

# Exploring Crime Trends in Washington D.C.: A Tableau Data Analysis



A project authored by  
Waqar Shaozab

25 August, 2023

## Introduction

In an increasingly complex urban landscape, understanding crime patterns is pivotal to enhancing public safety and shaping effective law enforcement policies. My data analyst project, "Exploring Crime Trends in Washington D.C.," delves into the intricate web of crime incidents within the nation's capital. Leveraging the power of Tableau, this project endeavors to provide a comprehensive visual narrative that uncovers underlying insights, temporal dynamics, and spatial correlations inherent in the crime landscape of Washington D.C.



National Mall - Washington D.C.

By harnessing a rich dataset encompassing crime incidents over a multi-year period, we embark on a data-driven journey to uncover trends, patterns, and potential drivers behind crime occurrences. The tableau visualizations I present illuminate the relationship between crime rates and various factors, shedding light on the intersections between societal dynamics and criminal activities.

Crime rates have experienced a noteworthy surge, with data from the Metropolitan Police Department (MPD) indicating a substantial **30 percent** escalation compared to the previous year. This rise is characterized by a **38 percent** upswing in violent crimes and a **29 percent** elevation in non-violent offenses.

This project is not merely about data; it is a proactive effort to empower stakeholders,

policymakers, and law enforcement agencies with actionable insights. Our tableau-based analysis opens avenues for evidence-based decision-making, allowing for targeted interventions, resource allocation, and the development of strategic crime prevention strategies.

Through a meticulous amalgamation of data exploration, visual representation, and informed interpretation, "Exploring Crime Trends in Washington D.C." seeks to foster a more informed understanding of the crime landscape, ultimately contributing to safer communities and a more secure urban environment. Join me as I unveil the hidden stories behind the data and unlock the potential for positive change through data-driven analysis.

# Table of Contents

<b>1</b>	<b>ASK</b>	<b>6</b>
1.1	Key Tasks . . . . .	6
1.2	Deliverable . . . . .	7
1.3	Research Task . . . . .	7
1.4	Research Objectives . . . . .	7
1.5	Project Deliverables . . . . .	7
1.6	Key Stakeholders . . . . .	7
<b>2</b>	<b>PREPARE</b>	<b>8</b>
2.1	Key Tasks . . . . .	8
2.2	Deliverable . . . . .	8
2.3	Information on Data Source . . . . .	9
2.4	Limitations of Data . . . . .	9
2.5	Is The Data ROCCC? . . . . .	9
<b>3</b>	<b>PROCESS/ANALYZE</b>	<b>11</b>
3.1	Key Tasks . . . . .	11
3.2	Deliverable . . . . .	11
3.3	Removed Duplicates . . . . .	12
3.4	Checked Spelling . . . . .	12
3.5	Searched for Null Values . . . . .	12
3.6	Deleted Column Called <i>X</i> and <i>Y</i> . . . . .	12
3.7	Deleted Column Called <i>OCTO_RECORD_ID</i> . . . . .	13

3.8	Ensured <i>REPORT_DAT</i> , <i>START_DATE</i> , and <i>END_DATE</i> Columns Had Consistent Timestamp Format . . . . .	13
3.9	Used <i>=TRIM()</i> Function . . . . .	13
3.10	Used <i>=LEFT()</i> and <i>=LEN()</i> Functions . . . . .	13
3.11	Created New Column called <i>CRIME_LENGTH</i> . . . . .	13
3.12	Created New Column Called <i>DAY_OF_WEEK</i> . . . . .	14
3.13	Created New Column Called <i>INCIDENT_PER_WARD</i> . . . . .	14
3.14	Created New Column Called <i>INCIDENTS_BY_DAY</i> . . . . .	14
3.15	Created New Column Called <i>TOTAL_INCIDENTS_PER_YEAR</i> . . . . .	14
3.16	Created New Column Called <i>TOTAL_POP_YEAR</i> . . . . .	14
3.17	Created New Column Called <i>CRIME_RATE</i> . . . . .	15
3.18	Created New Column Called <i>YEAR_OF_INCIDENTS</i> . . . . .	15
<b>4</b>	<b>SHARE - Visualization 1: Crime Distribution Map</b>	<b>16</b>
4.1	Key Tasks . . . . .	18
4.2	Deliverable . . . . .	18
4.3	Background/Key Findings . . . . .	18
4.4	Hypothesis . . . . .	19
<b>5</b>	<b>SHARE - Visualization 2: Crime Trend Over Time</b>	<b>20</b>
5.1	Key Tasks . . . . .	22
5.2	Deliverable . . . . .	22
5.3	Background/Key Findings . . . . .	22
5.4	Hypothesis . . . . .	23
<b>6</b>	<b>SHARE - Visualization 3: Crime Type Comparison</b>	<b>24</b>
6.1	Key Tasks . . . . .	26
6.2	Deliverable . . . . .	26
6.3	Background/Key Findings . . . . .	26
6.4	Hypothesis . . . . .	27
<b>7</b>	<b>SHARE - Visualization 4: Crime Heat-map</b>	<b>29</b>
7.1	Key Tasks . . . . .	31

7.2	Deliverable . . . . .	31
7.3	Background/Key Findings . . . . .	31
7.4	Hypothesis . . . . .	32
<b>8</b>	<b>SHARE - Visualization 5: Day of Week Comparison</b>	<b>33</b>
8.1	Key Tasks . . . . .	35
8.2	Deliverable . . . . .	35
8.3	Background/Key Findings . . . . .	35
8.4	Hypothesis . . . . .	36
<b>9</b>	<b>SHARE - Visualization 6: Crime Length</b>	<b>37</b>
9.1	Key Tasks . . . . .	39
9.2	Deliverable . . . . .	39
9.3	Background/Key Findings . . . . .	39
9.4	Hypothesis . . . . .	40
<b>10</b>	<b>SHARE - Visualization 7: D.C. Population</b>	<b>41</b>
10.1	Key Tasks . . . . .	43
10.2	Deliverable . . . . .	43
10.3	Background/Key Findings . . . . .	43
10.4	Hypothesis . . . . .	44

# Chapter 1

## ASK



Posing pertinent inquiries is of utmost importance during the initial stages of data analysis. When your questions align with the criteria of being **Specific**, **Measurable**, **Action-Oriented**, **Relevant**, **Time-Bound**—essentially conforming to the **S.M.A.R.T.** framework—you can be assured that the resulting answers will exhibit a comprehensive level of precision and accuracy. This meticulous approach invariably paves the way for a comprehensive and actionable analytical outcome.

