DATA ANALYTICS/DATA/crime-in-atlanta-2017/COBRA-YTD2017.csv'
 7945 dispo_code a double    ADM 'G:/DATA ANALYTICS/DATA/crime-in-atlanta-2017/COBRA-YTD2017.csv'
 8527 dispo_code a double    ADM 'G:/DATA ANALYTICS/DATA/crime-in-atlanta-2017/COBRA-YTD2017.csv'
10145 dispo_code a double    ADM 'G:/DATA ANALYTICS/DATA/crime-in-atlanta-2017/COBRA-YTD2017.csv'
11912 dispo_code a double    ADM 'G:/DATA ANALYTICS/DATA/crime-in-atlanta-2017/COBRA-YTD2017.csv'
..... .......... ........ ...... ..........................................................................
See problems(...) for more details.

> library(data.table)
> library(Amelia)
> library(Rcpp)
> data<-COBRA_YTD2017
> data[4:10,3] <- rep(NA,7)
> data[1:5,4] <- NA
> data <- data[-c(5,6)]
```

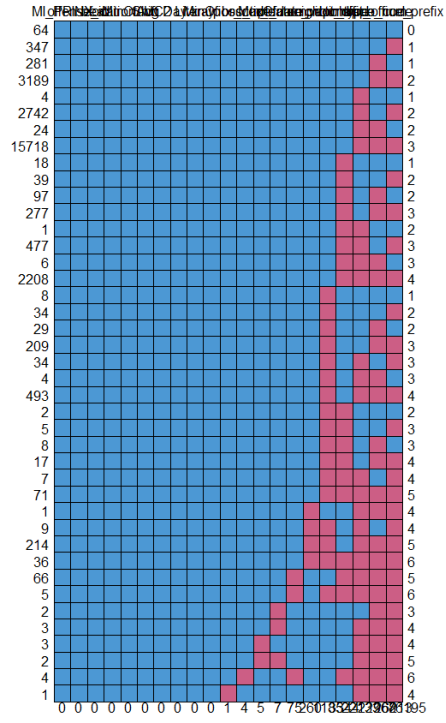
```
> summary(data)
   MI_PRINX        offense_id        rpt_date          occur_date         poss_time            beat       apt_office_prefix  apt_office_num
Min.   :8838438   Min.   :1.608e+08   Length:26759       Length:26759       Length:26759       Min.   :101.0   Length:26759       Length:26759
1st Qu.:8904204   1st Qu.:1.711e+08   Class :character   Class :character   Class1:hms         1st Qu.:208.0   Class :character   Class :character
Median :8910894   Median :1.720e+08   Mode  :character   Mode  :character   Class2:difftime    Median :312.0   Mode  :character   Mode  :character
Mean   :8910851   Mean   :6.523e+08                                         Mode  :numeric     Mean   :355.6
3rd Qu.:8917584   3rd Qu.:1.728e+08                                                            3rd Qu.:505.0
Max.   :8924410   Max.   :1.735e+11                                                            Max.   :710.0

   location           MinOfucr         MinOfibr_code        dispo_code     MaxOfnum_victims     Shift              Avg Day             loc_type
Length:26759       Length:26759       Length:26759       Min.   :10.00   Min.   : 0.00    Length:26759       Length:26759       Min.   : 1.00
Class :character   Class :character   Class :character   1st Qu.:10.00   1st Qu.: 1.00    Class :character   Class :character   1st Qu.:13.00
Mode  :character   Mode  :character   Mode  :character   Median :10.00   Median : 1.00    Mode  :character   Mode  :character   Median :18.00
                                                         Mean   :13.32   Mean   : 1.16                                           Mean   :20.76
                                                         3rd Qu.:10.00   3rd Qu.: 1.00                                           3rd Qu.:20.00
                                                         Max.   :60.00   Max.   :27.00                                           Max.   :99.00
                                                         NA's   :22968   NA's   :75                                              NA's   :3344
 UC2 Literal        neighborhood           npu                 x                y
Length:26759       Length:26759       Length:26759       Min.   :-84.55   Min.   : 0.00
Class :character   Class :character   Class :character   1st Qu.:-84.43   1st Qu.:33.73
Mode  :character   Mode  :character   Mode  :character   Median :-84.40   Median :33.76
                                                         Mean   :-83.69   Mean   :33.47
                                                         3rd Qu.:-84.37   3rd Qu.:33.79
                                                         Max.   :  0.00   Max.   :33.88
```

