Assignment 2

igal tsirulnik, Yoni Elhanati

April 11, 2016

## Description of the dataset-

data contained in this set pertain to vehicular accident and crash locations from 2010 to 2014 in the District of Columbia. Data is provided by the District of Columbia's Metropolitan Police Department (MPD) along with the Department of Transportation (DDOT) and the University of Colorado at Boulder. It is a partial data and conaints the following fields: LATITUDE| LONGITUDE| ACCIDENTDATE YEAR| DAYOFWEEK| ACCIDENTTYPE| FEETFROMINTERSECTION| ROADSURFACE| ROADCONDITION| STREETLIGHT| LIGHTCONDITION| |EATHER| TRAFFICCONDITION

## Here you can see a small tase of the dataset-

f<- read.csv('C:\\Users\\igal\\Desktop\\data\\v\_smal.csv')  
head(f,5)

## LATITUDE LONGITUDE ACCIDENTDATE YEAR DAYOFWEEK ACCIDENTTYPE  
## 1 38.90485 -77.06614 31-Oct-12 2012 Wednesday Prop. Damage  
## 2 38.91712 -77.04189 23-Oct-10 2010 Saturday Prop. Damage  
## 3 38.91204 -77.04525 21-Aug-10 2010 Saturday Injury  
## 4 38.89431 -77.01654 25-May-13 2013 Saturday Injury  
## 5 38.83598 -76.99574 6-Jul-11 2011 Wednesday Injury  
## FEETFROMINTERSECTION ROADSURFACE ROADCONDITION STREETLIGHT  
## 1 5 Asphalt Dry Street Lights Off  
## 2 75 Concrete Dry Street Lights On  
## 3 50 Asphalt Dry Street Lights Off  
## 4 0 Concrete Dry   
## 5 0 Asphalt Dry Street Lights Off  
## LIGHTCONDITION WEATHER TRAFFICCONDITION  
## 1 Daylight Clear Medium  
## 2 Dark (Lighted) Clear Medium  
## 3 Daylight Clear Light  
## 4 Daylight Other Light  
## 5 Daylight Clear Light

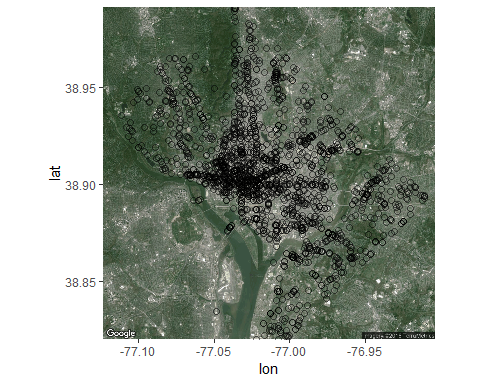
## Dataset on a map-

f<- read.csv('C:\\Users\\igal\\Desktop\\data\\v\_smal.csv')  
# getting the map  
map <- get\_map(location = c(lon = mean(f$LONGITUDE), lat = mean(f$LATITUDE)), zoom = 12,maptype = "satellite", scale = 2)

## Map from URL : http://maps.googleapis.com/maps/api/staticmap?center=38.906093,-77.013928&zoom=12&size=640x640&scale=2&maptype=satellite&language=en-EN&sensor=false

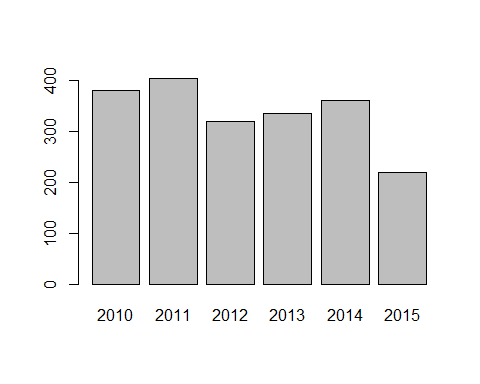
# plotting the map with some points on it  
ggmap(map) +  
 geom\_point(data = f, aes(x = LONGITUDE, y = LATITUDE , alpha = 0), size = 2, shape = 21) + guides(fill=FALSE, alpha=FALSE, size=FALSE)

## Warning: Removed 4 rows containing missing values (geom\_point).



## analysis of accident by year -

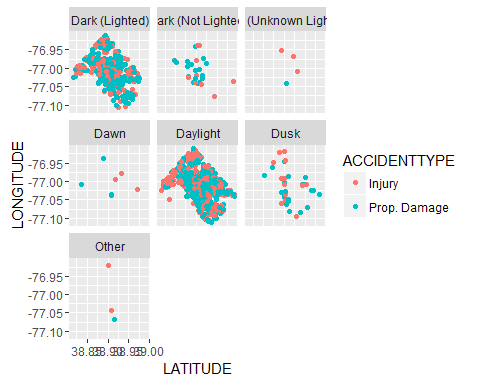
f<- read.csv('C:\\Users\\igal\\Desktop\\data\\v\_smal.csv')  
barplot(table(f$YEAR))



#we can see that after a slow rise in accidents 2015 was the year where they went down for the first time

## analysis of sevitity of accident by road light condition-

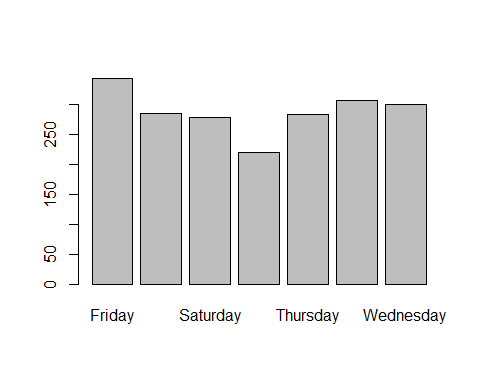
f<- read.csv('C:\\Users\\igal\\Desktop\\data\\v\_smal.csv')  
qplot(LATITUDE, LONGITUDE, data=f, color = ACCIDENTTYPE,facets = ~LIGHTCONDITION)



# we can see that accidents that happen with problematic light condition preduce more injeries than other accidents in precentage

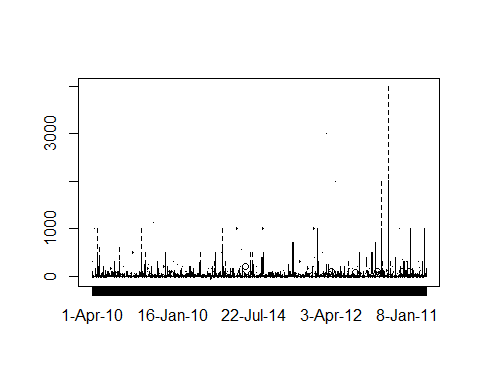
## analysis of accident by year day of the week-

f<- read.csv('C:\\Users\\igal\\Desktop\\data\\v\_smal.csv')  
barplot(table(f$DAYOFWEEK))



#as expected sunday is the day with the lest accidents  
  
## analysis of sevitity of accident by roud light condition-

f<- read.csv('C:\\Users\\igal\\Desktop\\data\\v\_smal.csv')  
with(f,plot(ACCIDENTDATE,FEETFROMINTERSECTION))



#almost all of the accidents happen within 10 feet from intersections

## Summury

Most of the accidents happen near inter section, we should be concentrating our efforts of providing the bes?? light and roud conditon around those areas