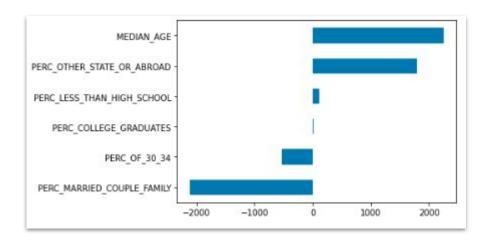
Boston Crimes

Les Experts: L'analyse statistique

Ce qu'on a trouvé

The two most significant regressors:

- MEDIAN_AGE
- PERC_OTHER_STATE_OR_ABROAD



Plus des découvertes

MEDIAN_AGE

There is a 12.8% probability that the median_age is not correlated to the number of incidents.

OTHER_STATE

There is a 2% probability that the population from other states of abroad are not related to incidents.

| | | | egression Resu | | | | | |
|----------------------------|------------------|------------|-------------------|---------------------|-------|-----------|-----------|--|
| Dep. Variable: | NB_INCIDENTS | | | R-squared: | | | 0.899 | |
| Model: | OLS | | | Adj. R-squared: | | | 0.777 | |
| Method: | Least Squares | | | F-statistic: | | | 7.391 | |
| Date: | Tue, 29 Jun 2021 | | | Prob (F-statistic): | | | 0.0222 | |
| Time: | 15:41:10 | | | Log-Likelihood: | | | -114.59 | |
| No. Observations: | 12 | | | AIC: | | | 243.2 | |
| Df Residuals: | | | | BIC: | | | 246.6 | |
| Df Model: | | | | | | | | |
| Covariance Type: | nonrobust | | | | | | | |
| | | coef | std err | | 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 | |
| Omnibus: | 6.933 | | Durbin-Watson: | | | 2.255 | | |
| Prob(Omnibus): | 0.031 | | Jarque-Bera (JB): | | | 3.155 | | |
| Skew: | 1.148 | | Prob(JB): | | | 0.206 | | |
| Kurtosis: | 4.020 | | Cond. No. | | | 2.42e+03 | | |

Pourquoi nous ne brillons pas?

- Primary reason
 - Very small sample size

Other explanations

- High-correlation
- Multicollinearity
- Combined too many parameters
- Combined together districts that may not be related

Comment ameliorer?

- Examine districts one-by-one
- Examine incidents without grouping by district but by indicators