

Analysis of London Crime Dataset using Tableau

Descriptives and Visualisation

1. Showing where crimes and stop-and-search events are occurring using tableau

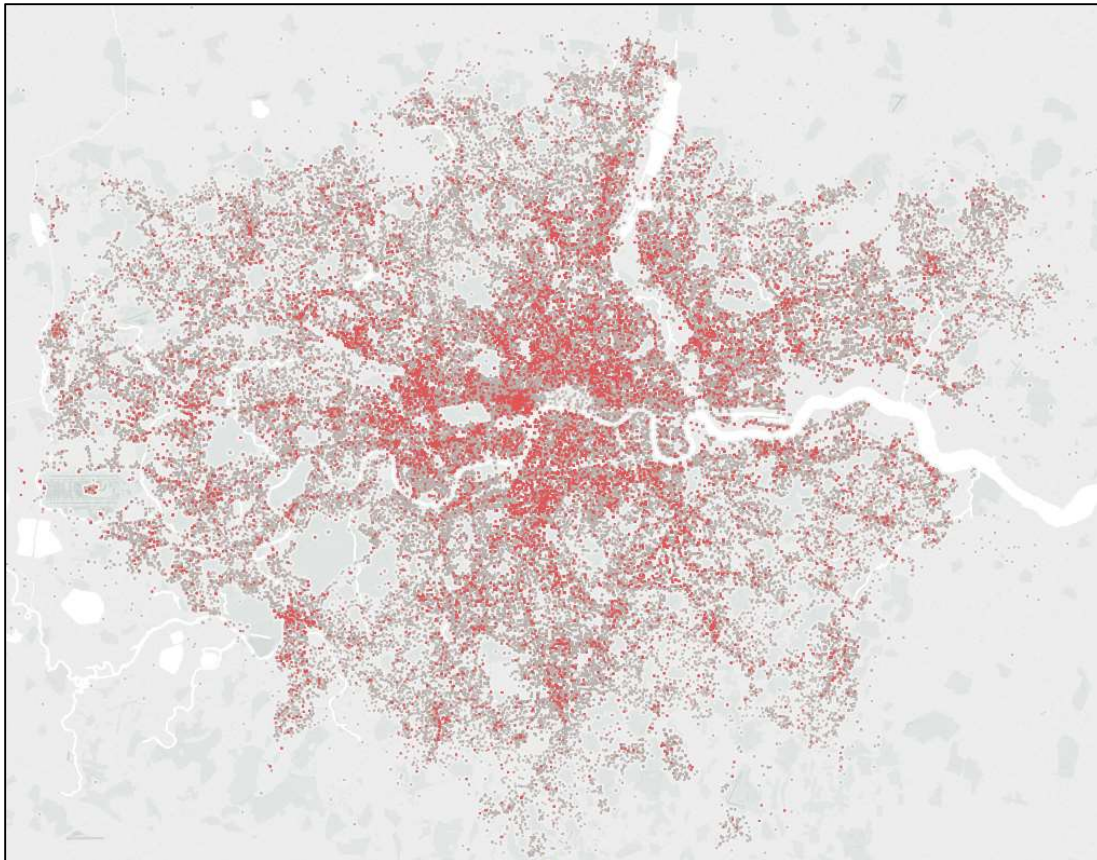
Please Note: I have focused on Greater London area as this is where high majority of the events occur.



Stop and Search results



Crime results

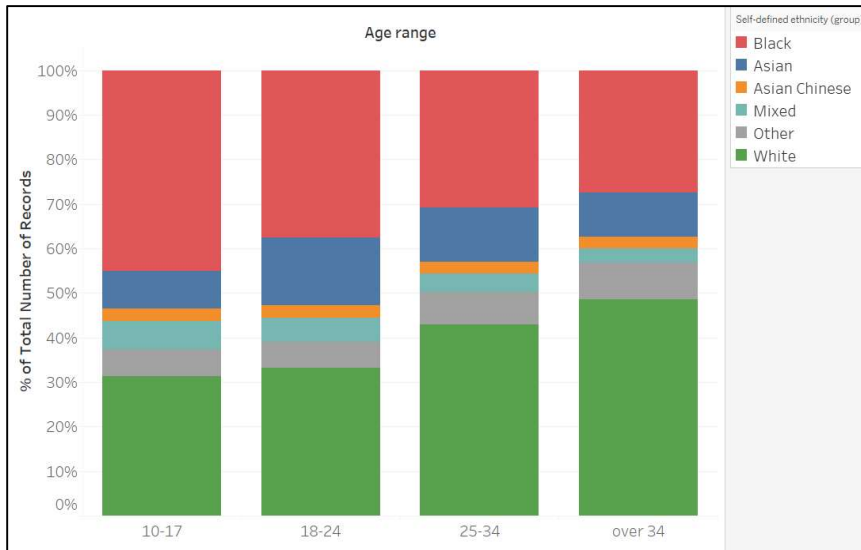


Data for Stop and Search (red) and Crime (grey)