```
pMiss <- function(x){sum(is.na(x))/length(x)*100}
apply(data,2,pMiss)
apply(data,1,pMiss)
```

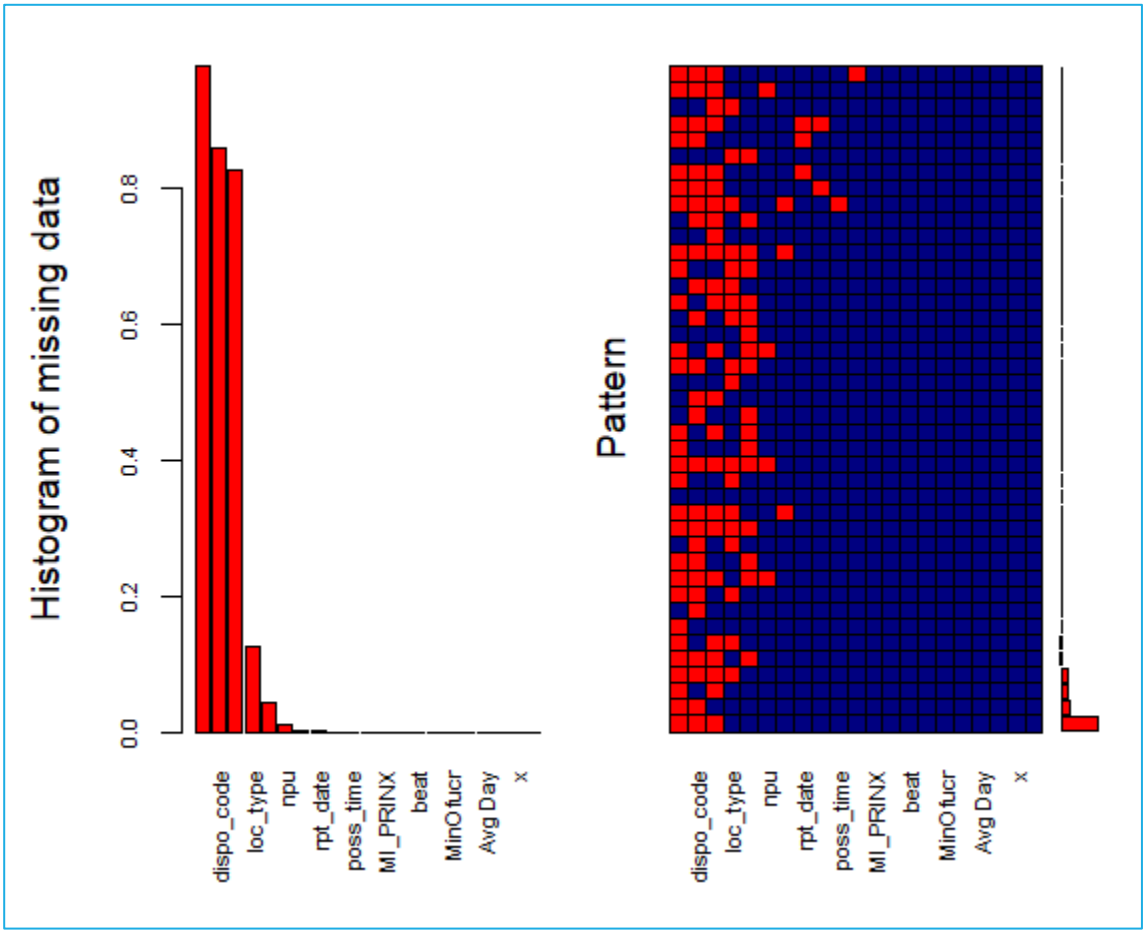
```
> pMiss <- function(x){sum(is.na(x))/length(x)*100}
> apply(data,2,pMiss)
        MI_PRINX        offense_id          rpt_date        occur_date         poss_time              beat apt_office_prefix    apt_office_num
      0.00000000        0.00000000        0.02615942        0.01868530        0.01494824        0.00000000        97.95956501        82.71235846
        location          MinOfucr     MinOfibr_code        dispo_code  MaxOfnum_victims             Shift           Avg Day           loc_type
      0.00000000        0.00000000        0.00373706       85.83280392        0.28027953        0.00000000        0.00000000       12.49673007
      UC2 Literal      neighborhood               npu                 x                 y
      0.00000000        4.42841661        0.97163571        0.00000000        0.00000000
> apply(data,1,pMiss)
  [1] 19.047619 19.047619 19.047619 23.809524 23.809524 19.047619 14.285714 19.047619 19.047619 14.285714 14.285714 19.047619 14.285714  9.523810 14.285714
 [16]  4.761905 14.285714 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714 14.285714 14.285714 14.285714
 [31]  9.523810 14.285714 14.285714  9.523810 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714 19.047619 14.285714
 [46]  4.761905  4.761905  9.523810  9.523810  9.523810 14.285714 14.285714 14.285714 14.285714 19.047619 14.285714 14.285714 14.285714 14.285714 14.285714
 [61] 14.285714 14.285714 14.285714  9.523810  9.523810 14.285714 23.809524 19.047619 19.047619 14.285714 19.047619  9.523810 14.285714 14.285714 14.285714
 [76] 14.285714 14.285714 19.047619 14.285714 14.285714  9.523810 14.285714 14.285714 14.285714 14.285714 19.047619 14.285714 19.047619 14.285714 14.285714
 [91]  9.523810 19.047619 14.285714 14.285714 14.285714 19.047619  9.523810 19.047619 14.285714 14.285714 19.047619 14.285714 14.285714 14.285714  9.523810
[106] 19.047619  9.523810 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714  9.523810  9.523810 14.285714 14.285714 14.285714 14.285714
[121] 19.047619 14.285714 14.285714 14.285714 19.047619 14.285714 14.285714 14.285714  9.523810 14.285714  9.523810 14.285714 14.285714 14.285714 19.047619
[136] 23.809524 14.285714 14.285714 14.285714 14.285714  4.761905 19.047619 14.285714 14.285714 14.285714 14.285714  4.761905 14.285714 14.285714 14.285714
[151] 14.285714  9.523810 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 19.047619 14.285714  9.523810 14.285714 19.047619
[166] 14.285714 23.809524 19.047619 14.285714 23.809524 14.285714  9.523810 19.047619 14.285714  9.523810 19.047619 14.285714 23.809524 14.285714 14.285714
[181] 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714
[196] 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 19.047619 14.285714 14.285714 14.285714  4.761905 14.285714 14.285714 14.285714 14.285714
[211] 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714  9.523810 14.285714
[226] 14.285714 14.285714  9.523810 14.285714 14.285714  4.761905 14.285714 14.285714  9.523810 14.285714  9.523810 14.285714 23.809524 14.285714  9.523810
[241] 14.285714 14.285714 14.285714 14.285714 14.285714 19.047619 19.047619 14.285714  9.523810 14.285714 19.047619  4.761905 14.285714 19.047619 14.285714
[256] 19.047619 14.285714 14.285714 14.285714  9.523810  9.523810 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714
[271] 14.285714  9.523810 14.285714  4.761905 14.285714 19.047619 14.285714 14.285714 14.285714 19.047619  9.523810 14.285714  9.523810 14.285714 14.285714
[286] 14.285714 14.285714  9.523810 19.047619 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714  9.523810 14.285714 14.285714 14.285714 14.285714
[301] 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714  9.523810 14.285714  9.523810 14.285714 19.047619  9.523810 14.285714
[316] 14.285714  9.523810 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 19.047619 14.285714 14.285714
[331] 14.285714 14.285714 14.285714 14.285714 19.047619 14.285714 14.285714 19.047619 14.285714 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714
[346] 14.285714  9.523810  9.523810 14.285714 19.047619 19.047619 14.285714  9.523810 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 19.047619
[361] 14.285714 14.285714 14.285714 23.809524 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714  9.523810  9.523810 14.285714  9.523810  9.523810
[376]  9.523810 14.285714 14.285714 19.047619 14.285714  9.523810 14.285714 14.285714 14.285714 14.285714 19.047619 14.285714 19.047619 14.285714  9.523810
[391] 14.285714 14.285714  9.523810 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 19.047619 14.285714 14.285714 14.285714 14.285714
[406] 14.285714  9.523810 14.285714 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714 14.285714 14.285714  9.523810 14.285714  4.761905 14.285714
[421] 14.285714 14.285714  4.761905 14.285714 19.047619 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714
[436]  4.761905 23.809524 19.047619 14.285714  9.523810 19.047619 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714  4.761905 14.285714 14.285714
[451] 14.285714  9.523810 14.285714 19.047619 19.047619 14.285714 14.285714 14.285714 14.285714  9.523810 14.285714 19.047619 14.285714  9.523810  9.523810
[466]  9.523810 19.047619  9.523810 14.285714 14.285714  9.523810 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714  9.523810 19.047619
[481] 14.285714 14.285714  9.523810 14.285714 14.285714 19.047619 19.047619 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714
[496]  4.761905 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714 14.285714 14.285714 14.285714  9.523810 14.285714  9.523810 14.285714 19.047619
[511] 14.285714 14.285714 23.809524  9.523810  9.523810 14.285714 14.285714 14.285714 19.047619  9.523810 14.285714 14.285714 14.285714 14.285714 14.285714
[526] 19.047619 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714 14.285714
[541] 14.285714 14.285714 14.285714 19.047619  9.523810  9.523810  9.523810 14.285714 19.047619 14.285714 19.047619 14.285714 14.285714 14.285714 14.285714
[556] 14.285714  9.523810 14.285714 14.285714 14.285714 19.047619  9.523810 19.047619 14.285714  9.523810 14.285714 14.285714 14.285714  9.523810  9.523810
[571] 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714 14.285714 14.285714 14.285714 19.047619 23.809524 14.285714
[586]  9.523810 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714 14.285714 14.285714 19.047619 14.285714 23.809524 14.285714 14.285714 14.285714
[601]  4.761905 14.285714 14.285714 14.285714 14.285714 14.285714 19.047619 14.285714 14.285714 14.285714 14.285714 14.285714 19.047619 14.285714 14.285714
[616] 14.285714 14.285714 14.285714 19.047619 14.285714  9.523810 14.285714  9.523810 14.285714 14.285714 14.285714  9.523810  9.523810 14.285714 14.285714
[631]  9.523810 19.047619 19.047619  9.523810 14.285714 14.285714 14.285714 19.047619 19.047619  9.523810 14.285714 14.285714 14.285714 14.285714 19.047619
[646] 14.285714  9.523810 14.285714 14.285714  9.523810 14.285714 14.285714 14.285714  9.523810 14.285714  9.523810  0.000000  9.523810  4.761905 14.285714
[661]  9.523810  9.523810 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714
[676] 14.285714 14.285714 19.047619 14.285714 14.285714 14.285714 14.285714 23.809524 14.285714  9.523810  9.523810 14.285714 14.285714 14.285714 14.285714
[691] 19.047619 14.285714 19.047619 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714 19.047619 14.285714  9.523810
[706]  4.761905 14.285714 14.285714 14.285714  9.523810  9.523810  9.523810 14.285714 14.285714 14.285714 19.047619 14.285714 14.285714  9.523810 14.285714
[721] 19.047619 14.285714 14.285714 14.285714 14.285714  9.523810 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714
[736]  9.523810  9.523810 14.285714 14.285714 14.285714 14.285714 14.285714 14.285714 19.047619 14.285714 14.285714  9.523810  9.523810 14.285714 14.285714
```

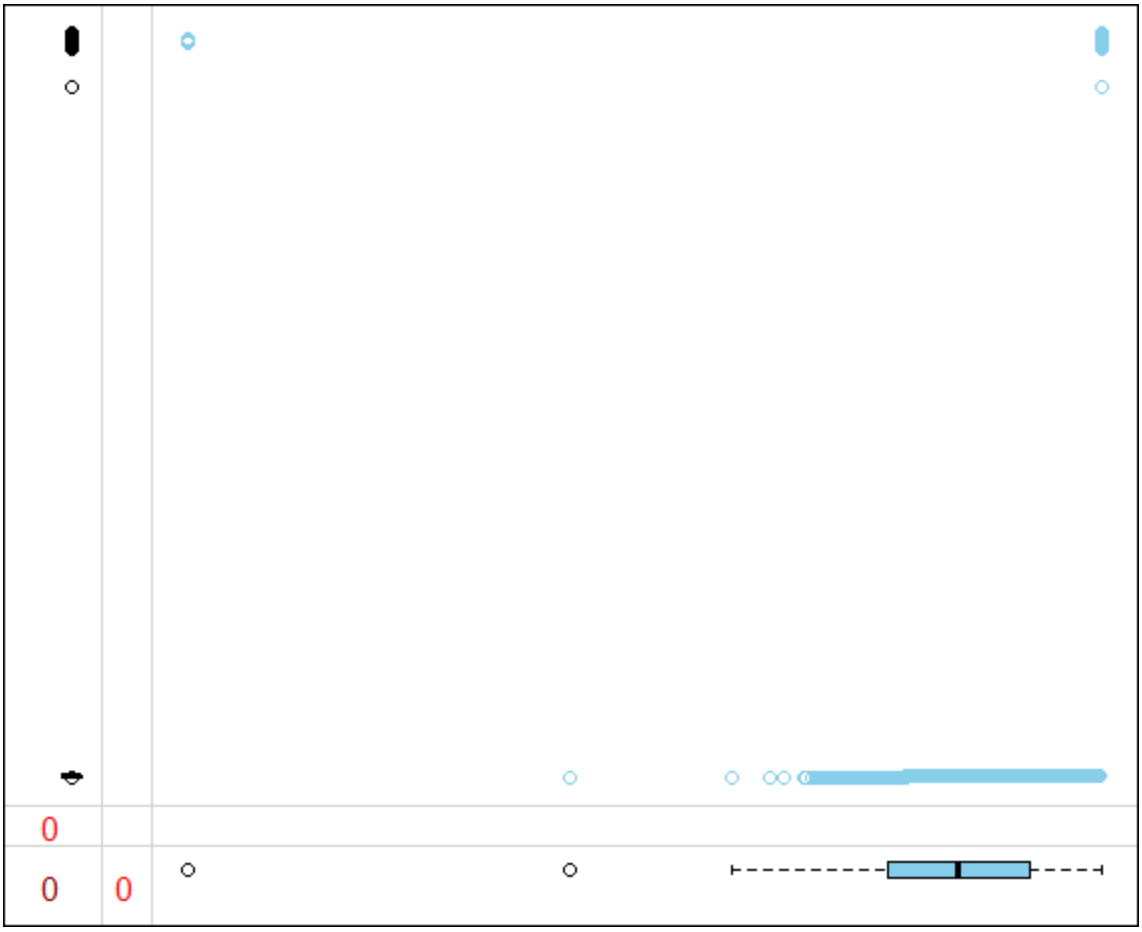
```
library(mice)
md.pattern(data)
library(VIM)
aggr_plot <- aggr(data, col=c('navyblue','red'), numbers=TRUE, sortVars=TRUE, labels=names(data), cex.axis=.7,
gap=3, ylab=c("Histogram of missing data","Pattern"))
marginplot(data[c(1,2)])
# All below charts provide the visualization of missing data in the data set
m <- matrix(data=cbind(rnorm(30, 0), rnorm(30, 2), rnorm(30, 5)), nrow=30, ncol=3)
apply(m, 1, mean)
apply(m, 2, function(x) length(x[x<0]))
apply(m, 2, function(x) is.matrix(x))
apply(m, 2, is.vector)
apply(m, 2, function(x) mean(x[x>0]))
sapply(1:3, function(x) x^2)
lapply(1:3, function(x) x^2)
sapply(1:3, function(x) mean(m[,x]))
sapply(1:3, function(x, y) mean(y[,x]), y=m)
```

