Lucie Dumolin & Elsa Scheidler Autumn 2025

Final Paper - Introduction to Data Analysis with Python

# Criminality and Socio-Demographic Changes in France (2016–2024): A Data-Driven Analysis

Outline :

I. Introduction
II. Revue de littérature
III. Data and Methods
IV. Findings and discussion
V. Conclusion

*Note : Artificial intelligence tools were used for redactional and linguistic support in the drafting of this paper.*

## I. Introduction

In recent years, public discourse in France has increasingly emphasized the idea that criminality is on the rise and that it is closely linked to specific socio-demographic factors, such as immigration or poverty. These claims are frequently echoed in the media and political debates, yet they are often supported by limited empirical evidence or by selective interpretations of official statistics.

Within academia, criminality has long been studied through various disciplinary lenses, with a strong focus on its causes and consequences. Classical and contemporary works—such as those by Hans J. Eysenck and Gisli H. Gudjonsson—have explored psychological, social, and environmental determinants of criminal behavior. However, the systematic analysis of recent trends in criminality, particularly in relation to evolving socio-demographic characteristics, remains comparatively underdeveloped. While public authorities regularly collect and publish detailed crime statistics for administrative and monitoring purposes, relatively few studies mobilize these data to examine how changes in population size, composition, and age structure may relate to observed crime dynamics in France.

This paper seeks to address this gap by applying data analytical methods to compare official records of registered criminal offenses with national socio-demographic statistics to answer key questions:

- Has criminality been on a rise between 2016 and 2024?
- For which groups is it possible to see an increase in committing crimes?
- What variables are most closely linked to committing a crime?

The analysis focuses on France and draws on data from the Ministry of the Interior covering the evolution of criminality between 2016 and 2024. These figures are compared with population data published by the National Institute of Statistics and Economic Studies (INSEE), including indicators of population growth, demographic composition, and age distribution.

A key methodological issue concerns the definition of criminality itself. Criminality is generally understood as behavior that is prohibited or punishable under criminal law—a broad definition that encompasses a wide range of offenses and situations. Given the constraints of data availability and the scope of this paper, the analysis is restricted to the category of *mis en cause* (MEC), referring to individuals suspected of having participated in an offense without having been formally charged. While this indicator does not measure criminal behavior in a

strict legal sense, it provides a consistent and operational proxy for examining trends in recorded involvement in crime over time.

The paper proceeds as follows. After reviewing the relevant literature, the methodology and data sources are presented in detail. The results of the analysis are then discussed, with particular attention to the relationship between criminality and socio-demographic variables. The paper concludes with a critical discussion of the findings, their limitations, and their implications for public debate and future research.

## II. Literature review

**Criminality has long been a central object of research in the social sciences and has been examined through a wide range of disciplines : sociology, economics, political science and geography.**

In sociology, research has focused on the social construction of crime and deviance, as well as on the impact of different social factors in the experiences of criminality. Main contributions include studies on *The sociology of crime* (Hester & Eglin, 2017) and early analyses of gendered patterns, such as *The Criminality of Women* (Pollak, 1950). More recent sociological work has explored the relationship between crime and migration, notably through empirical analyses of large immigration waves and their effects on crime rates (Bell, Fasani & Machin, 2013).

From an economic perspective, crime has been approached as a rational behavior influenced by incentives, labor market conditions, and institutional constraints. Foundational work in this field, such as Freeman's *The Economics of Crime* (1999), has contributed to understanding criminal activity through cost–benefit frameworks and has informed public policy debates on dissuasion and prevention.

Political science has examined crime through its implications for democratic participation and governance. Studies such as Bateson (2012) have shown that crime victimization can significantly affect political behavior, including levels of participation and trust in institutions. This body of research highlights the broader political consequences of criminality beyond its immediate social and economic impacts.

Geographical approaches have emphasized the spatial distribution of crime and the role of place-based factors. Early contributions, such as Harries' *The Geography of Crime and*

*Justice* (1974), laid the groundwork for understanding how urban form, neighborhood characteristics, and spatial inequalities shape crime patterns.

**Despite this extensive body of literature across sub-fields, and the complementarity of these disciplines, relatively few studies adopt a comprehensive, country-level perspective on criminality over time.** In the French context, research on criminality is largely based on descriptive statistical outputs produced by public institutions. The Ministry of the Interior and the Ministry of Justice publish regular reports documenting trends in recorded crime, while the *National Institute of Statistics and Economic Studies* (INSEE) provides complementary socio-demographic data. However, these sources are primarily descriptive statistics and rarely accompanied by in-depth analytical academic research.

