Socio-Economic Indicators and Crime: A Regression Analysis of DV and Violent Crime (Queensland)

**Author:** Davi Santos Meloni **Date:** April 2025

**Note**: This analysis was conducted as part of a self-guided learning project. While the methods and interpretations are grounded in established statistical practices, this is the author's first attempt at modeling real-world datasets. As such, results should be viewed as exploratory and subject to the typical limitations of an individual learning exercise.

Executive Summary

*This report investigates the relationship between socio-economic conditions and crime rates, with a focus on domestic violence (DV) and violent crimes. These categories were intentionally selected due to their severe social impact and stronger ties to interpersonal and community dynamics. By excluding property and drug-related offences—often influenced by economic gain or substance use—the analysis aims to isolate how broader socio-economic disadvantage relates to crimes involving personal harm.*

*Using multiple linear regression models, we examined how three SEIFA indexes—IRSD (disadvantage), IER (education and employment), and IEO (economic opportunity)—predict log-transformed crime rates across local government areas in Queensland. The findings suggest that these socio-economic indicators are statistically significant predictors of both DV and violent crime patterns.*

*Interestingly, while disadvantage (IRSD) was negatively associated with crime—as expected—higher scores in IER and IEO (representing greater education, employment, and economic opportunity) showed a slight positive association with crime rates. This highlights the complexity of these relationships and suggests that socio-economic advantage does not always correlate with lower levels of DV or violent crime.*

*On a more personal note, I’ve recently started diving deeper into the world of statistical and predictive data analysis—and I’ve found myself really enjoying the process. This report was born from that curiosity and enthusiasm. Despite being very much a beginner-level project, creating it has been both a rewarding and highly educational experience. I look forward to building on this foundation with more advanced work in the future.*

## 1. Introduction

*The purpose of this analysis is to better understand how socio-economic factors relate to violent and interpersonal forms of crime. Domestic violence-related and violent offences were chosen due to their severe impact and strong ties to social and community dynamics. In contrast, property and drug-related crimes were intentionally excluded to provide a clearer focus on the role of community-level disadvantage without potential noise from crimes more related to economic gain or substance use.*

## 2. Data Sources and Preparation

* ***Crime Data****: Crime data was collected from 2018 through to the most recent available reporting period. Raw offence types were grouped into broader categories, with this analysis focusing specifically on:*
  + ***Violent Crime***
  + ***DV-Related Crime (Breach Domestic Violence Protection Order)***
* ***Socio-Economic Data****: SEIFA indexes (2021) were obtained from the Australian Bureau of Statistics. The following indexes were included:*
  + ***IRSD****: Index of Relative Socio-Economic Disadvantage*
  + ***IER****: Index of Education and Employment*
  + ***IEO****: Index of Economic Opportunity*

*This analysis is based on data from* ***234 Queensland police divisions****, each matched to corresponding suburb-level socio-economic indexes from the SEIFA dataset. Not all police divisions were included, as some could not be reliably matched to SEIFA suburbs. Rather than attempting to manually combine or interpolate SEIFA scores—which could have introduced bias or compromised the accuracy of the data—those unmatched divisions were excluded. After excluding divisions with minimal population, the final dataset of 234 divisions was deemed sufficient to produce meaningful insights while maintaining the integrity and transparency of the analysis.*

## 3. Methodology

***3.1 Regression Approach***

*Multiple linear regression was used to model the relationship between SEIFA scores and crime rates. Both DV-related crime and violent crime were log-transformed to improve normality and interpret results in terms of proportional change.*

***3.2 Model Design***

* ***Dependent Variables****:*
  + *log\_DV\_Crime*
  + *log\_Violent\_Crime*
* ***Independent Variables****:*
  + *IRSD Score*
  + *IER Score*
  + *IEO Score*

***3.3 Interpretation Strategy***

*The primary focus of interpretation was on:*

* *The* ***direction*** *of each coefficient (positive or negative)*
* *The* ***magnitude****, which reflects the change in log-crime rates per one-unit increase in SEIFA score*
* *The* ***p-values****, used to assess whether each socio-economic factor is a statistically significant predictor of crime rates (threshold: p < 0.05)*