```
> library(mice)
> md.pattern(data)
      MI_PRINX offense_id beat location MinOfucr Shift Avg Day UC2 Literal x y MinOfibr_code poss_time occur_date rpt_date MaxOfnum_victims npu neighborhood
64           1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            1
347          1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            1
281          1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            1
3189         1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            1
4            1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            1
2742         1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            1
24           1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            1
15718        1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            1
18           1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            1
39           1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            1
97           1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            1
277          1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            1
1            1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            1
477          1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            1
6            1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            1
2208         1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            1
8            1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            0
34           1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            0
29           1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            0
209          1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            0
34           1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            0
4            1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            0
493          1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            0
2            1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            0
5            1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            0
8            1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            0
17           1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            0
7            1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            0
71           1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   1            0
1            1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   0            1
9            1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   0            0
214          1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   0            0
36           1          1    1        1        1     1   1           1 1 1              1         1         1        1                 1   0            0
66           1          1    1        1        1     1   1           1 1 1              1         1         1        1                 0   1            1
5            1          1    1        1        1     1   1           1 1 1              1         1         1        1                 0   1            0
2            1          1    1        1        1     1   1           1 1 1              1         1         1        0                 1   1            1
3            1          1    1        1        1     1   1           1 1 1              1         1         1        0                 1   1            1
3            1          1    1        1        1     1   1           1 1 1              1         1         0        1                 1   1            1
2            1          1    1        1        1     1   1           1 1 1              1         1         0        0                 1   1            1
4            1          1    1        1        1     1   1           1 1 1              1         0         1        1                 0   1            1
1            1          1    1        1        1     1   1           1 1 1              0         1         1        1                 1   1            1
             0          0    0        0        0     0   0           0 0 0              1         4         5        7                75 260         1185
      loc_type apt_office_num dispo_code apt_office_prefix
64           1              1          1                 1     0
347          1              1          1                 0     1
281          1              1          0                 1     1
3189         1              1          0                 0     2
4            1              0          1                 1     1
2742         1              0          1                 0     2
24           1              0          0                 1     2
15718        1              0          0                 0     3
18           0              1          1                 1     1
39           0              1          1                 0     2
97           0              1          0                 1     2
```



```
> library(VIM)
> aggr_plot <- aggr(data, col=c('navyblue','red'), numbers=TRUE, sortVars=TRUE, labels=names(data), cex.axis=.7, gap=3, ylab=c("Histog
n"))

 Variables sorted by number of missings:
          Variable        Count
 apt_office_prefix 0.9795956501
        dispo_code 0.8583280392
    apt_office_num 0.8271235846
         loc_type 0.1249673007
      neighborhood 0.0442841661
               npu 0.0097163571
  MaxOfnum_victims 0.0028027953
         rpt_date 0.0002615942
        occur_date 0.0001868530
         poss_time 0.0001494824
     MinOfibr_code 0.0000373706
          MI_PRINX 0.0000000000
        offense_id 0.0000000000
              beat 0.0000000000
          location 0.0000000000
          MinOfucr 0.0000000000
             Shift 0.0000000000
           Avg Day 0.0000000000
       UC2 Literal 0.0000000000
                 x 0.0000000000
                 y 0.0000000000
Warning message:
In plot.aggr(res, ...) :
  not enough vertical space to display frequencies (too many combinations)
> marginplot(data[c(1,2)])
>
```





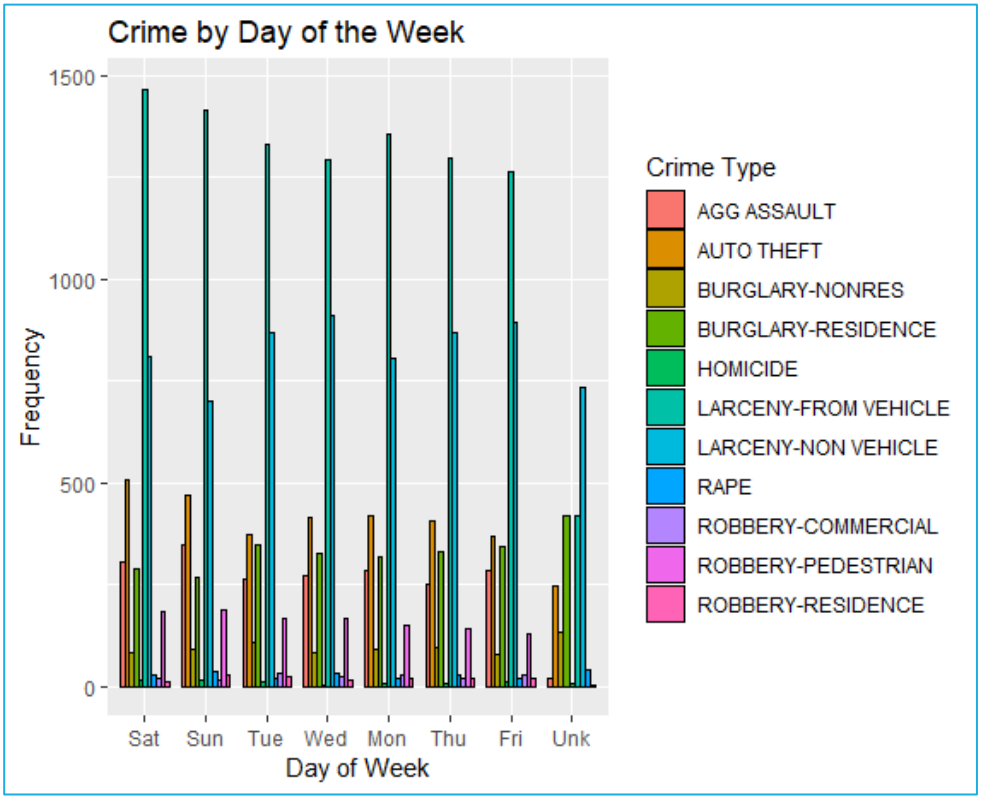