### 1.1 Key Tasks

- Identify the research task
- Consider key stakeholders

## 1.2 Deliverable

- ☐ Clear statement of the research task

## 1.3 Research Task

Identify relationships and trends that show how crime changed in D.C.

## 1.4 Research Objectives

- Create a geographic visualization (map) displaying crime incidents.
- Visualize the overall crime trend over a specific time period
- Use bar charts or stacked area charts to compare different crime types
- Generate a heat-map to visualize crime density across different areas

## 1.5 Project Deliverables

1. A clear statement of the research task.
2. A description of all data sources used.
3. Documentation of any cleaning or manipulation of data.
4. A summary of analysis.
5. Supporting visualizations and key findings.

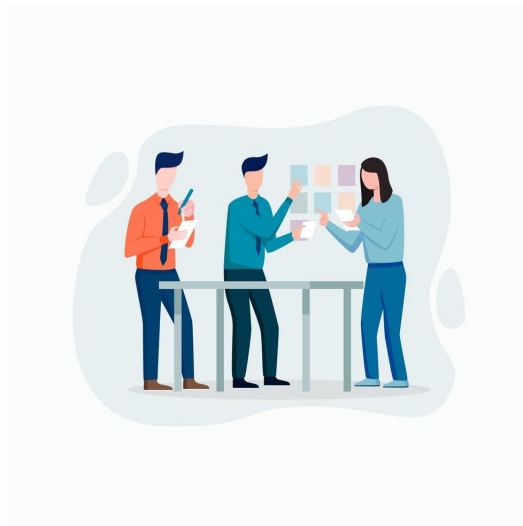
## 1.6 Key Stakeholders

- Current D.C. Residents
- Prospective D.C Residents
- Policymakers
- Law Enforcement Agencies



## Chapter 2

# PREPARE



### 2.1 Key Tasks

- Download data and store it appropriately
- Identify how it's organized
- Sort and filter data

### 2.2 Deliverable

- Description of all data sources used.

## 2.3 Information on Data Source

- The data is stored **here** via 5 datasets saved as **.CSV (comma-separated values)** files ranging from years **2018-2022**.
- The datasets are publicly available for this case-study by Creative Commons under this **license**.
- The original attributes that were captured include: *X, Y, CCN, REPORT\_DAT, SHIFT, METHOD, OFFENSE, BLOCK, XBLOCK, YBLOCK, WARD, ANC, DISTRICT, PSA, NEIGHBORHOOD\_CLUSTER, BLOCK\_GROUP, CENSUS\_TRACT, VOTING\_PRECINCT, LATITUDE, LONGITUDE, BID, START\_DATE, END\_DATE, OBJECTID, OCTO\_RECORD\_ID* with a total of 25 columns.
- Housed the original files in a separate folder.
- Created sub-folders for .CSV files and .XLSX files with appropriate naming conventions.

## 2.4 Limitations of Data

- One limitation would be that all the data are incidents reported and just that. We only have what people had reported leaving out anything that went under reported.
- Of the historical crime data, we are only using datasets that correspond to the last **5 years**.

## 2.5 Is The Data ROCCC?

A credible data source is **ROCCC** which stands for **R**eliable, **O**riginal, **C**omprehensive, **C**urrent, and **C**ited.

1. **Reliable** - YES - Large sample size; representative of population.
2. **Original** - YES - Primary source of data comes internally from opendatadc.

3. **Comprehensive** - MED - Data can be easily deciphered; does contain null values but not enough to lead to an incomplete analysis.
4. **Current** - YES - The most relevant data is available; entirety of the 2018-2022 calendar years.
5. **Cited** - YES - Sources comes directly from internal databases; no third-party was involved.

## Chapter 3

# PROCESS/ANALYZE



### 3.1 Key Tasks

- Check the data for errors
- Choose your tools and format your data
- Transform the data so you can work with it effectively
- Document the cleaning process and perform calculations

### 3.2 Deliverable

- ☐ Documentation of any cleaning and manipulation of data

The phase of processing stands as a pivotal component in the realm of high-level decision-making. Within this phase lies the indispensable task of refining and reshaping

data, transforming it from a state of disorder to one of pristine clarity—a prerequisite for subsequent analysis. It is crucial to recognize that the degree of cleanliness exhibited by your data corresponds directly to the quality and efficacy of your ensuing analytical endeavors. Data beset by inaccuracies or inconsistencies can undeniably cast a shadow of doubt upon the reliability of the insights derived.

Within the toolkit of data processing, an array of tools stands at one's disposal, including the versatile trio of spreadsheets, SQL, and R. In the context of our ongoing project, our journey commences with the meticulous refinement and manipulation of data utilizing the prowess of **Microsoft Excel**. Subsequently, the spotlight turns towards **Tableau**, selected for its remarkable capacity to craft visually captivating and dynamic representations, epitomized by its adeptness in forging intricate visual narratives like interactive dashboards.

