THREE MONTHS OF MURDER: A COMPUTATIONAL ANALYSIS OF THE REPRESENTATION OF MURDER IN SOUTH AFRICAN NEWS MEDIA

A research report presented to the School of Computer Science and Applied Mathematics at the University of Witwatersrand in partial fulfilment of the requirements for the degree of Master of Arts (M.A.) in e-Science

by:

Inonge Sepo Lupinda

Supervised by:

Dr Nechama Brodie

Dr Biandri Joubert

Johannesburg, South Africa

2025

Abstract

The way crime, particularly murder, is covered in news media plays an important role in shaping public perception and influencing policy debates. However, biases in reporting can mislead the understanding of crime trends and their broader societal impact. This research examines South African news media coverage of murder over a three-month period, using computational methods to analyse the factors that influence media attention. Specifically, it investigates the demographic and circumstantial characteristics of murder reporting, including race, age, gender, methods of murder, geographic location, victim-perpetrator relationships, and legal outcomes. These findings are then compared with actual crime statistics to assess differences in reporting. The results indicate instances of both over- and underreporting by media outlets, suggesting potential biases in coverage. This research highlights the value of computational techniques in criminology research in extracting meaningful insights from annotated news data. It contributes to a more nuanced understanding of how murder is represented in South African media.

Keywords: South African news media; R programming; murder reporting; quantitative analysis; computational analysis.

Acknowledgements

This journey has been one of growth, challenge, and learning, and I am incredibly grateful to everyone who has supported me along the way. This research would not have been possible without the work of journalists, researchers, and law enforcement officials who document and investigate crime in South Africa. I also acknowledge the South African Police Service (SAPS) and crime analysts whose data and investigations formed the foundation of this research. Their efforts, often under difficult conditions, help shape our understanding of crime and public safety. Most importantly, I want to acknowledge the victims of murder whose stories are reflected in this research. Behind every statistic are real people, families, and communities who have experienced loss and are seeking justice. I hope this research contributes to important conversations about crime, justice, and media representation.

I give all glory and thanks to God, whose grace, guidance, and strength have carried me through this journey. To my family, my parents, Miaze Maboshe-Lupinda and Justin Lupinda, my siblings, Mwika Lupinda and Samuel Lupinda, and my family friends, Sheba Kuwani-Mulenga and Kalenga Mulenga, Anyway Shonhiwa, Henry Kambafwile and Reegan-Mwansa Kambafwile, thank you for your love, sacrifices, guidance, encouragement, and unwavering belief in me. This achievement is as much yours as it is mine. A special acknowledgement to my grandfather, Patrick Maboshe, and my grandmothers, Zondiwe Maboshe and Catherine Sikazwe. Your wisdom and support have been a guiding light in my life, and I carry your lessons with me every day.

To my incredible friends, special mention to Chona Simfukwe, Fhatuwani Tshikungulu, Gloria Matondo, Joel Amoni, Keamogetswe Molefe, Kalaba Kambafwile, Lisa Sinkala, Salifya Sinkala, Sitwala Mundia, and Vanessa Seme, thank you for being my constant cheerleaders, my sounding boards, and my source of strength. Your unwavering support, laughter, and encouragement have made this journey all the more meaningful.

I would also like to thank my sponsors, The National e-Science Postgraduate Teaching and Training Platform (NEPTTP) and the Department of Science and Innovation (DSI), for making this opportunity possible. To the University of the Witwatersrand, thank you for providing the space and resources that have shaped my academic journey.

Finally, my gratitude to my supervisors, Dr. Nechama Brodie and Dr. Biandri Joubert, for your time, guidance, patience, and invaluable mentorship. Thank you for pushing me to think critically and engage deeply with this research. I also extend my thanks to Professor Rod Alence, my course coordinator, and my mentors, Dr. Colin Tredoux and Anelda van der Walt, for showing me how valuable these technical skills are in creating meaningful impact and for fostering an academic environment that allowed me to grow and thrive.

Coming from a social sciences and humanities background, I never imagined myself transitioning into technology, data analysis, and coding in R. This journey has opened new opportunities and created a learning space I never thought possible. I am deeply grateful for all the support that made this transition possible. To other mentors, colleagues, and everyone who has supported me in ways big and small, thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

News media has a well-documented role in shaping public perceptions of crime, both globally and within the South African context (Brodie, 2021). Murder reporting is a key area where news media coverage can influence public understanding of crime trends, often affecting public discourse, policy-making, and societal attitudes toward violence. However, the reporting of crimes varies due to the selective and subjective nature of news coverage (Silva & Guedes, 2022). The media's focus on murders with specific characteristics is sometimes driven by their attractiveness to news consumers, representing a potential profit for news reporters (Huff-Corzine et al., 2014). Thus, there is selective reporting on crime, and murder stories attract more attention than other crimes (Liska & Baccaglini, 1990). Crime news, commodified for commercial appeal and readership engagement (Robinson, 2011; Mayr & Machin, 2011), prioritises mass murder and murder as attention-grabbing content, resulting in selective coverage due to newsworthiness considerations (Schildkraut et al., 2018).

In this research, ‘murder’ should be understood to imply an unlawful killing in the context of the South African legal system (Brodie, 2019). This definition does not include ‘culpable homicide’, which in South African law, occurs when a person unlawfully and negligently causes the death of another (Snyman, 2014). South Africa remains among the countries with the highest incidents of murder (Bowman et al., 2022). Violent crime, such as murder, is prevalent across South Africa, with certain areas bearing a heavier burden than others (SAPS, 2023). According to the SAPS 2022/2023 Annual Crime Statistics Report, the Eastern Cape reported 71 murders per 100,000 people, making it the province with the highest murder rate in South Africa. KwaZulu-Natal and the Western Cape both reported 56 murders per 100,000 people.

Violence does not occur equally across South Africa, but is more frequent within certain race, gender, and age groups. In South Africa, young adult males are more likely to be both perpetrators and victims of violent crimes (Thomson, 2016). Classified according to race, the Coloured[1] population has one of the highest murder rates in South Africa, with most murders being intra-racial and committed by someone known to the victim or living within the community (Thomson, 2016). Given the high rates of murder in South Africa, it is important to analyse how the media portrays violence, as the media can impact and influence public perception (Jewkes, 2015). It is also important to investigate which types of murders are mostly covered and selected by South African news media outlets (e.g. infanticide, femicides, murders, or celebrity or high- profile murders). When it comes to murder itself, the media tends to sensationalise stories, or choose which stories are more newsworthy than others, thus contributing to either an inflated or deflated perception of their frequency and severity (Kappeler & Potter, 2017; Thomas et al. 2021). Understanding the influence of the media coverage of murder helps to shape policy and practice in crime prevention and victim support, as well as identify areas that need more media attention, such as marginalised groups (i.e. individuals or communities who are socially, economically, or politically disadvantaged, including women, children, LGBTQI+, and ethnic minorities (Posel & Casale, 2019)). This research analyses the news coverage of murders over a three-month period to explore characteristics of factors shaping news media portrayals of murder.

1 In South Africa, racial categories such as Black, White, Coloured, and Indian remain in use across official statistics, legal frameworks and academic research. These classifications are rooted historically in apartheid-era systems and continue to be recognised by the South African government and institutions like Statistics South Africa and the Constitution of the Republic of South Africa (1996).

Problem Statement

Murders occur frequently in South Africa, yet they are not equally covered in news media. Research regarding the representation and quantification of how many murders are reported within the time period of this study, as well as the reported characteristics and features of these murders in South Africa, is limited. The lack of coverage creates a gap in our understanding of how and why certain cases receive media attention while others are ignored. By analysing media coverage over a specific period of three months from January 1st to March 31st, 2023, this research aims to identify patterns in how different murder cases are portrayed.

Research Aims

- To investigate South African news media murder reporting through computational analyses using R Programming, including data visualization, regression modelling and text mining techniques.
- To determine the type of murders covered by the press within three months, as accessed from NewsBank[2] (from January 1st to March 31st, 2023).
- To investigate factors and characteristics such as race, age, gender, modes of murder, geographical location and victim-perpetrator relationship as reported in the media stories.

[2] NewsBank is an online archive and database that has a collection of news sources from a variety of topics and issues. It contains both international and local news titles, and includes advanced search features such as date, location, and topic.

Research Objectives

1. What specific themes, narrative, and contextual factors emerge in South African news Media reporting of murders obtained from NewsBank between January 1st and March 31st, 2023?
2. What are the demographic and circumstantial characteristics of murder victims as reported in the media, and how do these characteristics relate to the geographical location of the murders and the subsequent legal proceedings? (i.e. the gender, age, and race distribution of the victims, the geographical location of the murders, victim-perpetrator relationships, modes of murder, dates of murder and justice events such as police investigation arrests, bail hearings, court appearances, pretrial proceedings, and trial proceedings)

Relevance and Importance

This research project used data from NewsBank, an online database containing newspaper resource records from South Africa, to analyse the news coverage of murders over a three-month period. This research aimed to explore how different demographic factors (such as age, race, and gender of murder victims) are represented in murder news coverage, as well as the prevailing modes and locations of murders, during a three-month period from January 1st to March 31st, 2023. Through a systematic analysis of news reports, the research sought to identify patterns and themes in the portrayal of these factors, providing insights into the ways in which media coverage reflects or shapes public perceptions of crime. The findings of this study will contribute to ongoing studies on news reporting concerning murder, considering various demographic factors and contextual elements, the identification of patterns and the dynamic influential factors in media reporting may be valuable data for other researchers or further research.

The research also includes a computational dimension, using R Programming, which is underutilised in South African criminology studies, which will add to the existing body of research that is often conducted using traditional tools (such as manual content analysis, interviews, qualitative coding or Excel).

CHAPTER 2: LITERATURE REVIEW

This chapter is a review of literature that explores murder reporting in South African news media, its historical context, trends, and demographic factors that shape coverage. It discusses the role of the media post-apartheid, the portrayal of crime and violence, and the influence of societal issues such as racial inequality, gender biases, and socioeconomic disparities on media reporting. By analysing both qualitative and quantitative studies, this review aims to provide an understanding of how murders are represented in the South African media and the implications of this coverage for public perception.

The Role and History of South African News Media

The South African media landscape has undergone significant transformations, particularly following the transition from apartheid to a democratic government in 1994 (Wasserman, 2020). This historical shift has had implications for the role of the media in society. Apartheid was characterised by media censorship and restrictions on reporting, particularly regarding resistance to the regime (Lloyd, 2013). However, the advent of democracy brought about constitutional protections for freedom of expression, including press freedom, as detailed in Section 16 of the Constitution of the Republic of South Africa, 1996 (1996, s. 16). Wasserman (2020) highlights the importance of understanding these historical factors in analysing the media's role in contemporary South African society. The role of the media is not limited to only informing the public but also contributing to the consolidation of democracy (Arendt et al., 2023). The South African Press Council advocates for ethical journalism and self-regulation, emphasizing the role of the media in realizing the democratic promise (Wasserman, 2020).

In this context, crime reporting plays an important role. Du Plessis (2022) argues that crime reporting serves as a surveillance function, providing the audience with practical information about threats and crime events. This form of reporting is essential for public awareness and safety, as it helps inform and educate the public about criminal activities. The media's role in reporting crime is not just about relaying facts; it also involves shaping public perception and contributing to the societal discourse on crime and safety (Anand & Taneja, 2024).

Recent studies have explored the complexities of crime reporting in South Africa. For example, Chiumbu and Moyo (2022) examine the challenges faced by journalists in reporting crime in a context marked by high crime rates and social inequalities. They highlight the ethical dilemmas and pressures that journalists encounter, including the need to balance sensationalism with responsible reporting. Hunt and Jaworska (2019) analyse the representation of crime in South African newspapers, particularly pertaining to gender, nationality and race, revealing patterns of sensationalism and bias. These studies show the multifaceted nature of crime reporting in South Africa and its implications for public perceptions of crime and safety. Ethical journalism and responsible reporting are essential in this context, as they contribute to the media's role in supporting democracy and social cohesion (Kovach & Rosenstiel, 2014; Christians et al., 2009). Responsible reporting on sensitive issues like murder is particularly critical, as it shapes public perceptions of crime, justice, and safety, ultimately influencing policy and community relations (Surette, 2015).

Trends in the Media Coverage of Murder in South Africa

Violent crime has been dominant in news reporting (Berns, 2017) and is a significant contributor to unnatural death (Matzopolous, 2015). How the media presents victims and perpetrators is influenced by several demographic factors, including age, race, gender, and the nature of crime (Surette, 2015). The historical context of apartheid and ongoing disparities in the treatment of different races and genders are also evident in media coverage (Phaswana, 2021; Brodie, 2019).

Intimate femicide, the killing of women by their intimate partners, is a pervasive and critical issue in South Africa. Research indicates that it is the leading cause of female homicide in the country, highlighting the urgent need for addressing gender-based violence (Mathews et al., 2014). Media coverage of such cases often highlights the perpetrators' attempts to assert dominance and control, reflecting broader societal issues of gender inequality (Mathews et al., 2014). Similarly, sexual murders of women and children are widespread, linked to societal norms that tolerate sexual and gender-based violence (Abrahams et al., 2017). The media's role in reporting these crimes is important for raising awareness and advocating for change.

Various factors, including age, socioeconomic status, and educational attainment, were found to independently predict the risk of being a victim of murder (Seedat et al., 2009). For example, individuals living in urban informal settlements, often characterised by higher crime rates and limited access to resources, are at greater risk of homicide (Shaw, 2002; Seedat et al., 2009). Women in South Africa are six times more at risk of being killed compared to the world average (Lindegaard, 2017). However, it is important to note that men in South Africa are also significantly vulnerable, both as victims and offenders. In South Africa, some murders often occur

within the context of arguments, in public spaces, and between strangers (Lindegaard, 2017). Specifically, dispute-related killings of men in the country are driven by a desire to assert dominance and control over individuals who challenge the offender in public (Lindegaard, 2017).

Additionally, factors such as low standards of education, widespread alcohol abuse, a lack of social and vocational skills, inadequate housing and living conditions, and poor parenting skills contribute significantly to the socioeconomic environment that fosters violent crime (Christodoulou et al., 2019; Seedat et al., 2014). These factors collectively contribute to the challenging socio-economic environment that fosters criminal behaviour. In 2009, a study of female and child murders in South Africa revealed patterns in sexual homicide. Out of 2,670 cases of adult female murders, 494 (19.8%) were classified as sexual murders. Similarly, among 1,277 child murders, 104 (8.7%) were identified as sexual murders (Abrahams et al., 2017). Strangulation emerged as the most common cause of death in both adult female and child sexual murders. Notably, among child victims, sexual murders were overwhelmingly gendered: only one boy was identified as a victim of sexual murder, compared to 92 girls (Abrahams et al., 2017). The study also found that perpetrators of sexual murders were rarely strangers highlighting the prevalence of familiar or known offenders. Interestingly, there was no significant disparity in conviction rates between sexual murders and non-sexual murders for both adult females and children, suggesting that the legal system treated these cases with similar seriousness (Abrahams et al., 2017).

The trends in media coverage of murder in South Africa are influenced by multiple factors, including social, historical, and cultural factors. The portrayal of violent crime in the media often reflects broader societal norms and the country's historical legacies of inequality and violence. While the media serves a critical role in raising public awareness and shaping discourse around

crime, its coverage is frequently criticised for sensationalism, bias, and ethical shortcomings. These challenges not only distort public perceptions of crime but also risk perpetuating stereotypes and undermining the public's trust in media institutions.

Newsworthiness

Newsworthiness refers to the criteria used to determine whether a story, incident or event is considered worthy of being reported in the news (Boukes et al., 2020). Shoemaker (2006) argues that "news" is a social construct and a commodity, while "newsworthiness" is a cognitive construct and a mental judgment. Newsworthiness is not necessarily a reliable predictor of which events make it to the news and how they are covered, but one of the multiple factors that influence news coverage and how incidents are reported (Shoemaker, 2006). The concept of newsworthiness, which determines the worthiness of a story for reporting, is influenced by various factors including proximity, prominence, human interest, and sensationalism (Boukes et al., 2020; Shoemaker, 2006). Building on these ideas, Harcup and O'Neill (2017) propose an updated framework that includes additional dimensions such as relevance, entertainment, and surprise, offering a more nuanced understanding of how news values shape media coverage.

Audience preferences drive newsworthiness and influence the coverage of murder stories based on location. For example, *Pretoria News* may prioritise or emphasise murder cases occurring in Pretoria and its surrounding areas to align with the interests of its target audience (Shoemaker, 2006). Another factor is the level of violence and brutality involved in the crime, as seen in the prominent murder cases of Uyinene Mrwetyana and Anene Booysen (Brodie, 2022),

as well as the identity of the victim, with media coverage being intensified when the victim is a well-known figure (Brodie, 2021).

Research has explored the dynamics of newsworthiness in the South African media. For instance, Wasserman (2020) investigates the role of race and class in the media coverage of crime, highlighting the disparities in reporting based on the victim's background. Another study by Sutherland et al. (2016) examines the portrayal of violence against women in South African newspapers, revealing a focus on sensational and extreme cases, often at the expense of more nuanced reporting. Additionally, murders committed by perpetrators who are unknown to the victim receive more coverage. In Los Angeles, Petersen (2014) found that the level of economic disadvantage and minority residents around the crime scene neighbourhoods negatively affect the presence or absence and rate of newspaper coverage, and the amount of coverage given to victims of murder (cited in Brodie, 2019). Media research also suggests that certain characteristics make victims more newsworthy than others (Gruenewald et al., 2013). Less than 20% of female murders or femicides in South Africa are reported in the media, whereas more than 70% of (White) farm murders receive press coverage (Brodie, 2020).

The term ‘ideal victim’ refers to victims that hold specific characteristics which remove the label of blame in their victimization (Bosma et al., 2018; Bradford-Clarke et al, 2022; Christie, 1986). This means portraying victims in a way that encourages sympathy from readers, often emphasizing their innocence and undeserved suffering (Bosma et al., 2018). Ideal victims are typically White, elderly, or female. For example, elderly women are ‘blameless and passive’, through which society may label them as a ‘legitimate victim’ (Brodie, 2020). On the other hand, unworthy victims are usually blamed for their victimization.

These include victims that are involved in risky behaviour such as “drinking, using drugs, dressing provocatively or not conservatively, and especially if she engages in sex for money” (Gilchrist, 2010). These victims are portrayed as being complicit in what happens in their victimization (Brodie, 2020). The media's portrayal of these murders also fails to convey the systemic nature of violence against women, which is further entrenched by racial and class-based oppression.

Responsible reporting is important for accurately portraying intimate femicide and addressing biases in crime coverage (Spies, 2020). While most murders of women are committed by intimate partners, the media often fails to highlight this reality (Spies, 2020). Instead, crime stories that make the news tend to focus on familial killings (Du Plessis, 2022) or violent mass murders, especially those involving a single gunman and multiple victims. Crimes that seem unusual or involve vulnerable victims are more likely to get attention, while more common but equally serious cases, like intimate partner femicide, often receive far less coverage.

Newsworthiness in South African murder reporting is a complex construct influenced by a range of factors, including societal norms, media biases, and audience preferences. Addressing the challenges of sensationalism and inequality in reporting is crucial for a more balanced and accurate portrayal of crime in the media.

Quantitative and Computational Studies on the News Coverage of Murder

Quantitative research in crime reporting often involves content analysis, where researchers systematically categorise and analyse the content of media reports to identify patterns in the portrayal of crime (Riffe et al., 2019). A study by Cornelissen et al. (2019) used topic modelling to analyse South African news articles, uncovering biases in media coverage and identifying topics

that were over- or under-reported. Additionally, recent studies in South Africa have used geospatial analysis and visualization techniques to map murder rates and change-point analysis to detect abrupt changes in murder trends, thereby revealing patterns in crime distribution across different neighbourhood and providing early warning for potential spikes in violent crimes (Monyeik et al., 2020; Peterson et al., 2021).

Computational methods are techniques used to analyse complex systems and predict outcomes based on existing data. These approaches typically use mathematical models or tools to simulate, investigate and forecast intricate systems (Kryzwanski et al., 2024). Computational methods, including natural language processing (NLP) and machine learning (ML)[3], have been used to process large datasets of media content, enabling the analysis of complex patterns and trends over time and across different media outlets (Conway & O'Connor, 2016). Besides ML and NLP, computational methods consist of a broad range of techniques, such as text mining, statistical modelling, and content analysis. For distinction, in this research, computational analysis primarily refers to the use of statistical tools and text processing techniques in R to analyse crime-related news reports, rather than advanced machine learning models.

Media coverage serves as a valuable repository of murder data, enabling researchers to examine the evolution of crime and public perceptions surrounding it (Brodie, 2020). While media reports on crime have limitations and cannot solely address data challenges, they present a significant and easily accessible resource (Brodie, 2020). Collecting comprehensive murder data

[3] Natural Language Processing (NLP) and Machine Learning (ML) are subfields of artificial intelligence (AI). NLP focuses on enabling computers to understand and generate human language, while ML involves developing algorithms that learn from data to make predictions or decisions (Shoenbill, Kasturi, & Mendonca, 2023). Both are used in various applications, including speech recognition, sentiment analysis, and language translation.

is inherently challenging, as researchers depend on post-incident records from judicial systems, law enforcement, medical facilities, and mortuaries. However, such data is often unreliable due to systemic issues like underreporting, incomplete records, and data erasure, particularly in high-crime countries where marginalised groups may be disproportionately affected (Brodie, 2020).

In terms of analysis, traditional approaches to data analysis in crime-related studies have been explored in South Africa, but the potential of computational methods approaches has not been fully used. Extensive centralised databases are maintained by SAPS and other crime-related agencies in South Africa, encompassing a wide range of information that can be used to analyse criminal activities across the provinces. Contemporary researchers in South Africa, such as Cornelissen et al. (2019) and Monyeki et al. (2020), are increasingly using modern digital techniques, including statistical programs and computational methods, to enhance traditional manual coding and content analysis. These approaches, such as document term matrices, sentiment analysis, and topic modelling, enable the computational analysis of news reporting on a larger scale. For instance, Cornelissen et al. (2019) explored visibility and tonality bias in South African news articles, acknowledging both the advantages and limitations of computational methods. In this research, a total of 38 topics were extracted from the corpus using topic modelling, and sentiment analysis was performed for each of these topics. The findings of the study revealed instances of both over- and under-reporting by media houses on specific topics, which suggests the presence of potential biases in their coverage (Cornelissen et al., 2019). The identification of various tonality biases by these media houses further indicates a potential lack of objectivity in their reporting.

Consequently, the results of the study challenge the assumption of unbiased and objective reporting by news media in South Africa, highlighting the need for caution (Cornelissen et al., 2019).

The presence of biases in reporting and tonality raises concerns regarding the potential influence of external political or corporate interests on the news coverage provided by these media houses. To gain a deeper understanding of the underlying factors contributing to these biases and their implications for the media landscape in South Africa, further analysis and investigation are warranted.

The research collectively highlights the potential usefulness of computational techniques in yielding meaningful information from vast datasets, offering new perspectives and crucial insights into the complex landscape of murder in South Africa.

Demographic and Circumstantial Characteristics of Murder Reporting

While this research refers to categories such as “Black”, “White”, “Coloured”, and “Indian”, it does so with full recognition that these are socially constructed racial classifications rooted in South Africa’s apartheid history. These labels are used by official institutions like SAPS and Stats SA for statistical purposes and do not reflect biologically essential characteristics. Research indicates that murders involving Black victims are less likely to be covered extensively in the media compared to those involving White victims, reflecting broader societal disparities (Wasserman, 2020). In South Africa, there is a gap in research to refute or support this. Brodie (2020) has highlighted the challenges in obtaining reliable data on murders and the potential of media coverage to fill in some of these gaps.

The way murders are reported in South Africa shows clear gender biases, with female victims of gender-based violence often receiving significant media attention (Brodie, 2019).

However, this coverage is frequently sensationalised, framing women as vulnerable and helpless rather than addressing the larger social issues that contribute to gender-based violence (Gqola, 2015). This pattern reflects what Boonzaier (2020) describes as the spectacularisation of violence, where crimes against women are turned into dramatic headlines, often focusing on the gruesome details rather than the systemic factors behind the violence. This is not just a South African issue. Globally, the media tends to sensationalise cases of sexual violence and domestic abuse, drawing attention to the crime itself while overlooking the broader issues of gender inequality and the failures of the justice system (Jewkes, 2015).

The nature of the crime and the relationship between victim and perpetrator significantly influence media coverage in South Africa. High-profile cases, such as the murder of Reeva Steenkamp by Oscar Pistorius, receive extensive international and local media attention, highlighting issues of domestic violence and gun control (Langa et al., 2018). Such cases highlight the media's role in shaping public discourse around crime and justice, like international coverage patterns where sensational cases often dominate headlines, potentially skewing public perceptions of crime prevalence and severity (Surette, 2015). Geographical location plays a significant role in the coverage of murders in South Africa, with urban areas, particularly those perceived as tourist or economic hubs, receiving more attention than rural or impoverished regions (Peterson, 2020).

The media coverage of murders in South Africa is shaped by a complex interplay of demographic and circumstantial factors, including race, gender, age, socioeconomic status, and geographical location. These factors not only influence the representation of murder cases but also reflect broader societal issues such as racial inequality, gender biases, and socioeconomic disparities. Addressing these biases and ensuring a more balanced and equitable portrayal of

murders in news media is crucial for fostering a more informed and nuanced understanding of crime and safety in South African society.

Additionally, this research acknowledges the relevance of intersectionality, the interconnected nature of race, gender, class, and geographic location, in shaping both the risk of victimisation and the likelihood of media coverage. Applying an intersectional lens enables an understanding of how overlapping systems of inequality influence which murders are reported, and how victims are represented in news narratives (Brodie, 2019; Crenshaw, 1989).

CHAPTER 3: METHODOLOGY

This chapter outlines the methodological approach used to investigate how murder is represented in South African news media. The research was guided by a dual method design that combines qualitative content analysis with quantitative statistical tools. This approach enables the research to capture both narrative detail and statistical trends: qualitative content analysis systematically identifies themes, framing, and relationships within articles (Neuendorf, 2025), while quantitative methods such as chi-squared tests and regression allow for testing of associations and general patterns across the dataset (Ahmed et al., 2024; Knappertsbusch et al., 2023). By integrating these methods, the research benefits from the depth of narrative insight and the rigour of measurable evidence, ensuring a more robust and reliable analysis. The aim was to go beyond surface-level description to uncover patterns and gaps in how these murder incidents are reported. This research draws from both manual and computational techniques to make sense of a large body of news articles. Manual coding was used to extract rich, contextual information from the data, such as the relationship between victim and perpetrator, or how race and age were framed. These annotations were then analysed using descriptive statistics and inferential models to explore broader trends. The decisions made at each stage, from selecting data sources to defining what counts as a “unique case” were shaped by both practical constraints and conceptual considerations. This chapter begins with an overview of the research process before describing the data sources, sampling strategy, annotation procedure, and analysis methods in detail. Where appropriate, it also reflects key methodological trade-offs and how these may have influenced the findings.

Research Approach, Overview and Rationale

The research followed a multi-phase process, moving from literature review through to statistical analysis. Insights from the literature review (Chapter 2) guided the selection of analytical variables such as race, gender, geographic region, and mode of violence. Prior studies also informed the choice to use a combined approach of manual coding and computational analysis, responding to gaps in prior South African media studies (e.g. Brodie, 2019; Cornelissen, 2020). Following the review of the literature, the main data sources were selected: news articles were drawn from NewsBank's archive of South African publications, while official crime data were sourced from the South African Police Service's (SAPS) quarterly statistics for January to March 2023. A clear set of sampling criteria was applied to collect and narrow down relevant articles from NewsBank, ensuring that only cases of murder occurring within the defined period and within South Africa were included.

Once the dataset was assembled, an annotation framework was developed through an iterative process of reading and coding by the researcher. This framework included both descriptive variables (such as age, race, and gender) and more relational or contextual ones (such as the relationship between victim and perpetrator, or the presence of names or photos). Each article was then reviewed and manually coded using this framework. With the full dataset coded, the analysis proceeded in two stages. First, exploratory data analysis (EDA) was used to describe general trends and distributions in the reporting. This was followed by formal statistical testing, including chi-square tests and logistic regression models, to evaluate specific hypotheses about representational patterns and media bias. This structured approach reflects a multi-methods design, balancing the interpretive depth of qualitative content analysis with the generalisability and statistical clarity of computational techniques.

Both qualitative (including manual coding) and computational analyses were employed to provide an understanding of the data. In qualitative terms, ‘coding’ refers to the process of categorising and labelling data based on predefined variables such as murder modes, victim demographics and legal outcomes (Chowdhury, 2014). In quantitative terms, coding refers to the use of computational tools to annotate the analysis of datasets using statistical models and machine learning techniques. It used the strength of human interpretation and algorithmic interpretation to provide insight into how murder incidents are portrayed and quantified in South African news media (Kamiri & Mariga, 2021).