2. *Organising tableau visualisation in order to be able to filter by gender, age range (6:01-11am; 11:01 - 1pm, 1:01-4:30pm; 4:31-8pm; 8:01-11pm; 11:01pm-2:30am; 2:30am-6am), race (as in “self-defined”), and outcome. Presenting screenshots for most interesting 2 patterns*

a) Summary of Searches based on Ethnicity and Age

Please note that here we use searches data. The outcomes where there was nothing found on suspect (i.e. there was no further action to take) were excluded. Therefore these results show where police took some action towards the suspects (arrests, warnings, cautions, etc.)



Self-defined ethnicity (group)	Age range			
	10-17	18-24	25-34	over 34
Black	45.03%	37.63%	30.87%	27.42%
White	31.35%	33.11%	42.96%	48.53%
Asian	8.42%	15.18%	12.20%	9.91%
Asian Chinese	2.84%	2.84%	2.54%	2.70%
Mixed	6.37%	5.37%	4.07%	3.07%
Other	6.00%	5.88%	7.35%	8.38%

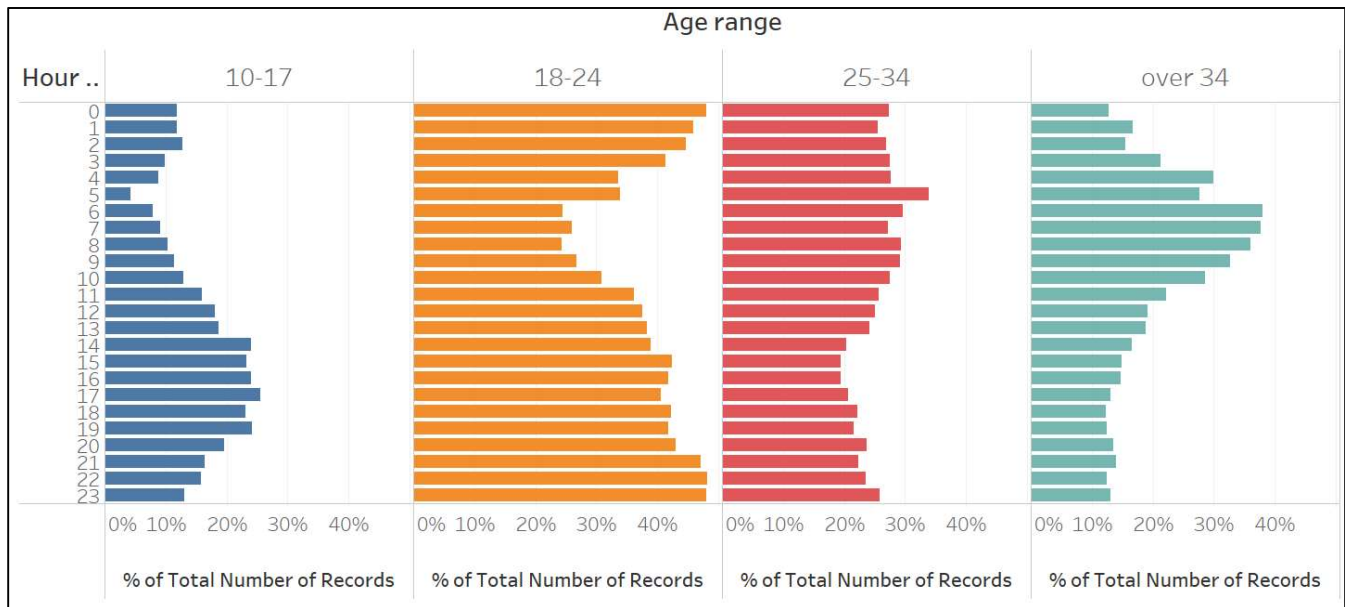
Self-defined ethnicity (group)	10-17	18-24	25-34	over 34	Grand Total
Black	856	2,439	1,167	661	5,123
White	596	2,146	1,624	1,170	5,536
Asian	160	984	461	239	1,844
Asian Chinese	54	184	96	65	399
Mixed	121	348	154	74	697
Other	114	381	278	202	975
Grand Total	1,901	6,482	3,780	2,411	14,574

