Analysis of the Crimes in Chicago during the Super Bowl

Josean Velazquez

March 19, 2018

Reads the Chicago crime into a data frema called vdat

```
vdat = read.table("chicago_crimes_by_day_and_hour_2001_to_2017.csv",header=T,as.is=T,sep=",")

head(vdat)
```

```
##   month day year day_of_year weekday   jul hour assaults assaults_domestic
## 1     1   1 2001           1  Monday 11323    0        5                 1
## 2     1   1 2001           1  Monday 11323    1        6                 1
## 3     1   1 2001           1  Monday 11323    2        5                 1
## 4     1   1 2001           1  Monday 11323    3        3                 0
## 5     1   1 2001           1  Monday 11323    4        6                 2
## 6     1   1 2001           1  Monday 11323    5        3                 0
##   batteries batteries_domestic burglaries thefts mv_thefts prostitution
## 1        17                 8          2    121         9            0
## 2        23                 8          1      9         2            0
## 3        27                11          2      3         2            0
## 4        16                 9          1     13         3            0
## 5        19                11          2      3         2            0
## 6        13                 4          1      7         2            0
##   narcotics
## 1         3
## 2         1
## 3         1
## 4         2
## 5         0
## 6         0
```

```
#Reads the chigaco crime into a data frame called vdat

vdat$batteries_nodomestic = vdat$batteries - vdat$batteries_domestic

#Calculates the number of batteries assaults that are nor domestic

superbowl = read.csv("Superbowl_Data_Velazquez.csv")
head(superbowl)
```

```
##   Month Day Year
## 1     1  15 1967
## 2     1  14 1968
## 3     1  11 1970
## 4     1  17 1971
## 5     1  16 1972
## 6     1  14 1973
```

```
# Reads the Superbowl Sunday data
```

```r
superbowl$jul = julian(superbowl$Month,superbowl$Day,superbowl$Year)


#uses the function julian to turn the month/day/year in to a single code for reference later

sunday = subset(vdat,jul%in%superbowl$jul)
#Selects the number of crimes on Super Bowl Sunday
head(sunday)
```

```
##     month day year day_of_year weekday   jul hour assaults
## 649     1  28 2001          28  Sunday 11350    0        2
## 650     1  28 2001          28  Sunday 11350    1        2
## 651     1  28 2001          28  Sunday 11350    2        5
## 652     1  28 2001          28  Sunday 11350    3        3
## 653     1  28 2001          28  Sunday 11350    4        2
## 654     1  28 2001          28  Sunday 11350    5        1
##     assaults_domestic batteries batteries_domestic burglaries thefts
## 649                 1         8                  4          1      6
## 650                 0        18                  8          1      