Text

Description automatically generated with medium confidence

**A**

**REFLECTIVE DIARY**

ON

**PROJECT 3: US CRIME ANALYSIS**

**SUBMITTED**

**IN PARTIAL FULFILLMENT TO THE**

**MODULE**

**BUSINESS INTELLIGENCE AND DATA VISUALIZATION**

**BY**

**OKOLI, CHUKWUEMERIE OKECHUKWU**

**21062209**

**MAY 2022**

**Table of Contents**

1. Introduction…………………………………………………..................
2. Literature Review of Related works………………………....................
3. Milestone 1
4. Data Analysis…………………………………………...............
5. Milestone 2
6. Business requirement and identify stakeholders
7. Milestone 3
8. Data visualization and answer to business questions……………
9. Milestone 4
10. Storytelling around the data……………………………………
11. Milestone 5
12. Recommendation for future improvement……………………
13. Challenges....................................................................................
14. Knowledge Achievement……………………………………………...
15. Reference…………………………………………………………………

**Introduction**

This report examines the growth of crime trends as an investigative tool in various states across the United States of America, as well as the need for crime analysis in the corporate world among potential partners and its approach. To demonstrate various crime rates, a visualisation tool for visual representation is employed, and lastly, my perspective assisted in the development of some recommendations that can contribute to enhancing the outcome of crime analysis among law enforcement officers. The goal of this project was to gather the crime analytics dataset, analyse it, and prepare a report that employs many visualisations to offer a specific information narrative by telling stories with the data. I was able to investigate the dataset using a visualization tool to outline some of the areas with chances of strong crime events and visualizing crime vulnerable areas. Crime data analyst can assist the law enforcement officers accelerate in solving of these crimes with the help of these modern technologies.

**16th Feb**

**Project Selection**

Today our team gathered to choose which of the three projects namely Guardian League Table Analysis, Large Scale North American Retailer Analysis, Crime Analytics that was allocated to us by the module leader in the course specification. After deliberations our team unanimously chose to analyse the US crime dataset because it examines the progression of the crime trends in various states across the United States of America. We liked to think that it would be an amazing adventure seeing us perform as an instrument in this research, trying to aid various departments that are interconnected to reduce these crimes by evaluating these crime patterns. The team after choosing the project agreed to use one of the leading data visualization tool Tableau to analyse the dataset.

**17th February**

**First Industry Advisor Meeting**

Today we had our first industry advisor meeting with Aman Azhar where he explained the dataset for us and how to go about it. He also lectured about the crime analysis and different types of crime analysis. I asked him if its possible to create a different dashboard for different stakeholders or just one dashboard to for different stakeholders when visualizing our work. He responded that we should create specific dashboards for specific needs, and each should have what they are all addressing or speaking to.

**18th February**

This is our first team meeting, and we learnt a little more about the topic. We realised that the dataset had two degrees of crimes: negligent manslaughter and murder or manslaughter. After that, we researched for materials to learn more about how this crime happens. Negligence of the highest order Manslaughter is a crime committed when a person in a position of authority conducts an illegal act that they should have known would result in the death of another person. (Enable Law, 2019). Murder and manslaughter are two of the offenses that constitute homicide (Homicide: Murder and Manslaughter | The Crown Prosecution Service, 2021). The murder of one person by another is known as homicide and it’s an expression that refers to both lawful and unlawful executions. While Murder is a type of homicide in which someone is killed without their consent, Manslaughter is a type of homicide in which someone is killed unintentionally. (Stein & Markus, 2019).

**20th February**

I looked over some more literatures today were (Wright, et al., 1981) gave a detailed review and analysis of the literature on a broad array of topics, including firearm ownership and use, as well as the relationship between weapons and crime and violence in the United States. The research looks at data on the number of guns in private hands, as well as current trends in weapon ownership. The public demand for guns as defensive weapons is addressed, as well as trends in the usage of guns for sport and enjoyment. An examination of police gun ownership is also presented. The characteristics of private firearm owners are described in great depth. In a chapter focusing on three hypotheses, the relationship between violent crime and the incidence of private weapon ownership is examined: (1) that weapons in private hands may be an important cause of criminal violence, (2) that this may be an important effect of criminal violence, and (3) that this may be an important deterrent to criminal violence. The data on the incidence of crime and violence, the characteristics of victims and offenders, the use of firearms in crime, factors of criminal motivation, and the ways of dealing with weapons offenders in the criminal justice system are all reviewed and analysed in this paper.

