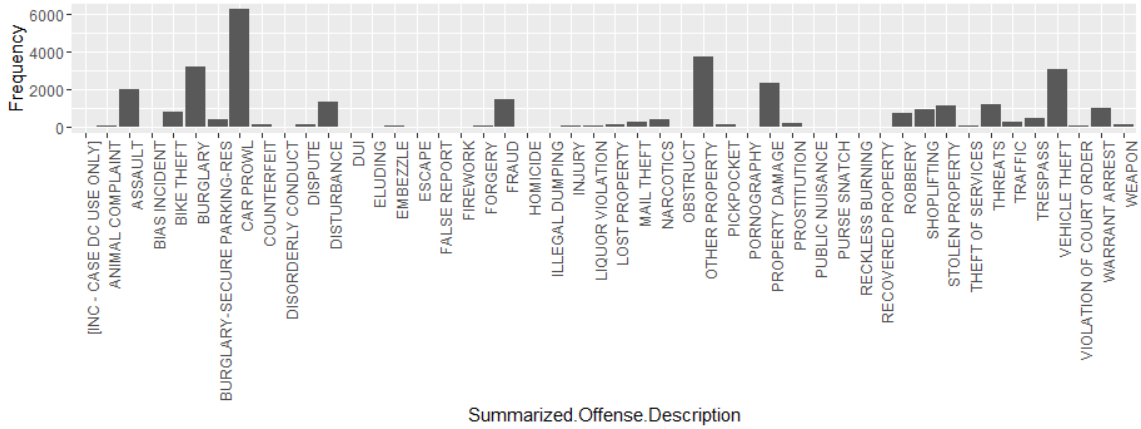
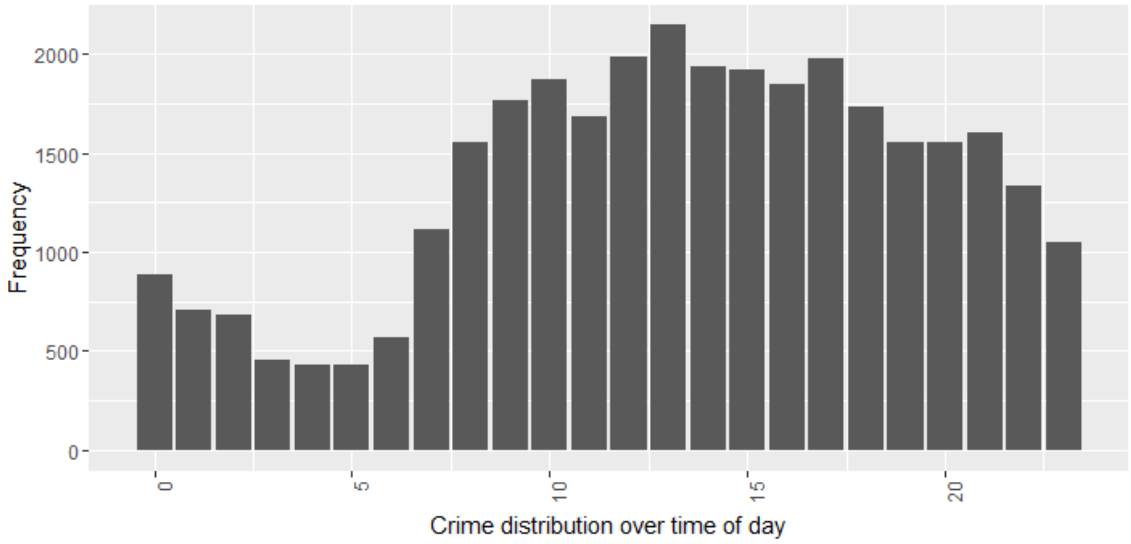
Crime Analytics: Visualization of Incident Reports

(1) Summary of the Seattle_incident_summer_2014 (***Car prowl***)