This research draws on ideas and methods from both digital humanities and computational social science. The digital humanities offer useful tools for working with text-based data, including manual annotation and computational techniques for identifying patterns and structures. As Berry (2012) points out, these methods range from simple data processing to more advanced algorithmic approaches that help researchers scale their analysis while maintaining rigour. Ramsay (2011) refers to this as “algorithmic criticism,” where computation is used not to replace interpretation, but to support it, helping researchers move through large datasets more effectively and uncover deeper meaning. Schreibman et al. (2008) also describe computational methods as useful for revealing trends, relationships, and hidden structures in cultural materials.

These ideas informed the researcher’s approach to building and analysing the dataset, particularly in how the researcher designed the annotation framework and applied statistical tools. However, the overall orientation of the project is more closely aligned with computational social science. The research aims to test patterns and hypotheses across a broad set of cases, using a structured dataset derived from news media texts.

It is less about interpreting individual narratives and more about identifying broader representational trends, especially those relating to race, gender, location, murder mode and victim-perpetrator relationships in the reporting of murder.

In that sense, computational analysis in this study supports (not replaces) the manual work. It adds another layer to the analysis, enabling both qualitative attention to language and tone, and quantitative exploration of how those patterns show up across hundreds of cases. This combination strengthens the findings and allows for a more balanced understanding of how murders are portrayed in the South African media landscape.

Data Sources

NewsBank is a digital news archive that provides access to an extensive collection of newspapers and other news sources. It offers a large archive of news articles, including content from a diverse range of local, national, and international publications. This database is not open source. Users must either pay for or gain access through an institution. Access was gained through the institutional University of Witwatersrand online library for this research. Users with access can use its advanced search features to explore and retrieve information on various topics, making it a valuable resource for researchers, journalists, and anyone seeking to access historical or current news coverage.

NewsBank includes a search function that allows users to filter by media type and location (e.g., South African news titles), an advanced search option for entering multiple keywords (e.g., "murder"), and a date filter to specify a date range (e.g., January 1, 2023, to March 31, 2023) to refine results based on user preferences. While this study is limited to data available through NewsBank, the database was selected for its extensive collection of news articles from various South African titles across multiple provinces.

The three months' timeframe was chosen to ensure a manageable and focused dataset, given the extensive volume of news articles available on murder cases in South Africa, as well as the fact that it paints a recent and relevant representative picture of murder reporting in South Africa. Additionally, time constraints with data collecting, manual coding and computational analysis meant that the three months' timeframe was the ideal balance between the feasibility of the study and the depth of analysis. These articles were then subjected to a systematic manual coding process, following a predefined schema that includes variables related to victim demographics and incident details. The resulting structured dataset, built through manual coding, was further analysed computationally. This approach ensured a comprehensive examination of the complex narratives surrounding murder incidents in South African media.

A key decision in the research design involved how to handle cases that were reported multiple times across different articles. Instead of removing all duplicates at the point of data collection, the research created two separate datasets: one that retained all articles referring to murders within the study period (including repeat reports on the same victim), and another that filtered these down to unique incidents. The first version of the dataset created by the researcher comprises all news articles recorded, including instances where the same victim is reported in multiple articles. This includes coverage of significant events such as anniversaries of notable murder cases (e.g., the case involving Oscar Pistorius) and updates on legal proceedings involving prominent figures such as vice chancellors, political leaders, celebrities (e.g. the murder investigation of Kiernan Forbes (aka)), or cases of child murders. The database contains a total of 968 entries, sourced from 22 different newspaper titles and compiled from 536 news articles written by a total of 109 journalists. This is important in understanding the media's focus and editorial choices.

The second version, the unique database, filtered to remove duplicates and include only unique murder cases, and accounts for 559 unique victims. This figure also includes incidents that occurred before January 1, 2023, but were reported during the period of January 1 to March 31, 2023, indicating that 559 individual victims were covered in NewsBank during that timeframe. Lastly, the number of incidents outside the specified date range (January 1 to March 31, 2023) is 292. While the unique database was used throughout the overall descriptive analysis, only the murder incidents that took place within the reporting period (resulting in 267 victims) were used when comparing with data to actual crime statistics provided by the South African Police Service (SAPS) central database, ensuring a one-to-one comparison. This filtered dataset was used for statistical comparison and pattern analysis to avoid over-representing high-profile cases and skewing the results. These two versions were necessary to ensure analytical consistency and comparability across datasets.

Additionally, the researcher developed a dataset of SAPS statistics, sourced from the "First Quarter 2022/2023 Crime Statistics Presentation" (SAPS, 2024), where data between January 1st and March 31st, 2023, was captured. This dataset encompasses all reported murders recorded in each province from January 1st and March 31st, 2023. These records were compiled for a comparative analysis with corresponding data on murders reported in news articles over the same period, which were extracted from the NewsBank database. It is important to note that detailed demographic information about the victims, such as race and gender, is not available in the SAPS database. Instead, the available data is aggregated by provinces and on a countrywide basis, and it is categorised into various types of crimes. These categories include contact crimes (i.e. murder), property-related crimes, and other serious crimes. This categorisation allows for a focused analysis of crime trends by type but limits the ability to conduct a more detailed demographic analysis.

Data Collection

The process of data collection involved selecting news media that are relevant to the outcomes of the study. This included classifying all news media coverage and incidents of murders that were reported in the NewsBank database in South Africa during the three-month period, from January 1st 2023 to March 31st 2023.

The news media articles were filtered using advanced search functions on the platform. The retrieved data were then organised into a structured dataset using Microsoft Excel as a tool for storage and preliminary coding. While Excel is not a relational database, it was used in this context to create a flat-file database with clearly defined variables, as outlined in Table 1. Each column represented a distinct attribute or factor relevant to the analysis, following the structure of the femicide database developed by Brodie (2019).

Below is a step-by-step description of how the data collection process was conducted using NewsBank:

1. The database was accessed through the institution's library here: http://0infoweb.newsbank.com.innopac.wits.ac.za/?db=AWNB , with last access on the 8th August 2023.

2. On the homepage of the NewsBank database, “South African Titles” on the right-hand side was selected, as this study contextually focuses on murder incidents in South Africa.
3. The keyword “murder” was used to identify and search for potential murders reported in the database during the three-month period.

4. After this search, the database brought 1,253 articles that were eligible. From there, the researcher downloaded these articles (it allowed 20 at a time in a PDF) and prepared to manually read, code and annotate for information to add to the Excel spreadsheet database.

Coding/Category Scheme Development

Manual coding using content analysis[4] consists of segmenting data into codes or labels that describe the data (Krippendorff, 2013). This method allows for the systematic categorization of qualitative information, ensuring consistency and reliability in the interpretation of news articles. In this research, content analysis is applied to develop a coding scheme that captures key variables, such as victim demographics, murder modes and legal outcomes, as reported in South African news media.

This method sorts through the information from the news articles into pre-existing categories to add to the Excel spreadsheet, where the researcher built the database. The study aims to investigate the different categories of murders that were portrayed in the media during the three-month period captured in NewsBank. This includes thematic coding for important variables and developing categories from the data. Below are the pre-determined codes for incident data, using Brodie's (2019) data structure, and adding code categories (i.e. “victim id”, “article id”, “newspaper title”, and “journalist name”). This data structure includes incident data, which has information about the demographics of the victim, information about the murder and the victim-perpetrator relationship (Brodie, 2019). The manual coding scheme will serve as a guide for reading and categorising the articles, forming the foundation of the database that will later be used for computational analysis.

[4] Content analysis is a systematic research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use (Krippendorff, 2013). It involves the systematic categorization and interpretation of textual, visual, or multimedia data to identify patterns, themes, and trends. In this study, content analysis is used to manually code news articles based on predefined variables such as victim demographics, murder modes, and legal outcomes. It is widely used in social sciences to analyse communication and media representations.

Table 1: *Pre-Determined Categories and Descriptions for Incident Data Analysis (Brodie, 2019) Used in Manual Annotation*

Table 2: *Categories of Murder Modes and their Detailed Descriptions.*

Category	Description
incident_id	The incident identification
article_id	The number of articles within the same newspaper on the same date
journalist_name	The name of the journalist
victim_id	The victim identification
newspaper	The name of the South African English newspaper title
age	The age of the victim
gender	The gender of the victim (i.e., male or female)
race	The race of the victims (i.e., Black, White, Coloured) as described in the news article
murder_mode	How the victim was killed (i.e., the mode of death)
location	The location of the murder (i.e. the province)
area	The area within the province
location_type	Whether the location is rural, urban or peri-urban
relationship	The victim-perpetrator relationship (the relationship the victim has to the perpetrator)
legal_outcome	Whether the case is an ongoing investigation, a pretrial action (i.e. arrest), an active trial, a special case (i.e. suspect dying in police custody), or a post-trial proceeding (i.e. bail hearing, conviction)
date_of_incident	The date the murder took place
date_of_report	The date of the news report
news_headline	The headline of the news story

Table 3*: Categories of Legal Outcomes and their Detailed Descriptions.*

Category	Description
Pre-trial actions	Suspect granted bail, Suspect's bail denied, Suspect denied bail, Suspect released on bail, Suspect applies for parole, Suspect abandons bail, Suspect released due to insufficient evidence, Suspect released, Suspect arrested
Active Trials	Suspect/s on trial
Ongoing investigations	Ongoing investigations of the case
Post-trial proceedings	Convict, Convict to be released, Convict's parole hearing, Convict's parole denied, Suspect sentenced, Convict Acquitted
Special cases	Attempted murder suicide, justice attack, suspect killed in justice attack, Murder being investigated, Suspect fatally wounded, Psychiatric interventions, Suspect killed in police custody,
Murder suicide	Murder suicide
Unspecified	Legal outcomes unspecified or unknown

Table 4: *Categories of Murder Modes and their Detailed Descriptions*

Category	Description
Stabbing	Murder by stabbing of any object
Gunshot	Murder by gunshot
Physical Assault	Murder by being beaten, or bludgeoned
Strangulation	Murder by asphyxiation or strangulation
Fire related	Murder by fire or being burned
Mass Shooting	Mass shooting
Drowning	Murder by drowning
Medical Malpractice	Murder by medical malpractice
Unspecified	Unspecified or unknown

Table 5: *Categories of Victim-Perpetrator Relationships (i.e. Who the Perpetrator was to the Victim) and their Detailed Descriptions Developed by the Researcher.*

Category	Description
Intimate partner	Wife, Husband, Boyfriend, Girlfriend, Ex-Boyfriend, Ex-girlfriend
Family member	Father, Mother, Son, Parents, Stepfather, Mother-in-law
Assassination	Assassination, Hitman
Gang violence	Gang violence
Taxi violence	Taxi violence, Uber driver, Bolt driver
Acquaintance	Friend, Neighbour, Teacher
Stranger	Road rage, Kidnapping, Drive-by shooting, Robbery
Unspecified	Not specified or unknown

By implementing these steps, the study attempts to provide an accurate and dependable analysis of murder reporting, contributing valuable insights into public discourse and policy development concerning crime and safety in South Africa.

In some cases, surname-based inference was cautiously used to approximate racial classification, following established conventions in South African research (Brodie, 2020). In instances where race was not explicitly mentioned in the news reports, the race variable was marked as “unspecified” to avoid misclassification and acknowledge ethical concerns. Where feasible, surnames were used as indicators, for example, “Naidoo” was classified as likely Indian, while “Ndlovu” was classified as likely Black, drawing on linguistic and cultural naming patterns commonly used in local South African demographic research (Brodie, 2020).

This method was used sparingly and only when necessary to support comparative analysis with existing crime statistics. However, this approach has clear limitations. It risks oversimplifying complex racial identities, reinforcing assumptions, and misclassifying individuals, particularly in cases of mixed heritage or surname adoption. Additionally, race is a socially constructed and context-dependent category that may not align neatly with surname-based assumptions. While this method offered a practical compromise in the absence of explicit information, its use was always guided by caution, transparency, and the recognition that it cannot capture the full complexity of racial identity.

Data Analysis

To extract new and meaningful insights from the dataset, RStudio (an interface for R) was employed to facilitate reproducible statistical analysis and visualization of the characteristics of murder stories (Posit, 2023). This environment serves as the primary platform for statistical analysis and data visualization. RStudio enhances the functionality of R by providing an integrated development environment that facilitates script writing, data manipulation, and package management (Posit, 2023).

To analyse the dataset effectively, this research relied on a combination of specialised R packages, each playing a role in data cleaning, statistical analysis, visualisation, and modelling. Given the complexity of the dataset, different tools were needed to process, structure, and interpret the information in a meaningful way. The core packages used in this research included “Tidyverse”, “ggplot2”, “Base R”, “readxl”, “knitr”, “kableExtra”, “lubridate”, “broom”, “Tidytext”, “caret”, “cluster”, “gridExtra”, and “Stargazer”. To clean and prepare the data, “Tidyverse” (Wickham et al., 2019) was particularly useful, with “dplyr” playing a role in filtering, grouping, and summarising crime reports.

This made it possible to examine trends based on variables such as victim demographics, crime locations, and reporting timelines. Structuring the dataset in this way ensured that the information was well-organised and ready for analysis.

For visualisation, “ggplot2” (Wickham & Grolemund, 2016) was used to generate bar charts, line graphs, and facet plots, making it easier to see changes in how murder cases were reported over time, and visualising other variables such as the time to report against victim demographics (age, race, gender). It also allowed for comparisons between different provinces, highlighting variations in media coverage of crime. These visual tools provided a clearer way to interpret trends rather than relying on raw data alone. Basic statistical analysis was conducted using Base R, which was used to calculate means, medians, standard deviations, and other descriptive statistics. It also allowed statistical testing, such as chi-square tests and t-tests, to determine whether differences in crime reporting patterns were statistically significant (Posit, 2023). This explanation is included to add clarity and understanding for readers who may be unfamiliar with statistical testing and modelling.

Since the dataset was initially recorded in Excel spreadsheets, “readxl” (Walker, 2019; Wickham & Bryan, 2019) was used to import the data into RStudio. These packages made the transition from raw media reports to a structured dataset more efficient, ensuring that the data could be processed and analysed without formatting issues. To present findings in a clear and structured format, “knitr” and “kableExtra” (Xie, 2015; Zhu, 2019) were used to create well-organised tables and reports in RMarkdown. These tools helped generate summary tables and format outputs, making it easier to present statistical results and key trends in a professional manner.

Since time-based analysis was an important aspect of this study, particularly in assessing delays between crime occurrences and media reporting, as well as trends (daily, weekly and monthly), “lubridate” (Grolemund & Wickham, 2011) was used to handle date-related variables. This made it possible to analyse seasonal trends in murder reporting and track changes in media attention over time. This research also explored predictive modelling techniques using “caret”, and “cluster” (Liaw & Wiener, 2002; Kuhn, 2008; Maechler et al., 2019). These machine learning models were used to classify cases based on factors such as victim profile, crime location, and media source, helping to identify patterns in media attention. This analysis provided insight into whether certain types of cases were more likely to be widely reported.

To organise multiple visual outputs and statistical summaries, “GridExtra” and “Stargazer” (Auguie, 2017; Hlavac, 2018) were used. These tools helped structure the final report by arranging graphs, tables, and statistical results in a clear and readable format. By integrating these computational tools, this research was able to analyse crime reporting patterns efficiently and systematically. The use of RStudio, combined with these packages, allowed for a comprehensive examination of media representation of crime, revealing trends in reporting frequency, geographic disparities, and potential biases in coverage. The ability to reproduce and replicate statistical analysis and visualization provided deeper insights than traditional tools like Excel, making the research more robust and data driven.

Validity and Reliability

Ensuring validity and reliability was a priority for this research, as this research relied heavily on manual coding, extensive reading of news articles, and human interpretation. The goal was to make sure that the data accurately reflected crime reporting patterns while maintaining consistency throughout the coding process.

Given the amount of manual work involved, there was always the risk of errors, subjectivity, and inconsistencies. Therefore, the following steps were taken to reduce these risks and make the dataset as robust and reliable as possible.

One of the first challenges was the way the data was structured. Initially, a simple manual coding format was used, where articles, victims, and incidents were labelled with broad, alphanumeric codes such as A1_V2_I1 (Article 1, Victim 2, Incident 1). This structure, however, made it difficult to analyse specific patterns that would be easily readable in R, as R is not intuitive for tracking individual victims or incidents. To fix this, the categories were separated into more complex, distinct, clearly labelled variables, making it easier to work with the data. A Victim ID (i.e. V1 (victim 1), V2 (victim 2)) was also introduced, which ensured that each victim had a unique identifier. This not only improved organisation but also made it possible to analyse trends related to specific victim demographics, which in the previous system would have been more prone to errors.

Another major change was the removal of full news article text. Initially, sentiment analysis was considered, but it became clear that this approach would not be feasible without creating a custom sentiment dictionary. Since sentiment analysis relies on predefined word associations, longer texts require more than just identifying positive or negative words, they involve interpreting context, sarcasm, and nuanced language. Given the complexity of crime reporting, where tone is shaped by phrasing rather than isolated words, a predefined dictionary would have been insufficient. Developing a tailored dictionary would have required extensive manual annotation and validation, which was beyond the scope of this study. Instead, the dataset was refined to focus on structured, quantifiable variables, such as crime frequency, location, and victim demographics. This shift ensured a more objective and reliable analysis, avoiding potential inconsistencies in sentiment interpretation and strengthening the validity of the findings.

To ensure consistency and reliability, a cross-verification process was put in place. To reduce the risks of misclassifying a case, making a typographical error, or interpreting something differently on different days, a random selection of entries was reviewed and cross-checked in Excel using the filtering system. This also meant revisiting the original articles and ensuring that the coding was correct and consistent. Additionally, to fix duplicate entries, a de-duplication process was applied in Excel. Cases with the same victim, location, and crime details were flagged, reviewed, and merged into a single record. This ensured that every crime was counted once, preventing overrepresentation of high-profile cases. To make it easier for the researcher, victim IDs were also colour coded in Excel. If the same victim was reported again, it was easy for the researcher to identify and assign the victim their ID.

The dataset evolved through an ontology-building process which involved clearly defining categories and the relationships between them. Ontologies are commonly used in computational social science to organise complex qualitative ideas, especially when shifting from manual coding towards computational methods (Borra & Rieder, 2014; Kaplan et al., 2023). In this research, categories such as victim–perpetrator relationships, legal outcomes, and murder modes were gradually refined into a structured classification scheme. This ensured consistency during manual coding and allowed computational techniques to effectively highlight patterns in murder reporting across demographic groups, geographic locations, and case outcomes. Some variables, such as the victim-perpetrator relationship and legal outcome, started as open-ended descriptions, which made it difficult to analyse quantitatively. Over time, these were reclassified into standardised categories, making it easier to compare cases and identify trends. For example, rather than having a mix of vague terms such as “extortion”, “wife” or “unknown assailants”, relationships were coded into structured groups like “intimate partner”, “family member”, or “stranger”.

Similarly, legal outcomes were refined into distinct categories like “ongoing investigation”, “pretrial actions”, or “post-trial proceedings”, which made it easier to track how cases progressed through the justice system.

A systematic approach was employed to manually collect data from the NewsBank database, focusing on articles related to murders published between January 1st and March 31st, 2023. Articles were downloaded starting from the earliest to the latest within each month. Throughout the collection period, an increasing number of murder-related articles were observed, indicating a rising trend in media reporting. In total, 536 articles were collected (where some had multiple victims reported) and were deemed eligible for further analysis.

Initially, a generic set of categories (age, race, and gender) was used to code the victim data. However, as the research progressed, additional categories were introduced to enhance the dataset’s granularity and statistical clarity. These included victim ID, article ID, and incident ID as separate columns for easier analysis and reading in the R environment, as opposed to the original linking of codes (i.e. DD_2023-01-15_A1_I1 (Daily Dispatch, 2023-01-15, Article 1, Incident 1). Each victim was assigned a unique identifier (e.g., V1 for Victim 1) to maintain consistency across articles, particularly in cases where multiple victims were reported in a single piece. Additionally, broader descriptions were created for victim-perpetrator relationships, murder mode and legal outcome codes (refer to tables 2, 3, and 4).

The manual coding process was also important in handling special cases such as celebrity murders or well-known murders. When key information, such as the age or date of the incident, was missing in the article, additional research was conducted using reliable sources to fill in these gaps. This ensured the accuracy and completeness of the dataset. In instances where multiple victims, such as in cases of gender-based violence (GBV), were reported within a single article, each victim was recorded and categorised appropriately. Articles detailing follow-ups on arrests and court proceedings, especially in high-profile cases, were also captured. International murder

cases were excluded from the dataset to focus solely on murders occurring within South Africa. Cases involving foreign victims were clearly noted; these would yield interesting results as these details were often highlighted in the headlines or body of the article.

The manual collection of data plays an important role in ensuring the accuracy and integrity of datasets, particularly when dealing with complex and context-sensitive information such as murder stories reported in the media. Unlike automated data collection, manual processes allow for the nuanced interpretation of articles, accounting for the contextual subtleties that automated systems may overlook.

This is especially important when analysing crime-related data, where victim details, geographic locations, and legal outcomes often require careful interpretation and categorisation by the researcher, as evidenced in the context of this study. The manual collection process ensured that critical information such as victim demographics, modes of murder, and legal outcomes were accurately captured. Through this process, the researcher was able to systematically code variables that might not have been clearly defined in the articles or that required additional investigation, such as celebrity or high-profile cases where certain details (e.g., ages or dates) were omitted. The manual approach allowed for the inclusion of these details, ensuring that the dataset was as complete and consistent as possible before proceeding to computational analysis.

The importance of this manual process cannot be overstated when preparing data for computational analysis. Clean, well-structured data is the foundation upon which accurate and reliable analysis is built. Manual data collection ensures that the dataset is free of duplicate entries, inconsistencies, and errors that have the potential to skew results or lead to misleading conclusions. For example, in this study, the manual process involved the careful removal of duplicates manually (using Excel filter tools). Excel was used instead of R at this step because it allowed for a more intuitive visual comparison of entries. Manually reviewing these cases in Excel was more reliable

and practical than automating the process with R, which would have been less effective due to the nuanced nature of the data. These steps were important in producing a dataset that is well-suited for further computational processing through R, such as statistical analysis and detecting patterns.

However, manual data collection is also a time-consuming and labour-intensive process. It requires a significant amount of attention to detail, with the researcher often having to consistently cross-reference multiple sources to ensure accuracy. This can lead to potential issues such as human error in coding or interpretation, particularly when dealing with large datasets over extended periods. Moreover, the manual process is slower than automated methods, which can limit the scope of the data that can be collected within a given timeframe.

Despite these drawbacks, the benefits of manual data collection, particularly in a context like this, outweigh the cons. By engaging in a meticulous review of each article, the researcher can ensure that the data is reflective of real-world complexities, something that automated systems may struggle to achieve without sacrificing accuracy. In cases where automated tools may misinterpret or overlook important details (such as specific victim characteristics or legal nuances), the manual approach allows for the inclusion of these subtle but significant factors.

Additionally, manual data collection offers flexibility in terms of adjusting and refining the coding scheme as new patterns or issues emerge. For instance, in this study, additional categories such as victim ID and incident ID were introduced to add further clarity to the dataset, as well as the broader description under the victim-perpetrator relationship, legal outcome, and murder mode. This adaptability is a key advantage over automated processes, which often rely on predefined rules that are harder to adjust during the data collection phase. These steps improved both the validity and reliability of the study. Instead of handling unclean, inconsistent, or overly complex data, the refined dataset was clean, structured, and well-suited for analysis in R. More importantly, the research process was transparent, replicable, and grounded in logical decision-making,

ensuring that the results accurately reflected crime reporting patterns rather than inconsistencies in data entry. By taking the time to refine, verify, and document the process, this study not only strengthened its findings but also set a standard for data integrity in media analysis research.

While the manual data collection process is tedious and time-consuming, it provides a level of detail, accuracy, and flexibility that is essential for producing a high-quality dataset. This clean and carefully curated dataset is crucial for effective computational analysis, as it allows for the identification of meaningful patterns and relationships in the data without being clouded by inconsistencies or inaccuracies. Therefore, despite the effort involved, manual data collection is a valuable method for studies that require detailed and reliable data, such as this analysis of murder reporting in South African media.

Ethical Considerations and Methodological Reflections

The use of secondary data, particularly media reports covering real cases of murder requires careful ethical reflection. These reports often included the names of victims, details of violent acts, and descriptions of traumatic events. While this information is already in the public domain, it is important to handle it sensitively and avoid sensationalism. To ensure ethical integrity, the study followed the guidelines set out by the University of the Witwatersrand and received ethical clearance from the departmental ethics committee. All data was stored securely, and only publicly available content was used. When coding sensitive attributes, such as race, age, or relationship between victim and perpetrator, care was taken to avoid reinforcing stereotypes or making assumptions beyond what was directly supported by the data. Where uncertainty existed (for example, in cases of surname-based racial inference), variables were marked as “unspecified” to reduce the risk of misrepresentation.

Several important methodological decisions were made in this research, each with implications for the scope, reliability, and interpretability of the study's findings. The choice to manually annotate the dataset allowed for rich, context-sensitive insights into how murder is reported in the news. However, this decision came at the cost of scalability and inter-coder validation, as the manual nature of the process limited the volume of data that could be reviewed and increased the risk of subjectivity (a challenge widely acknowledged in qualitative research (Jiang et al., 2021)). Similarly, the decision to maintain two separate datasets, one comprehensive and one filtered for unique incidents, allowed the research to preserve media dynamics while still aligning with official crime statistics for analytical clarity. This dual-dataset strategy follows recommendations in multi-method research (Guetterman & Fetters, 2018) but also introduced complexity in reporting and required careful explanation to avoid confusion. Excluding incidents that occurred outside the defined reporting period (1st of January to 31st March 2023) helped ensure comparability with SAPS quarterly data, though it likely resulted in the undercounting of late-reported cases. These trade-offs were made in response to specific analytical goals, data constraints, and the need for methodological transparency. While not without limitations, they were necessary to ensure that the patterns revealed were both meaningful and empirically grounded.

CHAPTER 4: FINDINGS AND RESULTS

This chapter presents the findings from the computational analysis of the murder reporting in South African news media, focusing on articles retrieved from NewsBank, published over a three-month period (January 1, 2023, to March 31, 2023). The database contains a total of 968 entries, sourced from 22 different newspaper titles and compiled from 536 news articles written by a total of 109 journalists. The unique database, filtered to remove duplicates and include only unique murder cases, accounts for 559 unique victims. This figure also includes incidents that occurred before January 1, 2023, but were reported during the period of January 1 to March 31, 2023, indicating that 559 individual victims were covered in NewsBank during that timeframe. Lastly, the number of incidents outside the specified date range (January 1 to March 31, 2023) is 292. While the unique database was used throughout the overall descriptive analysis, only the murder incidents that took place within the reporting period (resulting in 267 victims) were used when comparing with data to actual crime statistics provided by the South African Police Service (SAPS) central database, ensuring a one-to-one comparison. This filtered dataset was used for statistical comparison and pattern analysis to avoid over-representing high-profile cases and skewing the results. These two versions were necessary to ensure analytical consistency and comparability across datasets. The analysis aims to investigate how murders are reported in South African news media, identifying patterns in the types of murders covered, the demographics of victims, and the geographical distribution of reported cases.

Specifically, this research examines the prevalence of different modes of murder and how frequently certain victim characteristics, including race, age, and gender, appear in South African news media. It also explores the relationship between victims and perpetrators, assessing whether certain case profiles receive more media attention than others.

Through computational methods, this research also examines geospatial trends in murder reporting, identifying regional differences in media coverage.

By comparing media reporting patterns to official SAPS crime statistics, this research seeks to understand whether certain types of murders are disproportionately covered and whether media biases influence public perceptions of crime. Given the role of news coverage in shaping societal attitudes, policy discussions, and law enforcement priorities, analysing these patterns provides insight into the broader impact of media representation on crime narratives in South Africa.

South African News Titles

The distribution of articles published by different newspaper titles provides an understanding of how murder cases are covered across various media outlets. Newspapers such as the “Daily Dispatch” (95 articles), “Daily News” (73 articles), “Herald” (79 articles), and “Cape Argus” and “Mercury” (56 articles) have published a higher number of articles on murder cases during this study period. This indicates a stronger focus or a higher incidence of relevant events in the regions or communities these newspapers serve. These figures suggest that these newspapers might be in areas with higher crime rates, or they may have editorial policies that prioritise extensive crime reporting. Newspapers like “Cape Times”, “Mercury”, and “Sowetan” with 34, 56, and 36 articles respectively, also show considerable coverage, suggesting their active role in reporting crimes within their respective locales. Some newspapers, such as “Business Day” (published on weekdays), “Independent on Saturday” (Saturday), “Saturday Star" (Saturday) and “Sunday Independent” (Sunday) follow different publication schedules.

