# Incidents involving booked juveniles are most common in the Bayview area of San Francisco

This document investigates the distribution of incidents in which a juvenile was booked thoroughout San Francisco. The code used to generate the plots can be found at the end of the document so as to avoid breaking up the discussion.

First, here is a plot showing the ditribution of crimes across different areas of San Francisco.
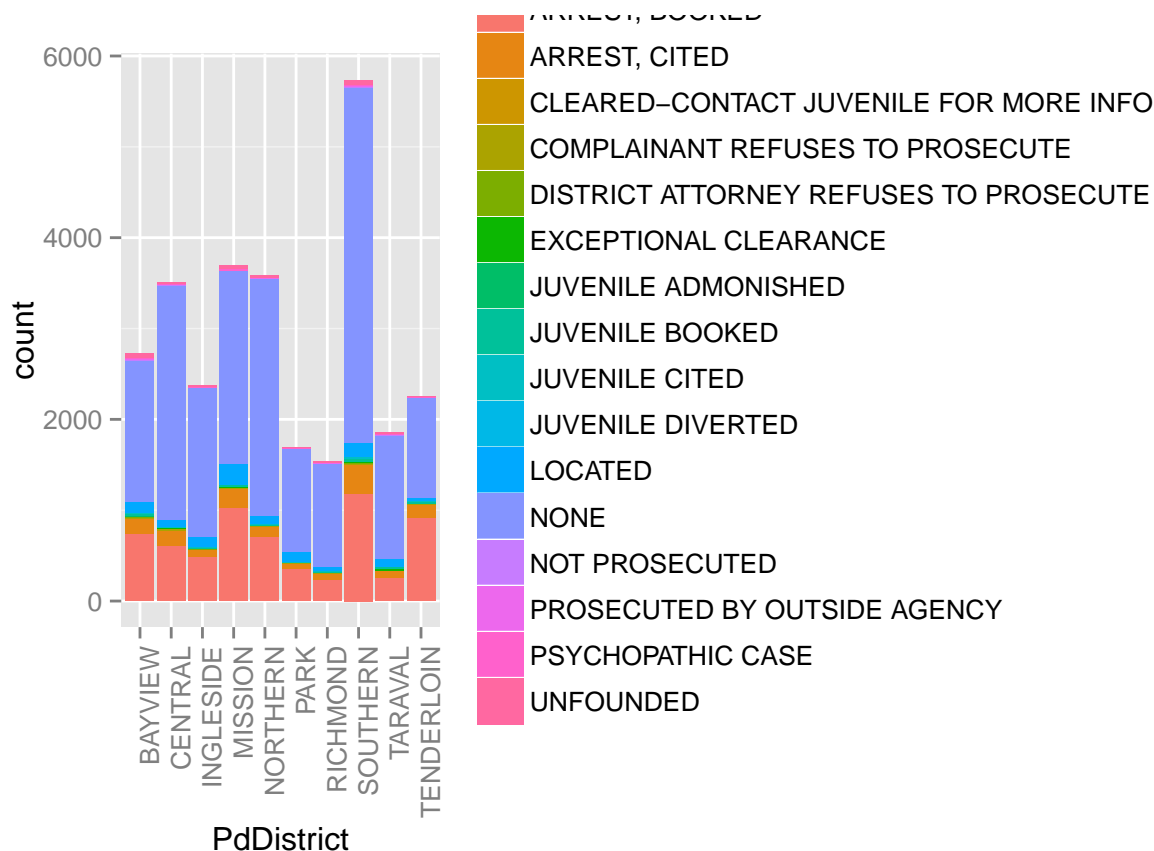
The plot clearly shows that the Southern district has the highest number of crimes reported during the given period. It is interesting to ask whether any particular types or categories of crimes are actually more prevalent in other areas. The plot includes colour coding to illustrate the resolution to each case.