By analysing the searches data from the perspective of their self-identified Ethnicity and Age, we can see a few interesting trends:

- For the group of people identified as 'Black': for ages between 10 and 17, this group represents 45% of all searches. As we observe this measure for other age groups for this ethnic group we can clearly see that the proportion of searches goes down gradually for other age groups we get with 37.6%, 30.9% and 27.4% for '18-24', '25-34' and 'over 34' age groups respectively. Based on the observation we could assume that, as a proportion of total searches per age category, people in this ethnic group are less likely to represent higher proportion in criminal activity older they get;
- For the group of people identified as 'White' we can see an opposite effect where for age group '10-17' this group represents 31.4% of all searches within this group. This measure goes up for older age groups with 33.1%, 43.0% and 48.5% for '18-24', '25-34' and 'over 34' age groups. Based on this observation we can assume that, as a proportion of total searches per age category, people in this ethnic group are more likely to represent higher proportion of people involved in criminal activity older they get.

Authorities could use this finding to identify which segments of population to focus on in to reduce crime. For example, authorities could focus on 'educating' late teens/early 20s in 'white' ethnic category to reduce potential offenders.

b) Summary of searches by age groups and time of search



We can observe a few interesting tendencies by looking at proportion of arrests split by time of the day and different age groups.

For '10-17' age group most of the searches happen during 13-20 hours. This is not very surprising as one would expect representatives of this age group to be in school from 8am until 1-2pm and therefore stay 'out of troubles' until they are out of schools. As a result, we can see that most of the number of searches occur at the time when teenagers are not in school and before their 'home time' at around 8-9pm.

For '25-34' age group there are few outstanding values with most of the results are being around mid-20s with a low variation.

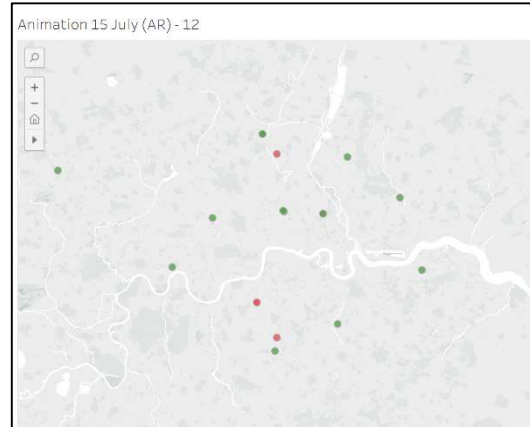
For '18-24' and 'over 34' age groups we can observe an interesting tendency where searches for the former group are at its highest from 15pm until 3am with lowest number of searches come between 6am and 10am. Interestingly there is opposite results for 'over 34' age group, where its highest values occur between 6am and 9am. After 9am the number of searches for 'over 34' gradually reduces until 2pm where it stays around the same region until 3am when it starts growing again.

These findings are somewhat surprising as it is difficult to explain why the results for these two groups move in opposite directions to each other and require further analysis to explain the phenomenon.

These findings, once better understood, could be used to better 'target' offenders throughout the day using crime-prevention activities and therefore use police personnel more effectively.

3. With the stop-and-search data for July only, I used tableau to create an animation showing stop and search events by hour for the day of 15th July. Using the "outcome" column to determine whether the outcome is "arrested" or not, I coded each data point by colour to distinguish crimes from stop-and-search events.

The Animation summary of the searches on 15th July could be viewed [here](#). The colours of the results represent the following:



- **Green**- No further action taken by police;
- **Red**- Police action. These include local resolution, offender given drugs possession warning, offender given penalty notice, suspect arrested, suspect summoned to court.

The animation includes the breakdown of the searches and their outcomes per hour, i.e. 24 slides in total. Example of hour 12 is shown above.

Analytics

To analyse the effect and impact of stop searches (referred to as 'searches' below) on crime a sample of 1,000 random occurrences have been used. Every search in the sample has then been assigned a geographical area (grid reference) based on given Latitude and Longitude. The searches have then been combined in groups/clusters based on their geographical area to enable a more focused analyses of searches effect per area. Finally, geographical areas with less than 10 searches have been removed from the sample as they are outliers. As a result, the final dataset consisted of 916 searches across 30 geographical areas.