5
## 651                 2        14                  4          1      2
## 652                 0         9                  5          1      4
## 653                 0        12                  7          2      3
## 654                 0         2                  1          0      3
##     mv_thefts prostitution narcotics batteries_nodomestic
## 649         5            0         2                   4
## 650         4            1         0                  10
## 651         4            0         1                  10
## 652         0            0         0                   4
## 653         2            0         1                   5
## 654         1            8         0                   1
```

```r
weekbefore = subset(vdat,(jul-7)%in%sunday$jul)
#Selects the number of crimes a week before Super Bowl Sunday
head(weekbefore)
```

```
##     month day year day_of_year weekday   jul hour assaults
## 817     2   4 2001          35  Sunday 11357    0        4
## 818     2   4 2001          35  Sunday 11357    1        6
## 819     2   4 2001          35  Sunday 11357    2       11
## 820     2   4 2001          35  Sunday 11357    3        6
## 821     2   4 2001          35  Sunday 11357    4        3
## 822     2   4 2001          35  Sunday 11357    5        2
##     assaults_domestic batteries batteries_domestic burglaries thefts
## 817                 1         7                  5          4     10
## 818                 0        22                  9          2      8
## 819                 2         6                  3          4      8
## 820                 1        15                  4          0      4
## 821                 0         7                  4          0      4
## 822                 1         7                  4          2      0
##     mv_thefts prostitution narcotics batteries_nodomestic
## 817        11            2         3                   2
## 818         4            0         2                  13
## 819         1            0         2                   3
## 820         0            0         4                  11
```

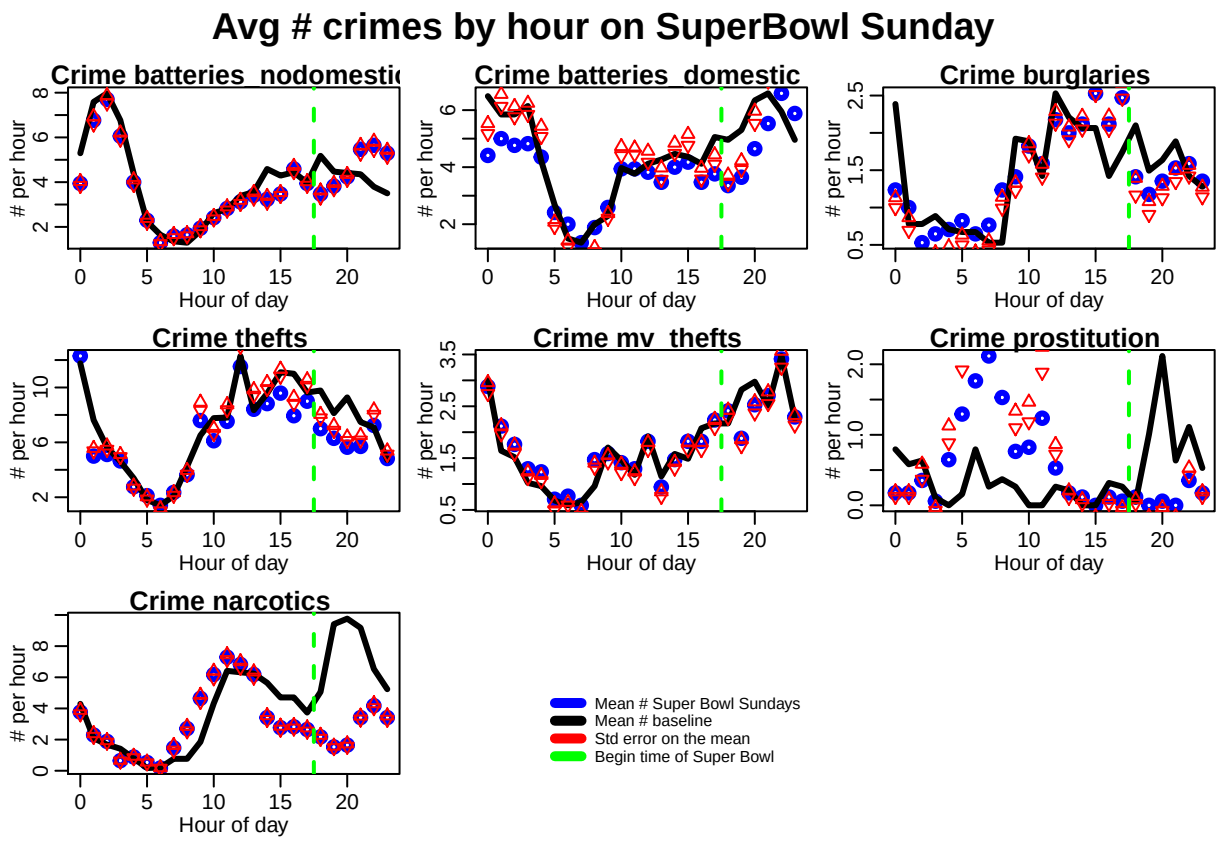
```
## 821         1            0         1                    3
## 822         0            0         0                    3
```

```
weekafter = subset(vdat,(jul+7)%in%sunday$jul)
#Selects the number of crimes a week after Super Bowl Sunday
head(weekafter)
```

```