I also looked at some research in which (N James., 2018) looked at current crime trends in the United States. This research examines trends in violent crime rates since 1960, with an emphasis on trends in violent crime and murder rates in the major cities in the United States which the information was provided to the Federal Bureau of Investigation's Uniform Crime Reporting Program from 2014 to 2016. The statistics show that violent crime and murder rates went up from 2014 to 2015 then from 2015 to 2016, yet both remain around record lows. In addition, while violent crime and murder rates in urban areas have received a lot of attention recently, they also increased in smaller places from 2014 to 2015 and again from 2015 to 2016, albeit not as much as in the major cities. In some of the country's major cities, violent crime, murder, and both rates climbed from 2014 to 2015 and 2015 to 2016. In some of these places, violent crime and murder rates were at all-time highs in the last 3 decades.

I learned quite a bit from related research, and that helped me realise how important it is to work on this topic because it means that with appropriate evaluation, I might prevent future incidents.

**21st February**

**Data collection**

The crime data that was collected on crimes that occurred in various parts of the United States of America was used to meet the purpose of this reflective report.  From 1980 to 2014, the data includes above 638,455 crimes perpetrated in various states.

**Milestone 1 Week 4 of February**

**24th February 2nd Team meeting**

1. **Data Analysis**

This is the first stage in the process of analysing and comprehending the dataset; python programming was used to see if there were any discrepancies in the data, as well as to see if there were any null values or anomalies that needed to be resolved.

1. **Data Cleaning**: - I tried to check if the dataset if it needs cleaning. After running the code, I discovered that there was null value which shows the dataset is cleaned already and can proceed to be visualized.

**Fig 1. Screenshot of Jupyter notebook showing no missing values**

Graphical user interface, text, application

Description automatically generated

**4th March 3rd Team meeting**

In the data, I found far too many unknown values and related to my team. we saw that many of the data's value fields had "unknown parameters." Hence many unidentified values were recorded in the Dataset, indicating that those values contained information that had not been recorded, so utilising tableau to extract unknown significant information from unstructured data was necessary. I deleted all the records that included unknown values like the perpetrator's race and ethnicity.

**Fig 2. Graphical representation showing the high number of unknown values**

**Chart, bubble chart

Description automatically generated**

**9th March 4th Team Meeting**

I showed Dr Haixia a piece of my work on the crime pattern, and she offered a suggestion, which is that I should consider visualising it on map to show the crime pattern in various locations. But this showed differently from what I visualised, and she offered to review it and give me feedback in the next class.

**Fig 3 and 4. Screenshots showing the two different visualisations**

Graphical user interface, chart

Description automatically generated

Graphical user interface, application, website, map

Description automatically generated

**16th March**

Today I reminded the tutor about the unresolved problem and asked if she had come up with an explanation as to why the results were different, she gave me a good response and explained that there were parameters which needed to be put in place before I can get it right. she suggested adding longitude and latitude to the column and row and the crime type. The suggestion was helpful after I applied it.

**Fig 5. Screenshot showing the resolved crime pattern**

Map

Description automatically generated

**21st March 5th Team meeting and 2nd Industry Advisor meeting**

Today we had our 2nd meeting with our industry advisor Awan Azhar during which I asked question which I was confused about concerning the dataset. I asked him about the zeros (0) in the some of the columns like perpetrator count and victim count if we can remove it when doing calculations and if it will affect the result of the calculation, he replied with yes that I can remove it and won’t affect it but says it depends on the calculation I’m trying to do. This helped me going forward in the project.

**Fig 6. Screenshot showing the victim and perpetrator column.**

Graphical user interface, application, table, Excel

Description automatically generated

I also discovered that there were some columns which showed unknown in the dataset and was curious as to why it was like that, and he said we should make some research to ascertain why it is like that. I carried out some research and discovered that the reason was because some victims might refuse to disclose their identities, or the perpetrators of the crime could not be identified or maybe the weapon which was used to commit these crimes could not be found at the crime scene. (Chun, 2019) explained that modern life has it easier for serial killers to thrive.

**Fig 7. Screenshot showing the columns with unknown values**

**Graphical user interface, application, table, Excel

Description automatically generated**

**Milestone 2 Week 3 of March**

**30th March 6th Team meeting**

I had a discussion today with my tutor Haixia about the second milestone, from which I wanted a few explanations on what's in it and purpose of business settings, as well as spot potentially relevant parties who might be willing to participate in the data analysis. After thoroughly examining the data, I can conclude that the individuals who will be willing to participate in this analysis are the security forces, insurance firms, legal professionals, leaders of societies, government organisation and other various departments.  For the sake of this analysis, all these parties are part of this research, but my target group would be the security agency. The business world must be comprehended by the aforementioned key parties. It is critical to make appropriate and solid judgments as a police officer. This proper leadership will aid in decision-making and leadership success. It is important to use crime data to improve our knowledge and development of a response to criminal behaviour.