```
> marginplot(data[c(1,2)])
> # All below charts provide the visualization of missing data in the data set
> m <- matrix(data=cbind(rnorm(30, 0), rnorm(30, 2), rnorm(30, 5)), nrow=30, ncol=3)
> apply(m, 1, mean)
 [1] 1.099147 2.971292 2.510675 2.430693 2.643626 2.378484 2.331804 1.257356 3.150575 1.273153 3.292220 1.768180 2.267829 2.930361 2.250306 2.695800 2.081050
[18] 1.777808 1.879934 3.723583 3.050432 2.460920 1.894280 2.854330 1.331913 2.276818 1.768319 1.832987 2.196240 1.690000
> apply(m, 2, function(x) length(x[x<0]))
[1] 15  0  0
> apply(m, 2, function(x) is.matrix(x))
[1] FALSE FALSE FALSE
> apply(m, 2, is.vector)
[1] TRUE TRUE TRUE
> apply(m, 2, function(x) mean(x[x>0]))
[1] 0.7588953 1.9745714 4.9578357
> sapply(1:3, function(x) x^2)
[1] 1 4 9
> lapply(1:3, function(x) x^2)
[[1]]
[1] 1

[[2]]
[1] 4

[[3]]
[1] 9

> sapply(1:3, function(x) mean(m[,x]))
[1] -0.1253956  1.9745714  4.9578357
> sapply(1:3, function(x, y) mean(y[,x]), y=m)
[1] -0.1253956  1.9745714  4.9578357
```

```
library(tidyverse)
library(ggmap)
library(readxl)
library(kableExtra)
library(knitr)
str(COBRA_YTD2017)
COBRA_YTD2017$long <- COBRA_YTD2017$x %>%
  as.numeric()

COBRA_YTD2017$lat <- COBRA_YTD2017$y %>%
  as.numeric()

COBRA_YTD2017$loc_type <- COBRA_YTD2017$`UC2 Literal` %>%   as.factor()

COBRA_YTD2017$days <- COBRA_YTD2017$`Avg Day` %>%
  as.factor()

kable(count(COBRA_YTD2017, loc_type, sort=TRUE), "html", col.names=c("Crime Type", "Frequency")) %>%
  kable_styling(bootstrap_options="striped", full_width=FALSE)

COBRA_YTD2017 %>%
  group_by(days, loc_type) %>%
  summarize(freq=n()) %>%
  ggplot(aes(reorder(days, -freq), freq)) +
  geom_bar(aes(fill=loc_type), position="dodge", stat="identity", width=0.8, color="black") +
  xlab("Day of Week") +
  ylab("Frequency") +
  labs(fill="Crime Type") +
  ggtitle("Crime by Day of the Week")

kable

atlanta_map <- qmap("atlanta",
                    zoom=12,
                    source="stamen",
                    maptype="toner",
                    color="bw")
atlanta_map
```

```
> COBRA_YTD2017$days <- COBRA_YTD2017$`Avg Day` %>%
+   as.factor()
> kable(count(COBRA_YTD2017, loc_type, sort=TRUE), "html", col.names=c("Crime Type", "Frequency")) %>%
+   kable_styling(bootstrap_options="striped", full_width=FALSE)
<table class="table table-striped" style="width: auto !important; margin-left: auto; margin-right: auto;">
 <thead>
  <tr>
   <th style="text-align:left;"> Crime Type </th>
   <th style="text-align:right;"> Frequency </th>
  </tr>
 </thead>
<tbody>
  <tr>
   <td style="text-align:left;"> LARCENY-FROM VEHICLE </td>
   <td style="text-align:right;"> 9840 </td>
  </tr>
  <tr>
   <td style="text-align:left;"> LARCENY-NON VEHICLE </td>
   <td style="text-align:right;"> 6589 </td>
  </tr>
  <tr>
   <td style="text-align:left;"> AUTO THEFT </td>
   <td style="text-align:right;"> 3197 </td>
  </tr>
  <tr>
   <td style="text-align:left;"> BURGLARY-RESIDENCE </td>
   <td style="text-align:right;"> 2635 </td>
  </tr>
  <tr>
   <td style="text-align:left;"> AGG ASSAULT </td>
   <td style="text-align:right;"> 2024 </td>
  </tr>
  <tr>
   <td style="text-align:left;"> ROBBERY-PEDESTRIAN </td>
   <td style="text-align:right;"> 1126 </td>
  </tr>
  <tr>
   <td style="text-align:left;"> BURGLARY-NONRES </td>
   <td style="text-align:right;"> 758 </td>
  </tr>
  <tr>
   <td style="text-align:left;"> RAPE </td>
   <td style="text-align:right;"> 226 </td>
  </tr>
  <tr>
   <td style="text-align:left;"> ROBBERY-COMMERCIAL </td>
   <td style="text-align:right;"> 157 </td>
  </tr>
  <tr>
   <td style="text-align:left;"> ROBBERY-RESIDENCE </td>
   <td style="text-align:right;"> 132 </td>
  </tr>
  <tr>
   <td style="text-align:left;"> HOMICIDE </td>
   <td style="text-align:right;"> 75 </td>
  </tr>
</tbody>
</table>
>
```

| Crime Type | Frequency |
|---|---|
| LARCENY-FROM VEHICLE | 9840 |
| LARCENY-NON VEHICLE | 6589 |
| AUTO THEFT | 3197 |
| BURGLARY-RESIDENCE | 2635 |
| AGG ASSAULT | 2024 |
| ROBBERY-PEDESTRIAN | 1126 |
| BURGLARY-NONRES | 758 |
| RAPE | 226 |
| ROBBERY-COMMERCIAL | 157 |
| ROBBERY-RESIDENCE | 132 |
| HOMICIDE | 75 |

