Our group decide to focus on crime data in Washington D.C.

Data source: http://opendata.dc.gov/datasets

Background: As we known, crimes can cost our society dearly and threaten our safety in our daily life[1]. And the influenced juveniles can make the situation worse. For the society, the distribution of crimes can help police departments deploy their men to prevent the crimes or add monitors at the best location so that they can use their funds more efficiently. Besides, for we ordinary people, analysis of crime data can show us which area is risky and we can keep away from these areas or be more careful when passing through these zones to avoid being harmed by criminals.

Historically, the work of dealing with crimes is the prerogative for the criminal justice and law enforcement specialists. However, the development of computer science and data science can help the law enforcement officers and detectives solve the problems by using algorithms in data analysis[1].

In most of the articles, they focused on how to deal with the crimes in the sight of government and aimed to prevent the crimes. But we will focus on how to avoid being harmed by crimes and try to know where is safe and where is risky. Besides we want to figure out the relations between crime data and juvenile arrests data.

Dataset: The dataset we are using is D.C. crime incidents in 2016 and 2017 to get criminal situation in region of D.C. The dataset contains a subset of locations and attributes of incidents reported in the ASAP (Analytical Services Application) crime report database by the District of Columbia Metropolitan Police Department (MPD).

Useful variables:

|  |  |
| --- | --- |
| REPORT\_DAT | When crimes happened |
| SHIFT | DAY/EVENING |
| METHOD | How criminal happens |
| OFFENSE | What kind of crime |
| BLOCK | Where crimes happens |
|  |  |

The juvenile arrests dataset represents bi-annual reports on Juvenile Arrests in the District of Columbia from 2011 to 2017. We want to find the relationship between crime and juvenile crime and know how crimes influence the young people.

Useful variables:

|  |  |
| --- | --- |
| ARREST\_DATE | When they arrested |
| TOP\_CHARGE\_DESC | Why they arrested |
| CRIME\_PSA | Where they do crimes |
| HOME\_PSA | Where they live |

Reference

[1] Nath, S. (n.d.). Crime Pattern Detection Using Data Mining. Web Intelligence and Intelligent Agent Technology Workshops, 2006. WI-IAT 2006 Workshops. 2006 IEEE/WIC/ACM International Conference on (pp. 41–44). IEEE. doi:10.1109/WI-IATW.2006.55