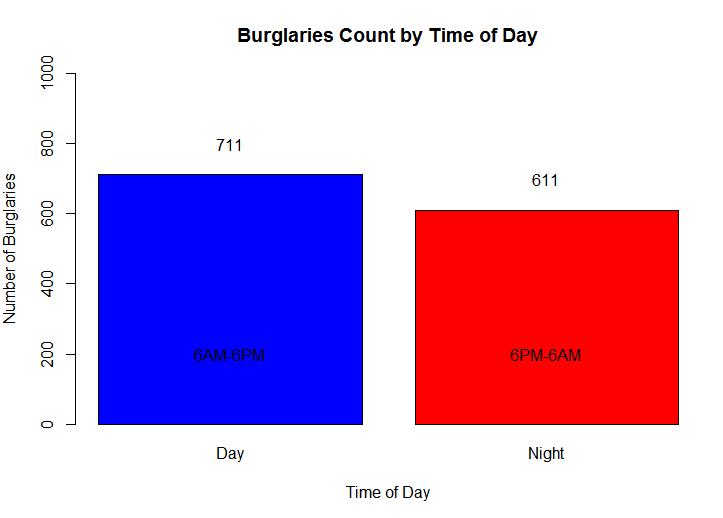
**DRAFT R PLOT**

**Question 1:** Are more 911 calls related to burglary made during the night? Or is that just a preconceived notion?



This graph displays the count of burglaries made during the day time (6AM – 6PM) and the night time (6PM – 6AM) in the city of Baltimore. The burglaries considered here consist of entries from the 911 emergency calls database related to burglaries such as ‘Burglary’, ‘Attempted Theft’, ‘Possible Stolen’ etc. To create the graph, I created a subset of all the burglary related calls and their date/time. I converted the call-date column from the database into a time format appropriate to perform computations in R. I then added a new column ‘Time\_Of\_Day’. If the time was between 6PM-6AM, I entered the value “Night”, into the column. Otherwise, I entered the value “Day”.

*finalburglary$Time\_Of\_Day[finalburglary$CallDateTime < '06:00:00'] <- 'Night'*

*finalburglary$Time\_Of\_Day[finalburglary$CallDateTime >= '18:00:00'] <- 'Night'*

I got the count of burglaries during the Day and the Night, as 711 and 611 respectively, out of 1322 total.

*counts<- table(finalburglary$Time\_Of\_Day)*

I used a bar plot to plot the ‘Time\_Of\_Day’ column, to see the difference in the count for Day and Night. I labeled the X and Y axis, and gave different colors for Day and Night. I then added text values to display the count and the timings for Day and Night

*counts<- table(finalburglary$Time\_Of\_Day)*

*barplot(counts, main = "Burglaries Count by Time of Day", xlab = "Time of Day", ylab = "Number of Burglaries", ylim = c(0, 1000), col = c("blue", "red"))*

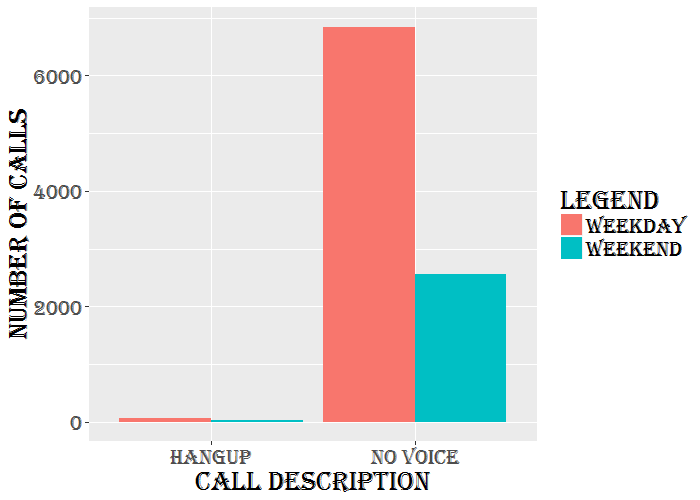
*text(0.7, 800, "711")*

*text(1.9, 700, "611")*

*text(0.7, 200, "6AM-6PM")*

*text(1.9, 200, "6PM-6AM")*

**Question 2:** Whether the number of blank calls increase/decrease with the arrival of the weekend?



The above ggplot shows the number of “NO VOICE” and “HANGUP” calls that occur during the weekends and weekdays. This was done by first splitting the date from the “CallDateTime” and converting it into date format. Then the “isWeekend” command from the “timeDate” library is used to categorize all the dates into weekdays and weekends. We then use the “ggplot”.

library(ggplot2)

*library(xkcd)*

*windowsFonts(F = windowsFont('Algerian'))* #Create a custom font type*.*

*ggplot(q,aes(x=Description,y=Count,fill=factor(Index)))+* #Specify values for dataset to be considered and the input values for the x and y axes.

*geom\_bar(stat="identity",position="dodge")+* #Specify bar type and the position of the bars relative to each other.

