

Sanfrancisco at summer 2014: August has more Criminal at Weekend

Q

Sunday, November 29, 2015

This is a practice analysis for coursera lesson: Communicating Data Science Result.

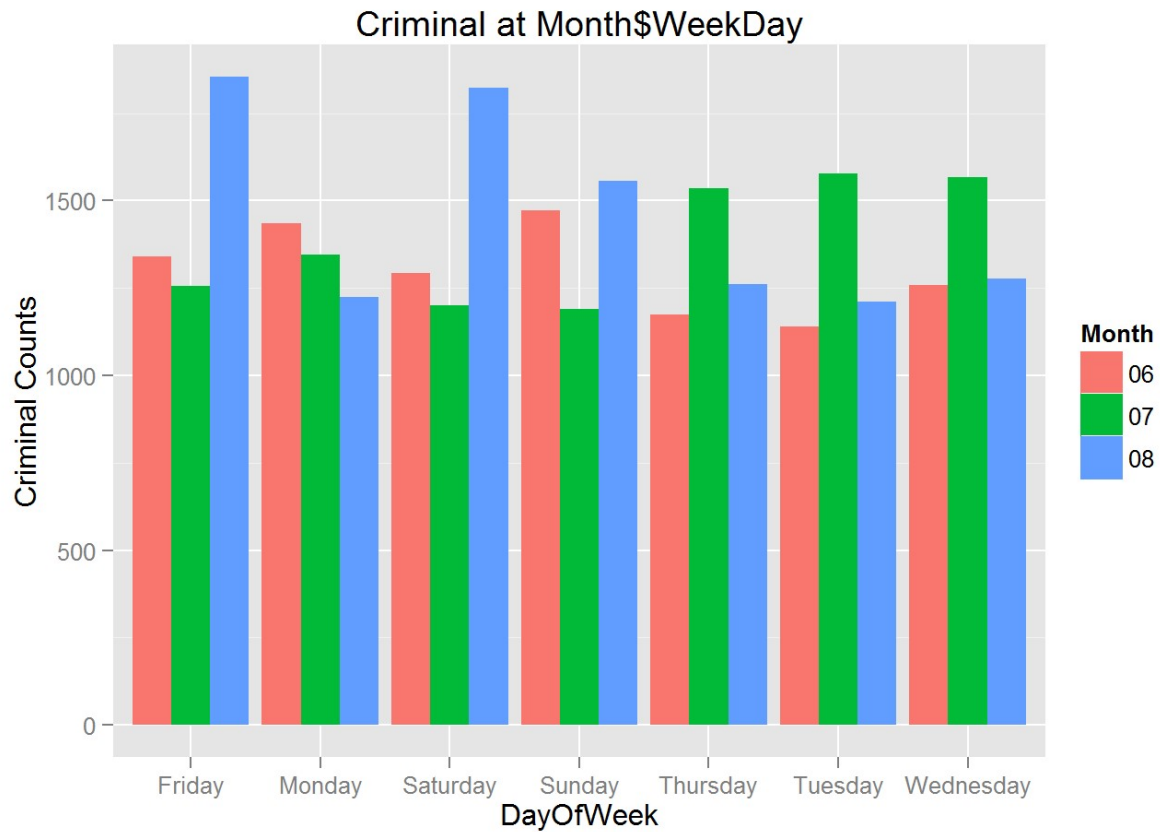
Prepare data:

```
library(ggplot2)
setwd('C:/Users/QQ/Documents/GitHub/datasci/datasci_course_materials/assignment6')
san_data <- read.csv('sanfrancisco_incidents_summer_2014.csv')
san_data$FTime <- paste(san_data$Date, san_data$Time)
san_data$FTime <- strptime(san_data$FTime, '%m/%d/%Y %H:%M')
san_data$Month <- format(san_data$FTime, format='%m')
san_data$Hour <- format(san_data$FTime, format='%H')
head(san_data)
```

```
##      IncidntNum      Category      Descript DayOfWeek
## 1  140734311      ARSON      ARSON OF A VEHICLE      Sunday
## 2  140736317  NON-CRIMINAL      LOST PROPERTY      Sunday
## 3  146177923  LARCENY/THEFT  GRAND THEFT FROM LOCKED AUTO      Sunday
## 4  146177531  LARCENY/THEFT  GRAND THEFT FROM LOCKED AUTO      Sunday
## 5  140734220  NON-CRIMINAL      FOUND PROPERTY      Sunday
## 6  140734349  DRUG/NARCOTIC  POSSESSION OF MARIJUANA      Sunday
##      Date  Time PdDistrict      Resolution      Address
## 1 08/31/2014 23:50    BAYVIEW      NONE  LOOMIS ST / INDUSTRIAL ST
## 2 08/31/2014 23:45    MISSION      NONE    400 Block of CASTRO ST
## 3 08/31/2014 23:30    SOUTHERN      NONE  1000 Block of MISSION ST
## 4 08/31/2014 23:30    RICHMOND      NONE    FULTON ST / 26TH AV
## 5 08/31/2014 23:23    RICHMOND      NONE  800 Block of LA PLAYA ST
## 6 08/31/2014 23:13    SOUTHERN  ARREST, BOOKED    11TH ST / MINNA ST
##      X      Y      Location      PdId
## 1 -122.4056 37.73832 (37.7383221869053, -122.405646994567) 1.407343e+13
## 2 -122.4350 37.76177 (37.7617677182954, -122.435012093789) 1.407363e+13
## 3 -122.4098 37.78004 (37.7800356268394, -122.409795194505) 1.461779e+13
## 4 -122.4853 37.77252 (37.7725176473142, -122.485262988324) 1.461775e+13
## 5 -122.5099 37.77231 (37.7723131976814, -122.509895418239) 1.407342e+13
## 6 -122.4166 37.77391 (37.773907074489, -122.416578493475) 1.407343e+13
##      FTime Month Hour
## 1 2014-08-31 23:50:00    08    23
## 2 2014-08-31 23:45:00    08    23
## 3 2014-08-31 23:30:00    08    23
## 4 2014-08-31 23:30:00    08    23
## 5 2014-08-31 23:23:00    08    23
## 6 2014-08-31 23:13:00    08    23
```

Following plot shows that weekend (Friday, Saturday and Sunday) of August has more criminal incident than other time period. Possibly the weather is hot and people has more time at weekend.

```
ggplot(san_data) + geom_bar(aes(x=DayOfWeek,fill=Month),position='dodge') +
labs(y='Criminal Counts',title='Criminal at Month$WeekDay')
```



Following plot shows that there are more criminal incident especially at midday (12:00), compare with 11:00 and 13:00. Possibly it's because the temperature is especially high. And 4:00~5:00 in the morning has less criminal, because most of people now goes to rest. However, the happened criminal could be serious ones, which worth further analysis.

```
ggplot(san_data) + geom_bar(aes(x=Hour,fill=Month)) + labs(y='Criminal Counts',title='Criminal at Month$WeekDay')
```