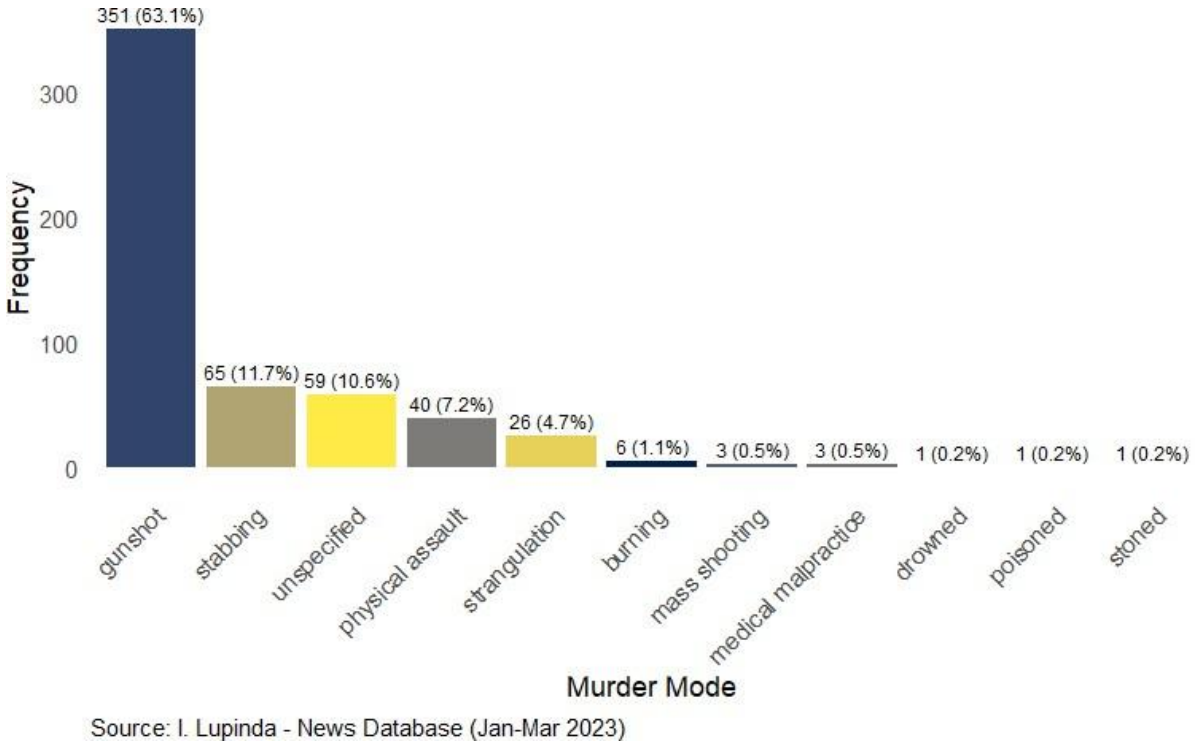
This publication schedule inherently results in fewer articles per week (typically 1-2 each) compared to daily newspapers, which likely accounts for the lower number of murder stories reported by these titles. Smaller newspapers like “GroundUP” and ‘Diamond Fields’’ (2 and 4, respectively) also have fewer articles.

Overview of Murder Types and Victims Demographics

Murder Modes

This analysis used the captured news report data (including reported incidents that occurred outside the date range but were reported in the date range). The only time filtered data was used (i.e. murder incidents that took place within the date range of January 1st to March 31st) was in comparison with the SAPS central database. This section uses the unique database that includes unique victim counts. The analysis of murder modes in Figure 1 shows that gunshot-related deaths are the most common type reported in the news media during the three-month period, with 351 incidents recorded.

Figure 1: *The Distribution of Murder Modes*

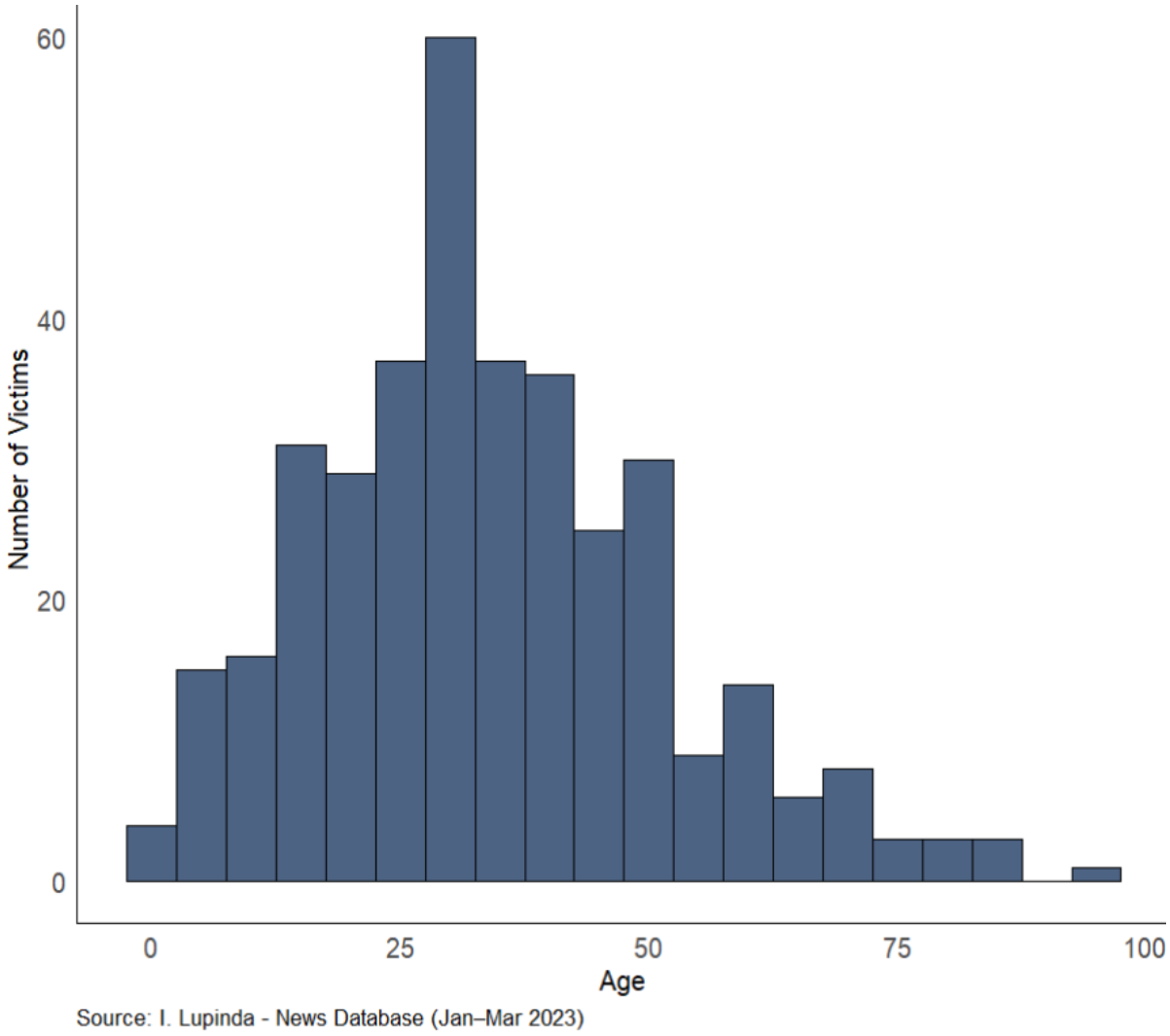


Gun-related murders account for nearly 64% of the total reported incidents. Following gun-related deaths, stabbings are the second most common mode of murder, with 65 cases, followed by physical assault at 40 cases. The next common category was labelled as “unspecified” with 59 incidents. Less frequent modes reported were strangulation (26), burning (6), medical malpractice (3), drowning (1), and poisoning (1).

Age, Gender and Race

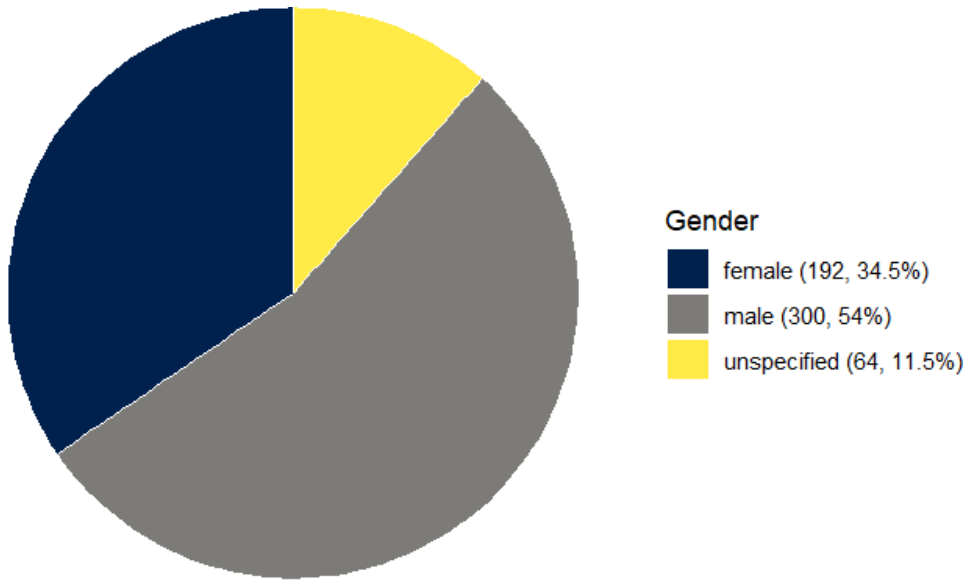
The demographic characteristics of murder victims offer a preliminary lens through which to explore broader social patterns and potential structural inequalities. The news data in Figure 2 shows a wide age range from 1 to 97 years, with a mean age of approximately 34 years, a median age of 32, and a mode of 28, indicating that young adults are disproportionately represented in media reports of murder.

Figure 2: *The Age Distribution of Murder Victims*



While age alone does not capture the full complexity of the social context, these figures may point to factors such as employment challenges, social mobility, or exposure to violence that can disproportionately affect younger populations (Abrahams et al., 2016). Similarly, while race and gender are not definitive explanations for victimisation patterns, observed differences may reflect underlying social and economic influences or systemic issues that warrant further investigation.

Figure 3: *The Proportion of Murder Victims by Gender*



Source: I. Lupinda - News Database (Jan-Mar 2023)

The gender analysis in Figure 3 shows a difference in victimisation rates between males and females. Male victims are the most reported, with 300 cases compared to 192 female victims and 64 unspecified cases during the three months. A total of 34.5% of the victims were female, 54% were male, and 11.5% were victims for whom their gender was unspecified, as seen in Figure 3. SAPS data mirrors this trend, showing that most murder victims are men, about 81% in 2013/14 (SAPS, 2014). Similarly, a national mortuary study found that 87% of homicide victims in 2017 were male (Matzopoulos et al., 2020). Therefore, the news media sample in this research gives female victims relatively more visibility than their share among victims.

Figure 4: *The Distribution of Murder Victims by Race*

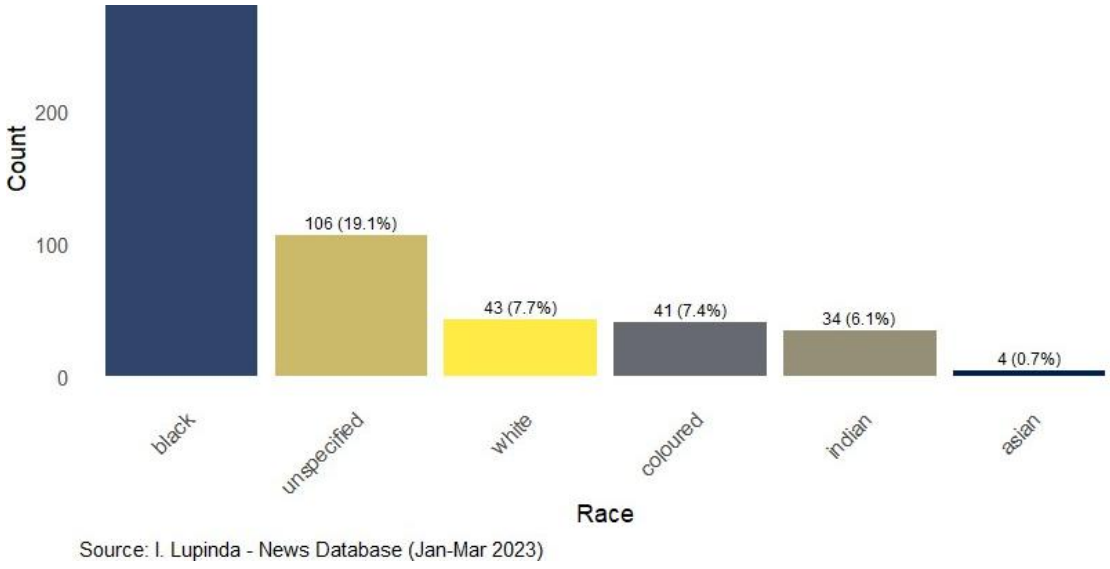


Figure 4 shows the racial breakdown of murder, where Black individuals constitute the most reported murder victims, with 328 cases, followed by White (43), Coloured (41), Indian (34) and 106 unspecified victims. Race data was recorded when explicitly mentioned in the article or inferred from accompanying details such as victim names or community references (see Brodie, 2019). Where race could not be reasonably determined, it was marked as 'unspecified'. Racial details of murder victims are not reported in SAPS crime statistics, limiting direct racial comparisons

The findings show that gun-related murders dominated media reports during this three-month period, making up 64% of cases, followed by stabbings and physical assaults. Young adults, particularly men, were the most reported victims, with a median age of 32 years. Black individuals accounted for most cases, though many reports lacked racial data. These patterns highlight key trends in murder reporting but also suggest gaps in demographic detail that may reflect broader social or reporting biases.

Comparative Analyses: Case Characteristics

Gender and Mode of Murder

This analysis focuses on comparing the modes of murder by gender to understand if certain methods of murder are more prevalent for males or females (using the unique database, with unique victim counts).

Figure 5: *Comparison of Murder Modes for Male, Female and all Victims*

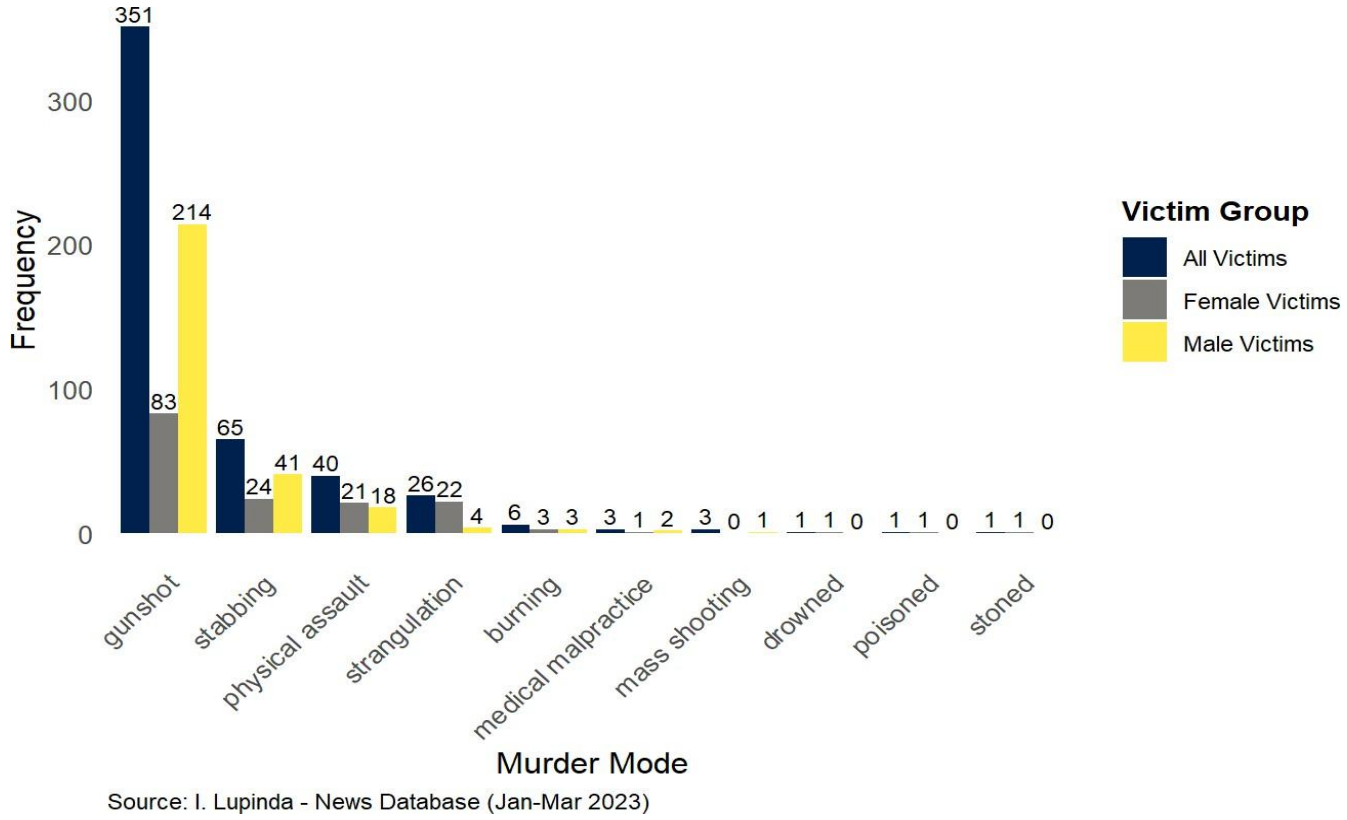


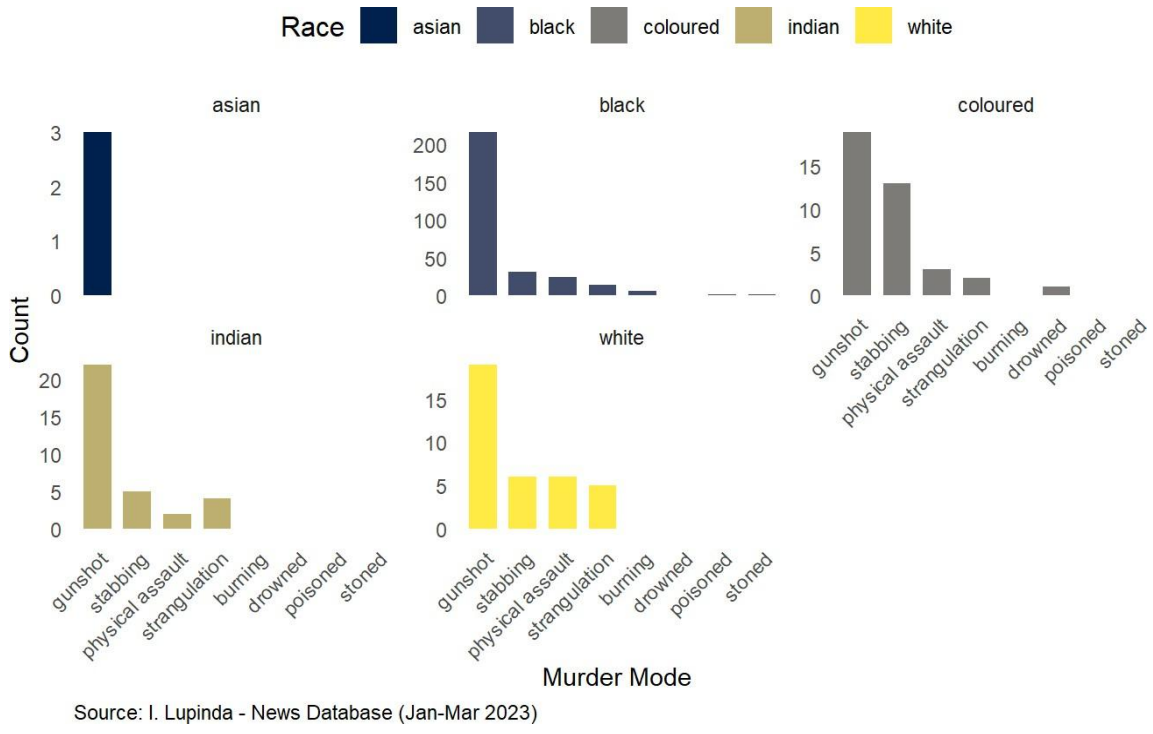
Figure 5 provides a comparison of murder modes across genders. The table shows differences in murder methods between genders. Gunshot is the leading mode of murder for both males and females, but it is significantly more prevalent among male victims (214 cases) compared

to female victims (83 cases). Strangulation, on the other hand, is more common among female victims, with 22 cases compared to 4 male victims. Similarly, physical assault shows a relatively even distribution, with 21 female victims and 18 male victims. Less common methods such as burning, poisoning, and drowning predominantly involve female victims.

Racial Patterns

Figure 6 shows how different murder modes are represented across racial groups in media reporting, illustrating variations in how violence is portrayed along racial lines. Cases with unspecified murder modes were excluded from this analysis, as their inclusion could distort the overall representation of the data. This refinement ensures that the comparisons more accurately reflect the distribution of known murder modes among the racial groups.

Figure 6: *Murder Mode Distribution by Race*



Source: I. Lupinda - News Database (Jan-Mar 2023)

Gunshot-related deaths are the most common across all racial groups, but Black victims are disproportionately affected, with 217 cases compared to 19 for White victims and 22 for Indian victims. Stabbings and physical assaults are also prevalent among Black victims. Strangulation is more frequent among Black and Coloured victims compared to other groups during this period.

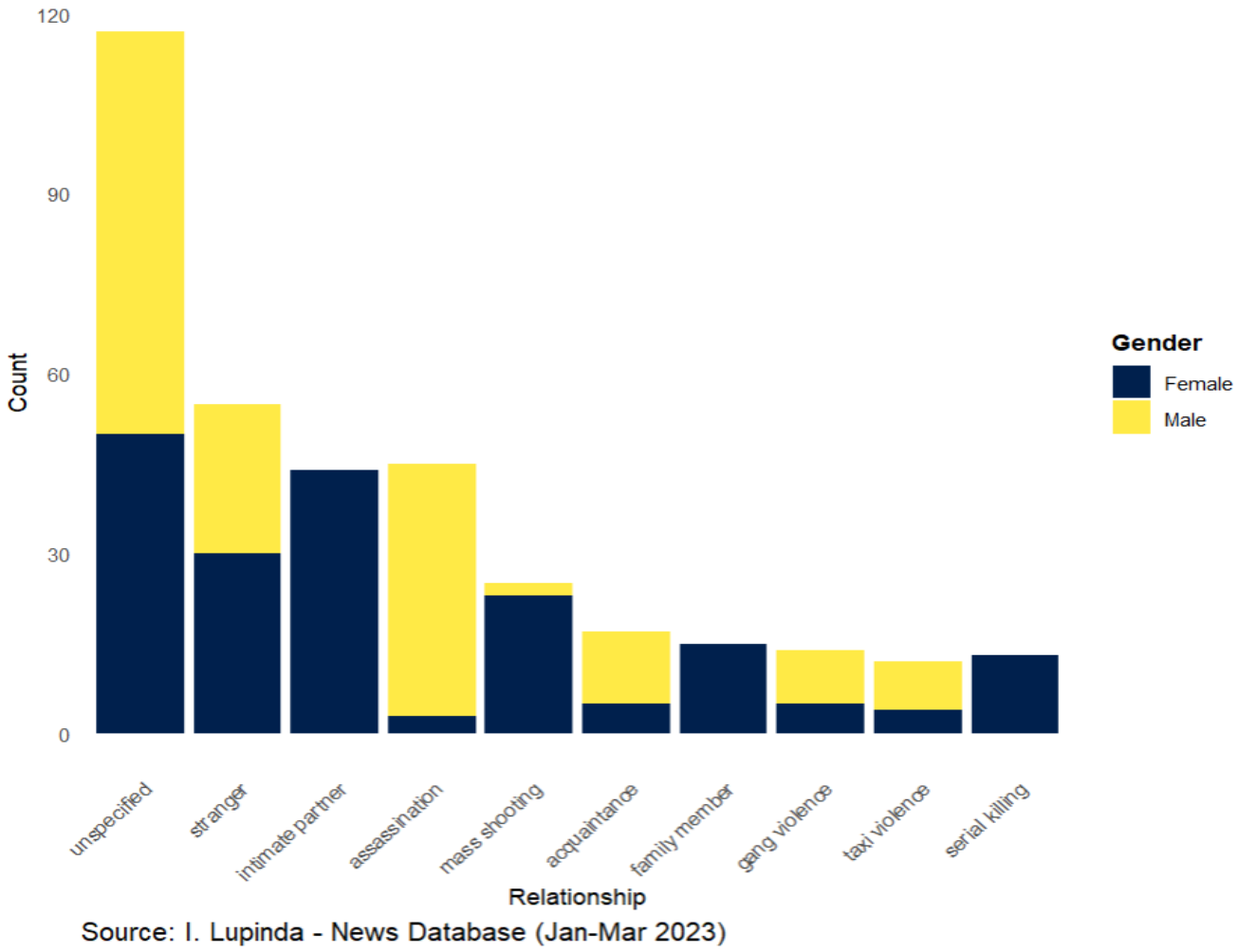
The news data from the unique database for the three-month period indicates that gun violence is the leading cause of murder among the reported cases, accounting for nearly two-thirds of all incidents. Stabbings and physical assaults are also observed, though at much lower frequencies. In terms of victim demographics, young adults are predominantly represented, with an average age of around 34 years. Men, particularly Black men, account for most gunshot cases. However, given that Black men constitute a significant proportion of the overall population (around 81%), these raw numbers should be interpreted with caution regarding disproportionality. These findings, based solely on the three-month reporting period, highlight trends in the data but require further analysis of a larger dataset over a larger period to fully understand underlying social dynamics.

As shown in the news data, gun violence is the most common method of murder, with men more frequently killed by strangers, assassins, or in gang-related violence, while women are more likely victims of intimate partner and family-related murders. Strangulation, burning, and poisoning appear more in female cases, and serial killers exclusively targeted women. Racial patterns show Black victims make up most reported cases, though reporting gaps may affect the accuracy of these comparisons. The high number of cases with unspecified victim-perpetrator relationships suggests that much of this context is missing in media reports, highlighting the need for a deeper look into patterns of violent crime and how they are documented.

Victim-Perpetrator Relationships

Figure 7 displays the count of murder cases from the unique incident dataset, categorised by the relationship between the perpetrator and the victim. The figure further differentiates these counts by gender.

Figure 7: *The Relationship Between Perpetrators and Victims by Gender*



The most frequent category (unspecified relationships) with 117 (39%) incidents involves male victims, and 50 (26%) incidents involving female victims, indicating a significant number of cases where the relationship context is not determined or reported.

Next is the ‘stranger’ category that ranks high, with 55 (18.3%) incidents involving males and 30 (15.6%) involving females, suggesting that a substantial portion of murders occur between individuals who do not know each other.

Assassinations account for 45 (15%) incidents involving males and 3 (1.6%) involving females, indicating targeted killings predominantly affecting males. Mass Shootings show a nearly balanced count with 25 (8.3%) male-related incidents and 23 (12%) female-related incidents, reflecting the indiscriminate nature of these violent events. Acquaintances consist of 17 (5.7%) incidents involving males and 5 (2.6%) involving females, pointing to murders occurring within known but possibly informal relationships. Gang Violence involves 14 (4.7%) male-related incidents and 5 (2.6%) female-related incidents. Taxi violence totals 16 incidents (12 (4%) male-related, and 4 (2.1%) female-related incidents).

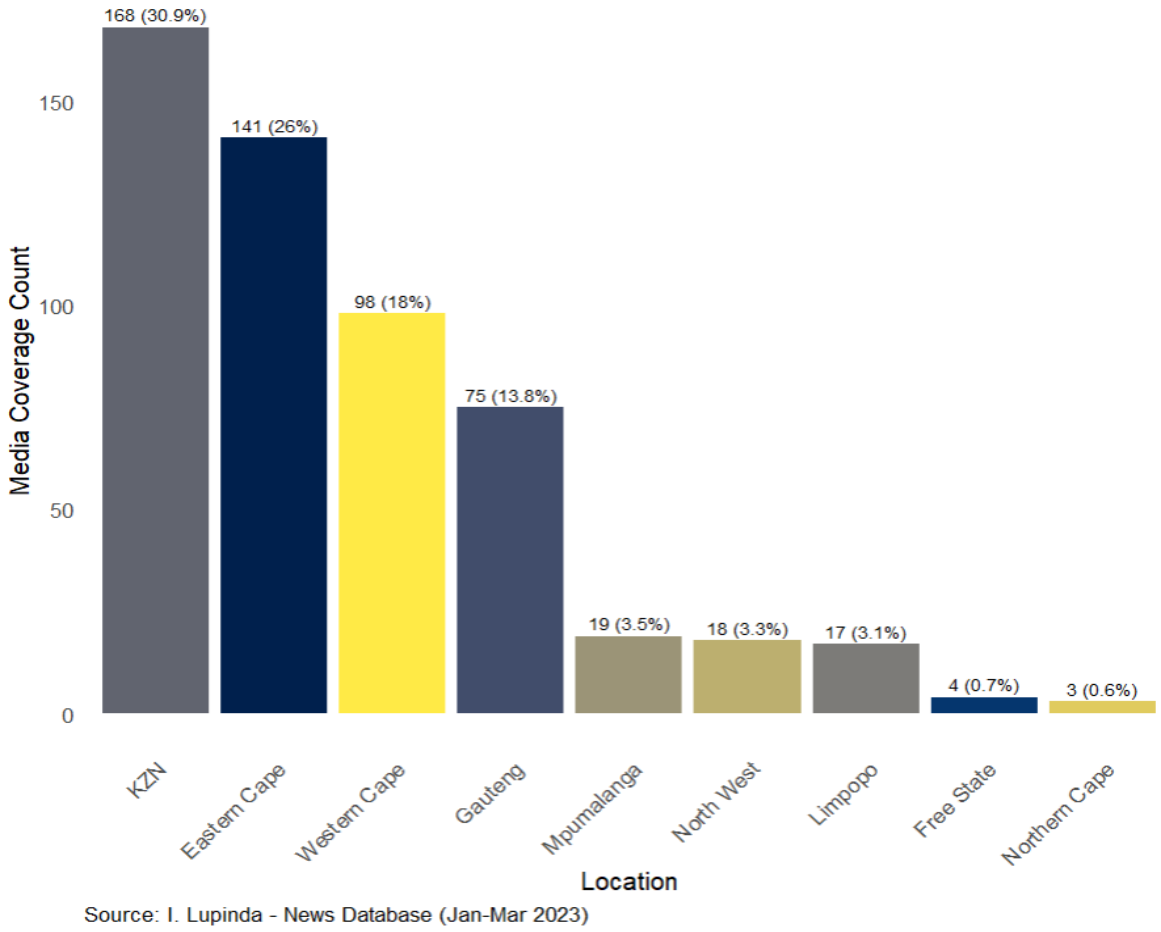
Intimate Partner relationships are higher in incidents involving female victims, with 44 (22.9%), compared to 8 (2.7%) involving male victims, clearly indicating a gendered pattern where females are more frequently victims in intimate partner murders. Family Member relationships involve 7 (2.3%) male victims and 15 (7.8%) female victims, showing more frequent female victimisation within familial contexts. Serial killings are unique in that they solely involve female victims, with 13 (6.8%) reported incidents, suggesting a specific victim profile in serial murder cases.

