Objective of our study:

Conduct an analysis of crime and identify relevant factors in order to find potential ways to reduce it

The data set:

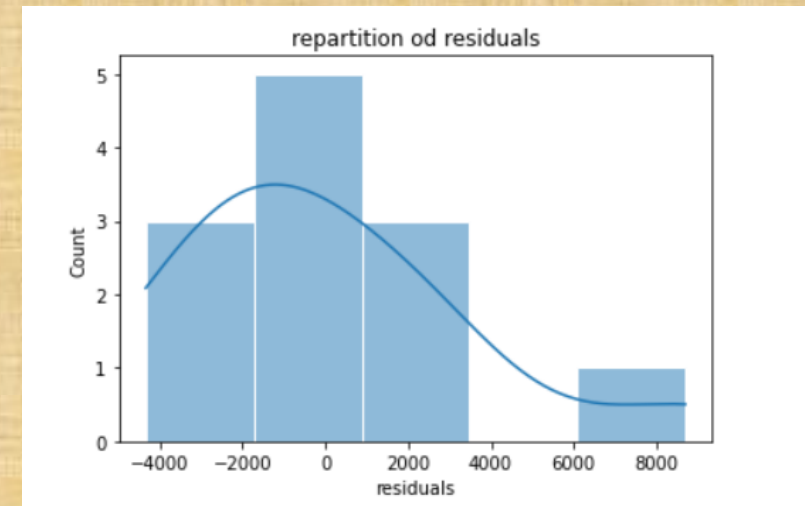| CODE	NAME	NB_INCIDENTS	MEDIAN_AGE	TOTAL_POP	PERC_OF_30_34	PERC_MARRIED_COUPLE_FAMILY	PER_CAPITA_INCOME	PERC_OTHER_STATE
A1	Downtown	26260	33.5	18306	32.5	35.8	80057	
A15	Charlestown	4765	35.7	19890	28.2	36.4	75339	
A7	East Boston	9691	30.6	47263	31.1	30.4	31473	
B2	Roxbury	38877	32.5	54161	27.8	17.8	20978	
B3	Mattapan	28331	36.7	26659	20.9	29.8	28356	

This section aims at giving you a first overview of the database we get:
- Data provided by the Police Department of Boston report criminal incident that has required an intervention by the police in the municipality of Boston
- Data provided by Urbanism Department of Boston. It gathers geographical information about the various police districts of Boston;
- shared by the Economics Department of Boston, which keeps track of various indicators of the social and economic activity of Boston neighborhoods. Each observation corresponds to a police district.

Does the crime is linked to social and economic indicators ?

We tried to find correlation between crime and social and economic indicators. The result seems good, but in fact residuals are really not random so R-squared remains perfecly valid (deterministic coef) However, inferencial coefs cannot be trusted
p-values and confidence intervals may be smaller than they should be
Don't be too confident that your model generalizes well.

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	1.358e+04	6.16e+04	0.220	0.834	-1.45e+05	1.72e+05
MEDIAN_AGE	2252.7344	1237.522	1.820	0.128	-928.416	5433.885
PERC_OF_30_34	-532.9973	468.940	-1.137	0.307	-1738.447	672.452
PERC_MARRIED_COUPLE_FAMILY	-2115.1839	408.344	-5.180	0.004	-3164.866	-1065.502
PERC_OTHER_STATE_OR_ABROAD	1787.6939	532.432	3.358	0.020	419.034	3156.354
PERC_LESS_THAN_HIGH_SCHOOL	112.7375	493.977	0.228	0.829	-1157.072	1382.547
PERC_COLLEGE_GRADUATES	19.4699	458.775	0.042	0.968	-1159.850	1198.790

repartition od residuals

What is the most common offense code group ?

	NAME	OFFENSE_CODE_GROUP	INCIDENT_NUMBER
73	South End	Larceny and vandalism	14012
59	Roxbury	Larceny and vandalism	10471
24	Downtown	Larceny and vandalism	9256
17	Dorchester	Larceny and vandalism	9230
61	Roxbury	Police investigation procedure	8030
19	Dorchester	Police investigation procedure	7428
54	Mattapan	Police investigation procedure	7001
52	Mattapan	Larceny and vandalism	6598
66	South Boston	Larceny and vandalism	5637
75	South End	Police investigation procedure	5540
62	Roxbury	Violence and harassment	5495
26	Downtown	Police investigation procedure	5464
3	Brighton	Larceny and vandalism	5274
58	Roxbury	Fraud and law violations	4820
45	Jamaica Plain	Larceny and vandalism	4546
20	Dorchester	Violence and harassment	4067
55	Mattapan	Violence and harassment	3910
16	Dorchester	Fraud and law violations	3705
49	Mattapan	Disputes	3573

As you can see here the Larceny and vandalism is the most common offense group.