Below will outline data cleaning process in Excel:

### 3.3 Removed Duplicates

1. *Data → Data Tools → Remove Duplicates → Select Columns → Ok*

### 3.4 Checked Spelling

1. *Select Column → Review → Proofing → Spelling*

### 3.5 Searched for Null Values

1. *Home → SelectSheet → Editing → Find and Select → GoTo Special → Blanks → Highlight Blanks Yellow*

### 3.6 Deleted Column Called *X* and *Y*

1. This Data is the Same as *LONGITUDE* and *LATITUDE*
2. *Select Column → Delete*

### 3.7 Deleted Column Called *OCTO\_RECORD\_ID*

1. This Column Contained No Data
2. *Select Column → Delete*

### 3.8 Ensured *REPORT\_DAT*, *START\_DATE*, and *END\_DATE* Columns Had Consistent Timestamp Format

1. *Home → Number → Format Cells → Time/Date → dd - mm - yyyyhh : mm : ss → Ok*

### 3.9 Used *=TRIM()* Function

1. This Is To Eliminate Extra Spaces
2. *Insert Column → = TRIM(A2, B3 etc...) → Drag Values Down → Copy New Column → Use Paste Special on Old Column → Values → Ok → Delete New Column*

### 3.10 Used *=LEFT()* and *=LEN()* Functions

1. This Is To Delete The Trailing Zeros (+00) in DateTime Columns
2. *Right Click Column → Insert → = LEFT(E2, LEN(E2) - 3) → Drag Values Down → Copy New Column → Use Paste Special on Old Column → Values → Ok → Delete New Column*

### 3.11 Created New Column called *CRIME\_LENGTH*

1. Calculated Length of Each Crime by Subtracting *START\_DATE* From the *END\_DATE* Columns
2. *=(W2-V2)*
3. Formatted Time as HH:MM:SS (37:30:55)
4. *Home → Number → Format Cells → Time → HH : MM : SS(37 : 30 : 55)*

### 3.12 Created New Column Called *DAY\_OF\_WEEK*

1. Calculated Day of Week Each Crime Was Committed by Using the *=WEEKDAY* Command.
2. *=WEEKDAY(B2,1)*
3. Formatted as Number With No Decimals
4. *Home* → *Number* → *Format Cells* → *Number* → *Decimal Places* : 01 = *Sunday*, 7 = *Saturday*

### 3.13 Created New Column Called *INCIDENT\_PER\_WARD*

1. Calculated How Many Incidents Had Occurred in Each Ward Using the *=COUNTIF()* Function
2. *Right Click Column* → *Insert* → *=COUNTIF(L:L,"1")* → *Repeat For All Wards* (2 – 7)

### 3.14 Created New Column Called *INCIDENTS\_BY\_DAY*

1. Calculated the Frequency of Each Crime Committed by Day of Week Using the *=COUNTIF()* Function
2. *=COUNTIF(Z:Z,"1")* → *Repeat For Each Day of Week* (2 – 7)

### 3.15 Created New Column Called *TOTAL\_INCIDENTS\_PER\_YEAR*

1. Determined How Many Incidents Had Occurred Each Year Using the *=COUNT()* Function
2. *=COUNT(Z:Z)* → *Repeat For Each Year*

### 3.16 Created New Column Called *TOTAL\_POP\_YEAR*

1. Input the Total Population of D.C. by Year.

2. Data was obtained through the U.S. Census Bureau

### 3.17 Created New Column Called *CRIME\_RATE*

1. Calculated Crime Rate by Dividing *TOTAL\_POP\_YEAR* by *TOTAL\_INCIDENTS\_PER\_YEAR*
2. *Create New Column*  $\rightarrow = (AA2/AB2) \rightarrow$  Repeat For Each Year

### 3.18 Created New Column Called *YEAR\_OF\_INCIDENTS*

1. Input Years 2018-2022

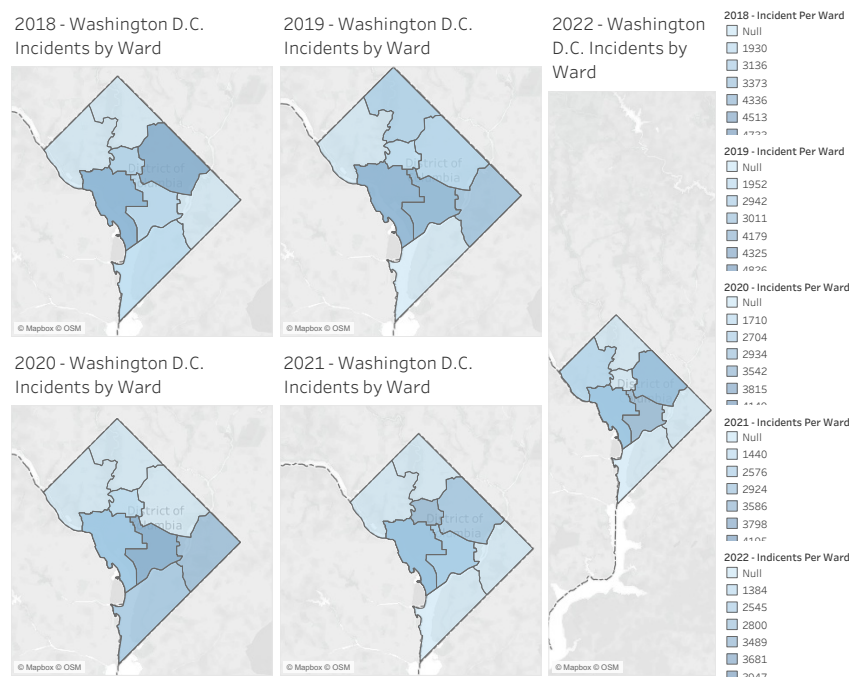
Subsequently, I consolidated the five datasets into a singular Excel document named *Crime\_Incidents\_2018\_to\_2022.xlsx*. The forthcoming step entails employing **Tableau** to craft visualizations based on the manipulated data.





## Chapter 4

# SHARE - Visualization 1: Crime Distribution Map



A link to the interactive dashboard can be found **here**. For a more specific view, you can access each sheet by clicking sheets 7-11.