Geographical Analyses

Location of Incidents

The geographic analysis identifies the areas with the highest number of murders. Figure 8 provides the distribution of murder incidents by geographical location.

Figure 8: *The Distribution of Murders by Geographical Location*

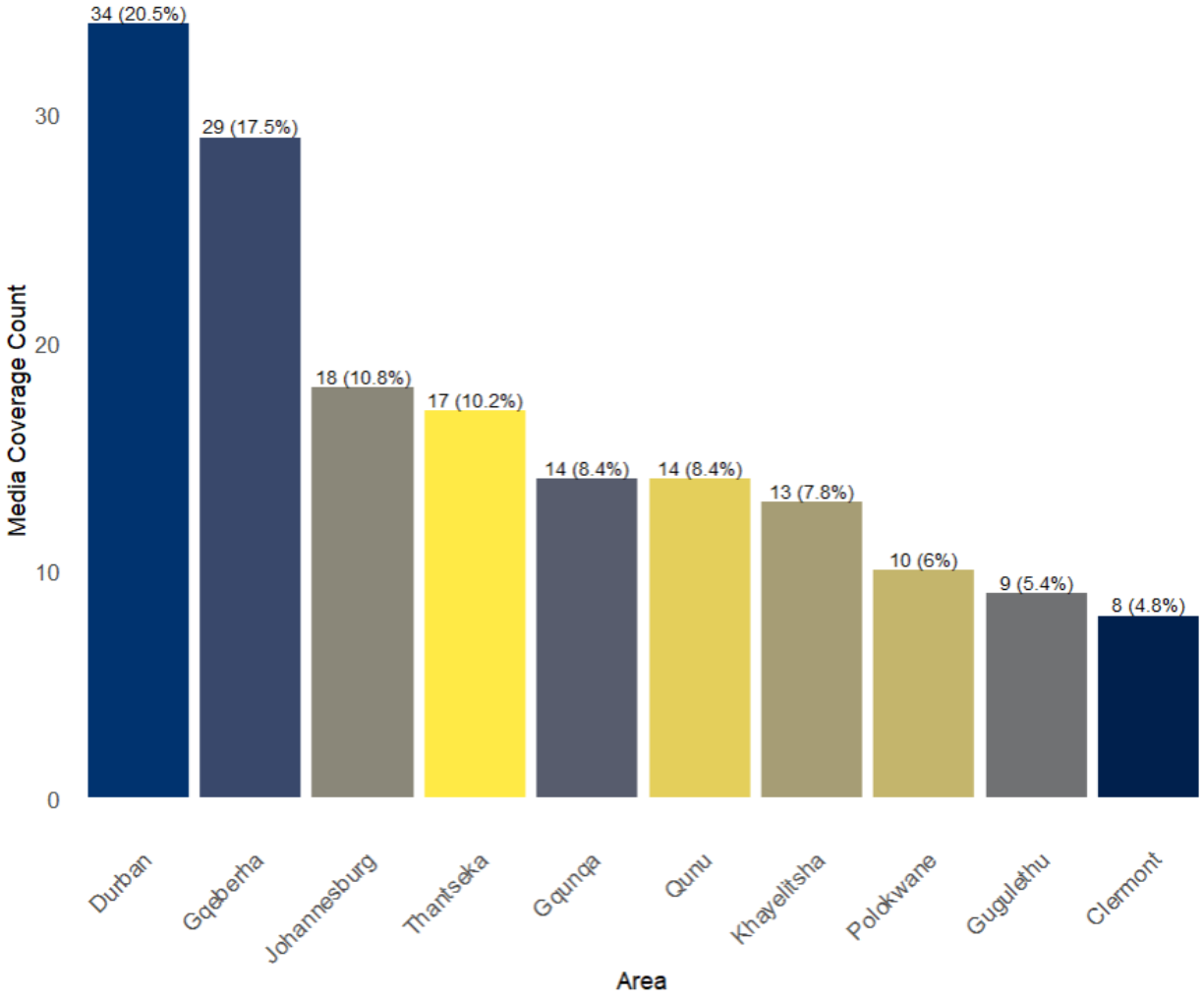


Source: I. Lupinda - News Database (Jan-Mar 2023)

Figure 8 shows the top 10 specific location areas with the highest frequency of murders, based on the data collected. KwaZulu-Natal (KZN) has the highest number of murders reported in the press, with 168 incidents, followed by the Eastern Cape with 141 incidents, and the Western Cape with 98 incidents. Gauteng also records a significant number of incidents (75 cases).

Mpumalanga, North West, Limpopo, Free State, and Northern Cape have considerably lower numbers of murders reported in the media. This may also reflect proportionally lower media coverage and limited media presence in these provinces.

Figure 9: *Top 10 Areas with the Highest Murder Rates*



Source: I. Lupinda - News Database (Jan-Mar 2023)

The area with the most murders reported in the news data during the three-month period is Durban, KZN where 34 (20.5%) incidents were recorded. Following Durban is Gqeberha in the Eastern Cape, with a count of 29 (17.5%) murders. Johannesburg comes next in the list with 18 incidents, while Thantseka in the Eastern Cape has 17 deaths (10.2%), although these are related to mass killings and are outliers in the data. Cape Town, Western Cape, accounted for a combined

22 murders (13.2%), drawn from its two most affected areas, Khayelitsha (13 incidents) and Gugulethu (9 incidents). Both Gqunqa and Qunu, Eastern Cape, recorded 14 murders each (8.4%), while Polokwane, Limpopo, reported 10 murders (6%), and Clermont, KZN, recorded 8 murders (4.8%).

Urban, Peri-Urban and Rural Distribution

The distribution of murder incidents across urban, peri-urban, and rural areas provides insights into the spatial dynamics of violence. Table 5 presents the distribution by area type.

Table 6: *Media Coverage of Murders by Area Type with Proportions*

Location type	Media Count	Proportion
Peri urban	121	21.762590
Rural	150	26.978417
Unspecified	24	4.316547
Urban	261	46.942446

Urban areas account for most murders, with 261 reported cases. Peri-urban areas report 121 reported cases. Rural areas have the lowest number of murders, with 150 incidents. Urban areas receive most of the media coverage, with nearly 47% of reported murders occurring in urban settings.

Murder Mode by Location

The analysis of murder modes by location provides insight into the types of violence predominant in different regions. The heatmap in Figure 10 details the count of incidents by location and type.

The analysis shows that gunshot-related murders are the most common in all high-incident areas, with KZN, Eastern Cape, and Gauteng reporting the highest frequencies.

Figure 10: *Incidents by Location and Murder Mode*

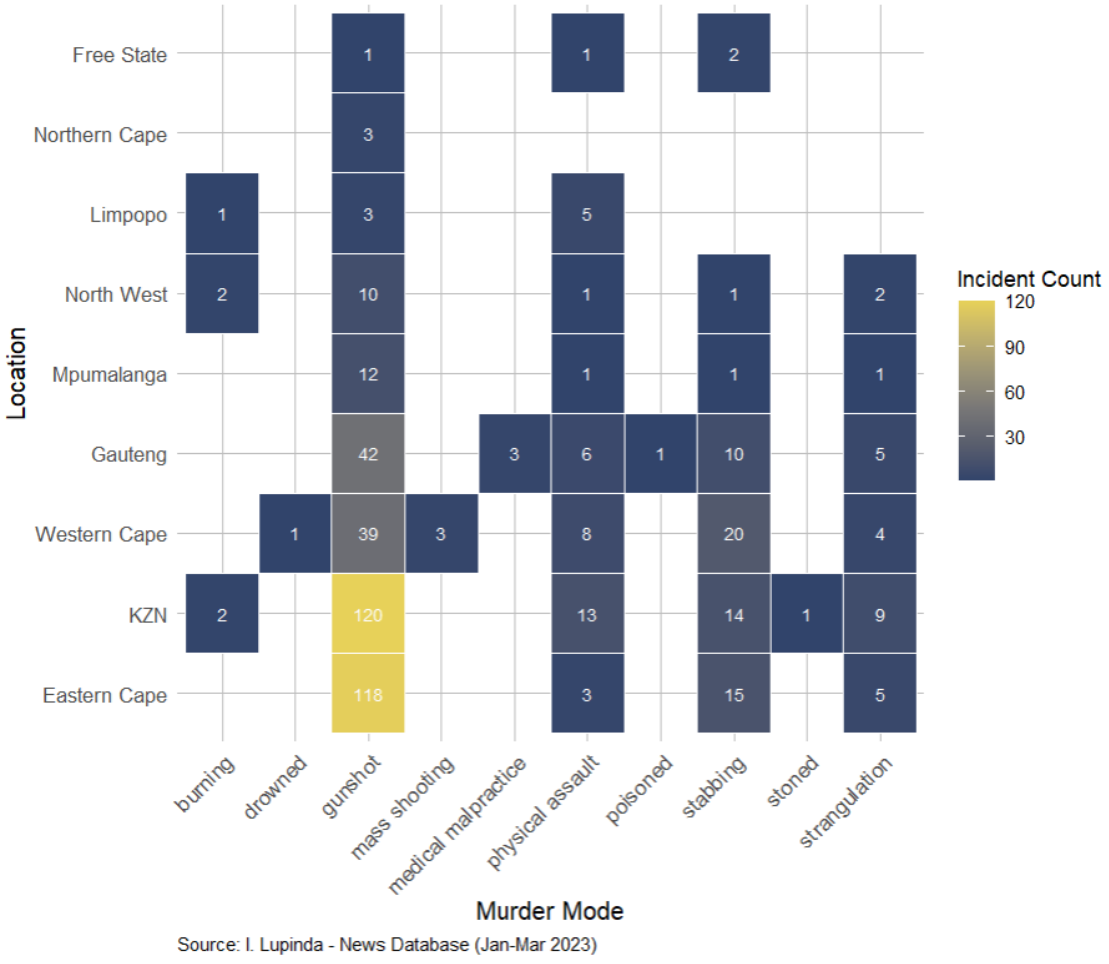
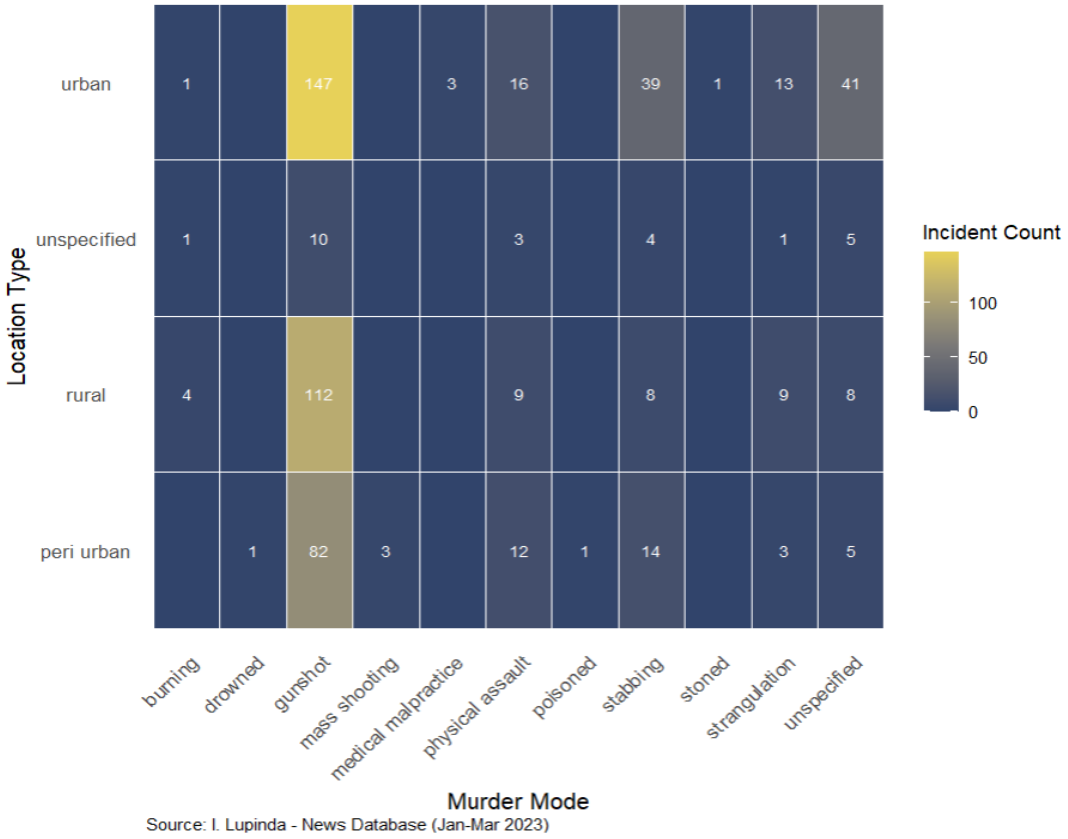


Figure 10 shows the distribution of murder incidents by location and murder type between January and March 2023. Gun-related murders are the most common, with KwaZulu-Natal (120 incidents) and Eastern Cape (118 incidents) showing the highest numbers. Other locations, such as Gauteng (42 incidents) and Western Cape (39 incidents), also report high gun-related deaths but at lower levels. Beyond gun violence, stabbings and physical assaults appear frequently in several provinces, particularly in Western Cape (20 stabbings) and Gauteng (10 stabbings, 10 physical assaults).

Less common methods such as burning, drowning, medical malpractice, and strangulation are scattered across different regions but occur at much lower frequencies.

Overall, the news data highlights regional differences in murder patterns, with KwaZulu-Natal and Eastern Cape experiencing the highest number of gun-related murders reported on, while other provinces report a mix of murder types.

Figure 11: *Association Between Location and Murder Mode*



The heatmap in Figure 11 presents the distribution of different murder modes across various location types over a three-month period in the news dataset. The vertical axis categorises locations as urban, rural, peri-urban, and unspecified, while the horizontal axis lists the types of murder, including gunshot, stabbing, physical assault, strangulation, burning, drowning, and others.

Each cell represents the number of recorded incidents, with darker shades indicating higher frequencies. The data highlights firearm-related murders as the most common, particularly in urban areas (147 cases), followed by rural areas (112 cases) and peri-urban areas (82 cases). Gun violence dominates across all location types, but it is most concentrated in cities. Stabbings are also prevalent, with 39 cases in urban areas, 14 in peri-urban areas, and 8 in rural areas. Less frequent murder methods, such as burning, poisoning, and drowning, occur sporadically across all location types, with each recording only a handful of cases. The "unspecified" category appears consistently across locations, suggesting gaps in reporting or classification.

Overall, the heatmap provides a clear visualisation of the patterns of violence, emphasising the dominance of gun-related murders in urban settings while showing that murder methods vary slightly depending on location type.

Chi-Square Test: Media Attention vs Location Type

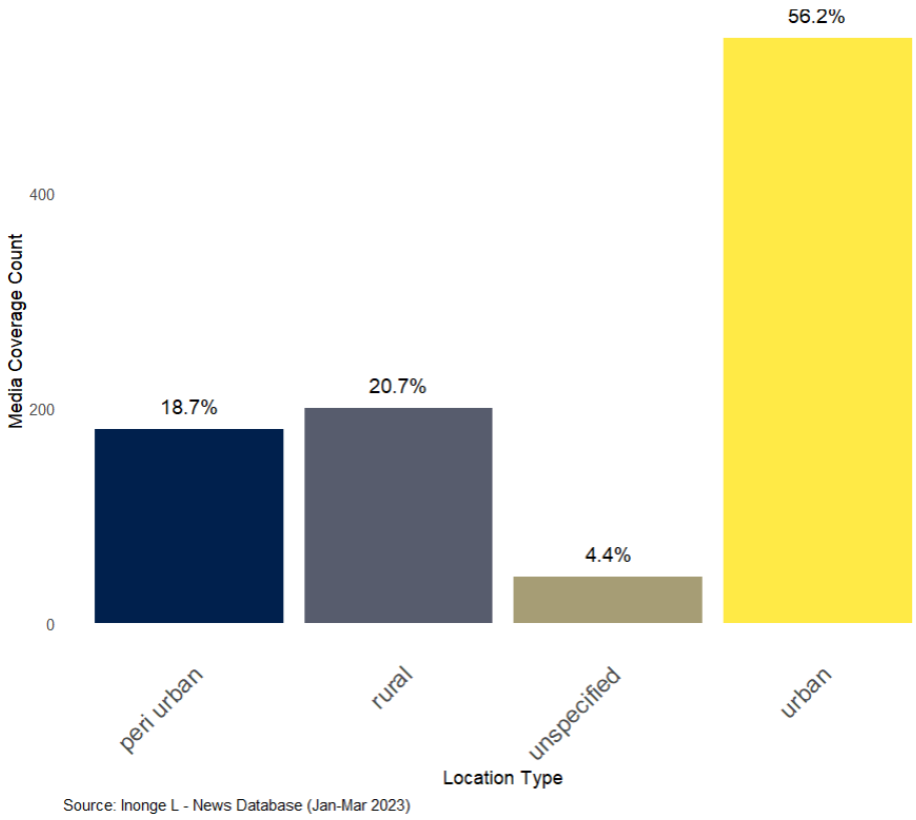
Understanding how media coverage varies across different location types is important in assessing potential biases in reporting. A Chi-square test in Figure 12 was conducted to examine whether murder cases in certain location types (i.e. urban, rural, peri-urban), are more likely to receive media attention than others. A Chi-square test is a statistical method used to determine whether there is a significant association between two categorical variables, in this case, location type and media attention (Anitha, Savarimuthu, & Bhanu, 2025). By comparing the observed distribution of media reports across these locations to what would be expected if location had no influence, this analysis helps reveal whether geography plays a role in shaping how murders are reported in the news.

The dataset contained two key variables for this analysis. As shown in Figure 12, the first variable categorised each murder report by location type, distinguishing between urban, rural, peri-urban, and unspecified locations. The second variable captured the level of media attention, measured by the number of news articles covering each case. The goal was to test whether certain location types were disproportionately reported on, suggesting that geography might influence the level of media coverage a case receives. To conduct the analysis, the data was first structured into a contingency table, counting how many murder reports appeared in the media for each location type. The Chi-square test was then performed using the "chisq.test()" function in R, which compares the observed distribution of media coverage across different locations to the expected distribution if no relationship existed between location and media attention.

A Chi-square test was conducted to examine the relationship between location type (urban, rural, peri-urban, and unspecified) and media attention. The test was statistically significant, X^2 (3, N = 556) = 563.18, $p < .0001$, indicating that murders occurring in certain location types were more likely to receive media coverage than others. This supports the conclusion that media reporting is not evenly distributed across geographic areas. These findings provide strong evidence that urban areas are disproportionately represented in murder reporting, with rural and peri-urban areas receiving considerably less attention.

These findings in Figure 12 indicate that urban areas received the highest media coverage, followed by rural areas, then peri-urban areas, and lastly, unspecified locations. This suggests that where a murder occurs plays a significant role in whether it is widely reported in the media. Various factors could explain this pattern, including higher population density in urban areas, the presence of major news outlets, or a perceived difference in the "newsworthiness" of crimes depending on location.

Figure 12: *Media Coverage of Murders by Area Type*

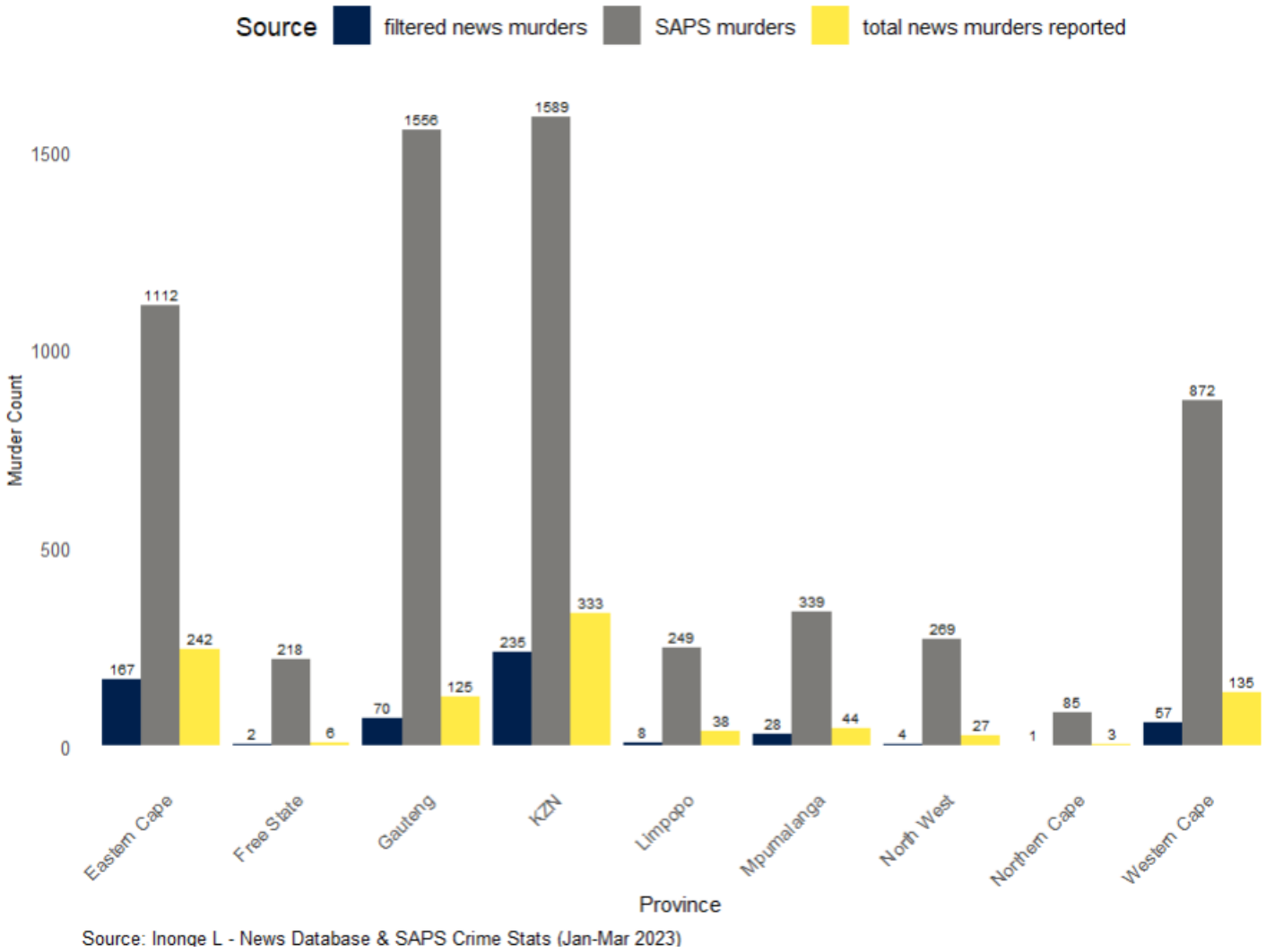


Comparative Analyses: SAPS Statistics and News Media

SAPS Statistics and News Media

When comparing provincial murder counts, it is important to consider differences in population size. Provinces with larger populations, such as Gauteng and KwaZulu-Natal, naturally record more murder cases, both in SAPS data and media coverage. To address this, raw counts (Figure 13) are complemented by per capita murder rates (Figure 14), allowing for a more meaningful comparison across provinces. Normalising by population offers insight into which provinces experience higher murder rates relative to their size. Figure 13 compares murder counts across South African provinces, using data from SAPS crime statistics and media reports between January 1st and March 31st, 2023. The figure highlights differences in officially recorded murders (SAPS data), total news-reported cases, and filtered news murders, showing variations in how murders are documented and covered by the media.

Figure 13: *Provincial Murder Counts (SAPS vs News Media (Raw Totals))*



Source: Inonge L - News Database & SAPS Crime Stats (Jan-Mar 2023)

It distinguishes between filtered news murders, which only include murders that happened between January and March 2023, and total news murders, which include all murder cases reported in the media during this period, even if they occurred earlier. The three datasets in this section provide different perspectives on murder reporting in South Africa. The “SAPS murder” data represents the official number of murders recorded by the police between January and March 2023, serving as a benchmark for comparison with media reports. The “total news murders” dataset includes all unique murder cases reported in the media, regardless of when the crime occurred, but was specifically reported in this period. This dataset captures the broader trends in media coverage, showing which cases receive attention, even if they happened outside the timeframe of the study. The “filtered news murders” dataset, on the other hand, focuses only on murders that took place between January 1st and March 31st, 2023, as reported in the news. By filtering out older cases that were only reported later, this dataset allows for a more direct comparison with SAPS records, highlighting any patterns or potential biases in media coverage. Together, these datasets help shows the differences between official crime records and media reporting, to see how murder cases are documented and represented in the public domain.

In every province, SAPS recorded far more murders than the media reported. For example, KwaZulu-Natal had 1589 SAPS-recorded murders, but the media only covered 235 filtered incidents and 333 total cases. Gauteng showed a similar pattern, with 1,556 murders according to SAPS, yet only 70 filtered and 75 total cases appeared in the news. In smaller provinces, media coverage was even lower. Northern Cape had 85 murders recorded by SAPS but only one filtered and three total news reports. Free State and Limpopo also had very limited media coverage compared to SAPS records.

Overall, media reports do not fully represent the actual number of murders occurring across South Africa, even in high-crime areas like the Eastern Cape (1112 SAPS murders, 167 filtered news reports) and Western Cape (872 SAPS murders, 57 filtered news reports). Figure 14 shows how selective media coverage shapes public perceptions, potentially giving a distorted view of murder trends. Based on the comparison shown in the graph, the murders reported in the news media are not proportionally representative of the official SAPS murder statistics. The media coverage during the three-month period tends to highlight certain cases, while leaving out many others. Provinces like KwaZulu-Natal and Gauteng have high murder counts according to SAPS, yet their representation in the news data is much lower. In smaller provinces, media coverage is even less proportional. This uneven reporting indicates that media attention is not purely based on the number of murders that occur, and that the media prioritises murders that seem especially shocking, involve prominent individuals, or are expected to generate public interest (Jewkes, 2015; Du Plessis, 2022), as seen in the news data (e.g. the extensive coverage of the murder of Kiernan Forbes or the murder of the University of Fort Hare Vice-Chancellor's bodyguard). Therefore, relying on news reports alone would provide an incomplete and skewed understanding of the overall murder trends across South Africa.