In the corporate world, crime statistics are crucial in determining our awareness of and response to criminal behaviour. Considering the many continuing discussions around its restrictions, security agencies depend on a grain of credibility in crime reports. To seek solutions to the increasing amount of crime, we must be imaginative in our representations, irrespective of how well the data is described. Public crime data give us a feeling of general community prosperity as well as an insight of general crime trends, allowing us to assess the level of security on the streets. protracted coordination is vital, and data act as a framework for developing strategies that are then investigated on. Countries would not be able to disperse their resources to their full potential without it, nor would they be able to discern between them.

The following are the points that I feel the security agencies requires responses to from the dataset.

1. Which month does this crime mostly happen?
2. Where did the majority of the crimes occur in the United States?
3. Which of the weapons have shown to be the most lethal in terms of its usage?
4. What kind of crimes were committed?
5. Is the country's rate of crime increasing or decreasing in a designated time?
6. What effect do the age ranges most frequently encountered in such incidents possess?
7. Patterns in crime incidence rates by level of relationship with the casualty?
8. What are perpetrators demographic attributes?
9. Have the victims involved in this incident been disclosed to the relevant authorities and how many of the crimes have been solved?

**4th April**

Prior to beginning milestone 3, I had to first deliberate on Dr. Paul Matthews' slides- Visual Perception, Cognition and Design, in which he tried to teach us the cognitive ethics and having regard to the need of effectively gasping the concept before commencing my visualisation. I try and make sure that my visualisation adheres to the criterion of validity, that states that we should convey somewhat details than is required by a viewer when designing our visualisation (Hegarty, 2011). In addition, there is the Capacity Restriction Concept. If this concept is applied correctly, visuals will be developed to give an explanation for limits imposed in working memory and cognitive, as compared to The Principle of Compatibility, which states that a graphic representation is easier to comprehend if its pattern is compliant with its definition.

**Milestone 3 Week 1 of April**

The goal of this milestone was to visualise the dataset while also attempting to answer business questions. From my perspective, the following are the types of questions that should be addressed from the dataset.

**Pertinent Business Question: Which month does this crime mostly happen?**Chart, bar chart

Description automatically generated

**Fig 8. Monthly Crime rate­­**

The visualisation above shows the highest period in the year when these crimes are committed mostly which is in July. I used the bar chart instead of the line chart because line chart defines unlimited time while bar chart shows a limited time frame.

**Pertinent Business Question: Where did the majority of the crimes occur in the United States?**

Map

Description automatically generated

**Fig 9. Map of USA**

The above map is the map of United States of America. The map depicted the region of the country with the highest number of criminal occurrences. The colour used was custom sequential which shows area with the hottest crime incidence and areas with least crime rates.

**Pertinent Business Question: Which of the weapons have shown to be the most lethal in terms of its usage?**

**Chart, bubble chart

Description automatically generated**

**Fig 10. Packed Bubbles displaying the percentage of weapons used**

Here, I used bubble chart to represent the weapons used. Bubble chart is commonly used to visualize relationships between three or more numeric variables. (Ricks, 2021). I got the inspiration to use this chart during one of the lectures when the tutor Dr Haixia lectured on ways to visualise a question and when to use it. The various shades of red here depict harm and danger. The text description shows the weapons usage by perpetrators. The thickest red bubble demonstrated that the weapon used mostly in perpetrating this crime was handgun. I attempted to use of precautionary characteristic regarding colour to consolidate the memory of the public.

**Persistent Business Question: What kind of crimes were committed?Chart, pie chart

Description automatically generated**

**Fig 11. Pie chart showing the kind of crimes committed.**

The above visualization depicts the type of crime committed. The red shaded portion shows the total number of crimes committed by murder or manslaughter while the peach shaded part shows the number of crimes perpetrated through manslaughter by negligence.

**Persistent Business Question: Is the country's rate of crime increasing or decreasing in a designated time?**

**Chart, line chart

Description automatically generated**

**Fig 12. Line graph representing the progression of the crime**

I used the line graph to visualise the business question because according to my lecturer Dr Haixia, line graph can be used to visualise a continuous trend. The above visualisation shows the period when the crime was at its peak and when it was at its lowest level. This will clearly show my audience how the crimes were progressing and regressing.