Results

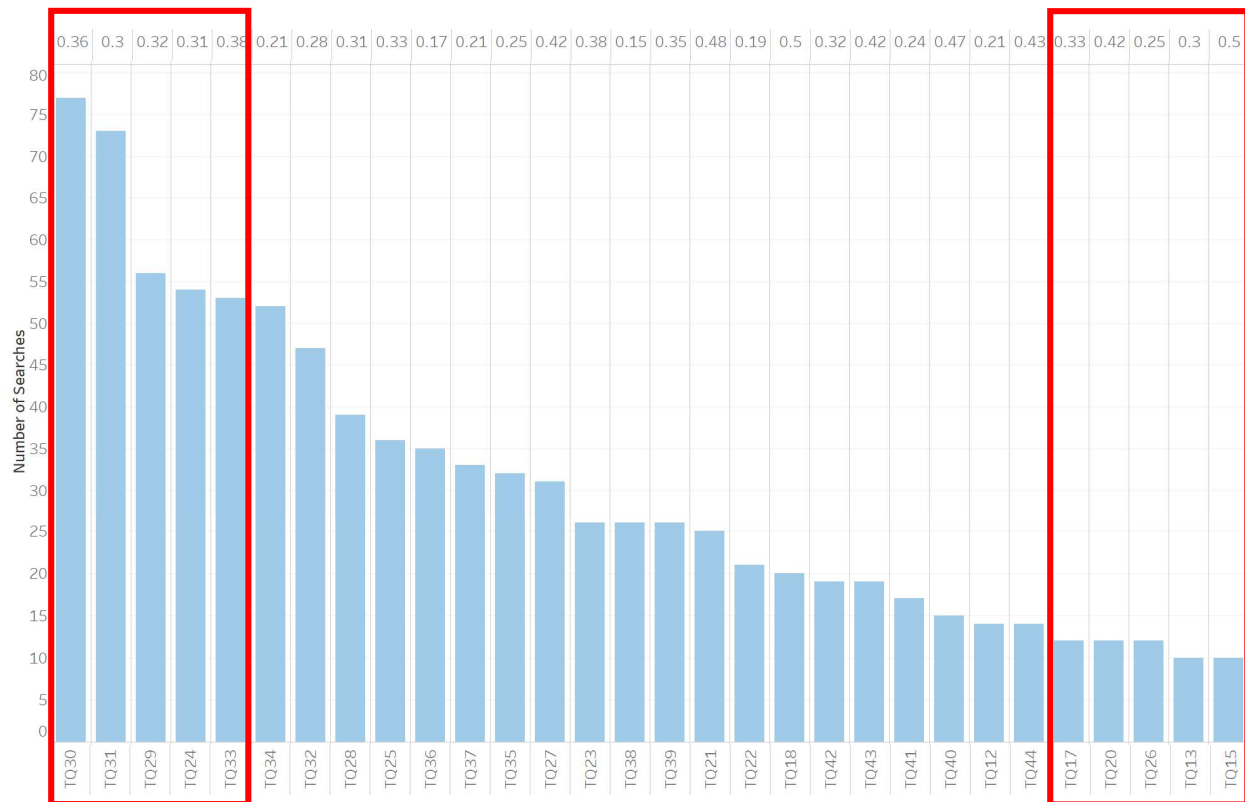
Grid Ref	Searches	Police Action	Ratio
TQ18	20	10	50.0%
TQ15	10	5	50.0%
TQ21	25	12	48.0%
TQ40	15	7	46.7%
TQ44	14	6	42.9%
TQ43	19	8	42.1%
TQ27	31	13	41.9%
TQ20	12	5	41.7%
TQ23	26	10	38.5%
TQ33	53	20	37.7%
TQ30	77	28	36.4%
TQ39	26	9	34.6%
TQ25	36	12	33.3%
TQ17	12	4	33.3%
TQ29	56	18	32.1%

Grid Ref	Searches	Police Action	Ratio
TQ42	19	6	31.6%
TQ24	54	17	31.5%
TQ28	39	12	30.8%
TQ31	73	22	30.1%
TQ13	10	3	30.0%
TQ32	47	13	27.7%
TQ26	12	3	25.0%
TQ35	32	8	25.0%
TQ41	17	4	23.5%
TQ12	14	3	21.4%
TQ37	33	7	21.2%
TQ34	52	11	21.2%
TQ22	21	4	19.0%
TQ36	35	6	17.1%
TQ38	26	4	15.4%

Based on the results above, we can see that on average, 31.7% of searches resulted in some form of a police action (local resolution, offender given drugs possession warning, offender given penalty notice, suspect arrested, suspect summoned to court) with 'success' ratio varies from 15.4% to 50%. Essentially

the ratio shows a probability of a person stopped and searched in a given geographical area to be arrested (given a warning, summon to court, etc.).