## 4.1 Key Tasks

1. Determine the best way to share your findings
2. Create effective data visualizations
3. Present your findings
4. Ensure your work is accessible

## 4.2 Deliverable

- ☐ Supporting visualizations and key findings

## 4.3 Background/Key Findings

The objective of this graphic is to craft a geographical visualization (map) showcasing instances of criminal activity. The dashboard I designed serves to accentuate regions with high crime rates or clusters, as well as the distribution of these crimes across the city, categorized by year and segmented by ward. Upon glancing at the legends, one can observe that the color gradients are predicated on the frequency of incidents. Deeper shades of blue correspond to higher incident counts, while lighter shades indicate lower occurrences. These maps reveal two compelling trends. Firstly, a decline in overall incidents becomes apparent, and secondly, a noticeable decrease in incidents emerges from 2020 onwards, particularly in wards 7-8, which encompass Southwest (SW) and Southeast (SE) D.C.

The dashboard has facilitated the discernment of several noteworthy insights. The drop in reported incidents beginning in 2020 can potentially be attributed to the George Floyd protests of June 2020, during which a surge in mistrust towards law enforcement was witnessed. This shift may have led to fewer individuals reaching out to the police in situations that might warrant their involvement. This trend appears to have persisted into 2022. Additionally, a consistent pattern emerges in which wards 7-8 consistently

report fewer incidents compared to other wards, such as wards 3–4, which are notably affluent and situated deep within Northwest (NW). This phenomenon can be traced back to historical factors, such as long-standing community distrust towards law enforcement, spanning multiple generations. This pattern seems to become even more pronounced after 2020. The central area of the city encompasses wards 1, 2, 5, and 6, encompassing a blend of economically disadvantaged and affluent communities. This juxtaposition might contribute to the consistent reporting of incidents in this region.

## 4.4 Hypothesis

Prior to commencing my analysis, I formulated a hypothesis that projected a substantial surge in reported incidents within wards 7–8, particularly post-2020. Conversely, I anticipated a minimal occurrence in wards 3–4, with the city center falling within a moderate range of reported cases. This conjecture was primarily derived from an amalgamation of anecdotes, news coverage, and the prevailing sentiment associated with that particular city area.

Upon conducting the analysis, it became evident that I had not accounted for the potential dearth of reported incidents in this specific region. My initial assumption of an elevated number of reports was indeed inaccurate, prompting a realization of the broader societal influences that can inadvertently distort data collection processes – in this instance, crime reporting.

This experience underscores the significance of comprehending the intricate societal dynamics that can significantly impact data patterns, serving as an enlightening lesson on the complex interplay between perception and empirical evidence.

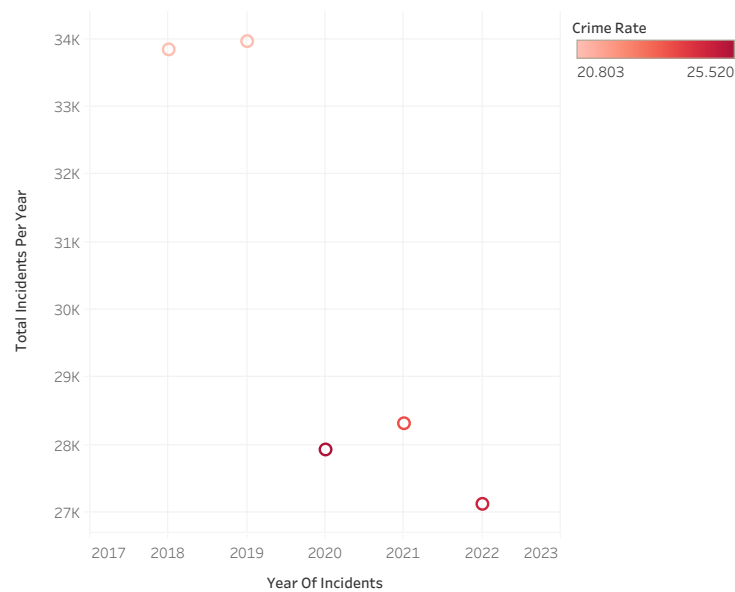
However, a pertinent question arises in light of the reduced incident reports: How can the crime rate be purportedly on the rise? The prevailing narrative repeatedly highlights an increase in crime, which seems incongruent with the lower number of reported incidents. Allow me to offer a brief explanation.



## Chapter 5

# SHARE - Visualization 2: Crime Trend Over Time

Total Incidents Per Year vs Year of Incidence



A link to the interactive sheet can be found **here**

## 5.1 Key Tasks

1. Determine the best way to share your findings.
2. Create effective data visualizations.
3. Present your findings.
4. Ensure your work is accessible.

## 5.2 Deliverable

- ☐ Supporting visualizations and key findings

## 5.3 Background/Key Findings

The concept behind this graphical representation aims to construct a line chart or area chart that effectively portrays the overarching trajectory of crime over a specified time-frame, spanning from 2018 to 2022. Within this context, our focus is directed towards gauging the crime trend over the mentioned period, comparing the total incidents per year against their respective years of occurrence.

By employing this approach, we facilitate an intuitive and comprehensible visualization of the crime trend. This graphical depiction readily unveils a distinct and notable drop in reported incidents, particularly from the year 2020 onward. A closer examination of the plotted data points reveals a discernible transition in hue, where the shade of red transforms into a deeper and richer tone between the years 2020 and 2022. This transformative shift is attributed to the incorporation of a filtering mechanism, which effectively portrays the crime rate for each individual year. In this context, the intensification of red hue signifies an escalated crime rate.

This visualization unfolds two intriguing insights. Firstly, the conspicuous decline in reported incidents becomes unmistakably apparent within the 2020-2022 time-frame. Paradoxically, despite the dwindling number of reported incidents during this period, the

crime rate exhibited a surge. This intriguing observation underscores the reality that even in instances of decreased incident reporting to law enforcement, a discernible elevation in the crime rate persists.

## 5.4 Hypothesis

Prior to crafting this visualization, I formulated a hypothesis grounded in the notion that a decline in incident reporting to law enforcement would paradoxically coincide with a rise in overall crime. This conjecture stemmed solely from online news reports. Gratifyingly, my theory found validation as it harmonized seamlessly with corroborative accounts of crime trends in D.C. This alignment of my analysis with independent crime reports lends substantial credence to the accuracy of my assessment.