Figure 14*: Provincial* Murder Rates per 100,000 Population *(*SAPS*,* Total News-Reported Murders, and Filtered News- Reported Murders (January–March 2023)*)*

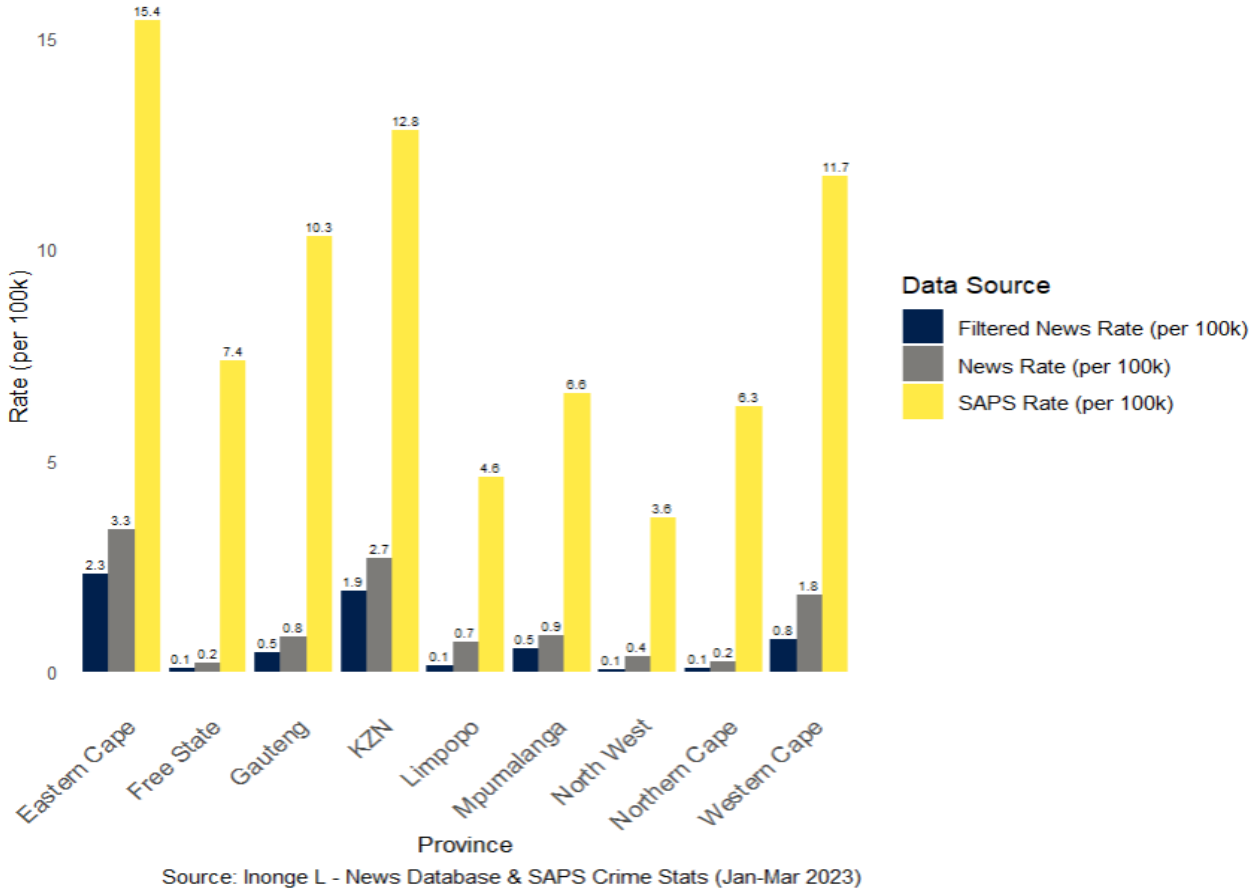


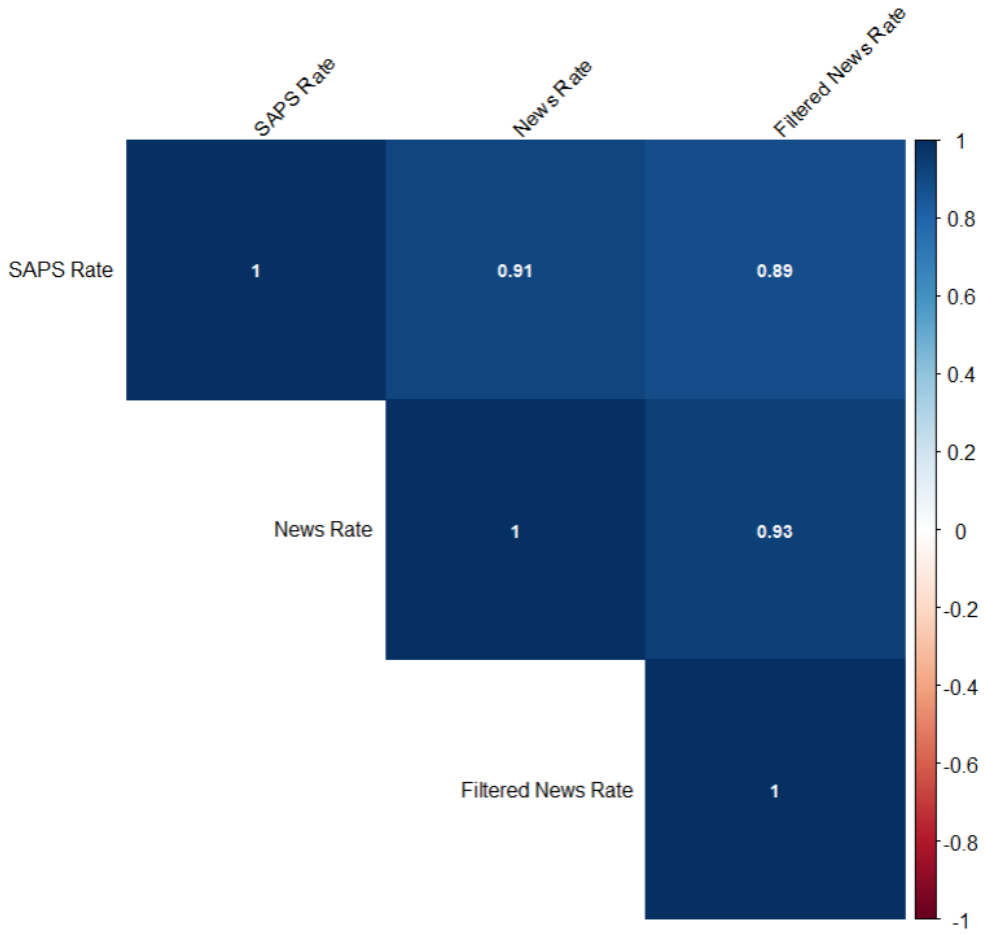
Figure 14 shows a comparison of murder rates per 100,000 population across South African provinces using the three distinct data sources: SAPS official crime statistics, total news-reported homicides, and filtered news-reported murders for January 1st to March 31st, 2023. The figure was created by merging murder counts from these sources with official provincial population data, calculating rates per 100,000 people, and visualising the results.

The findings show patterns across data sources. SAPS-reported murder rates consistently exceed those calculated from the news sources for all provinces, indicating potential underreporting or selective coverage by media outlets.

Provinces such as the Eastern Cape (15.4 per 100,000), KwaZulu-Natal (12.8 per 100,000), and Western Cape (11.7 per 100,000) have higher official murder rates. Conversely, news-reported murder rates are markedly lower, with the highest total news rate observed in Eastern Cape (3.3 per 100,000). Filtered news murder rates, on the other hand (i.e. murder incidents that took place within the specified timeframe), are even lower, as expected, reinforcing the impact of data validation processes on reported numbers. The lower coverage seen in certain provinces (e.g., Limpopo, North West, Northern Cape) compared to their official SAPS figures suggests geographical disparities in media attention. Despite normalisation by population, media coverage remains uneven. Some high-crime provinces like Gauteng are underrepresented in the news, while provinces like the Western Cape appear overrepresented. This suggests that population alone does not account for differences in media reporting. Provinces with the highest SAPS-recorded rates, such as Eastern Cape, KwaZulu-Natal, and Western Cape, also show the highest media-reported rates, though at much lower levels. This suggests that while the absolute number of covered cases is limited, the geographical patterns remain directionally consistent. The normalisation by population provides a clearer comparison of relative murder burdens and reveals how news reporting may skew public perception of crime trends by underreporting cases in certain regions.

Figure 15 shows a correlation matrix showing the strength of the relationships among murder rates derived from the three different data sources specified above.

Figure 15: *Comparison and Correlation Matrix of Murder Rates by Province*



All correlation coefficients are strongly positive and close to 1, indicating very high positive correlations among the three data sources. Specifically, the correlation between SAPS rates and total news murder rates is 0.91, while the correlation between SAPS rates and filtered news murder rates is slightly lower at 0.89. The highest correlation occurs between total news rates and filtered news rates, as expected, due to their shared source of information.

These high positive correlations suggest that, despite differences in absolute values (as indicated in Figure 14), the underlying patterns and provincial rankings remain relatively consistent across all three sources.

In practical terms, provinces identified as having high murder rates according to official SAPS data generally also appear with relatively higher murder rates in news reports (e.g., Eastern Cape, KwaZulu-Natal, Western Cape), even though the absolute figures differ significantly.

This correlation matrix reinforces the notion that media-reported data, while less comprehensive, may still reflect general trends in murder patterns. However, given the limited timeframe, these results should be interpreted with caution. Therefore, while news reports alone may not accurately capture the full extent of murder occurrences, they nevertheless reliably represent broader geographical patterns of violence. Thus, media-based murder data may be a complementary source to official statistics, particularly for identifying relative differences between provinces.

Legal Outcomes and Media Representation

Frequency of Reported Legal Outcomes of Murder Cases

To examine how murder cases reported in news media translate into legal actions, the frequency of different legal outcomes reported in murder cases was analysed. This analysis helps contextualise media representation of murder incidents by highlighting how often various legal proceedings are mentioned or documented in news reports within the studied period. Figure 16 provides insight into the legal outcomes of reported murders.

Figure 16: *Frequency of Reported Legal Outcomes*

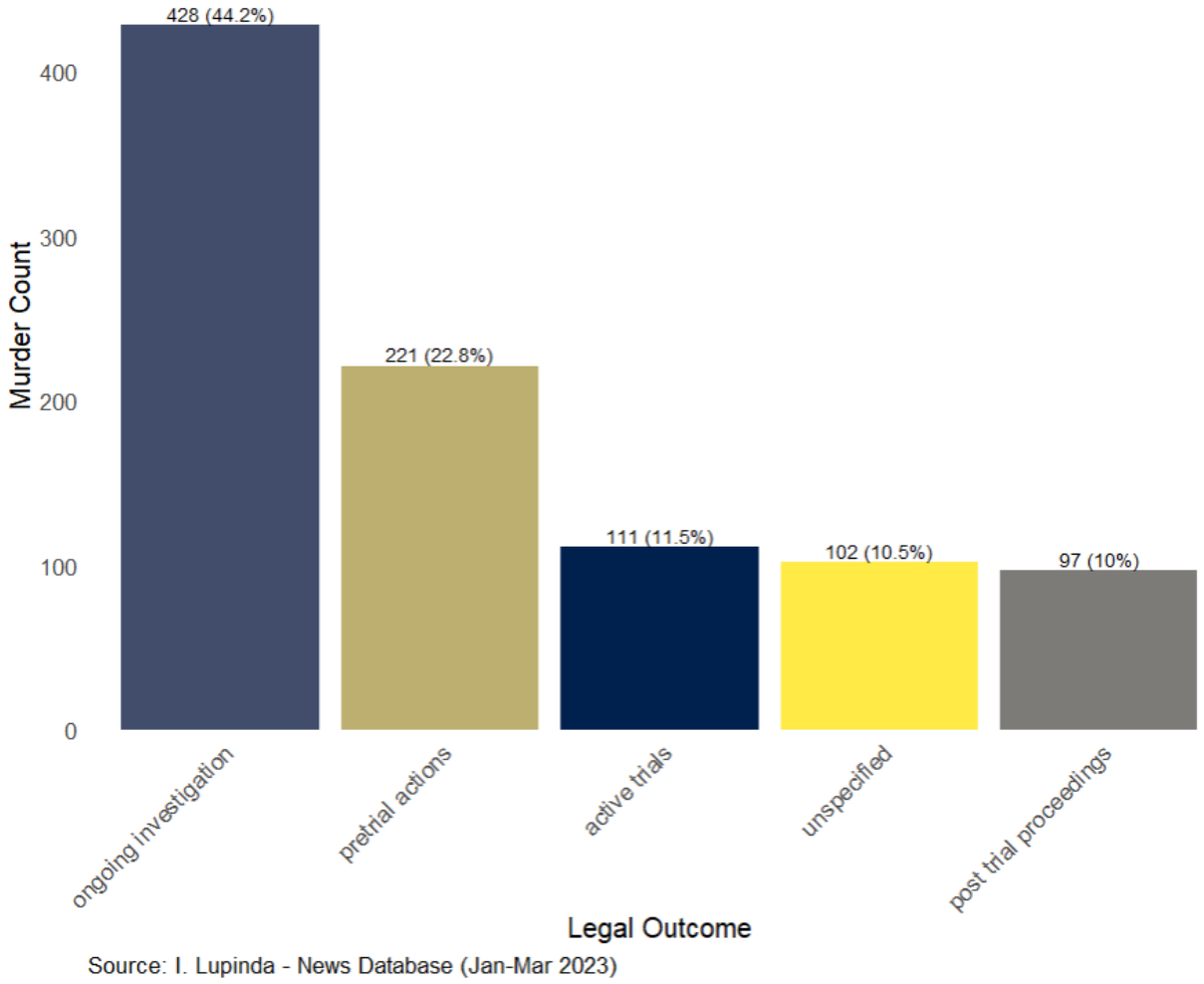


Figure 16 uses the total news-reported murder dataset, not filtered by date of incident. This approach captures all reported legal outcomes mentioned during the three-month media analysis period, regardless of when the murder occurred. Since this analysis covers only a three-month period, it is too short a timeframe to conclude delays or inefficiencies in the justice system. Murder investigations and court proceedings are inherently lengthy, therefore, the significant number of cases (428) remaining under investigation is expected within such a short timeframe. Similarly, the figures for pretrial actions (221 cases), active trials (111 cases), and post-trial proceedings (97 cases) likely reflect the standard progression of legal processes rather than indicating systemic delays.

Additionally, the 102 cases with unspecified legal proceedings highlight the need for a longer observation period to accurately assess the efficiency of the justice system.

Media Representation and Bias

Top 6 Most Reported Victims

The analysis identifies the top 6 most reported victims during the three months, from January 1st to March 31st, 2023. The top reported victim, V230, was mentioned in 61 articles, which is significantly higher than the other victims. The disproportionate coverage of certain victims may reflect media biases due to the circumstances of the murder. The demographic details of these victims show a mix of racial backgrounds, with Coloured and Black males being among the most reported. Selecting the top 6 most reported victims provides a focused and in-depth look into the factors driving media coverage of murders. The top 6 were chosen because they represent cases with the highest level of media interest, allowing the exploration of commonalities among these victims that might explain the nature of media bias. They were also selected based on a natural cut-off point in the dataset, there was a significant drop in media attention after the sixth most-reported victim. While a Top 10 or Top 3 format was considered, choosing the top 6 allowed for focused qualitative insight into the highest-profile cases without diluting the patterns of media bias. Each murder victim was assigned a unique identifier (e.g., V230) within the dataset to ensure consistency across analyses, particularly in visualisations like Figure 17. These codes are used in figures to maintain traceability to the database, but where possible, real names are provided in the text to contextualise the cases. By examining these highly reported cases, it becomes possible to identify trends, such as celebrity status, prominence in public or political roles, and emotionally engaging narratives, which may not be as apparent when considering victims with lower levels of coverage.

This approach allows for a more meaningful exploration of media priorities and the specific characteristics that elevate certain cases to national attention. Factors that contributed to the heightened level of media attention for each victim:

1. V230: AKA (Kiernan Forbes) (mentioned in 61 news articles) – Well-known local South African Artist: AKA was a well-known and influential figure in the South African music industry. As a prominent celebrity, his death attracted significant news media attention, largely due to his popularity and the shock it caused to his fanbase and the public.

2. V231: Tebello Motsoane (mentioned in 30 news articles) – Celebrity Chef and Friend of AKA: Tebello Motsoane was a celebrity chef and close friend of AKA, who was shot alongside AKA. His connection to AKA and the circumstances of his death contributed to the high media attention his case received. The fact that Tebello was killed in the same incident as AKA, a high-profile musician, likely contributed to the media's focus on his death. This kind of narrative (two prominent figures being killed together) adds a layer of drama that tends to attract more coverage. The tragedy of losing two well-known figures in the same event makes the story more newsworthy.

3. V214: 21-Year-Old University Student (mentioned in 12 news articles): The victim was a young university student, aged 21, whose death drew considerable news media attention.

4. V41: University of Fort Hare Vice Chancellor's Bodyguard (mentioned in 12 news articles): This victim was the bodyguard of the Vice Chancellor at the University of Fort Hare and was killed in an attempted hit targeting the Vice Chancellor. The nature of this incident (an assassination attempt involving a high-profile academic institution) likely contributed to the level of media coverage.

5. V357: Cloete Murray – Insolvency Specialist (mentioned in 11 news articles): Cloete Murray, an insolvency specialist, was fatally shot two days after filing an urgent application with the South Gauteng High Court. The timing of Cloete Murray's death, days after a high-profile legal action,

adds a layer of intrigue and speculation, which often draws media attention. Cases involving legal professionals, particularly those handling high-stakes insolvency matters, are likely to attract public interest due to the potential for controversy or foul play.

6. V359: Son of Cloete Murray (Insolvency Specialist) (mentioned in 5 news articles): The victim was the son of Cloete Murray; he was killed alongside his father. The tragedy of a father and son being killed together creates an emotional story that captures public interest and leads to extensive media coverage.

Figure 17: Legal Outcomes of the Top 6 Most Reported Victims of Murder

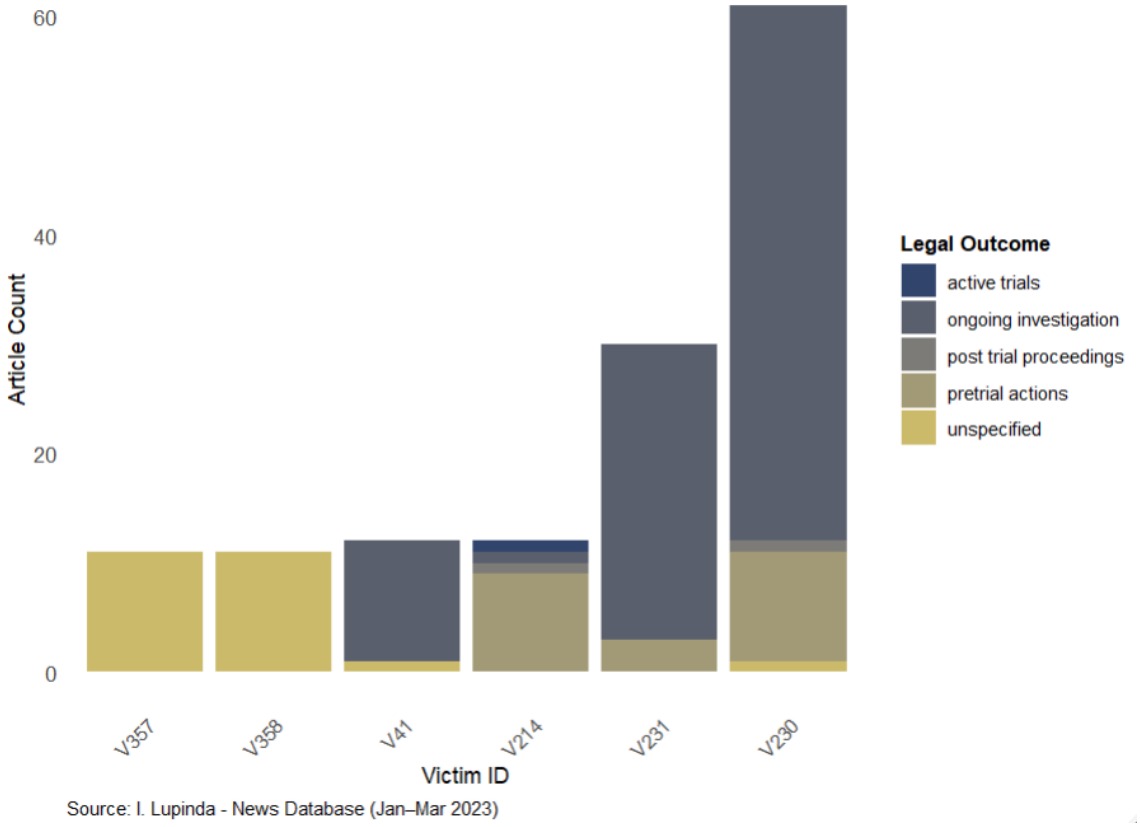


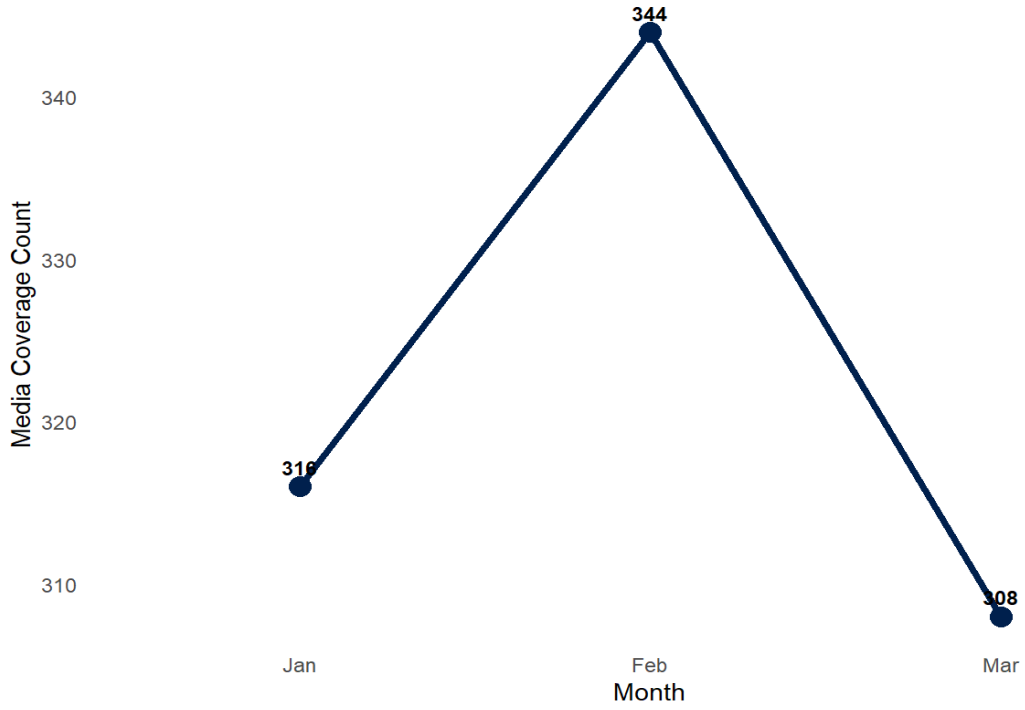
Figure 17 shows the dynamics of legal outcomes reported for the top 6 most reported victims, focusing on patterns in the progression of legal processes. For the top six most reported victims, the predominant legal outcome recorded is ongoing investigation.

V230 (AKA) and V231 (Tebello Motsoane), both high-profile individuals, had 49 and 27 instances, respectively, classified as ongoing investigations. V41 (Fort Hare Vice Chancellor's Bodyguard) also has 11 instances under ongoing investigation. Pretrial Action legal outcomes of victims like V230, V214 (University Student), and V231 each have some records of pretrial actions (e.g., V230 has 10 instances). Active trials and post-trial proceedings; only V214 has one instance recorded under active trials, and both V214 and V230 have post-trial proceedings reported once each.

Temporal Trends in Media Coverage

The temporal analysis of media coverage of murder cases in Figures 18, 19 and 20 reveals distinct patterns in how murder stories were reported over the first quarter of 2023. Figure 18 shows the monthly variation in murder coverage. The February spike corresponds to the murder of Kiernan Forbes, demonstrating how high-profile cases can skew overall reporting volume

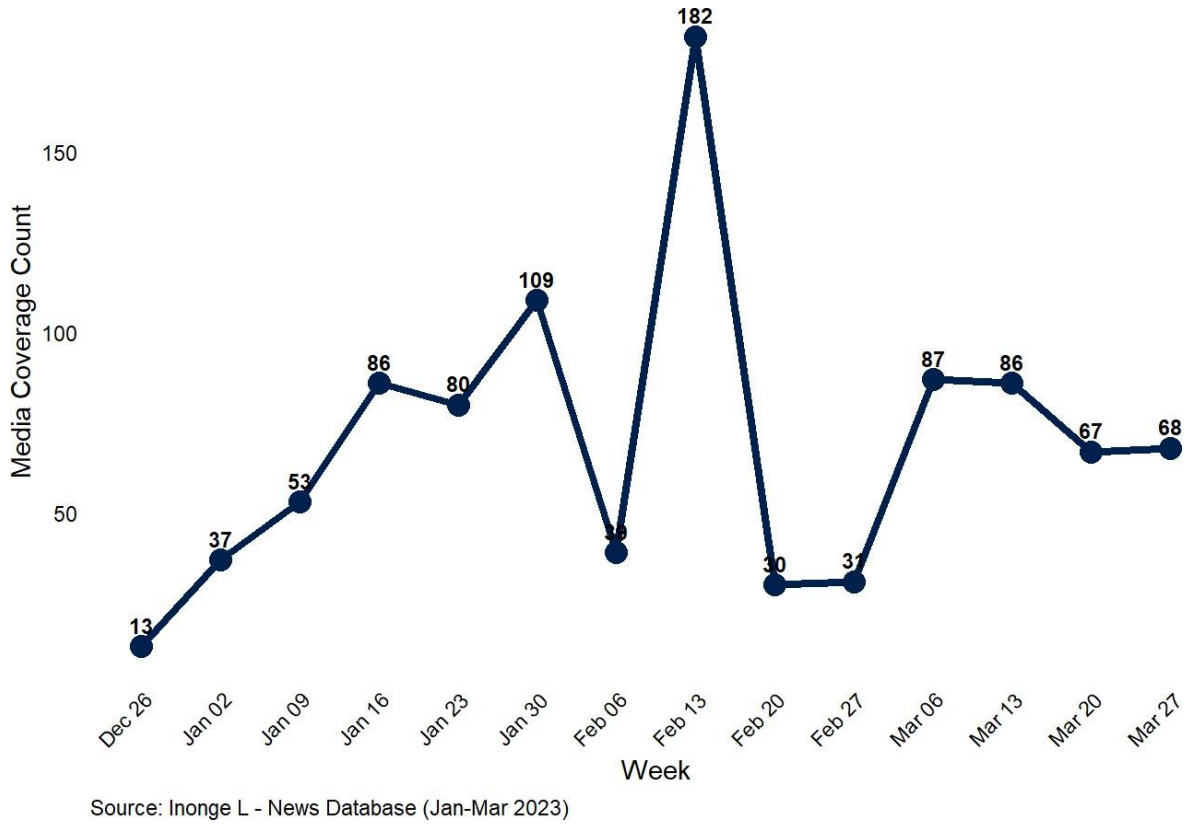
Figure 18: *Monthly Trends in the Media Coverage of Murder Cases*



Source: Inonge L - News Database (Jan-Mar 2023)

On a monthly level, coverage slightly increased from January (318 reports) to a peak in February (344 reports), before declining again in March (308 reports). The rise in February can largely be attributed to the murder of Kiernan Forbes (commonly known as AKA, a Coloured male) and his friend Tebello Motsoane (a Black male). Their deaths generated widespread media attention, with AKA's murder alone being mentioned in 61 reports.

Figure 19: *Weekly Trends in Media Coverage of Murder Cases*

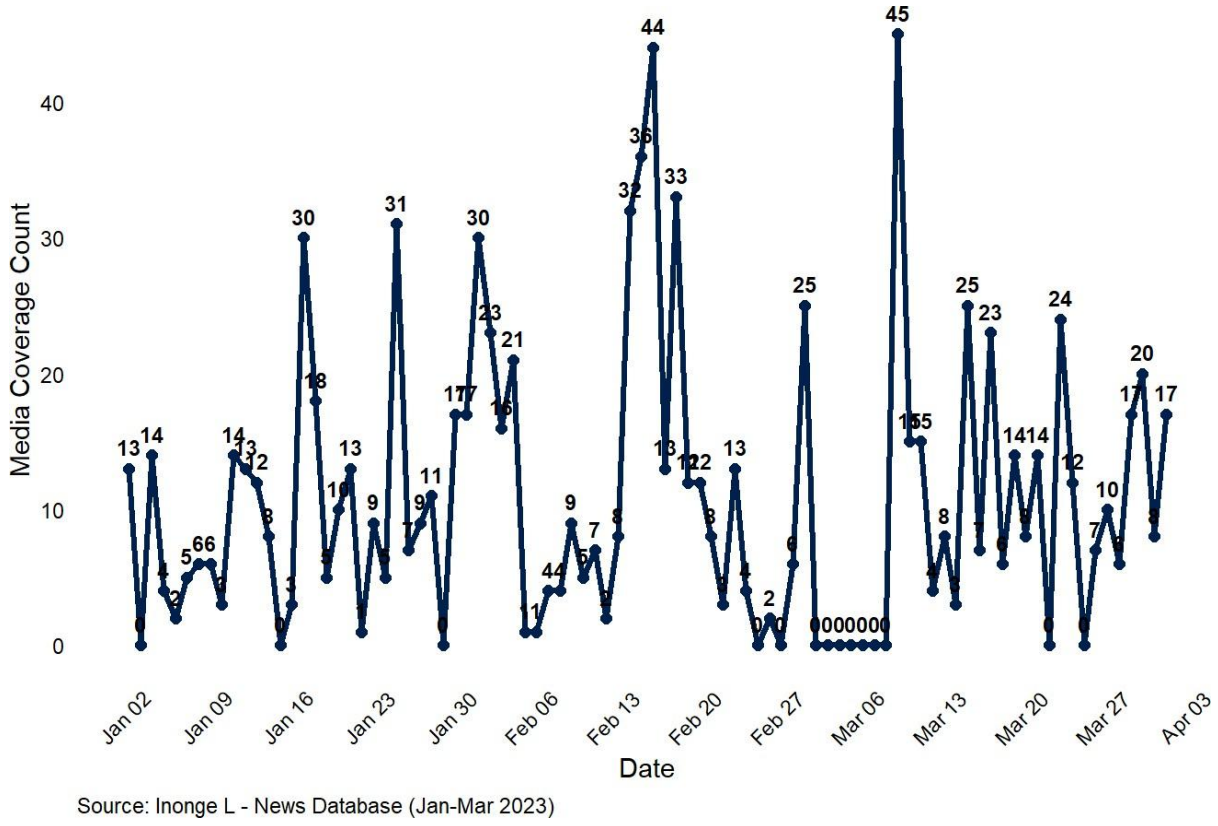


Source: Inonge L - News Database (Jan-Mar 2023)

The weekly breakdown in Figure 19 offers finer temporal resolution. It confirms the surge during the week of February 6-12, with a sharp drop-off afterwards, indicating how coverage is driven by episodic, high-profile events. At the weekly level, coverage followed a gradual increase from early January before spiking dramatically during the week of February 6-12, with 182 reports, the highest in the period. Immediately after this peak, coverage dropped sharply, reaching a low of 30 reports the following week.

