# Project

*Aaron*

*3/25/2018*

```r
library(plyr)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:plyr':
##
##     arrange, count, desc, failwith, id, mutate, rename, summarise,
##     summarize

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(magrittr)
library(reshape2)
library(ggplot2)
library(lubridate)
```

```
##
## Attaching package: 'lubridate'

## The following object is masked from 'package:plyr':
##
##     here

## The following object is masked from 'package:base':
##
##     date
```

```r
a <- read.csv(file="/Users/aaron/Desktop/R/Project/TrafficCollision.csv", header=TRUE, sep=",")

count(a, Victim.Sex, Date.Occurred)
```

```
## # A tibble: 10,528 x 3
##    Victim.Sex Date.Occurred     n
##        <fctr>        <fctr> <int>
## 1              01/01/2010     4
## 2              01/01/2011     1
## 3              01/01/2013     2
## 4              01/01/2014     4
## 5              01/01/2015     2
## 6              01/01/2016     1
## 7              01/02/2010     1
## 8              01/02/2011     1
## 9              01/02/2012     1
```

```
## 10                    01/02/2013      1
## # ... with 10,518 more rows
Male <<- subset(count(a, Victim.Sex, Date.Occurred), Victim.Sex == "M")
Female <<- subset(count(a, Victim.Sex, Date.Occurred), Victim.Sex == "F")
num <- inner_join(Male, Female, by = "Date.Occurred")

d <- data_frame(date = as.Date(num$Date.Occurred, format = "%m/%d/%Y"),
                Number.of.Male = num$n.x,
                Number.of.Femlae = num$n.y)

dmelt <- melt(d, id.vars = "date", variable.name = "type")

ggplot(dmelt)+
aes(date, value, colour = type)+
geom_line()+
theme(
  plot.title = element_text(size = 13, hjust = 0.5, face = "bold"),
  axis.title = element_text(size = 12, face = "bold"),
  strip.text.x = element_text(size = 12, face = "bold"),
  legend.title = element_text(size = 8),
  legend.key.size = unit(0.5, "cm"),
  legend.text = element_text(size = 6)
)+
labs(
  x = "Date",
  y = "Number of Incidents Per Day",
  shape = "Transmission"
)+
scale_colour_manual(
  values = c("darkblue", "orange"),
  name = "Gender",
  labels = c("Male", "Female")
)+
ggtitle("Time Series Plot of Traffic Collision Incidents from 2010")
```
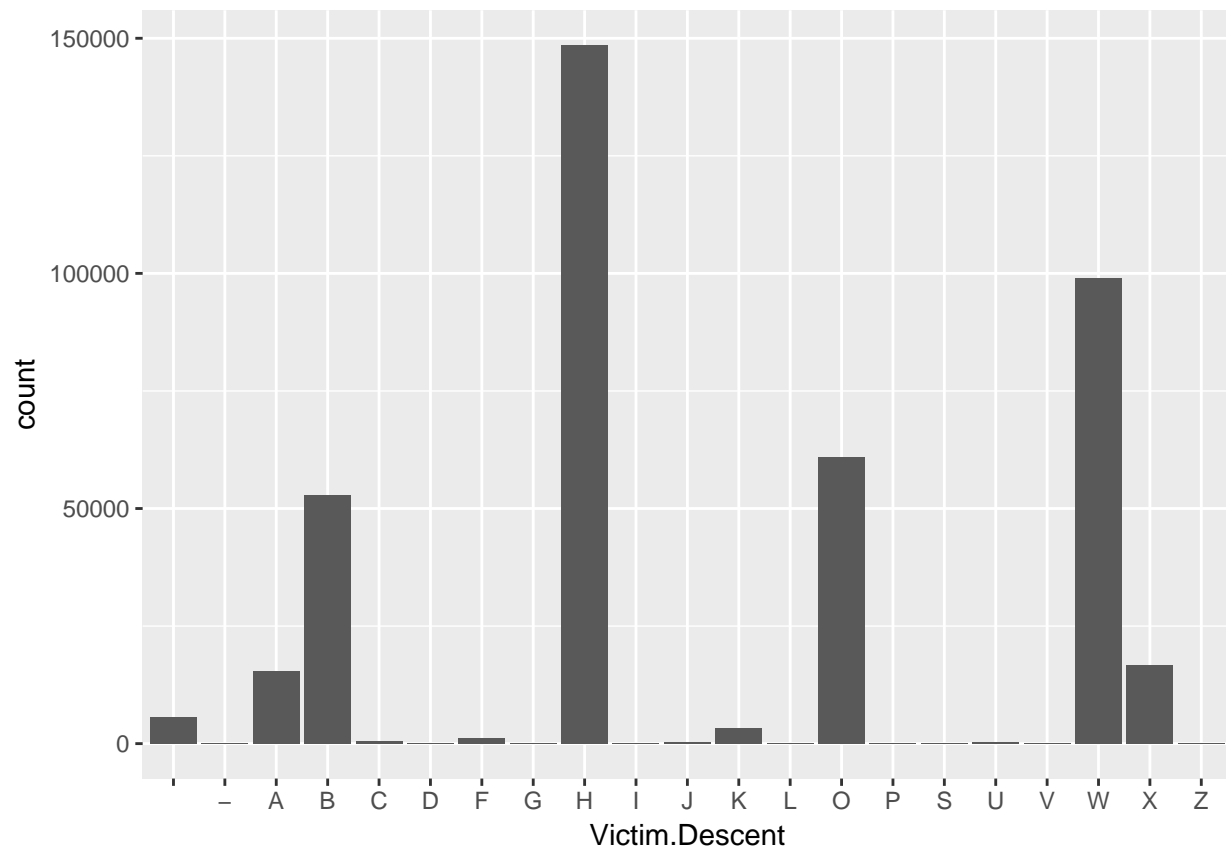