*By log-transforming the dependent variables, the coefficients can be interpreted approximately as the* ***percentage change in crime rates*** *associated with a one-unit increase in each SEIFA index.*

## 4. Results

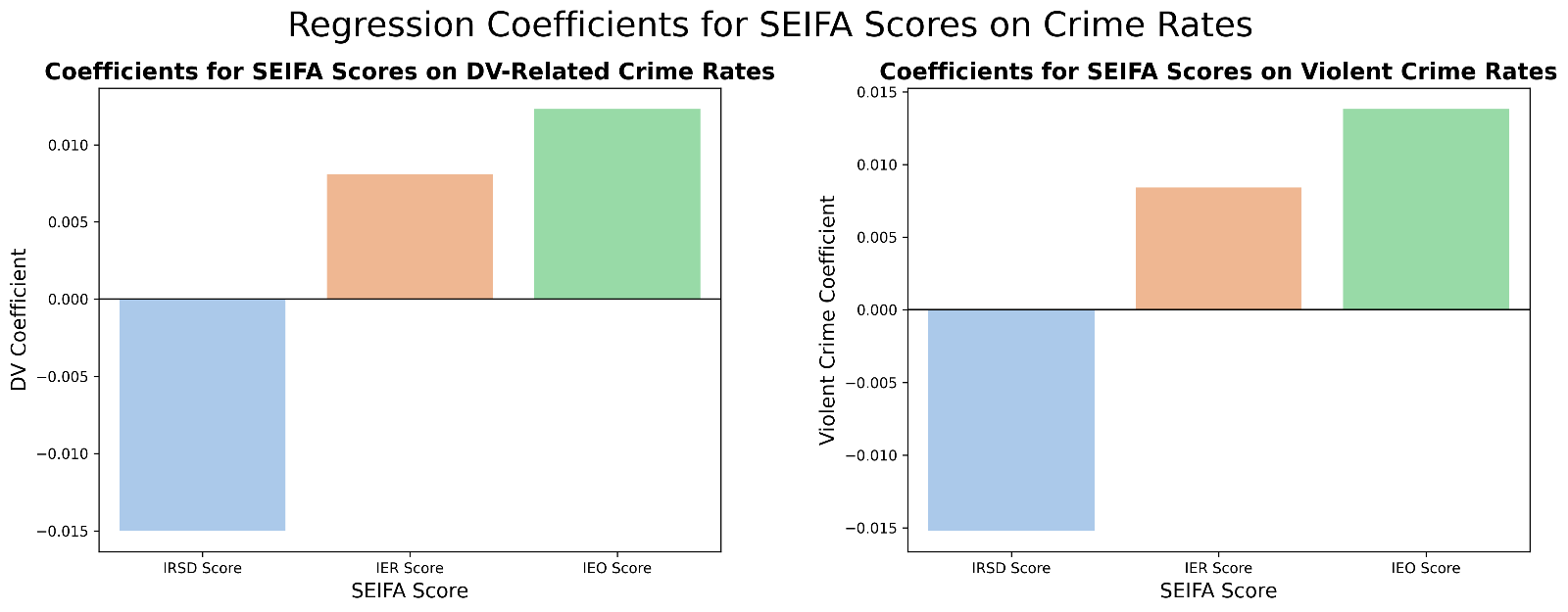
***4.1 Model Fit***

| ***Crime Type*** | ***R-squared*** | ***Adjusted R-squared*** |
| --- | --- | --- |
| *DV-Related Crime* | *0.966* | *0.960* |
| *Violent Crime* | *0.980* | *0.975* |

*The models explain over 96% and 98% of the variance in DV-related and violent crime, respectively.*

***4.2 Coefficients***

*The table below shows the regression coefficients for each SEIFA score:*

**

| ***SEIFA Score*** | ***DV Crime Coefficient*** | ***Violent Crime Coefficient*** |
| --- | --- | --- |
| *IRSD Score* | *-0.0150* | *-0.0152* |
| *IER Score* | *0.0081* | *0.0084* |
| *IEO Score* | *0.0123* | *0.0138* |