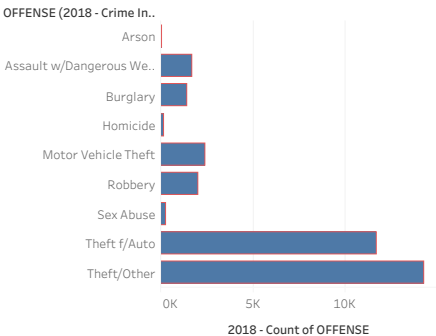




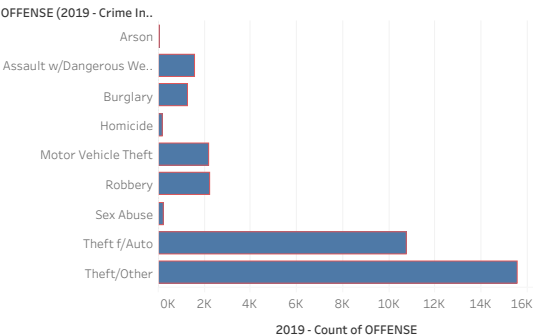
# Chapter 6

## SHARE - Visualization 3: Crime Type Comparison

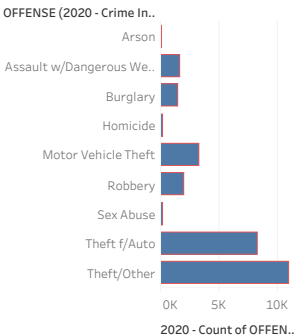
2018 - Frequency of Offense



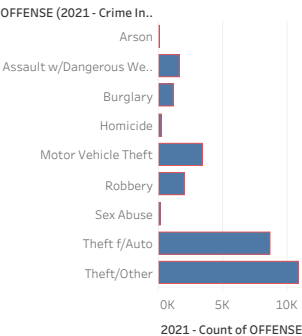
2019 - Frequency of Offense



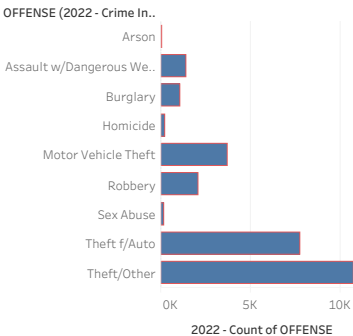
2020 - Frequency of Offense



2021 - Frequency of Offense



2022 - Frequency of Offense



A link to the interactive dashboard can be found **here**. For a more specific view, you can access each sheet by clicking sheets 7-11.

## 6.1 Key Tasks

1. Determine the best way to share your findings.
2. Create effective data visualizations.
3. Present your findings.
4. Ensure your work is accessible.

## 6.2 Deliverable

- ☐ Supporting visualizations and key findings

## 6.3 Background/Key Findings

The primary objective behind this visualization is to employ bar charts for the purpose of comparing distinct categories of crimes. The aim is to underscore the prevalence or proportion of specific criminal activities, facilitating the identification of both the most and least frequent types of crimes. To provide a temporal perspective, the data is partitioned by year, affording the opportunity to observe fluctuations across time.

This dashboard yields compelling insights that are readily interpretive. Across each year, a consistent pattern emerges, with theft and auto theft registering as the predominant crimes occurring in D.C. Following these, instances of robbery, assault with a dangerous weapon, and burglary are notable. In contrast, the least frequently occurring crimes encompass arson, sex abuse, and homicide.

One particularly salient trend that stands out pertains to the surge in motor vehicle theft, commencing in 2020. This observation becomes intriguing when juxtaposed with earlier data visualizations, which highlighted a substantial reduction in reported incidents from 2020 onward, in comparison to the preceding years of 2018-2019. This implies that

despite the diminished number of reported incidents, a notable increase in this specific crime statistic persists.

Furthermore, a discernible insight is gleaned from the constancy of reported theft occurrences across each successive year. This persistence underscores the enduring nature of this specific crime category, which remains a consistent concern in D.C., even amid the decline in police reports.

Another discernible trend that caught my attention involves the incremental rise in homicides on an annual basis. This observed pattern closely aligns with corroborative reports indicating an ongoing surge in homicides, particularly within the year 2023 (Hermann, 2023).

In sum, this visualization substantiates the multifaceted nature of crime trends, offering insights into both overarching patterns and counter intuitive variations, underscoring the intricate dynamics that shape the criminal landscape.

## 6.4 Hypothesis

Prior to crafting this data visualization, I found myself in a realm of uncertainty, unsure of the patterns that might unfold. Earlier graphics had illuminated the notion that low incident reporting might not unveil the comprehensive narrative we seek. Notwithstanding this, I ventured to formulate hypotheses, anchored in the anticipation that theft-related crimes would dominate the landscape, and Motor Vehicle Theft would emerge as a focal point. This supposition was rooted in the incessant reports from local news and social media platforms, underscoring the pressing nature of this offense.

Simultaneously, I harbored the expectation that Homicides would surge, commencing from 2020. This projection was predicated on analogous narratives stemming from news outlets and social media sources, shedding light on an alarming escalation in stabbings and shootings.

Following the unveiling of the crime comparison graphs, a sense of validation sweeps over my hypotheses. The data affirmatively indicates that theft holds sway as the most pervasive criminal activity. Additionally, the projection pertaining to Homicide trends finds resonance in the visualized data, revealing a consistent uptick in homicides year