Moreover, to the best of our knowledge, France has not been the subject of a comprehensive, long-term study of criminality synthesizing trends across offense types, population, and time horizons for over 10 years. The most recent examples of such type of studies are *La criminalité en France* (Soullez, 2013) and a report from the French observatory for delinquency from 2007.

This gap in the literature suggests a need for updated, systematic analyses that move beyond institutional reporting to provide an integrated understanding of criminality dynamics.

## III. Methods and data

The starting point of this analysis is the 2025 publication of official crime statistics by the *Service statistique ministériel de la sécurité intérieure* (SSMSI), the ministerial statistical office for internal security in France. Compiled into a centralized database, these data document the evolution of several key indicators of recorded criminality over the period 2016–2024, including the number of victims, the number of *mis en cause*, and the geographical distribution of crime.

For the purposes of this study, we focus on the number of *mis en causes*, defined as individuals suspected of having participated in a criminal offense without having been formally charged. This methodological choice allows for an individual-centered analysis, which is particularly relevant when examining the relationship between criminality dynamics and socio-demographic variables.

However, this indicator also entails a number of important limitations. First, the number of *mis en cause* does not measure “actual” criminality. Not all offenses are reported, and not all offenders are identified or apprehended by law enforcement authorities. As a result, this

indicator partly reflects police activity, operational priorities, and targeting strategies rather than criminal behavior alone. Such mechanisms may introduce structural biases, which have been documented in the literature, including differential exposure to police controls related to age, gender, or perceived ethnic origin.

Second, *mis en cause* are presumed innocent and have not necessarily committed the offenses for which they are suspected. The indicator therefore captures not only recorded involvement in crime but also the dynamics of police suspicion. Finally, certain forms of criminality are likely to be underrepresented in this measure, particularly offenses that are either underreported or difficult to solve, such as violence without witnesses, sexual violence, or cybercrime.

Despite these limitations, the use of *mis en cause* does not aim to capture criminality in its entirety. Rather, it provides a consistent and individual-level indicator of recorded involvement in crime, making it particularly suitable for analyzing the links between criminality trends and socio-demographic change, provided that its institutional and procedural biases are explicitly acknowledged.

Among the seventeen offense categories available in the SSMSI database, this analysis focuses on six specific types of infractions:

- homicides;
- sexual violence;
- domestic and intrafamilial violence;
- non-violent theft;
- drug trafficking;
- drug use.

These six categories allow for a broad and contrasted overview of criminality in France, encompassing both crimes and misdemeanors, frequent and rarer offenses, and a wide range of profiles among both victims and *mis en cause*.

A final methodological consideration concerns the age of *mis en cause*. Due to data accessibility constraints and statistical confidentiality rules, the analysis of the relationship between criminality and age is restricted to adults aged 18 to 59. While this limitation excludes minors and older individuals, it ensures a sufficient level of statistical robustness and comparability across the period under study.

## IV. Data Analysis

The details of our analysis can be found under the following link:

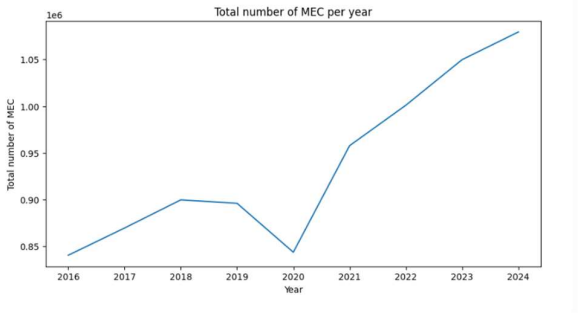
https://github.com/luciedum-dot/Intro-to-data-analysis-with-Python---Final-Project

Part 1 : Has criminality increased between 2016 and 2024?

In this section, we analyze the evolution of the MEC, first overall and then by type of offence, and compare it with trends in the total French population. While an increase in the absolute number of MEC may indicate a rise in criminal activity, comparing these figures with population dynamics allows us to assess whether this growth also corresponds to an increase in relative terms, that is, in proportions rather than in absolute values.