*All coefficients are statistically significant (p < 0.001).*

## 5. Visual Diagnostics

*The scatter plots below illustrate the model fit for actual vs. predicted crime rates:*

*A diagram of a crime

AI-generated content may be incorrect.*

## 6. Interpretation of Key Findings

Although the SEIFA indexes were statistically significant predictors, the overall strength of the relationships was moderate. This was evident in the spread of data points around the regression lines—while the general trend followed the model’s predictions, the points were not tightly clustered around the line. This indicates that, although socio-economic disadvantage plays a role, it is not the sole factor influencing crime rates, and there is substantial variability that may be explained by other variables not captured in this analysis.

***6.1 IRSD Score (Disadvantage)***

* *A* ***negative relationship*** *was found: areas with higher disadvantage (lower IRSD scores) tend to have higher crime rates.*
* *This aligns with expectations, as disadvantaged areas may have weaker social cohesion or fewer resources for crime prevention.*

***6.2 IER & IEO Scores (Education, Employment, Economic Opportunity)***

* *A* ***positive relationship*** *emerged: surprisingly, areas with better education, employment, and opportunity had* ***slightly higher*** *crime rates.*
* *This may be due to:*
  + *Higher population density in economically active areas*
  + *Better reporting and policing infrastructure*
  + *Socio-economic inequality or micro pockets of disadvantage in otherwise advantaged areas*

## 7. Model Validity Checks

* ***Durbin-Watson statistic****: ~2.09 — indicates no significant autocorrelation*
* ***Omnibus & Jarque-Bera tests****: suggest residuals are approximately normally distributed*

*These diagnostics confirm that the assumptions of linear regression are sufficiently met.*

## 8. Conclusion

*This analysis reveals a complex yet statistically significant relationship between socio-economic factors and* ***violent and interpersonal crime patterns*** *across Queensland. The SEIFA indexes used—IRSD (Index of Relative Socio-economic Disadvantage), IER (Index of Education and Employment), and IEO (Index of Economic Opportunity)—were all found to be significant predictors of both domestic violence-related and violent crime rates.*

*As expected, areas with higher levels of disadvantage (lower IRSD scores) tended to have higher crime rates, while crime rates decreased as IRSD scores increased. However, unexpectedly, crime rates showed a slight upward trend as IER and IEO scores increased—that is, areas with higher levels of education, employment, and economic opportunity exhibited somewhat slightly higher crime rates than those with lower scores. While the pattern was not strongly concentrated, the overall direction of the relationship was positive, suggesting a more complex social dynamic behind these forms of crime. This counterintuitive pattern may reflect underlying structural dynamics, such as differences in reporting behaviour, service availability, or population density, rather than a direct causal relationship.*

*It is important to note that this analysis focused exclusively on* ***violent and domestic violence-related offences****, deliberately excluding property, drug, and other crimes to isolate community-level social dynamics from economic or substance-related factors.*

*The model’s statistical significance and reasonable fit demonstrate that SEIFA indexes can serve as meaningful predictors in understanding these specific types of crime at a community level. While this report is not exhaustive or definitive, it offers insights that may contribute to more targeted discussions around social policy, support services, and crime prevention strategies.*

***Data Sources***

* ***SEIFA Indexes (2021)*** *Australian Bureau of Statistics*[*https://www.abs.gov.au/statistics/people/people-and-communities/socio-economic-indexes-areas-seifa-australia/latest-release#data-downloads*](https://www.abs.gov.au/statistics/people/people-and-communities/socio-economic-indexes-areas-seifa-australia/latest-release#data-downloads)
* ***Crime Data by Police Division (2018–Latest)*** *Queensland Government Open Data Portal*[*https://www.data.qld.gov.au/dataset/offence-numbers-police-divisions-monthly-from-july-2001*](https://www.data.qld.gov.au/dataset/offence-numbers-police-divisions-monthly-from-july-2001)
* ***Code Repository*** *GitHub – SEIFA & Crime Rate Analysis (Queensland)*[*https://github.com/davimeloni/qld-seifa-crime-analysis.git*](https://github.com/davimeloni/qld-seifa-crime-analysis.git)