The second plot of this report looks only at incidents where the resolution was the booking of a juvenile. Again, an histogram plotting incidents against district is presented with colour coding this time to show the category of the crime.
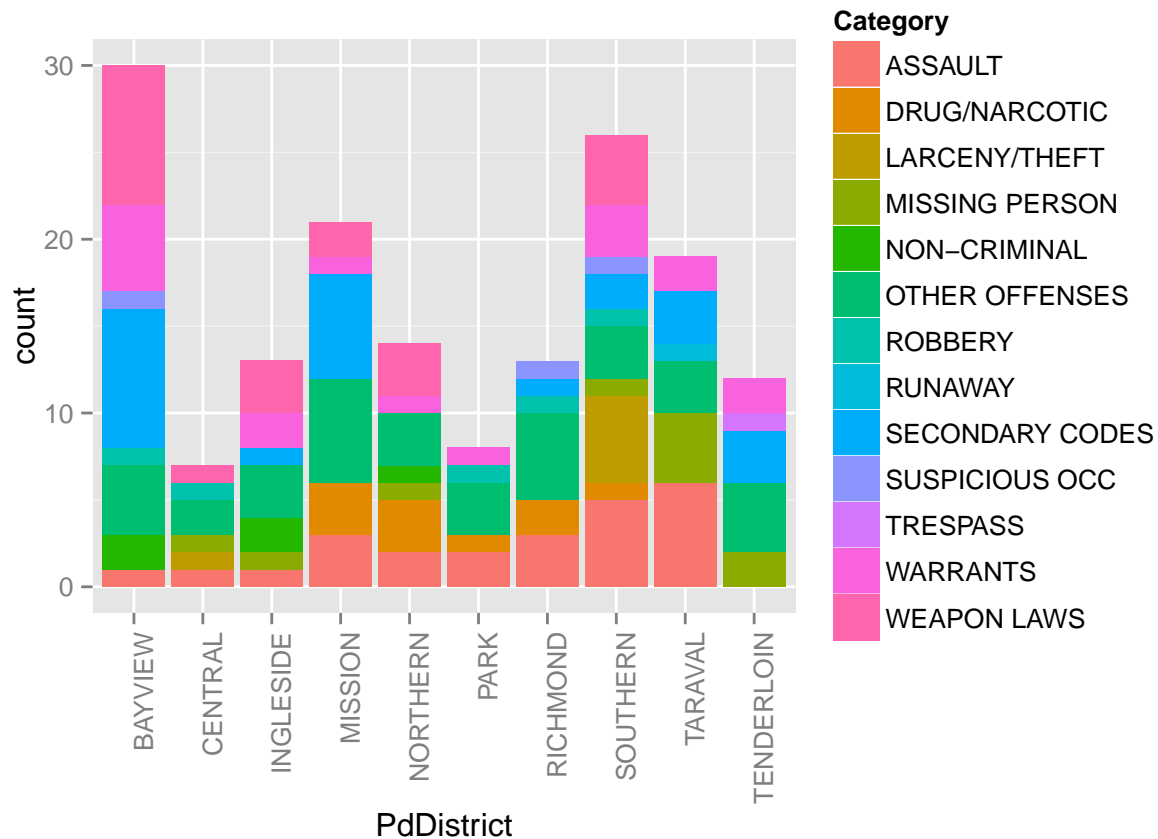
Now we notice that the district with the highest number of incidents during Summer 2014 is Bayview. Southern still has a high number of incidents but, it is second to Bayview in the number of incidents where a juvenile is booked.

The colour-coding reveals some interesting information. First of all, across all districts we see that there is a reduction in the variety of crime committed by juveniles. In particular, during Summer 2013 no juveniles were booked for a large number of crimes including more severe crimes and also what might be called 'whitecollar' crimes such as fraud and embezzlement.

Secondly, we might ask why Bayview has more crime involving juveniles. The distriction of crimes is different to those in, for example, Southern. We see the there are many more crimes categorised as "Secondary codes" or "Weapon laws". There are also far fewer crimes categorised as "Assault".