**Persistent Business Question: What effect do the age ranges most frequently encountered in such incidents possess?**

**Chart, bar chart

Description automatically generated**

**Fig 13. Bar chart showing the Age group involved in the crime**

Here I used bar chart to visualise the question as this will clearly show the age range of crime perpetrators and colour red was used to show the measure of danger and from the image, those in their twenties mostly commit this crime.

**Persistent Business Question: Patterns in crime incidence rates by level of relationship with the casualty?**

**Chart

Description automatically generated**

**Fig 14. Vertical bar chart illustrating the relationship count between perpetrator and victim**

In this visualization, I illustrated the relationship between the perpetrator and the victim. The idea of using orange colour to represent this visualisation was because inmates in US prison mostly wear it and if eventually these perpetrators are caught and behind bars, this is the colour they will be putting on.

**Persistent Business Question: What are perpetrators demographic attributes?**

**Chart, waterfall chart

Description automatically generated**

**fig 15. Screenshot showing Top States with solved and unsolved crimes.**

The visualisation above shows the percentage of the crime solved and those not yet solved. I had to filter out the top 15 states with the highest crime incident and the red colour shows the crimes not yet solved. This will give my audience clearer view if more effort needs to be put to mitigate these crimes or not.

**Persistent Business Question: Have the victims involved in this incident been disclosed to the relevant authorities and how many of the crimes have been solved?**

**Graphical user interface

Description automatically generated with medium confidence**

**Fig 16. Vertical Bar chart showing the Accomplishment of Various Agencies**

The performance of different agencies in solving crime and the number of criminal proceedings they get are depicted in the above bar chart. Among the other agency types, its proven that municipal police had the most recorded and solved crimes.

**Milestone 4 Week 2 of April**

**Chart

Description automatically generated**

**Fig 17. Progression of unsolved crimes**

**Chart

Description automatically generated**

**Fig 18. Progression of Crime with respect to Relationship and Age**

**6th April 7th Team meeting**

Today at school, I discovered a few of new ideas from Dr. Haixia's lectures. She suggested us to read the book storytelling with data and try to apply our findings to whatever visualisation we have accomplished with the data thus far. I checked out a lecture video from Pavan Lalwani on YouTube about storytelling which I found interesting because it gave me more insight about how to tell a story with data.

I also asked her a question regarding decluttering to get a better understanding of it and how to apply it in my work. She gave a detailed explanation of it by saying if we remove the gridlines from the background, then, we say we do declutter, and the reason is to increase data-ink ratio. The idea of data-ink ratio was founded by Edward Tufte where he explained that the data-ink ratio is the percentage of ink utilised to portray real data in comparison to the overall amount of ink used in the presentation. (Tufte, 1985), even though there were criticisms about it, I agreed with him because the background colour, gridlines are all unnecessary data ink.

Chart

Description automatically generated

**Fig 19. Screenshot of Dashboard showing decluttering**

I also asked her about having to remove or keep legends on the dashboard, she explained to me removing legend doesn’t count as declutter, that in interactive dashboard, we know the legend by hovering on the marks. So, removing legend is to save and give more space for visuals. This helped me going forward in my work.

A picture containing text, electronics

Description automatically generated

**Fig 20. Screenshot of dashboard showing the legend**

**7th April 8th Team meeting and Storytelling Deliberation**

I constructed a narrative on the data by telling a story about some unique elements that captured my interest surrounding the data after visualizing the questions. Increases in violent crime rates were seen in some of the country's major cities in 1993, accompanied by a reduction until 1999, then a spike in 2014, before crime figures stabilised in recent years. In high-crime areas like California, New York, Texas, and Florida, violent crime rates were at their peak in 34 years in several of these cities. My investigation proved that municipal police receive the most violent incidents, although the crime-solving record is good, one intriguing fact is that the majority of unsolved cases are perpetrated by black males, whereas crimes perpetrated by whites are solved. This implies that when it comes to addressing crimes, these bureaucracies are racially motivated.

I was able to find out that several weapons were used to commit each of these crimes, but the weapon used most in perpetrating these crimes is the handgun. Also, from data visualized, I was about to find out that those who perpetrates these violent incidents are mostly in their twenties which tells that there is something substantial and societally improper that occurs in this age range that law enforcement officials should investigate.

With this narrative, I'm pleased I was sufficient to convey meaning to my research by creating a story that explains the findings so that law enforcement agents may go deeper into it.

**25th April 3rd Industry Advisor Meeting**

Today we had our third industry advisor meeting Awan Azhar, and I had some questions which I had challenges in my work where I needed some clarifications to it. I needed to find out about criteria needed when trying to visualize states with highest crime rate. He outlined that to be able to do that, the victim count should be present, we can also filter the age, race, and weapons used in committing these crimes.