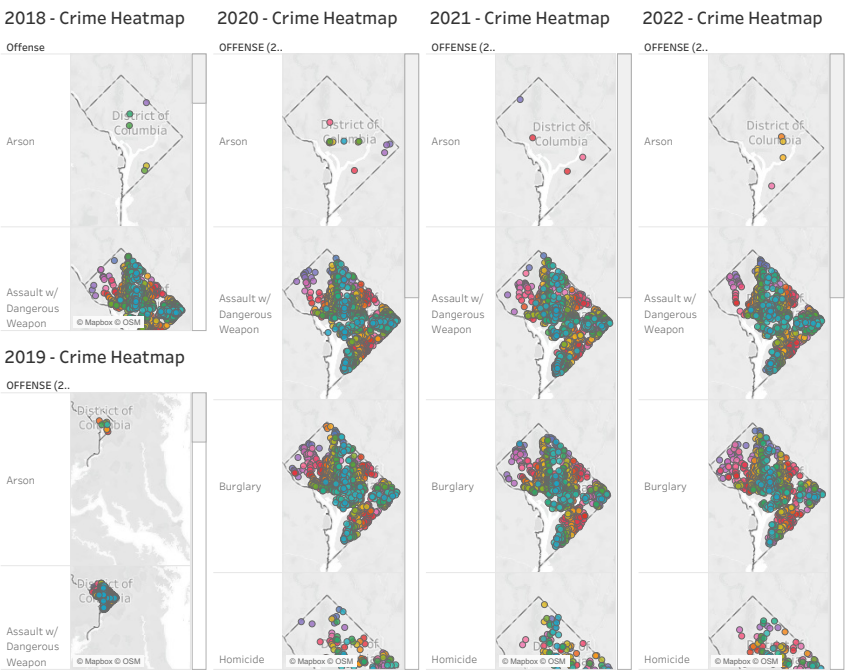
after year.

A thought-provoking avenue for exploration arises, one that entails comparing these contemporary yearly trends against historical data spanning over two decades. Such a retrospective analysis could yield fascinating insights into the evolving nature of crime dynamics and their enduring impact on society.



# Chapter 7

## SHARE - Visualization 4: Crime Heat-map



A link to the interactive dashboard can be found **here**. For a more specific view, you can access each sheet by clicking sheets 2-6.

## 7.1 Key Tasks

1. Determine the best way to share your findings.
2. Create effective data visualizations.
3. Present your findings.
4. Ensure your work is accessible.

## 7.2 Deliverable

- ☐ Supporting visualizations and key findings

## 7.3 Background/Key Findings

The fundamental objective driving this visualization is the creation of a comprehensive heat map, meticulously delineating the density of criminal activities across diverse urban areas. Employing a spectrum of color gradients, the visualization masterfully renders the spatial distribution of crime incidents. A technical limitation emerged during the implementation phase, as Tableau grappled with accommodating each individual incident within a legend capable of encompassing the entirety of incident types. Nonetheless, this challenge was navigated with innovative solutions, culminating in the identification of localized high and low crime density zones, strategically poised for targeted analysis.

Within this dynamic dashboard, the geographical distribution of each crime unfolds vividly across the cityscape, categorized according to crime type. This spatial alignment of crime types on a map provides an exceptional geographical dimension that augments the visualization's impact. Embedded within this heat map are a multitude of insights, though for the sake of brevity, we shall spotlight the salient patterns.

Drawing our focus to the crime type of Homicides, a fascinating revelation materializes. Despite the lower count of incident reports, the Homicide occurrences cluster conspicuously



toward the Southwest (SW) and Southeast (SE) areas, setting this region apart from the city's remaining sectors. This concentration underscores an elevated likelihood of life-threatening harm for individuals residing within this geographical sphere. Importantly, this clustering of Homicides persists consistently across each passing year.

Notably, the prevalence of Theft and Motor Vehicle Theft emerges prominently, with incidents clustering ubiquitously throughout the city, transcending quadrant divisions. An analogous trend unfolds with the crime category of Assault with a Dangerous Weapon. This observed pattern illuminates a crucial insight – these specific crime types possess an inherent likelihood of transpiring regardless of one's location within the city.

Collectively, this visualization not only empowers the audience with a profound grasp of crime density distribution but also furnishes invaluable insights into crime dynamics that transcend geographical boundaries, ultimately enhancing our understanding of urban safety considerations.

## 7.4 Hypothesis

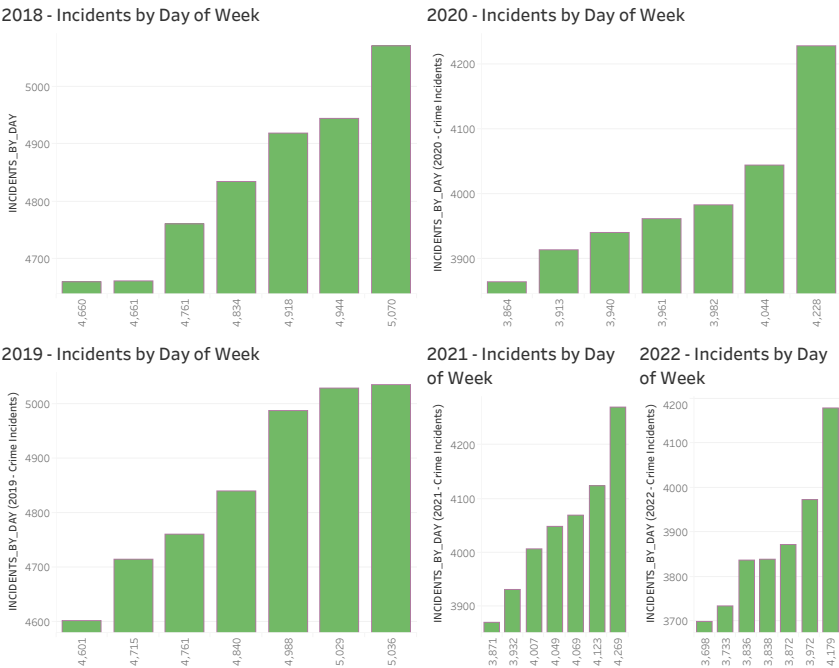
Antecedent to constructing this map, I entertained a hypothesis that an aggregation of crime incidents would materialize in the southern sector of the city, owing to an established precedent. As previously demonstrated, despite the diminished tally of reported incidents, an escalation in the overall crime rate persisted. This observation led me to surmise that heightened activity might still be observable in this specific geographical locale.

Upon executing the analysis and realizing the visual representation, a notable discovery emerged. The crime type that displayed an augmented trend commencing in 2020 was Homicide. This revelation resonates with the earlier conjecture and substantiates the hypothesis that an elevated prevalence of this specific crime would indeed manifest within the southern region of the city.