The significant surge in early February coincides with the reporting of AKA and Motsoane's murder, showing how media coverage spikes in response to high-profile cases before declining once the immediate news cycle moves on.

Figure 20: *Daily Media Coverage of Murder*



The daily analysis further supports this trend, showing several distinct peaks throughout the three-month period. The highest daily coverage occurred around mid-February, particularly on days when AKA and Motsoane's deaths were widely reported, with 44 and 45 reports on consecutive days. Meanwhile, a noticeable period of very low or zero coverage appears in early March, suggesting either a temporary lull in murder news or a shift in media priorities away from crime reporting.

Overall, the data highlights that media coverage of murders is highly episodic, reacting more to specific, high-profile events rather than maintaining a consistent reflection of crime trends. Certain cases, such as AKA's murder, receive extensive attention, while others, despite occurring in the same period, may go largely unreported. This selective focus shapes public perceptions of crime, often amplifying narratives while overlooking broader trends in murder incidents.

Victim Demographics and Media Attention

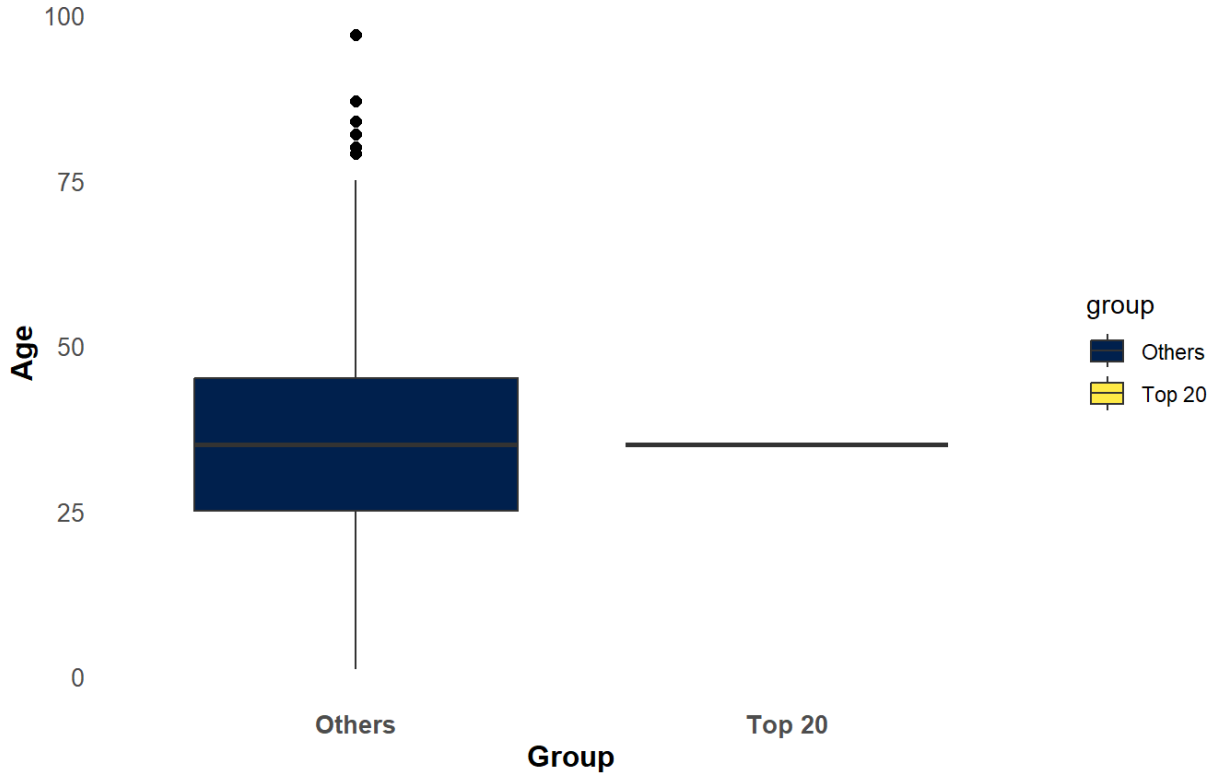
Age and Significance

Figure 21 and Table 7 examine whether the age of murder victims influences the level of media coverage they receive. The box plot compares the age distribution of victims in the 20 most covered cases against all other cases. The Top 20 comparison was used to examine whether certain victim characteristics, such as age, gender, or race, are associated with a higher likelihood of receiving extensive media coverage. By isolating the most covered cases, this analysis helps identify patterns of media selectivity (who is most visible in the public eye, and why). Comparing this group with all other murder victims provides insight into whether media attention reflects broader crime patterns or is influenced by biases and social perceptions of victim worthiness. This is not meant to be a predictive model, but rather an exploratory framing to test whether "high visibility" cases differ systematically from those that are less reported.

To explore whether demographic characteristics influence media attention, murder victims were divided into two groups: the Top 20 most-mentioned individuals in news reports, and all others. The Top 20 were selected based on frequency of media mentions. This threshold allowed for a focused comparison of whether factors such as age, gender, and race differ between high-visibility and lower-visibility cases.

This comparison helps identify whether certain characteristics increase the likelihood of elevated media attention. Figure 21 presents this comparison using age as the predictor.

Figure 21: *Age Distribution Between Top 20 Victims and Others*



Source: Inonge L - News Database (Jan-Mar 2023)

Table 7: *Welch Two-Sample T-Test Results for Age Between Top 20 Victims and Others*

	T.Statistic	Degrees.of.Freedom	P.Value	Confidence.Interval	Mean.Difference
t	0.739	727	0.46	(-0.749, 1.653)	0.452

While the Top 20 group has a slightly lower median age and a more concentrated distribution, the overall variation is minimal. A Welch Two Sample[5] t-test was conducted to determine if these differences are statistically significant. The t-statistic of 0.739 and a p-value of 0.46 indicates that any difference in average age between the two groups is likely due to random chance. The confidence interval (-0.749 to 1.653) further confirms that age does not significantly impact media coverage during this period of review.

Overall, the analysis suggests that the age of a murder victim does not determine the level of media attention their case receives. Both groups show a similar median age of around 35 years, indicating that other factors, rather than age alone, drive media coverage.

Gender and Significance

Table 8: *Chi-Square Test Results for Gender Distribution Between Top 20 Victims and Others*

	Chi.Square.Statistic	Degrees.of.Freedom	P.Value
X-squared	15.521	2	< .0001

Table 7 examines whether the gender of murder victims influences media coverage by comparing the Top 20 most-covered cases to all other cases. The chi-square statistic of 15.521 suggests a notable difference between the observed and expected gender distributions, with a p-value of $p < .0001$ indicating that this difference is highly unlikely to be due to chance. The result

[5] Welch Two-Sample t-test is a statistical test used to compare the means of two independent groups when their variances are unequal and/or when the sample sizes differ. Unlike the traditional Student's t-test, which assumes equal variances, Welch's test adjusts the degrees of freedom to provide a more reliable result under these conditions. It is used where group variances might differ, ensuring that comparisons of means are robust even when the data are heteroscedastic (i.e., have unequal variances (Delacre et al., 2017).

was evaluated at $\alpha = 0.05$, confirming that, in the media set prepared for analysis, gender distribution differs significantly between high-profile cases and others. Since the p-value is statistically significant, the analysis rejects the assumption that the gender distribution in the top 20 cases is the same as in other cases. The statistically significant difference in gender distribution among top-covered cases points to potential media bias in favouring certain gender narratives, particularly in high-profile murders; however, this effect may also be an artefact resulting from a single high-profile case involving two male victims. Different results might be found if a similar analysis were performed on a three-month period where there had been a high-profile killing of a female victim (for example, Reeva Steenkamp).

Race and Significance

Table 9: *Chi-Square Test Results for Race Distribution Between Top 20 and Others*

	Chi.Square.Statistic	Degrees.of.Freedom	P.Value
X-squared	144.315	5	0

Table 8 provides a statistical examination of whether there's a significant difference in the race distribution of murder victims among the top 20 most-covered cases compared to all other cases. This test helps to identify if media coverage varies based on the race of the victims, potentially pointing to biases or disparities in media attention. The significant Chi-square statistic of 144.315 with a p-value of 0 indicates that the race distribution in the top 20 most covered murder cases is not the same as in the others. This statistically significant result suggests that certain racial groups may be more likely to be represented in media reports of murder cases, either due to the nature of the cases themselves or due to media biases in which cases are deemed newsworthy.

Murder Modes and Significance

Table 10: *Chi-Square Test Results for Murder Mode Distribution Between Top 20 and Others*

	Chi.Square.Statistic	Degrees.of.Freedom	P.Value
X-squared	10.466	10	0.401

Table 9 provides insights into whether there's a statistically significant difference in the murder modes of murder victims among the top 20 most-covered cases compared to all other cases. This test is used to explore potential variations in media coverage based on the method of murder.

The Chi-square statistic of 10.466 with a p-value of 0.401 suggests that the differences in murder mode distributions between the Top 20 most-covered cases and the Others are not statistically significant. This implies that the type of murder method used does not significantly influence whether a murder case receives media coverage relative to other cases. This analysis indicates that murder mode alone does not appear to be a defining factor in the level of media attention a murder case receives.

Predictive Analyses

Predictors of High Media Coverage of Murder

To identify which factors influence whether a murder case receives high media coverage, two logistic regression models were run using gender, race, murder mode, and location type as predictors. Odds ratios and confidence intervals were calculated from the first model, while a second model was refined by excluding high-profile outliers that disproportionately influenced estimates.

The base model found that location type was the strongest predictor. Murder cases that occurred in urban areas were over twice as likely to receive high media coverage compared to unspecified areas (OR = 2.08, 95% CI [1.06, 4.26], p = .038), while rural cases were significantly less likely to be covered (OR = 0.30, 95% CI [0.10, 0.80], p = .020). Gender showed a marginal effect, with male victims slightly less likely to be covered (OR = 0.58, p = .063), although this was not statistically significant. Race and murder mode were not significant predictors, and many of their coefficients were highly unstable due to data sparsity.

To improve model reliability, a refined logistic regression was run using a deduplicated dataset and removing high-leverage outliers (based on Cook's Distance and residual diagnostics). In this second model, urban location remained statistically significant, with an even stronger effect (Estimate = 3.5064, p = .0022), confirming that urban cases consistently attract more media attention. However, rural location was no longer significant, suggesting that the effect in the base model may have been driven by a few extreme cases. Gender and race remained non-significant in the refined model, with racial variables especially unstable due to missing data in the media reports.

These models support the conclusion that urban location is the most consistent predictor of media visibility, while other demographic characteristics like race and gender showed no clear influence in this dataset. The strongest takeaway is that urban murders receive significantly more media attention, reinforcing the role of location in shaping news coverage. However, gender and race are not consistent predictors in this dataset, indicating that media focus may be influenced by other factors, such as case sensationalism, victim status, or crime circumstances. Overall, these comparisons underline the importance of addressing outliers and ensuring balanced data representation for reliable interpretations. Addressing data imbalances and high-profile case distortions is crucial to improving the reliability of models assessing media bias in crime reporting.

Predictors of Time to Report a Murder based on Demographic Variables and Murder Modes

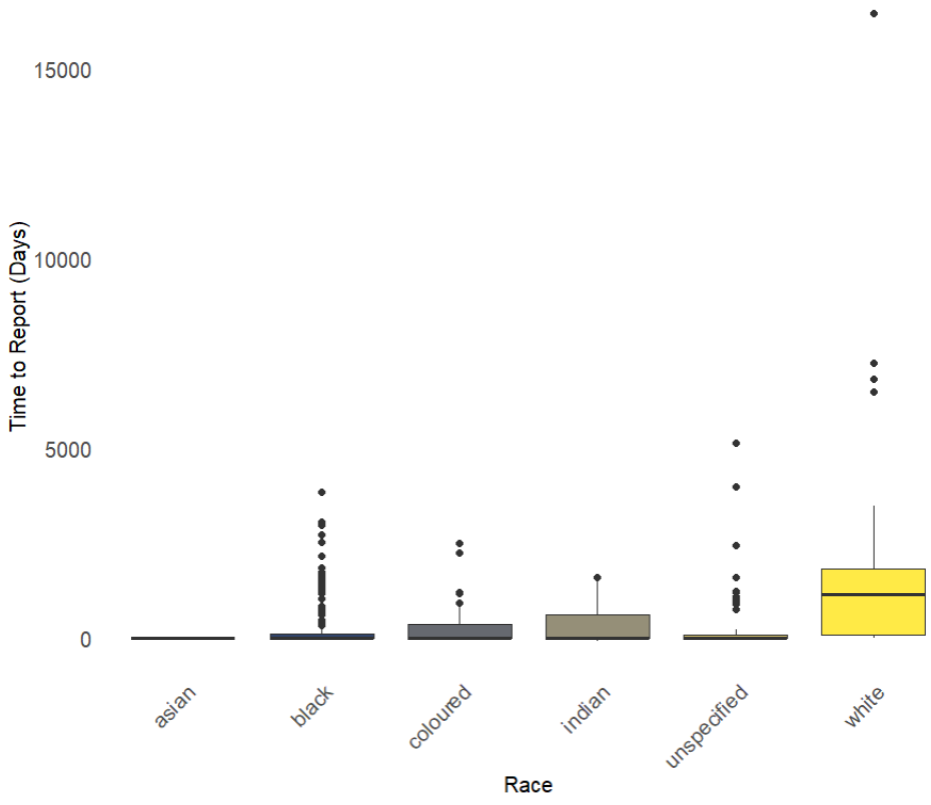
This section examines whether demographic factors and murder modes influence the time it takes for a murder to be reported in the media, with a focus on racial disparities in reporting delays.

Race

Figure 22 provides a visual representation of how quickly different racial groups have their murder cases reported in the media. It visualises potential racial disparities in the time taken for a murder to be reported in the news.

The wider range and outliers among White victims suggest less consistent coverage timelines, though this requires cautious interpretation given small subgroup sizes. The plot shows significant variability in the time to report across different racial groups.

Figure 22: *Time to Report and Race*

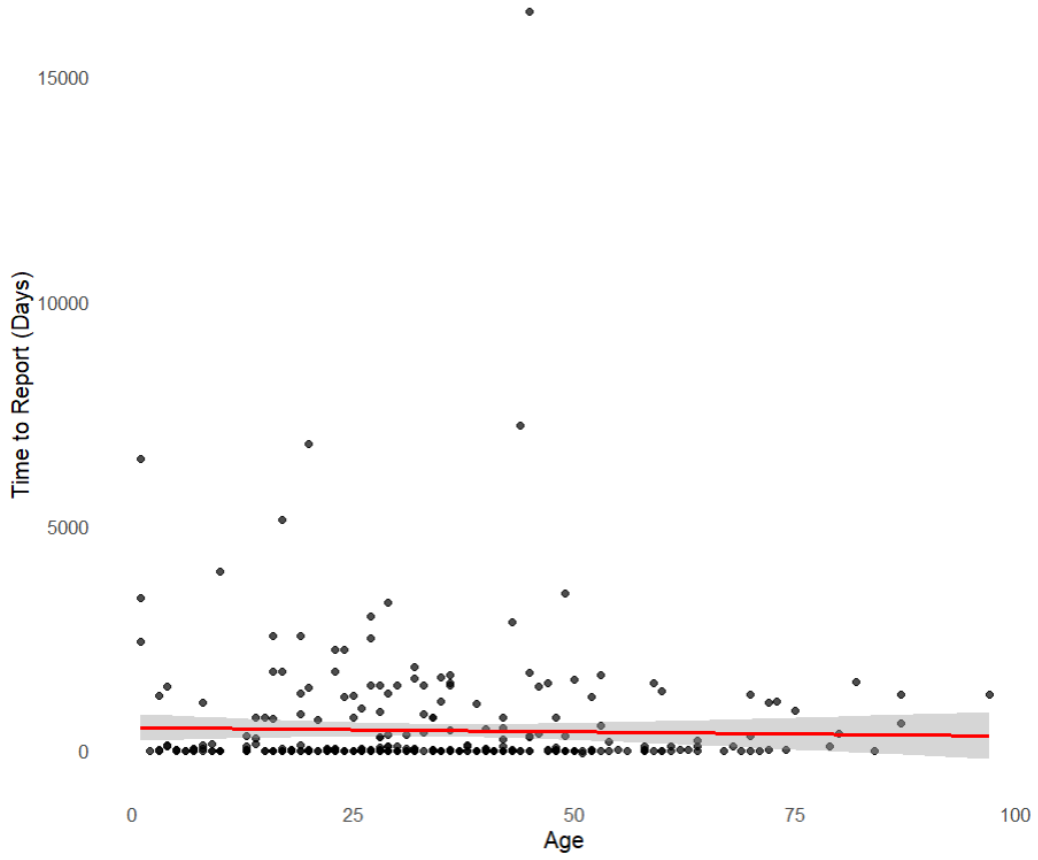


This indicates a potential disparity in media attention or the efficiency of the reporting process based on the race of the victim. The White category has a notably wider interquartile range, suggesting greater variability in the time taken to report murders involving White victims, compared to other races. All racial categories except "Asian" exhibit outliers, indicating that there are some cases where the time to report is exceptionally high. This might reflect complex cases or cases with extensive legal or investigative processes. The presence of outliers in most racial groups except for "Asian" might indicate either underreporting within this category or more consistent reporting times.

Age

To further explore patterns of media attention, the analysis now considers the relationship between the victim's age and the time taken for their murder to be reported in the news media. Figure 23 shows that there is no clear trend relating age to the time taken to report a murder.

Figure 23: *Time to Report and Age*

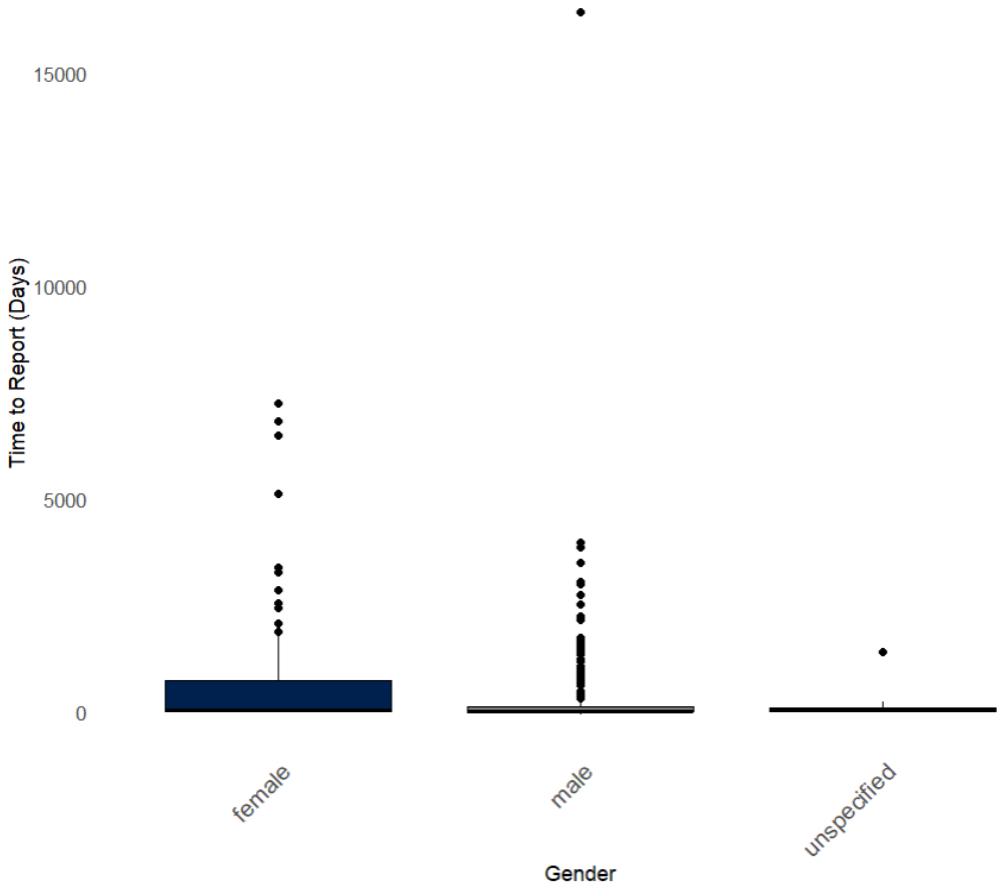


The time to report is scattered across different ages without showing any apparent pattern, indicating that age, as a standalone variable, may not have a significant impact on the delay in reporting.

Gender

Next, the analysis examines whether gender plays a role in influencing how quickly murder cases are reported in the media. Figure 24 reveals that there are outliers in the reporting time across all genders, but particularly among females.

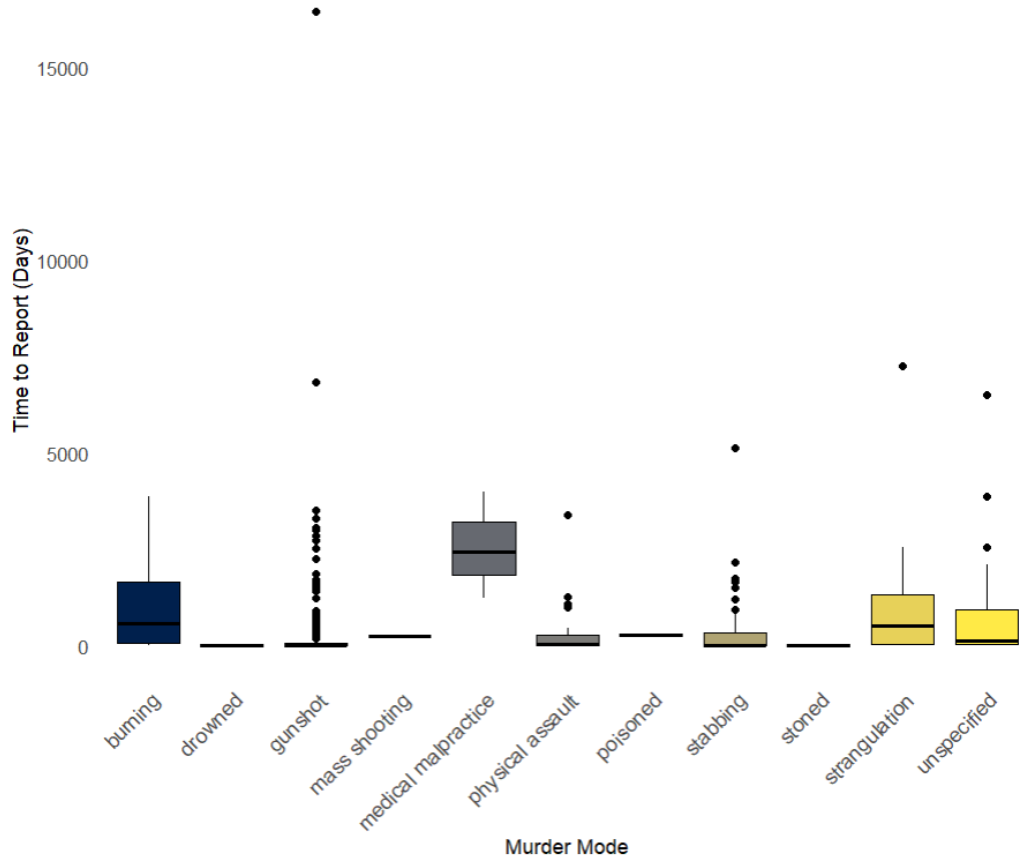
Figure 24: *Time to Report and Age*



This might suggest that cases involving female victims can sometimes have exceptionally long reporting times. However, the bulk of the data for each gender category clusters near zero, suggesting that for most cases, gender does not heavily influence reporting delay.

The boxplot grouped by Murder Mode in Figure 25 shows varying levels of reporting times across different modes of murder.

Figure 25: *Time to Report and Murder Mode*



Certain modes like gunshot and strangulation show more consistency in reporting times (tighter boxes and fewer outliers), whereas medical malpractice and physical assault have wider ranges. This could suggest that the nature of the crime influences how quickly it is reported, potentially due to the clarity or ambiguity around the event's classification as a murder.

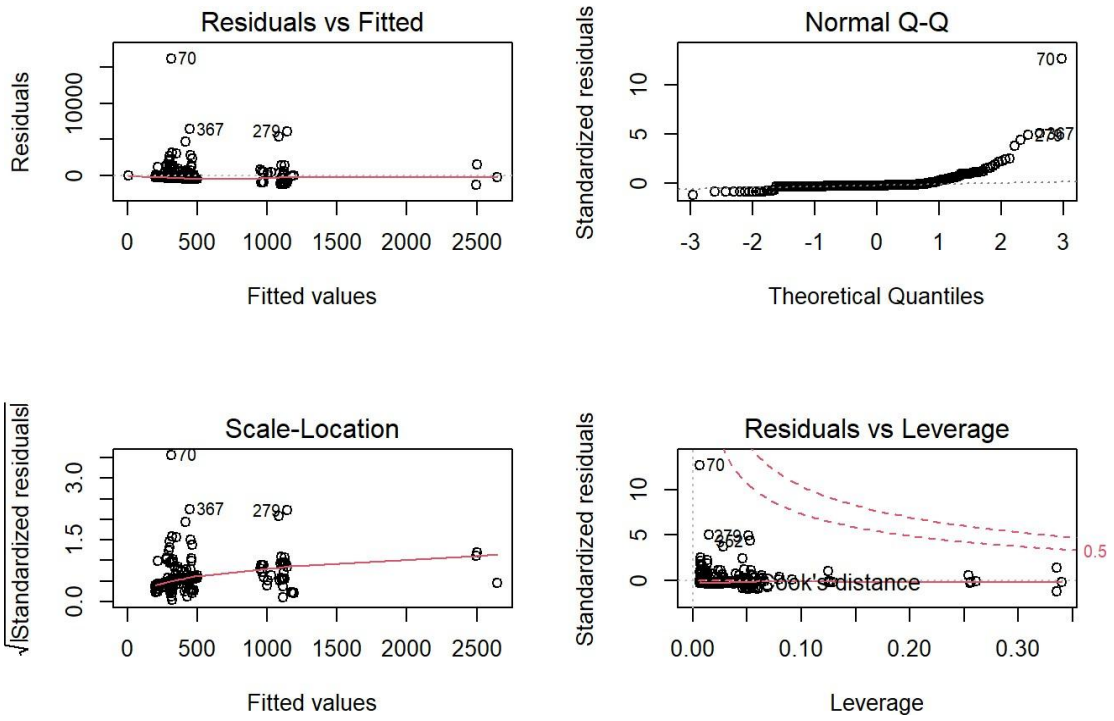
The analysis suggests that while age does not significantly impact reporting time, factors like murder type and the victim's gender may play a role. The presence of outliers indicates that some cases experience unusual reporting delays, likely influenced by multiple interacting factors. These complexities are explored further in the discussion section.

To ensure the reliability of the regression models, diagnostic tests were conducted, including an assessment of multicollinearity using the Generalised Variance Inflation Factor (GVIF). A GVIF close to 1 indicates minimal multicollinearity, meaning predictors do not overly influence each other. The results, shown in Table 11 below, confirm that all predictors fall within acceptable limits, ensuring that the model's estimates remain stable and reliable.

Table 11: *GVIF Values for Predictors*

Predictor	GVIF	GVIF^(1/(2*Df))
Age	1.194	1.093
Gender	1.446	1.097
Race	1.628	1.050
Murder Mode	1.646	1.032

Figure 26: *Model Diagnostics*



To evaluate the reliability of the multiple linear regression model predicting the time taken to report murders, diagnostic tests were conducted. Table 11 shows Generalised Variance Inflation Factor (GVIF) values for predictors included in the regression model. GVIF values below 5 indicate low multicollinearity (Hair et al., 2019), confirming that predictor relationships did not compromise the model's estimates.