I tried it and this helped in making my visualization meaningful. Also, I asked him about the unknown in some of the columns like perpetrator age, ethnicity, and victim ethnicity if we can visualize it, and he responded that if you have significant number of unknowns in the column, its advisable to visualize the unknowns as this will show that there are reasons behind why they are left unknown.

A picture containing graphical user interface

Description automatically generated

**Fig 21. Barchart showing the Victim Ethnicity**

Chart

Description automatically generated

**Fig 22. Barchart showing the Perpetrator Ethnicity**

Chart, pie chart

Description automatically generated

**Fig 23. Piechart showing the Perpetrator Race**

**Milestone 5 Week 3 of April**

**9th April Recommendation for future improvements**

After performing experimental and analytical evaluation on my data, I arrived with some recommendations that could serve as an approach to help enhance the law enforcement agency's efforts. Below are some of the recommendations I have been able to put up.

* I recommend that the law enforcement agencies should review the license issuance process for firearm bearing. They should also outlaw the unauthorized use of handguns which accounted for 49.73 percent of all weapons used in these incidents.
* In order to promote better problem-solving, the law enforcement agencies should treat all crimes been perpetrated by these criminals with equality without being biased of ethnicity, race or gender.
* From the data, people in their twenties were found to be the most perpetrators of these crimes, I strongly recommend that the government engage this age group and create an empowerment programme which will keep them occupied.
* To minimize terrible fatalities such as homicide, I advocate a more proactive strategy to catching criminals. In high-crime areas like California, New York, Texas, the law enforcement authorities should concentrate more on providing specific projects to confront imminent threats.

**11th April Challenges and Knowledge Achievements**

While taking on this project I had some challenges as well as also gaining some new learnings.

**Challenges**

* I had the challenge of decluttering my dashboards while trying to find solution to the questions given to us
* Also, there were many unknown values in the dataset, and this got me frustrated in the first instance.
* Trying to find the right visualization chart to use in visualizing my work was also a challenge for me.
* How to choose the right color to use was a problem for me as I needed to put many people into consideration

**Knowledge Achievement**

* I was able to know when to use a particular type of chart to visualize my work as my lecturer Dr Haixia hinted in one of the lectures that we should be careful when using pie chart to visualize our work as it sometimes cannot convey the true meaning of what you trying to reveal. I also made sure I used it smartly in my work.
* I learnt the usefulness of font size. Using higher font size for numbers and lower font size for letter.

# References

Enable Law, 2019. *What Is Gross Negligence Manslaughter and How Can You Make A Claim?* [Online]

Available at: <https://www.enablelaw.com/news/expert-opinion/what-is-gross-negligence-manslaughter-and-how-can-you-make-a-claim/>

[Accessed 18 February 2022]

Chun, R., 2019. *Modern Life Has Made It Easier for Serial Killers to Thrive .* [Online]   
Available at: https://www.theatlantic.com/magazine/archive/2019/10/are-serial-killers-more-common-than-we-think/596647/  
[Accessed 21 March 2022].

Hegarty, M., 2011. *The cognitive science of visual-spatial displays: Implications for design, Topics in Cognitive Science,.* [Online]   
Available at: https://onlinelibrary.wiley.com/doi/10.1111/j.1756-8765.2011.01150.x  
[Accessed 2 April 2022].

Markus & Stein, 2019. *What’s the Difference Between Homicide, Murder, and Manslaughter?* [Online]   
Available at: https://www.steinandmarkuslaw.com/whats-the-difference-between-homicide-murder-and-manslaughter/  
[Accessed 18 February 2022].

Ricks, E., 2021. *Storytelling with data.* [Online]   
Available at: https://www.storytellingwithdata.com/blog/2021/5/11/what-is-a-bubble-chart  
[Accessed 8 April 2022].

The Crown Prosecution Service., 2021. *Homicide: Murder and Manslaughter.* [Online]   
Available at: https://www.cps.gov.uk/legal-guidance/homicide-murder-and-manslaughter  
[Accessed 18 February 2022].

Tufte, E., 1985. *The Visual Dpspaly of Quantitative Information*.

Wright, J. D., Rossi, P. H., Daly, K. & Weber-Burdin, E., 1981. *Weapons, Crime, and Violence in America - A Literature Review and Research Agenda.* [Online]   
Available at: https://www.ojp.gov/ncjrs/virtual-library/abstracts/weapons-crime-and-violence-america-literature-review-and-research  
[Accessed 20 February 2022]