~Script~

DC Metro Crime Data:

Examining Homicides (2007 - 2017)

**[Opening Slide]**

Good morning, everyone. Thank you for joining me today. My name is Joel Himes, and I would like to present my project on DC Metro Crime Data from 2007 to 2017. This presentation will explore various aspects of the dataset and address several critical questions about homicides in the DC Metro area.

**[Project Goals]**

 The primary goals of my project were to analyze the DC Metro Crime Data from 2007 to 2017 and answer the five following questions:

1. Is there any correlation between the time of day and the occurrence of homicides?
2. Have homicides increased over this decade?
3. Which wards or neighborhoods had the highest number of homicides?
4. Are homicides more likely to occur with a gun or a knife?
5. Which ward and police service area had the highest number of homicides over the last decade?

**[Benefit of Exploratory Data Analysis]**

Before we delve into the findings, let's discuss the benefits of exploratory data analysis using this dataset:

1. Resource allocation and planning: By understanding crime patterns, we can allocate resources effectively to combat homicides.
2. Community engagement and collaboration: Armed with data-driven decisions, law enforcement can work collaboratively with the community to address crime-related issues.
3. Targeted law enforcement strategies: Exploratory data analysis helps develop targeted strategies to address specific crime trends.
4. Evaluation of crime reduction initiatives: We can track specific metrics to evaluate the effectiveness of crime reduction initiatives.

**[Describing the Dataset: 3 Slides]**

Now, let's take a closer look at the dataset itself.

* It consists of 325,340 rows and 30 columns.
* Some of the interesting features include:
* {'REPORT\_DAT,' 'SHIFT,' 'OFFENSE,' 'METHOD,' 'BLOCK,' 'DISTRICT,' 'WARD,' 'NEIGHBORHOOD\_CLUSTER,' 'BLOCK\_GROUP,' 'VOTING\_PRECINCT,' 'START\_DATE,' 'END\_DATE,' 'XBLOCK,' 'YBLOCK,' 'year,' and 'crimetype.'}
* The dataset primarily contains object/string data with boolean values, float64 coordinates, and int64 date breakdowns.
* There are missing data points in a few columns, accounting for less than 5% of the dataset.
* The relationship between most columns is strong(ish), except for coordinate-based and time-based columns (time and space).

**[Correlation between Time of Day and Homicides]**

Let's address our first question: Is there any correlation between the time of day and the occurrence of homicides? After analyzing the dataset, I found a weak correlation, at best, between time of day, homicides, and shift. This was the case even when favoring the 'Evening/Midnight Shift.' Therefore, we can conclude that the time of day over the last decade does not significantly impact the occurrence of homicides.

**[Homicides Over the Decade]**

Moving on to our second question: Have homicides increased over this decade? The dataset suggests fluctuations in homicides over the years, indicating that some policing initiatives may have worked at certain times over the last decade. However, it is essential to note that this dataset covers a specific period from up to 2017, and further analysis beyond this timeframe is required to draw definitive conclusions about long-term trends.

**[Lay of the Land]**

This slide I just a quick refresher for some or a quick intro for others regarding the “Lay of the Land” in terms of how DC is broken out. Smack dab in between Maryland and VA.

* 8 voting Wards (shown on the left)
* 7 Police Districts (basically just combine voting ward 7 and 8 and you have the 7th police district)
* With 57 Police Service Areas (shown on the right)

**[Wards and Neighborhoods with Highest Homicides]**

Let's explore our third question: Which wards or neighborhoods had the highest number of homicides? According to the data, Ward 8 of Police District 7 in the Southeastern (SE) corridor had the highest number of homicides. Accounting for 355 homicides over the last decade.

**[Weapon Preference for Homicides]**

Now, let's address the fourth question: Are homicides more likely to occur with a gun or a knife? Comparing the data, we found that homicides were committed roughly 75% of the time with a gun, as opposed to other methods, over the specified decade.

**[Ward and Police Service Area with Highest Homicides: 2 Slides]**

Lastly, let's determine the ward and police service area with the highest number of homicides over the last decade.

* Based on the dataset, Ward 8 emerged as the ward with the highest number of homicides – further proving our point in slide 10.
* Additionally, PSA 604 of Police District 7 recorded the highest number of homicides.
* Showing that ward 8 had more homicides over a spread-out area during the specified decade, while PSA 604 in Capital Heights DC/MD had more murders per square mile than any other PSA. This PSA is roughly 4 or 5 square miles.

**[Contact Information]**

Thank you for your attention during the presenetation. If you have further questions or want to explore the project in more detail, please contact me. Here are my contact details:

* Student: Joel P Himes
* GitHub: Joel H
* Email: [himejoel2107@gmail.com](mailto:himejoel2107@gmail.com)
* Link to Repository: <https://github.com/joelphimes/DC-Metro-Crime-Data-2007---2017>
* Link to dataset: <https://www.kaggle.com/datasets/vinchinzu/dc-metro-crime-data>