1. Trend of the total number of MEC per year, 2016-2024

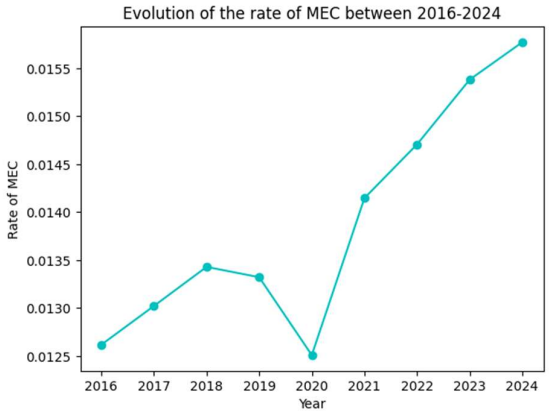
By processing the data, we can examine the evolution of the total number of MEC between 2016 and 2024. A first observation is that, despite a generally stable upward trend over the period, there is a marked decline in the number of MEC in 2020. This sharp decrease can reasonably be associated with the COVID-19 pandemic and the lockdown measures implemented during that year, which substantially altered social interactions and mobility patterns.




2. Comparison between the total number of MEC and the population

The evolution of the rate of MEC reveals a clear and sustained upward trend between 2016 and 2024. Since this rate corresponds to the number of MEC relative to the total population, its increase indicates that the number of MEC is rising even as the population grows. Although a temporary decrease is observed in 2020, likely due to the Covid-19 pandemic, the rate rises sharply from 2021 onward. This post-pandemic increase more than compensates for the earlier decline, leading to a higher MEC rate in 2024 than in all previous years. Consequently,

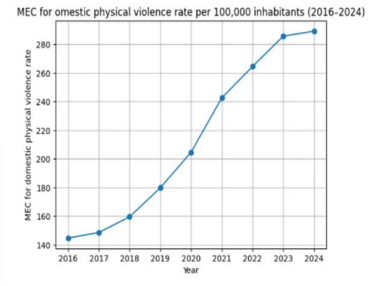
the increase in the MEC rate cannot be explained solely by population growth and instead suggests a genuine rise in the prevalence of criminality over time.

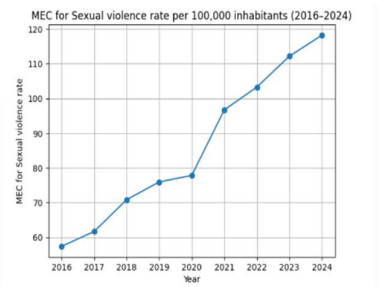


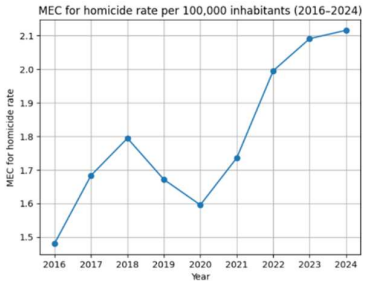

3. Evolutions of the number of MEC by types of offenses

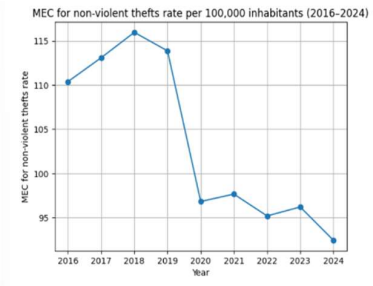
To understand better the evolution of criminality, we looked at the evolution of 6 different types of infractions over the eight years. To observe if the evolution in the number of MEC for different types of offenses is rising regardless of the evolution of the population (and thus to avoid this bias), we calculated the number of MEC for each offense as a rate per 100,000 inhabitants in France.

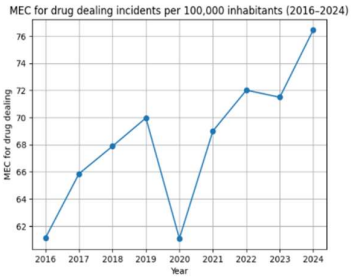
An analysis of the evolution of different types of offences reveals several key patterns. Trends vary substantially across offence categories. While drug-related offences massively dropped during Covid, they increased sharply afterwards. At the same time property offences (non-violent thefts) follow an opposite trajectory. By contrast, offences against physical persons, and especially sexual and domestic violence kept a rather stable rate of growth.