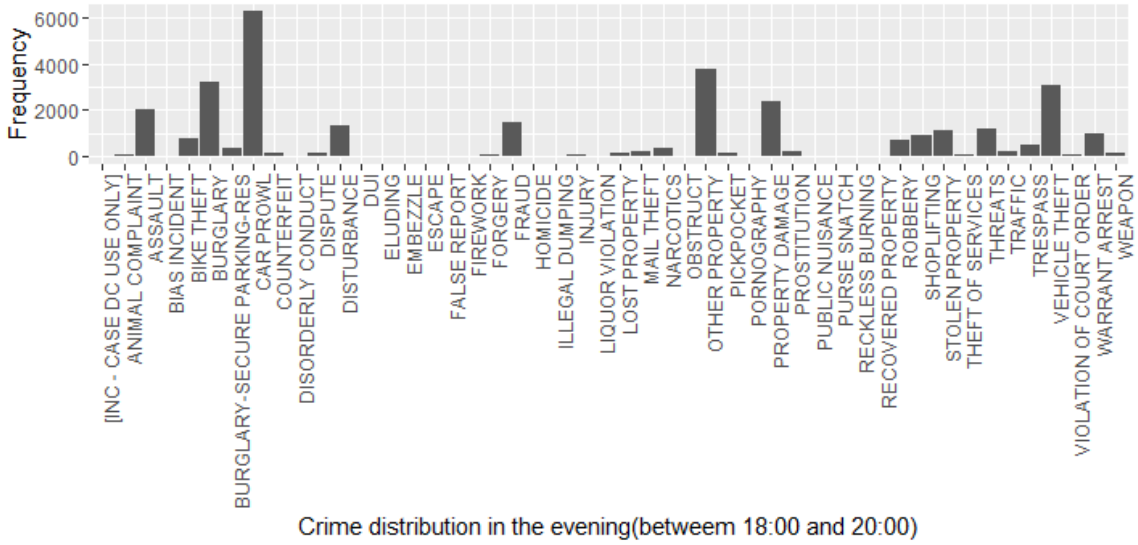
The above histogram figure shows that CAR PROWL is the most frequent crime.

(2) Crime distribution over time (***13***)



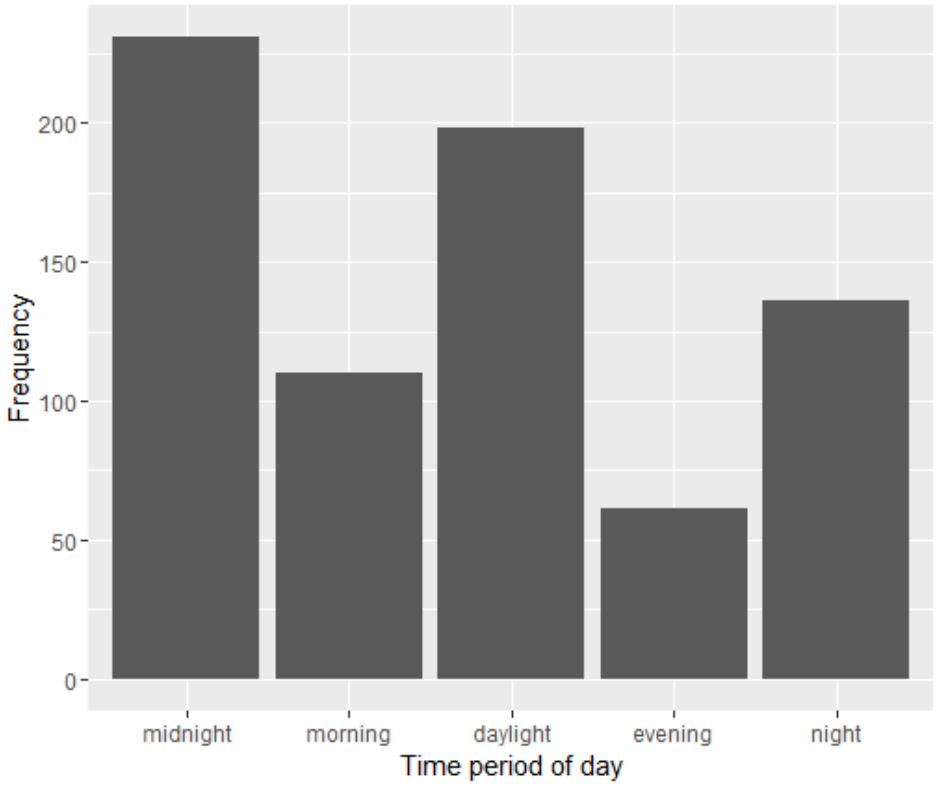
The figure above shows the crime distribution over time of day. The hourly crime frequency is relatively high during the daytime and evening from 8:00 to 21:00. The majority of crimes have been committed during broad daylight.