*scale\_fill\_discrete(name="Legend",* #Specify Legend to be displayed

*breaks=c(1, 2),*

*labels=c("Weekday", "Weekend"))+*

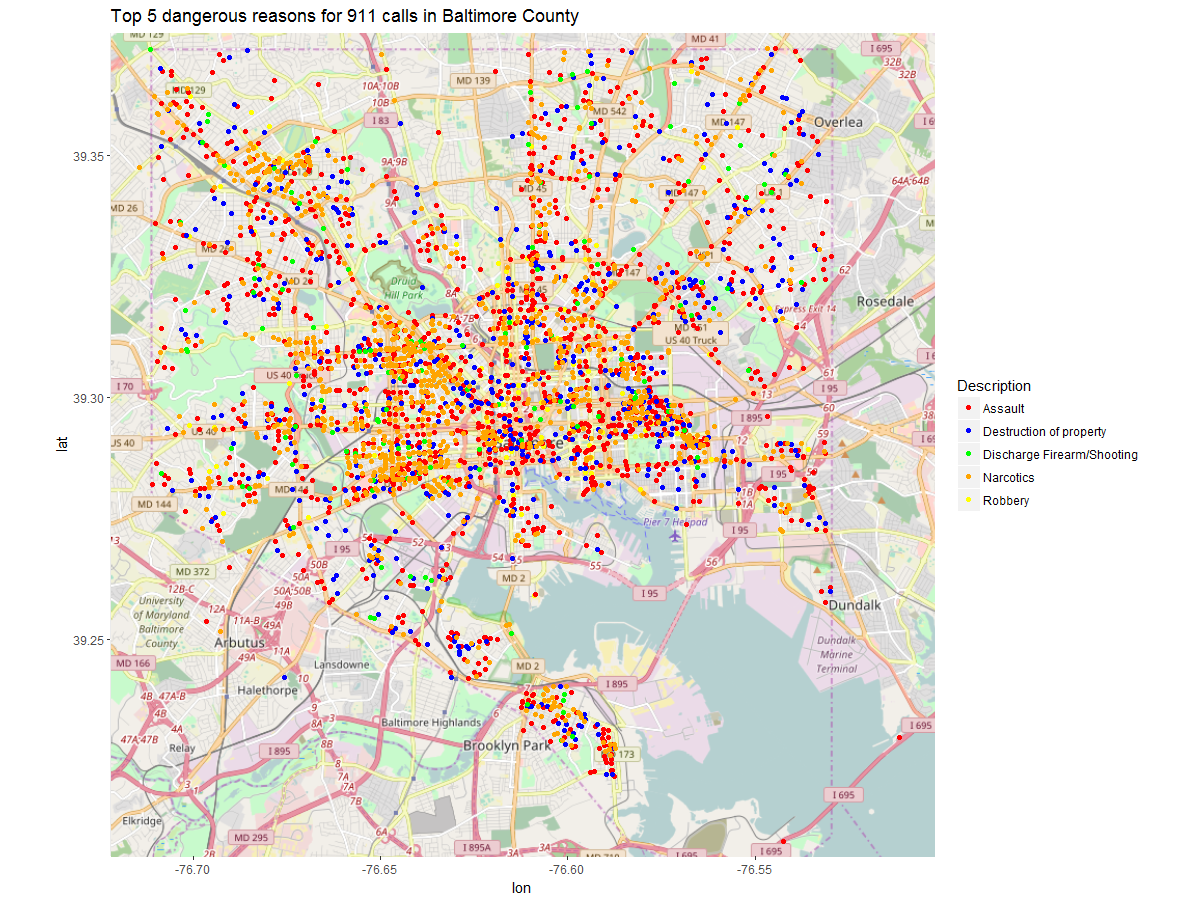
*xlab("Call Description")+* #Specify label for x axis

*ylab("Number of Calls")+* #Specify label for y axis

*theme(text=element\_text(family="F", size = 20))*  #Call the custom font type created before and assign it to the text. The font size is also specified.