```
> COBRA_YTD2017 %>%
+   group_by(days, loc_type) %>%
+   summarize(freq=n()) %>%
+   ggplot(aes(reorder(days, -freq), freq)) +
+   geom_bar(aes(fill=loc_type), position="dodge", stat="identity", width=0.8, color="black") +
+   xlab("Day of Week") +
+   ylab("Frequency") +
+   labs(fill="Crime Type") +
+   ggtitle("Crime by Day of the Week")
>
```

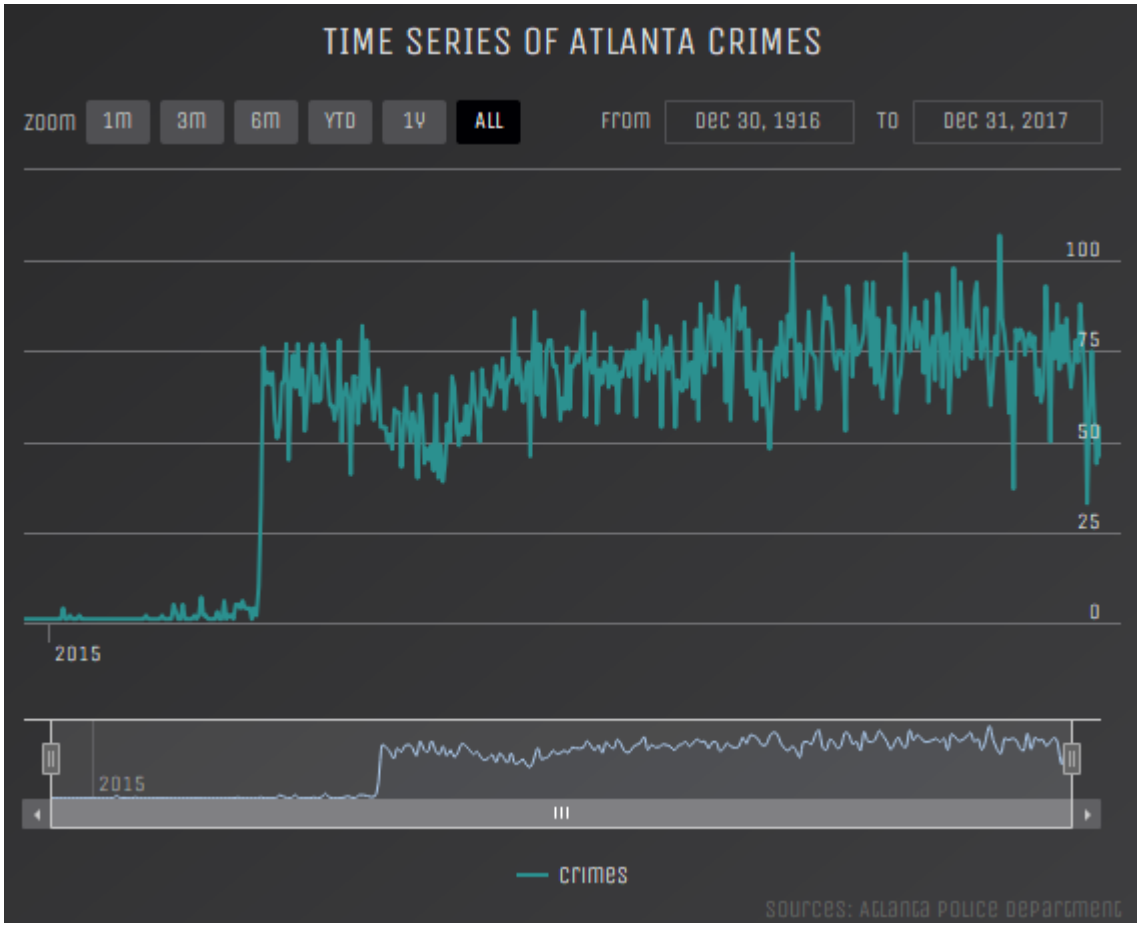


```
================================================================================
> kable
function (x, format, digits = getOption("digits"), row.names = NA,
    col.names = NA, align, caption = NULL, format.args = list(),
    escape = TRUE, ...)
{
    if (missing(format) || is.null(format))
        format = getOption("knitr.table.format")
    if (is.null(format))
        format = if (is.null(pandoc_to()))
            switch(out_format() %n% "markdown", latex = "latex",
                listings = "latex", sweave = "latex", html = "html",
                markdown = "markdown", rst = "rst", stop("table format not implemented
yet!"))
        else if (isTRUE(opts_knit$get("kable.force.latex")) &&
            is_latex_output()) {
            "latex"
        }
        else "pandoc"
    if (is.function(format))
        format = format()
    if (format != "latex" && !missing(align) && length(align) ==
        1L)
        align = strsplit(align, "")[[1]]
    if (!is.null(caption) && !is.na(caption))
        caption = paste0(create_label("tab:", opts_current$get("label"),
            latex = (format == "latex")), caption)
    if (inherits(x, "list")) {
        if (format == "pandoc" && is_latex_output())
            format = "latex"
        res = lapply(x, kable, format = format, digits = digits,
            row.names = row.names, col.names = col.names, align = align,
            caption = NA, format.args = format.args, escape = escape,
            ...)
        res = unlist(lapply(res, paste, collapse = "\n"))
        res = if (format == "latex") {
            kable_latex_caption(res, caption)
        }
        else if (format == "html" || (format == "pandoc" && is_html_output()))
            kable_html(matrix(paste0("\n\n", res, "\n\n"), 1),
                caption = caption, escape = FALSE, table.attr = "class=\"kable_wrapper\
""")
        else {
            res = paste(res, collapse = "\n\n")
            if (format == "pandoc")
                kable_pandoc_caption(res, caption)
            else res
        }
        return(structure(res, format = format, class = "knitr_kable"))
    }
    if (!is.matrix(x))
        x = as.data.frame(x)
    if (identical(col.names, NA))
        col.names = colnames(x)
    m = ncol(x)
    isn = if (is.matrix(x))
        rep(is.numeric(x), m)
    else sapply(x, is.numeric)
    if (missing(align) || (format == "latex" && is.null(align)))
        align = ifelse(isn, "r", "l")
    digits = rep(digits, length.out = m)
```

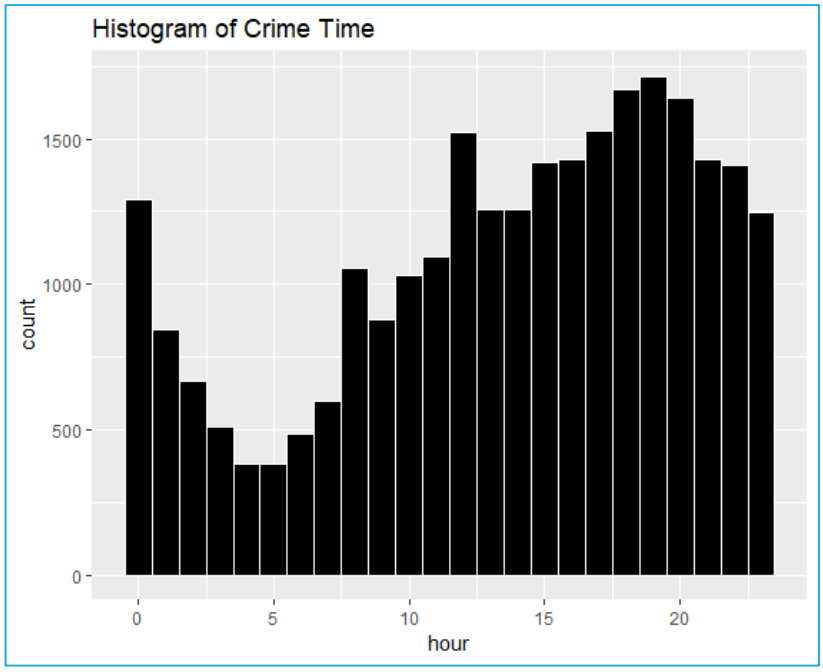
```
    for (j in seq_len(m)) {
        if (is_numeric(x[, j]))
            x[, j] = round(x[, j], digits[j])
    }
    if (any(isn)) {
        if (is.matrix(x)) {
            if (is.table(x) && length(dim(x)) == 2)
                class(x) = "matrix"
            x = format_matrix(x, format.args)
        }
        else x[, isn] = format_args(x[, isn], format.args)
    }
    if (is.na(row.names))
        row.names = has_rownames(x)
    if (!is.null(align))
        align = rep(align, length.out = m)
    if (row.names) {
        x = cbind(` ` = rownames(x), x)
        if (!is.null(col.names))
            col.names = c(" ", col.names)
        if (!is.null(align))
            align = c("l", align)
    }
    n = nrow(x)
    x = replace_na(to_character(as.matrix(x)), is.na(x))
    if (!is.matrix(x))
        x = matrix(x, nrow = n)
    x = trimws(x)
    colnames(x) = col.names
    if (format != "latex" && length(align) && !all(align %in%
        c("l", "r", "c")))
        stop("'align' must be a character vector of possible values 'l', 'r', and 'c'")
    attr(x, "align") = align
    res = do.call(paste("kable", format, sep = "_"), list(x = x,
        caption = caption, escape = escape, ...))
    structure(res, format = format, class = "knitr_kable")
}
<bytecode: 0x000000001f56f938>
<environment: namespace:knitr>
================================================================================
```

```
library(dplyr)
library(data.table)
library(ggplot2)
at <- COBRA_YTD2017
str(at)
at$MI_PRINX <- at$apt_office_prefix <- at$apt_office_num <- at$location <- at$dispo_code <- at$loc_type <-
at$npu <- NULL
library(chron)
library(lubridate)
at$lon <- at$x
at$lat <- at$y
at$occur_date <- mdy(at$occur_date)
at$rpt_date <- mdy(at$rpt_date)
at$occur_time <- chron(times=at$occur_time)
at$lon <- as.numeric(at$lon)
at$lat <- as.numeric(at$lat)
at$x <- at$y <- NULL
library(xts)
by_Date <- na.omit(at) %>% group_by(occur_date) %>% summarise(Total = n())
tseries <- xts(by_Date$Total, order.by= by_Date$occur_date)
library(highcharter)
hchart(tseries, name = "Crimes") %>%
  hc_add_theme(hc_theme_darkunica()) %>%
  hc_credits(enabled = TRUE, text = "Sources: Atlanta Police Department", style = list(fontSize = "12px")) %>%
  hc_title(text = "Time Series of Atlanta Crimes") %>%
  hc_legend(enabled = TRUE)
hchart
#Graph provides the data spread of the crime during the year
at$dayofWeek <- weekdays(as.Date(at$occur_date))
at$hour <- sub(":.*", "", at$occur_time)
at$hour <- as.numeric(at$hour)
ggplot(aes(x = hour), data = at) + geom_histogram(bins = 24, color='white', fill='black') +
  ggtitle('Histogram of Crime Time')
```