##     month day year day_of_year weekday   jul hour assaults
## 481     1  21 2001          21  Sunday 11343    0        6
## 482     1  21 2001          21  Sunday 11343    1        4
## 483     1  21 2001          21  Sunday 11343    2        2
## 484     1  21 2001          21  Sunday 11343    3        3
## 485     1  21 2001          21  Sunday 11343    4        1
## 486     1  21 2001          21  Sunday 11343    5        1
##     assaults_domestic batteries batteries_domestic burglaries thefts
## 481                 2        16                 10          3     19
## 482                 1        12                  3          1      7
## 483                 1        13                  4          0      4
## 484                 0         7                  4          3      6
## 485                 0         7                  3          0      6
## 486                 1         4                  1          0      2
##     mv_thefts prostitution narcotics batteries_nodomestic
## 481         4            0         5                    6
## 482         6            0         2                    9
## 483         2            0         6                    9
## 484         0            0         1                    3
## 485         1            1         3                    4
## 486         1            0         0                    3
```

```
mult.fig(9,oma=c(0,0,3,1),mar=c(3,3,1,0),main="Avg \043 crimes by hour on SuperBowl Sunday")
# Opens a field to plot out the graphs

for (y in c("batteries_nodomestic",
            "batteries_domestic","burglaries",
"thefts","mv_thefts","prostitution" ,"narcotics")){
  # Loops through the crimes



  sig = (aggregate(sunday[,y],by=list(sunday$hour),FUN="sd",na.rm=TRUE)[,2])
  #Calculates the sigma or the standard deviation

  SCrime = (aggregate(sunday[,y],by=list(sunday$hour),FUN="mean",na.rm=TRUE)[,2])
  #Calculates the mean of the crimes on SuperBowl Sunday


  BCrime = (aggregate(weekbefore[,y],by=list(sunday$hour),FUN="mean")[,2])
  #Calculates the mean of the crimes the week beore SuperBowl Sunday

  ACrime =  (aggregate(weekafter[,y],by=list(sunday$hour),FUN="mean") [,2])
  #Calculates the mean of the crimes the week after SuperBowl Sunday

  averagecrime = ((BCrime + ACrime)/2)
  #Calculates the average of the crime means of the week before and the week after Superbowl Sunday
  SE = (sig/sqrt(408)) #Takes the Standar error
```

```
  ymin = min(c(SCrime, averagecrime))
  #Sets the min limit of the y axes
  ymax = max(c(SCrime, averagecrime))
  #Sets the max limit of the y axes

  plot(0:23,SCrime,xlab="Hour of day",ylab="\043 per hour",col = "blue",main=paste("Crime", y,
   split = ""),ylim=c(ymin,ymax), lwd=3)
  par(new=T)
  plot(0:23,averagecrime,type="l",col="black",axes=F,xaxt="n",yaxt="n",xlab="",ylab="",lwd = 3)
  #Plots the hours of the day(x) and the number of crimes commited per hour(y) for all crimes



  points(0:23,(SCrime-SE),col = "red", pch = 25)
  points(0:23,(SCrime+SE), col = "red", pch = 24)
  # Plots the standar deviation



abline(v=17.5, col="green", lty=2, lwd=2) #Sets the start of the superbowl
}

plot(0,axes=F,xaxt="n",yaxt="n",col=0,xlab="",ylab="",ylim=c(-15,15), xlim=c(-15,15))
legend("bottomright",