We can dig deeper in to the records for these crimes by examing the description of each crime. Our final plot looks at the description of each crimes where a juvenile was booked that was categorised as a secondary code. However, for this final plot, we will use the extended data set recording crimes in San Francisco from Jan 2003 to 7th Febrary 2016 inclusive.
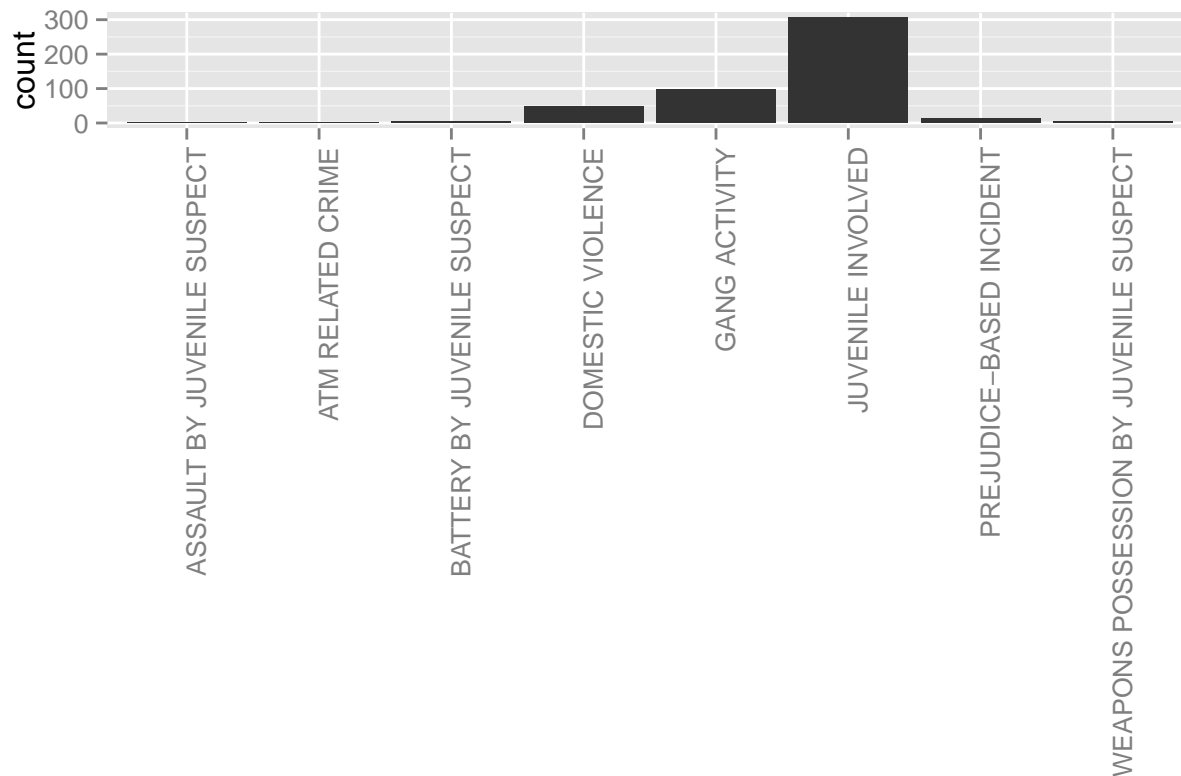
Descript

The description "Juvenile involved" is not particularly informative given that we have selected crimes where a juvenile was booked. However, the second largest contributor is has description "Gang Activity". The author has no knowledge of the the San Francisco area, but it is interesting to ask whether the increased level of crimes involving juveniles in the Bayview district is caused by a larger presence of gangs and gang activity there, in particular in contrast to the Southern district.

Code to generate the plots:

```r
#Load ggplot2 an R package used for plotting
library(ggplot2)

#Load the data. Notice that we will use both the Summer 2014 data
#and the data from Jan 2003 up until present
SF <- read.csv("sanfran.csv")
FullSF <- read.csv("fullsanfran.csv")


JB <- SF[SF$Resolution == "JUVENILE BOOKED",]
JBSC <- JB[ JB$Category == "SECONDARY CODES", ]

FullJB <- FullSF[FullSF$Resolution == "JUVENILE BOOKED",]
FullJBSC <- FullJB[ FullJB$Category == "SECONDARY CODES",]

#Plot 1
qplot(PdDistrict, data = SF, fill = Resolution) + theme(axis.text.x = element_text(angle = 90, hjust =
```

```r
#Plot 2
qplot(PdDistrict, data = JB, fill = Category) + theme(axis.text.x = element_text(angle = 90, hjust = 1))

#Plot 3
qplot(Descript, data = FullJBSC) + theme(axis.text.x = element_text(angle = 90, hjust = 1))
```