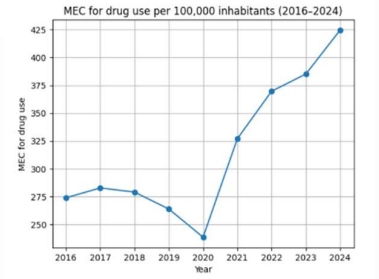





It is also important to underline a difference between the trends for the 3 types of offenses against persons: while a decrease in the rate of MEC for homicides can be noted for 2020, as for the other types of offenses analyzed, it is not the case for sexual nor physical domestic violence. The lockdown measures implemented during the pandemic can here again reasonably explain these findings: while offenses which most often happen outside of one's home, such as drug dealing or theft, decrease, sexual and physical domestic violence occur within the home, with thus their trend remaining stable or even increasing.

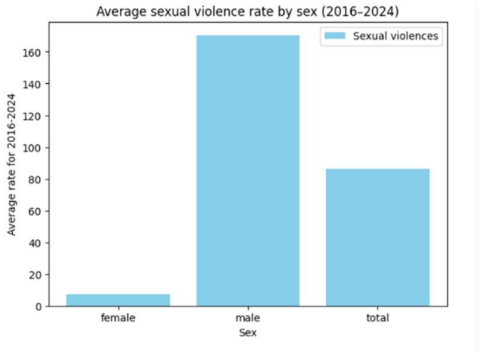
The rise in the global rate of MEC can largely be explained by an increase in most categories of offenses, with the notable exception of thefts. However, it is important to stress that an increase in the MEC rate for a specific offense does not necessarily imply that the underlying offense rate itself has risen. In the case of sexual violence, for instance, the emergence of the #MeToo movement during the period under study likely heightened awareness among victims and within society more broadly, which may have led to increased reporting rather than a genuine increase in the incidence of such offenses.

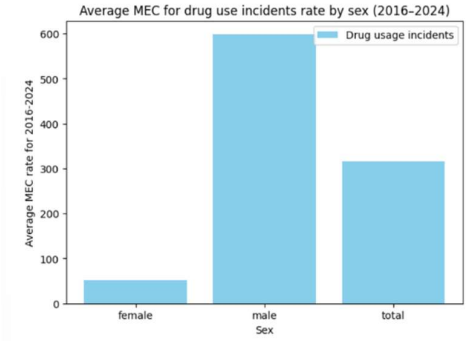
Part 2 : What variables are most closely linked to committing a crime?

1) Analyzing the differences in committing offenses according to sex and nationality

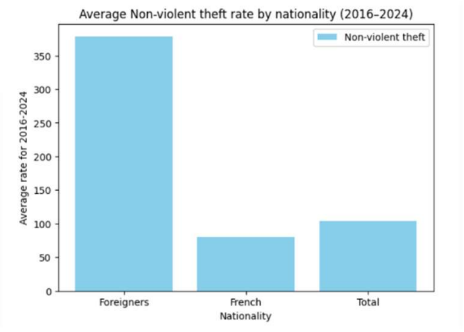
Marked disparities in MEC rates are observed across offense categories when stratified by gender, nationality and age. Two robust patterns emerge from the data.

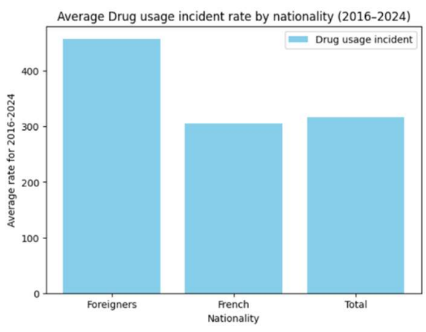
First, MEC rates are consistently higher for men than for women, with particularly pronounced gender gaps in offenses involving severe violence and drug-related activities. For sexual violence, the MEC rate for men is nearly 25 times higher than that for women, while homicide rates are more than six times higher for men. Similarly, men exhibit substantially higher involvement in non-violent theft (approximately four times higher), physical domestic violence (around five times higher), and drug-related offenses, with MEC rates exceeding those of women by more than a factor of eleven for both drug dealing and drug use.