# Chapter 8

## SHARE - Visualization 5: Day of Week Comparison



A link to the interactive dashboard can be found **here**. For a more specific view, you can access each sheet by clicking sheets 7-11.

## 8.1 Key Tasks

1. Determine the best way to share your findings.
2. Create effective data visualizations.
3. Present your findings.
4. Ensure your work is accessible.

## 8.2 Deliverable

- ☐ Supporting visualizations and key findings

## 8.3 Background/Key Findings

This bar chart effectively captures the incident frequency distribution across different days of the week, spanning the years from 2018 to 2022. Organized chronologically from Sunday to Saturday along the x-axis, each bar delineates the occurrences for their respective day.

The resulting dashboard intricately dissects the reported incidents, affording us a comprehensive breakdown of occurrences for each day. This dissection offers an illuminating insight into the data. Notably, the trend exhibits a progressive surge in crime incidents as the week advances, culminating with the apex observed on Saturday, which records the highest number of reports. A discernible pattern emerges, whereby the week's commencement appears comparatively tranquil, gradually transitioning into a period of heightened activity. This phenomenon beckons an exploration into the underlying factors driving such fluctuations.

One plausible interpretation for this pattern revolves around the dynamics of weekends and the conclusion of the workweek. Weekends often bestow individuals with additional leisure time, potentially contributing to an increased propensity for engaging in criminal

activities. Moreover, the spike in reported incidents during Friday and Saturday could be attributed not solely to criminals but also to the general public enjoying their weekend freedom, partaking in various activities. This augmentation in pedestrian presence creates an environment ripe for both criminal elements and ordinary individuals to interact, thereby elevating the potential for criminal acts to transpire.

In essence, this insightful visualization unravels the intricate interplay between temporal dynamics and crime trends, inviting a deeper contemplation into the contextual factors shaping these patterns

## 8.4 Hypothesis

In anticipation of crafting this graph, I held the belief that weekends, especially Friday and Saturday, would exhibit the highest tally of reported incidents. This inference stemmed from the premise that weekends afford individuals more leisure time, prompting them to venture beyond their residences for recreational pursuits. Upon the completion of the visual representation, our hypothesis was compellingly affirmed.

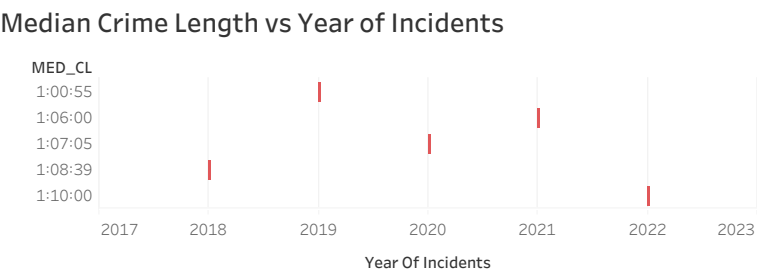
However, one intriguing revelation surfaced that had not been initially considered. The gradual uptick in reported incidents throughout the week emerged as an unexpected pattern. While I had indeed envisioned lower incident counts during weekdays in contrast to weekends, I had postulated a more pronounced divergence. As the graph materialized, it became evident that while a discernible gap does exist between weekdays and weekends, it might not be immediately conspicuous to the observer.

This observation underscores the intricacies inherent in crime trends, where the interplay of factors can yield both anticipated outcomes and unforeseen nuances. In essence, this visual analysis underscores the value of data-driven insights in augmenting our understanding of human behavior and its correlation with temporal dynamics.



# Chapter 9

## SHARE - Visualization 6: Crime Length



A link to the interactive sheet can be found **here**. For a more specific view, you can access the sheet by clicking to sheet 1.

## 9.1 Key Tasks

1. Determine the best way to share your findings.
2. Create effective data visualizations.
3. Present your findings.
4. Ensure your work is accessible.

## 9.2 Deliverable

- ☐ Supporting visualizations and key findings

## 9.3 Background/Key Findings

The impetus driving the development of this Gantt bar graph was to explore potential intriguing revelations that might emerge from analyzing the duration between the start and end times of various crimes. The focal objective was to discern whether any noteworthy patterns or insights could be gleaned from these crime duration intervals, measured in terms of median values across different years. In adopting the median as the metric of choice, the intention was to mitigate the impact of outliers, which might skew results, particularly when crime lengths extend to weeks or even months.

Upon completion of the visual representation, a distinct observation emerges: there exists no substantial correlation or discernible insights that can be readily extracted. The crime duration lengths appear quite consistent across the years, characterized by an even distribution with a marginal range variation of approximately 10 minutes.

It's essential to acknowledge that this visualization was conceived with the intent of transparency, revealing the thought process behind the exploration. It underscores a cardinal truth in data analysis: not every measured variable will yield a significant result or profound insight. This candid demonstration serves as a reminder that the pursuit of



knowledge inherently involves both illuminating discoveries and instances where patterns might elude our grasp.