We can assume that geographical areas with a low ratio are examples of 'poor' policing. This is because in a large number of stop and searches there were no further action taken against the suspect. However, it is also possible to suggest that a high number of stop and searches acts as a deterrent and therefore could be example of a 'good' policing. Therefore the results are inconclusive.



Top 5 results
(high number of searches)

Grid Ref	Searches	Police Action	Ratio
TQ30	77	28	36.4%
TQ31	73	22	30.1%
TQ29	56	18	32.1%
TQ24	54	17	31.5%
TQ33	53	20	37.7%
Average			33.6%

Bottom 5 results
(low number of searches)

Grid Ref	Searches	Police Action	Ratio
TQ20	12	4	33.3%
TQ17	12	5	41.7%
TQ26	12	3	25.0%
TQ15	10	3	30.0%
TQ13	10	5	50.0%
Average			36.0%

We can see that by sorting the results by the number of searches there is no obvious conclusion to be made whether increase in number of searches necessarily leads to significantly better ratio of police action: average value for top 5 values is only marginally bigger than average across all 30 areas. Similarly, for 5 areas with the smallest number of searches we the results are on average are around 36%, which again is only marginally higher than average value.

Community Impact

Although we do not observe major changes in the ratio because of the increase in stop and search events, it will be logical to suggest that with increase in searches will result in increase in police action (arrests, etc.) as the ratio stays roughly the same regardless of number of searches. However, it is important to consider effect of the searches might have on the community. Large number of searches might have a counter-productive effect by alienating members of the community, particularly within ethnic minorities. This impact could also have an adverse effect on relationship between police and members of community and therefore limit impact of police work in the area in the long term.

Limitation of the research and considerations for future analysis

The analysis shown that based on the provided data it is difficult to make any valid conclusions regarding the impact and effectiveness of searches on reduction in crime and whether this method of crime prevention could equate as 'good' or 'bad' policing. The difficulty could be explained by the following:

- a) Lack of conviction data. Not all arrests will necessarily lead to a prosecution (i.e. the potential offender is found guilty), therefore only relying on arrests and other types of police action might lead to wrong assumptions;
- b) Point in time. It is difficult to judge the effectiveness of the activities without being able to analyse its results on an ongoing basis;
- c) Sample size. The original dataset for the stop-and-search police activity contained over 45,000 rows. Therefore by taking a small random sample of 1,000 instances we can expose ourselves to potential misleading results and assumptions based on the limited data input;
- d) Isolation of data. It is impossible to judge the effectiveness of the searches as we do not know how many members of police forces are patrolling the selected areas. The presence of police conducting searches could act as a deterrent for potential offenders. However, based on the available data it is impossible to confirm or reject this assumption;
- e) Cause and effect. Here we have an example of activity where increase in one variable (searches) is likely to have an increase in another (police action). Therefore, we can have a wrong impression of the positive impact of the crime prevention technique without looking at other impacts of the activities (such as cost, impact on community, etc.).

In order to eliminate these limitations and to better analyse the effectiveness and impact of stop-and-search crime prevention technique, the following changes are suggested for the future analyses:

- a) Additional information should be considered to be included:
 - Number of police officers involved
 - Other crimes in the area (not directly resulted from searches)
 - Cost of the operations
 - Population number and general demographics of the observed areas
 - Reason for search
- b) As mentioned previously, it is essential to analyse whether police action lead to a conviction (i.e. the suspect was found guilty or not);
- c) Observations should be made over a longer period to see how changes or differences in practices over time impact the crime prevention and reduction;
- d) Where possible it is advisable to avoid using sampling and instead include the total data population;

- e) Qualitative analysis of the impact on the community must be conducted to judge the sentiment of the public towards this crime prevention activity. This analysis should include information gathered from a wide range of ethnic groups within communities;
- f) Qualitative analysis of the technique from the offenders point of view could also be used. This could include interviews or/and surveys of offenders to establish whether searches could act as a deterrent of potential crime and to better understand when searches might be unwarranted.

All the above information should lead to better understanding of implications of stop-and-search crime prevention technique and whether it could be characterised as 'good' or 'bad' policing.