       legend = c("Mean # Super Bowl Sundays",
                  "Mean # baseline","Std error on the mean",
                  "Begin time of Super Bowl"),
       col=c("blue","black","red","green"),
       bty = "n",lwd=5,cex=0.7)

#Creates the legend for the graphs
```

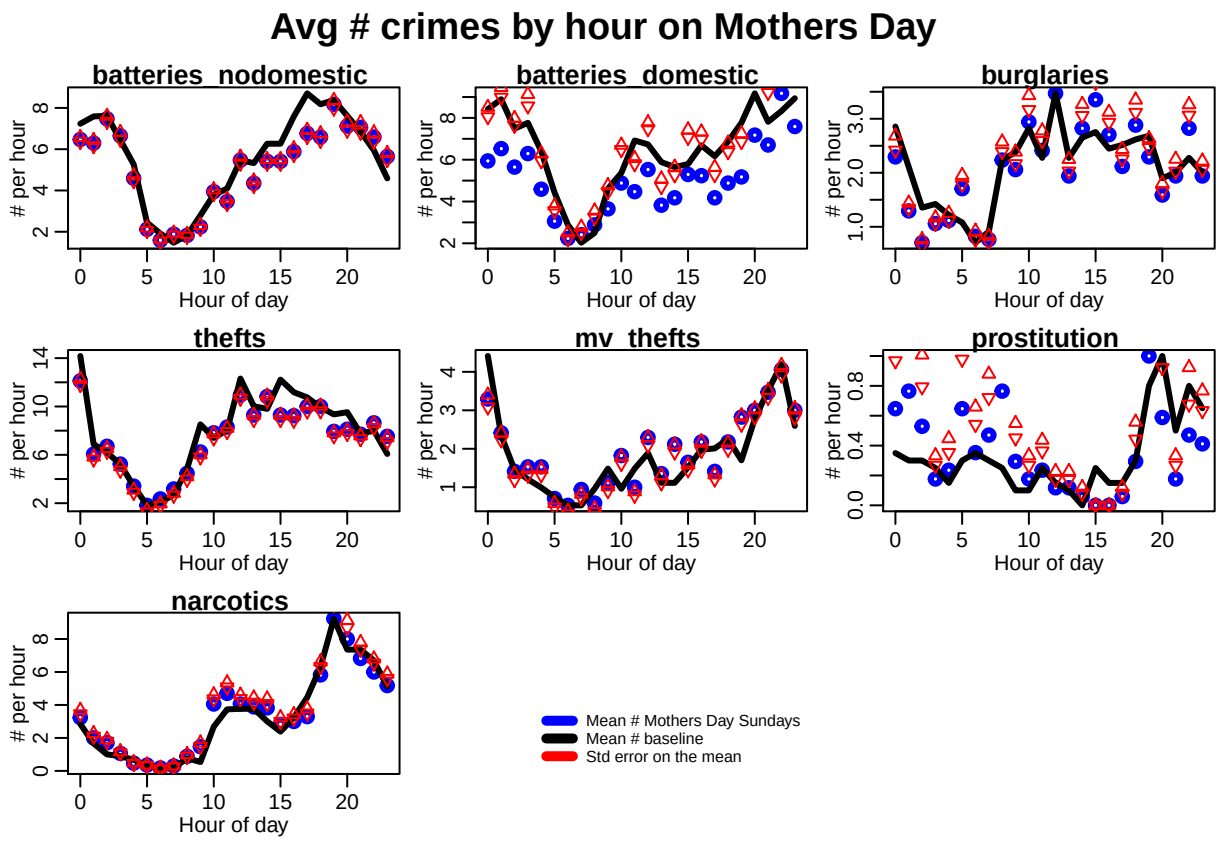



```
#########
#Part 1: Mothers Day.
#The following code is very similar to the previous one. The main difference
#is the Holiday we are considering
#and the data set of the holiday, in this case is Mothers Day
#
########

mothersday_data = read.csv("mothersday_Velazquez.csv",header=T,as.is=T,sep=",")
mothersday_data
```

```
##    Month Day Year
## 1      5  13 2001
## 2      5  12 2002
## 3      5  11 2003
## 4      5   9 2004
## 5      5   8 2005
## 6      5  14 2006
## 7      5  13 2007
## 8      5  11 2008
## 9      5  10 2009
## 10     5   9 2010
## 11     5   8 2011
## 12     5  13 2012
## 13     5  12 2013
## 14     5  11 2014
```

```
## 15     5  10 2015
## 16     5   8 2016
## 17     5  14 2017
```

```
mothersday_data$jul = julian(mothersday_data$Month,mothersday_data$Day,mothersday_data$Year)


Msunday = subset(vdat,jul%in%mothersday_data$jul)

Mweekbefore = subset(vdat,(jul-7)%in%mothersday_data$jul)

Mweekafter = subset(vdat,(jul+7)%in%mothersday_data$jul)

mult.fig(9,oma=c(0,0,3,1),mar=c(3,3,1,0),main="Avg \043 crimes by hour on Mothers Day")

for (y in c("batteries_nodomestic",
            "batteries_domestic",
            "burglaries","thefts","mv_thefts","prostitution" ,"narcotics")){


  Msig = (aggregate(Msunday[,y],by=list(Msunday$hour),FUN="sd",na.rm=TRUE)[,2])

  MCrime = (aggregate(Msunday[,y],by=list(Msunday$hour),FUN="mean",na.rm=TRUE)[,2])


  MBCrime = (aggregate(Mweekbefore[,y],by=list(Msunday$hour),FUN="mean")[,2])
  MACrime =  (aggregate(Mweekafter[,y],by=list(Msunday$hour),FUN="mean") [,2])

  Maveragecrime = (MBCrime + MACrime)/2
  Mymin = min(c(MCrime, Maveragecrime)) #Sets the min limit of the y axes
  Mymax = max(c(MCrime, Maveragecrime)) #Sets the max limit of the y axes


  plot(0:23,MCrime,xlab="Hour of day",ylab="\043 per hour",col = "blue",
       main=paste( y, split = ""),
       ylim=c(Mymin,Mymax), lwd=3)
  par(new=T)
  plot(0:23,Maveragecrime,type="l",col="black",axes=F,xaxt="n",yaxt="n",xlab="",ylab="",lwd = 3)

  MSE = (Msig/sqrt(408))


  points(0:23,(MCrime-MSE),col = "red", pch = 25)
  points(0:23,(MCrime+MSE), col = "red", pch = 24)

}

plot(0,axes=F,xaxt="n",yaxt="n",col=0,xlab="",ylab="",ylim=c(-10,10), xlim=c(-10,10))
legend("bottomright",
       legend = c("Mean # Mothers Day Sundays",
                  "Mean # baseline","Std error on the mean"
                  ),
       col=c("blue","black","red"),
       bty = "n",lwd=5,cex=0.7)
```




```
#########################################################################
#Part 2: Fathers Day.