## 9.4 Hypothesis

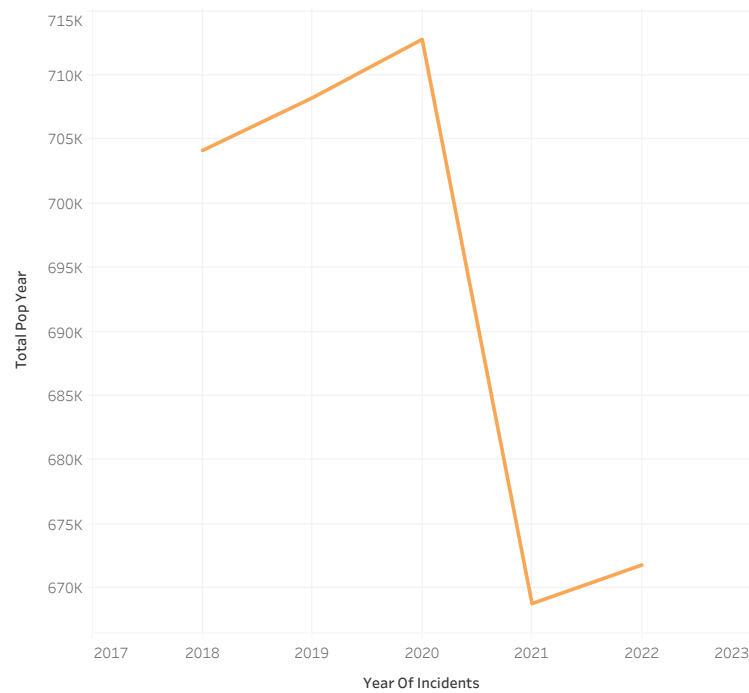
My original hypothesis rested upon the notion of uncovering a positive correlation between crime duration and certain factors, specifically the potential impact of police workload on case closure efficiency. I surmised that a greater workload might hinder officers from promptly resolving cases, thereby elongating crime duration. However, upon rigorous examination, this hypothesis did not withstand scrutiny, as no discernible correlation materialized from the analysis. It's noteworthy that as revealed earlier, the caseload notably diminished post-2020, indicating a potential shift in dynamics that contributed to this outcome.



## Chapter 10

# SHARE - Visualization 7: D.C. Population

Population of D.C. 2018-2022



Change in D.C. population from 2018-2022

A link to the interactive sheet can be found **here**. For a more specific view, you can access the sheet by clicking to sheet 12.

## 10.1 Key Tasks

1. Determine the best way to share your findings.
2. Create effective data visualizations.
3. Present your findings.
4. Ensure your work is accessible.

## 10.2 Deliverable

- ☐ Supporting visualizations and key findings

## 10.3 Background/Key Findings

The underlying purpose driving the creation of this line graph is to shed light on the population dynamics observed throughout the designated time-frame. In the context of pervasive crime, the exploration seeks to ascertain if any correlation exists between population shifts and the prevalent criminal activities. To achieve this, the population data was meticulously sourced from the U.S. Census Bureau.

The line graph effectively aligns population fluctuations with reported incidents by year and the corresponding crime rates. In the period spanning from 2018 to 2020, a steady population increase of approximately 8,000 individuals was discernible. However, a notable transformation transpired beyond 2020, marked by a substantial population decline, tallying around 44,000 fewer residents. This captivating observation gains even more significance as it coincides with a decrease in reported incidents and a simultaneous increase in the crime rate post-2020. This intricate interplay reveals a range of insights.

Firstly, it becomes apparent that the onset of the 2020 pandemic appears to have prompted a mass exodus of residents from the city. This phenomenon could potentially be attributed to the concurrent surge in crime, prompting a migration away from the urban

center. This population exodus might also provide a plausible explanation for the dip in reported incidents, as fewer people remained in the city to lodge complaints with law enforcement.

Moreover, a subsequent minor increase in population by around 3,000 individuals from 2021 to 2022 emerges as a noteworthy occurrence. This parallel mirrors the population growth observed during the years 2018 to 2019. It indicates a semblance of population rebound, although the full recuperation from the pandemic-induced decline might be an extended process.

In summation, this insightful line graph unravels the intricate relationship between population shifts, crime dynamics, and external factors such as the pandemic, underscoring the nuanced interplay that shapes urban landscapes.

## 10.4 Hypothesis

My initial hypothesis postulated a decline in population subsequent to 2020. This projection was rooted in the observation of an escalating crime rate, leading me to conjecture that certain individuals might opt to relocate to perceived safer environments. While the resultant finding proved intriguing, I must confess that it didn't strike me as particularly surprising, given the circumstances.

## Conclusion

In the ever-evolving urban fabric of Washington D.C., this data analyst project has illuminated the multi-faceted landscape of crime through the lens of data-driven exploration. By harnessing the power of Tableau, we have dissected crime incidents, unraveled temporal trends, and charted the spatial nuances of criminal activities across the city.

Through an intricate interplay of visualizations, we have traversed the corridors of crime, delving deep into the relationship between crime rates and key factors. Our insights have offered a compelling narrative of how societal dynamics intertwine with crime occurrences, underscoring the importance of context in shaping the urban safety equation.



U.S. Capitol - Washington D.C.

The tableau visualizations we present are not mere data points; they represent a clarion call for informed action. Armed with evidence, our analysis empowers stakeholders and decision-makers to re imagine crime prevention strategies, allocate resources judiciously, and craft policies that resonate with the diverse communities of Washington D.C.

As we close the chapters of this project, we recognize that data is not an endpoint, but a powerful tool for change. Our exploration of crime trends in Washington D.C. is a stepping stone towards safer neighborhoods, stronger partnerships, and a more secure

future. By heeding the lessons learned from the data, we take strides towards a shared vision of vibrant, resilient communities—a vision where every line on a graph tells a story of proactive transformation.

I humbly pass the torch of knowledge to the guardians of our city, urging them to embrace these insights as catalysts for positive change. We extend our gratitude to the data, the technology, and the unwavering commitment of those who strive to make Washington D.C. a safer haven for all. Together, let us navigate the path to a future where the canvas of crime is painted over with the hues of progress, unity, and hope.

## References

Crime Incidents in 2018. (2015, January 1). Washington D.C. Retrieved August 1, 2023,.

Crime Incidents in 2019. (2015, January 1). Washington D.C. Retrieved August 1, 2023,.

Crime Incidents in 2020. (2015, January 1). Washington D.C. Retrieved August 1, 2023,.

Crime Incidents in 2021. (2015, January 1). Washington D.C. Retrieved August 1, 2023,.

Crime Incidents in 2022. (2015, January 1). Washington D.C. Retrieved August 1, 2023,.

Wards From 2012. (2012, January 28). Retrieved August 5, 2023,.

2023 Year-to-Date Crime Comparison. (2023, August 22). Washington D.C. . Retrieved August 22, 2023,.

Hermann, P. (2023, August 19). Homicides are falling in many big cities. in d.c., they're rising. The Washington Post, pp. 1–1.