Second, MEC rates are systematically higher for foreigners than for French nationals, although these differences are more moderate than those observed by gender and vary across offense types. The largest nationality-based disparities are found in non-violent theft, where the MEC rate for foreigners is slightly more than four times higher than that for French nationals, and in homicides, physical domestic violence, and drug dealing, for which rates are approximately 2.5 times higher. In contrast, the differential is smaller for sexual violence (less than a twofold difference) and for drug use (around 1.5 times higher).



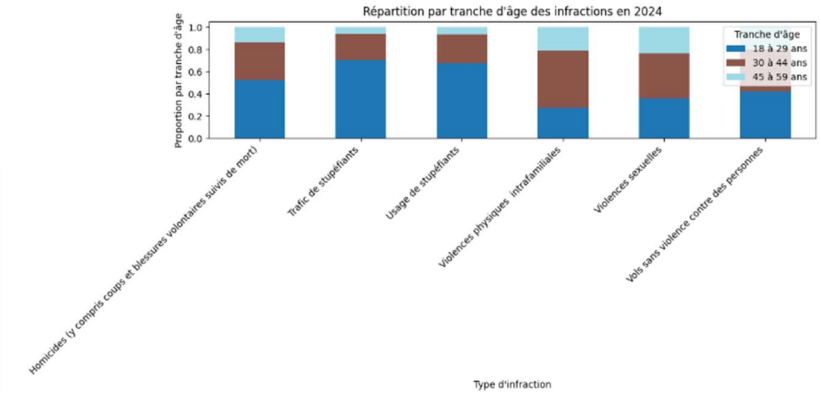
Finally, these results must be interpreted with caution. MEC captures recorded involvement in offenses rather than underlying criminal behavior and may therefore be influenced by a range of confounding factors, including unequal exposure to law enforcement, socioeconomic vulnerability, and potential discriminatory practices—factors that are particularly salient when comparing foreigners and French nationals[1].

2) Analyzing the difference in age categories

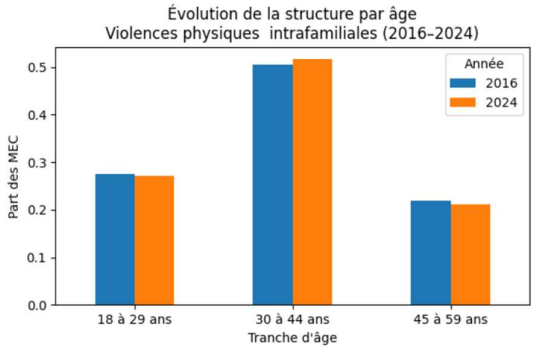
Due to discrepancies between the age group classifications used by the French Ministry of the Interior and the National Institute of Statistics, it was not possible to directly compare the age distribution of MEC with that of the general population. Consequently, the analysis focuses on temporal changes between the years 2016 and 2024.

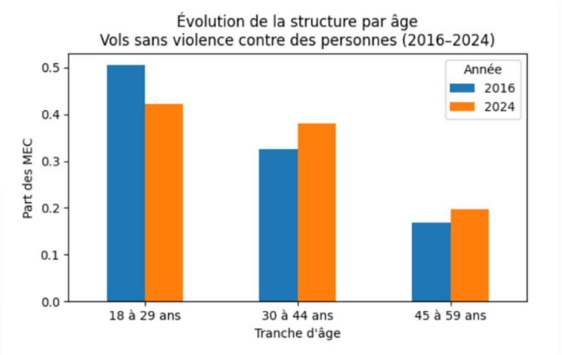
The age distribution of MEC varies substantially across offense types. While MEC involved in sexual violence are relatively evenly distributed across age groups, individuals aged 18–29 are markedly overrepresented in offenses related to drug trafficking and drug use. More generally, even in the absence of precise population data, younger age groups appear to be overrepresented across most offenses, with the notable exception of intrafamilial violence.

A disaggregated, offense-by-offense analysis further reveals that age distributions remain largely stable over time between 2016 and 2024. One notable exception concerns non-violent theft against persons, for which the share of MEC aged 18–29 declines over the period, while the proportions of individuals aged 30–44 and 45–59 increase. A similar, though less pronounced, trend is also observed for drug use offenses.




[1] https://www.cepii.fr/blog/bi/post.asp?IDcommunique=982 ; https://www.histoire-immigration.fr/societe-et-immigration/y-a-t-il-un-lien-entre-delinquance-et-immigration




Distinct offense profiles emerge when comparing groups by gender, nationality, and age. Among women, the highest MEC rate per 100,000 inhabitants is observed for physical domestic violence, a pattern that contrasts with men, for whom drug use offenses display the highest rate.