**Question 3:** Find the top 5 dangerous reasons for calls (e.g. Aggravated assault). Plot them based on the locations.

**Using open street maps**



# get\_map is a smart wrapper that queries the Google Maps, OpenStreetMap, Stamen Maps or

# Naver Map servers for a map. It returns a ggmap object. Source attribute tells us the type of map

# to be used. By defaut, it uses Google map.

# geom\_point is used to plot points on the map. It is generally used to create scatter plots

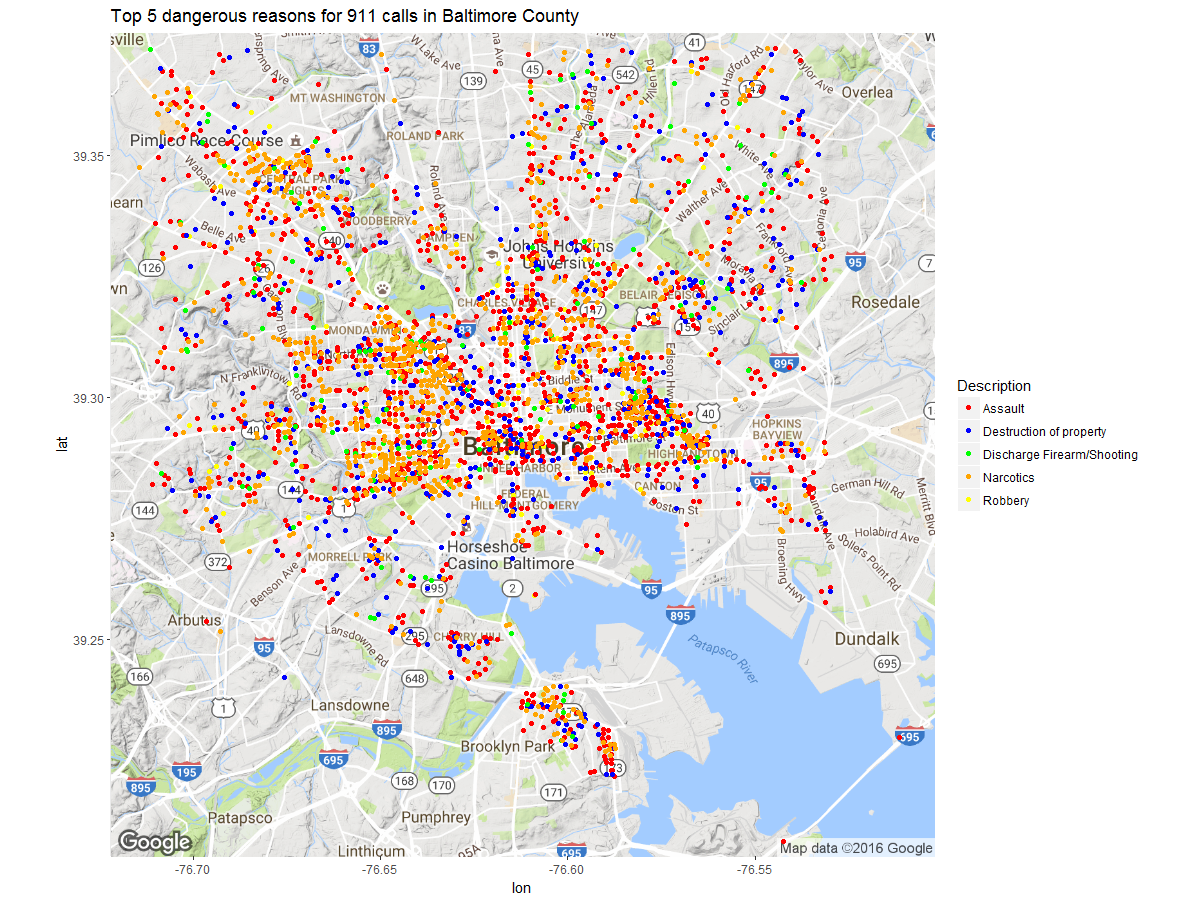
# scale\_color\_manual is used to map color values

# ggtitle For giving a title to the plot

*Use library(ggmap)*

*ggmap(get\_map(location = "baltimore",source="osm",zoom=12)) + geom\_point(data=subsetcalls,aes(x=subsetcalls$Longitude,y=subsetcalls$Latitude,colour=Description)) + scale\_color\_manual(values=c("red","blue","green","orange","yellow")) + ggtitle("Top 5 dangerous reasons for 911 calls in Baltimore County")*

**Using google maps**



*ggmap(get\_map(location = "baltimore",zoom=12)) + geom\_point(data=subsetcalls,aes(x=subsetcalls$Longitude,y=subsetcalls$Latitude,colour=Description)) + scale\_color\_manual(values=c("red","blue","green","orange","yellow")) + ggtitle("Top 5 dangerous reasons for 911 calls in Baltimore County")*

**Description**

The above plot represents the origin location of 911 calls (based on top 5 dangerous reasons) in Baltimore County for our sample. We have selected the top 5 reasons (‘Description’ column in our sample dataset) and used the latitude and longitude values to plot the locations from where these calls have originated. The plot gives us an overview of the dangerous localities in Baltimore based on the 911 calls. It shows that Assault (common+aggravated) and Narcotics are the most widely reported crimes in Baltimore, followed by Destruction of property, Discharge Firearm/Shooting and Robbery. The plot tells us that the concentration of 911 calls is more towards the center of Baltimore and reduces as we move towards the boundaries. It would be interesting to analyze the districts that fall in this area and study the socio-economic factors for this region to understand the high concentration of 911 calls. Also there are some areas, like the area near Brooklyn Park (bottom of the map) and Hopkins Bayview (right side of the map), that can be studied to understand the specific crime statistics there. The coastal area near the center (Inner Harbor) have more reported crime as compared to the remaining coastal areas. It would be exciting to find some patterns there relating to crime and proximity of coast. Finally we would like to plot the statistics for different samples of our dataset and see whether the observations change or more or less remain the same.