```
> library(dplyr)
> library(data.table)
> library(ggplot2)
> at <- COBRA_YTD2017
> str(at)
Classes 'spec_tbl_df', 'tbl_df', 'tbl' and 'data.frame':        26759 obs. of  26 variables:
 $ MI_PRINX          : num  8924155 8924156 8924157 8924158 8924159 ...
 $ offense_id        : num  1.74e+08 1.74e+08 1.74e+08 1.74e+08 1.74e+08 ...
 $ rpt_date          : chr  "12/31/2017" "12/31/2017" "12/31/2017" "12/31/2017" ...
 $ occur_date        : chr  "12/30/2017" "12/18/2017" "12/30/2017" "12/30/2017" ...
 $ occur_time        : 'hms' num  23:15:00 13:00:00 22:01:00 20:00:00 ...
  ..- attr(*, "units")= chr "secs"
 $ poss_date         : chr  "12/31/2017" "12/30/2017" "12/31/2017" "12/31/2017" ...
 $ poss_time         : 'hms' num  00:30:00 22:00:00 01:00:00 01:06:00 ...
  ..- attr(*, "units")= chr "secs"
 $ beat              : num  510 501 303 507 409 612 605 603 605 304 ...
 $ apt_office_prefix: chr  NA NA NA NA ...
 $ apt_office_num    : chr  NA NA NA NA ...
 $ location          : chr  "43 JESSE HILL JR DR NE" "1169 ATLANTIC DR NW" "633 PRYOR ST SW" "333 NELSON ST SW" ...
 $ MinOfucr          : chr  "0640" "0640" "0640" "0640" ...
 $ MinOfibr_code     : chr  "2305" "2305" "2305" "2305" ...
 $ dispo_code        : num  NA NA NA NA NA NA NA NA NA NA ...
 $ MaxOfnum_victims : num  2 1 1 1 2 1 1 1 1 1 ...
 $ Shift             : chr  "Morn" "Unk" "Morn" "Eve" ...
 $ Avg Day           : chr  "Sat" "Unk" "Sat" "Sat" ...
 $ loc_type          : Factor w/ 11 levels "AGG ASSAULT",..: 6 6 6 6 6 6 10 6 6 4 ...
 $ UC2 Literal       : chr  "LARCENY-FROM VEHICLE" "LARCENY-FROM VEHICLE" "LARCENY-FROM VEHICLE" "LARCENY-FROM VEHICLE" ...
 $ neighborhood      : chr  "Downtown" "Home Park" "Mechanicsville" "Castleberry Hill" ...
 $ npu               : chr  "M" "E" "V" "M" ...
 $ x                 : num  -84.4 -84.4 -84.4 -84.4 -84.5 ...
 $ y                 : num  33.8 33.8 33.7 33.8 33.7 ...
 $ long              : num  -84.4 -84.4 -84.4 -84.4 -84.5 ...
 $ lat               : num  33.8 33.8 33.7 33.8 33.7 ...
 $ days              : Factor w/ 8 levels "Fri","Mon","Sat",..: 3 7 3 3 4 4 4 4 3 4 ...
 - attr(*, "problems")=Classes 'tbl_df', 'tbl' and 'data.frame':        9 obs. of  5 variables:
  ..$ row     : int  3239 7945 8527 10145 11912 12629 13305 17684 20632
  ..$ col     : chr  "dispo_code" "dispo_code" "dispo_code" "dispo_code" ...
  ..$ expected: chr  "a double" "a double" "a double" "a double" ...
  ..$ actual  : chr  "COS" "ADM" "ADM" "ADM" ...
  ..$ file    : chr  "'G:/DATA ANALYTICS/DATA/crime-in-atlanta-2017/COBRA-YTD2017.csv'" "'G:/DATA ANALYTICS/DATA/crime-in-atlanta-2017/COB
TA ANALYTICS/DATA/crime-in-atlanta-2017/COBRA-YTD2017.csv'" "'G:/DATA ANALYTICS/DATA/crime-in-atlanta-2017/COBRA-YTD2017.csv'" ...
 - attr(*, "spec")=
  .. cols(
  ..   MI_PRINX = col_double(),
  ..   offense_id = col_double(),
  ..   rpt_date = col_character(),
  ..   occur_date = col_character(),
  ..   occur_time = col_time(format = ""),
  ..   poss_date = col_character(),
  ..   poss_time = col_time(format = ""),
  ..   beat = col_double(),
  ..   apt_office_prefix = col_character(),
  ..   apt_office_num = col_character(),
  ..   location = col_character(),
  ..   MinOfucr = col_character(),
  ..   MinOfibr_code = col_character(),
  ..   dispo_code = col_double(),
  ..   MaxOfnum_victims = col_double(),
  ..   Shift = col_character(),
  ..   `Avg Day` = col_character(),
  ..   loc_type = col_double(),
  ..   `UC2 Literal` = col_character(),
  ..   neighborhood = col_character(),
  ..   npu = col_character(),
  ..   x = col_double(),
  ..   y = col_double()
  .. )
> at$MI_PRINX <- at$apt_office_prefix <- at$apt_office_num <- at$location <- at$dispo_code <- at$loc_type <- at$npu <- NULL
> library(chron)
> library(lubridate)
> at$lon <- at$x
> at$lat <- at$y
> at$occur_date <- mdy(at$occur_date)
> at$rpt_date <- mdy(at$rpt_date)
> at$occur_time <- chron(times=at$occur_time)
> at$lon <- as.numeric(at$lon)
> at$lat <- as.numeric(at$lat)
> at$x <- at$y <- NULL
> library(xts)
> by_Date <- na.omit(at) %>% group_by(occur_date) %>% summarise(Total = n())
> tseries <- xts(by_Date$Total, order.by= by_Date$occur_date)
> library(highcharter)
> hchart(tseries, name = "Crimes") %>%
+   hc_add_theme(hc_theme_darkunica()) %>%
+   hc_credits(enabled = TRUE, text = "Sources: Atlanta Police Department", style = list(fontSize = "12px")) %>%
+   hc_title(text = "Time Series of Atlanta Crimes") %>%
+   hc_legend(enabled = TRUE)
>
```

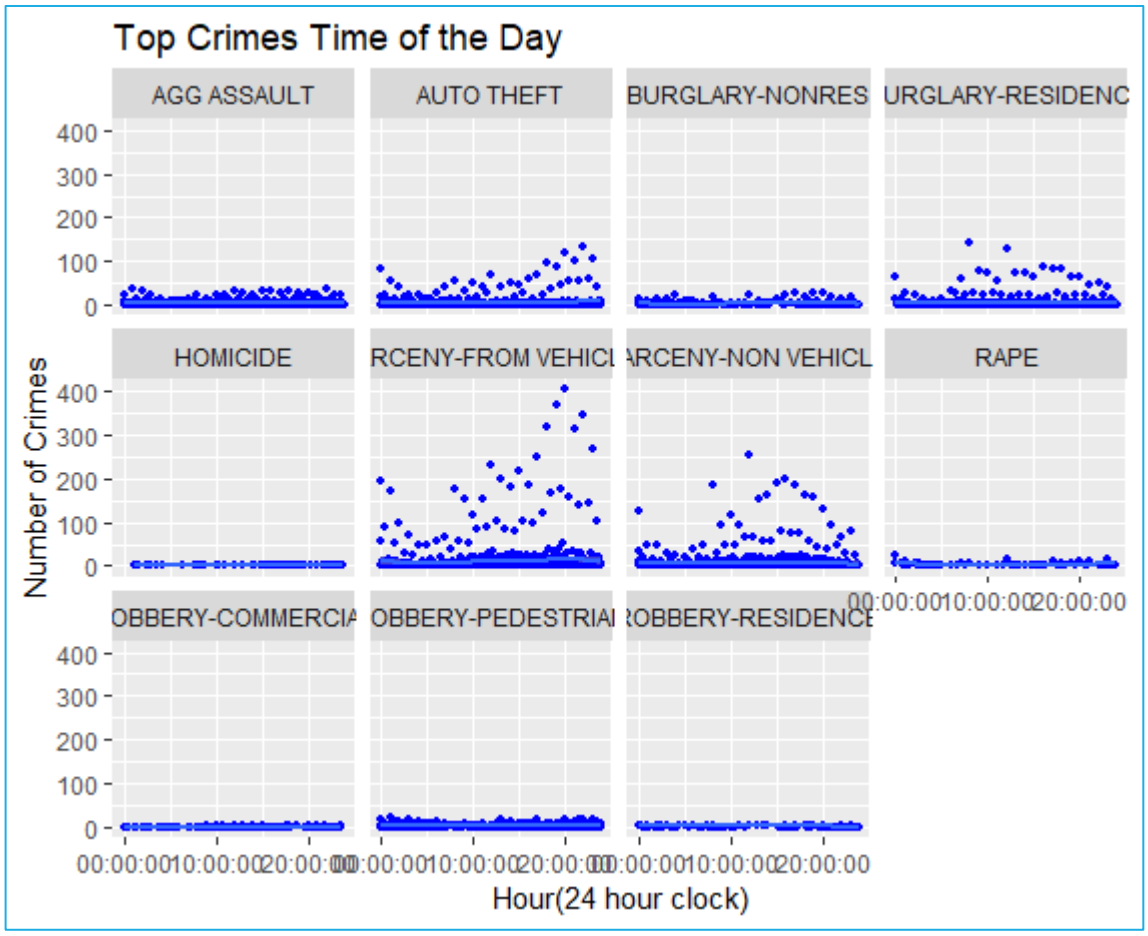