Diagnostic plots (Figure 26) assessed the key regression assumptions. The Residuals vs. Fitted Values plot demonstrates that residual variance remains generally constant across fitted values, suggesting homoscedasticity holds. The normality assumption was examined using the Q-Q plot of residuals, which shows minor deviations at the extremes; however, these deviations are modest and do not significantly violate normality assumptions (Hair et al., 2019). Additionally, the residuals versus leverage plot identified a few influential points, but their impact remained within acceptable thresholds, indicating the overall stability and robustness of model estimates.

In sum, diagnostic evaluations confirm that model assumptions hold sufficiently well, with no severe violations observed. However, the presence of minor outliers and modest deviations from normality suggests that results should be interpreted carefully, considering these potential limitations.

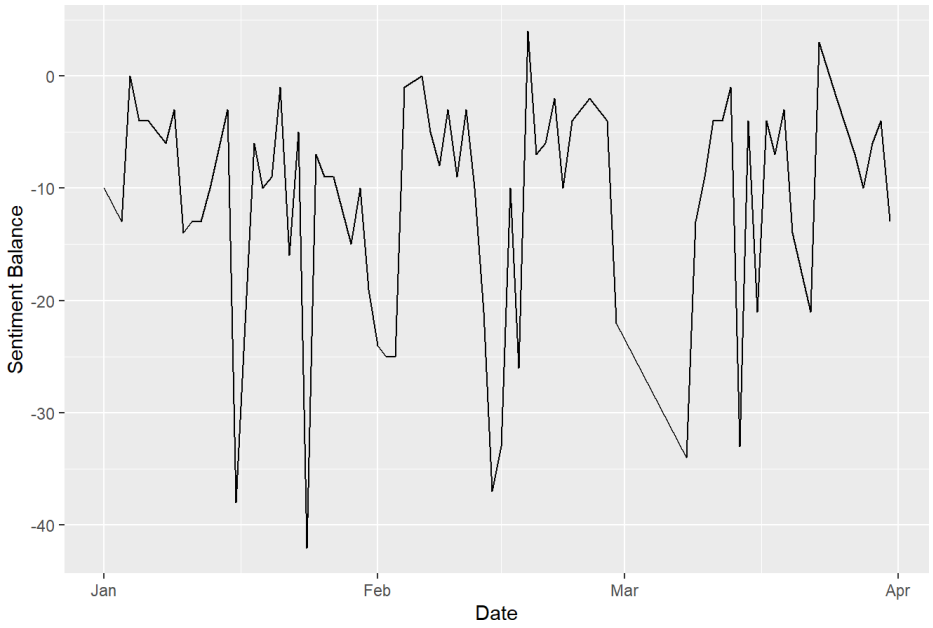
Text Mining

Topic Modelling and Sentiment Analysis

This section presents results from a text mining analysis conducted on news headlines covering murder cases in South Africa from 1st January to March 31st, 2023. The analysis includes a sentiment trend analysis and topic modelling to identify key themes and patterns in how murder is portrayed by the media. These methods provide insight into the language, tone, and focus areas prevalent in murder-related reporting, helping to reveal underlying media narratives and framing of violent crime.

Figure 26 tracks the sentiment balance of murder-related headlines over time, showing how the tone of reporting changes from January to early April.

Figure 26: *Sentiment Balance of News Murder Headlines Over Time*



Sentiment Analysis was performed using the 'bing' lexicon from the tidytext package. Each headline was tokenised into individual words, joined with the lexicon, and then scored based on the number of positive and negative words it contained. The sentiment balance was computed by subtracting the number of negative words from the number of positive words per date. This allows for a general sense of the emotional tone of media coverage over time, as shown in Figure 26.

Topic Modelling was conducted using Latent Dirichlet Allocation (LDA) from the topicmodels package. The document-term matrix was created by tokenising the headlines and counting word frequency per victim_id. A topic model with k = 3 was chosen for interpretability and based on preliminary testing using coherence and visual interpretability. The model identified three dominant topics, each representing a distinct cluster of words. This helped uncover recurring narrative themes in murder reporting.

The sentiment remains largely negative, which is expected given the nature of crime reporting. However, there are sharp declines on certain dates, linked to major murder cases or violent incidents that received heavy media coverage. While there are brief periods where sentiment is less negative, the overall trend suggests that news coverage rarely includes more neutral or positive aspects, such as justice being served or crime prevention efforts. As anticipated in the introduction of this study, the sample size and method of analysis provided extremely limited utility for sentiment analysis, and these results should be considered within these limitations.

For the topic modelling, one dominant topic includes words like “murder”, “cop”, “AKA”, and “KZN”, indicating media focus on high-profile murder cases, police-related incidents, and regional crime patterns. The presence of words like “trial” and “justice” suggests that some coverage extends beyond the crime itself to legal proceedings. This analysis shows that while murder reporting primarily focuses on the crime, it also includes elements of law enforcement and judicial outcomes.

Together, these findings show that media coverage of murder is overwhelmingly negative and focuses on crime events, victims, and legal proceedings. Certain regions and high-profile cases receive more attention, raising questions about how crime is framed in the media and whether some incidents are prioritised over others in public discourse.

CHAPTER 5: DISCUSSION

This research sought to analyse the representation of murder incidents in South African news media over a defined three-month period, from 1st January to 31st March 2023. By using computational techniques alongside manual data collection, the researcher investigated how different demographic and circumstantial characteristics, including age, gender, race, mode of murder, and geographical location, influence media coverage. The objective was to understand the interactions and selective nature of reporting and uncover potential biases, shedding light on how the media shapes public perception of violent crime in South Africa, over a three-month period, from January 1st to March 31st, 2023.

Patterns in Murder Victim Demographics

Age Distribution

The findings in this research confirm that young adults, particularly men in their late 20s and early 30s, are disproportionately represented among murder victims. With an average victim age of 34 years and a mode of 28 years, this aligns with global patterns where individuals in this age range face higher exposure to violence (World Health Organization, 2022). This is consistent with prior research showing that young men in South Africa are at greater risk of both perpetration and victimisation (Brodie, 2019; Langa et al., 2018; Brookman et al., 2017). The high levels of youth unemployment, income inequality, and urbanisation likely contribute to this pattern, increasing the likelihood of exposure to violence (Peterson, 2020). These findings reinforce the need for targeted policy interventions, particularly those focused on economic stability, community-based programs, and youth engagement strategies that could help mitigate risk factors.

While young adults make up most of the recorded victims, at times, media narratives do not always align with statistical realities. The selective reporting of elderly and child victims suggests a media bias in crime coverage, reinforcing the “ideal victim” framing (Brodie, 2019). Elderly victims, though less frequently murdered, tend to receive disproportionate attention, particularly in non-intimate femicide cases. Similarly, child victims are often overrepresented in the media, creating the perception that they are at higher risk than statistics suggest. Meanwhile, domestic violence and neglect-related murders, which disproportionately affect both the elderly and children, are often underreported because they do not fit into dramatic or sensationalised news narratives (Brodie, 2020).

These findings highlight that if the media overrepresents certain victim groups while neglecting others, this can create misconceptions about crime patterns and skew resource allocation. A more balanced approach to crime reporting is needed to ensure that coverage reflects actual risks rather than reinforcing existing narratives.

Gender Dynamics

The gendered nature of murder in South Africa is further compounded by racial and socioeconomic factors. Black and Coloured women, for instance, face higher risks of intimate partner violence due to systemic inequalities, while White women may receive more media attention, reinforcing racialised narratives of victimhood. The findings in this analysis reinforce the gendered nature. According to the Centre for the Study of Violence and Reconciliation (CSVR, 2008), 94% of murder suspects are male, with nearly half between the ages of 20 and 29. Male victimisation follows similar patterns, particularly among young men aged 15 to 29, who experience murder rates of up to 184 per 100,000, with even higher rates in disadvantaged urban

areas (Groenewald et al., 2008). The news dataset reflects these trends, with men accounting for 54% of murder victims, far outnumbering women (34.5%). The most common cause of death for male victims is gun-related violence (214 cases compared to 83 for women), which aligns with studies showing that men are more likely to be killed in public spaces following altercations, robberies, or gang-related conflicts (CSVR, 2008).

While male victims dominate overall numbers, female murders are more likely to occur in intimate or familial settings. South African women face one of the highest femicide rates globally, with 24.7 per 100,000, and over half of these murders are linked to IPV (Mathews et al., 2015; Seedat et al., 2009). The findings support this, showing that strangulation, a common method in domestic violence cases, is far more prevalent among women (22 cases compared to 4 for men). Additionally, 44 women were killed by intimate partners, and 15 by family members, reinforcing the strong link between gender-based violence and murder (Abrahams et al., 2013). The persistence of these patterns reflects deeply entrenched patriarchal norms and power dynamics, where women are more vulnerable to violence in private spaces (Jewkes et al., 2011; Morrell, 2001; Wood & Jewkes, 2001).

Masculinity and socio-economic marginalisation also play a key role in male victimisation. The General Strain Theory (GST) suggests that economic hardship, social exclusion, and blocked opportunities can fuel frustration and aggression, leading to higher crime rates (Agnew, 1992). In South Africa, high youth unemployment and income inequality exacerbate these issues, making violent conflict among men more common (Peterson et al., 2020). The concept of hegemonic masculinity further reinforces violence as a means of asserting dominance and status, especially when traditional roles as providers are unattainable (Ratele,

2008; Wood, Lambert, & Jewkes, 2007). As a result, many male-on-male murders occur in public spaces, often involving strangers or acquaintances.

Media coverage also shapes public perceptions of gendered violence, often amplifying sensational cases while overlooking domestic violence and familial murders (Mawby, 2010). Male victimisation is frequently linked to public violence, while female murders gain attention primarily in cases of IPV, mass shootings, or serial killings (Mathews et al., 2015; Lindegaard, 2010). This selective framing reinforces gendered stereotypes, portraying men as active participants in violent conflicts and women as passive victims (Jewkes et al., 2011).

The findings in this study align with global trends, where men make up around 80% of murder victims, with firearms as the leading cause of death (UNODC, 2019). However, the nature of murder in South Africa is strongly shaped by local socio-economic conditions. The high rates of male-on-male violence in public spaces contrast with the prevalence of intimate partner and domestic femicides, highlighting the structural factors driving different types of gendered violence.

These trends point to the relevance of broader structural and social factors in shaping patterns of violence. Prior literature suggests that economic opportunity, education, and community-based programmes may play a role in reducing male-on-male violence by providing alternatives to conflict (Peterson et al., 2020; Ratele, 2008). Similarly, addressing gender-based violence may benefit from legal frameworks that support victims of IPV public awareness initiatives, and access to support services (Mathews et al., 2015; Abrahams et al., 2013). While this research does not assess interventions directly, the persistence of such patterns in both data and literature underscores the importance of considering how structural conditions and media representations may intersect to shape public narratives around violence.

Race Representation

The findings in this analysis highlight the relationship between race, socio-economic inequality, and media coverage in shaping perceptions of murder in South Africa. While the racial distribution of murder victims in the dataset largely mirrors national demographics, patterns in the news database emerge in how different racial groups experience and are represented in violent crime. Black victims accounted for approximately 49.6% of murder cases in the dataset, despite comprising around 81% of the national population (StatsSA, 2022). This significant underrepresentation in media coverage is not necessarily indicative of lower victimisation rates, but rather reflects media selection patterns, particularly the prioritisation of urban, high-profile, or sensational cases. Prior research confirms that Black South Africans face the highest rates of violent crime due to structural inequalities, yet these cases are often underreported in mainstream media (Seedat et al., 2009; Brodie, 2019). This suggests a media bias in which the deaths of Black individuals are rendered less newsworthy unless tied to unusual or high-status circumstances.

Coloured individuals, who represent 8.8% of the population, account for 6.2% of murder victims in the dataset. While this appears proportional, past studies indicate that Coloured men are overrepresented in dispute-related murders, reflecting long-standing socio-economic marginalisation and a culture of interpersonal violence in certain communities (CSVR, 2008; Brookman et al., 2017). While Coloured individuals are not significantly overrepresented in gang-related murders, their exposure to street disputes and social pressures increases their risk of victimisation (Abrahams & Jewkes, 2005). The relatively muted media coverage in this dataset may reflect the normalisation of violence in these communities, where recurring murder incidents attract less media urgency unless tied to extraordinary circumstances.

White South Africans make up 6.5% of the dataset, closely matching their 7.3% share of the population, but their victimisation is disproportionately amplified in the media, particularly in cases where White females, children and public figures are involved (Brodie, 2019). Media narratives frequently prioritise White victims, reinforcing perceptions of their heightened vulnerability despite statistically being at lower risk of murder compared to Black and Coloured individuals (Brodie, 2019; Mawby, 2010). The Oscar Pistorius case (which was reported as an anniversary murder in the dataset) exemplifies how racialised fears of Black criminality can influence public narratives, with Pistorius' defence constructing an imagined Black intruder as a justification for violence (Langa et al., 2018). White victims were often featured in high-profile, heavily covered stories, such as the Reeva Steenkamp case, suggesting a media bias in favour of White victim visibility. While not overrepresented numerically, White victimisation appears amplified in narrative prominence and follow-up coverage. Brodie (2019) notes this is especially true in Afrikaans-language media, though similar patterns are visible in the English-language dataset used in this study.

Indian and Asian individuals represent 5.1% of murder victims, nearly double their 2.6% share of the population. This overrepresentation may be due to concentrations of Indian communities in urban areas, where news reporting is more frequent and detailed. However, it also raises questions about socio-economic and spatial factors influencing victimisation within this group, suggesting a need for further research.

Notably, 16% of cases lacked racial information, highlighting the inconsistency of media reporting and limiting the precision of racial analysis. This data gap reflects a broader issue in news-based crime research, where editorial omissions and source ambiguity create blind spots in demographic representation.

The racial disparities in murder methods also highlight unequal exposure to different types of violence. Black victims experience a wider range of murder methods, including gunshots, stabbings, physical assaults, and strangulations, reflecting their heightened exposure to systemic violence and gang-related crime (Matzopoulos et al., 2006). In contrast, White and Indian victims are more commonly killed by gunshot, suggesting a different pattern of risk linked to armed crime rather than interpersonal violence (Kagawa et al., 2018).

One of the major limitations in racial analysis is the high percentage (16%) of unspecified racial data, which underscores the lack of standardised reporting practices. These data gaps reduce the ability to fully understand how race, structural inequalities, and violence intersect in South Africa.

The findings reinforce how race and socio-economic status shape patterns of violence, but also how media reporting selectively frames these narratives. Addressing these disparities requires structural interventions, including reducing inequality, improving access to economic opportunities, and challenging racial stereotypes in crime reporting. Targeted violence prevention programs, particularly for marginalised Black and Coloured communities, are crucial for mitigating the disproportionate burden of violence on these groups. Without a more balanced media representation of crime, public perceptions will continue to reinforce racialised fears and misconceptions about victimisation and criminality.

Murder Modes and Trends

The analysis of murder methods in South Africa highlights the dominance of firearm-related violence, which accounts for 64% of all recorded incidents. This aligns with concerns about the availability of firearms, organised crime, and systemic socio-economic inequalities (Matzopoulos et al., 2006; Peterson, 2020).

Between 16 and 18 people are shot and killed daily in South Africa, making firearms the most common weapon used in murders (Langa et al., 2018) The dataset reflects this trend, with KwaZulu-Natal (KZN) and the Eastern Cape recording the highest numbers of gun-related murders, regions that are known for high levels of gang violence, political assassinations, and social unrest (Matzopoulos et al., 2006). The easy access to firearms, both legal and illegal, plays a central role in this crisis, with an estimated 4.5 million guns in circulation, many in civilian hands (Langa et al., 2018). While firearms are often acquired for self-defence, studies show that their presence increases the likelihood of violence, particularly in domestic settings, where many intimate femicide-suicides involve legally owned guns (Brodie, 2019).

Gang violence further fuels South Africa's murder crisis, particularly in marginalised communities. Gangs are often heavily armed, contributing to high murder rates among young men through drug-related violence, territorial disputes, and organised crime (Brookman et al., 2017). Firearms are also frequently used in carjackings and armed robberies, reinforcing public fears and the perceived necessity of gun ownership for protection. This creates a cycle of violence, where high crime rates drive demand for firearms, further escalating murder rates (Sierra-Arévalo & Papachristos, 2017; Pyrooz, Turanovic, & Wu, 2015).

IPV and femicide remain major concerns, with firearms frequently used in these crimes (Brodie, 2019). Research shows that more than 80% of intimate femicide victims in South Africa are killed with firearms, often with a single gunshot (Brodie, 2019). The dataset reflects this trend, showing that women are more likely to be shot by their intimate partners, reinforcing the links between gun ownership and domestic violence. These findings support broader research indicating that patriarchal norms, emphasising male control over women, continue to drive IPV-related murders (Jewkes et al., 2011).

Although the Firearms Control Act of 2000 initially led to a decline in gun-related murders, firearm violence remains widespread due to illegal trafficking and corruption within law enforcement (Brodie, 2019; Langa et al., 2018). Stolen and illegally sold police and military firearms often end up in the hands of criminal networks, undermining gun control efforts (Brodie, 2021). Additionally, deep socio-economic inequalities continue to fuel violence, with blocked economic opportunities and social marginalisation leading to interpersonal and gang-related conflicts (Brookman et al., 2017).

Gun violence in South Africa is also shaped by racialised perceptions of crime and self-defence. Among White South Africans, gun ownership is often framed as necessary protection against perceived threats from Black perpetrators, reinforcing apartheid-era fears and racial stereotypes (Langa et al., 2018).

Reducing firearm violence requires comprehensive policy interventions, including stricter gun control enforcement, targeting both legal and illegal firearms, and holding corrupt officials accountable for enabling the illegal arms trade (Brodie, 2021). Addressing gang violence requires community-based programs that focus on economic empowerment, youth intervention, and conflict resolution. Efforts to combat IPV and femicide must consider the role of firearms in domestic violence, while also challenging deeply ingrained societal norms around masculinity and power (Jewkes et al., 2011). Given the high rates of gun-related murders, tackling South Africa's firearm problem is critical to reducing overall murder rates and fostering safer communities.

Geographical Analysis of Murders

The geographical distribution of murders in South Africa reflects deep-rooted social, economic, and historical inequalities, with violence concentrated in disadvantaged urban areas and townships.

The dataset aligns with prior research, showing KwaZulu-Natal (KZN) and the Eastern Cape as the provinces with the highest murder counts, with 168 and 141 cases, respectively. These figures, though significantly lower than official SAPS statistics (1,556 in KZN and 1,112 in the Eastern Cape), indicate the disparity in media coverage, where urban and high-profile cases receive more attention than rural or underprivileged communities (Brodie, 2019; Peterson, 2020).

The Western Cape, particularly Cape Town, also reflects high levels of gang-related violence, with 98 reported murders, reinforcing existing findings that gang activity and territorial disputes play a key role in murder rates (Peterson, 2020). In contrast, provinces such as the Free State and Northern Cape reported far fewer cases (4 and 3, respectively), a pattern consistent with broader trends showing that rural areas experience lower overall murder volumes, despite localised spikes in violence (Peterson, 2020).

The prevalence of firearm-related murders, which account for 64% of all cases, is particularly pronounced in urban and peri-urban areas, where gun access, gang conflicts, and economic disparities drive violent crime (Matzopoulos et al., 2006). The dataset reveals that KZN and the Eastern Cape lead in gun-related murders, reflecting the high availability of firearms and entrenched organised crime networks in these regions.

The historical and structural legacy of apartheid continues to shape these geographical patterns. Townships, which were originally designed as racially segregated areas for Black and Coloured populations, remain disproportionately affected by violent crime, with high rates of unemployment, poverty, and gang-related disputes (Brookman et al., 2017). Overcrowding in these communities pushes social interactions into public spaces, increasing the likelihood of conflicts escalating into violence. Additionally, a culture of retaliation and street justice often normalises violence, making it a recurring feature in some communities (Gravel, Valasik, Mulder, & Tita, 2023).

Addressing geographical patterns in murder rates requires region-specific interventions that target the root causes of violence. Socioeconomic upliftment programs, including education, job creation, and improved access to social services, are crucial in breaking cycles of poverty and crime. Community-based initiatives, such as violence prevention programs and youth engagement strategies, can provide alternatives to gang involvement and promote conflict resolution skills. Additionally, community policing initiatives that rebuild trust between law enforcement and marginalised communities can enhance crime reporting and prevention efforts (Peterson, 2020).

The logistic regression analysis confirmed that urban areas were far more likely to receive media coverage than rural locations, reinforcing media biases that prioritise high-profile, city-based murders over those occurring in more remote areas. This imbalance in reporting limits public awareness of rural crime trends and skews policy responses toward urban-centric solutions (Peterson et al., 2021). Moreover, the dataset's high proportion of unspecified locations and missing data points highlights gaps in crime reporting that hinder a comprehensive understanding of South Africa's murder landscape.

Ultimately, addressing South Africa's geographic disparities in violence requires a dual approach: tackling the socio-economic conditions that drive crime while improving crime data collection and media representation of underreported regions. A more balanced reporting system, coupled with targeted interventions in high-risk areas, is critical in creating safer communities and reducing the country's persistently high murder rates.

Media Coverage and Reporting Bias

The media coverage of murder in South Africa is heavily shaped by biases related to race, gender, socioeconomic status, and the perceived "newsworthiness" of cases (Brodie, 2019; Spies, 2020). These biases influence which stories receive extensive coverage and which are ignored, creating a distorted picture of crime in the country.

One of the most apparent biases is racial disparity in reporting. White victims receive disproportionate attention, while murders of Black and Coloured individuals, who make up most victims, are often underreported (Brodie, 2019). The media also reinforces racial stereotypes in crime reporting by frequently depicting Black and Coloured individuals as suspects, fuelling racialised perceptions of crime (Spies, 2020). Additionally, murders of White farmers often receive special coverage, while killings of Black and Coloured farmworkers in similar rural settings do not receive the same attention (Brodie, 2019).

Gender bias is also evident in reporting, particularly in cases of intimate femicide. Despite IPV being one of the leading causes of female murder, these cases are often underreported, with the media instead focusing on murders by strangers (Brodie, 2019). This selective framing skews public perceptions, making it seem as though women are primarily at risk from unknown attackers, rather than partners or family members. The media also frames victims differently based on race and gender, often searching for complex backstories to justify or explain the actions of White perpetrators, while quickly villainising Black suspects (Brodie, 2019).

Socioeconomic status also plays a role in coverage bias. High-profile or "mega-cases" dominate the news, while murders in low-income communities often go unreported (Brodie, 2019; Spies, 2020). Media outlets prioritise urban stories, reinforcing the perception that crime is an urban issue while neglecting violence in rural and marginalised areas (Brodie, 2019). Additionally, newsroom dynamics and a lack of diversity among journalists contribute to these

one-sided narratives, shaping which stories make it into the news and how they are framed (Brodie, 2019).

The framing of murder cases also influences public perception. Media narratives frequently focus on episodic events, portraying murders as individual incidents rather than symptoms of systemic problems like poverty, inequality, and gender-based violence (Aldrete & Fernández-Ardèvol, 2023; Spies, 2020). This narrow focus obscures the broader patterns behind violent crime, preventing meaningful public discussion and policy interventions. The media's tendency to sensationalise certain cases, particularly those involving celebrities, public figures, or high-status professionals, reinforces the idea that some lives are more valuable than others.

The murder of Kiernan Forbes (AKA) received extensive coverage, with 61 reports and ongoing updates in February, far outpacing coverage of other murder cases that occurred during the same period. This trend is consistent with Brodie (2019), who found that the social role, occupation, and public visibility of a victim heavily influence the extent of media attention.

The legal outcomes of murder cases also influence media interest. Cases that remain open, with ongoing investigations, receive more attention than those that have concluded (Brookman et al., 2017). The high number of reports on high-profile victims with ongoing investigations, such as Kiernan Forbes (AKA), aligns with this trend. The media thrives on uncertainty, keeping audiences engaged with updates on arrests, court proceedings, and speculation. This was evident in the coverage of AKA's murder, which remained in the headlines for weeks due to investigations, potential suspects, and public interest in his legacy. By contrast, ordinary murder cases, particularly those that conclude without controversy, receive little to no follow-up coverage.

This selective reporting creates the impression that justice is pursued more vigorously for some individuals than for others, reinforcing public scepticism about law enforcement and judicial fairness (Brookman et al., 2017).

Media coverage also shapes public narratives about crime, justice, and policy priorities. When high-profile cases dominate headlines, they can skew law enforcement priorities, directing resources and public pressure toward specific incidents rather than addressing broader systemic drivers of crime (Peterson et al., 2021). This imbalance in attention can result in distorted crime prevention strategies, where isolated cases drive policy changes instead of evidence-based approaches.

Another notable bias in media coverage is xenophobia. Whenever national identity was explicitly mentioned in the dataset, it referred to foreign nationals (e.g., individuals from Mozambique, Zimbabwe, Botswana, or Asian backgrounds). Studies show that South African media have contributed to xenophobic tensions, with headlines and reporting styles portraying foreign nationals as criminals or threats (Mgogo & Osunkunle, 2021). This framing leads to anti-immigrant sentiments and contributes to misguided public perceptions about crime and nationality.

To address these biases, media outlets must take a more responsible approach to crime reporting. This includes accurate representation of data, ensuring all victims, regardless of race, gender, or socioeconomic status, receive fair coverage, and contextualising murder cases within broader societal issues (Brodie, 2019). The media should also challenge harmful stereotypes, particularly in the way they portray Black and Coloured victims and suspects and avoid framing murders of White individuals as exceptional tragedies while downplaying violence in marginalised communities (Spies, 2020). Finally, journalists should engage with diverse sources beyond institutional figures like the police, incorporating perspectives from violence prevention experts, social workers, and affected communities (Spies, 2020).

By shifting toward more balanced and responsible reporting, the media can play a constructive role in shaping public understanding of crime, rather than reinforcing fear, stereotypes, and social divisions.

Predictors of Media Coverage

The predictors of media coverage in murder cases reveal how demographic factors, case specifics, and broader societal biases shape which cases receive extensive attention and which go largely unnoticed. Using two logistic regression models, this analysis examined the likelihood of a murder case being widely covered in the news. The comparison between the models emphasised the importance of refining statistical analysis by addressing outliers and redundancies, ultimately strengthening the framework for understanding what drives media interest in certain cases.

One of the most significant findings is the role of urban bias in murder reporting. The significance of urban location suggests that murders in urban areas are far more likely to receive coverage, reflecting the concentration of media houses, ease of access to information, and audience interest in city-based crime. This aligns with Peterson (2020) and Brodie (2019), who found that metropolitan areas such as Gauteng and the Western Cape dominate media reporting, while violence in rural and peri-urban areas is often underrepresented. Rural cases, by contrast, were significantly underreported. While gender showed a marginal trend toward lower coverage for male victims, this did not reach statistical significance. While regression results did not identify race or gender as statistically significant predictors, this may reflect data limitations, including missing race labels and an overrepresentation of urban or high-profile cases in media data, rather than the absence of bias. Qualitative trends and prior literature continue to suggest that race shapes media narratives.

Research consistently shows that White victims receive disproportionate attention, particularly in cases that are framed as rare, unusual, or particularly brutal (Brodie, 2021). The murder of Reeva Steenkamp by Oscar Pistorius serves as an example; Steenkamp's case dominated headlines, reinforcing the media's tendency to prioritise White victims who fit the "ideal victim" stereotype (Langa et al., 2018). By contrast, murders of Black women, even in cases of IPV, are often overlooked (Brodie, 2019).

This selective coverage shapes public perceptions of who is most at risk, often minimising the violence experienced by Black and Coloured communities. Additionally, cases with readily available information, such as official police statements, witness accounts, and visual evidence, are more likely to be reported since news agencies rely on accessible data to produce quick, compelling stories (Brodie, 2019). The role of newsroom dynamics in shaping crime coverage is also crucial. Journalists and editors make decisions based on editorial priorities, audience interest, and resource availability (Brodie, 2019). The reliance on wire services like SAPA (South African Press Agency) further reinforces selective reporting, as pre-packaged news tends to focus on high-profile or easily reportable cases, rather than offering a comprehensive view of crime across different communities.

This selective approach to murder coverage contributes to a distorted perception of crime in South Africa. By prioritising White victims, high-profile cases, and urban violence, the media reinforces racial and class-based biases, leaving Black, Coloured, and rural communities underrepresented in crime reporting. This skewed coverage has real-world consequences, influencing public opinion, law enforcement priorities, and policy decisions.

While statistical analysis suggests that race and gender are not significant predictors of media coverage, qualitative findings from Brodie (2021) and Jewkes (2004) show that race plays a major role in how murder cases are framed. Black victims, despite being the majority in the dataset, are less likely to receive widespread media attention, unless their cases involve unusual

circumstances or sensational elements. Meanwhile, Coloured victims tend to be overrepresented in dispute-related murders, a trend also noted in Brookman et al. (2017).

The findings of this study align with broader literature on media bias in crime reporting. Jewkes (2004) highlights that media attention is driven by news values, prioritizing cases based on rarity, social proximity, and the victim's status. Lindegaard (2010) further explores how gender and race shape crime narratives, reinforcing stereotypical depictions of suspects and victims. O'Hear (2020) confirms the overrepresentation of urban crime in media, reflecting logistical advantages and audience demand for city-based news.

