

Q-Step Internship: The predictive factors of youths engaging in crimes - Home Office

Marilena Onisiforou, BA Linguistics



In summer 2016 I completed an 8 week placement at the Home Office in London with the Crime and Policing Analysis Unit. This unit

undertakes research on crime and policing issues in order to facilitate policy development and operational activity.

Summary

Over the two months that I worked in the Home Office I worked in the Crime Patterns team and my main project was on the predictive factors of youths engaging in crime.

The data I used was comprised of a sample of twins, and was taken from a longitudinal study.

My analysis additionally drew on matched data from the Police National Computer (PNC). My findings were compared to a previous analysis based on self-reported data.

As well as doing some basic/ descriptive statistics on the data from the longitudinal study, I also undertook a more complex analysis on the predictive variables. Such analyses included a chi-square analysis and logistic regression.

Methods

In the first part of my report I completed some descriptive statistics by using Excel to find percentages and proportions of numbers of offenders and offences according to various crime types.

I also analysed the top 10% most prolific offenders and compared those in the PNC data with those in the self-reported data. I created several charts, tables and Venn diagrams showing the results.

In the second section I analysed the raw numbers of the predictive factors as well as applying various chi square analyses and a logistic regression using R.

Key Skills:

Microsoft Excel:

- Calculations: % changes, proportions and rates
- Logical formulae: If statements, OR statements, lookup, Index/match functions
- Data manipulation formulae: Countifs, Sumifs,
- Chart creation and manipulation
- Creating inputs for Venn diagrams

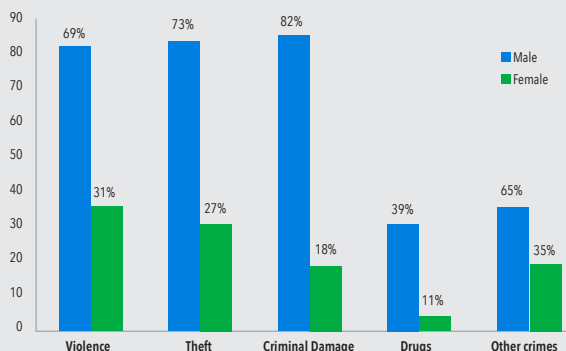
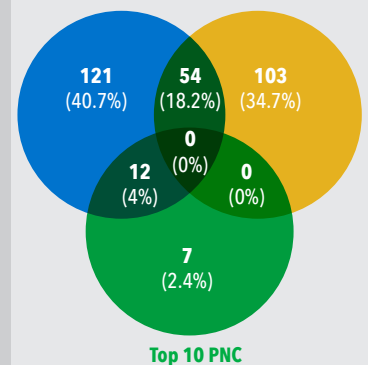
Microsoft Word and Powerpoint:

- Reporting and writing up the analysis and findings of my project.
- Presenting my project at the end of the two months using PowerPoint, showcasing some graphs and diagrams that I had created.

R:

- Applied a logistic regression to the predictor variables, including using a Wald's t test and finding Nagelkerkes R^2

Top 10% Self-rep Some PNC Crime



Proven offences received by males and females, according to crime type

Conclusion

Since the datasets that I used have not yet been published, and much of the data is confidential, I am not able to discuss the majority of my findings.

The first part of the report was completed; the second part still had some analysis left to do concerning regression as well as twin ACE modelling.