```
> #Graph provides the data spread of the crime during the year
> at$dayofweek <- weekdays(as.Date(at$occur_date))
> at$hour <- sub(":.*", "", at$occur_time)
> at$hour <- as.numeric(at$hour)
> ggplot(aes(x = hour), data = at) + geom_histogram(bins = 24, color='white', fill='black') +
+   ggtitle('Histogram of Crime Time')
```

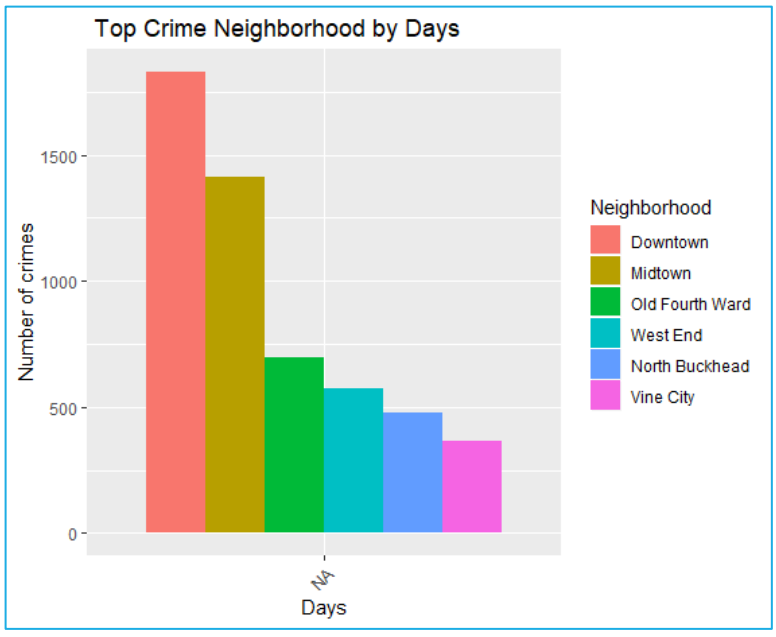


```
#The crime time distribution appears bimodal with peaking around midnight and again at the noon, then again between 6pm and 8pm.
topCrimes_1 <- COBRA_YTD2017 %>% group_by(`UC2 Literal`,occur_time) %>%
  summarise(total = n())

ggplot(aes(x = occur_time, y = total), data = topCrimes_1) +
  geom_point(colour="blue", size=1) +
  geom_smooth(method="loess") +
  xlab('Hour(24 hour clock)') +
  ylab('Number of Crimes') +
  ggtitle('Top Crimes Time of the Day') +
  facet_wrap(~`UC2 Literal`)


#Downtown and midtown are the most common locations where crimes take place, followed by Old Fourth Ward and West End.
topLocations <- subset(at, neighborhood =="Downtown"|neighborhood =="Midtown" | neighborhood=="Old Fourth Ward" | neighborhood=="West End" | neighborhood=="Vine City" | neighborhood=="North Buckhead")

topLocations <- within(topLocations,  neighborhood <- factor(neighborhood, levels = names(sort(table(neighborhood), decreasing = T))))

topLocations$days <- ordered(topLocations$days,
                             levels = c('Monday', 'Tuesday', 'Wednesday', 'Thursday', 'Friday', 'Saturday', 'Sunday'))

ggplot(data = topLocations, aes(x = days, fill = neighborhood)) +
  geom_bar(width = 0.9, position = position_dodge()) + ggtitle(" Top Crime Neighborhood by Days") +
  labs(x = "Days", y = "Number of crimes", fill = guide_legend(title = "Neighborhood")) + theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

```
> #The crime time distribution appears bimodal with peaking around midnight and again at the noon, then again between 6pm and 8pm.
> topCrimes_1 <- COBRA_YTD2017 %>% group_by(`UC2 Literal`,occur_time) %>%
+   summarise(total = n())
> ggplot(aes(x = occur_time, y = total), data = topCrimes_1) +
+   geom_point(colour="blue", size=1) +
+   geom_smooth(method="loess") +
+   xlab('Hour(24 hour clock)') +
+   ylab('Number of Crimes') +
+   ggtitle('Top Crimes Time of the Day') +
+   facet_wrap(~`UC2 Literal`)
```




```
> #Downtown and midtown are the most common locations where crimes take place, followed by Old Fourth Ward and West End.
> topLocations <- subset(at, neighborhood =="Downtown"|neighborhood =="Midtown" | neighborhood=="Old Fourth Ward" | neighborhood=="West End" | neighborh
ity" | neighborhood=="North Buckhead")
> topLocations <- within(topLocations,  neighborhood <- factor(neighborhood, levels = names(sort(table(neighborhood), decreasing = T))))
> topLocations$days <- ordered(topLocations$days,
+                              levels = c('Monday', 'Tuesday', 'Wednesday', 'Thursday', 'Friday', 'Saturday', 'Sunday'))
> ggplot(data = topLocations, aes(x = days, fill = neighborhood)) +
+   geom_bar(width = 0.9, position = position_dodge()) + ggtitle(" Top Crime Neighborhood by Days") +
+   labs(x = "Days", y = "Number of crimes", fill = guide_legend(title = "Neighborhood")) + theme(axis.text.x = element_text(angle = 45, hjust = 1))
> |
```



**b. What is the difference between covariance and correlation, take an example from this dataset and show the differences if any?**

- Covariance and Correlation are two mathematical concepts which are quite commonly used in business statistics.
- Both of these two determine the relationship and measures the dependency between two random variables.
- Despite, some similarities between these two mathematical terms, they are different from each other. Correlation is when the change in one item may result in the change in another item.
- Correlation is considered as the best tool for for measuring and expressing the quantitative relationship between two variables in formula.
- Covariance is when two items vary together. Read the given article to know the differences between covariance and correlation.

| BASIS FOR COMPARISON | CORRELATION | COVARIANCE |
|---|---|---|
| Meaning | Correlation is a statistical measure that indicates how strongly two variables are related. | Covariance is a measure indicating the extent to which two random variables change in tandem. |
| What is it? | Scaled version of covariance | Measure of correlation |
| Values | Lie between -1 and +1 | Lie between -∞ and +∞ |
| Change in scale | Does not affects correlation | Affects Covariance |
| Unit free measure | Yes | No |

