

Analysis of criminal situation in Chicago from 2001-2019

Authors: Ken Pan, Sicong Chang

The research Questions:

1. What is the ranking of safety of different district of Chicago?

The safety is computed by the number of occurrence and weight of different types of crime.

Crime whose type is more severe should have higher weigh.

The result to this question can be a reference for people who want to live in Chicago.

2. Predict the possibility of the criminal being arrested at a given time of a day and a given criminal type.

Use machine learning model to predict the possibility of solving a crime.

Give the police an overview of the difficulty of the case and remind them to concentrate more on cases that are hard to be solved.

Can also be a reference to the victim to see how likely can their case can be solved and can provide an safety education for them.

3. Does the increase or decrease in rate of solved case affect the crime rate? Does the result vary from different criminal types?

We want to use the data from earlier years.

First compute the proportion of solved cases for each type of crime for each year.

Then see if the change of the police performance can affect the amount of crimes in later years.

It could be a motivation for police to keep improving their performance.

Data set:

The data set we choose is crimes from 2001 to present in Chicago.

The url is below

<https://data.cityofchicago.org/api/views/ijzp-q8t2/rows.csv?accessType=DOWNLOAD>

It's a csv file almost without missing data. The frame size is 6866498 * 30. The index of the frame is ID for each case. The columns include some useless data to me such as case number and FBI number. However, most of the columns in the dataset are useful. The time of the case include hour and minutes. It allows us to analyze annually changing rate or the situation in a specific time period of a day. There are also lots of location information such as the block, position coordinate, district number and police beats, which allow us to graph map. The type of crimes is also well stated for most of the cases. There is also a column of Boolean which shows if the case is solved or not. The column is extremely useful for analyzing the performance of police on different type of cases.