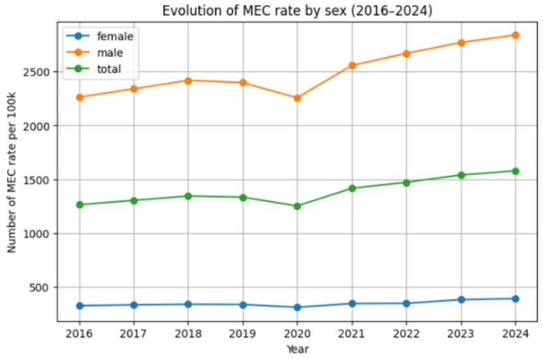
Similarly, nationality-based differences are evident: among foreigners, physical domestic violence is the offense category associated with the highest MEC rate, whereas among French nationals, drug use offenses dominate.

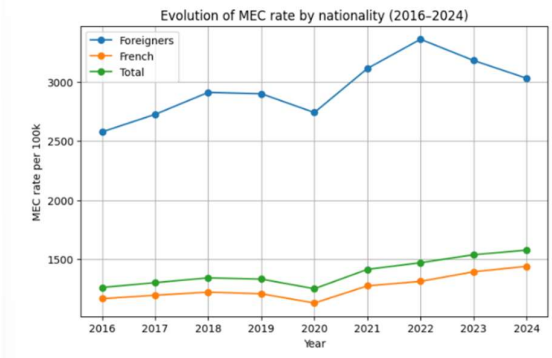
Age-related patterns further highlight differentiated profiles of recorded criminal involvement. While younger individuals are disproportionately represented in drug-related offenses, middle-aged groups show the highest rates in intrafamilial violence, and older individuals (aged 45–59) exhibit the highest rates for sexual violence.

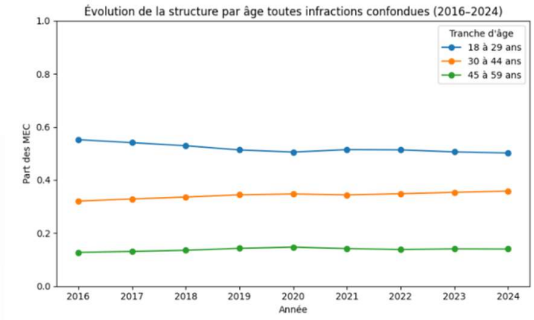
Part 3 : For which groups is it possible to see an increase in committing crimes?

The increase in the MEC rate for sexual violence over the period is primarily driven by a rise in the rate among men. While the MEC rate for women remains relatively flat and stable throughout the period, a clear upward trend is observed for men. Moreover, the rate for women appears less sensitive to short-term contextual shocks, such as the COVID-19 pandemic, whereas fluctuations are more pronounced among men.

More broadly, the increase in MEC rates across offenses is more pronounced among foreigners than among French nationals over most of the period. However, this trend reverses in the final two years of observation, during which MEC rates for foreigners decline, in contrast to the continued increase observed among French nationals. This divergence suggests distinct temporal dynamics by nationality, potentially reflecting differences in exposure, enforcement practices, or broader social and institutional factors. There is no significant evolution observed for the age-variable.




## V. Conclusion

This study provides a nuanced picture of recorded criminal involvement in France between 2016 and 2024, showing that changes in criminality are neither uniform across offenses nor evenly distributed across social groups. Overall, while population growth contributes to the increase in the number of *mis en cause*, the rise in the rate of *mis en cause* indicates a generalized increase in recorded criminal involvement in relative terms over the period. However, important heterogeneities emerge by offense type, gender, nationality, and age. Men, younger individuals, and foreigners are disproportionately represented among *mis en cause*, though with distinct offense profiles and temporal dynamics. In particular, the increase in recorded sexual violence is largely driven by men, while nationality-based trends diverge in recent years. These findings underscore the importance of interpreting *mis en cause* as an

indicator of recorded involvement shaped by institutional, social, and reporting mechanisms rather than as a direct measure of underlying criminal behavior. Future research would benefit from integrating victims data to better capture the full dynamics of crime and social exposure.