#The following code is very similar to the SuperBowl Crime program. The main difference is
#the Holiday we are considering
#and the data set of the holiday,
#in this case is Fathers Day
#
#
#########################################################################
fathersday = read.csv("fathersday_Velazquez.csv")
fathersday
```

```
##    Month Day Year
## 1      5  18 2000
## 2      5  17 2001
## 3      5  16 2002
## 4      5  15 2003
## 5      5  20 2004
## 6      5  19 2005
## 7      5  18 2006
## 8      5  17 2007
## 9      5  15 2008
## 10     5  21 2009
## 11     5  20 2010
## 12     5  19 2011
## 13     5  17 2012
```

```
## 14     5  16 2013
## 15     5  15 2014
## 16     5  21 2015
## 17     5  19 2016
## 18     5  18 2017
```

```
fathersday$jul = julian(fathersday$Month,fathersday$Day,fathersday$Year)


Fsunday = subset(vdat,jul%in%fathersday$jul)

Fweekbefore = subset(vdat,(jul-7)%in%fathersday$jul)

Fweekafter = subset(vdat,(jul+7)%in%fathersday$jul)


mult.fig(9,oma=c(0,0,3,1),mar=c(3,3,1,0),main="Avg \043 crimes by hour on Fathers Day")

for (y in c("batteries_nodomestic",
            "batteries_domestic",
            "burglaries","thefts","mv_thefts","prostitution" ,"narcotics")){


  Fsig = (aggregate(Fsunday[,y],by=list(Fsunday$hour),FUN="sd",na.rm=TRUE)[,2])

  FCrime = (aggregate(Fsunday[,y],by=list(Fsunday$hour),FUN="mean",na.rm=TRUE)[,2])


  BFCrime = (aggregate(Fweekbefore[,y],by=list(Fsunday$hour),FUN="mean")[,2])
  AFCrime =  (aggregate(Fweekafter[,y],by=list(Fsunday$hour),FUN="mean") [,2])

  Faveragecrime = (BFCrime + AFCrime)/2

  Fymin = min(c(FCrime, Faveragecrime)) #Sets the min limit of the y axes
  Fymax = max(c(FCrime, Faveragecrime)) #Sets the max limit of the y axes



  plot(0:23,FCrime,xlab="Hour of day",ylab="\043 per hour",col = "blue",
       main=paste( y, split = ""),ylim=c(Fymin,Fymax), lwd=3)
  par(new=T)
  plot(0:23,Faveragecrime,type="l",col="black",axes=F,xaxt="n",yaxt="n",xlab="",ylab="", lwd=3)

  FSE = (Fsig/sqrt(408))


  points(0:23,(FCrime-FSE),col = "red", pch = 25)
  points(0:23,(FCrime+FSE), col = "red", pch = 24)

}

plot(0,axes=F,xaxt="n",yaxt="n",col=0,xlab="",ylab="",ylim=c(-10,10), xlim=c(-10,10))
legend("bottomright",
       legend = c("Mean # Fathers Day Sundays",
```

```
                "Mean # baseline","Std error on the mean"),
       col=c("blue","black","red"),
       bty = "n",lwd=5,cex=0.7)
```