ID	Case Number	Date	Block	IUCR	Primary Type	Description	Location Des	Arrest	Domestic	Beat	District	Ward	Community #	FBI Code	X Coordinate	Y Coordinate	Year	Updated On	Latitude	Longitude	Location	Historical Wz	Zip Codes	Community #	Census Tract	Wards	Boundaries -	Police District	Police Beats
11678704	JC252900	05/06/2019	0180X W HO	610	BURGLARY	FORCIBLE EN RESTAURANT	FALSE	FALSE	FALSE	2424	24	49	1	5	1162562	1950347	2019	05/13/2019	42.0194111	-87.677125	42.019411121, -87.67712	21853							
11678911	JC253154	05/06/2019	1000X W OH	1152	DECEPTIVE P	ILLEGAL USE AIRPORT TER	FALSE	FALSE	FALSE	1651	16	41	76	11	1100658	1994241	2019	05/13/2019	41.9762904	-87.905227	41.9762904	34	16197	75	668	29	38	12	24
11678735	JC252879	05/06/2019	0630X S MAF	610	BURGLARY	FORCIBLE EN RESIDENCE	FALSE	FALSE	FALSE	725	7	16	67	5	1166412	1862674	2019	05/13/2019	41.7787485	-87.665469	41.7787484	44	22257	65	280	3	23	17	204
11679131	JC252877	05/06/2019	0010X W 10S	1030	ARSON	POS. EXPLOS RESIDENCE	FALSE	FALSE	FALSE	513	5	34	49	9	1177310	1832676	2019	05/13/2019	41.696191	-87.62642	41.6961910	45	21861	45	524	22	19	10	260
11679387	JC253416	05/06/2019	0030X S ALB	810	THEFT	OVER \$500 RESIDENCE	FALSE	FALSE	FALSE	1124	11	28	27	6	1155765	1898312	2019	05/13/2019	41.8767639	-87.703544	41.8767638	11	21184	28	737	23	28	16	123
11678669	JC252872	05/06/2019	0070X W WA	1310	CRIMINAL DA	TO PROPERT APARTMENT	FALSE	FALSE	FALSE	1925	19	46	6	14	1170748	1924887	2019	05/13/2019	41.9493722	-87.647752	41.9493722	37	21186	57	726	39	53	5	18
11681539	JC256186	05/06/2019	0220X W NO	820	THEFT	\$500 AND UN STREET	FALSE	FALSE	FALSE	1424	14	1	24	6	1160993	1910597	2019	05/13/2019	41.9103681	-87.694007	41.9103680	24	22535	25	516	41	4	7	200
11678753	JC252959	05/06/2019	0690X S DOR 031A		ROBBERY	ARMED; HAN SIDEWALK	FALSE	FALSE	FALSE	321	3	5	43	3	1186735	1859307	2019	05/13/2019	41.7690518	-87.591071	41.7690518	32	22260	39	416	33	60	18	206
11679864	JC252869	05/06/2019	0040X S CLAF	530	ASSAULT	AGGRAVATE; RESIDENCE P	FALSE	FALSE	FALSE	122	1	4	32	04A			2019	05/13/2019	04:13:25 PM										
11678677	JC252870	05/06/2019	0050X W 119 143A		WEAPONS VI	UNLAWFUL P STREET	FALSE	FALSE	FALSE	524	5	34	53	15	1174681	1825981	2019	05/13/2019	41.6778775	-87.636207	41.6778774	45	21861	50	255	22	19	10	220
11678692	JC252868	05/06/2019	1080X S SAN	560	ASSAULT	SIMPLE RESIDENCE	FALSE	TRUE	FALSE	2234	22	34	75	08A	1171912	1832962	2019	05/13/2019	41.6970957	-87.646175	41.6970957	45	22212	74	315	22	13	9	263
11678736	JC252897	05/06/2019	0340X S OAK	2825	OTHER OFFE	HARASSMEN RESIDENCE	FALSE	FALSE	FALSE	912	9	12	59	26	1161585	1881747	2019	05/13/2019	41.8311887	-87.682636	41.8311886	26	14920	56	2	1	43	23	165
11678717	JC252854	05/06/2019	0080X W WA	460	BATTERY	SIMPLE STREET	FALSE	FALSE	FALSE	1923	19	46	6	08B	1170017	1924784	2019	05/13/2019	41.9491056	-87.650442	41.9491056	37	21186	57	727	39	53	5	12
11678918	JC252850	05/06/2019	0050X S PUL	420	BATTERY	AGGRAVATE CTA PLATFOI	FALSE	FALSE	FALSE	1132	11	24	26	04B	1149812	1897228	2019	05/13/2019	41.873907	-87.72543	41.8739070	36	21572	27	675	14	30	16	142
11678667	JC252857	05/06/2019	1100X S WEN	860	THEFT	RETAIL THEF SMALL RETA	TRUE	FALSE	FALSE	513	5	34	49	6	1176911	1831518	2019	05/13/2019	41.6930223	-87.627915	41.6930222	45	21861	45	524	22	19	10	260
11678685	JC252855	05/06/2019	0700X S RACI	820	THEFT	\$500 AND UN RESIDENCE	FALSE	TRUE	FALSE	734	7	6	67	6	1169534	1858205	2019	05/13/2019	41.7664179	-87.654153	41.7664178	17	22257	65	21	32	23	17	216
11678953	JC253268	05/06/2019	0130X N SPRI	820	THEFT	\$500 AND UN RESIDENCE-C	FALSE	FALSE	FALSE	2535	25	26	23	6	1150132	1908770	2019	05/13/2019	41.9055733	-87.723954	41.9055732	27	4299	24	454	49	5	6	193
11678697	JC252847	05/06/2019	0380X W IOV	486	BATTERY	DOMESTIC B STREET	TRUE	TRUE	FALSE	1112	11	37	23	08B	1150308	1905727	2019	05/13/2019	41.8972196	-87.723387	41.8972195	41	4299	24	456	45	5	16	66
11679088	JC253174	05/06/2019	0300X N NOT	1310	CRIMINAL DA	TO PROPERT RESIDENCE	FALSE	FALSE	FALSE	2511	25	29	18	14	1128306	1919448	2019	05/13/2019	41.9352726	-87.803888	41.9352725	39	22254	18	397	7	52	6	179

Useful libraries that may be used:

Pandas

Geopandas

Sci-kit learn

sea-born

matplotlib.pyplot

plotly

requests