(3) Crime distribution in the evening (***Car Prowl***)



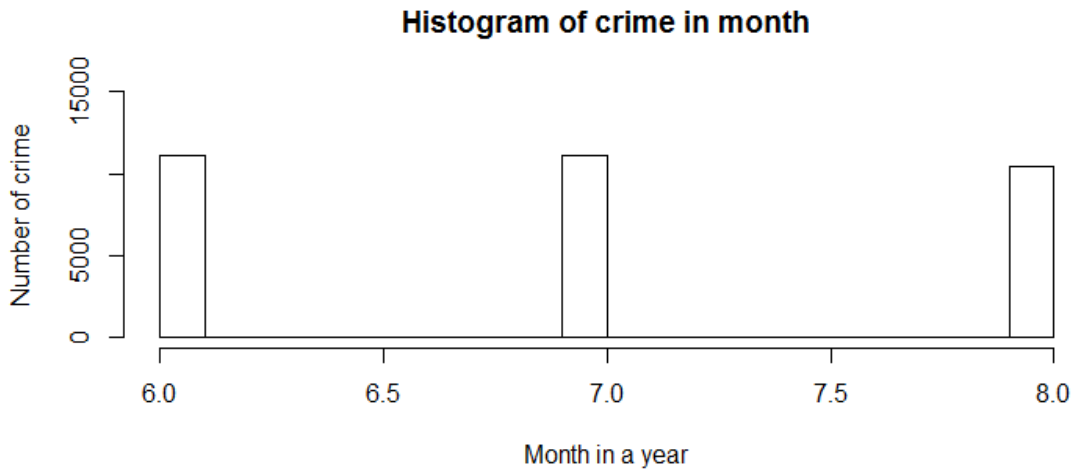
The figure above shows the crime distribution in the evening, which shows the most frequency crime type is CAR PROWL.

(4) Robbery distribution in time period of day (***Midnight***)



The above histogram shows that during midnight robberies are most common.

(5) Crime distribution over month (***Similar over 6,7,8***)



The above figure shows the crime distribution over month in the summer. Here we can see that the crime frequencies in June, July and August are very close, all of which are above 10,000.

Appendix 1: R code for Seattle_incident_summer_2014

```
library('ggplot2')

sea=read.csv(file.choose(),head=T)
# plot the frequency of each crime
sort(table(sea$Summarized.Offense.Description), decreasing=T)
xm=qplot(Summarized.Offense.Description, data=sea, StatSmooth='count')
xm+theme(axis.text.x = element_text(angle = 90, hjust = 1))+labs(y='Frequency')

# the frequency of crimes over time
xx <- as.POSIXct(sea$Date.Reported, format="%m/%d/%Y %I:%M:%S %p")
xy=as.POSIXlt(xx);
hour=xy$hour;
hour.df=as.data.frame(hour)
xh=ggplot(hour.df, aes(hour))
xh+geom_bar()+theme(axis.text.x = element_text(angle = 90, hjust = 1))+labs(x='Crime
distribution over time of day', y='Frequency')

# distribution in evening
seaNew=cbind(sea, hour)
subEvening=subset(seaNew,hour>=18 & hour<=20, select =
Summarized.Offense.Description)
C=sort(table(subEvening$Summarized.Offense.Description),decreasing = T)
xe=qplot(Summarized.Offense.Description, data=subEvening, StatSmooth='count')
xe+theme(axis.text.x = element_text(angle = 90, hjust = 1))+labs(x='Crime distribution in
the evening(betweem 18:00 and 20:00)',y='Frequency')

# distribution of robbery
seaNew=cbind(sea, hour)
subRob=subset(seaNew,Summarized.Offense.Description=='ROBBERY', select
=c(Summarized.Offense.Description, hour))
in the evening(betweem 18:00 and 20:00)',y='Frequency')
subhour=subRob[,2]
subhourd=as.data.frame(subhour)
xr=ggplot(data=subhourd,aes(subhourd$subhour))
xr+geom_histogram(breaks=c(0,6,12,18,20,24), col="red")
subhour=subRob[,2]

breaks <- c(0,6,  12,  18, 20, 24) # times are internally fractions of a day
labels <- c("midnight","morning", "daylight", "evening","night")
subhd=cut(subhour, breaks, labels, include.lowest = TRUE)
```

```
subdd.rob=as.data.frame(subhd)

#  distribution over month
Month=sea$Month
xname='crime in month'
hist(Month, ylim=c(0,15000),xlab = 'Month in a year', ylab='Number of crime', main =
paste("Histogram of" , xname))
```