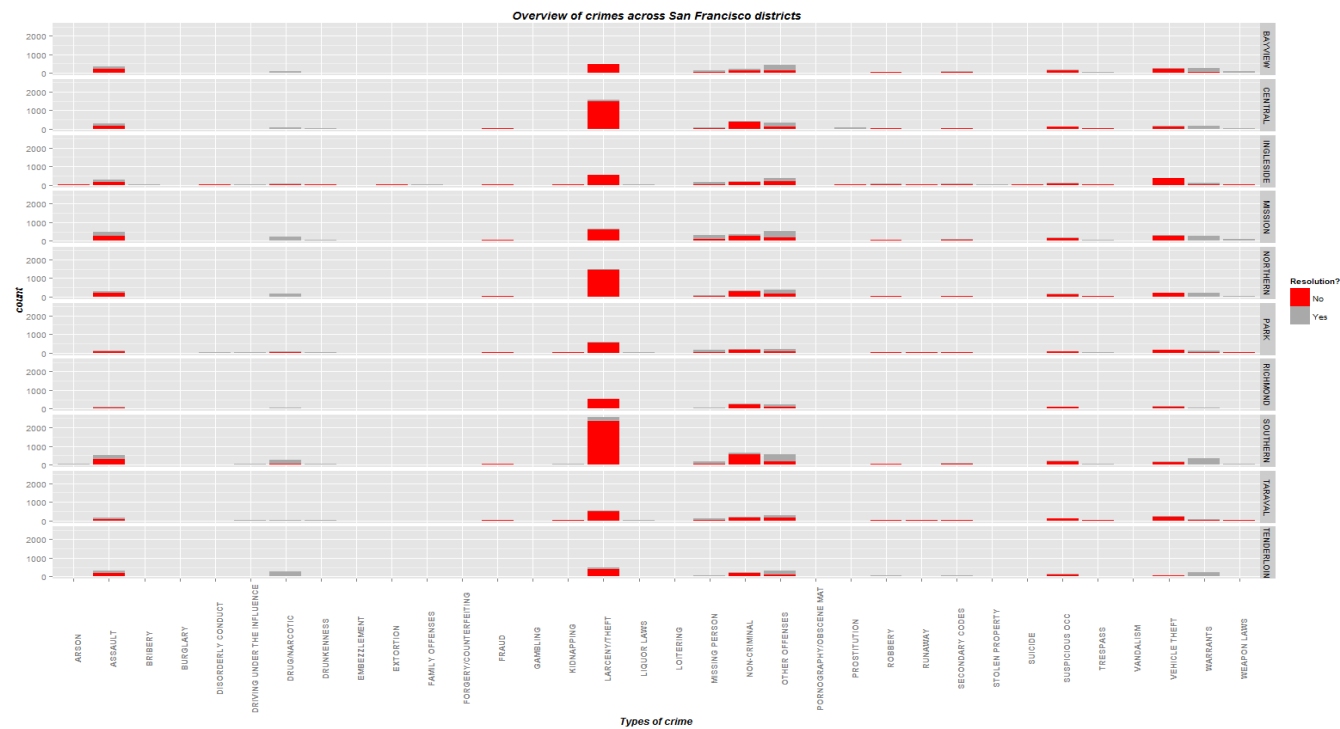


# SF's Southern district has the highest number of thefts in SF ...

Monday, December 14, 2015 11:10 AM

Based on the reduced San Francisco Summer 2014 dataset, here's an overview of the types of crime reported across the districts of SF.