**Similarities**

- Both measures only linear relationship between two variables, i.e. when the correlation coefficient is zero, covariance is also zero. Further, the two measures are unaffected by the change in location.
- Correlation is a special case of covariance which can be obtained when the data is standardized. Now, when it comes to making a choice, which is a better measure of the relationship between two variables, correlation is preferred over covariance, because it remains unaffected by the change in location and scale, and can also be used to make a comparison between two pairs of variables.
- correlation is preferred over covariance, because it remains unaffected by the change in location and scale, and can also be used to make a comparison between two pairs of variables.

```
#Correlation & covariance
cor(COBRA_YTD2017$x,COBRA_YTD2017$y)
cov(COBRA_YTD2017$x,COBRA_YTD2017$y)
cor.test(COBRA_YTD2017$x,COBRA_YTD2017$y)
cor(COBRA_YTD2017$long,COBRA_YTD2017$lat)
cor.test(COBRA_YTD2017$long,COBRA_YTD2017$lat)
cov(COBRA_YTD2017$long,COBRA_YTD2017$lat)
plot(COBRA_YTD2017$x,COBRA_YTD2017$y)
mod=lm(COBRA_YTD2017$long~COBRA_YTD2017$lat)
summary(mod)
predict(mod)
pred= predict(mod)
COBRA_YTD2017$predicted=NA
COBRA_YTD2017$predicted=pred
COBRA_YTD2017$error=COBRA_YTD2017$residuals
library(car)
dwt(mod)
plot(COBRA_YTD2017$long,COBRA_YTD2017$lat,abline(COBRA_YTD2017$long~COBRA_YTD2017$lat),col='red')
```

```
> #Correlation & covariance
> cor(COBRA_YTD2017$x,COBRA_YTD2017$y)
[1] -0.9998355
> cov(COBRA_YTD2017$x,COBRA_YTD2017$y)
[1] -23.86342
> cor.test(COBRA_YTD2017$x,COBRA_YTD2017$y)

        Pearson's product-moment correlation

data:  COBRA_YTD2017$x and COBRA_YTD2017$y
t = -9017.2, df = 26757, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.9998394 -0.9998315
sample estimates:
       cor
-0.9998355

> cor(COBRA_YTD2017$long,COBRA_YTD2017$lat)
[1] -0.9998355
> cor.test(COBRA_YTD2017$long,COBRA_YTD2017$lat)

        Pearson's product-moment correlation

data:  COBRA_YTD2017$long and COBRA_YTD2017$lat
t = -9017.2, df = 26757, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.9998394 -0.9998315
sample estimates:
       cor
-0.9998355

> cov(COBRA_YTD2017$long,COBRA_YTD2017$lat)
[1] -23.86342
>
```

```
> cov(COBRA_YTD2017$long,COBRA_YTD2017$lat)
[1] -23.86342
> plot(COBRA_YTD2017$x,COBRA_YTD2017$y)
> mod=lm(COBRA_YTD2017$long~COBRA_YTD2017$lat)
> summary(mod)

Call:
lm(formula = COBRA_YTD2017$long ~ COBRA_YTD2017$lat)

Residuals:
     Min       1Q   Median       3Q      Max
-0.36967 -0.08504  0.01124  0.08245  0.35407

Coefficients:
                   Estimate Std. Error   t value Pr(>|t|)
(Intercept)       -0.0219717  0.0093186    -2.358   0.0184 *
COBRA_YTD2017$lat -2.4996054  0.0002772 -9017.210   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1401 on 26757 degrees of freedom
Multiple R-squared:  0.9997,    Adjusted R-squared:  0.9997
F-statistic: 8.131e+07 on 1 and 26757 DF,  p-value: < 2.2e-16

> predict(mod)
          1           2           3           4           5           6           7           8           9          10          11          12
-84.39820118 -84.47548898 -84.35265837 -84.38755286 -84.31231474 -84.35880740 -84.37128043 -84.44121939 -84.37462990 -84.32313803 -84.34905894 -84.38300358
         13          14          15          16          17          18          19          20          21          22          23          24
-84.21475514 -84.30094154 -84.33486118 -84.26084787 -84.40442520 -84.50041005 -84.27562054 -84.62726502 -84.40257549 -84.43532032 -84.26789675 -84.39457676
         25          26          27          28          29          30          31          32          33          34          35          36
-84.37482987 -84.47993828 -84.43552029 -84.40914946 -84.59552003 -84.62551530 -84.36765601 -84.67755708 -84.54125360 -84.40897448 -84.49868532 -84.62551530
         37          38          39          40          41          42          43          44          45          46          47          48
-84.32171326 -84.14966542 -84.42074762 -84.51900711 -84.37140541 -84.47193954 -84.39317698 -84.50428444 -84.54177852 -84.49201137 -84.39730133 -84.47231448
         49          50          51          52          53          54          55          56          57          58          59          60
-84.51813225 -84.36078209 -84.25194927 -84.64116283 -84.37765443 -84.35113361 -84.38260365 -84.51833222 -84.62579025 -84.40095075 -84.20833116 -84.44846825
         61          62          63          64          65          66          67          68          69          70          71          72
-84.29301779 -84.27956991 -84.40617492 -84.19038399 -84.39255208 -84.39640147 -84.11082155 -84.14291649  -0.02197167 -84.38292859 -84.61489197 -84.46364085
         73          74          75          76          77          78          79          80          81          82          83          84
-84.39747630 -84.42889634 -84.44316908 -84.38570316 -84.25712346  -0.02197167 -84.30936521 -84.38977751 -84.38400343 -84.38660301 -84.19140883 -84.59844457
         85          86          87          88          89          90          91          92          93          94          95          96
-84.38185376 -84.49811041 -84.21468016 -84.53642936 -84.61341721 -84.36613125 -84.45734185 -84.47238947 -84.36185692 -84.29046819 -84.56599969 -84.40050082
         97          98          99         100         101         102         103         104         105         106         107         108
-84.54245341 -84.54635280 -84.44779335 -84.43102100 -84.47436416 -84.40442520 -84.64116283 -84.64041295 -84.40732474 -84.46421576 -84.38490328 -84.35868242
        109         110         111         112         113         114         115         116         117         118         119         120
-84.51833222 -84.31738894 -84.22980277 -84.51833222 -84.52545609 -84.64113783 -84.40110073 -84.38357849 -84.38812777 -84.38822776 -84.62551530 -84.27489565
        121         122         123         124         125         126         127         128         129         130         131         132
-84.28179456 -84.63016456 -84.36115703 -84.37765443 -84.52678088 -84.39345193 -84.47336432 -84.43879478 -84.56487487 -84.21035584 -84.20810619 -84.34655934
        133         134         135         136         137         138         139         140         141         142         143         144
-84.34865900 -84.14526611 -84.35815751 -84.28024481 -84.15543951 -84.69247973 -84.29954176 -84.40075078 -84.19860769 -84.69242973 -84.38810278 -84.40050082
        145         146         147         148         149         150         151         152         153         154         155         156
-84.36903079 -84.42579683 -84.19813277 -84.46636542 -84.41444862 -84.53415472 -84.63821329 -84.54130359 -84.46306594 -84.25667353 -84.42139752 -84.42579683
        157         158         159         160         161         162         163         164         165         166         167         168
-84.51468279 -84.35395817 -84.32176325 -84.62601522 -84.24112598 -84.34355981 -84.61686666 -84.52210662 -84.55457650 -84.41107415 -84.52540610 -84.43749498
        169         170         171         172         173         174         175         176         177         178         179         180
-84.36698111 -84.53340484 -84.31936363 -84.41764811 -84.43677009 -84.36185692 -84.47736369 -84.42814646 -84.39302700 -84.11039662 -84.14436626 -84.41507352
        181         182         183         184         185         186         187         188         189         190         191         192
-84.41789807 -84.39345193 -84.35360822 -84.39540163 -84.39000248 -84.31583919 -84.30746551 -84.54732764 -84.49833538 -84.40007589 -84.57079894 -84.27072131
        193         194         195         196         197         198         199         200         201         202         203         204
-84.38625307 -84.52508115 -84.29791702 -84.38047898 -84.51438284 -84.19998248 -84.40202558 -84.27777020 -84.52418130 -84.35438310 -84.42687166 -84.39625149
```

```
> pred= predict(mod)
> COBRA_YTD2017$predicted=NA
> COBRA_YTD2017$predicted=pred
> COBRA_YTD2017$error=COBRA_YTD2017$residuals
Warning message:
Unknown or uninitialised column: 'residuals'.
> library(car)
> dwt(mod)
 lag Autocorrelation D-W Statistic p-value
   1      0.02809992       1.943799       0
 Alternative hypothesis: rho != 0
> plot(COBRA_YTD2017$long,COBRA_YTD2017$lat,abline(COBRA_YTD2017$long~COBRA_YTD2017$lat),col='red')
>
```