Ultimately, the predictors of media coverage in South Africa highlight systemic inequalities in crime reporting. Sensationalised cases, racialised narratives, and urban bias distort public understanding of who is most at risk of violent crime. Addressing these issues requires greater media accountability, balanced reporting practices, and a commitment to representing all victims, regardless of race, gender, or location.

This research aimed to investigate patterns in murder reporting in South African news media and assess the influence of demographic, geographical, and situational factors on media attention. The findings reveal a clear urban and sensationalism bias, with high-profile, city-based murders receiving disproportionate coverage. While statistical models did not find race and gender to be significant predictors, qualitative patterns and literature indicate persistent media framing biases. The gap between official crime rates and media representation reinforces that murder reporting is shaped less by incidence than by editorial choices, public interest, and systemic inequalities.

CHAPTER 6: STUDY LIMITATIONS

While this research provides valuable insights into the representation of murder in South African media, several limitations need to be acknowledged, as they may have influenced the findings and interpretations. These limitations span across data coverage, timeframes, media biases, methodological concerns, and generalisability. Recognising these constraints provides context for the study's findings and highlights areas for further research.

Language and Media Outlet Coverage

The analysis was limited to English-language South African newspapers included in the NewsBank platform. This means that coverage from online news platforms, Afrikaans newspapers and other language publications was not included, potentially missing important perspectives, especially given the multilingual nature of South African society. As Brodie (2020) indicates, different language media can have varied focuses, with Afrikaans publications often highlighting different narratives compared to English ones. Consequently, the findings may not fully represent how murder is portrayed across the wider South African media landscape, thereby limiting the generalisability of the study's results.

Temporal Scope

Additionally, the research focused on a limited time frame, specifically January to March 2023. This short duration may not be representative of broader trends in media coverage over the entire year, which could include fluctuations based on factors like seasonal crime trends, political events, or economic conditions. Longer-term studies would be beneficial to establish whether the patterns observed are consistent throughout the year or years.

Media Bias and Reporting Constraints

The study relied on data from news media reports, which may be subject to journalistic bias and reporting constraints. Certain types of murders, particularly those involving marginalised communities or occurring in rural areas, were underreported and portrayed with limited depth. The selective nature of media reporting means that the dataset likely underrepresents certain types of murders, particularly those involving marginalised groups or occurring in rural areas. This creates potential bias in understanding the true nature of murder in South Africa, as some incidents may not be reported or may be reported without sufficient depth.

Data Quality and Completeness

Further, the analysis was constrained by the availability of data, particularly concerning legal outcomes. Several cases in the dataset were marked as “ongoing investigation” or “unspecified,” particularly in relation to legal outcomes. Additionally, some articles lacked complete demographic or case-specific details, such as the victim’s age, race, or relationship to the perpetrator. These data gaps limited the depth of statistical analysis and may have introduced bias or weakened interpretative strength. Furthermore, the reliance on secondary data from media outlets introduces the risk of inaccuracies or inconsistencies in how information was recorded and categorised.

Qualitative Methodological Limitations

Although the study was primarily quantitative, certain qualitative insights were drawn from news media narratives, especially in interpreting framing patterns and thematic elements. These qualitative components are inherently subjective and may be influenced by the researcher's interpretive lens. The absence of a formal qualitative coding protocol or inter-coder reliability measures reduces the transparency and reproducibility of these interpretations.

Future research employing structured qualitative content analysis or discourse analysis would be better equipped to offer rigorous insights into media framing.

Future Research Directions

To address these limitations, future research could expand the dataset to include a more diverse set of media outlets, languages, and timeframes. Additionally, incorporating official crime statistics or conducting audience reception studies could help triangulate findings. Qualitative analyses, including narrative or discourse analysis, may also enrich the understanding of media framing. Despite the limitations noted above, this study contributes a valuable foundation for ongoing exploration of media representation of crime in South Africa

CHAPTER 7: CONCLUSION

This research examined how murder incidents are reported in South African news media, identifying patterns in demographic, geographic, and contextual representation. By combining manual data collection with computational analysis, the research sheds light on biases and trends in media coverage, emphasising the selective nature of news reporting. Drawing on a multi-method approach that combined manual data extraction with computational analysis, including descriptive statistics, chi-square tests, and logistic regression models, this research explored both the frequency and framing of murder reports across a diverse set of news articles.

Summary of Findings

The analysis revealed multiple trends in how murder is reported in South African news media. Gun-related murders were most frequently represented, reflecting their high incidence in official crime statistics. However, other types of murder, such as IPV and strangulation, were more frequently highlighted in cases involving women, suggesting gendered media narratives, even if statistical significance was not always achieved. Urban-centric reporting dominated the dataset, with rural and peri-urban areas notably underrepresented. While visual inspection and raw percentages suggested differences in gender-based reporting patterns, statistical analyses, particularly logistic regression, did not find gender to be a significant predictor of high media coverage. Although gendered patterns may appear in the data, they were not supported by statistically significant results. Therefore, caution is warranted in asserting definitive gender-based disparities in coverage. Racial and locational disparities, however, were more clearly observed. Urban areas were more likely to receive high media coverage, and the framing of victims appeared to differ across racial categories. These trends highlight how historical and societal inequalities continue to shape media narratives in post-apartheid South Africa.

The Value of Computational Techniques

A key contribution of this research lies in its use of computational social science methods. Through structured coding, EDA and formal statistical modelling, the researcher was able to interrogate a large dataset efficiently and systematically. These techniques enabled prediction and the detection of patterns, such as the interaction between location type and coverage levels that might otherwise have gone unnoticed in a purely qualitative study. The use of computational methods offered advantages such as scalability and reproducibility

These methods underscore the potential for future criminology and media studies to take advantage of computational tools. They also bridge the gap between traditional social science and data science approaches, offering a path forward for interdisciplinary research.

Ultimately, this research contributes to a deeper understanding of crime reporting in South Africa, highlighting both the patterns of representation and the underlying mechanisms that drive them. The integration of computational methods in media research holds great promise for advancing the field. As media environments become more data-rich and complex, the ability to critically and systematically analyse large volumes of content will be key to ensuring accountability, fairness, and equity in representation.

REFERENCES

Abrahams, N., Jewkes, R., & Mathews, S. (2010). Guns and gender-based violence in South Africa. SAMJ. South African Medical Journal/South African Medical Journal, 100(9), 586. https://doi.org/10.7196/samj.3904

Abrahams, N., Mathews, S., Lombard, C., Martin, L. J., & Jewkes, R. (2017). Sexual murders in South Africa: A national cross-sectional epidemiological study of adult women and children. PloS One, 12(10), e0186432. https://doi.org/10.1371/journal.pone.0186432

Abrahams, N., Mathews, S., Martin, L. J., Lombard, C., Nannan, N., & Jewkes, R. (2016). Gender differences in murder of neonates, infants, and children under 5 y in South Africa: Results from the cross-sectional 2009 National Child Murder Study. PLoS medicine, 13, e1002003. https://doi.org/10.1371/journal.pmed.1002003

Agnew, R. (1992). Foundation for a General Strain Theory of Crime and Delinquency. Criminology, 30, 47-88. http://dx.doi.org/10.1111/j.1745-9125.1992.tb01093.x

Aldrete, M., & Fernández-Ardèvol, M. (2023). Framing femicide in the news, a paradoxical story: A comprehensive analysis of thematic and episodic frames. *[Journal Name]*, *20*(3). https://doi.org/10.1177/1741659023119977

Anitha, M., Savarimuthu, N., & Bhanu, S. M. S. (2025). Chi-square target encoding for categorical data representation: A real-world sensor data case study. *SN Computer Science, 6*(228).

Anand, A., & Taneja, M. (2024). *The Influence of Media on Public Perception of Crime*. https://www.ijfmr.com/papers/2024/2/15227.pdf

Arendt, F., Northup, T., Forrai, M., & Scheufele, D. (2023). Why we stopped listening to the other side: How partisan cues in news coverage undermine the deliberative foundations of democracy. *Journal of Communication, 73*(5), 413–426. https://doi.org/10.1093/joc/jqad007

Arije, O., Oyedun, O. O., & Omisile, I. (2021). Synthesizing qualitative research evidence from complex interventions by inductive thematic content analysis using modified in-vivo coding. Academia Letters. https://doi.org/10.20935/al2527

Auguie, B. (2017). gridExtra: Miscellaneous Functions for "Grid" Graphics. R package version .3.

Berns, N. (2017). Making victims: Victimization, media, and social problems (2nd ed.). Routledge.

Berry, D. M. (2012). Understanding Digital Humanities. Palgrave Macmillan.

Bivand, R., & Piras, G. (2015). Comparing Implementations of Estimation Methods for Spatial Econometrics. Journal of Statistical Software, 63(18), 1-36.

Boonzaier, F. & Jekwes, C. (2023). Spectacularising narratives on femicide in South Africa: A decolonial feminist analysis. 71, 78-96.

Bosma, A., Mulder, E., & Pemberton, A. (2018). The ideal victim through other(s) ' eyes. In Policy Press eBooks (pp. 27–42). https://doi.org/10.1332/policypress/9781447338765.003.0003

Boukes, M., Jones, N. P., & Vliegenthart, R. (2020). Newsworthiness and story prominence: How the presence of news factors relates to upfront position and length of news stories. Journalism, 23(1), 98–116. https://doi.org/10.1177/1464884919899313

Bowman, B., Kramer, S., Salau, S., & Matzopoulos, R. J. (2022). Trends, Correlates, and Contexts of Robbery-Murder in South Africa. 10887679211070230. https://doi.org/10.1177/10887679211070230

Bradford-Clarke, L., Davies, R., & Henley, A. (2022). When 'ideal victim' meets 'criminalised other': Criminal records and the denial of victimisation. Probation Journal, 69(3), 353–372. https://doi.org/10.1177/02645505221095068

Brodie, N. (2020). Femicide in South Africa. Kwela Books.

Brodie, N. (2021). Ideal victims and familiar strangers: Non-Intimate Femicide in South African news media. African Journalism Studies, 42(3), 82–99. https://doi.org/10.1080/23743670.2021.1933559

Brodie, N. (2022). How to read South Africa's quarterly crime statistics after the Covid pandemic. Africa Check. Available: https://africacheck.org/fact-checks/blog/how-read-south-africas-quarterly-crime-statistics-after-covid-pandemic [Accessed August 5, 2023].

Brodie, N. R. (2019). Using mixed-method approaches to provide new insights into media coverage of femicide. University of the Witwatersrand, Faculty of Humanities.

Brookman, F., Maguire, E. R., & Maguire, M. (Eds.). (2017). The handbook of homicide. Wiley-Blackwell.

Centre for the Study of Violence and Reconciliation (CSVR). (2008). A state of sexual tyranny: The prevalence, nature and causes of sexual violence in South Africa. Johannesburg: CSVR.

Chiumbu, S. H., & Moyo, D. (2022). "South Africa belongs to all who live in it." Communicare, 37(1), 136–152. https://doi.org/10.36615/jcsa.v37i1.1586

Christenson, D. (2014). Victim Worthiness: The Effect of Media Coverage on the Portrayal of Murder Victims. Follow this and additional works at: https://vc.bridgew.edu/undergrad_rev

Chowdhury, M. F. (2014). Coding, sorting and sifting of qualitative data analysis: Debates and discussion. *Quality & Quantity, 49*(3), 1135–1143. https://doi.org/10.1007/s11135-014-0039-2

Christians, C. G., Fackler, M., Richardson, K. B., Kreshel, P. J., & Woods, R. H. (2009). Media ethics: Cases and moral reasoning (9th ed.). Pearson.

Conway, M., & O'Connor, D. (2016). Social media, big data, and mental health: current advances and ethical implications. Current Opinion in Psychology, 9, 77–82. https://doi.org/10.1016/j.copsyc.2016.01.004

Cornelissen, L. A., Daly, L. I., Sinandile, Q., De Lange, H., & Barnett, R. J. (2019). A Computational Analysis of News Media Bias. Research Gate. https://doi.org/10.1145/3351108.3351134

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. University of Chicago Legal Forum, 1989(1), 139–167.

Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's *t*-test instead of Student's *t*-test. *International Review of Social Psychology, 30*(1), 92–101. https://doi.org/10.5334/irsp.82Download RStudio | The Popular Open-Source IDE from Posit. (2023, July 26). Posit. https://posit.co/products/open-source/rstudio/

Du Plessis, D. (2022). An analysis of crime reporting (and audience perceptions of it) in selected South African media. Communicare, 22(1), 169–186. https://doi.org/10.36615/jcsa.v22i1.1805

Fox, J., & Monette, G. (1992). Generalised collinearity diagnostics. *Journal of the American Statistical Association, 87*(417), 178–183. https://doi.org/10.2307/2290467

Fox, J., & Weisberg, S. (2018). *An R companion to applied regression* (3rd ed.). Sage Publications.

Gilchrist, K. (2010). “Newsworthy” victims? Feminist Media Studies, 10(4), 373–390. https://doi.org/10.1080/14680777.2010.514110

Gravel, J., Valasik, M., Mulder, J., & Tita, G. (2023). Rivalries, reputation, retaliation, and repetition: Testing plausible mechanisms for the contagion of violence between street gangs using relational event models. *Network Science*. https://doi.org/10.1017/nws.2023.8

Grolemund, G., & Wickham, H. (2011). Dates and Times Made Easy with lubridate. Journal of Statistical Software, 40(3), 1-25.

Groenewald, P., Bradshaw, D., Daniels, J., Matzopoulos, R., Bourne, D., & Naledi, T. (2008). Cause of death and premature mortality in Cape Town, South Africa, 2001–2006. South African Medical Journal, 98(9), 674–680.

Gruenewald, J., Chermak, S. M., & Pizarro, J. M. (2013). Covering victims in the news: What makes minority murders newsworthy? Justice Quarterly, 30(5), 755–783. https://doi.org/10.1080/07418825.2011.628945

Gqola, P. D. (2015). Rape: A South African Nightmare. Jacana Media.

Guetterman, T. C., & Fetters, M. D. (2018). Two methodological approaches to the integration of mixed methods and case study designs: A systematic review. *American Behavioural Scientist, 62*(7), 900–918. https://doi.org/10.1177/0002764218772641

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review, 31*(1), 2-24. https://doi.org/10.1108/EBR-11-2018-0203

Harcup, T., & O’Neill, D. (2017). What is news? News values revisited (again). Journalism Studies, 18(12), 1470-1488. https://doi.org/10.1080/1461670X.2016.1150193

Hlavac, M. (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables. R package version 5.2.2.

Hossain, A., Karimuzzaman, M., Hossain, M. M., & Rahman, A. (2021). Text mining and sentiment analysis of newspaper headlines. Information, 12(10), 414. https://doi.org/10.3390/info12100414

Huff-Corzine, L., McCutcheon, J. C., Corzine, J., Jarvis, J. P., Tetzlaff-Bemiller, M. J., Weller, M., & Landon, M. (2014). Shooting for accuracy: Comparing data sources on mass murder. Murder studies, 18, 105-124. https://doi.org/10.1177/1088767913512205

Hunt, S., & Jaworska, S. (2019). Intersections of nationality, gender, race and crime in news reporting: The case of Oscar Pistorius – Olympian and murderer. Discourse, Context & Media, 30, 100299. https://doi.org/10.1016/j.dcm.2019.100299

Jewkes, R., Sikweyiya, Y., Morrell, R., & Dunkle, K. (2011). Gender inequitable masculinity and sexual entitlement in rape perpetration South Africa: Findings of a cross-sectional study. PLoS ONE, 6(12), e29590.

Jewkes, Y. (2015). Media and crime. SAGE.

Kamiri, J., & Mariga, G. (2021). Research methods in machine learning: A content analysis. *International Journal of Computer and Information Technology, 10*(2). https://doi.org/10.24203/ijcit.v10i2.79

Jiang, M., Lee, S., & Brantingham, P. J. (2021). Challenges in qualitative annotation at scale: A case study of crime data. *arXiv*. https://arxiv.org/abs/2008.00977

Kaplan, R. M., et al. (2023). Ontologies in behavioural science: Accelerating research and the spread of knowledge. *Behavioural Sciences & Policy*. (Based on NASEM consensus findings)

Kappeler, V. E., & Potter, G. W. (2017). The mythology of crime and criminal justice. Waveland Press.

Kagawa, R. M. C., Cerdá, M., Rudolph, K. E., Pear, V. A., Keyes, K. M., & Wintemute, G. J. (2018). Firearm involvement in violent victimization and mental health: An observational study. *Annals of Internal Medicine, 169*(8). https://doi.org/10.7326/M18-0365

Knappertsbusch, F., Schreier, M., Burzan, N., & Fielding, N. (2023). Mixed methods and multimethod social research. *Forum: Qualitative Social Research / Forum: Qualitative Sozialforschung, 24*(1). https://www.qualitative-research.net/index.php/fqs/article/view/3866

Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. Journal of Statistical Software, 28(5), 1-26.

Kovach, B., & Rosenstiel, T. (2014). The elements of journalism: What newspeople should know and the public should expect (3rd ed.). Three Rivers Press.

Langa, M., Kirsten, A., Bowman, B., Eagle, G., & Kiguwa, P. (2018). Black masculinities on trial in absentia: The case of Oscar Pistorius in South Africa. Men And Masculinities,

23(3–4), 499–515. https://doi.org/10.1177/1097184x18762523

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. R News, 2(3), 18-22.

Libby, L. (2012). South Africa's Media 20 Years After Apartheid A Report to the Center for International Media Assistance. In The Center for International Media Assistance (CIMA). https://www.cima.ned.org/wp-content/uploads/2015/02/final_4.pdf

Lindegaard, M. R. (2017). Murder in South Africa. Research Gate, 499–514. https://doi.org/10.1002/9781118924501.ch28

Liska, A.E., Baccaglini, W. (1990). Feeling safe by comparison: Crime in the newspapers. *Social Problems*, 37(3), 306-374. doi. 10.2307/800748

Lloyd, L. (2013). South Africa's media 20 years after apartheid: A report to the Center for International Media Assistance. Center for International Media Assistance, National Endowment for Democracy.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2019). cluster: Cluster Analysis Basics and Extensions. R package version 2.1.0.

Maluleke, W., & Dlamini, S. (2019). The prevalence of organised cross-border crimes in South Africa: A non-empirical statistical data analysis on stock theft and hijacking of motor vehicles. International Journal of Social Sciences and Humanity Studies, 11(1), 116-145.

Mathews, S., Abrahams, N., Martin, L. J., Lombard, C., & Jewkes, R. (2019). Murder pattern among adolescents: A national epidemiological study of child murder in South Africa. Plos one, 14, e0221415.

Mathews, S., Jewkes, R., & Abrahams, N. (2014). 'So now I'm the Man':Intimate partner femicide and its interconnections with expressions of masculinities in South Africa. British Journal of Criminology, 55(1), 107–124. https://doi.org/10.1093/bjc/azu076

Mathews, S., Jewkes, R., & Abrahams, N. (2015). "I had a hard life": Explaining patterns of intimate partner violence in South Africa. Social Science & Medicine, 128, 142–150. https://doi.org/10.1016/j.socscimed.2015.01.023

Matzopoulos, R., Marineau, L., Mhlongo, S., Ketelo, A., Prinsloo, M., Dekel, B., Martin, L. J., Jewkes, R., Lombard, C., & Abrahams, N. (2024). Who is killing South African men? A retrospective descriptive study of forensic and police investigations into male murder. BMJ Global Health, 9(4), e014912. https://doi.org/10.1136/bmjgh-2023-014912

Matzopoulos, R. (2015). Mortality and the epidemiology of violence in South Africa. South African Medical Research Council.

Mawby, R. C. (2010). Police corporate communications, crime reporting and the shaping of

policing news. *Policing & Society, 20*(1), 124-139. https://doi.org/10.1080/10439461003611526

Mayr, A., & Machin, D. (2011). The language of crime and deviance: An introduction to critical linguistic analysis in media and popular culture. Bloomsbury Publishing.

McHugh, M. L. (2013). The Chi-square test of independence. *Biochemia Medica*, 23(2), 143-149. https://doi.org/10.11613/BM.2013.018

Mgogo, Q., & Osunkunle, O. (2021). Xenophobia in South Africa: An insight into the media representation and textual analysis. *Global Media Journal, 19*(230), 37.

Monyeki, P., Naicker, N., & Obagbuwa, I. C. (2020). Change-Point analysis: an effective technique for detecting abrupt change in the murder trends in a democratic South Africa. The Scientific World Journal, 2020, 1–10. https://doi.org/10.1155/2020/4158472

Morrell, R. (2001). The times of change: Men and masculinity in South Africa. In R. Morrell (Ed.), Changing men in Southern Africa (pp. 3–37). University of Natal Press.

Neuendorf, K. A. (2025). *The content analysis guidebook* (2nd ed.). SAGE Publications.

O'Connor, B. T. (2011). Be careful with dictionary-based text analysis. Social Science Webpage. Published Oct, 5, 2011.

O'Hear, M. (2020). Violent crime and media coverage in one city: A statistical snapshot. *Marquette Law Review, 103*(3), 1007. Retrieved from https://scholarship.law.marquette.edu/mulr/vol103/iss3/14

Palmer, P. B., & O'Connell, D. G. (2009). Regression analysis for prediction: understanding the process. Cardiopulmonary physical therapy journal, 20(3), 23–26.

Paulsen, D. J. (2003). Murder in Black and White. Murder Studies, 7(3), 289–317. https://doi.org/10.1177/1088767903253707

Petersen, N. (2014). Neighborhood context and media constructions of murder. Murder Studies, 20(1), 25–52. https://doi.org/10.1177/1088767914554616

Petersen, N. (2016). Neighborhood context and media constructions of murder: A multilevel analysis of murder newspaper coverage in Los Angeles County. Murder studies, 20, 25-52.

Peterson, S. K., Cooke, A. M., & Metcalf, S. S. (2021). Interrogating the "Murder Centre of South Africa": The Spatial Distribution of Murder Risk in Cape Town. Geographical Analysis, 53, 499-519. https://doi.org/10.1111/gean.12256

Phaswana, E. D. (2021). Women, gender, and race in post-apartheid South Africa. The Palgrave Handbook of African Women's Studies. Springer.

Posel, D., & Casale, D. (2019). Gender and the economy in post-apartheid South Africa: Changes and challenges. *Agenda, 33*(4), 3–10. https://doi.org/10.1080/10130950.2019.1679439

Pyrooz, D. C., Turanovic, J. J., & Wu, J. (2015). Taking stock of the relationship between gang membership and offending: A meta-analysis. *Criminal Justice and Behaviour, 43*(3). https://doi.org/10.1177/0093854815605528

Ratele, K. (2008). Masculinity and male mortality in South Africa. African Safety Promotion, 6(2), 19–41.

Republic of South Africa. (1996). Constitution of the Republic of South Africa, *1996.* Retrieved from https://www.gov.za/documents/constitution-republic-south-africa-1996

Rieder, B., & Röhle, T. (2014). Programmed Method: Developing a Toolset for Capturing and Analysing Tweets. https://doi.org/10.1108/AJIM-09-2013-0094

Riffe, D., Lacy, S., Watson, B. R., & Fico, F. (2019). Analysing media messages. In Routledge eBooks. https://doi.org/10.4324/9780429464287

Robinson, M. B. (2011). Media coverage of crime and criminal justice. Carolina Academic Press Durham, NC.

Robinson, D., & Hayes, A. (2020). broom: Convert Statistical Objects into Tidy Tibbles. R package version 0.7.0.

SAPS. (2021). Police Recorded Crime Statistics: Republic of South Africa. South Africa Police Service. Available: https://www.saps.gov.za/services/downloads/april_june_2021_22_quarter1_presentation.pdf [Accessed August 5, 2023].

SAPS. (2022). Police Recorded Crime Statistics; Republic of South Africa. South Africa Police Service. Available: https://www.saps.gov.za/services/downloads/Annual-Crime-2021_2022-web.pdf [Accessed August 5, 2023].

SAPS. (2023). 2022-2023 Q4 Crime Statistics. South Africa Police Services. Available: https://www.saps.gov.za/services/crimestats.php [Accessed August 5, 2023].

Schildkraut, J., Elsass, H. J., & Meredith, K. (2018). Mass shootings and the media: Why all events are not created equal. Journal of Crime Justice Quarterly, 41, 223-243. https://doi.org/10.1080/0735648X.2017.1284689.

Schreibman, S., Siemens, R., & Unsworth, J. (Eds.). (2008). A companion to digital humanities. Blackwell.

Seedat, M., Van Niekerk, A., Jewkes, R., Suffla, S., & Ratele, K. (2009). Violence and injuries in South Africa: Prioritising an agenda for prevention. The Lancet, 374(9694), 1011–1022. https://doi.org/10.1016/S0140-6736(09)60948-X

Seedat, M., Van Niekerk, A., Jewkes, R., Suffla, S., & Ratele, K. (2014). Violence and injuries in South Africa: Prioritising an agenda for prevention. The Lancet, 374(9694), 1011-1022

Shaw, M. (2002). Crime and policing in post-apartheid South Africa: Transforming under fire. Indiana University Press.

Shoemaker, P. J. (2006). News and newsworthiness: A commentary. Communications, 31(1), 105–111. https://doi.org/10.1515/commun.2006.007

Shoenbill, K. A., Kasturi, S. N., & Mendonca, E. A. (2023). Artificial Intelligence, Machine Learning, and Natural Language Processing. Chronic Illness Care, Query date: 2024-05-23 08:20:46, 469–479. https://doi.org/10.1007/978-3-031-29171-5_36

Sierra-Arévalo, M., & Papachristos, A. V. (2017). Social networks and gang violence reduction. *Annual Review of Law and Social Science, 13*(1). https://doi.org/10.1146/annurev-lawsocsci-110615-085017

Silge, J., & Robinson, D. (2016). tidytext: Text Mining and Analysis Using Tidy Data Principles in R. JOSS, 1(3).

Silge, J., & Robinson, D. J. (2017). Text Mining with R: A Tidy Approach. https://cds.cern.ch/record/2269044

Silva, C., & Guedes, I. (2022). The role of the media in the fear of Crime: A Qualitative study in the Portuguese context. Criminal Justice Review, 48(3), 300–317. https://doi.org/10.1177/07340168221088570

Spies, A. (2020). The portrayal of victims of intimate femicide in the South African media. Journal of African Media Studies, 12(1), 41–59. https://doi.org/10.1386/jams_00010_1

Statistics South Africa. (2022). *Experience of crime in SA increased over the 2021/22 period.* Retrieved from https://www.statssa.gov.za/?p=15700

Sutherland, G., McCormack, A., Pirkis, J., Vaughan, C., Dunne-Breen, M., Easteal, P., & Holland, K. (2016). Media representations of violence against women and their children.

https://anrows.intersearch.com.au/anrowsjspui/bitstream/1/19851/1/FINAL-Co-branded-Media-Representations_WEB.pdf

Surette, R. (2015). Media, crime, and criminal justice: Images, realities, and policies (5th ed.). Cengage Learning.

Snyman, C. R. (2014). *Criminal law* (6th ed.). LexisNexis.

Thomas, M. D., Jewell, N. P., & Allen, A. M. (2021). Black and unarmed: statistical interaction between age, perceived mental illness, and geographic region among males fatally shot by police using case-only design. Annals of epidemiology, 53, 42-49. e3.

Thomson, J. (2016). A murderous legacy: Coloured murder trends in South Africa. South African Crime Quarterly. https://doi.org/10.17159/2413-3108/2004/v0i7a1050.

United Nations Office on Drugs and Crime. (2019). Global study on homicide 2019 edition. https://www.unodc.org/unodc/en/data-and-analysis/global-study-on-homicide-2019.html

Walker, A. (2019). openxlsx: Read, Write and Edit xlsx Files. R package version 4.1.4.

Wasserman, H. The state of South African media: a space to contest democracy.

Publizistik 65, 451–465 (2020). https://doi.org/10.1007/s11616-020-00594-4

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L. B., Hester, J., Kühn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D. M., Seidel, D. P., Spinu, V., & Yutani, H. (2019). Welcome to the Tidyverse. Journal of Open Source Software, 4(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., & Grolemund, G. (2016). R for Data Science: import, tidy, transform, visualise, and model data. https://openlibrary.org/books/OL26836898M/R_for_Data_Science_Import_Tidy_Transform_Visualise_and_Model_Data

Wickham, H., & Bryan, J. (2019). readxl: Read Excel Files. R package version 1.3.1.

Wood, K., & Jewkes, R. (2001). "Dangerous" love: Reflections on violence among Xhosa township youth. In R. Morrell (Ed.), Changing men in Southern Africa (pp. 317–336). University of Natal Press.

Wood, K., Lambert, H., & Jewkes, R. (2007). "Showing roughness in a beautiful way": Talk about love, coercion, and rape in South African youth culture. Medical Anthropology Quarterly, 21(3), 277–300. https://doi.org/10.1525/maq.2007.21.3.277

Xie, Y. (2015). Dynamic Documents with R and knitr (2nd ed.). Chapman and Hall/CRC.

Zhu, H. (2019). kableExtra: Construct Complex Table with 'kable' and Pipe Syntax. R package version 1.1.0.