**Time Series Plot of Traffic Collision Incidents from 2010**
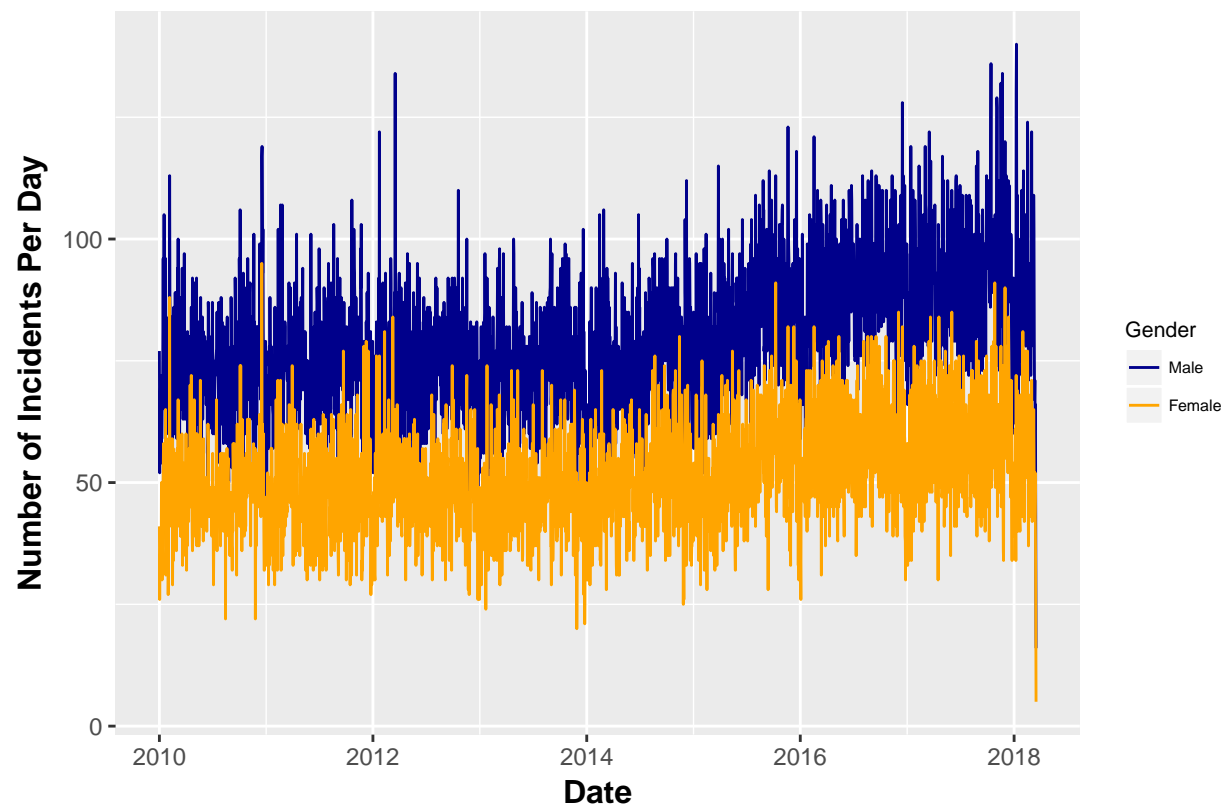
```r
ggplot(a) +
aes(Victim.Descent) +
geom_bar()
```

```
youth<-within(filter(a, a$Victim.Age <= 30),{Time<- floor(Time.Occurred/100)})
youth1<-count(youth, Time)

strong<-within(filter(a, Victim.Age > 30 & Victim.Age < 60),{Time<- floor(Time.Occurred/100)})
strong1<-count(strong, Time)

old <-within(filter(a, Victim.Age >= 60),{Time<- floor(Time.Occurred/100)})
old1<-count(old, Time)

AGEH<-data.frame(Time = youth1$Time,
                 Youth.Number = youth1$n,
                 Strong.Number = strong1$n,
                 Old.Nimber = old1$n)
MAGEH <- melt (AGEH, id.vars = "Time", variable.name = "age")
ggplot(MAGEH) +
aes(Time, value, colour = age) +
geom_line()+
scale_colour_manual(values = c("darkblue","red","darkgreen"))
```
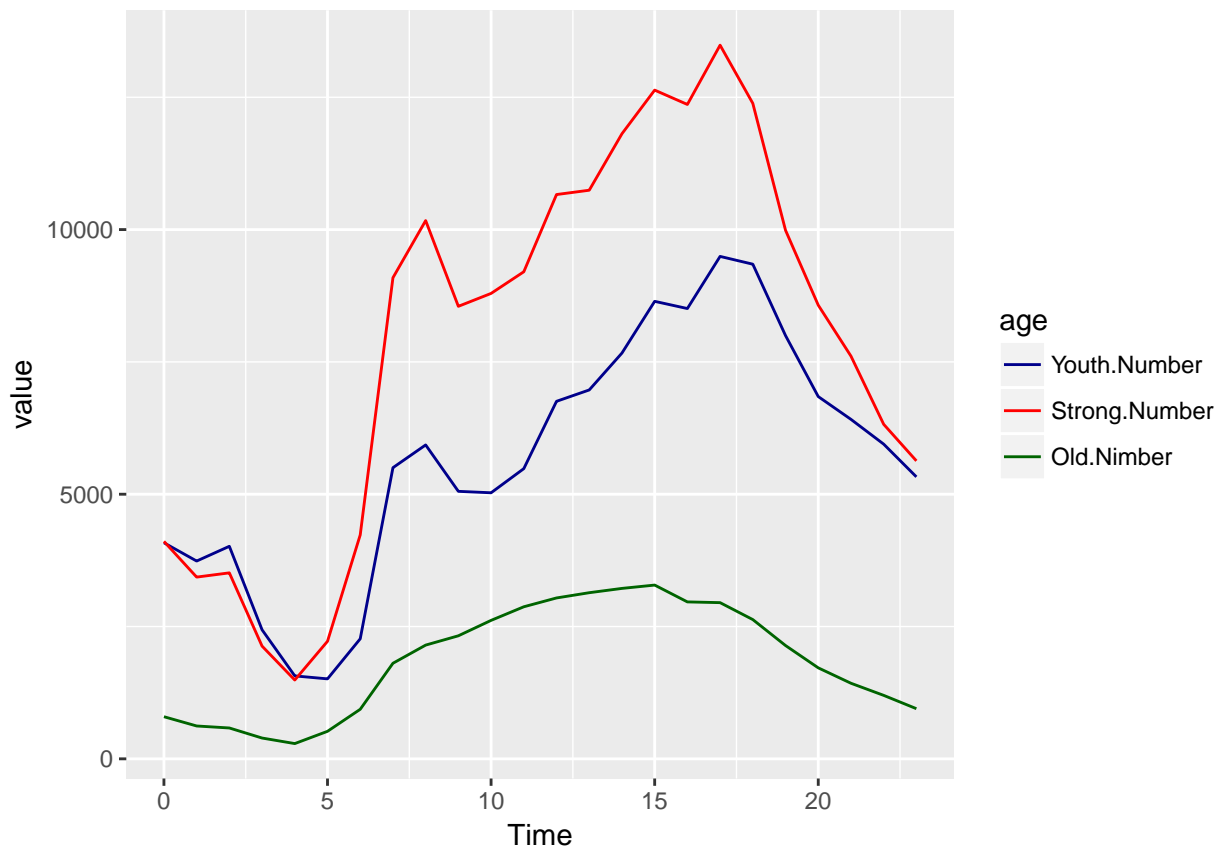
```
TCS<- filter(a, Victim.Sex !="" , Victim.Sex != "H", Victim.Sex !="X" )
h<-hour(as.POSIXct(sprintf("%04.0f", TCS$Time.Occurred), format="%H%M"))
##Bar graph
ggplot(TCS, aes(h, fill = Victim.Sex))+ geom_bar()
```