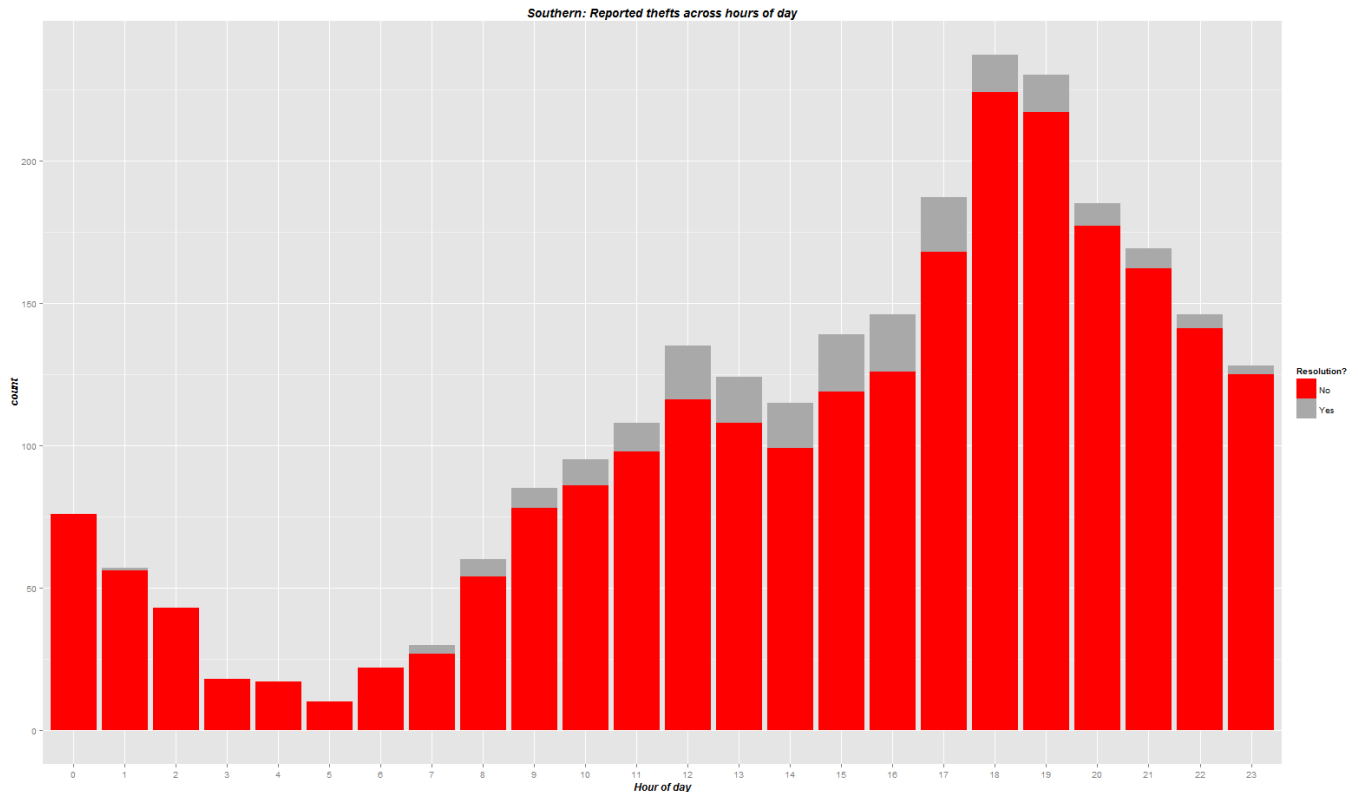


The most common type of crime by far is larceny/theft or simply theft. The rest like non-criminal, assault and other offenses are too similar visually to judge their rankings. To be precise, we list the frequency counts below.

Category freq	Pd district freq
1 LARCENY/THEFT 9466	1 SOUTHERN 5739
2 OTHER OFFENSES 3567	2 MISSION 3700
3 NON-CRIMINAL 3023	3 NORTHERN 3589
4 ASSAULT 2882	4 CENTRAL 3513
5 VEHICLE THEFT 1966	5 BAYVIEW 2725

Back to the visual overview, we can see Southern district being reported with the most theft cases, followed by Central and Northern districts. And it is not always the case that there is a significantly higher number of thefts than other types of crimes. Counter example would be Bayview, Mission and Tenderloin which have more even distribution of types of crimes.

Now, let's take a closer look at the most common crime in the district with highest number of crimes: theft in Southern district.



We plot the count of thefts across the hour of the day in Southern. The most number of thefts are reported between 1800H - 2000H. And the most common sub types of thefts are grand theft auto (1,081 cases) and petty thefts of property (359 cases).

## And it is not due to lack or difficulty of enforcement.

Admittedly, as I was doing this assignment and plotting the numbers, I made the lazy assumption that thefts are the most prevalent in Southern district due to the difficulty in enforcement which should show up with lower resolution rate than the overall resolution rate for all theft crimes in SF. A crime is considered resolved if its resolution status is not 'None'.

However I found that the average resolution rate for theft in SF is 5.64% while that of Southern is 7.61%. To be more statistically accurate, I ran a bootstrap of 10,000 resamples and Southern's resolution rate of 7.61% is still more than 99% of the upper confidence interval of the resampled distribution.

In summary, we found (i) the most common crime in SF and (ii) the district with the most number of crimes. We took a deeper look at Southern thefts and also found that it has a surprisingly higher rate of resolution than the average SF theft (no puns intended).